

Universidade de São Paulo
Escola de Artes, Ciências e Humanidades

**Análise de dados não estruturados:
contribuições usando
Inteligência Computacional e Aprendizado de Máquina**

*Unstructured data analysis:
contributions using
Computational Intelligence and Machine Learning*

Dr. Sarajane Marques Peres

Texto sistematizado apresentado à Escola de Artes, Ciências e Humanidades da
Universidade de São Paulo para obtenção do título de Livre Docente.
Área de concentração: Computação
Área de conhecimento: Inteligência Computacional e Mineração de Dados

Systematized text presented to the School of Arts, Sciences and Humanities of the
University of São Paulo to obtain Habilitation title.
Concentration area: Computer Science
Knowledge field: Computational Intelligence and Data Mining

São Paulo, agosto de 2017

Abstract

This text has been elaborated with the objective of meeting the requirements of the author's application for habilitation in Information Systems at the School of Arts, Sciences and Humanities of the University of São Paulo (EACH-USP). The content of this document systematizes her research contributions, focused on the analysis of unstructured data, considering the strategies provided by the Computational Intelligence and Machine Learning fields. The data covered in the research presented refers to: natural gestures and sign language gestures, texts in Portuguese language (representing news) and event logs coming from the execution of business processes. The strategies that are being applied in this research work include, mainly: artificial neural networks, support vector machines and coclustering algorithms. This document is organized aiming to present discussions about related problems, research hypotheses, obtained results and future directions for the author's research related to such unstructured data types. Moreover, the scientific publications resulting from her research are cited throughout the text. The document concludes with a critical analysis of this research, including a brief explanation of the didactic contributions derived from the knowledge resulted from these research efforts.

Keywords: Data analysis. Unstructured data. Computational Intelligence. Machine Learning.

Resumo

O presente texto foi elaborado como requisito para participação no Concurso Público para Professor Livre Docente junto à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH-USP). O conteúdo organizado neste documento sistematiza as contribuições da autora para a análise de dados não estruturados considerando o uso de estratégias oferecidas pelas áreas de Inteligência Computacional e Aprendizado de Máquina. Mais especificamente, os dados que formam o contexto para pesquisa da autora se referem a: gesticulação natural e gestos de línguas de sinais, textos (notícias) em língua portuguesa, e *logs* de eventos provenientes da execução de processos de negócio. As estratégias sob estudo nesse contexto incluem, principalmente: redes neurais artificiais, máquinas de vetores suporte e algoritmos de coagrupamento. O documento está organizado de forma a apresentar discussões sobre problemas e hipóteses de pesquisa, resultados obtidos e direcionamentos futuros para a pesquisa no contexto de cada um dos tipos de dados não estruturados citados. Além disso, no decorrer do texto, as publicações científicas decorrentes da pesquisa da autora são citadas. O texto é finalizado com uma análise crítica sobre a pesquisa discutida, e com uma rápida explanação sobre a contribuição didática decorrente da produção de conhecimento via tais esforços de pesquisa.

Palavras-chaves: Análise de dados. Dados não estruturados. Inteligência Computacional. Aprendizado de Máquina.

Contents

1	Introdução	4
1.1	Contexto de pesquisa	7
1.2	Objetivos de pesquisa	10
1.3	Organização do texto	12
2	Contribuições em análise de gestos	14
2.1	Gestos	15
2.2	Reconhecimento de gestos em língua de sinais	16
2.2.1	Reconhecimento de <i>queremas</i> e dos sinais da datilologia da Libras	18
2.2.2	Segmentação e reconhecimento das expressões faciais gramaticais da Libras	25
2.3	Análise das fases do gesto	33
2.4	Objetivos atuais e limitações da pesquisa	43
3	Contribuições em análise de textos	46
3.1	Textos, pré-processamento e representações	47
3.2	Análise de texto via agrupamento e coagrupamento	48
3.3	Área de aplicação: Sistemas de Recomendação baseados em conteúdo textual	58
3.4	Objetivos atuais e limitações da pesquisa	61
4	Contribuições em análise de <i>logs</i> de eventos	63
4.1	Processos e <i>log</i> de eventos	64
4.2	Mineração de Dados <i>versus</i> Mineração de Processos	66
4.3	A tarefa de classificação sobre um processo não estruturado	72
4.4	Objetivos atuais e limitações da pesquisa	75
5	Considerações finais	78
5.1	Análise do aspecto científico	79
5.2	Análise do aspecto técnico	82
5.3	Parceria em pesquisa: análise de dados biométricos	84
5.4	Contribuição didática	84

1 Introdução

Em sua definição mais básica, dados são valores que documentam, sob algum aspecto, um fato. Esses valores podem assumir os mais diferentes formatos e representações a depender de como eles serão observados ou usados. A partir da observação ou uso de um dado, atribui-se a ele algum significado, gerando informação. Atualmente, principalmente por conta da evolução tecnológica na qual são produzidos artefatos capazes de gerar dados em quantidades imensas e em uma velocidade praticamente imensurável, o consumo dos dados e da informação presente neles está, claramente, em um estágio rudimentar. Para ilustrar como o conceito de produção de dados e informação mudou durante a história da humanidade, compare os seguintes fatos: a obra completa de Shakespeare pode ser armazenada em 5 megabytes, duas horas de filme podem ser compactadas em 2 gigabytes e a Google processa cerca de 1 pentabyte de dados por hora [1]. Ainda que avanços científicos e tecnológicos estejam sendo rapidamente produzidos, é senso comum dentro da comunidade de analistas de dados que a capacidade de organizar e analisar dados ainda está em seus estágios iniciais. Sabe-se inclusive que, em certas áreas do conhecimento, parte dos dados produzidos é dispensada, haja vista a incapacidade atual de persistir o volume de dados que se pode gerar. Um popular resultado de pesquisa da IDC (*International Data Corporation*), publicado em diferentes reportagens, palestras e notas de aulas sobre o fenômeno de produção de dados, mostra que a capacidade de armazenamento de dados tem se desenvolvido em uma velocidade menor do que a capacidade de criação de dados, sendo que em meados de 2011, havia cerca de 800 exabytes de capacidade de armazenamento e 2.000 exabytes de dados criados – veja o gráfico em [1]. Neste contexto, uma das questões ainda mais prementes é que dentro da produção de dados, a maioria apresenta-se em um formato comumente conhecido como “não estruturado”, o qual em uma primeira análise, diz respeito a textos, documentos, áudios, vídeos, imagens, séries temporais, dados geográficos, cadeias de proteínas, etc.

Tipicamente, segundo [2], dados estruturados são resultantes de processos de geração de dados inerentes a sistemas transacionais, e são armazenados e gerenciados sob um conjunto de relações e de relacionamentos entre relações. Os exemplos de dados supracitados são caracterizados por não permitirem esse tipo de modelo de organização formal, e embora possuam estruturas relevantes para diferentes áreas do conhecimento (um texto, por exemplo, está organizado em estruturas sintáticas e semânticas; áudio está organizado em um sistema de ondas sonoras e, se expressa uma música, há uma organização em termos de escalas e ritmos), em termos de armazenamento e gerenciamento computacional, são considerados como dados não estruturados. Corroborando com esse conceito, o disposto em [3] sobre dados estruturados, não estruturados e semiestruturados. Para esses autores e também para [4], dados não estruturados carecem de uma organização exigida na análise

automática e, ainda, na fronteira dessas definições, os autores defendem que se há uma maneira de descrever os dados não estruturados, esses se tornam semiestruturados (a exemplo do uso de linguagens de marcação como XML - *eXtensible Markup Language*).

A preocupação com o armazenamento e gerenciamento de dados não estruturados, assim como o estabelecimento de formalismos que permitam representá-los e manipulá-los de maneira adequada, é presente principalmente na pesquisa na área de Banco de Dados [5]. Estratégias de gerenciamento de dados alternativas ao modelo relacional, como o modelo orientado a objetos [6, 7], o modelo objeto-relacional [8] e o NoSQL [9], foram propostas para atender às demandas de armazenamento e gerenciamento de dados não estruturados.

Entretanto, a manipulação desse tipo de dados e a extração de informação a partir deles foi, ao longo do tempo, exigindo mecanismos alternativos a linguagens estruturadas de manipulação de dados como SQL (*Structured Query Language*) e derivadas, que pudessem implementar eficientemente a extração e análise de informação e conhecimento de bases de dados (KDD - *Knowledge Discovery in Database* [10] e Mineração de Dados [11]), que agora são grandes e não estruturadas. Diferentes conjuntos de ferramentas são necessários para realizar a análise de dados não estruturados, incluindo pelo menos ferramentas da Matemática e da Estatística, do Processamento de Imagens e Processamento de Sinais, da Inteligência Artificial e Computacional e do Aprendizado de Máquina.

Segundo o Capítulo de Montreal da Sociedade de Inteligência Computacional (IEEE *Computational Intelligence Society*), Inteligência Computacional é um ramo da Inteligência Artificial no qual pesquisadores estão interessados em usar aprendizado, adaptação e computação evolutiva para criar soluções que, em algum nível de abstração, possam ser consideradas inteligentes. Dentro dos interesses da área estão os métodos de resolução de problemas relacionados à redes neurais artificiais [12], computação evolutiva [13, 14] e teoria de conjuntos fuzzy [15]. Já Aprendizado de Máquina, por outro lado, pode ser vista como uma abordagem de aprendizado de software (um aprendizado do tipo indutivo), ao invés de um conjunto de técnicas. Diferentes técnicas, incluindo aquelas que compõem o espectro de interesse da Inteligência Computacional, podem ser aplicadas na implementação do Aprendizado de Máquina, sendo que a interface mais comum entre as duas áreas é representada pelas redes neurais artificiais. Uma definição clássica para Aprendizado de Máquina foi apresentada por [16]. Esse autor entende que um software aprende a partir de um conjunto de experiências E , considerando algumas classes de tarefas T e uma medida de desempenho D , se o seu desempenho nas tarefas T , avaliado sob a medida D , melhora a partir das experiências em E . As classes de tarefas T podem ser as mais diversas. Especificamente no contexto da pesquisa discutida aqui, essas classes foram delimitadas para aquelas definidas nas clássicas taxonomias da Mineração de Da-

dos [2], abrangendo tanto as tarefas preditivas quanto as tarefas descritivas, restritas à subdivisão de tarefas de classificação (ou predição categórica) e tarefas de agrupamento e coagrupamento¹.

Os dois campos de estudos - Inteligência Computacional e Aprendizado de Máquina, e as ferramentas por eles oferecidas, aplicadas a um recorte de problemas modelados como tarefas de Mineração de Dados e caracterizados por demandar a análise de dados não estruturados (especificamente movimentos ou gestos, textos e *logs* de eventos), compõem o interesse de pesquisa, e desenvolvimentos correlatos, discutido neste documento. Os dados não estruturados a que se faz referência são resultantes de duas grandes áreas de estudo:

1. A primeira, da qual dados de movimentos ou de gestos e dados textuais são obtidos, diz respeito à comunicação e expressão humana, manifestada por humanos e/ou estabelecida entre humanos, e a como sistemas computacionais se inserem nesse contexto de comunicação. Os dados referentes a gestos e textos, provenientes da comunicação ou expressão em linguagem gestual, corporal ou escrita, compõem o interesse da maior parte da pesquisa aqui sistematizada. A comunicação e expressão humana é atualmente um dos principais agentes, senão o principal, de produção de dados e informação; por consequência desse fato, o entendimento e análise automática desses dados está assumindo um lugar de destaque nos esforços de pesquisa científica.
2. A segunda área de estudo, da qual *log* de eventos são obtidos, é ainda embrionária no contexto da pesquisa aqui discutida e concentra-se na análise de como sistemas computacionais são usados e como o uso desses sistemas pode suportar a análise de processos organizacionais dos mais variados tipos. *Logs* de eventos ocorridos no contexto da automação de processos são tecnicamente muito fáceis de obter. Não raramente, ferramentas computacionais de propósito geral, ou desenvolvimentos sob demanda, incorporam módulos que registram todas as operações realizadas durante o uso dos sistemas de informação. Inicialmente com a motivação de possibilitar apenas operações de auditoria, esses dados eram processados de maneira muito simplificada; atualmente, com o avanço das técnicas de análise de dados, informação de valor agregado e conhecimento podem ser extraídos desse grande montante de dados.

Este documento, portanto, apresenta uma sistematização dos trabalhos científicos e acadêmicos, realizados e em andamento, que delineiam a pesquisa em *análise de dados*

¹Na literatura correlata, outras nomenclaturas podem ser encontradas para fazer referência à mesma tarefa ou a tarefas similares, por exemplo, biagrupamento (*biclustering*), agrupamento bidirecional (*two-way clustering*), agrupamento por similaridade parcial, agrupamento por partes.

não estruturados e as contribuições apresentadas para essa área por meio da aplicação da Inteligência Computacional e do Aprendizado de Máquina. A partir de uma revisão das estratégias adotadas para resolução de problemas referentes à análise de dados não estruturados, com o apoio de técnicas da Inteligência Computacional e do Aprendizado de Máquina, na pesquisa desenvolvida pela autora a partir de seu doutorado, esse documento sintetiza:

- Uma série de definições de problemas, os requisitos envolvidos em cada problema e suas aplicações práticas.
- Os desafios encontrados na busca pela resolução dos problemas.
- As oportunidades que essa busca abre para o desenvolvimento tanto de soluções inovadoras quanto para o estudo e aprimoramento das técnicas de Inteligência Computacional e Aprendizado de Máquina.

1.1 Contexto de pesquisa

O suporte automatizado da análise de movimentos e gestos provenientes de atividades de comunicação humana tem se tornado cada vez mais eficiente, possibilitando que os movimentos do corpo sejam interpretados para permitir tradução entre línguas e linguagens, geração de comandos de operação de sistemas ou até mesmo o reconhecimento biométrico. Todo esse movimento em busca da automação cada vez mais eficiente vem criando formas inovadoras de interação multimodal, como discutido por [17]. Contudo, para modelar sistemas nos quais os gestos serão usados como meio de interação, não basta capturar os dados e interpretá-los sob um ponto de vista puramente sintático. Segundo [18], é necessário levar em consideração o que está sendo comunicado por meio do uso do gesto (qual é o tipo de mensagem utilizada: mensagem linguística, sinal não linguístico, emoção, atitude), como a informação é passada (expressão facial, gesto manual, movimentos de cabeça ou poses) e por que a informação é passada (o contexto da ação: onde o usuário está, o que está fazendo, se há outras pessoas e objetos envolvidos). Dessa forma, entende-se que não é suficiente apenas uma interpretação superficial do padrão gestual para composição de aplicações de interesse da sociedade, mas sim uma interpretação que seja capaz de gerar informação e conhecimento útil sobre a comunicação expressada por meio de ações do corpo humano. *É nesse contexto que a autora desenvolve pesquisa de análise de dados referentes à linguagem gestual ou corporal, com ênfase principal nos movimentos, ou gestos, realizados com as mãos e com a face, na linguagem espaço-gestual e na gesticulação natural; e com alguma atenção para a execução da marcha humana (caminhada).*

A fim de brevemente ilustrar os dados referentes aos movimentos ou gestos que são objetos de estudo na pesquisa aqui discutida, seguem algumas figuras nas quais contextos reais de execução do movimento ou gesto são apresentados e também alguns exemplos da representação gráfica gerada para os dados que são submetidos à análise automatizada (figuras 1 e 2).

Figure 1: Exemplos de execução de gestos da Língua Brasileira de Sinais: imagem da mão captada na execução de uma letra da datilologia e o resultado do pré-processamento aplicado sobre a imagem para permitir a extração de características referentes a um dedo da mão; imagem da face captada na execução de uma sentença e o resultado da captura em termos de pontos na face, possibilitando a representação da expressão facial em termos de coordenadas espaciais [19, 20]

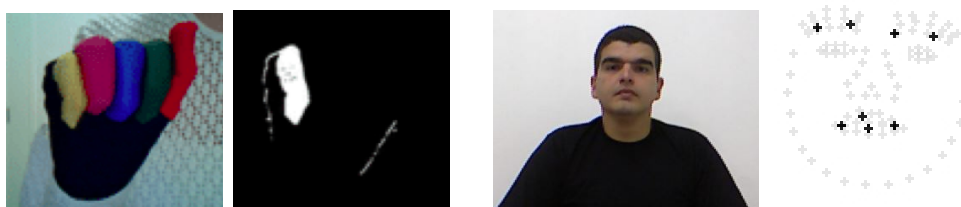
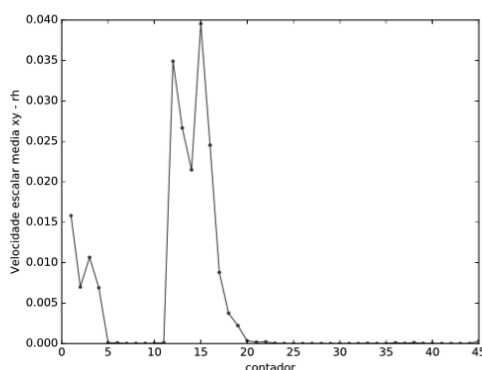
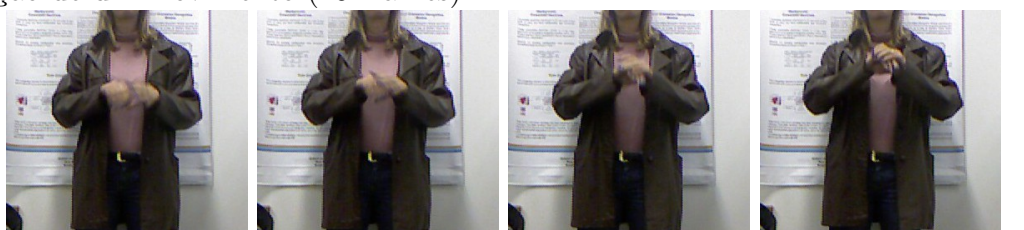


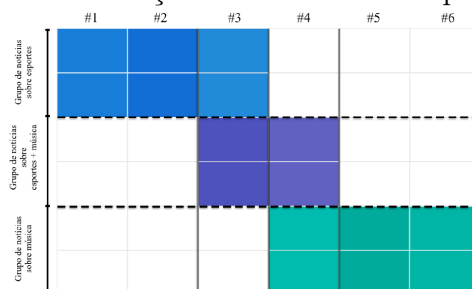
Figure 2: Exemplo de execução de gesto natural: sequência de frames iniciais da execução de um movimento na gesticulação natural; plotagem da velocidade da mão direita na execução de um movimento (45 frames)



Outro aspecto da comunicação humana é aquele caracterizado pela expressão escrita. Como destacado por [21] em seu artigo sobre desafios, anseios e tendências da análise de textos, esse tipo de dado não estruturado tem se proliferado em velocidade e quan-

tidade sem precedentes, principalmente porque a tecnologia de informação pervasiva e ubíqua fornece ferramentas eficientes, tanto do ponto de vista tecnológico quanto do ponto de vista psicossocial, para que o ser humano se expresse livremente por meio de textos publicados *online*. Na pesquisa em análise de texto é comum encontrar iniciativas que focam em análises baseadas em verossimilhança e em análise de similaridade, e que obtêm resultados promissores em termos das clássicas métricas usadas na Mineração de Dados. Entretanto, iniciativas que buscam atender necessidades de mais alto nível, como eficiência em termos de geração de novidade, ou levantamento de informação serendipitosa, são menos comuns e, sob o ponto de vista da autora, tais necessidades são as que mais bem representam as facetas da comunicação humana. *A busca pelo atendimento dessas necessidades de alto nível contextualiza os resultados almejados pela autora no que se refere à análise de dados textuais, com ênfase na comunicação via publicação de notícias e na proposição de recomendação para esse tipo de artefato textual.* Os dados textuais são geralmente representados pela relação entre textos (documentos) e palavras (termos), de forma que cada texto pode ser posicionado no espaço vetorial. A figura 3 expressa graficamente essa relação. Essa figura ainda ilustra o conceito de similaridade parcial entre os documentos, representado pela existência de relações fortes entre palavras e grupos temáticos de textos: o grupo de documentos que tratam sobre esportes é parcialmente similar ao grupo de documentos que tratam de esporte e música devido à existência de uma relação entre os documentos desses grupos e a palavra #3.

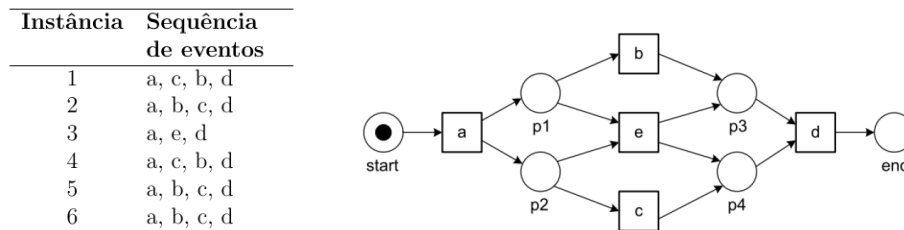
Figure 3: Representação gráfica para dados do tipo texto: cada linha é um texto (documento) e cada coluna é uma palavra (termo); as cores representam valores reais (quanto mais escura maior o valor) que estabelecem a relação entre os textos e as palavras, e cores brancas representam que não há relação entre o texto e a palavra [22]



Já em um caminho marginal à comunicação, mas ainda seguindo a linha da análise do comportamento humano, está a análise de dados provenientes de *logs* de eventos. A Mineração de Processos, segundo [23], tem por objetivo a extração de conhecimento dos *logs* de eventos provenientes do trabalho realizado nas diferentes fases de um processo de negócio. Esse trabalho é, majoritariamente, decorrente do comportamento dos agentes humanos envolvidos no processo e tem potencial para mostrar o perfil de comportamento

humano associado, por exemplo, ao uso de um sistema de informação. *Os dados que são analisados na área de Mineração de Processos são armazenados em arquivos de ações sequenciais, e representam o contexto de análise de dados não estruturado de interesse da autora.* Embora não estruturados, esses dados indiretamente representam um comportamento que pode ser estruturado (como ilustrado na figura 4), o qual está sendo buscado na mineração.

Figure 4: Exemplo de dados na forma de *log* de eventos e representação em Rede de Petri do processo estruturado descoberto a partir da análise dos dados de *log*. Adaptado de [24]



Nos três contextos ilustrados nesta seção, e em outros nos quais também a autora tem investido em colaboração de pesquisa (reconhecimento biométrico), os problemas a serem resolvidos são modelados como tarefas de Mineração de Dados, e essas, por sua vez, são resolvidas a partir da aplicação das técnicas de Inteligência Computacional e Aprendizado de Máquina.

De forma geral, os problemas de análise de gestos têm sido modelados como tarefas de classificação e tarefas de regressão, e resolvidos a partir da aplicação de redes neurais artificiais e máquinas de vetores suporte, com algum suporte da teoria de conjuntos fuzzy para tomada de decisão. Já os problemas referentes à análise de dados textuais têm sido prioritariamente atacados a partir da definição da tarefa de coagrupamento, e algoritmos especializados para tratar essa tarefa têm sido desenvolvidos. Por fim, no caso da Mineração de Processos, a tarefa de classificação, resolvida a partir da técnica de redes neurais artificiais, foi o primeiro alvo de atuação da autora. Contudo, a fase de pré-processamento de dados e de otimização para descoberta de processos também estão em estudo por meio da aplicação de técnicas de busca local e busca heurística usando computação evolutiva.

1.2 Objetivos de pesquisa

Considerando o contexto apresentado nesta introdução, os esforços em pesquisa da autora, desde o término de seu doutorado, estão concentrados, em termos técnicos, no uso e aprimoramento das técnicas de Inteligência Computacional e Aprendizado de Máquina,

à luz da resolução de problemas da análise de dados modelados como tarefas de Mineração de Dados.

Como objetivo geral de pesquisa, a autora segue a linha de contribuir para definir formalmente, ou refinar a definição formal, problemas e soluções na área de análise da comunicação e expressão humana, primeiramente a partir da análise da linguagem gestual (ou gestual-espacial) e mais recentemente a partir da análise da linguagem escrita. Ainda de maneira embrionária, problemas de mineração de processos começaram a ser tratados também na linha de refinamento de definições e apresentação de soluções inteligentes. A estratégia de abordagem dos problemas inerentes às áreas citadas está sempre pautada em uma área de aplicação que demanda soluções para problemas de ordem prática.

Para clarificar o objetivo geral de pesquisa em termos de contribuição científica, os seguintes objetivos específicos são apresentados:

- Os problemas que geram os dados não estruturados explorados nesta pesquisa não apresentam, necessariamente, uma definição formal adequada para o estudo de soluções automatizadas buscadas sob a luz de tarefas de Mineração de Dados. Então, é um objetivo da pesquisa propor definições formais ou semiformais que estabeleçam a aderência entre a área de aplicação e a área técnica de resolução do problema (Inteligência Computacional e Aprendizado de Máquina), e proporcionem tanto a especificação formal de um conjunto de requisitos funcionais e não funcionais dos problemas, quanto a escolha de métricas de avaliação intrínsecas e extrínsecas para as soluções encontradas. Para isso, o entendimento conceitual do problema e das variáveis envolvidas e de como elas influenciam no comportamento dos dados que caracterizam o ambiente de experimentação das soluções, é essencial e constitui-se como o objetivo específico fundamental da pesquisa.
- Esclarecer a aplicabilidade das soluções apresentadas em suas formas genéricas, ou seja, mostrar as situações práticas nas quais os problemas se manifestam e os agentes reais (pessoas e processos) que podem se beneficiar das soluções, é um objetivo específico buscado nesta pesquisa, tanto para clarear sua importância para a sociedade, bem como para apresentar meios de ilustrar o benefício que tal pesquisa está gerando ou tem o potencial de gerar, inclusive em termos de inovação tecnológica.
- Propiciar o estreitamento da interface entre o problema da análise de dados não estruturados com os esforços gerais de desenvolvimento científico e tecnológico, por meio da criação de ambientes de experimentação refinados (conjuntos de dados de referência), orientados às características e variáveis específicas dos problemas sob exploração. Embora, como já discutido neste texto, haja uma enorme massa de

dados sendo continuamente criada pelos mais diferentes recursos tecnológicos, esses dados não se apresentam devidamente organizados para que problemas ainda pouco conhecidos e “mal” definidos possam ser adequadamente explorados, permitindo que soluções eficientes e eficazes sejam desenvolvidas, e aplicações de fato úteis para o consumo inteligente desses dados sejam disponibilizadas para uso geral da sociedade.

Enfim, a contribuição acadêmica buscada com a realização do trabalho de pesquisa que pretende alcançar os objetivos já citados pode ser descrita como um caminho para:

- Elevação da maturidade da área de análise de dados não estruturados em termos de formalização, localização da pesquisa para o contexto brasileiro e melhoramento de interfaces inerentes à interdisciplinaridade.
- Estabelecimento de um grupo local de pesquisa com colaborações nacionais e internacionais e com interesses voltados para a análise de dados não estruturados no contexto da análise automática da comunicação humana.
- Criação de material sistematizado (bibliográfico e conjunto de dados) para suportar e direcionar os resultados comparativos na área de análise automática da comunicação humana.
- Estabelecimento de uma linha de atuação em mineração de processos de negócios, capaz de aproximar as áreas de Inteligência Artificial e de Gerenciamento de Negócios, dentro da área de Sistemas de Informação.
- Produção de material didático, para uso no ensino de graduação e pós-graduação, baseado nos conhecimentos científicos e avançados adquiridos e produzidos no trabalho com pesquisa científica.

1.3 Organização do texto

Este documento é composto de quatro capítulos, além desta introdução e da seção de referências que inclui tanto os itens de produção científica da autora, citados ou discutidos no decorrer do texto, quanto os itens usados como referências bibliográficas. Em relação ao conteúdo de cada capítulo, tem-se:

- Capítulo 2: o conteúdo deste capítulo retrata os investimentos em pesquisa mais antigos da autora. O interesse por análise de gestos executados por humanos iniciou no estudo da automação do reconhecimento de sinais da Língua Brasileira de Sinais, e se estendeu para a análise de gestos suportada pela Teoria de Gestos usada

principalmente na Linguística e na Psicolinguística. Nessa linha de pesquisa está a maioria da produção científica obtida ao longo de dez anos de trabalho. O capítulo traz as informações referentes aos diferentes problemas estudados, tanto em termos de conceitos quanto em termos de aplicações correlatas, as contribuições técnicas e científicas produzidas, as limitações das soluções desenvolvidas, os desafios que estão atualmente sendo abordados e as oportunidades futuras.

- Capítulo 3: a análise de dados textuais é a segunda linha de pesquisa em foco. Nessa linha o interesse está voltado para a exploração da interpretabilidade dos resultados obtidos a partir da aplicação das técnicas automáticas de análise de dados. Especificamente para a análise de textos, há cerca de três anos apenas, a autora tem investido em análise de coagrupamentos, e tem explorado o problema de recomendação como principal motivação para ilustrar as potencialidades das soluções alcançadas. Os resultados produzidos até o presente momento apresentam contribuições diretamente ligadas à melhoria de técnicas de coagrupamento. O capítulo relata tais resultados e apresenta a organização atual da pesquisa em desenvolvimento.
- Capítulo 4: esforços de pesquisas preliminares na área de Mineração de Processos são relatados nesse capítulo. Para Mineração de Processos, são explorados dados de *logs* de eventos obtidos a partir da execução de sistemas de informação. As pesquisas nessa área também foram iniciadas há cerca de três anos. Como primeiro resultado de pesquisa, um mapeamento sistemático da área, que permitiu conhecer seus pontos fortes e fracos e também as lacunas existentes, é o principal resultado produzido. Em uma escala menor do que os outros dois focos de pesquisa, a Mineração de Processo é atualmente assunto de algumas orientações e coorientações em andamento, tanto no nível de mestrado quanto no nível da graduação. A principal dificuldade em pesquisar nessa área é a escassez de conjuntos de dados. Os dados são, em sua maioria, corporativos e sensíveis, como explicado durante o capítulo.
- Capítulo 5: as considerações finais sobre o texto sistematizado referente à pesquisa desenvolvida pela autora incluem uma reflexão crítica em termos de aspectos científicos e técnicos relacionados à pesquisa em discussão, incluindo uma análise crítica do que foi possível realizar e do que ficou pendente durante todos esses anos, e as perspectivas futuras para o trabalho científico pretendido na linha de análise de dados não estruturados. O capítulo também conta com algumas considerações sobre parcerias de pesquisa e contribuições didáticas na área de análise de dados. Essa finalização do texto inclui, permeadas às demais discussões, algumas observações referentes a desafios e oportunidades de cunho econômico, social e operacional da

pesquisa percebidas pela autora no decorrer de todo o processo relatado.

2 Contribuições em análise de gestos

Este capítulo é destinado a apresentar a pesquisa que vem sendo realizada pela autora na área de análise de gestos. Essa linha de pesquisa vem sendo trabalhada sob duas perspectivas: (a) análise de gestos no contexto de línguas de sinais – um foco de interesse mais antigo dentro do trabalho desenvolvido pela autora; (b) análise de gestos no contexto do discurso e comunicação natural, suportada pela Teoria de Gestos usada principalmente na Linguística – ponto que é atualmente o principal interesse da autora dentro dessa linha de pesquisa.

Como uma forma de delimitar a complexidade de análise de dados demandada em cada uma das duas perspectivas supracitadas, essa pesquisa tem adotado uma taxonomia simplificada para a análise de gestos. Essa taxonomia foi definida no escopo do trabalho desenvolvido pela autora, a partir da experiência adquirida no contato com diferentes linhas de trabalhos de análise de gestos adotada por pesquisadores da área:

- Análise de vocabulário predefinido: estudo de conjuntos finitos e predefinidos de gestos, como aqueles usados em grande parte dos estudos e desenvolvimentos tecnológicos para interface humano computador, ou considerando um conjunto reduzido e simplificado de sinais em línguas de sinais. As aplicações mais comuns que implementam a interação humano-computador seguem esse escopo de gestos. Alguns exemplos são: interação com sistema via telas de toque ou usando o espaço tridimensional em frente ao dispositivo [25, 26, 27]; interpretação de línguas de sinais [28, 29, 30]², e análise de comportamento humano ao tocar um instrumento [31] ou ao conduzir uma orquestra [32].
- Análise de gesticulação: análise de gestos que acompanham uma fala (oral ou gestual). Trata de estudar os gestos que as pessoas executam durante uma conversação (usando uma modalidade de comunicação em um determinado contexto), sob a visão de estudos linguísticos. Segundo [26], são gestos sem significado específico, incertos e que refletem diversidade local e cultural. Alguns exemplos de aplicação desse tipo de análise são: descoberta sobre se um gesto é ou não intencional [33], autenticação biométrica [34], e incorporação de comportamento natural para robôs [35, 36].

O conteúdo do presente capítulo está organizado de forma a apresentar a análise de gestos em termos dos principais conceitos teóricos referentes a gestos, a modelagem do

²Madeo, R. C. B.; **PERES, S. M.**; Biscaro, H. B.; Boscaroli, C. A committee machine implementing the pattern recognition module for fingerspelling applications. In: Proceedings of the 2010 ACM Symposium on Applied Computing. New York, NY, USA: ACM, 2010. (SAC'10), p. 954–958.

problema de análise de gestos incluindo o entendimento desses como dados não estruturados e a aplicabilidade decorrente da automação de análise de dados nessa área. O capítulo também mostra que a inserção da autora nesse campo de estudo se dá por meio da aplicação de Inteligência Computacional e Aprendizado de Máquina na resolução do problema de análise de gestos modelados como tarefas de Mineração de Dados (classificação e análise de agrupamento). Ainda, como uma forma de ilustrar a pesquisa já realizada, a produção científica, técnica e acadêmica decorrente da pesquisa realizada são citadas ao longo do capítulo sempre que o texto apresentado mencionar conceitos correlatos ao conteúdo da produção, sendo que dados bibliográficos das mesmas são apresentados em detalhes no rodapé. O capítulo é finalizado com a apresentação dos objetivos que estão atualmente sendo trabalhados e de limitações inerentes ao contexto da pesquisa.

2.1 Gestos

As definições encontradas em dicionários para a palavra *gesto*, no geral, fazem referência a movimentos realizados com uma parte do corpo com o objetivo de expressar, de forma voluntária ou involuntária, uma ideia, um atitude ou um estado psicológico³. Mais especificamente, as seguintes definições são relacionadas ao o escopo da análise de gestos discutida aqui:

- Movimento do corpo, principalmente das mãos, dos braços, da cabeça e dos olhos, para exprimir ideias ou sentimentos, na declamação e conversação⁴.
- Movimento expressivo de ideias; Aceno; Esgar, trejeito⁵.
- Modo de se expressar utilizando as mãos, os braços ou outras partes do corpo para substituir ou complementar a fala (GESTICULAÇÃO, MÍMICA); O mesmo que gesticulação: conjunto de gestos das mãos e dos braços e de expressões faciais de que se vale o orador para prender a atenção da audiência ou para enfatizar ou valorizar seu discurso; Manifestação de um sentimento por meio de expressão facial ou das mãos (EXPRESSÃO; GESTICULAÇÃO)⁶.

No escopo desta pesquisa, o gesto é estudado como um elemento de um conjunto finito e predefinido de gestos, ou como parte da gesticulação, caso em que tal conjunto não existe. Na definição do gesto como um elemento de um conjunto finito é necessário

³Gesto pode ainda ser definido como um *procedimento*. Em [37], o seguinte exemplo do uso da palavra gesto ilustra esse significado: “O bispo pediu gestos concretos em direção à paz”. Contudo, esse significado para a palavra *gesto* está fora do que é estudado pela autora.

⁴Dicionário Michaelis Online - michaelis.uol.com.br, acessado em janeiro de 2017.

⁵Dicionário Priberam da Língua Portuguesa - www.priberam.pt, acessado em janeiro de 2017.

⁶Projeto Caldas Aulete - www.aulete.com.br, acessado em janeiro de 2017.

estabelecer os parâmetros que o definem e esses parâmetros incluem, por exemplo, qual a parte do corpo que está envolvida na manifestação do gesto, quais configurações físicas são esperadas para a parte do corpo enquanto o gesto é executado, qual a forma do movimento que define o gesto. Os parâmetros escolhidos dependem da análise pretendida em termos de: objetivo de uso dos gestos; ou complexidade admitida para o estudo.

Os gestos executados com as mãos (gestos manuais) e aqueles executados com a face (um tipo de gesto não-manual) são objeto de estudo, na pesquisa aqui descrita, dentro do contexto de línguas de sinais (Seção 2.2); gestos manuais são também estudados nesta pesquisa sob o ponto de vista da área de Estudo dos Gestos, considerando a gesticulação no contexto da conversação natural (Seção 2.3). A área de Estudo dos Gestos tem o objetivo de entender como os gestos são usados para fins comunicativos, dentro de uma visão interdisciplinar que abrange, principalmente, a Linguística, a Psicologia e as Ciências Sociais [38]. A comunicação pode ser essencialmente construída por meio dos gestos, porém, se os gestos são usados junto à comunicação oral (fala), diz-se que eles compõem a gesticulação [39].

Em ambos os contextos (língua de sinais ou gesticulação na conversação natural) pressupõe-se o estudo de *gestos naturais* tanto quanto possível⁷ e a sua inserção em um sistema que organiza os elementos de uma língua (oral ou gestual), buscando a realização de uma comunicação plena.

2.2 Reconhecimento de gestos em língua de sinais

As línguas de sinais são sistemas linguísticos, compostos por regras e padrões, estabelecidos naturalmente ao longo do tempo, que coexistem sobre um sistema. Segundo [40] e [41], tal sistema é composto por sinais arbitrários decorrentes de um processo criativo colaborativo, executado pelos usuários da língua, caracterizado por uma estrutura bem definida – os parâmetros, e bastante influenciado por questões culturais. Na linha de esclarecimento sobre a aderência da língua de sinais a um sistema linguístico humano se encontram alguns estudos primários. Dentre eles destaca-se [42], um estudo originalmente publicado em 1965, no qual o autor apontou, dentro da manifestação das línguas de sinais, a presença de características que permitem que um número restrito de elementos (fonemas e morfemas) seja capaz de representar um número irrestrito de significados, o que foi considerado uma propriedade das línguas humanas segundo [43, 44].

A língua de sinais pode ser estudada a partir de parâmetros. Três parâmetros foram inicialmente associados ao estudo da língua de sinais americana por William C. Stokoe

⁷A necessidade de construção de *corpora* de gestos ou do teste sistemático de sistemas influencia na “naturalidade” da execução dos gestos, seja no contexto de uma língua de sinais ou no contexto da conversação natural.

Jr., e mais tarde outros pesquisadores, em diferentes estudos para diferentes línguas de sinais, passaram a defender a inclusão de outros parâmetros básicos [45, 46, 47, 41]⁸:

- configuração das mãos (*designator*): formato que a mão assume durante a execução do sinais;
- ponto de articulação (*tabula*): local onde o sinal é realizado no espaço tridimensional – ancorado ao corpo ou realizado no espaço neutro em frente ao corpo do interlocutor;
- movimento (*signation*): excursão realizada pelas mãos durante a sinalização, no espaço tridimensional;
- orientação e direção: orientação que a palma da mão assume durante a produção do sinal, e a direção na qual esse sinal é executado;
- expressões não-manuais: posição e movimentação da cabeça, posição e movimentos do corpo, configuração e movimentação dos elementos da face formando uma expressão facial e olhar.

Esses parâmetros foram chamados de *queremas* por [42]. Segundo [44], essa denominação veio de uma metodologia de estudo estruturalista que permitiu identificar os aspectos dos sinais que por si só não possuíam qualquer significado, mas que se relacionados poderiam compor itens lexicais. Nas línguas orais, tais parâmetros são chamados fonemas. Na língua brasileira de sinais, [48] classificam os sinais por meio dos valores de seus *queremas*, também chamados por esses autores de “primitivas”.

Seguindo tal perspectiva de estudo da língua de sinais, a autora investiu esforços para extrair informação primeiramente de imagens e depois de dados (coordenadas espaciais) sensoreados para compor uma representação dos *queremas*. Os *queremas* trabalhados pela pesquisa aqui discutida são: configuração de mão, movimento, direção e expressão não-manual (facial). A partir dessas representações, uma série de trabalhos de pesquisa, suportados por orientações realizadas no âmbito de iniciação científica, trabalhos de conclusão de cursos e de dissertações de mestrado, deram origem a resultados documentados em um conjunto de publicações científicas e técnicas, os quais seguem discutidos no restante desta seção.

⁸Ainda outros parâmetros podem ser considerados: arranjo das mãos, contato, movimentos locais e contornos, planos e qualidade do movimento, etc.

2.2.1 Reconhecimento de *queremas* e dos sinais da datilologia da Libras

Os primeiros trabalhos desenvolvidos pela autora no contexto de análise de gestos se deram mediante orientações de alunos de graduação, são relacionados ao reconhecimento de configurações de mãos da datilologia em Libras, e foram inspirados pelo trabalho de [49]. Essas primeiras iniciativas, embora simples, permitiram conhecer em mais detalhes as particularidades e dificuldades envolvidas no problema de reconhecimento das configurações de mão. O contexto de estudo foi limitado a algumas variações de execução de sinais estáticos da datilologia, gerando imagens de fundo uniforme que foram processadas para extração do objeto de interesse - a mão. O problema de reconhecimento da configuração de mão foi modelado como um problema de classificação, resolvido a partir do uso de algoritmos de Aprendizado de Máquina baseados em quantização vetorial (*Learning Vector Quantization* - LVQ) [12]. A codificação usada para representar vetorialmente uma configuração de mão foi a assinatura de bits, a qual embora simples, foi usada em modelagens mais complexas deste mesmo problema em outros trabalhos comentados neste texto. Os resultados desses estudos foram apresentados em eventos da área de redes neurais artificiais [50, 51]^{9,10}.

Com o avanço do estudo das particularidades de uma língua de sinais, o passo seguinte para melhorar o escopo dos trabalhos foi aumentar o número de *queremas* sob análise. Então a pesquisa passou a abranger também alguns movimentos e a direção assumida pela mão, além de aumentar o conjunto de configurações de mão sob análise para o conjunto completo usado na Libras¹¹. E, ainda, a fim de isolar o problema de reconhecimento do padrão referente aos elementos gestuais (os *queremas*), o problema de processamento da imagem e segmentação da mão – que executa o *querema* – foi simplificado a partir do uso de luvas coloridas. Assim, a segmentação da mão passou a ser um problema simples de reconhecimento de cores em uma imagem, no qual redes neurais *Multilayer Perceptrons* (MLP) [12] foram aplicadas como um operador capaz de realizar um mapeamento que implementa a segmentação de um objeto dentro de uma imagem (como definido por [52]¹²).

Com o uso das luvas coloridas, o objeto de interesse (a mão ou os dedos da mão) passa

⁹PERES, S. M.; Flores, F. C.; Veronez, D.; Olguin, C. J. M. LIBRAS signals recognition: a study with learning vector quantization and bit signature. In: Proceedings of the Ninth Brazilian Symposium on Neural Networks - SBRN. Ribeirão Preto, 2006. p. 119–124.

¹⁰Neris, M. N.; Silva, A. J.; PERES, S. M.; Flores, F. C. Self organizing maps and bit signature: A study applied on signal language recognition. In: Proceedings of the International Joint Conference on Neural Networks - IJCNN, part of the IEEE World Congress on Computational Intelligence - WCCI. Hong Kong:IEEE, 2008. p. 2934–2941.

¹¹Não existe um consenso sobre o conjunto completo de configurações de mão usadas na Libras.

¹²Flores, F. C.; PERES, S. M.; Zuben, F. J. V. Automatic design of W-operators using LVQ - application to morphological image segmentation. In: Proceedings of the International Joint Conference on Neural Networks. Hawaii, USA: IEEE, 2002, v. 1, p. 755–760.

a ser a região colorida de interesse, e a MLP é usada para construir um modelo que é capaz de classificar um pixel da imagem como pertencente ou não à região colorida (a classe) de interesse. A figura 5 ilustra o processo de construção de uma representação vetorial para uma janela de pixels que representará um pixel de interesse a ser analisado pela rede neural MLP. A MLP, por sua vez, classificará o pixel como pertencente à região colorida de interesse ou ao fundo da imagem. Todos os pixels da imagem assumem a posição de pixel de interesse, com exceção dos pixels da borda. O resultado da segmentação pode ser observado na figura 6, considerando dois tipos de luvas; a luva de seis cores tem a vantagem de permitir a segmentação de cada um dos dedos das mãos, fornecendo mais informação para o modelo classificador que fará o reconhecimento da configuração de mão e, portanto, permitindo a obtenção de resultados de reconhecimento de maior qualidade.

Figure 5: Processo de construção de uma representação vetorial para um pixel de interesse em uma imagem. O círculo vermelho é o pixel de interesse, o qual será representado pelos valores RGB dele e dos pixels vizinhos a ele na janela [19]

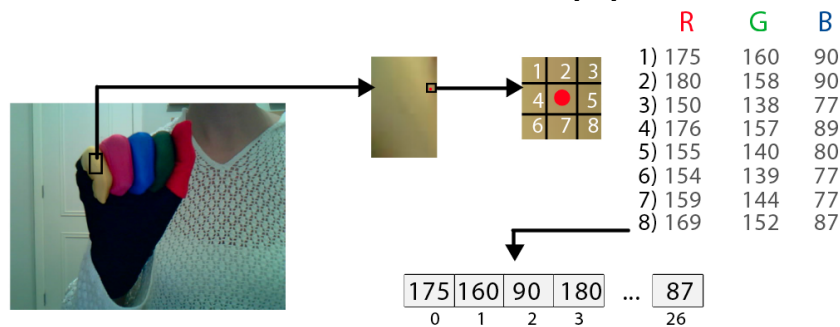


Figure 6: Resultados de segmentação de imagem considerando regiões coloridas de interesse. O resultado da segmentação é apresentado em termos de uma imagem binária



O *querema* configuração de mão foi também estudado em termos do conjunto completo de configurações de mão assumidas na Libras. Esse conjunto é composto por 63 configurações de mão, sendo que algumas delas são bastante similares e de difícil reconhecimento mesmo para humanos. Ainda, a similaridade de tais configurações é tal que, eventualmente, os próprios falantes da Libras não as utilizam necessariamente da mesma maneira. A figura 7, veiculada em diversos sítios de divulgação sobre Libras, mostra as 63 configurações de mão. As configurações de mão mais similares entre si estão destacadas

por meio de quadrados coloridos.

Figure 7: As 63 configurações de mão usadas na Libras, com destaque para as mais similares entre si. Disponível em diversos sítios na Internet



Na pesquisa aqui discutida, o problema de reconhecimento dessas configurações de mão foi modelado como uma tarefa de classificação e a assinatura de bits foi usada na representação dos dados. Uma solução para essa tarefa foi implementada usando a rede neural *Fuzzy LVQ* [53, 54, 55] e uma arquitetura de múltiplas camadas formada por várias instâncias de modelos classificadores também implementados com *Fuzzy LVQ*. Esse estudo revelou que há uma complexidade bastante alta para fazer o reconhecimento de cada uma dessas configurações. A conclusão a partir desta análise é que a representação usada não foi suficientemente discriminante para permitir resultados considerados adequados para, por exemplo, implementar aplicações com os modelos obtidos.

O estudo do *querema* de movimentos dentro do escopo da Libras foi a primeira iniciativa da autora na linha de trabalhar com análise de dados não estruturados que possuem informação variante no tempo e no espaço. Esta linha da pesquisa exigiu a construção de um conjunto de dados com os movimentos que se desejava analisar, o qual foi disponibilizado publicamente e segue descrito ainda nesta seção. A construção do conjunto de dados foi necessária devido à inexistência de conjuntos de dados públicos que permitissem a análise dos movimentos desejados.

Para implementação da análise, os movimentos executados por uma das mãos de um sinalizador humano durante um período do tempo foram representados como curvas bidimensionais [56]¹³. A análise dos movimentos foi realizada, primeiramente, a partir da

¹³Dias, D. B.; Madeo, R. C. B.; Rocha, T.; Bísvaro, H. H.; **PERES, S. M.** Hand movement recognition

aplicação da rede neural artificial *Self-Organizing Map* (SOM) [57], que permitiu realizar a visualização do conjunto de dados mediante a representação escolhida; e então a partir de um problema de classificação resolvido com modelos construídos usando LVQ e variações da *Fuzzy LVQ*. Os resultados do estudo estão disponíveis em [56].

Uma solução mais abrangente quanto a quantidade de *queremas* a ser usada no reconhecimento de sinais da Libras foi objeto dessa pesquisa. Nessa solução, os dois *queremas* já estudados são usados juntamente com o *querema* de direção da mão, a fim de propor uma arquitetura modular de reconhecimento de sinais da Libras. A hipótese defendida nesse estudo é que a construção de uma arquitetura modular, na qual os resultados de modelos classificadores especializados por *querema* pudessem ser combinados, poderia aumentar o poder de reconhecimento de padrões e proporcionar a resolução de problemas mais complexos. A arquitetura é composta de módulos de reconhecimento de configurações de mão e de movimentos baseados em *Fuzzy LVQ*, um módulo de reconhecimento da direção da mão baseado em regras heurísticas e um módulo de integração dos resultados dos módulos anteriores baseados em uma gramática *fuzzy*. Nessa gramática *fuzzy* cada *querema* é considerado uma variável linguística *fuzzy* cuja regra semântica (função que define os conjuntos *fuzzy* que estabelece o grau de pertinência de uma entrada aos termos linguísticos da variável) é o módulo de reconhecimento do *querema*.

A prova de conceito para essa solução foi a datilologia da Libras, cujos sinais envolvem apenas combinações dos três *queremas* sob estudo. A gramática *fuzzy* estabelecida para a datilologia permitiu alcançar resultados na ordem de 85% de acerto nos casos de teste, representando uma taxa de acerto maior (10 pontos percentuais maior) do que a obtida para as configurações de mão, apenas. O modelo de gramática *fuzzy* para os gestos é

$$\begin{aligned} \text{Variáveis-Linguísticas} &= \{ Conf, Mov, Orient \} \\ \text{Termos-Linguísticos} &= \{ conf_i, mov_j, orient_k \} \\ \text{Símbolo-inicial} &= \{ Conf \} \\ Conf &\xrightarrow{\alpha} conf_i \oplus Mov \\ \text{Regras-Produção} &= \{ Mov \xrightarrow{\beta} mov_j \oplus Orient \} \\ &\quad Orient \xrightarrow{\gamma} orient_k \end{aligned}$$

em que *Variáveis-Linguísticas* é o conjunto de símbolos não terminais em um alfabeto, e correspondem aos *queremas*; *Termos-Linguísticos* é o conjunto de símbolos terminais em um alfabeto, e correspondem às variações consideradas para cada *querema*; α , β e γ são os graus de pertinência para os termos linguísticos gerados a partir da apresentação de uma entrada (execução de um sinal da datilologia da Libras); e \oplus é uma t -norma adequadamente escolhida¹⁴. Os resultados desse estudo, assim como mais detalhes sobre

for Brazilian sign language: A study using distance-based neural networks. In: International Joint Conference on Neural Network - IJCNN. Atlanta:IEEE, 2009. p. 697–704.

¹⁴Experimentos para esse problema mostraram que a t -norma adequada é o produto algébrico.

a arquitetura de reconhecimento de padrões proposta, estão disponíveis em [58]¹⁵.

Embora os resultados obtidos até este ponto da pesquisa tenham sido promissores, a acurácia da solução ainda não seria suficiente para, por exemplo, suportar o desenvolvimento de uma aplicação que usasse o reconhecimento de sinais da Libras, ou da datilologia da Libras, para implementação de alguma funcionalidade. Então, finalmente, o último trabalho desenvolvido nesta linha pela autora fez uso de uma abordagem de reconhecimento baseada no trabalho de [59], e de uma abordagem de extração de informação das mãos usada em [60] e [61]. Nesses trabalhos, os autores sugerem o uso de cores para extrair informação sobre partes da mão, e mais especificamente em [59], a informação sobre partes da mão é usada para reconhecimento de sinais de uma língua de sinais. Assim, a representação da configuração de mão via assinatura de bits foi substituída por informações de largura, comprimento, compacidade e curvatura de cada um dos dedos da mão. Essa representação, usada em uma tarefa de classificação solucionada a partir do uso de MLP, permitiu alcançar uma acurácia na ordem de 92% na datilologia da Libras. Esse resultado motivou a construção de uma aplicação - o Jogo da Força em Libras, discutido nesta seção.

Conjunto de dados - *Libras Movement Data Set*

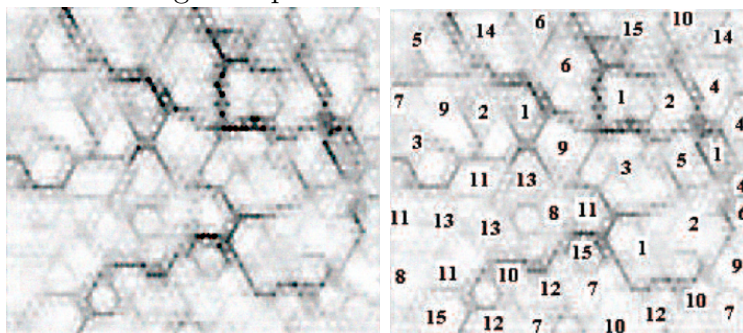
O conjunto de dados *Libras Movement Data Set* [62]¹⁶ contém 15 classes com 24 instâncias em cada uma delas. Cada classe é referente a um tipo de movimento que ocorre na fala em Libras. Nos procedimentos de pré-processamento dos vídeos, foi executada uma normalização para selecionar 45 quadros de cada um dos vídeos, seguindo uma amostragem uniforme. Em cada um dos quadros do vídeo, a mão é segmentada da imagem e o pixel central é encontrado. A posição deste pixel na sequência de 45 quadros compõe a versão discreta de uma curva. Todas as curvas estão normalizadas no espaço unitário. A representação vetorial adotada para preparar esses dados para serem analisados por algoritmos que trabalham no espaço vetorial consiste de 90 dimensões, representando as coordenadas x, y de cada ponto na curva. A figura 8 fornece uma noção da complexidade da distribuição das classes de movimentos projetados no espaço de duas dimensões a partir da apresentação do conjunto de dados para uma rede neural artificial SOM.

Os arquivos textos com a representação vetorial das curvas referentes aos movimentos estão disponibilizados no repositório *UCI - Machine Learning Repository* [63] para servir como um conjunto de referência para os pesquisadores interessados em estudar dados não

¹⁵Madeo, R. B. C.; **PERES, S. M.**; Lima, C. A. M.; Boscaroli, C. Hybrid architecture for gesture recognition: Integrating fuzzy-connectionist and heuristic classifiers using fuzzy syntactical strategy. In: Proceedings of International Joint Conference on Neural Networks – IJCNN. Brisbane: IEEE, 2012. p. 1–8.

¹⁶Dias, D. B., **PERES, S. M.**, Biscaro, H. H. *Libras Movement Data Set. 2009*. Disponível em <https://archive.ics.uci.edu/ml/datasets/Libras+Movement>

Figure 8: Projeção do conjunto de dados *Libras Movement Data Set* por meio de um mapeamento realizado pela rede neural artificial SOM: (a) projeção sem rótulos; (b) projeção rotulada. Cada número na figura representa o rótulo de uma classe de movimentos [56]



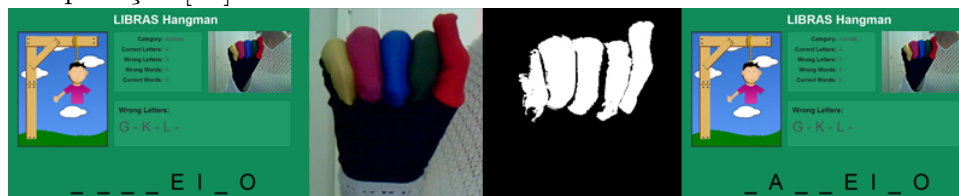
estruturados relacionados a gestos manuais. O conjunto está disponível na sua versão completa e em versões compostas por subconjuntos de movimentos.

Jogo da Forca em Libras – Libras *Hangman Game*

A partir dos estudos realizados no reconhecimento de configurações de mãos e dos sinais da datilologia, uma produção técnica foi gerada [19]¹⁷¹⁸, a qual embora esteja em condições de uso, está ainda em fase de refinamento técnico, isto é, ainda não se constitui como uma aplicação com maturidade para ser disponibilizada para o público em geral.

A figura 9 ilustra a tela principal de interação com a aplicação e traz uma imagem referente à entrada esperada pelo jogo, seguida da imagem processada da qual as características são extraída para possibilitar o processo de reconhecimento de padrões. A sequência apresentada nessa figura é uma abstração da lógica do Jogo da Forca em Libras. A figura 10 mostra uma arquitetura simplificada com a descrição das principais funcionalidades da aplicação. São três funcionalidades principais:

Figure 9: Ilustração da tela principal do Jogo da Forca em Libras e abstração da lógica usada pela aplicação [19]

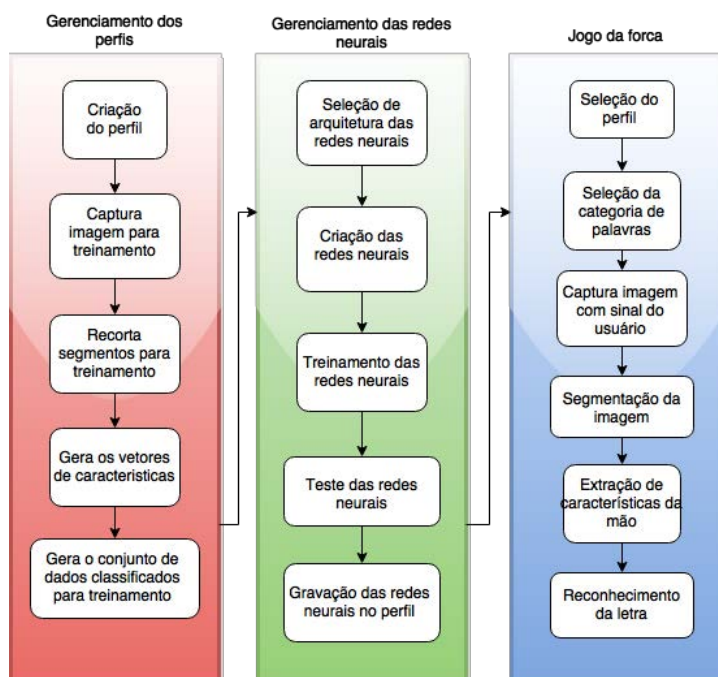


- preparação do conjunto de dados: o programa possui funções de captura de imagem

¹⁷Souza, R. G.; PERES, S. M.; Lima, C. A. M. Adaptação do jogo da forca para língua de sinais usando luvas coloridas e redes neurais multilayer perceptron. In: Anais do XIV Simpósio Brasileiro de Jogos e Entretenimento Digital - SBGames. 2015. p.283-286.

¹⁸Todo o processo de desenvolvimento dessa aplicação esteve sempre relacionado à orientação de alunos em nível de graduação.

Figure 10: Arquitetura simplificada com as principais funcionalidades da aplicação Jogo da Forca em Libras [19]



usando uma *webcam*, manipulação de imagens, armazenamento de dados referentes às imagens manipuladas e organização desses dados na forma de conjunto de dados para treinamento das redes neurais artificiais que permitem a construção de modelos de reconhecimento de padrões usados no jogo;

- construção de modelos de reconhecimento de padrões: o programa também oferece uma interface gráfica para operação de funções de criação, treinamento e teste de modelos reconhecedores de padrões, a partir do uso de redes neurais artificiais; tais modelos são utilizados para a segmentação dos dedos da mão na imagem de entrada do programa e para reconhecimento da configuração de mão apresentada nessa entrada;
- regras do jogo da forca: o programa apresenta funcionalidades para lidar com listas de categorias de palavras que podem ser escolhidas pelo jogador para cada sessão de uso do jogo da forca. O programa permite que uma palavra seja sorteada (da lista) e organiza a interface gráfica de forma que a palavra possa ser usada na sessão do jogo. No uso, a *webcam* do computador é manipulada pelo programa, com apoio do usuário via teclado e mouse, para capturar a imagem do jogador soletrando uma letra em língua de sinais. Com a imagem capturada, o programa aplica os modelos de redes neurais artificiais de segmentação para extrair informações sobre a configuração da mão. Essas informações são codificadas para serem usadas em

um outro modelo de rede neural artificial para reconhecimento de qual letra foi apresentada como entrada (via imagem). Com a letra identificada, o programa verifica se ela faz parte da palavra sorteada e, caso sim, as posições nas quais a letra ocorre são reveladas ou, caso contrário, a letra entra para a lista de letras erradas e o jogador fica mais próximo de ter seu personagem “enforcado”.

A motivação para a adaptação de um jogo clássico para o contexto do uso da língua de sinais é, principalmente, proporcionar a aproximação entre crianças surdas e ouvintes. A criança ouvinte, conhecedora do jogo clássico, pode se sentir atraída pela inovação do jogo no qual a interação se dá a partir de uma forma diferente de comunicação - uma língua gestual-visual. A criança surda, detentora da habilidade em se comunicar em língua de sinais, é então colocada numa posição de destaque, já que o seu modo de comunicação é o centro das atenções na interação com o jogo. Ainda, o jogo da forca para língua de sinais serve como base para o ensino da soletração e da própria língua para ambos os grupos de crianças.

2.2.2 Segmentação e reconhecimento das expressões faciais gramaticais da Libras

Recentemente a identificação e análise das expressões faciais têm recebido atenção especial de grupos de pesquisas de diferentes áreas. O rosto humano tem sido objeto de estudo para a compreensão de aspectos fisiológicos e psicológicos do comportamento das pessoas durante a sua interação com seu entorno, pois parte dessa interação é baseada nas expressões faciais [64, 65]. Pesquisas na área do reconhecimento automático de expressões faciais têm sido úteis para o desenvolvimento de muitas aplicações [66, 67]. Também, as expressões faciais são usadas como um tipo de gesto não manual que expressa informação referente ao discurso e à comunicação interpessoal. Segundo [68], as expressões faciais compõem 55% da comunicação estabelecida entre os seres humanos, o que evidencia a sua importância nas relações interpessoais. Alguns estudos têm mostrado melhorias em seus resultados no estudo do uso de uma língua (neste caso, língua de sinais) e no uso da linguagem gestual, quando o estudo das expressões faciais está incluído dentro de uma abordagem multimodal de análise. Como exemplos estão os estudos de [69], que trabalham na interpretação da Língua de Sinais Americana, e os estudos de [70] que trabalham na interpretação da linguagem gestual usada por alemães.

Especificamente no contexto das línguas de sinais, as expressões faciais assumem um papel extra. Além de possibilitar a expressividade da emoção de um indivíduo, ela também é usada na formação da estrutura gramatical da língua, assumindo um papel de expressão da prosódia (expressa por meio da voz em línguas orais). Segundo [71], as expressões faciais assumem um papel de grande destaque nas línguas de sinais, pois elas

são essenciais para dar sentido ao que é dito. Na entanto, nas línguas de sinais, determinadas expressões são usadas na composição de sua estrutura, e essas são então chamadas de expressões faciais gramaticais. Essas se encontram presentes no nível morfológico da língua, no qual têm atribuição de adjetivação, e no nível sintático da língua [41]. Elas têm o papel de modificar os sinais, alterando alguns traços mínimos na gestualidade e alterando o sentido da frase. Assim, as expressões faciais gramaticais têm a função de executar marcações não-manuais e, em cada nível no qual atuam, são usadas da seguinte forma, segundo [41]:

- nível morfológico: nesse nível as expressões faciais gramaticais são usadas para impor o grau de adjetivação, fazendo com que seja determinado um grau de intensidade para um objeto; e quando associadas a substantivos permite a construção de superlativos e comparativos de superioridade e inferioridade.
- nível sintático: nesse nível as expressões faciais gramaticais são responsáveis por determinar a construção de frases interrogativas, determinar polaridades (negativas e afirmativas), condicionais relativas, com tópicos e com foco.

Os diferentes contextos de construção de frases relacionados ao uso de expressões faciais gramaticais são:

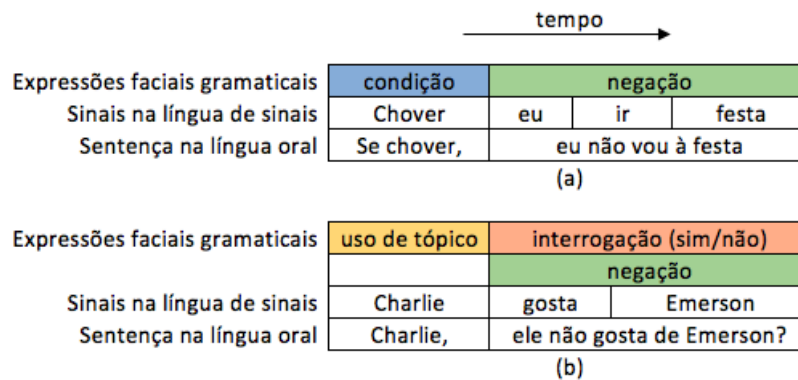
- interrogação: nesse contexto, a intenção é obter alguma informação. Há diferentes tipos de frases interrogativas – tipo *qu* que inclui perguntas envolvendo os advérbios de interrogação “que”, “quem”, “quando”, “onde”, “como”, “por que”; tipo *sim/não*; tipo *dúvida*, que expressa uma desconfiança;
- negação: contextos nos quais há um elemento negativo explícito como “não”, “nada” ou “nunca”¹⁹;
- afirmação: contextos que expressam ideias afirmativas;
- condição: contextos que exigem o estabelecimento de uma condição associada a uma ação, consequência ou conclusão;
- construções relativas: contextos em que se insere informação dentro de uma frase para tornar o contexto mais claro; ou quando é necessário encaixar uma segunda questão relacionada ao que está sendo dito;
- uso de tópicos: trata-se de uma forma diferente de organizar o discurso, explicitando o seu tópico importante.

¹⁹Há manifestações linguísticas, nas línguas de sinais, que admitem o uso de marcação manual para expressão das negações [72].

- **foco:** contexto no qual é necessário introduzir uma informação nova no que está sendo dito para estabelecer contraste, para informar algo adicional ou colocar ênfase em algum objeto, contexto ou informação.

É importante notar que a sinalização dos mesmos sinais que compõem uma determinada construção em uma língua de sinais, se repetida sem a execução da expressão facial gramatical, pode tornar a frase agramatical. Para ilustração desse efeito, considere os exemplos discutidos em [20]²⁰, referentes à construção de sentenças em língua oral e suas respectivas modelagens em língua de sinais, usando expressões faciais gramaticais (figura 11). A construção das sentenças na língua de sinais é, em um primeiro nível, feita a partir da execução de sinais para a composição básica da sentença. Concomitantemente, o aspecto gramatical e prosódico é conferido ao discurso por meio da execução de expressões faciais gramaticais. No exemplo da figura 11(b) há a execução de duas expressões faciais ao mesmo tempo, o que confere um grau de complexidade alto à tarefa de reconhecimento de padrões, discutida mais à frente neste texto.

Figure 11: Construção de sentenças em línguas de sinais usando expressões faciais gramaticais [20]



Neste contexto, o problema de análise de dados consiste em estudar um período de sinalização em língua de sinais, encontrar os segmentos desse período que dizem respeito à execução de uma expressão facial gramatical, e então reconhecer qual expressão facial gramatical está sendo executada. A proposta de resolução para esse problema, apresentada no trabalho de pesquisa da autora, é modelá-lo como uma tarefa de classificação e gerar um modelo de reconhecimento de padrões baseado em Aprendizado de Máquina supervisionado.

Para desenvolver tal modelagem, uma expressão facial $EF_i \in \{EF_1, EF_2, \dots, EF_n\}$ é representada como um conjunto de pontos $P = \{p_1, p_2, \dots, p_n\}$ extraídos de uma imagem (um dado do tipo não estruturado) na qual há uma face humana. Esses pontos estão

²⁰Freitas, F. A. *Reconhecimento automático de expressões faciais gramaticais na língua brasileira de sinais*. 112 p. Tese (Mestrado) – Universidade de São Paulo, 2015. Orientação da autora.

localizados no espaço bidimensional ou tridimensional, com coordenadas (x, y) ou (x, y, z) , a depender se as imagens disponíveis são capturadas apenas por câmeras do tipo RGB, ou também com o apoio de câmeras de profundidade. A figura 12 mostra os pontos disponíveis para uso em cinza; em preto estão os pontos usados no estudo em discussão. O período de sinalização em língua de sinais é representado por uma sequência de quadros de vídeo $\mathcal{V} = \{f_1, f_2, \dots, f_n\}$, sendo que cada quadro corresponde a uma imagem da qual os pontos da EF são obtidos. Um segmento do período de sinalização é uma subsequência $\mathcal{S} = \{f_1, f_2, \dots, f_l\}$, com $l \leq n$ e $\mathcal{S} \subseteq \mathcal{V}$. Os períodos de sinalização \mathcal{V} e \mathcal{S} podem ser entendidos como uma série temporal multidimensional, na qual as dimensões dizem respeito às coordenadas de cada um dos pontos P extraídos da imagem de face humana.

Figure 12: Representação bi-dimensional dos pontos da face. Todos os pontos disponíveis para uso (em cinza) e os pontos usados no estudo em discussão (em preto) [20]



O problema de encontrar os segmentos \mathcal{S} que correspondem a expressões faciais gramaticais consiste em encontrar uma função de mapeamento \mathcal{F} capaz de associar cada $f_i \in \mathcal{V}$ a uma expressão facial gramatical do conjunto $EFG = \{EFG_1, EFG_2, EFG_n\}$, em que n é a cardinalidade de expressões faciais gramaticais existentes na língua de sinais sob estudo, ou no recorte de estudo objetivado dentro do contexto de uma língua de sinais. Sendo assim, esse problema está modelado como um problema de classificação de n classes.

O estado da arte em reconhecimento de expressões faciais gramaticais mostra que o principal objetivo dos estudos na área é apoiar o reconhecimento automático da língua de sinais, considerando uma análise de múltiplos parâmetros. Nenhum dos estudos encontrados até o presente momento consideraram o conjunto completo de expressões faciais gramaticais que compõem o contexto completo da língua de sinais sob estudo. Esses estudos usam como entrada de dados os quadros de vídeos provenientes de captura realizadas com câmeras do tipo RGB [70, 73, 74, 75, 76, 69, 77, 78].

Uma das características mais críticas para o desenvolvimento da área de reconhecimento de expressões faciais gramaticais é a disponibilização de bases de dados para realização de testes e comparação de resultados. Cinco bases de dados foram identificadas no estudo da autora:

- **UWB-07-SLR-P** [79]: é composta por 378 sinais da língua de sinais tcheca, gravados cinco ou mais vezes, em três perspectivas diferentes, em ambiente de laboratório, com controle de iluminação e da variação do fundo. Essa base de dados foi usada por [80].
- Base de dados do **National Center for Sign Language and Gestures Resources**²¹: é composta por 873 enunciados executados na Língua de Americana de Sinais. Essa base de dados foi usada por [76];
- Base de dados de [70]: é composta por 135 sinais isolados e 780 sentenças completas executadas na Língua Alemã de Sinais, capturadas em laboratório, com condições controladas.
- Base de dados de [81] **apud** [73]: é composta por 5 repetições de 8 sinais não manuais da Língua Turca de Sinais, executadas por 11 indivíduos. Essa base foi usada por [73].
- Base de dados da **Boston University American Sign Language Linguistic Research Project** [82]: é composta por 15 pequenas narrativas espontâneas e 400 expressões adquiridas com a colaboração de falantes da Língua Americana de Sinais. Essa base foi usada por [75].

O primeiro estudo realizado no âmbito da pesquisa da autora abordando este problema para a Libras teve o objetivo de explorar a dificuldade de análise referente a cada uma das expressões faciais gramaticais usadas nessa língua. Essa exploração foi necessária por não ter sido encontrado nenhum estudo anterior que já tivesse abordado esse problema. Assim, dois objetivos foram buscados no estudo: (a) a criação de um conjunto de dados para suportar o desenvolvimento de modelos de classificação para reconhecimento; (b) a criação de classificadores binários ($n = 2$) a fim de verificar a complexidade inerente ao problema de encontrar cada uma das expressões faciais gramaticais no vídeo. O conjunto de dados criado constitui-se como uma produção técnica e está descrito em uma subseção específica, mais a frente neste texto. O estudo de classificadores binários foi realizado a partir da arquitetura de rede neural artificial MLP. Tal estudo ainda não considera a presença de expressões faciais gramaticais diferentes em uma mesma sentença, e, portanto, não abrange situações em que há coocorrência de expressões faciais gramaticais. O escopo de análise de dados referentes a expressões faciais, em termos de características da face que se modificam durante a execução da expressão, está mostrado na tabela 1, e exemplos de expressões faciais gramaticais são mostrados na figura 13. O seguinte esquema de símbolos

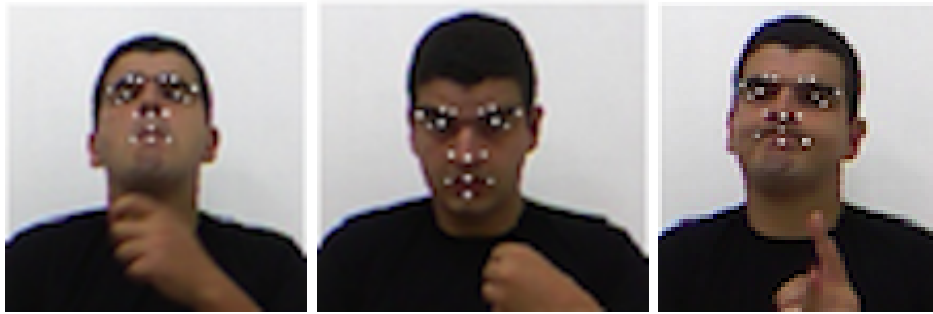
²¹<http://www.bu.edu/asllrp/cslgr/>

é usado nessa tabela: \uparrow movimento para cima; \downarrow movimento para baixo; \leftrightarrow movimento para direita e para esquerda; \updownarrow movimento para cima e para baixo; * compressão; \diamond abertura; \ominus afastamento; e \cap cantos da boca para baixo.

Table 1: Caracterização das expressões faciais gramaticais [20]

Funções sintáticas	Sobrancelhas	Olhos	Boca	Cabeça
Interrogação (<i>qu</i>)	\downarrow			\uparrow
Interrogação (<i>sim/não</i>) / condição	\uparrow			\downarrow
Interrogação (dúvida)	\downarrow	*	*	\ominus
Negação	\downarrow		\cap	\leftrightarrow
Afirmação				\updownarrow
Construção relativa	\uparrow			
Uso de tópico / foco	\uparrow	\diamond		\downarrow

Figure 13: Exemplos de expressões faciais gramaticas: interrogação-*qu*, interrogação-*sim/não* e negativa [83]



Os resultados quantitativos obtidos no primeiro estudo realizado estão apresentados em [20]. Parte desses resultados foram divulgados em uma conferência internacional [83]²². As considerações mais importantes que puderam ser delineadas a partir dos experimentos realizados em tal estudo são:

- considerando os melhores resultados para cada uma das expressões faciais gramaticais, o menor F-score obtido foi na ordem de 0,68 em um experimento com características de independência de usuário (modelos treinados com os dados de um sinalizador e testados com os dados de outro sinalizador). Em média, os experimentos alcançam um F-score de 0,87, indicando a plausibilidade da abordagem utilizada para encontrar as expressões faciais gramaticais no vídeo;
- a expressão facial gramatical usada para a interrogação do tipo *sim/não* é mais complexa e difícil de encontrar dentro do vídeo do que os outros dois tipos de

²²Freitas, F. A.; PERES, S. M.; Lima, C. A. M.; Barbosa, F. V. Grammatical facial expression recognition with machine learning. In.: *27th Florida Artificial Intelligence Research Society Conference (FLAIRS)*. Palo Alto: The AAAI Press, 2014. p. 180–185.

interrogação (*qu* e *dúvida*). Um dos pontos que levam a essa conclusão é que os erros de borda (quadros que estão próximos às transições entre face neutra e expressão facial gramatical na sequência de quadros do vídeo) são mais presentes no reconhecimento da primeira expressão facial gramatical do que no reconhecimento das outras duas;

- as expressões faciais gramaticais mais difíceis de serem encontradas com precisão são aquelas usadas com as funções sintáticas de negação e afirmação. Um dos problemas encontrados nesse estudo exploratório é que há uma diferença no padrão de execução de expressões com função de negação. Ainda, o movimento contínuo da cabeça exigido na execução dessas expressões parece não ser bem capturado pela representação vetorial usada para representar a entrada dos modelos classificadores.
- para todas as expressões faciais gramaticais, o uso de janelas para representar vetorialmente características que variam no tempo²³ (i.e., que variam em uma sequência de quadros) levou a melhoria nos índices de F-score. Contudo, para cada expressão gramatical facial, um tamanho de janela específico é mais adequado. As melhorias mais significativas foram observadas nas expressões que exigem movimento de sobrancelhas e movimento da cabeça para baixo.

Conjunto de dados: *Grammatical Facial Expressions Data Set*

A fim de propiciar o estudo referente a identificação e análise das expressões faciais gramaticais no contexto da Libras, foi necessário projetar e construir um conjunto de dados que contivesse vídeos com sinalização em Libras envolvendo sentenças com determinadas construções gramaticais. O conjunto de dados foi construído com dezoito vídeos, gravados usando o sensor Microsoft Kinect. Em cada vídeo, um sinalizador sinaliza cinco vezes cinco sentenças em Libras. Todas essas sentenças requerem o uso de expressões faciais gramaticais e estão listadas na tabela 2.

Usando o sensor Microsoft Kinect foi possível obter: (a) uma imagem de cada quadro, identificada por um selo de tempo; (b) um arquivo texto contendo 100 coordenadas (x,y,z) de pontos dos olhos, nariz, sobrancelhas, contorno da face e iris²⁴. As imagens propiciaram a rotulação manual de cada arquivo por um especialista em Libras. A rotulação se fez necessária para que modelos classificadores (aprendizado supervisionado) pudessem ser construídos. Os arquivos texto referentes às coordenadas foram disponibilizados publicamente no repositório *UCI - Machine Learning Repository* [63], sob o nome de *Grammatical*

²³Veja uma explicação para esse tipo de representação vetorial na seção 2.3.

²⁴O ponto referente à iris não diz respeito ao rastreamento olhar.

Table 2: Sentenças que compõem o contexto de sinalização no conjunto de dados *Grammatical Facial Expressions Data Set* [20]

Interrogação (qu)	Interrogação (sim/não)	Interrogação (dúvida)
Quando a Waine pagou?	Waine comprou um carro?	Waine comprou UM CARRO?
Por que a Waine pagou?	Isso é seu?	Isso é SEU?
O que é isso?	Você se formou?	Você se FORMOU?
Como faz isso?	Você gosta de mim?	Você gosta DE MIM?
Onde você mora?	Você vai embora?	Você vai EMBORA?
Negação	Afirmação	Condição
Eu não vou.	Eu vou.	Se chover, eu não vou.
Eu não fiz nada.	Eu quero.	Se você faltar, você vai perder.
Eu nunca fui preso.	Eu gosto.	Se você não quiser, ele aceita.
Eu não gosto.	Eu comprei.	Se você não comprar, ele vai querer.
Eu não tenho.	Eu trabalho lá.	Se fizer sol, eu vou para praia.
Construção relativa		
Menina que caiu de bicicleta ... ela está no hospital.		
A Universidade Unifei ... ela fica em Itabubá.		
Aquela empresa ... ela trabalha com tecnologia.		
A Waine, amiga do Lucas, é formada em Pedagogia		
A Celi, escola de surdos, fica em São Paulo		
Uso de tópico	Foco	
Universidade, eu estudo na USP.	Foi a WAINE que fez.	
Frutas, eu gosto de abacaxi.	Eu gosto de AZUL.	
Trabalho, eu trabalho com informática.	A WAINE que pagou.	
Computador, eu tenho um <i>notebook</i> .	A bicicleta QUEBROU.	
Esporte, eu gosto de vôlei.	VOCÊ que está errado.	

Facial Expressions Data Set [84]²⁵ para servir a comunidade que pesquisa tanto na área de aprendizado de máquina quanto na área de análise de gestos e língua de sinais.

2.3 Análise das fases do gesto

A principal motivação para desenvolver pesquisa em automação da análise de gestos é dar suporte à interpretação automática da comunicação humana e da interação entre humanos e agentes computacionais (sistemas de software ou hardware). Contudo, para que essa interpretação alcance a qualidade desejada, e seja tão natural quanto possível, é importante considerar os gestos como parte de um sistema linguístico, ou seja, como ações que possuem propósitos comunicativos.

Diferentes disciplinas estudam os gestos sob a perspectiva da comunicação dando origem a um campo de estudo denominado Estudos dos Gestos. Segundo a *International Society for Gesture Studies* [85], entre os tópicos tipicamente pesquisados na área de Estudos de Gestos estão: o papel e a organização dos gestos na conversação face a face e a relação do gesto com o pensamento e a linguagem. Esses dois tópicos vêm ao encontro dos estudos de Kendon [86, 38], McNeill [39, 87] e Kita [88], que fornecem os subsídios para o problema de análise das fases dos gestos objetivado pela pesquisa em discussão neste texto.

As fases dos gestos foram organizadas hierarquicamente em um *framework* proposto primeiramente por Kendon [86] e posteriormente endossado por McNeill e Kita [39, 88]. Neste *framework*, no momento em que o gesto não está em execução diz-se que as mãos (ou o corpo) estão na posição de descanso, e para organizar o gesto, dois importantes conceitos são estabelecidos: *G-unit*, ou unidade gestual, que diz respeito ao período de tempo que inicia quando, por exemplo, o braço se distancia do corpo (um movimento começa) e termina quando retorna ao que se conhece como posição de descanso; *G-phrase*, ou frase gestual, que ocorre dentro de uma unidade gestual e é composta por uma ou mais fases de movimentos. Então, cinco fases do gesto são estabelecidas, e elas estão ilustradas na figura 14:

- preparação: as mãos vão para uma posição na qual o *stroke* começa;
- pré-*stroke hold*: uma pequena pausa no fim da preparação, segurando a posição e a configuração da mão;
- *stroke*: o pico de esforço no gesto, que expressa o significado do gesto;

²⁵Freitas, F. A.; Barbosa, F. V.; PERES, S. M. *Grammatical Facial Expressions Data Set*. 2014. Disponível em <https://archive.ics.uci.edu/ml/datasets/Grammatical+Facial+Expressions>.

- *pós-stroke hold*: uma pequena pausa no fim do *stroke*, segurando a posição e a configuração da mão;
- *retração*: retorno das mãos para a posição de descanso.

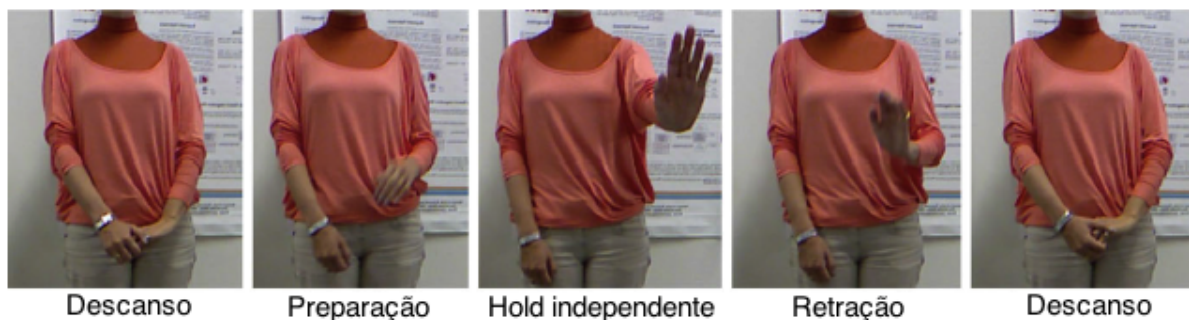
Figure 14: Quadros de um vídeo ilustrando as fases do gesto estabelecidas no *framework* de Kendon [86]. Traduzido de [89]



Além das fases originalmente estabelecidas por Kendon, Kita [88] expandiu a hierarquia usando o conceito de “fase expressiva”. Esta é uma fase na qual ocorre: um *pré-stroke hold* opcional, um *stroke* obrigatório e um *pós-stroke hold* também opcional, nessa ordem; ou um *hold* independente, que se constitui como uma pausa, ao invés de um movimento, para expressar o significado do gesto, substituindo o *stroke* na fase expressiva do gesto. A figura 15 ilustra o uso de um *hold* dentro de um gesto.

A fim de semiformalizar a hierarquia das fases do gesto, Kita [88] propôs a seguinte gramática, que insere mais alguns elementos ao *framework* de Kendon. Esta gramática

Figure 15: Quadros de um vídeo ilustrando a fase *hold* independente, estabelecida por Kita [88]. Traduzido de [89]



tem sido usada na automação da análise das fases do gesto em estudo na pesquisa aqui discutida:

$G\text{-unit} = G\text{-Phrase}^*$

$G\text{-Phrase} = (\text{Preparação}) \Rightarrow \text{Fase expressiva} \Rightarrow (\text{Retração})$

$\text{Fase expressiva} = \text{Hold independente}$

$\text{Fase expressiva} = (\text{Hold dependente}) \Rightarrow \text{Stroke} \Rightarrow (\text{Hold dependente})$

$\text{Preparação} = (\text{Liberação do movimento}) \Rightarrow \text{Preparação da localização} \gg \text{Preparação interna da mão}$

$\text{Retração (quando seguida de outra frase gestual)} = \text{Retração parcial}$

O estudo de gestos, sob os olhos de diferentes disciplinas ou áreas de aplicação (antropologia, linguística, psicologia, comunicação, música, teatro e dança), geralmente requer a análise de gravações de pessoas conversando, discutindo, discursando, dançando ou tocando um instrumento [89]²⁶. Essas análises envolvem a transcrição de diferentes aspectos dos gestos, incluindo suas fases. Esse processo de transcrição é atualmente realizado a partir do uso de softwares que permitem a manipulação do vídeo e o registro de anotações [90, 91, 92]. Porém, desde que tais softwares não dão apoio a tomada de decisão via análise automática de gestos, o custo da tarefa de análise de gestos é alto [93], principalmente quando mais de um aspecto do gesto está sendo analisado. Experiências particulares da autora com a tarefa de anotação de vídeos para fases dos gestos, que corresponde a apenas um aspecto do gesto, duram cerca de duas horas para anotar um vídeo de cerca de dois minutos. Na literatura é possível encontrar diferentes aplicações que fazem uso da análise dos gestos seguindo a visão da área de Estudos dos Gestos. Alguns exemplos são: estudo da sincronia entre fala e gestos [94]; mineração de conceitos

²⁶Madeo, R. C. B.; Lima, C. A. M.; PERES, S. M. Studies in automated hand gesture analysis: an overview of functional types and gestures phases. Language Resources and Evaluation, Springer:Verlag, p. 1-33, 2016. ONLINE FIRST.

semânticos expressos com apoio dos gestos [95]; análise do comportamento de professores e alunos durante o processo de aprendizado [96]; síntese de gesticulação para robôs [35, 97, 36].

A análise das fases do gesto, objeto de estudo da autora, é modelada como uma tarefa de segmentação de um vídeo, com base na teoria de gestos explorada por Kendon, McNeill e Kita. A resolução dessa tarefa foi proposta, no trabalho de [98]²⁷, por meio do desenvolvimento de um classificador seguindo a definição: uma sequência de quadros $\mathcal{S} = \{f_1, f_2, \dots, f_n\}$ de tamanho n , contendo imagens de uma pessoa gesticulando, é apresentada a um modelo classificador cujo objetivo é classificar um quadro de interesse nessa sequência de acordo com a fase do gesto correspondente. As fases são modeladas como classes $c_i = \{Descanso, Preparação, Stroke, Hold, Retração\}$, que correspondem a fases do *framework* de Kendon, McNeill e Kita. Ao receber sequências \mathcal{S} , o modelo classifica cada um dos quadros, fornecendo a segmentação do vídeo nas fases.

A escolha por fazer a análise do gesto usando sequências de quadros é devido à hipótese de que a informação sobre a fase do gesto tem relação com a execução do gesto, ou seja, a dinâmica de um gesto varia no tempo, e a caracterização de cada fase se daria por meio de tal variação. Então, técnicas capazes de processar a informação temporal ou representar vetorialmente a informação temporal foram escolhidas para realizar a construção dos classificadores.

Para o primeiro caso, métodos baseados em máquinas de vetores de suporte (SVM - *Support Vector Machines*) e que incorporam o processamento temporal foram estudados e aplicados ao problema [99]²⁸. Mais especificamente, SVM com *Kernel* Recursivo e *Support Vector Echo-State Machines* (SVESM) foram usados. SVM com *Kernel* Recursivo [100] usa dados anteriores (tempo $t - 1$) no cálculo do *kernel* para o dado atual (tempo t). Dessa forma, o mapeamento construído pelo kernel incorpora informação contida em uma série histórica composta pelos dados anteriores ao que está em análise. A estratégia SVESM [101] utiliza um reservatório (*reservoir*), implementado por uma rede neural com neurônios conectados de maneira esparsa e aleatória, capaz de processar informação temporal considerando estados anteriores (tempo $t - 1$) no cálculo do estado atual (tempo t).

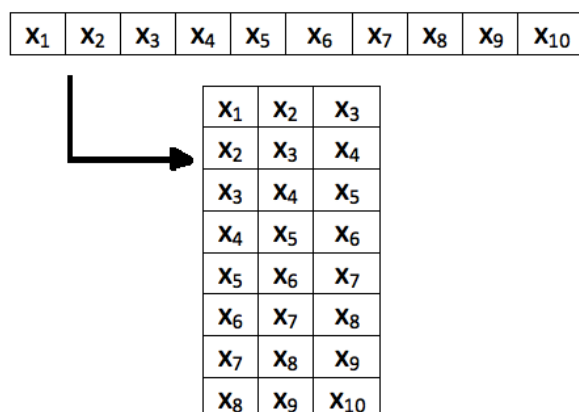
Para o segundo caso, a representação vetorial para a informação temporal é realizada a partir da aplicação de procedimentos de pré-processamento que criam um espaço de características alternativo àquele que originalmente representa um dado, o qual é capaz

²⁷Madeo, R. C. B. Máquinas de Vetores Suporte e a Análise de Gestos: incorporando aspectos temporais. 112 p. Dissertação (Mestrado) - Universidade de São Paulo, 2013. Orientação da autora.

²⁸Madeo, R. C. B.; Lima, C. A. M; **PERES, S. M.** A review on temporal reasoning using support vector machines. In: Proceedings of 19th International Symposium on Temporal Representation and Reasoning. Washington, DC, USA: IEEE Computer Society, 2012. p. 46–52.

de representar o tempo vetorialmente [102]²⁹. Dessa forma, técnicas que tratam dados independentes e identicamente distribuídos, como a SVM tradicional ou uma rede neural MLP, podem ser aplicadas. Uma técnica comum nesse contexto é a Reconstrução do Espaço de Fases (REF) – cada vetor nesse espaço corresponde a um estado de um sistema dinâmico, e vice-versa [103]. Na REF, um ponto de uma série é composto por uma função aplicada a um estado no tempo, considerando uma dimensão desejada de reconstrução (μ) e um atraso de tempo (τ). A figura 16 ilustra o conceito para $\mu = 3$ e $\tau = 1$, que na pesquisa em discussão é chamado de “uso de janelas deslizantes de tamanho μ ” ou “janelamento de dimensão μ ”.

Figure 16: Ilustração do conceito de reconstrução do espaço de fases considerando dimensão $\mu = 3$ e atraso de tempo $\tau = 1$

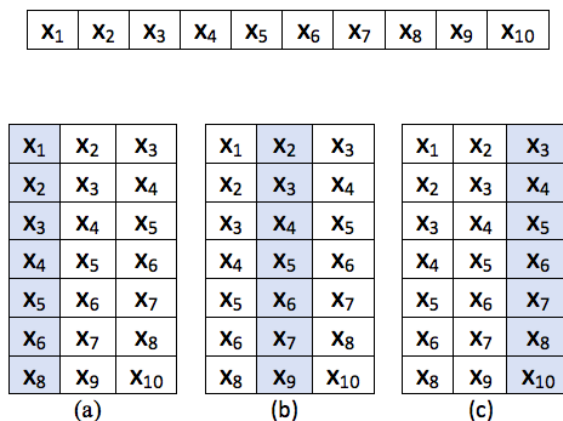


Para usar o conceito de janelamento no problema de análise de gestos, cada x na figura 16 é valorado com o conjunto de características que descrevem um quadro da sequência de vídeo \mathcal{S} . O quadro em descrição é chamado de quadro de interesse e é o objeto em avaliação pelo classificador, ou seja, o resultado do classificador dirá respeito à fase do gesto em execução no tempo de vídeo desse quadro. A figura 17 mostra como são construídas as janelas a depender da localização determinada para o quadro de interesse. O quadro de interesse pode estar no início da janela (o futuro é considerado na representação da informação temporal relacionado ao quadro), no meio da janela (informação do passado e do futuro são consideradas), e no fim da janela (apenas informação do passado é considerada).

As características usadas na proposta de análise em discussão, extraídas para cada um dos quadros do vídeo e de cada ponto de interesse nesses quadros, são: coordenadas espaciais (x, y, z) , velocidade e aceleração vetoriais e escalares, ou características no domínio do tempo, no domínio da frequência e no domínio do tempo-frequência. Assim, como exem-

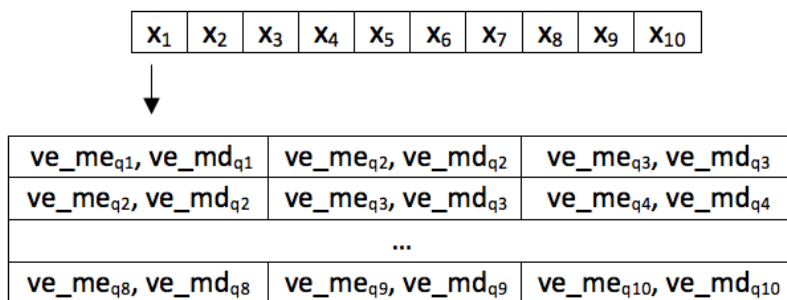
²⁹Madeo, R. C. B.; **PERES, S. M.**; Lima, C. A. M. Overview on Support Vector Machines applied to temporal modeling. In: Anais do IX Encontro Nacional de Inteligência Artificial. 2012. p. 1–6.

Figure 17: Representação dos quadros de interesse: (a) o quadro de interesse é no início da janela, a representação não é capaz de representar alguns quadros no final do vídeo; (b) quadros do início e do fim do vídeo não são representados quando o quadro de interesse é no meio da janela; (c) o quadro de interesse é no fim da janela e assim alguns quadros no início do vídeo não são representados



plo da representação vetorial do tempo aplicada para dados gestuais, segue a ilustração da figura 18.

Figure 18: Representação vetorial do tempo usando janelamento. Do quadro (q_i) são extraídas a velocidade escalar da mão esquerda (ve_me) e a velocidade escalar da mão direita (ve_md). Assumindo um janelamento de dimensão 3 e essas duas características tem-se uma representação vetorial de 6 dimensões para o gesto ($\vec{x}_q = [ve_me_{q1}, ve_md_{q1}, ve_me_{q2}, ve_md_{q2}, ve_me_{q3}, ve_md_{q3}]$)



Desde que a técnica escolhida para resolução do problema de segmentação das fases do gesto é preparada para implementar classificadores binários, estratégias “um contra todos” ou “um contra um” podem ser adotadas. Adicionalmente, como há ainda a possibilidade de trabalhar com o problema seguindo uma estratégia “dividir-para-conquistar”, uma terceira maneira de tratar o problema multiclasse a partir de classificadores binários é possível. As três abordagens foram estudadas no trabalho de [98]. Na terceira estratégia, o problema foi dividido em problemas menores da seguinte maneira:

1. Segmentação de fases de descanso: dada uma entrada $q_i \in \mathcal{S}$, o classificador binário terá como possíveis saídas $c_i = \{\textit{Descanso}, \textit{Unidade Gestual}\}$, em que $\textit{Unidade Gestual} \subset \{\textit{Preparação}, \textit{Stroke}, \textit{Hold}, \textit{Retração}\}$;
2. Segmentação de fases *hold*: dada uma entrada $q_i \in \mathcal{S}_{\textit{Unidade Gestual}}$ em que $\mathcal{S}_{\textit{Unidade Gestual}}$ é o subconjunto de quadros de \mathcal{S} classificados como *Unidade Gestual* pela etapa de segmentação do descanso, o classificador binário terá como possíveis saídas $c_i = \{\textit{Hold}, \textit{Frase Gestual}\}$, em que $\textit{Frase Gestual} \subset \{\textit{Preparação}, \textit{Stroke}, \textit{Retração}\}$;
3. Segmentação de fases *stroke*: dada uma entrada $q_i \in \mathcal{S}_{\textit{Frase Gestual}}$ em que $\mathcal{S}_{\textit{Frase Gestual}}$ é o subconjunto de quadros de \mathcal{S} classificados como *Frase Gestual* pela etapa de segmentação do *hold*, o classificador binário terá como possíveis saídas $c_i = \{\textit{Stroke}, \textit{Transição}\}$, em que $\textit{Transição} \subset \{\textit{Preparação}, \textit{Retração}\}$ e correspondem às fases que podem ser vistas como a transição entre o descanso e a fase expressiva do gesto;
4. Segmentação de fases de preparação e retração: dada uma entrada $q_i \in \mathcal{S}_{\textit{Transição}}$ em que $\mathcal{S}_{\textit{Transição}}$ é o subconjunto de quadros de \mathcal{S} classificados como *Transição* pela etapa de segmentação do *stroke*, o classificador binário terá como possíveis saídas $c_i = \{\textit{Preparação}, \textit{Retração}\}$.

Os estudos realizados pela autora consideraram como descritores dos dados gestuais as coordenadas (x, y, z) de mãos e pulsos e as velocidades e acelerações vetoriais e escalares extraídas a partir das coordenadas considerando uma janela de três quadros de vídeos. Além disso, características no domínio do tempo, frequência e tempo-frequência foram exploradas. Os melhores resultados foram obtidos com o uso de velocidades e acelerações e com a combinação dessas características com coordenadas espaciais.

O uso das coordenadas espaciais na representação vetorial do gesto facilita o reconhecimento dos padrões das fases, principalmente quando a posição de descanso é frequentemente executada na mesma região do espaço. Contudo, essa estratégia leva à construção de modelos de reconhecimento com menos poder de generalização: se forem testados a partir de dados de um usuário que possui o padrão de posição de descanso em uma região do espaço diferente, o classificador perderá desempenho em uma escala maior do que os classificadores construídos sobre a representação vetorial sem coordenadas espaciais.

Os resultados obtidos com a aplicação de SVM, janelamento e representação baseada em velocidade, aceleração e características no domínio do tempo, frequência e tempo frequência, considerando cinco fases do gesto, estão documentados em [98, 104]³⁰, e os

³⁰Madeo, R. C. B.; **PERES, S. M.**; Lima, C. A. M. Gesture phase segmentation using Support Vector Machines. Expert Systems with Applications, Elsevier, V. 56, p. 100–115, 2016.

resultados referentes ao uso de MLP, janelamento e representação baseada em velocidade, aceleração e coordenadas espaciais, e considerando apenas descanso e unidades gestuais, estão documentados em [105]³¹. As tabelas 3 e 4 organizam alguns resultados obtidos a partir de uma avaliação quantitativa do desempenho dos modelos classificadores analisados.

Table 3: Resultados em termos de F-score e matriz de confusão referentes ao desempenho de classificadores SVM (abordagens: um contra um, um contra todos e hierárquica) na tarefa de segmentação das fases do gesto. Representação de dados em janelas, considerando velocidades e acelerações escalares das mãos e pulsos. [98, 104]

Fase do gesto classe esperada	F-score	Matriz de confusão					F-score	Matriz de confusão				
	<i>um contra um</i>						<i>um contra todos</i>					
<i>Descanso</i>	0,78	413	8	15	1	33	0,81	427	6	25	0	1
<i>Preparação</i>	0,56	49	103	39	0	9	0,54	49	84	55	1	22
<i>Stroke</i>	0,65	82	49	251	0	49	0,69	78	22	286	4	41
<i>Hold</i>	0,28	32	0	12	10	0	0,29	30	0	3	8	1
<i>Retração</i>	0,33	8	4	18	5	32	0,29	18	0	24	0	22
Fase do gesto classe esperada	F-score	–					F-score	Matriz de confusão				
	<i>classificadores individuais</i>						<i>abordagem hierárquica</i>					
<i>Descanso</i>	0,87						0,84	445	8	21	7	8
<i>Preparação</i>	0,79						0,55	24	91	75	0	10
<i>Stroke</i>	0,78						0,75	35	18	336	5	37
<i>Hold</i>	0,57						0,33	34	3	4	13	0
<i>Retração</i>	–						0,22	29	9	30	0	18

Table 4: Resultados em termos de F-score e matriz de confusão referentes ao desempenho de classificadores MLP, na tarefa de segmentação de unidade gestual. Representação de dados em janelas, considerando coordenadas espaciais, velocidades e acelerações escalares das mãos e pulsos. Experimentos executados para diferentes vídeos, com dependência de gesticulador e envolvendo dois gesticuladores [105]

Vídeo	F-Score	Vídeo	F-Score
A1	0,93	B3	0,75
A2	0,82	C1	0,81
A3	0,87	C3	0,82
B1	0,38	A1/A2	0,87

Apenas a análise de desempenho do classificador não é suficiente para entender se a abordagem automática de segmentação das fases do gesto pode ou não ser útil para a área de estudos dos gestos. Algumas questões adicionais são interessantes nesse contexto:

1. O padrão gestual é replicável para diferentes usuários?

³¹Wagner P. K.; Madeo, R. C. B; Lima, C. A. M.; **PERES, S. M.** Segmentação de unidades gestuais com Multilayer Perceptron. In: Anais do X Encontro Nacional de Inteligência Artificial. 2013.

2. O padrão gestual é generalizável para contextos diferentes?
3. É necessário considerar os movimentos das duas mãos, ou apenas a mão dominante é suficiente?
4. O conjunto de fases de gestos estabelecidas por Kendon e McNeill são suficientes (ou necessárias) para a análise do padrão de fases de gesto?
5. A generalização obtida é suficiente para implementação de uma ferramenta computacional de apoio ao trabalho de análise dos gestos?

Os trabalhos desenvolvidos até o momento de escrita deste texto permitem delinear algumas conclusões para as questões 1 e 5. No que diz respeito a generalização do modelo para diferentes usuários, foi possível constatar que, considerando as condições estabelecidas para os experimentos realizados, os modelos de segmentação apresentam um desempenho razoável, mas que necessita de melhora. Isso mostra que para construção de uma ferramenta computacional de suporte ao trabalho de segmentação, seria interessante considerar a possibilidade de disponibilização da funcionalidade de geradores de modelos (ao invés de embutir um gerador universal na ferramenta), os quais usariam um trecho inicial de um vídeo de um usuário, rotulado por um especialista, para criar modelos orientados a usuários.

Entretanto, há que se considerar também que os estudos realizados constataram que existem dois tipos de erros cometidos pelos classificadores: o erro interno e o erro de borda. O erro de borda é aquele erro de classificação que ocorre em quadros que estão próximos à transição entre fases. Os erros internos são aqueles que ocorrem em quadros distantes da transição. Do ponto de vista do especialista de análise de gestos, o quadro exato no qual termina ou começa uma fase é uma questão de alta subjetividade e que não possui criticidade para o seu trabalho. O importante para o especialista é o reconhecimento da existência de uma fase (a correta). Ao considerar esta visão na avaliação do modelo dos classificadores, tanto SVM quanto MLP, há uma melhora de avaliação que motiva o uso desses classificadores para construção de ferramentas de apoio à segmentação das fases do gesto. Para ilustrar essa questão, [98] propôs avaliar o desempenho dos classificadores usando uma análise por segmento. A tabela 5 ilustra um cenário de análise por segmento apresentado em [104]. As regras usadas para considerar a corretude de reconhecimento de um segmento (uma fase) são descritas em [98, 104].

As questões 2, 3 e 4 supracitadas estão atualmente sendo estudadas pela autora.

Conjunto de dados - *Gesture Phase Segmentation Data Set*

Table 5: Análise por segmentos para avaliação de um classificador gerado para segmentação das fases do gesto. Resultados apresentados em termos de uma matriz de confusão [104]

Classe Real	Classe Predita				
	Descanso	Preparação	<i>Stroke</i>	<i>Hold</i>	Retração
Descanso	7	1	0	1	0
Preparação	1	8	5	0	0
Stroke	2	0	11	0	0
Hold	2	0	1	3	0
Retração	2	1	3	0	2

O conjunto de dados *Gesture Phase Segmentation Data Set* [106]³² é composto por sete vídeos gravados a partir de um sensor Microsoft Kinect. O conteúdo dos vídeos diz respeito à contação de histórias realizadas por três pessoas a partir de três histórias em quadrinhos. O conjunto de dados está disponível publicamente no repositório *UCI - Machine Learning Repository* [63].

Os arquivos texto que compõem o conjunto de dados contêm:

- informação das coordenadas espaciais x, y, z de seis pontos de interesse no corpo dos gesticuladores – mão direita, mão esquerda, pulso direito, pulso esquerdo, cabeça e tronco;
- informação sobre velocidade e aceleração, vetorial e escalar, de cada um dos pontos de interesse, calculadas a partir das coordenadas espaciais e considerando uma janela de tempo de deslocamento baseada em uma quantidade de quadros do vídeo.

Cada linha dos arquivos corresponde a um quadro do vídeo capturado pelo sensor e cada arquivo corresponde a uma combinação de usuário e história. O conjunto de dados está rotulado com as cinco fases do gesto: descanso, preparação, *stroke*, *hold* e retração. A tabela 6 mostra informações que permitem analisar a complexidade do conjunto de dados.

Uma complementação a esse conjunto de dados foi disponibilizada em www.each.usp.br/sarajane. Essa complementação conta com os dados referentes aos vídeos B2 e C2, além de uma série de rotulações (4 rotulações) realizadas para os nove vídeos. A partir das quatro rotulações, análises de concordância podem ser feitas para: permitir que a subjetividade do trabalho de rotulação seja considerada na análise do desempenho do classificador; permitir criar uma rotulação final a partir da combinação das rotulações individuais.

³²Madeo, R. C B., Wagner, P. K., PERES, S. M. *Gestures Phase Segmentation Data Set, 2014* <https://archive.ics.uci.edu/ml/datasets/Gesture+Phase+Segmentation>

Table 6: Informações sobre o conjunto de dados *Gesture Phase Segmentation Data Set*

Id.	Usuário	História	Tempo (segundos)	# Quadros	# Quadros por fase do gesto				
					D	P	S	H	R
A1	A	1	58	1747	698	163	656	39	191
A2	A	2	42	1264	493	200	431	54	86
A3	A	3	61	1834	662	279	535	150	208
B1	B	1	36	1073	74	412	287	217	83
B3	B	3	48	1424	194	469	390	201	170
C1	C	1	37	1111	286	236	262	193	134
C3	C	3	48	1448	362	338	389	144	215

2.4 Objetivos atuais e limitações da pesquisa

Diante dos estudos realizados pela autora, principalmente considerando os resultados obtidos com o reconhecimento de configurações, movimentos e direção da mão no contexto da Libras, observa-se que o reconhecimento de padrões para gestos dentro de um vocabulário finito pode alcançar alta eficiência, sendo possível criar modelos com boa capacidade de generalização a partir do uso de arquiteturas que combinem heurísticas, Inteligência Computacional e Aprendizado de Máquina. Nesse quesito, como questão de pesquisa, vislumbra-se apenas a necessidade de lidar com a escala com a qual se trabalha em termos de tamanho do vocabulário. Trabalhar com vocabulários grandes não faz parte dos interesses de pesquisa atuais da autora. Sendo assim, as pretensões de continuidade do trabalho junto a vocabulários finitos de gestos estão localizadas no desenvolvimento de aplicações, com o aprimoramento do Jogo da Força em Libras, e na produção de inovação tecnológica a partir da proposição de outros tipos de aplicações em temas similares.

No que diz respeito ao reconhecimento de expressões faciais gramaticais, os estudos realizados ainda são preliminares. Tendo sido, o estudo realizado, pioneiro para o caso da Libras, foi necessário realizá-lo com perspectivas preparatória e exploratória. Perspectiva preparatória porque não havia, até então, um conjunto de dados referente ao objeto de reconhecimento, que permitisse concentrar os esforços de estudo no problema de reconhecimento da expressão. Havia, e há, obviamente, muitos vídeos com falas em Libras disponíveis, e eventualmente, alguns vídeos anotados em relação à expressão facial gramatical. No entanto, usar esses vídeos implica em também resolver o problema de segmentação da face humana, em ambientes não controlados, com variações de iluminação e posicionamento, e problemas de oclusão. Como uma das principais contribuições dos esforços desta pesquisa está a criação e disponibilização de um conjunto de dados no contexto de expressões faciais gramaticais da Libras. Já a perspectiva exploratória é devido ao fato que, como o estudo realizado foi o primeiro na área em Libras, foi necessário despender algum tempo para compreender quais são as nuances do problema, ou seja, qual é a complexidade inerente à análise automática de cada uma das expressões faciais

gramaticais.

A partir do ponto atual da pesquisa em expressões faciais gramaticais da Libras, há alguns objetivos sendo buscados por meio de estudo atualmente em desenvolvimento. Esses objetivos incluem:

- Extensão do conjunto de dados *Grammatical Facial Expressions Data Set* de forma que ele contenha trechos de fala em Libras que incorporem várias expressões faciais gramaticais usadas de maneira natural. Ainda, a versão atual do conjunto de dados deve passar por uma revisão para receber rótulos novos. Há algumas sentenças nesse conjunto que embora tenham sido especialmente preparadas para o estudo de uma determinada expressão facial, incorporam também uma segunda expressão facial. Esses casos podem ser devidamente rotulados para suportar o estudo de classificadores multiclasse.
- Exploração de estratégia de comitê de máquinas (*ensembles* ou mistura de especialistas) para implementar arquiteturas eficientes para o reconhecimento de expressões faciais gramaticais modelado como um problema de classificação multiclasse. Uma das dificuldades encontradas na resolução desse problema é que expressões faciais gramaticais diferentes são mais bem analisadas por modelos gerados a partir de configurações de parâmetros específicas e com representação vetorial em janelas de tamanhos distintos. Com o uso da arquitetura, é pretendido combinar os diferentes modelos que resolvem bem o problema para diferentes expressões, e ainda automatizar, por meio de um último modelo de Aprendizado de Máquina, a escolha sobre qual dos modelos especialistas usar e em que momento do trabalho de reconhecimento.
- Estender todo o estudo realizado, e por realizar, para expressões faciais morfológicas.

Há ainda a intenção de trabalhar com a questão da fluência em Libras e com o suporte ao reconhecimento de problemas da fala em Libras. As expressões faciais gramaticais são consideradas como uma das características que define o nível de fluência de um usuário da Libras. Sendo assim, é possível colocar o objetivo de análise de fluência como um problema que poderia ser resolvido com o uso de modelos classificadores, os quais, na realidade, poderiam atuar como uma ferramenta de avaliação da execução das expressões faciais. Contudo, para que isso seja possível, os modelos de reconhecimento precisam alcançar um nível de eficiência alto. Da mesma forma, a análise das expressões faciais gramaticais pode revelar, ou ajudar a diagnosticar, distúrbios da fala, e assim, os mesmos modelos de reconhecimento poderiam ser usados com o fim de suporte ao diagnóstico desses distúrbios.

Na área de segmentação automática das fases dos gestos, a autora entende ter contribuído, de forma mais proeminente, de duas maneiras:

1. Construindo, e disponibilizando publicamente, um conjunto de dados especialmente preparado para ser a base de exploração de algoritmos para segmentação das fases dos gestos. A segmentação das fases do gesto é ainda muito pouco explorada, e até a construção do conjunto *Gesture Phase Segmentation Data Set* não havia nenhum conjunto de dados rotulado com cinco fases do gesto. Com a apresentação desse conjunto de dados, vários pesquisadores de análise de dados puderam ter contato com esse problema e puderam iniciar estudos em relação a ele.
2. Chamando a atenção para análises diferenciadas nesse contexto, como a análise por segmento, interessante para pesquisadores da área de estudos do gesto, e a análise considerando as variações nos gestos consequentes das variações naturais do comportamento humano.

A partir do que já foi realizado pela autora na área de segmentação automática das fases dos gestos, novos objetivos foram traçados, os quais quando alcançados permitirão responder as perguntas 2, 3 e 4, levantadas no final da seção 2.3, e também abrirão caminhos para dar início ao desenvolvimento de aplicações correlatas. Esses objetivos incluem:

- Estudo da influência do contexto do discurso e da conversação natural no desempenho dos classificadores que realizam a segmentação das fases do gesto. Esse estudo se dá com base na análise do desempenho de classificadores sobre gestos proveniente do mesmo usuário, porém executados em momentos diferentes e para objetivos diferentes. (*trabalho em andamento*)
- Exploração de diferentes descritores dos gestos para composição da representação vetorial usada na construção dos classificadores que realizam a segmentação das fases do gesto. (*trabalho em andamento*)
- Modelagem do problema de segmentação das fases do gesto como um problema de agrupamento, de forma que não seja necessário criar modelos guiados por uma estrutura de fases já preestabelecida. (*trabalho em andamento*)
- Exploração dos problemas de reconhecimento de gestos suportados pelo *corpus Natural Media Motion-Capture Corpus*. (*trabalho em andamento*)
- Aprofundamento dos estudos em outras teorias da área de estudos de gestos para identificar possíveis nichos de atuação para análise automática de gestos.

A pesquisa aqui discutida, nos três diferentes escopos, sofre de uma limitação referente aos ambientes de experimentação. Todos os resultados obtidos estão pautados em experimentos executados em *corpora* que não podem ser considerados como uma amostra robusta do comportamento humano. No entanto, essa é uma limitação comum em pesquisas que tangenciam a análise do comportamento humano.

3 Contribuições em análise de textos

A forma mais natural de persistir a informação é por meio da escrita. A humanidade tem na escrita uma ferramenta eficiente de fazer com que conhecimento, ideias e ideais sejam perpetuados, organizados na forma de textos. Trazendo para esse cenário o advento da criação e posterior popularização de computadores, e a evolução do seu uso com o desenvolvimento de mecanismos de produção, armazenamento e disseminação de artefatos digitais, tem-se um ambiente que favorece a produção de conteúdo digital, com destaque para o conteúdo digital textual.

Atualmente, a produção de conteúdo digital (não apenas, texto, mas também imagens, vídeos e som), e disponibilização dele na internet, tem como consequência um volume muito grande de informação que, com certeza, não tem sido usada com eficiência e eficácia. A real disponibilização do conteúdo digital textual existente no mundo ainda carece de mecanismos eficientes de organização e indexação, e de algoritmos que possam fazer uso dos dados/informação nele existente de maneira a orientar e bem informar o consumidor da informação. Obviamente que muito já se desenvolveu nas áreas de recuperação de informação e mineração e análise de textos, mas sem dúvida ainda há espaço para que sistemas computacionais sirvam os consumidores da informação textual com conteúdo mais sofisticado, em termos de provimento de informação, do que é atualmente oferecido.

Este capítulo é destinado a apresentar a pesquisa que vem sendo realizada pela autora na área de análise de textos. Esta é uma linha de pesquisa recente para a autora e está ainda em formação no que diz respeito às estratégias de pesquisa. No entanto, é claro para a autora que o principal objetivo nessa linha de pesquisa é explorar as potencialidades das técnicas de Inteligência Computacional e Aprendizado de Máquina, sob a linha de resolução de tarefas de Mineração de Dados (principalmente a tarefa de agrupamento), para melhorar a qualidade da informação que é possível descobrir de forma automática a partir de um conjunto de textos.

O desenvolvimento da automação que seja capaz de interpretar conteúdo textual e recuperar ou descobrir informação a partir dessa interpretação remete a, pelo menos, duas linhas de trabalho, a linguística e a baseada em contagem (ou estatística). A primeira se preocupa em analisar aspectos sintáticos e semânticos dos textos [107], eventualmente com apoio de representações computacionais que usam contagem de palavras. Inicial-

mente, não havia muito otimismo em relação a essa abordagem devido à necessidade de uma formalização de alto nível para os aspectos de automação da análise linguística [108], contudo, atualmente tem havido esforços nessa linha que vêm ganhando a atenção da comunidade acadêmica [109, 110]. A segunda é baseada na contagem de palavras existentes nos textos [111, 112], e é a linha mais amplamente usada dentro do contexto de Mineração de Dados [113]. A segunda linha é a adotada até o presente nos esforços de pesquisa da autora.

O presente capítulo está organizado de forma a introduzir os conceitos que embasam as hipóteses de pesquisa da autora, formuladas para o desenvolvimento dos trabalhos na área de análise de dados textuais. Tais hipóteses estão focadas na resolução da tarefa Mineração de Dados formulada como um problema de coagrupamento como uma forma de obter resultados de análise textual relevantes e que possuam serendipidade. Os primeiros resultados obtidos dentro desta linha de pesquisa são apresentados e o capítulo é finalizado com a discussão da área de aplicação almejada para o estudo em desenvolvimento: a área de sistemas de recomendação baseados em conteúdo textual. A produção científica obtida nesta linha de pesquisa, ainda inicial, é citada no texto e descrita em detalhes em rodapés. O capítulo é finalizado com a apresentação dos objetivos que estão no centro das atenções da pesquisa atualmente e algumas limitações associadas ao tema pesquisado.

3.1 Textos, pré-processamento e representações

Seguindo a notação apresentada em [2], um *corpus* Ψ é considerado um conjunto de documentos³³ (*doc*), ou seja, $\Psi = \{\vec{doc}_1, \vec{doc}_2, \dots, \vec{doc}_n\}$. Cada documento é representado por um conjunto de d termos (*te*) (por exemplo: radicais, palavras, sintagmas ou N-gramas), de forma que $\vec{doc}_i = (wte_{i1}, wte_{i2}, \dots, wte_{ij}, \dots, wte_{id})$, sendo que $i = (1, \dots, n)$, $j = (1, \dots, d)$ e wte_{ij} assume um valor referente ao peso do termo j para o documento i . Esta representação constitui-se como um modelo vetorial, conhecido como modelo de espaço vetorial (do inglês *Space Vector Model*), no qual cada eixo do espaço vetorial representa um termo de um dicionário de termos estabelecido para o *corpus* Ψ e o peso wte é a coordenada que localiza a importância do termo para um documento [114]. Tal representação é popularmente conhecida como *set-of-words* ou *bag-of-words*, a depender se o peso está limitado ao conjunto $\{0, 1\}$ ou se pode assumir valores no conjunto \mathfrak{R} [115].

O trabalho com dados textuais demanda a execução de procedimentos de pré-processamento que possam prepará-los para serem adequadamente representados como especificado nesta seção. De fato, a execução dos procedimentos de pré-processamento e posterior representação dos textos dentro de um modelo de espaço vetorial pode ser visto como um

³³Doravante neste capítulo, a palavra “documento” será usada para fazer referência a um item de dado não estruturado e do tipo texto.

processo de transformação de um *corpus* Ψ em um conjunto de dados \mathcal{X} que pode ser processado por técnicas de Inteligência Computacional e Aprendizado de Máquina, como feito na pesquisa aqui descrita. Esse processo envolve procedimentos de separação de *tokens*, eliminação de termos irrelevantes para o contexto (aplicação de listas de *stop-words*), redução de termos ao seu radical (*stemming* [116]) ou ao seu lema ou verbete [117], e criação da representação vetorial [2]. Especificamente sobre os procedimentos de eliminação de termos irrelevantes e redução de termos ao seu radical, faz-se interessante dizer que embora opcionais no processo, eles são classicamente vistos como procedimentos de redução de dimensionalidade, e conseqüentemente, redução de esparsidade, sendo, portanto, altamente desejáveis. Segundo [111], a eliminação de termos irrelevantes pode reduzir o tamanho de um documento em até 50%.

Para determinar os valores que devem entrar nas variáveis wte_{ij} , as opções mais frequentemente usadas na literatura são:

- representação binária: wte_{ij} assume valores do conjunto $\{0, 1\}$ sendo que 0 indica que o termo j não aparece no documento i e 1 indica que o termo j aparece no documento i ;
- representação tf : wte_{ij} assume valores no conjunto \mathfrak{R} de acordo com a frequência do termo j no documento i ;
- representação $tfidf$: wte_{ij} assume valores no conjunto \mathfrak{R} de acordo com a relação $tf(doc_i, te_j) * \log \frac{n}{nt}$, sendo n o número de documentos no *corpus* Ψ e nt o número de documentos no *corpus* Ψ em que o te_j aparece.

3.2 Análise de texto via agrupamento e coagrupamento

Os primeiros trabalhos desenvolvidos pela autora na área de análise de texto foram desenvolvidos como trabalhos de conclusão de curso e iniciações científicas, no nível da graduação. Desde essas iniciativas, a tarefa de agrupamento de texto [115, 118] tem sido o foco da pesquisa, com motivação no fato que a resolução dela confere à análise de texto uma maior liberdade de produção de conhecimento do que a resolução das tarefas de classificação (tarefas estudadas pela autora mais recentemente no contexto de orientações de trabalhos de conclusão de curso). A estratégia atualmente em uso na pesquisa aqui descrita é baseada em algoritmos de agrupamento por particionamento – *k-Means* e variações [119, 120, 121, 122], e baseados em modelos neurais – *Self Organizing Maps*, SOM [57, 12]. A primeira publicação nesta área foi resultante de um trabalho aplicado à análise de textos para produção de conhecimento de perfis de trabalhos de um curso de

especialização [123]³⁴.

Os principais esforços da autora na área de análise de textos estão representados pelo trabalho referente à resolução da tarefa de coagrupamento de textos [124, 125, 126, 127]. Nesta linha de pesquisa, há o intuito de estudar o efeito da análise de similaridades parciais dos dados textuais na descoberta de conhecimento novo sobre os textos. A fim de trabalhar nesse contexto, as tarefas de agrupamento e coagrupamento foram primeiramente estudadas como problemas formulados sob a teoria de fatoração de matrizes [124, 128, 125], e estratégias inéditas foram desenvolvidas no contexto de uma dissertação de mestrado [22]³⁵. Atualmente, estratégias heurísticas e meta-heurísticas estão também sendo aplicadas pela autora para trabalhar com coagrupamento de textos [129, 130].

Formalmente, assumindo uma estratégia de agrupamento por particionamento, como feito em [22], tem-se o conjunto de dados \mathcal{X} , representado por uma matriz $X \in \mathfrak{R}^{n \times m}$, como entrada em um algoritmo que resolverá o problema de agrupamento. Na matriz X , n é o número de linhas e corresponde ao número de instâncias no conjunto de dados \mathcal{X} , e m é o número de colunas e corresponde ao número de características (ou termos) no conjunto de dados \mathcal{X} . Ainda, a matriz X pode ser vista como um conjunto de n vetores linha $\mathcal{N} = \{\vec{x}_{1,j}, \dots, \vec{x}_{n,j}\}$, com $j = \{1, \dots, m\}$. O problema de agrupamento é resolvido a partir da determinação de uma partição com k partes de \mathcal{N} , formada por subconjuntos $\mathcal{K}_p \subseteq \mathcal{N}$, sendo $p \in \{1, \dots, k\}$. A solução para o problema de agrupamento, portanto, é o conjunto $\mathcal{K}_a = \{\mathcal{K}_1, \dots, \mathcal{K}_k\}$, que doravante neste capítulo será denominado conjunto de grupos de linhas.

Seguindo a mesma linha de trabalho, o problema de coagrupamento é considerado sob a visão de que se busca encontrar submatrizes de uma matriz de dados [131]. Formalmente, assumindo uma estratégia de coagrupamento por particionamento, como feito em [22], tem-se o conjunto de dados \mathcal{X} , representado pela matriz $X \in \mathfrak{R}^{n \times m}$, como entrada em um algoritmo que resolverá o problema de coagrupamento. Nesse caso, a matriz X pode ser vista como um conjunto de n vetores linhas $\mathcal{N} = \{\vec{x}_{1,j}, \dots, \vec{x}_{n,j}\}$, com $j = \{1, \dots, m\}$, e um conjunto de m vetores coluna $\mathcal{M} = \{\vec{x}_{i,1}, \dots, \vec{x}_{i,m}\}$, com $i = \{1, \dots, n\}$. O problema de coagrupamento é resolvido a partir da determinação de cogrupos, que são submatrizes de X denotadas por $X_{\mathcal{K}_p, \mathcal{L}_q}$, em que \mathcal{K}_p e \mathcal{L}_q são, respectivamente, os k subconjuntos $\mathcal{K}_p \subseteq \mathcal{N}$ e os l subconjuntos $\mathcal{L}_q \subseteq \mathcal{M}$, sendo $p \in \{1, \dots, k\}$ e $q \in \{1, \dots, l\}$, e que doravante neste capítulo serão denominados grupo de linha e grupo de colunas. Assim, $X_{\mathcal{K}_p, \mathcal{L}_q}$ é

³⁴Silva, B. B. C.; **PERES, S. M.**; Letti, B. C.; Brunialti, L. F.; Fantinato, M.; Lima, C. A. M. Análise de agrupamento com SOM aplicada a textos para construção de perfis de trabalhos de conclusão de curso. In: *Anais 12o. Congresso Brasileiro de Inteligência Computacional - CBIC 2015*. Curitiba-PR, 2015. p. 1-6.

³⁵Brunialti, L. F. Fatoração de matrizes no problema de coagrupamento com sobreposição de colunas. 122 p. Tese (Mestrado) - Universidade de São Paulo, 2016. Orientação da autora com a colaboração do Prof. Dr. Valdinei Freire da Silva.

um grupo de dados em \mathcal{K}_p perante as características em \mathcal{L}_q , ou seja, é um cogrupos. A solução para o problema de coagrupamento, portanto, é um conjunto K_{co} formado por $k \times l$ cogrupos $X_{\mathcal{K}_p, \mathcal{L}_q}$.

Algumas estruturas possíveis para organização de cogrupos são mostradas na figura 19. Nessa visão, os cogrupos de interesse estão representados em tons de cinza. A rigor, seguindo a definição de coagrupamento por particionamento apresentada neste texto, com exceção da terceira organização mostrada na figura, todos as demais apresentam mais cogrupos do que os que estão em destaque. Por exemplo, na segunda organização, pode-se verificar seis possíveis cogrupos sem sobreposições, como ilustrado na figura 20. Nessa figura, considere as marcações **A**, **B** e **C** como grupos de colunas, e as marcações **1**, **2** e **3** como grupos de linhas. Assim, duas interpretações de cogrupos possíveis presentes na organização, considerando coagrupamento por particionamento, são: $X_{1, A}$, $X_{2, BC}$, $X_{3, AC}$, $X_{4, B}$, $X_{5, AB}$ e $X_{6, C}$, ou $X_{1, A}$, $X_{2, 3, A}$, $X_{3, 1, 2, B}$, $X_{4, B}$, $X_{5, 1, 2, C}$ e $X_{6, C}$.

Figure 19: Organização de cogrupos. O número de cogrupos de interesse nas organizações apresentadas são, respectivamente: 1, 3, 9, 3, 3, 4, 4, 5, 4. As organizações permitem sobreposição de linhas e/ou colunas entre os cogrupos, e cogrupos contidos em cogrupos [131]

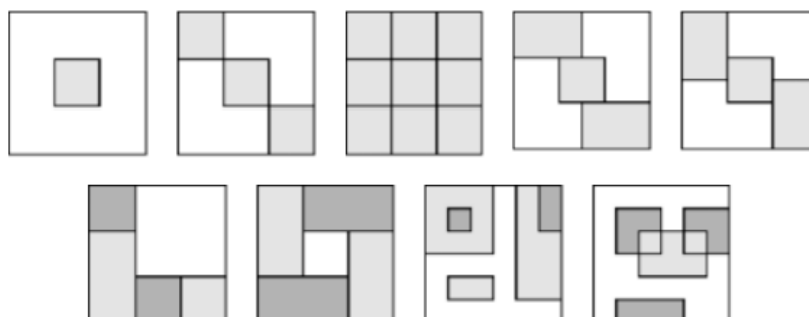
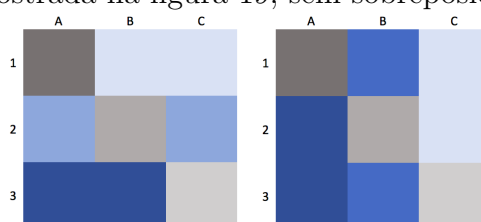


Figure 20: Interpretações de cogrupos como particionamento do conjunto de dados para a segunda organização mostrada na figura 19, sem sobreposições de linhas ou colunas

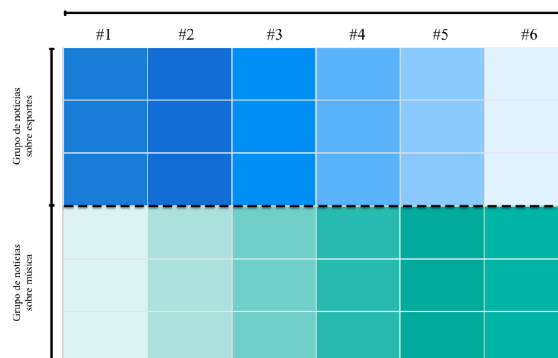


A principal base de pesquisa da autora em relação a análise de dados textuais está no fato que em um *corpus* existem textos “diferentes em algum grau” mas que tratam de assuntos correlatos. Ou seja, ainda que dois textos não sejam similares em todo o seu conteúdo, eles podem tratar de temas que fazem parte de um mesmo contexto de

interesse. Assim, tratar a análise desses textos a partir de abordagens que os comparam em sua totalidade pode não ser eficiente para trazer à tona tais relações existentes em seus conteúdos. Por isso, a autora atua na análise de texto a partir de estratégias que, por princípio, concordam com a ideia base da pesquisa, o que significa incorporar o conceito de que similaridades parciais entre elementos são úteis para suportar tomadas de decisão envolvendo interpretação textual. Para ilustrar essa ideia, [22] apresentou alguns elementos gráficos (Figura 21 e 22) que seguem reproduzidos e resumidamente discutidos aqui.

Na figura 21, o princípio de agrupamento de dados, que se baseia na análise da similaridade total, é aplicado. Nesse exemplo, dois grupos de textos são descobertos por um algoritmo hipotético capaz de resolver a tarefa de agrupamento clássico. O primeiro grupo (em azul) é formado por três textos (notícias) referentes ao tema “esporte”, e o segundo grupo (em verde) é formado por três textos referentes ao tema “música”. Os grupos foram formados a partir da análise da similaridade entre eles considerando seis atributos, e na figura, cores mais escuras indicam atributos que influenciaram com maior força no resultado da análise de similaridade. Em uma análise mais detalhada, poder-se-ia dizer que os atributos #1, #2 e #3 influenciaram com mais intensidade a formação do grupo de textos sobre esportes e que os atributos #6, #5 e #4 influenciaram com mais intensidade a formação do grupo de textos sobre música. Embora tal análise seja possível, e o algoritmo de agrupamento tenha usado a informação sobre ela intrinsecamente em suas operações, ela não é facilmente extraída do resultado e, portanto, não é usada para produzir conhecimento diferenciado sobre os grupos.

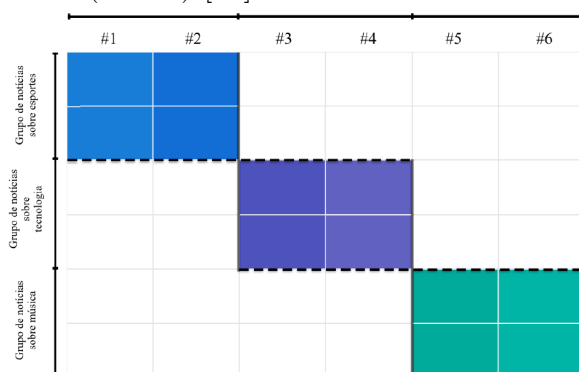
Figure 21: Representação gráfica para a análise de agrupamento: análise da similaridade total entre os dados (textos) [22]



Na figura 22, o princípio de coagrupamento de dados, que se baseia na análise de similaridades parciais, é aplicado. Nesse exemplo, a saída de um algoritmo hipotético, capaz de resolver a tarefa de coagrupamento, fornece informação sobre a existência de três grupos de textos (arbitrariamente ditos grupos de textos sobre “esporte”, “tecnologia” e

“música”) e três grupos de atributos ($\{\#1, \#2\}$, $\{\#3, \#4\}$ e $\{\#5, \#6\}$). As cores azul, roxo e verde indicam a interpretação de três cogrupos de interesse: o azul formado pelo primeiro grupo de linhas em associação ao primeiro grupo de colunas; o roxo é formado pelo segundo grupo de linhas em associação ao segundo grupo de colunas; e o verde é formado pelo terceiro grupo de linhas em associação ao terceiro grupo de colunas³⁶. Uma conclusão possível sobre essa organização é que os atributos $\{\#1, \#2\}$ indicam características que se referem a esportes e uma análise apenas sobre esses dois atributos permite agrupar dois textos e rotulá-los como tratando do tema “esporte”. O mesmo é dito sobre os outros cogrupos. Essa conclusão é diferente das conclusões possíveis a partir do resultado da aplicação de um algoritmo de agrupamento, e é mais informativa. Ainda, o exemplo mostra que mais relações podem ser descobertas com algoritmos de coagrupamento, já que um grupo novo de linhas (“tecnologia”) pode ser descoberto.

Figure 22: Representação gráfica para a análise de coagrupamento: análise da similaridade parcial entre os dados (textos) [22]



Ainda, o potencial de análise obtido a partir da resolução de tarefas de coagrupamento se estende para as possibilidades das sobreposições e interseções entre cogrupos, levando a um ambiente de análise ainda mais rico (como representado na figura 19).

A implementação de estratégias para resolução de tarefas de coagrupamento variam entre diferentes abordagens, passando por abordagens mais adequadas para uso em conjuntos de dados categóricos [132], em conjuntos de dados densos e volumosos [133, 134], para descoberta de determinados tipos de cogrupos [135, 134] ou baseados em meta-heurísticas [136, 137]. Considerando as características de dados textuais, como objetivado nos estudos da autora, alguns autores defendem o uso de fatoração de matrizes para implementação de coagrupamento [124, 138, 139]. Assim, a fatoração de matrizes foi estudada pela autora nas suas primeiras iniciativas de resolução da tarefa de coagrupamento considerando dados textuais.

³⁶Outras combinações de grupos de linhas com grupos de colunas, além das combinações em destaque com as cores na figura 22, podem ser consideradas, como discutido em [22].

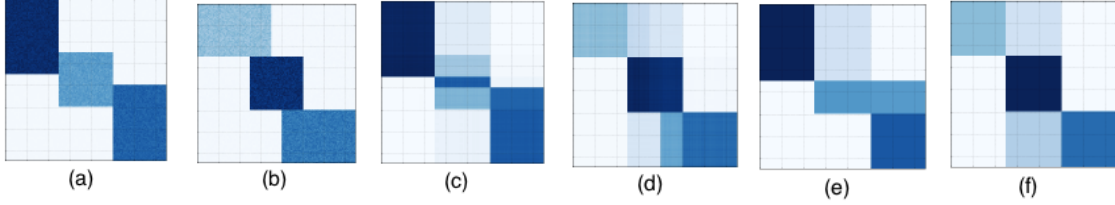
A fim de buscar a resolução para as tarefas de agrupamento e coagrupamento, via aplicação de fatoração de matrizes, vários problemas foram formulados na literatura, incluindo:

- agrupamento clássico (por *k-Means*) [140]: aproximar X pela multiplicação de duas matrizes U e C . Assim, o problema de otimização resolvido pelo algoritmo *k-Means* pode ser escrito como: $\mathcal{P}_1(U, C) = \min_{U, C} \sum_{i=1}^n \sum_{p=1}^k u_{ip} \|\vec{x}_i - \vec{c}_p\|^2 = \min_{U, C} \|X - UC\|^2$, sujeito a $U \in \{0, 1\}^{n \times k}$, $C \in \mathfrak{R}^{k \times m}$ e $\sum_{p=1}^k u_{i,p} = 1 \forall i$, sendo que $\|\cdot\|$ denota uma norma para matrizes;
- agrupamento *fuzzy* (por *Fuzzy-k-Means*): aproximar X pela multiplicação de duas matrizes U e C . Assim, o problema de otimização resolvido pelo algoritmo *Fuzzy-k-Means* pode ser escrito como: $\mathcal{P}_2(U, C) = \min_{U, C} \sum_{i=1}^n \sum_{p=1}^k u_{ip}^w \|\vec{x}_i - \vec{c}_p\|^2 = \min_{U, C} \|X - UC\|^2$, sujeito a $U \in \mathfrak{R}^{n \times k}$, $C \in \mathfrak{R}^{k \times m}$ e $\sum_{p=1}^k u_{i,p} = 1 \forall i$, sendo que $w \in [1, 2, \dots, \infty]$ e $\|\cdot\|$ denota uma norma para matrizes;
- coagrupamento por fatoração ortogonal tripla de matrizes não-negativas (*orthogonal non-negative matrix tri-factorization* - ONMTF) [139]: aproximar X pela multiplicação de três matrizes não-negativas U , S e V . Assim, o problema de otimização resolvido pelo algoritmo ONMTF pode ser escrito como: $\mathcal{P}_3(U, S, V) = \min_{U, S, V} \|X - USV^T\|^2$, sujeito a $U \geq 0, S \geq 0, V \geq 0, U^T U = I, V^T V = I$ em que $U \in \mathfrak{R}_+^{n \times k}$, $S \in \mathfrak{R}_+^{k \times l}$, $V \in \mathfrak{R}_+^{m \times l}$ e $\|\cdot\|$ denota uma norma para matrizes;
- coagrupamento por fatoração tripla rápida de matrizes não-negativas (*fast non-negative matrix tri-factorization* - FNMTF) [141]: aproximar X pela multiplicação de três matrizes não-negativas U , S e V . Assim, o problema de otimização resolvido pelo algoritmo FNMTF pode ser escrito como: $\mathcal{P}_4(U, S, V) = \min_{U, S, V} \|X - USV^T\|^2$, sujeito a $U \in \{0, 1\}^{n \times k}$, $V \in \{0, 1\}^{m \times l}$, $\sum_{p=1}^k u_{i,p} = 1 \forall i$ e $\sum_{q=1}^l v_{j,q} = 1 \forall j$, em que $S \in \mathfrak{R}^{k \times l}$ e $\|\cdot\|$ denota uma norma para matrizes.

Durante o estudo dos problemas supracitados, e dos respectivos algoritmos que os resolvem [22, 140, 139, 141], observou-se que, em relação à tarefa de coagrupamento, situações em que os cogrupos apresentam algum tipo de sobreposição, de linhas ou colunas, não são adequadamente solucionados. A figura 23 ilustra graficamente uma solução ideal para uma situação de coagrupamento com sobreposição de linhas (figura 23(a)) ou de colunas (figura 23(b)), e a solução obtida com a resolução dos problemas \mathcal{P}_3 (figuras 23(c,d)) e \mathcal{P}_4 (figuras 23(e,f)). Nessas figuras, os tons de azul indicam os cogrupos de interesse, nas figuras (a) e (b), e descobertos, nas figuras (c) a (f).

Na figura 23(c) observa-se que a fatoração de matrizes produziu três grupos de colunas, como desejado, porém, para atender a sobreposição de linhas, cinco grupos de linhas foram

Figure 23: Representação gráfica para conjuntos de dados com cogrupos com sobreposição de linhas e colunas e soluções a partir da otimização dos problemas \mathcal{P}_3 e \mathcal{P}_4 . As figuras (a) e (b) representam os cogrupos de interesse em um conjunto de dados hipotético considerando, respectivamente, sobreposição de linhas e sobreposição de colunas. As demais figuras representam soluções obtidas com os problemas de otimização \mathcal{P}_3 (figuras (c) e (d)) e \mathcal{P}_4 (figuras (e) e (f)). Adaptado de [22]

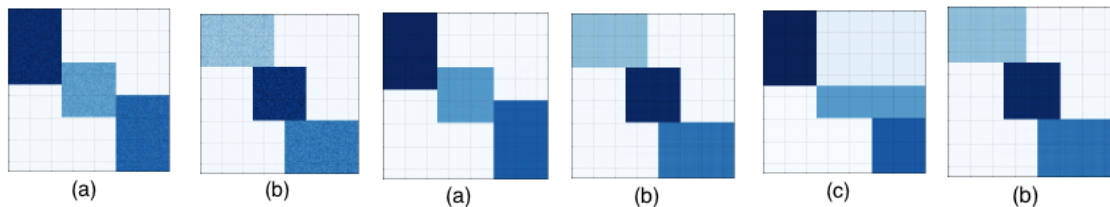


criados, dando origem a mais cogrupos do que o procurado. Na figura 23(e), a solução descobriu três grupos de colunas e três grupos de linhas, porém, apenas o primeiro grupo de linhas foi descoberto da forma correta, e os outros dois além de não possuírem o número de linhas desejado, também foram combinados com grupos de colunas de uma maneira diferente da esperada. Interpretações similares podem ser feitas para as figuras (d) e (f). Como uma maneira de melhorar o resultado de análise de coagrupamento em questão, dois novos problemas foram formulados, os quais seguem aqui apresentados. Para ambos, algoritmos de otimização foram derivados [22].

- coagrupamento por fatoração tripla de matrizes não negativas sobrepostas (*overlapped non-negative matrix tri-factorization* - OvNMTF) [22]: aproximar X pela multiplicação de U por uma composição de S e k matrizes V . Assim, o problema de otimização resolvido pelo algoritmo OvNMTF pode ser escrito como: $\mathcal{P}_5(U, S, V_{(1)}, \dots, V_{(k)}) = \min_{U, S, V_{(1)}, \dots, V_{(k)}} \|X - U \sum_{p=1}^k I_{(p)} S V_{(p)}^T\|^2$, sujeito a $U \geq 0$, $S \geq 0$ e $V_{(p)} \geq 0$, $\forall p$ em que $U \in \mathfrak{R}_+^{n \times k}$, $S \in \mathfrak{R}_+^{k \times l}$, $V_{(p)} \in \mathfrak{R}_+^{m \times l}$, $p \in \{1, \dots, k\}$ é o índice para o conjunto de matrizes $\{V_{(1)}, \dots, V_{(k)}\}$, $I_{(p)} \in \{0, 1\}^{k \times k}$ é uma matriz seletora com zeros em todos os seus elementos exceto no elemento $i_{(p)pp}$, e $\|\cdot\|$ denota uma norma para matrizes;
- coagrupamento por fatoração binária tripla de matrizes não-negativas com sobreposição (*overlapped binary non-negative matrix tri-factorization* - BinOvNMTF) [22]: aproximar X pela multiplicação de U por uma composição de S e k matrizes V . Assim, o problema de otimização resolvido pelo algoritmo BinOvNMTF pode ser escrito como: $\mathcal{P}_6(U, S, V_{(1)}, \dots, V_{(k)}) = \min_{U, S, V_{(1)}, \dots, V_{(k)}} \|X - U \sum_{p=1}^k I_{(p)} S V_{(p)}^T\|^2$, sujeito a $U \in \{0, 1\}^{n \times k}$, $V_{(p)} \in \{0, 1\}^{m \times l}$, $\sum_{p=1}^k u_{i,p} = 1 \forall i$ e $\sum_{j=1}^l v_{(p)j,q} = 1 \forall p, j$, em que $p \in \{1, \dots, k\}$ e $q \in \{1, \dots, l\}$ são os índices que iteram no número de linhas e colunas, respectivamente, $I_{(p)} \in \{0, 1\}^{k \times k}$ é uma matriz identidade seletora e $\|\cdot\|$ denota uma norma para matrizes.

Os resultados obtidos a partir da otimização dos problemas \mathcal{P}_5 e \mathcal{P}_6 são apresentadas na figura 24. Apenas um resultado encontrado não foi satisfatório (figura 24(c)). Esse problema se deve ao fato de que o problema \mathcal{P}_6 não permite que uma linha pertença a mais de um grupo de linhas. E embora também exista a restrição de pertinência a um único grupo de colunas, a existência de várias matrizes V independentes permite encontrar grupos de colunas (nesse caso, dois grupos) independentes para cada grupo de linhas.

Figure 24: Representação gráfica para conjuntos de dados com cogrupos com sobreposição de linhas e colunas e soluções a partir da otimização dos problemas \mathcal{P}_5 e \mathcal{P}_6 . As figuras (a) e (b) representam os cogrupos de interesse em um conjunto de dados hipotético considerando, respectivamente, sobreposição de linhas e sobreposição de colunas. As demais figuras representam soluções obtidas com os problemas de otimização \mathcal{P}_5 (figuras (c) e (d)) e \mathcal{P}_6 (figuras (e) e (f)). Adaptado de [22]



As estratégias de agrupamento e coagrupamento mencionadas neste texto foram aplicadas a dados textuais e avaliadas em relação à sua capacidade de descoberta de grupos, mediante uma avaliação via índice externo – índice de Rand [142]. Essa avaliação é pertinente no contexto em que se tem a informação de classes associadas aos dados, como é o caso dos conjuntos de dados textuais usados nos testes. Visto haver a informação de classes, as avaliações consideraram a aderência dos grupos de linhas descobertos em relação às classes associadas aos dados. Os resultados obtidos estão resumidamente apresentados nas tabelas 7 e 8. Detalhes completos sobre os experimentos e os resultados obtidos estão organizados em [22].

Table 7: Índice RAND máximo obtido na aplicação dos algoritmos de agrupamento e coagrupamento. Conjunto Toy-BR: 300 textos em português igualmente distribuídos em três classes; Conjunto-BR: 4.575 textos em português distribuídos em 13 classes; 555 textos em inglês distribuídos em nove classes. Baseado em [22]

Estratégia	Conjunto Toy-BR	Conjunto-BR	NIPS			
K-means	0,7086	0,3137	0,1573			
fuzzy-K-means	0,4970	0,2662	0,1882			
ONMTF	0,6479	1 = 5	0,1802	1 = 19	0,1579	1 = 6
FNMTF	0,2615	1 = 3	0,2399	1 = 19	0,2199	1 = 18
OvNMTF	0,7487	1 = 3	0,3554	1 = 16	0,1742	1 = 12
BinOvNMTF	0,4818	1 = 5	0,3784	1 = 16	0,2813	1 = 15

Table 8: Índice RAND máximo obtido na aplicação dos algoritmos de agrupamento e coagrupamento. Conjunto Toy-BR: 300 textos em português igualmente distribuídos em três classes; Conjunto-BR: 4.575 textos em português distribuídos em 13 classes; 555 textos em inglês distribuídos em nove classes. Baseado em [22]

Estratégia	Conjunto Toy-BR	Conjunto-BR	NIPS			
K-means	0,8152	0,3955	0,2006			
fuzzy-K-means	0,6669	0,3557	0,2728			
ONMTF	0,7885	1 = 5	0,2884	1 = 10	0,2704	1 = 18
FNMTF	0,4502	1 = 4	0,2813	1 = 16	0,3009	1 = 15
OvNMTF	0,8208	1 = 2	0,4251	1 = 10	0,1992	1 = 9
BinOvNMTF	0,8261	1 = 3	0,5743	1 = 10	0,3670	1 = 15

Para os algoritmos ONMTF e OvNMTF, foi realizada uma avaliação qualitativa baseada na interpretação semântica dos cogrupos descobertos, considerando o conjunto de dados “Conjunto Toy-BR”. Devido às características da matriz S (Tabela 9), foi possível semi-automatizar tal análise e um exemplo do tipo de resultado possível com essa interpretação é apresentado aqui.

Table 9: Matrizes S normalizadas obtidas em um resultados de coagrupamentos para o “Conjunto Toy-BR”. Os grupos de colunas GC 1 a GC 5 foram obtidos a partir do algoritmo ONMTF. Os grupos de palavras GC_{OV} 1 a GC_{OV} 6 foram obtidos a partir do algoritmo OvNMTF. Abreviações: “esp” - esportes; “are” - arena; “jov”: jovem. Baseado em [22]

	GC 1	GC 2	GC 3	GC 4	GC 5	GC _{OV} 1,3,5	GC _{OV} 2,4,6	
GL “esp”	0,0	0,5	0,1	0,0	0,4	GL “are”	0,38	0,62
GL “are”	0,0	0,05	0,05	0,9	0,0	GL “jov”	0,46	0,54
GL “jov”	0,4	0,1	0,5	0,0	0,0	GL “esp”	0,94	0,06

As matrizes S indicam uma relação entre grupos de linhas (documentos) e grupos de colunas (palavras). Números maiores na matriz S são interpretados como relações mais significativas, assim, seguindo a matriz S obtida com o algoritmo ONMTF tem-se que, por exemplo, o grupo de linhas GL “esportes”³⁷ está relacionado ao grupo de colunas GC 2 e GC 5. Os grupos de colunas encontrados pelo ONMTF, resultantes de um particionamento baseado na matriz V , estão ilustrados aqui a partir das cinco palavras mais relevantes em cada grupo, de acordo com os fatores (na matriz V) que relacionam as palavras aos grupos de palavras³⁸:

³⁷Os rótulos para os grupos de linhas foram estabelecidos seguindo o que se conhece sobre as classes de notícias do conjunto de dados e as palavras obtidas nos grupos de colunas.

³⁸Em [22] a análise está estendida para as 20 palavras mais relevantes.

- GP 1 (esportes radicais)³⁹: skate, surfe, bob, burnquist⁴⁰, circuito
- GP 2 (futebol): real, breno⁴¹, time, barcelona⁴², partida
- GP 3 (esportes em geral)): anos, mundial, brasil, etapa, brasileiro
- GP 4 (*games*): jogos, xbox⁴³, playstation⁴⁴, wii⁴⁵, jogo
- GP 5 (futebol): gol, madrid⁴⁶, gols, messi⁴⁷, euro

A interpretação do resultado para o OvNMTF exige um olhar sobre particionamentos em cada uma das matrizes V_i . Cada linha da matriz S estabelece uma relação com grupos de palavras correspondentes somente à aquela linha, ou seja, a um grupo de documentos específico. Assim, no caso do experimento em discussão, tem-se:

- os grupos de palavras GC_{OV} 1 e GC_{OV} 2 estão relacionados com o grupo de documentos GL “are”, mediante os valores, 0,38 e 0,62 respectivamente;
- os grupos de palavras GC_{OV} 3 e GC_{OV} 4 estão relacionados com o grupo de documentos GL “jov”, mediante os valores, 0,46 e 0,54 respectivamente;
- os grupos de palavras GC_{OV} 5 e GC_{OV} 6 estão relacionados com o grupo de documentos GL “esp”, mediante os valores, 0,94 e 0,06 respectivamente.

Para o caso da análise do resultado obtido com o OvNMTF, as cinco palavras mais relevantes para cada grupo de colunas, obtidos em cada uma das matrizes V_i foram⁴⁸:

- GP 1 (*games*): jogos, sony, ps⁴⁹, the, pessoas

³⁹Os rótulos pra cada grupo de palavras foram atribuídos com base na análise das 20 palavras mais relevantes em cada um deles.

⁴⁰Bob Burnquist é um skatista brasileiro-americano. É o maior medalhista da história do X Games, segundo Wikipédia, consultada em janeiro de 2017.

⁴¹Breno Vinícius Rodrigues Borges é um futebolista brasileiro, segundo Wikipédia, consultada em janeiro de 2017.

⁴²Futbol Club Barcelona é um clube de futebol profissional espanhol, segundo Wikipédia, consultada em janeiro de 2017.

⁴³Xbox é uma marca de consoles de videogame criada pela Microsoft, segundo Wikipédia, consultada em janeiro de 2017

⁴⁴PlayStation é uma marca de consoles de videogame criada pela Sony, segundo Wikipédia, consultada em janeiro de 2017

⁴⁵Wii é uma marca de consoles de videogame criada pela Nintendo, segundo Wikipédia, consultada em janeiro de 2017

⁴⁶Referência ao Real Madrid Club de Fútbol, um clube poliesportivo espanhol, segundo Wikipédia, consultada em janeiro de 2017.

⁴⁷Lionel Andrés Messi é um futebolista argentino, atuante na Espanha, segundo Wikipédia, consultada em janeiro de 2017.

⁴⁸Em [22] a análise está estendida para as 20 palavras mais relevantes

⁴⁹Referência ao PlayStation.

- GP 2 (*games*): jogo, of, playstation, game, novo
- GP 3 (esportes em geral): games, jovem, brasileiro, paulo, dia
- GP 4 (esportes radicais + música): anos, skate, mundial, brasil, surfe
- GP 5 (futebol): time, real⁵⁰, feira, final, gols
- GP 6 (futebol): breno, casa, gol, bayern⁵¹, minutos

3.3 Área de aplicação: Sistemas de Recomendação baseados em conteúdo textual

Segundo [143], sistemas de recomendação são ferramentas computacionais capazes de sugerir itens para um usuário. Os itens a serem sugeridos podem ser os mais variados como produtos para compra, notícias, livros e artigos científicos para leitura ou ainda lugares para visitaç o. A recomenda o   obtida a partir de uma fun o que se constitui como o n cleo do sistema, e sua implementa o est  diretamente ligada ao tipo de sistema em quest o: baseado em conte do, por filtro colaborativo, baseado em conhecimento ou h bridos, para citar os tipos mais conhecidos.

Sistemas de recomenda o baseado em conte do   definido por [143] como um sistema que recomenda itens escolhidos a partir da an lise de similaridade entre as caracter sticas associadas a eles. Por exemplo, se um usu rio indica, de alguma forma, interesse por filmes que pertencem ao g nero com dia, o sistema pode aprender a recomendar outros filmes do mesmo g nero [143].

Baseado em [144]⁵², o problema de recomenda o com base em conte do   definido como uma rela o de similaridade $s : \mathcal{I} \times \mathcal{I} \rightarrow \mathfrak{R}$, que permite aproximar uma fun o $\rho_s : \mathcal{H} \rightarrow \mathcal{I}$ de tal forma que a lista de recomenda es L gerada para um usu rio U_{s_n}   dada por $L_{U_{s_n}} = \rho_s(\mathcal{H}_{U_{s_n}})$, em que

- $\mathcal{I} = It_1, \dots, It_m$   um conjunto de itens;
- $\mathcal{U} = U_{s_1}, \dots, U_{s_n}$   um conjunto de usu rios;

⁵⁰Refer ncia ao Real Madrid Club de F tbol, um clube poliesportivo espanhol, segundo Wikip dia, consultada em janeiro de 2017.

⁵¹Fu ball-Club Bayern M nchen   uma agremia o poliesportiva alem , segundo Wikip dia, consultada em janeiro de 2017.

⁵²Brunialti, L. F., **PERES, S. M.**, Freire, V., Lima, C. A. M. Aprendizado de m quina em sistemas de recomenda o baseados em conte do textual: Uma revis o sistem tica. In: Anais do XI S mpo io Brasileiro de Sistemas de Informa o. Goi nia, GO, 2015. p. 203–210

- \mathcal{H} representa um histórico de interações do usuário com o sistema, sendo que uma interação $h \in \mathcal{H}$ é uma tripla $h = (U_{s_n}, It_m, \sigma)$ indicando que o usuário U_{s_n} interagiu com o item It_m e emitiu a avaliação σ sobre a experiência;

Assim, o problema na construção de sistemas de recomendação por conteúdo tem como núcleo o desenvolvimento de estratégias de implementação de ρ_s que tenham como entradas o histórico de interação do usuário com o sistema e o conteúdo que descreve, ou representa, os itens para recomendação.

As vantagens desse tipo de sistema, em comparação com o tipo mais popular de recomendação (a colaborativa) são a independência do usuário, a transparência e a capacidade de lidar com novos itens (*item cold-start problem*). Porém, ele sofre de algumas desvantagens relevantes: a análise de conteúdo apresenta limitações (conhecimento de domínio), a superespecialização (o problema da serendipidade) e a capacidade de lidar com novos usuários (*user cold-start problem*), visto que ele possui apenas a primeira interação com o sistema em seu histórico de interações.

Lidar com o problema de serendipidade é a base da motivação da pesquisa discutida aqui. A serendipidade, como definido por [143], pode ser vista como a experiência de receber uma recomendação inesperada, porém relevante. Na literatura da área, esse problema tem sido tratado com inserção de algum nível de aleatoriedade na tomada de decisão sobre recomendação ou com a penalização de recomendações referente a itens muito similares ou redundantes. Para a autora, o problema de serendipidade pode ser amenizado no que diz respeito à análise de conteúdo textual (quando este é o conteúdo sob análise) se similaridades parciais (como tratado no problema de coagrupamento) forem consideradas.

Como exemplo para ilustrar a proposta de tratamento do problema de serendipidade, considere um sistema de recomendação para notícias (um item de dado textual). A hipótese de que a análise de similaridades parciais pode ser usada para melhorar os cenários de recomendações no que diz respeito à serendipidade está relacionada à observação dos resultados obtidos com os experimentos mencionados na seção 3.2. Nesses experimentos, uma análise das palavras que caracterizam um cogrupamento descoberto no conjunto de dados “Conjunto Toy-BR” (Figura 25) mostrou que existe diversidade presente nas notícias nele agrupadas: “música” é um termo que aparece com relevância em meio a uma série de termos relacionados a esporte (“skate”, “mundial”, “surfe”, “final”, “pista”, “manobra”, para citar alguns exemplos), indicando que notícias relacionadas ao tema música tem alguma similaridade com notícias relacionadas a esportes, pelo menos no contexto do conjunto de notícias designadas para esse cogrupamento.

Com alguma liberdade de análise, sabe-se que em grandes eventos musicais é frequente a presença de subventos de esportes, enquanto o contrário é também razoável, uma vez

Figure 26: Representações de dois cogrupos a partir de palavras relevantes para suas formações. Quanto maior o destaque da palavra na “nuvem de palavras”, mais relevante ela é dentro do cogrupo. Baseado em [22]



de notícias em cada cogrupos) apresentam valores baixos quando um número maior de notícias é analisado (veja tabelas 7 e 8). Isso pode ser um indicativo de que o uso de similaridades parciais está levando à formação de grupos que diferem daqueles representados pelos clássicos “canais” de notícias, os quais se usados como critério de recomendação podem levar a recomendações óbvias e especializadas.

3.4 Objetivos atuais e limitações da pesquisa

No cenário atual da pesquisa em análise de dados textuais desenvolvida pela autora, tem-se como principal contribuição, a associação de resultados de análise de coagrupamento de textos (num contexto de notícias) com a proposição de melhorias para o problema de serendipidade em sistemas de recomendação baseada em conteúdo. Não se trata de uma contribuição definitiva para a área de sistemas de recomendação, visto que, embora já tenham sido produzidos alguns indícios de que tal associação pode ser produtiva, é necessário desenvolver estudos para verificação da hipótese formulada para esse contexto. Ainda, marginalmente à essa hipótese, resultados importantes foram obtidos no que diz respeito à resolução da tarefa de coagrupamento de dados via fatoração de matrizes. Dois novos problemas de otimização, e algoritmos para sua resolução, foram desenvolvidos e encontram-se em fase de inserção na comunidade científica via a publicação de artigos científicos.

A partir do ponto atual da pesquisa, há alguns objetivos sendo buscados por meio de estudos atualmente em desenvolvimento. Esses objetivos incluem:

- Estabelecimento de um ambiente para testes de análise de dados textuais baseado em textos em português brasileiro e que representem um meio dinâmico de comunicação humana: a publicação de notícias na internet. Para alcançar esse objetivo, a

criação, organização, documentação e publicação de um *corpus* de notícias em português brasileiro são ações em andamento. Essa iniciativa se justifica por não haver um *corpus* atual desse gênero disponibilizado para uso em trabalhos acadêmicos. Atualmente, o grupo de pesquisa da autora já possui autorização para trabalhar com textos publicadas no portal de notícias da Empresa Brasil de Comunicação (EBC - www.ebc.com.br).

- Estudo e desenvolvimento de algoritmos de aproximação, baseados em heurísticas e meta-heurísticas, que sejam capazes de lidar eficientemente com o problema de coagrupamento, considerando diferentes tipos e estruturas de cogrupos, em dados textuais. Dados textuais apresentam características de dados esparsos, e coagrupamento é uma tarefa que admite muitas soluções. Essas duas características são de difícil tratamento por algoritmos exatos e motivam o estudo referente a algoritmos de aproximação.
- Exploração de diferentes formas de representação dos dados textuais e das relações existentes entre as características (termos) que são usadas como descritores desses dados. Principalmente, para esse objetivo, a autora deve despender esforços no uso, e análise, de similaridade semântica de conteúdo textual aplicada na representação dos textos a serem submetidos a algoritmos de agrupamento e coagrupamento.
- Aplicação dos resultados dos estudos supracitados em contextos de sistemas de recomendação baseados em conteúdo textual, e avaliação de desempenho das estratégias que combinam algoritmos de coagrupamento e representações alternativas para dados textuais. Tal avaliação deve se dar no âmbito do problema da serendipidade. Métrica para avaliação de recomendações serendipitadas ainda são bastante raras e discutíveis, e assim, o estabelecimento de estratégias eficazes para realizar tais avaliações se torna também um objetivo da presente pesquisa.

A pesquisa aqui discutida, tanto em termos dos resultados já obtidos quanto em termos dos resultados pretendidos, sofre por estar limitada em três aspectos. O primeiro deles, inerente à natureza do contexto de comunicação humana, diz respeito à intenção de estabelecer a automação, em algum nível, de interpretações textuais. Ou seja, pretende-se interpretar o conteúdo de textos, sugerindo que tais textos são similares dentro de determinados contextos (similaridade parcial). No entanto, dentro da comunicação humana, incluindo a comunicação escrita, há aspectos como estilo literário e expressão de emoções/intenções que não estão sendo considerados dentro do escopo desta pesquisa, limitando a qualidade das interpretações possíveis.

A avaliação da serendipidade de recomendação é um problema considerado em aberto pela comunidade científica. Muitos trabalhos usam estratégias de avaliação da serendipi-

dade a partir de experimentos com usuários, incluindo a avaliação de serendipidade como um quesito de qualidade da recomendação a ser quantificado ou qualificado pelo próprio usuário. Uma exceção é o trabalho de [145] que estabelece uma métrica para realizar tal avaliação. A insegurança ou a fragilidade referente à avaliação do quesito serendipidade confere uma complexidade adicional para o alcance pleno dos objetivos referentes à contribuição científica pretendida nesta pesquisa.

Por fim, a terceira limitação desta pesquisa está relacionada à diversidade de repostas que a tarefa de coagrupamento admite. Assim como agrupamento, a depender das métricas de similaridades usadas, das estratégias implementadas nos algoritmos, e de condições iniciais e parametrização desses algoritmos, os resultados obtidos para a tarefa de coagrupamento podem ser bem diferentes, porém igualmente bem (ou mal) avaliados. Novamente, um fato de instabilidade de resultados se apresenta no contexto da pesquisa e complica a defesa dos resultados eventualmente obtidos.

4 Contribuições em análise de *logs* de eventos

Este capítulo é destinado a apresentar os esforços preliminares de pesquisa na área de Mineração de Processos desenvolvidos pela autora em sua linha de pesquisa mais recente. Embora a pesquisa esteja sendo feita há poucos anos, já foi possível alcançar resultados interessantes e estabelecer parcerias de pesquisa diferenciadas.

No início da pesquisa, o principal desafio para a autora foi compreender o papel da área de Mineração de Dados junto aos objetivos da Mineração de Processos, pois essa última é uma área predominantemente estudada por pesquisadores da área de Gerenciamento de Processos de Negócio, o que faz com que a taxonomia usada na literatura correlata, e mesmo a compreensão do papel das tarefas de “mineração”, sejam, eventualmente, diferentes do que é praticado na área de Mineração de Dados e reportado na literatura especializada da área. Além disso, o principal objetivo em Mineração de Processos é estudar a estrutura (se existente) de um processo de negócio, a qual está embutida em dados representados de forma não estruturada: as informações referentes à execução de um processo organizacional estão armazenadas em um *log* de eventos, que basicamente se refere a uma sequência de ações executadas dentro de um processo ao longo do tempo. Coube ainda à essa fase do estudo, constatar quais as estratégias que vinham sendo usadas para modelar as tarefas de Mineração de Processos e quais as suas relações com as tarefas de Mineração de Dados, e principalmente verificar se a Inteligência Computacional e o Aprendizado de Máquina estavam inseridos nesse meio, em que proporção e sob quais expectativas.

Do estudo realizado pela autora e colaboradores para levantar o referencial teórico e o estado da arte da área de Mineração de Processos, e sua relação com Mineração

de Dados, Inteligência Computacional e Aprendizado de Máquina, foi constatado que: (a) as tarefas clássicas de Mineração de Processos são diferentes daquelas estabelecidas em Mineração de Dados, porém, há algumas iniciativas que, explicitamente, modelam as tarefas de Mineração de Processos seguindo, nem sempre corretamente, o que está estabelecido pela área de Mineração de Dados; (b) embora Mineração de Processos esteja em ampla expansão, muito pouca atenção tem sido dada para o estudo e aplicação de Inteligência Computacional e Aprendizado de Máquina para resolução de seus problemas. Dentro desse contexto, a ideia de atuar na linha de pesquisa multidisciplinar se fortaleceu na possibilidade de potencializar os resultados para a área de Mineração de Processos, divulgando-a para pesquisadores das áreas de Mineração de Dados, Inteligência Computacional e Aprendizado de Máquina.

Também a partir dos estudos já realizados, ficou claramente constatado que a grande maioria dos esforços em pesquisa na área de Mineração de Processos é derivada dos estudos do grupo de pesquisa em Mineração de Processos da *Technische Universiteit Eindhoven*. A visão da pesquisa em Mineração de Processos pretendida pela autora, embora baseada na visão do grupo citado, está fortemente relacionada com a taxonomia já classicamente adotada em Mineração de Dados, de forma que abre um espaço para a evolução da intersecção entre as duas áreas, e principalmente, fortalece a Mineração de Processos como interesse em pesquisa da área de Mineração de Dados.

O presente capítulo está organizado de forma a introduzir os conceitos que embasam os estudos realizados pela autora na área multidisciplinar de Mineração de Processos. Esses estudos foram iniciados a partir de um trabalho de mapeamento sistemático sobre Mineração de Processos e suas relações com a área de Mineração de Dados e as técnicas da Inteligência Computacional e do Aprendizado de Máquina, cujos resultados são resumidos na seção 4.2. Na sequência, uma aplicação de redes neurais artificiais dentro do escopo de análise de *log* de eventos de um processo da área de Educação é apresentada. Essa é a primeira iniciativa prática da autora pautada no objetivo de aproximar a modelagem de uma clássica tarefa de Mineração de Dados dos objetivos de análise de *log* de eventos e, consequentemente, da Mineração de Processos. A produção científica obtida nesta linha de pesquisa, ainda inicial, é citada no texto e descrita em detalhes em rodapés. O capítulo é finalizado com a apresentação dos objetivos que estão norteando a pesquisa da autora atualmente e também algumas limitações associadas ao tema pesquisado.

4.1 Processos e *log* de eventos

A área de Mineração de Processos sobre a qual versam os estudos da autora está relacionada a processos de negócio, podendo ser mais especificamente nomeada como Mineração de Processos de Negócio. Ela deriva dos estudos em Gerenciamento de Pro-

cessos de Negócio (BPM - de *Business Process Management*) [146, 147]. Devido a isso, a definição de processos aqui adotada tem as características que são comumente adotadas em BPM.

Segundo [146, 148], um processo de negócio é um conjunto de tarefas (ou atividades) inter-relacionadas que deve ser executado para que uma determinada meta de negócio seja alcançada. Em um processo de negócio, uma atividade é vista como uma unidade semântica, em algum nível; ela pode ser entendida como uma função que modifica o estado de um processo [149]. Um modelo de processo BP é formalmente definido por [149] da seguinte forma: um conjunto de atividades $\mathcal{A}_{BP} = \mathcal{A}_1, \dots, \mathcal{A}_n$ e um conjunto de dependências entre atividades \mathcal{E}_{BP} organizados em um grafo direcionado $\mathcal{G}_{BP} = (\mathcal{A}_{BP}, \mathcal{E}_{BP})$, uma função de saída $o_{BP} : \mathcal{A}_{BP} \rightarrow \vec{\mathcal{B}}^k$, sendo k a dimensão do vetor $\vec{\mathcal{B}}$ de parâmetros de saída da atividade executada, e uma função Booleana $\theta_{(a_1, a_2)} : \vec{\mathcal{B}}^k \rightarrow \{0, 1\}$ definida $\forall (a_1, a_2) \in \mathcal{E}_{BP}$.

A execução de um processo de negócio segue o grafo de atividades definido. Sempre que uma atividade é executada, uma função de saída é computada. Sobre essa saída, a função booleana definida para as arestas que saem da atividade executada é avaliada, considerando o vetor de parâmetros instanciado, para decidir a próxima atividade que pode ser executada. Assim, de forma simplificada, cada instância de um processo é uma sequência de eventos que registram o início e o fim de cada atividade executada. Nesse contexto, um *log* pode ser definido como um conjunto de execuções seguindo um grafo de processo desconhecido. [149] definem o *log* de uma execução de um processo como uma lista de registros de eventos formados pela quintupla: identificador da execução do processo, nome da atividade executada, tipo de evento (de *início* ou de *fim*), momento de ocorrência do evento e a saída da atividade ($o(\mathcal{A})$ se o evento é de *fim*, e *null* caso contrário).

A definição supracitada pode ser simplificada para os fins pretendidos nos estudos em Mineração de Processos. A inclusão da informação de tipo de evento duplica os registros referentes à ocorrência da atividade, já que considera o evento de início e fim da mesma. A informação sobre a saída da atividade não é, necessariamente, útil para a análise do modelo do processo, pelo menos para o caso das tarefas mais básicas. Por outro lado, é comum encontrar algoritmos e ferramentas de Mineração de Processos que solicita a inclusão da informação referente ao “recurso” relacionado à execução da atividade. Ainda, outras informações podem ser úteis como custo da atividade, por exemplo. Um fragmento de *log* de eventos⁵⁶ clássico é mostrado na figura 27.

A natureza não estruturada dos dados no caso dos *logs* de eventos é principalmente

⁵⁶A partir deste momento neste texto, o termo *log* de eventos será usado para representar a lista de registros de eventos referentes ao *log* de uma ou várias execuções independentes de um processo.

Figure 27: Fragmento de um *log* de eventos referente a um sistema hipotético de pagamento de compensações. Neste exemplo, o “recurso” diz respeito ao usuário associado à execução da atividade no sistema a partir da qual o *log* de eventos foi gerado. Adaptado de [23]

IdCaso	Timestamp	Atividade	Usuário
1	30-12-2010:11.02	Registrar requerimento	Pete
2	30-12-2010:11.32	Registrar requerimento	Mike
2	30-12-2010:12.12	Verificar bilhete	Mike
2	30-12-2010:14.16	Analisar informalmente	Pete
3	30-12-2010:14.32	Registrar requerimento	Pete
3	30-12-2010:15.06	Analisar informalmente	Mike
3	30-12-2010:16.34	Verificar bilhete	Ellen
1	31-12-2010:10.06	Analisar o detalhe	Sue
2	05-01-2011:11.22	Verificar bilhete	Sara
1	05-01-2011:15.12	Verificar bilhete	Mike
3	06-01-2011:09.18	Decidir	Sara
1	06-01-2011:11.18	Decidir	Sara
1	06-01-2011:11.18	Rejeitar requerimento	Pete
3	06-01-2011:12.18	Reiniciar requerimento	Sara
3	06-01-2011:13.06	Analisar o detalhe	Sean
3	08-01-2011:11.43	Verificar bilhete	Pete
2	08-01-2011:12.05	Pagar compensação	Ellen
3	09-01-2011:09.55	Decidir	Sara
3	09-01-2011:09.55	Pagar compensação	Ellen

manifestada por conta da dependência existente entre cada um dos eventos. Na figura 27 há informação sobre três instâncias de processos, porém há 19 registros de *log* e os registros referentes a cada uma das instâncias estão intercalados no *log*.

4.2 Mineração de Dados *versus* Mineração de Processos

Como definido por [24], Mineração de Processos é uma área de estudo que tem como principal objetivo confrontar dados referentes a eventos (i.e., um comportamento observado) e modelos de processo, atuando: na descoberta de tais modelos a partir da análise do *log* de eventos; na reprodução e análise de dados referentes a eventos a partir de modelos de processo; e na geração de informação para a melhoria dos processos. O mesmo autor posiciona a disciplina de Mineração de Processos como um ponto de interesse entre Mineração de Dados, com atenção especial para a abordagem de Aprendizado de Máquina, e a área de Modelagem e Análise de Processos.

A partir dessa definição, as três tarefas básicas de Mineração de Processos podem ser apresentadas [24, 23]:

- Descoberta: essa tarefa diz respeito à produção de um modelo de processo, ainda desconhecido, com base somente em um *log* de eventos e sem usar qualquer outra

informação *a priori*. A tarefa de descoberta de modelos de processo pode ser vista como uma função que mapeia um *log* de eventos (como descrito na seção 4.1) para um modelo de processo, de forma que o modelo seja representativo para o comportamento observado no *log* de eventos.

- Conformidade: essa tarefa diz respeito à comparação entre um modelo de processo já conhecido com um *log* de eventos gerado sob o mesmo processo, com o objetivo de checar se a realidade registrada no *log* está alinhada com o modelo de processo associado, e vice-versa.
- Melhoria: essa tarefa diz respeito à extensão ou modificação (reparo) de modelos de processo existentes, com base no *log* de eventos referente ao mesmo processo, com o objetivo de adequar o modelo de processo e torná-lo mais aderente ao seu propósito, em algum sentido e em algum nível.

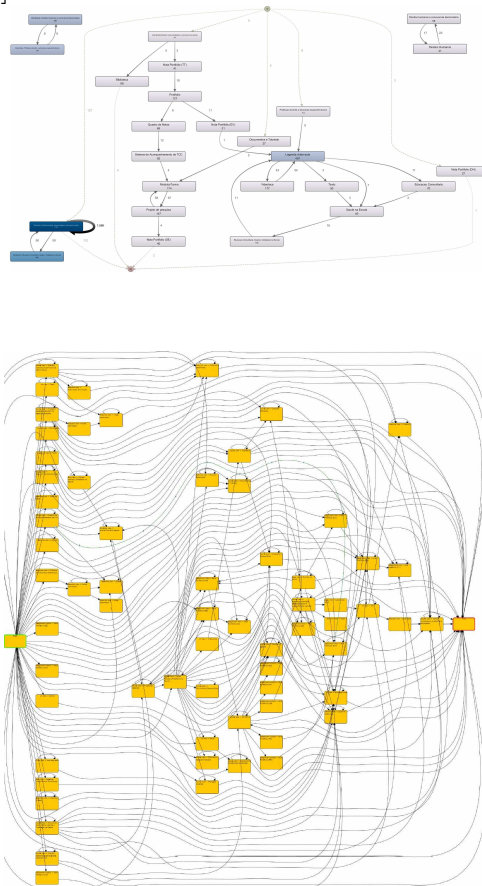
Para o estudo de Mineração de Processos, tipos de processos precisam ser discutidos: processos do tipo “lasanha” e do tipo “espaguete”; e a taxonomia considerada um *continuum* entre tais tipos, incluindo os tipos “estruturado”, “semiestruturado” e “não estruturado”. A figura 28⁵⁷ ilustra os modelos de processos do tipo lasanha e espaguete.

Processos do tipo lasanha tem uma clara estrutura, com relativamente poucas exceções, e os *stakeholders* possuem um claro entendimento do fluxo de trabalho associado ao processo. Segundo [24], é impossível apresentar uma definição formal nesse contexto e, por isso, sugere-se a seguinte heurística: um processo é do tipo lasanha se, considerando um limite de esforço, é possível criar um modelo para o processo que seja validado pelos *stakeholders* e no qual 80% dos eventos ocorram como planejado. O processo do tipo lasanha corresponde a um processo do tipo estruturado, no qual todas as atividades são repetíveis e possuem entradas e saídas bem definidas. Segundo o mesmo autor, os processos do tipo espaguete, ou não estruturado, apresentam características que não permitem definir pré e pós condições para as atividades, e são processos dirigidos à experiência, intuição, tentativa e erro, regras heurísticas e informação qualitativa e vaga. Em um espaço de transição entre as duas condições de processos está o processo do tipo semiestruturado. [24] define que nesse tipo de processo, os requisitos para as atividades são conhecidos e é possível esboçar modelos para ele, contudo algumas atividades no processo requerem julgamento humano, o que pode impedir o processo de seguir o curso modelado.

Mineração de Dados, como classicamente definida, tem o propósito de explorar e analisar dados, referentes a observações sobre fenômenos de diferentes naturezas, para

⁵⁷Maita, A. R. C. Um estudo da aplicação de técnicas de inteligência computacional e de aprendizado de máquina em mineração de processos de negócio. 129p. Dissertação (Mestrado) - Universidade de São Paulo, 2016. Coorientação da autora.

Figure 28: Modelos de processo com características do tipo “lasanha” (acima) e do tipo “espaguete” (abaixo) [150]



descobrir padrões relevantes que estão neles escondidos. Em uma relação direta, os **dados** analisados pela Mineração de Dados correspondem aos **logs de eventos** analisados na Mineração de Processos, e os **padrões relevantes** da Mineração de Dados correspondem aos **modelos de processo** da Mineração de Processos.

No entanto, a maneira como trabalham muitos dos algoritmos usados para resolver tarefas de Mineração de Dados não é diretamente aplicável para resolver as tarefas de Mineração de Processos, uma vez que esses algoritmos não são, necessariamente, orientados a processos. De fato, a área de Mineração de Dados possui uma boa maturidade em relação a associação de algoritmos e métodos à resolução de suas tarefas clássicas: predição categórica (ou classificação), predição numérica (ou regressão), agrupamento, descoberta de padrões frequentes, regras de associação ou de correlação; análise de ruído e de *outliers*; análise de evolução; e análise de estruturas (grafos, árvores, etc). A dificuldade, porém, está em transferir essa maturidade de aplicação de algoritmos e métodos para a área de Mineração de Processos, uma vez que as tarefas envolvidas em cada área possuem definições e propósitos diferentes. [24] defende que a Mineração de Dados possui, tipicamente, duas tarefas que se adequam ao contexto de Mineração de Processos: a min-

eração de sequências⁵⁸ e a mineração de episódios⁵⁹. Contudo, à parte da discussão em relação à modelagem e taxonomia de tarefas, e devido à alta aplicabilidade das técnicas de Inteligência Computacional e Aprendizado de Máquina na área de Mineração de Dados [151], é justificável levantar a hipótese de que tais técnicas podem ser úteis para área de Mineração de Processos, e assim, constituir um campo interessante e útil de atuação em pesquisa.

Diante do contexto discutido, a autora, junto a colaboradores, investiu esforços para realizar um mapeamento sistemático com o propósito de compreender como tarefas de Mineração de Dados e técnicas de Inteligência Computacional e Aprendizado de Máquina estão sendo usadas nos cenários pertinentes à Mineração de Processos. Nesse mapeamento, 705 estudos primários, publicados entre 2005 e 2014, foram analisados na busca de evidências sobre a pesquisa de modelagens e resoluções de tarefas de Mineração de Processos, considerando ou não a taxonomia das tarefas de Mineração de Dados, e na busca de evidências de uso de técnicas de Inteligência Computacional e Aprendizado de Máquina. O planejamento, condução e análise de resultados de tal mapeamento estão organizados na dissertação de mestrado de [150]. Ainda, um recorte do mapeamento foi aprofundado constituindo uma revisão sistemática de literatura sobre a aplicação de redes neurais artificiais e máquinas de vetores de suporte na resolução de tarefas de Mineração de Processos. Os resultados desta revisão sistemática estão publicados no artigo de [152]⁶⁰. Os gráficos apresentados nas figuras 29 e 30 apresentam, respectivamente, uma relação entre tarefas de Mineração de Dados e de Mineração de Processos, e uma relação entre técnicas usadas para resolução das tarefas de Mineração de Processos. Essas relações estão expressas em termos da quantidade de estudos primários que versam sobre iniciativas que as implementam ou discutem.

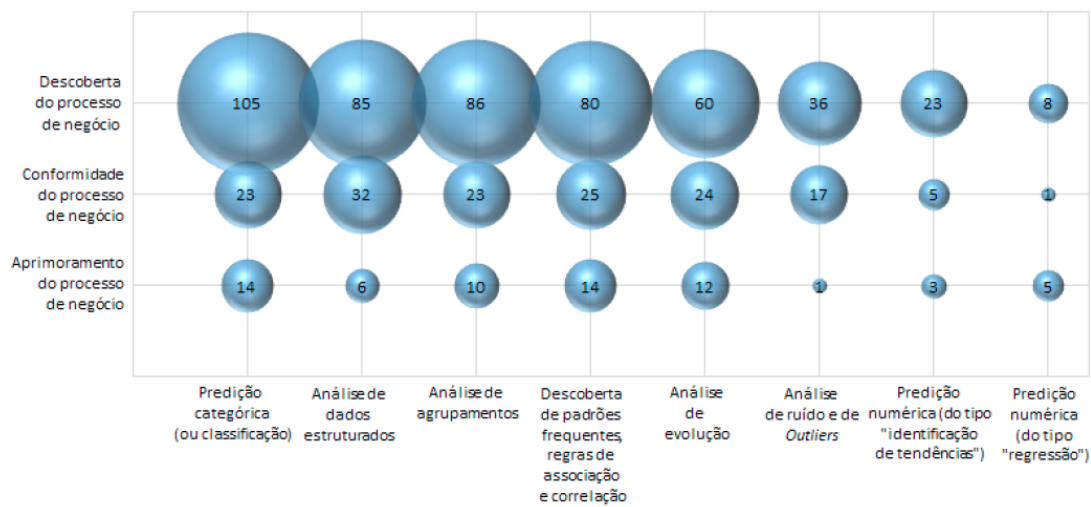
Os dados dispostos na figura 29 mostram que a distribuição de esforços na resolução das tarefas de Mineração de Processos se concentram, com algum destaque, da descoberta de processos. Essa constatação era esperada, uma vez que os interesses de mineração dizem respeito a busca por conhecimento que é inexistente dentro de um contexto. Também, foi constatado que as tarefas mais comuns em Mineração de Dados (predição categórica e análise de agrupamentos) aparecem com alta frequência, juntamente com tarefas que, em princípio, são as mais adequadas para o contexto de análise de processos

⁵⁸Mineração de sequências pode ser considerada como uma especificidade da tarefa de descoberta de regras de associação ou da análise de evolução.

⁵⁹Episódios podem ser vistos como conjuntos de sessões (em uma aplicação qualquer) que possuem um significado para o contexto em que estão inseridos. A área de mineração na WEB faz uso dessa nomenclatura para tarefas de Mineração de Dados.

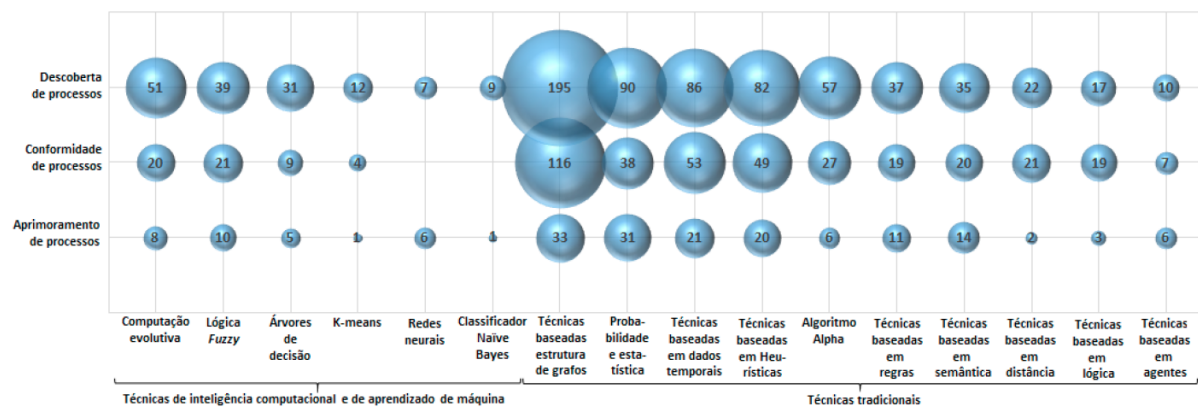
⁶⁰Maita, A. R. C., Martins, L. C., Paz, C. R. L., **PERES, S. M.**, Fantinato, M. Process mining through artificial neural networks and support vector machines: A systematic literature review. *Business Process Management Journal*, Emerald, p. 1391–1415, 2016.

Figure 29: Distribuição cruzada entre tarefas de Mineração de Processos e tarefas de Mineração de Dados [150]



(análise de estruturas, descoberta de associações e análise de evolução). Tais constatações não são reveladoras, mas dentro do processo sistemático usado no mapeamento, servem como confirmação de hipóteses e impressões.

Figure 30: Distribuição cruzada entre tarefas de Mineração de Processos e técnicas aplicadas para resolvê-las [150]



A partir da informação ilustrada pela figura 30, também são feitas algumas confirmações, como a predominância de técnicas de análise que são baseadas em algoritmos provenientes da teoria dos grafos, e uma concentração relativamente importante para o uso de um algoritmo em especial: o Algoritmo Alpha [153]. Esse algoritmo foi um dos primeiros especificamente desenvolvidos para uso em tarefas de Mineração de Processos. Seu objetivo é reconstruir as relações causais de um conjunto de seqüências de eventos. Ele é capaz de identificar um modelo de fluxo de trabalho baseado em Redes de Petri, e para ele já existem uma série de extensões e modificações para atendimento de questões

específicas do contexto de Mineração de Processos. Porém, percebe-se pelos dados organizados na figura 30 que as técnicas de Inteligência Computacional e de Aprendizado de Máquina não são amplamente exploradas nesse contexto. Naturalmente, tais técnicas podem ser, de fato, menos adequadas ao tipo de problema que a Mineração de Processos oferece, no entanto, percebe-se que há um nicho para pesquisa que merece receber atenção dos pesquisadores experientes nessas disciplinas.

A motivação para afirmar que existe potencial ainda não explorado no que diz respeito ao uso de técnicas de Inteligência Computacional e Aprendizado de Máquina, e também na modelagem de tarefas de Mineração de Processos sob a visão da organização em tarefas de Mineração de Dados, decorre também de uma constatação obtida com o desenvolvimento do mapeamento sistemático (tabelas 10 e 11). Foi possível verificar que a maioria das publicações da área ocorre em veículos (periódicos e conferências) que não são de áreas específicas das disciplinas citadas. Considerando apenas os dados apresentados nas tabelas, são 38 artigos em periódicos fora da área de Mineração de Dados, Inteligência Computacional ou Aprendizado de Máquina, contra 24 artigos publicados em periódicos dessas disciplinas. No caso das conferências, seguindo a mesma sequência, são 138 contra 45. Isso pode estar ocorrendo ou devido à improvável falta de interesse dos veículos de Inteligência Computacional e Aprendizado de Máquina em publicar estudos dessa área, ou devido ao provável desconhecimento, por parte dos pesquisadores nessas disciplinas, do interessante nicho de pesquisa que a área de Mineração de Processos está disponibilizando.

Table 10: Os sete periódicos que mais publicaram estudos relacionados a Mineração de Processos de 2005 a 2014 [150]

Sigla	Título do periódico	# de estudos
ESA	Expert System with Applications	19
JIS	Information Systems	17
DKE	Data and Knowledge Engineering	9
DSS	Decision Support Systems	8
IJBPM	Int. J. of Business Process Integration and Management	8
TKDE	IEEE Trans. on Knowledge and Data Engineering	6
SMC	IEEE Trans. on Systems, Man, and Cybernetics	5

Ainda com respeito à modelagem de problemas encontradas na literatura levantada no mapeamento sistemático, há que se fazer uma ressalva importante. Uma parte dos estudos tratam descoberta de conhecimento a partir de *log* de eventos como sendo uma tarefa ou uma aplicação da área de Mineração de Processos, ainda que a modelagem do problema tratado não esteja especificamente, ou diretamente, aderente ao entendimento sobre descoberta de um processo na forma de uma estrutura equivalente a um modelo de processo, ou na análise da conformidade entre processos a partir do uso de um modelo de

Table 11: As 17 conferências que mais publicaram estudos relacionados a Mineração de Processos de 2005 a 2014 [150]

Sigla	Título do periódico	# de estudos
BPM	Int. Conf. on Business Process Management	61
CAiSE	Int. Conf. on Advanced Information Systems Engineering	16
SIMPDA	Int. Symp. on Data-Driven Process Discovery and Analysis	11
Petri Nets	Int. Conf. on Application and Theory of Petri Nets and Concurrency	10
CIDM	IEEE Symposium on Computational Intelligence and Data Mining	9
SAC	ACM Symp. on Applied Computing	8
BPMDS	Int. WS on Business Process Modeling, Development and Support	8
CEC	IEEE Cong. on Evolutionary Computation	7
EDOC	IEEE Int. Enterprise Distributed Object Computing Conference	7
ICT&KE	Int. Conf. on ICT and Knowledge Engineering	7
BIS	Int. Conf. on Business Information Systems	6
BPI	Int. WS on Business Process Intelligence	6
ICDM	IEEE Int. Conf. on Data Mining	6
ICEIS	Int. Conf. on Enterprise Information Systems	6
CSCWD	IEEE Int. Conf. on Computer Supported Cooperative Work in Design	5
ICEBE	IEEE Int. Conf. on e-Business Engineering	5
PAKDD	Pacific-Asia Conf. on Advances in Knowledge Discovery and Data Mining	5

processo. É comum nesses estudos fazer uso da informação sobre o processo, extraída do *log* de eventos, para descobrir conhecimento útil para alguma tomada de decisão também ligada ao processo. Essa visão de Mineração de Processos é marginal à visão clássica estabelecida pelo grande grupo de pesquisa da *Technische Universiteit Eindhoven*, e se aproxima da visão mais abrangente de “mineração” adotada pela área de Mineração de Dados.

4.3 A tarefa de classificação sobre um processo não estruturado

Como primeiro esforço da autora em modelar uma tarefa de Mineração de Dados no contexto de Mineração de Processos, e resolvê-la usando uma técnica de Inteligência Computacional ou Aprendizado de Máquina, foi explorado um problema de análise de dados via *log* de eventos proveniente de um processo não estruturado inerente ao uso de um sistema de apoio ao ensino semipresencial. Os resultados obtidos nesse esforço estão

detalhados em [150] e em [154]⁶¹.

O domínio da aplicação usado nesse primeiro estudo aplicado é referente ao processo de uso de um ambiente virtual de aprendizado (AVA), cujos usuários são alunos de um curso de especialização semipresencial, e que foi implementado sobre a plataforma Moodle. Essa plataforma persiste dados sobre as interações dos usuários com os objetos de aprendizagem disponíveis. Na instância de uso do AVA analisada, cerca de 1.500 usuários interagiram com a plataforma, por cerca de um ano. O modelo de processo referente a esse contexto nunca tinha sido formalizado, embora os coordenadores pedagógicos do curso tinham uma expectativa sobre qual seria o processo seguido pelos alunos. No estudo aqui descrito, foi considerada apenas uma parte dos alunos matriculados em uma das ofertas do curso e o intervalo de tempo de cinco meses de registros de *log*. Informações mais específicas sobre o ambiente e seu processo de desenvolvimento foram descritos por [155]; mais detalhes do estudo realizado sobre os registros de *log* desse ambiente podem ser encontrados em [150] e [154]. A semântica dos dados que compõem o *log* de eventos usados no estudo são:

- identificador da instância de processo: uma instância do processo é uma sequência de eventos (ações do usuário) relacionada a uma sessão de uso do AVA, gerando várias entradas no *log*, cada uma referente à visita a um objeto de aprendizado disponível ou a uma seção do ambiente;
- marcador de tempo (*timestamp*): a data e a hora de execução de um evento;
- marcador de usuário: usuário que executa a ação;
- atividade: descrição da ação executada pelo usuário (na forma de um tipo enumerado).

Além da informação sobre o comportamento do usuário no sistema, o estudo também contou com a informação sobre o resultado de seu desempenho no curso. Esse resultado foi categorizado em: *aprovação* e *reprovação*; ou *ótimo*, *regular* e *insatisfatório*. Trata-se, portanto, de um estudo que estende a análise para além do *log* de eventos. A categorização foi usada indiretamente na primeira estratégia de análise, na qual o *log* de eventos completo foi dividido em partes referentes ao uso do sistema por alunos que se encaixaram em cada uma das categorias, e para cada um deles, modelos de processos foram descobertos e analisados; e foi usada diretamente na segunda estratégia de análise, na qual um modelo de predição categórica foi criado.

⁶¹Maita, A. R. C., Fantinato, M., **PERES, S. M.**, Thom, L. H., Hung, P. C. K. Mining Unstructured Processes: An Exploratory Study on a Distance Learning Domain. In: Proceedings of 30th International Joint Conference on Neural Networks. Anchorage, AK, USA: IEEE, 2017, p. 1–8. Accepted to.

O problema foi primeiramente estudado sob o ponto de vista da descoberta do processo (uma das tarefas clássicas de Mineração de Processos), a partir da aplicação de ferramentas de Mineração de Processos, a saber: Disco Fluxicon e ProM⁶². O resultado da aplicação dessas ferramentas para descoberta do processo presente no *log* de eventos resultou na apresentação da sequência de atividades mais comumente executadas pelos estudantes no AVA. Cada uma das ferramentas permite criar essas representações com mais ou menos detalhes, a depender da granularidade de informação (nesse caso, frequência mínima de ocorrência de sequências de atividades) desejada pelo usuário. Exemplos das representações criadas são mostrados na figura 28. Em termos de descoberta de conhecimento a partir da descoberta do modelo de processo presente no *log* de eventos, esses resultados permitiram: ilustrar características de um processo não estruturado; evidenciar um nível de estrutura de processo que emergiu do comportamento dos alunos na plataforma (suportando tomada de decisão de coordenadores de curso); analisar a conformidade entre processos descobertos para recortes de alunos (por exemplo, a conformidade entre o processo dos alunos que são reprovados no curso com o processo dos alunos que são aprovados no curso), e a conformidade entre os processos descobertos e as expectativas dos coordenadores do curso.

Contudo, analisando a situação sob o ponto de vista da Mineração de Dados, e considerando o contexto de tarefas preditivas, o problema de Mineração de Processos no contexto desse estudo foi também modelado como uma tarefa de classificação, e resolvido a partir do uso de redes neurais artificiais MLP. O *log* de eventos foi transformado em um conjunto de dados cuja distribuição segue uma representação de conjuntos de dados identicamente e independentemente distribuídos, e a informação de categoria foi incorporada ao procedimento de indução de um modelo classificador binário. O problema de classificação sobre o *log* de eventos transformado permitiu a indução de um modelo capaz de prever o resultado de aprovação ou não aprovação do aluno no curso a partir da análise de seu comportamento no sistema. Nesse caso, o processo real de uso do AVA, referente a cada categoria de alunos, suporta a criação de um mapeamento que implicitamente representa um modelo de processo.

O uso desse mapeamento para análise do resultado de aprovação ou não aprovação de

⁶²Essas são duas ferramentas populares na área de Mineração de Processos, como pode ser observado nos resultados do mapeamento sistemático [150]. A ferramenta Disco Fluxicon (<http://www.fluxicon.com>) foi usada em 12 artigos analisados no mapeamento. É uma ferramenta proprietária que disponibiliza versões simplificadas, de uso livre, para fins acadêmicos. Segundo seus fabricantes, seu principal algoritmo é chamado “*Fuzzy Miner*”, indicando o uso de uma técnica da Inteligência Computacional. A ProM é uma ferramenta extensível e de código aberto, que suporta diferentes tipos de, e técnicas para, Mineração de Processos, disponibilizados na forma de *plug-ins*. Essa ferramenta foi utilizada em 277 artigos analisados no mapeamento, muitos dos quais eram referentes à apresentação de *plug-ins*. Outras ferramentas disponíveis para uso em Mineração de Processos são mencionadas em [156] e [150], incluindo ferramentas que são de propósito geral para o contexto de Mineração de Dados.

um aluno está baseado na hipótese de que há um padrão de comportamento de acesso aos objetos de aprendizagem e às seções do AVA que leva um aluno a obter conhecimento suficiente para sua aprovação no curso. Considerando uma possibilidade real de avaliar um aluno durante o andamento do curso, o comportamento dele no sistema AVA pode servir como um indicativo de sucesso ou não no curso e pode ser usado para tomada de decisões referentes a aconselhamento ou instrução comportamental aos alunos.

Os resultados de classificação obtidos a partir da criação de modelos com MLP, sob uma estratégia de teste baseada em validação cruzada, alcançaram taxas em torno de 0,9 para precisão e revocação, e em torno de 0,5 para preditividade negativa e especificidade. Esses resultados precisam ser melhorados antes de efetivamente serem usados no contexto real, porém, no atual estágio da pesquisa, tais resultados já apontam para a razoabilidade de modelar tarefas de Mineração de Dados como suporte para a resolução de problemas de Mineração de Processos, num sentido, neste caso, de modelagem indireta do processo.

Os conjuntos de dados usados na indução do modelo classificador possuem uma representação orientada aos alunos e não orientada às atividades (como é originalmente um *log* de eventos). Devido a essa característica, a modelagem de classificação apresentada pode não ser aceita como aderente às definições clássicas de Mineração de Processos. Além disso, a transformação dos dados orientados a atividades para dados orientados a usuários causa alguma perda de informação semântica referente aos processos presentes no *log* de eventos. Isso mostra a necessidade de estudar com mais profundidade o problema de representação de informação e do conhecimento para que seja possível apresentar um uso de técnicas de Aprendizado de Máquina mais aderente às prerrogativas básicas da Mineração de Processos.

4.4 Objetivos atuais e limitações da pesquisa

Atualmente, pode ser considerado que a principal contribuição científica produzida pela autora no contexto de análise de dados do tipo *log* de eventos é o mapeamento sistemático que posiciona, dentro da área de Mineração de Processos, o arcabouço de definições estabelecido para a área de Mineração de Dados, de acordo com o que vem sendo desenvolvido pela comunidade de pesquisa internacional. Essa contribuição, no entanto, não diz respeito diretamente ao tratamento do dado de natureza não estruturada representado no *log* de eventos. No que diz respeito a esse objetivo, a autora tem já produzido um estudo que ilustra as dificuldades e potencialidades de trabalhar com uma representação da informação extraída do *log* de eventos junto à modelagem de uma tarefa de classificação resolvida por uma técnica de aprendizado indutivo. Os resultados desse estudo estão validados pela comunidade internacional da área de redes neurais artificiais, e foi, nesta análise, elogiado por trazer para uma comunidade especialista um problema que

tem muitas questões em aberto e que merece receber a atenção de estudos provenientes de tal área de especialidade.

A partir das ações de inserção na área de pesquisa multidisciplinar objeto de discussão neste capítulo, vários caminhos se abriram e estão no foco de atenção da autora para curto e médio prazo. Dentro desses caminhos, alguns objetivos estão traçados:

- Ao estudar problemas referentes a contextos reais, é comum encontrar situações nas quais os processos são essencialmente semiestruturados ou não estruturados. Para esses casos, diante do que a autora tem observado em trabalhos em andamento, é essencial considerar informações sobre os processos que estão fora do escopo do *log* de eventos. Tais informações são relacionadas ao processo sob mineração, mas estão persistidas em bancos de dados, geralmente relacionais. Assim, uma linha de atuação que está sendo seguida pela autora é referente a trabalhar com: o enriquecimento do *log* de eventos por meio da incorporação de atributos descritivos dos processos; a extração de características temporais; e o uso de representações alternativas para as características temporais ou o uso de raciocínio temporal.
- A partir do momento em que há a possibilidade de enriquecimento do *log* de eventos, problemas referentes ao tratamento de alta dimensionalidade de dados surgem e precisam ser tratados. Para isso, pretende-se trabalhar com o problema de seleção de atributos ou de seleção de características, a partir da aplicação assistida de filtros, técnicas de invólucro (*wrappers*) ou técnicas de seleção *embedded* [157, 158, 159]. Técnicas para seleção de atributos ou características são comumente usadas nos problemas tratados via modelagem de tarefas de Mineração de Dados. Especificamente em relação a técnicas de invólucro ou *embedded*, algoritmos e técnicas de Inteligência Computacional e Aprendizado de Máquina tem alta aplicabilidade e, portanto, estão no foco da atenção da autora para essa linha de trabalho.
- Especificamente para a implementação de mineração de processos não estruturados, há uma deficiência de desenvolvimento de soluções. Até onde foi possível analisar dentro da pesquisa da autora, os principais algoritmos, e portanto os mais usados e disseminados em artigos científicos e na implementação de ferramentas, têm limitações para atuar sobre processos que não possuem relações causais bem definidas, que possuem muitos caminhos ou sequência de ações fugindo do padrão mais frequente de comportamento de processo, que sofrem de influências externas de competência humana ou possuam dependências do comportamento de variáveis temporais. Para essas situações, algumas soluções heurísticas e *ad hoc* têm sido adotadas, com algum investimento em meta-heurísticas (principalmente algoritmos genéticos) que permitam a generalização de soluções. Esse é um nicho de pesquisa

que é foco de atenção da autora. Atualmente, há um trabalho em andamento que faz uso da computação evolutiva e de princípios de aprendizado indutivo (a base do aprendizado de máquina), para evoluir ou induzir padrões de comportamento (modelos de processo) que são relevantes (não necessariamente frequentes) e estão escondidos no *log* de eventos.

Como consequência futura do desenvolvimento dos trabalhos de pesquisa relacionados aos objetivos supracitados, duas contribuições técnicas são esperadas a médio prazo, com intenção de manter uma regularidade de evolução em relação aos artefatos relacionados a elas:

- Construção e publicação de *logs* de eventos de referência: para a evolução das áreas de Inteligência Computacional e Aprendizado de Máquina, em termos de validação de sua aplicação em contextos reais, assim como para teste e verificação de suas potencialidades em relação a problemas de diferentes características, é imprescindível a existência de conjuntos de dados de referência. A aceitação da resolução de um problema, sob algum aspecto, a partir do uso de técnicas dessas áreas, ou é provada teoricamente ou deve ser demonstrada por meio de experimentos reproduzíveis. Para alcançar a plena reprodutibilidade, os conjuntos de dados usados nas experimentações, principalmente quando derivados de fenômenos reais, precisam estar disponíveis publicamente. Há várias iniciativas de organização de conjuntos de dados de referência que recebem contribuições da comunidade científica. A disponibilização de conjuntos de dados reais nesses repositórios é uma contribuição técnica que naturalmente decorre dos trabalhos de pesquisa da autora em suas outras linhas de atuação e deverão também passar a ocorrer na linha de atuação em discussão neste capítulo, em termos de *log* de eventos.
- Desenvolvimento e publicação de *plug-ins* especialistas: por outro lado, para a comunidade de Mineração de Processos é comum encontrar artigos que apresentam *plug-ins*, principalmente para a ferramenta ProM, de forma que o algoritmo ou a técnica proposta seja facilmente usada por outros pesquisadores ou práticos, em seus problemas particulares, como uma forma de atestar a eficiência da estratégia proposta. Esse é um procedimento que a autora não tem adotado em suas outras linhas de atuação, mas que é pertinente para o caso da linha de atuação em discussão neste capítulo. As estratégias de seleção de atributos e características, bem como a estratégia de evolução de padrões de comportamentos relevantes, são passíveis de desenvolvimento de *plug-ins* de propósito geral para a área. Ambas estratégias têm potencial para serem aceitas pela comunidade que gerencia a ferramenta ProM e, portanto, os respectivos *plug-ins* deverão ser desenvolvidos e submetidos a ela.

A pesquisa aqui discutida está sujeita a duas grandes limitações, sendo uma delas de caráter operacional e a outra de caráter conceitual. A primeira diz respeito à dificuldade em encontrar *log* de eventos para uso nas experimentações. Para tratamento de alguns aspectos inerentes ao problema de modelagem de processo, quando do tratamento da descoberta do modelo relacionado, há alguns *logs* de eventos produzidos de maneira artificial. Para o tratamento da escalabilidade e da aderência ao mundo real, são poucos os *logs* de eventos publicamente disponíveis. Ainda, a busca por *log* de eventos reais tem alguns limitadores específicos: são poucos os sistemas computacionais que possuem a funcionalidade de persistência das sequências de eventos que ocorrem durante a execução dos processos que eles implementam; alguns sistemas que são instrumentados para tal são parametrizáveis e nem sempre são habilitados pelos usuários para persistir todas, ou algumas das, informações possíveis para construção do *log* de eventos; e quando o *log* de eventos existe, ele pode conter informações sensíveis sobre o negócio das empresa correlatas e por isso elas não tem interesse em disponibilizá-los para estudo. Diante de tais dificuldades, e aliado ao fato da pesquisa proposta pela autora ser altamente aplicada e dependente da existência de massa de dados, há um fator limitante referente à qualidade dos dados que são usados na pesquisa e também um fator de risco referente a possibilidade de inexistência de dados adequados para desenvolvimento da pesquisa.

A limitação de caráter conceitual diz respeito à aceitação das modelagens em proposição como sendo soluções para as tarefas clássicas de Mineração de Processos. Linhas de atuação (pesquisadores e veículos de disseminação científica) podem entender que a definição dos problemas está marginalmente associada à área de Mineração de Processos e esse fato pode eventualmente prejudicar a publicação dos resultados obtidos.

5 Considerações finais

Este texto teve o objetivo de sistematizar os trabalhos científicos e acadêmicos desenvolvidos pela autora no contexto de sua pesquisa em análise de dados não estruturados, considerando principalmente suas contribuições com uso de técnicas da Inteligência Computacional e do Aprendizado de Máquina. O conteúdo apresentado seguiu pela explanação da visão da autora sobre os dados com os quais trabalha (gestos, textos e *logs* de eventos), pela definição dos problemas que trata em relação a cada tipo de dado e pela proposição de soluções e apresentação dos principais resultados obtidos para tais problemas. Nos três capítulos que tratam cada uma das linhas de atuação (capítulos 2, 3, 4), os objetivos atuais em pesquisa assim como as limitações agora percebidas em cada contexto foram apresentados. Especificamente no capítulo 4, questões referentes à Mineração de Dados foram discutidas principalmente devido à natureza da área de aplicação tratada (Mineração de Processos).

Cabe ainda, para a sistematização construída neste texto, realizar uma reflexão crítica sobre o que já foi alcançado no tempo de pesquisa a que se refere o presente conteúdo, o que possui potencial para ser desenvolvido, e o que já poderia ter sido alcançado, mas que por diferentes motivos, não fora ainda possível. Para isso, duas seções foram criadas neste capítulo: uma dedicada a discutir o aspecto científico (Seção 5.1) e outra dedicada a discutir o aspecto técnico (Seção 5.2). Nessas seções, menções a desafios e oportunidades de cunho econômico, social e operacional da pesquisa são realizadas. Na sequência, completando o caráter acadêmico decorrente do trabalho realizado pela autora, duas seções são dedicadas a apresentar: o trabalho realizado em uma longa parceria em pesquisa no campo da análise de dados não estruturados aplicados a resolução de problemas em Biometria; e o reflexo do trabalho em pesquisa na atuação da autora como docente na área de análise de dados. Por fim, agradecimentos são devidamente realizados aos agentes que contribuem diretamente para a realização, com qualidade, de todo o trabalho sob discussão.

5.1 Análise do aspecto científico

A linguística, como definido por [160] e seguindo o discutido por [161], tem o objetivo de fornecer uma orientação teórica e formal às noções primitivas, operações elementares, regras e esquemas que organizam uma língua. Ainda segundo esta linha de pensamento, no estudo linguístico busca-se por invariantes e objetiva-se entender o que é generalizável de forma que seja possível identificar o que pode ser reconstruído sempre que necessário. Nesse sentido, a pesquisa aqui discutida colabora no momento em que tenta demonstrar, por meio do esforço computacional de geração de modelos de generalização, o que é, pelo menos computacionalmente, generalizável e sob quais condições e perspectivas.

Os resultados que têm sido obtidos na análise de gestos até o presente momento não estão, ainda, disponíveis para uso prático pela comunidade científica de qualquer área. Esse é um ponto fraco da pesquisa em desenvolvimento. Embora já exista um projeto de construção de uma aplicação computacional para disponibilizar os modelos de segmentação já criados aos pesquisadores interessados em análise das fases do gesto, ela ainda não foi desenvolvida. Há dois motivos para isso: a carência de recurso humano que esteja interessado em realizar esse trabalho; a necessidade de realizar mais estudos para delimitar as capacidades dos modelos de segmentação em diferentes cenários de uso.

Diferentes cenários de uso para os modelos de segmentação das fases do gesto fazem parte de um estudo em andamento orientado pela autora. O objetivo do estudo é delimitar o escopo no qual um modelo é capaz de generalizar a tomada de decisão sobre a segmentação das fases do gesto considerando: diferentes gesticuladores, diferentes sessões de gesticulação e diferentes objetivos de gesticulação. Em paralelo a esse estudo, a questão

relacionada à forma de representação dos dados gestuais também está em estudo com o intuito de encontrar novos atributos descritores, combinações dos mesmos que sejam úteis para os modelos classificadores, uso de extratores baseados em técnicas que atuam no domínio do tempo, frequência e tempo-frequência e uso de técnicas de discretização de séries temporais. Ainda, o estudo engloba a modelagem de tarefas de agrupamento sobre esses dados, pautando-se na hipótese de que há conhecimento novo sobre as fases incorporado nas representações usadas. Há uma expectativa inerente a todo esse esforço que está relacionada à descoberta de uma caracterização eficiente para cada uma das fases do gesto.

Um estudo mais amplo para a análise automática das expressões faciais gramaticais também é necessário e está nos planos da autora. No contexto da Libras há alguns aspectos que devem ser levados em consideração para atestar a capacidade de generalização dos modelos. Tais aspectos dizem respeito a: fluência na língua, regionalismos e dialetos, diferença natural entre usuários e deficiências na fala (que também se manifestam no uso da língua de sinais).

Uma deficiência da pesquisa em análise de gestos da autora é não considerar o uso de imagens RGB, as quais existem em abundância na internet e poderiam ser usadas para suprir a carência de dados para desenvolvimento dos modelos, tanto para a análise da gesticulação natural quanto para a análise das expressões faciais gramaticais da Libras. A justificativa para isso é que a autora não é especialista em processamento de imagens. No entanto, parcerias em pesquisa poderiam suprir essa lacuna e esta ação está nos planos da autora.

Já a continuidade da pesquisa em termos de reconhecimento de outras primitivas da Libras (configuração de mão, movimentos da mão e direção da mão) está prevista apenas em termos de aprimoramento e desenvolvimento de aplicações, a partir do estabelecimento de pequenos projetos de pesquisa para alunos de graduação. A maturidade obtida na pesquisa já realizada, tanto pela autora quanto pela comunidade científica em geral, permite a mudança de fase, partindo então para desenvolvimento de aplicações decorrentes do conhecimento obtido.

A pesquisa em análise de gestos discutida neste texto, tanto no que diz respeito à análise das fases do gesto quanto à análise das expressões faciais gramaticais da Libras, é pioneira na área de Computação ou Engenharia no Brasil. Até onde vai o conhecimento da autora, não há qualquer outro grupo brasileiro, ou pesquisador isolado, nessas áreas do conhecimento, que têm se debruçado sobre o entendimento dos gestos humanos sob a perspectiva da Teoria dos Gestos ou que tem dado atenção à primitiva “expressões faciais” da Libras.

A pesquisa em análise de dados textuais, embora mais recente, aponta para resultados

mais direcionados à evolução de algoritmos e técnicas computacionais. Os primeiros resultados obtidos, a partir da orientação de uma dissertação de mestrado, dizem respeito à proposição de alterações relevantes em algoritmos e também a demonstração de suas provas de convergência (a exemplo dos algoritmos originais). Essa linha de pesquisa, atrelada também à área de aplicação de Sistemas de Recomendação, está atualmente, recebendo os principais esforços da autora. A modelagem de tarefas de coagrupamento sobre textos seguida da extração de informação dos resultados obtidos para aplicação em Sistemas de Recomendação não é comumente encontrada na literatura correlata. A interpretação que se pretende dar para os cogrupos no contexto dessa pesquisa aparenta ter um bom potencial para uso em recomendação. Aliado a isso, esforços no sentido de explorar também representações textuais alternativas como as representações distribuídas de palavras [162] aumentam o escopo de exploração nessa linha de pesquisa.

A pesquisa em análise de textos e a aplicação em Sistemas de Recomendação é de interesse da indústria que desenvolve aplicativos para análise de dados textuais presentes na internet. Motivados pelo interesse da indústria, muitos alunos que procuram o curso de pós-graduação em Sistemas de Informação (PPgSI) se mostram muito interessados em atuar nessa área, para a qual, então, a autora tem investido na proposição de hipóteses de pesquisa. Outro aspecto que tem motivado o desenvolvimento dessa pesquisa é a proposição de soluções que trabalhem bem no âmbito da língua portuguesa. Há muito desenvolvimento para a língua inglesa, mas a eficiência de tais soluções na língua portuguesa carece de investigação.

No que diz respeito a futuros trabalhos na linha de análise de textos, para além das iniciativas baseadas em agrupamento e coagrupamento, com interpretação baseada em similaridade parcial, é intuito da autora investir na construção de modelos que fazem uso de técnicas de aprendizado profundo. Na indústria de software, há muito investimento e otimismo quanto ao uso dessas técnicas. Do ponto de vista científico, obviamente, há espaço para atuação em termos de proposição de melhorias, aplicações ou interpretações inéditas, contudo, há também o anseio por aplicar tais técnicas sem que seja necessário fazer uso de muito poder de processamento computacional.

A inserção da autora na área de Mineração de Processos foi motivada pela possibilidade de realização de uma pesquisa multidisciplinar. As hipóteses de pesquisa em Mineração de Processos estão localizadas na fronteira entre as duas linhas de pesquisa que caracterizam o PPgSI: a linha “inteligência de sistemas” e a linha “gestão e desenvolvimento de sistemas”. É uma área de pesquisa já bem difundida no campo de “gerenciamento de processos de negócio”, mas bem pouco disseminada no campo da “mineração de dados”. No Brasil, por exemplo, há alguns pesquisadores estudando Mineração de Processos, porém até onde foi possível apurar, são pesquisadores atuantes na

linha de gestão e desenvolvimento de sistema. Não é de conhecimento da autora, até o presente momento, nenhuma iniciativa brasileira em mineração de processos que venha de pesquisadores atuantes na linha de inteligência de sistemas (ou mineração de dados). Considerando esse cenário, a pesquisa multidisciplinar desenvolvida pela autora, em colaboração a outros pesquisadores do PPgSI e também externos ao programa (brasileiros e estrangeiros) cria uma linha de pesquisa inédita no contexto da Computação no Brasil.

Essa área de atuação tem atraído a atenção de alunos de graduação e pós-graduação que já atuam ou já atuaram na indústria. Em face desse cenário, os trabalhos atualmente em desenvolvimento pela autora e colaboradores tem um forte aspecto aplicado, estando todos inseridos em contextos referentes a problemas reais. Esse aspecto, porém, insere uma dificuldade operacional aos trabalhos, a qual é comentada na seção 5.2.

5.2 Análise do aspecto técnico

A análise de gesticulação natural pretendida na pesquisa em discussão possui alguns limitadores econômicos e operacionais. Na área da Linguística ou na Semiótica, nas quais a análise de gestos sob a luz da Teoria dos Gestos é mais difundida, costuma-se fazer uso de sensores de captação de movimentos de alto custo monetário. É comum que tais pesquisas façam uso de sistemas Vicon⁶³ o qual é capaz de sensores um ambiente 3D com alta precisão e acuidade, porém possui um custo elevado e uma operação não trivial. No contexto da pesquisa da autora, esse sistema de alta capacidade é substituído por sensores de baixo custo (sensores Microsoft Kinect) e de operação mais simples. Se por um lado, o uso de sensores simples de baixo custo aproxima os resultados da pesquisa à oportunidades de desenvolvimento de aplicações reais que de fato possam ser usadas em situações do cotidiano, e também permite escalar tais aplicações, por outro lado, os dados sob análise possuem uma qualidade inferior que pode afetar o desempenho das técnicas em uso e traz um dificultador a mais para o desenvolvimento dos modelos de análise. Já a limitação operacional da pesquisa diz respeito à necessidade de produção de conjuntos de dados de referência. Tal produção tem sido feita com o intuito de capturar a gesticulação natural, porém são poucos os cenários já produzidos que de fato permitem naturalidade. Além disso, os conjuntos de dados precisam ser anotados (ou rotulados) e além da anotação exigir conhecimento especialista, ela exige um tempo de dedicação que muitas vezes está além do que os especialistas podem dispor.

No que diz respeito à análise automática das expressões faciais gramaticais da Libras, destaca-se o mesmo limitador econômico relacionado ao sistema Vicon. Também, o sistema do sensor Microsoft Kinect tem suportado a aquisição dos dados da face para

⁶³Vicon é uma empresa que desenvolve produtos de captura de movimentos e presta serviços correlatos a seus sistemas. Mais informações em <https://www.vicon.com>.

análise. O limitador operacional referente à aquisição de dados e anotação dos mesmos é agravado no contexto da Libras, visto que nesse caso é necessário dispor de pessoas que conheçam a língua de sinais e que tenham disponibilidade para atuar na pesquisa. Esse é um cenário ainda mais complicado que o cenário das fases do gesto, visto que além de especialistas para anotação dos conjuntos de dados são também necessários especialistas para execução das falas em Libras.

Para a pesquisa referente a análise de dados textuais e sistemas de recomendação, no que concerne ao estudo sobre textos em língua portuguesa, um limitador da pesquisa é a inexistência de conjuntos de dados de referência. Existem poucos *corpora* textuais preparados para o desenvolvimento de algoritmos de Aprendizado de Máquina e, no que diz respeito a sistemas de recomendação, até o presente, a autora desconhece conjuntos de dados de referência que possam embasar avaliação de algoritmos, e comparações entre algoritmos, que façam uso de históricos de recomendação ou históricos de aceitação da recomendação baseado em itens que estão em língua portuguesa. Os trabalhos brasileiros desenvolvidos nessa área ou usam conjuntos de dados que não estão publicamente disponíveis ou tem sua contribuição atestada sobre conjuntos de dados provenientes de pesquisas ou aplicações em língua inglesa.

Os dados necessários para a realização de estudos em Mineração de Processos são provenientes de sistemas reais, ou seja, que de fato estão sendo usados dentro de algum processo de negócio. Por isso, obter dados para realização de estudos impõe duas dificuldades operacionais: a primeira diz respeito à sensibilidade dos dados, já que as empresas não querem ou não podem disponibilizar acesso aos dados completos; a segunda diz respeito ao problema de dados faltantes quando se quer usar *logs* de eventos enriquecido com informações transacionais, como é o caso de um dos estudos em desenvolvimento pela autora. Essa dificuldade tem sido superada com ações *ad hoc*, realizadas para cada um dos trabalhos em desenvolvimento. Contudo, há uma dificuldade em conseguir autorização para disponibilização dos dados para propiciar a reprodutibilidade dos experimentos realizados.

Por fim, uma das oportunidades técnicas que a pesquisa em análise de gestos e também em Mineração de Processos propicia, é a criação de aplicativos ou *plug-ins* que podem ser disponibilizados livremente para a comunidade como um todo, mas que também são passíveis de solicitação de registro de software. O registro de software é um objetivo que está na agenda da autora por entender que ele suporta, adequadamente, o registro da autoria decorrente de pesquisa realizada na Universidade de São Paulo.

5.3 Parceria em pesquisa: análise de dados biométricos

Esta seção trata de uma colaboração em pesquisa mantida junto a um professor pesquisador atuante no PPGSI, mesmo programa em que atua a autora. O pesquisador em questão possui uma forte atuação na área de sistemas biométricos, e no ano de 2015 iniciou seus trabalhos em uma linha de pesquisa referente a biometria comportamental usando movimentos da marcha (caminhar), se aproximando ainda mais dos interesses de pesquisa da autora. Devido à proximidade de interesses em análise de dados não estruturados provenientes de movimentos corporais ou características corporais, esta parceria se encaixa como parte inerente da pesquisa discutida neste documento. No entanto, por não se tratar de pesquisa liderada pela autora, o contexto da pesquisa e as produções geradas são apenas citados aqui. Esta parceria se dá em termos de coorientações em programa de mestrado, coautorias em artigos científicos e colaborações em grupos de discussão.

Os trabalhos desenvolvidos nessa linha de atuação podem ser divididos em duas sublinhas de pesquisa: biometria baseada em análise de sinais biomédicos (eletromiograma, eletrocardiograma e eletroencefalograma), e biometria baseada em análise de imagens (face, íris e forma de andar). No que diz respeito a dados biomédicos, os resultados obtidos com a contribuição da autora são principalmente referentes à aplicação de SVM e variações para resolução de problemas de predição [163, 164, 165]. Há também alguns resultados referentes a seleção de características com algoritmos genéticos [166, 167], e uso de representação de dados baseada em discretização de séries temporais usando *symbolic aggregate approximation* [168, 169]. Já na sublinha de pesquisa que envolve análise de imagens, os principais dados se referem a reconhecimento de íris [170, 171] e de face [172, 173], incluindo o uso em sistemas biométricos multimodais [174, 175], e mais recentemente alguns resultados na área de biometria com fusão de face e forma de andar [176].

5.4 Contribuição didática

Dos estudos realizados no desenvolvimento da pesquisa aqui discutida, deriva grande parte do conhecimento usado para o trabalho docente. A elaboração das aulas ministradas na graduação e pós-graduação pela autora tem tido sua qualidade elevada a cada ano, visto que diante de tantos problemas de pesquisas estudados, das soluções elaboradas, das buscas por soluções alternativas, da busca por métodos de avaliação, etc., aumenta a maturidade do entendimento dos meandros das técnicas de Inteligência Computacional e Aprendizado de Máquina, e também melhora a experiência com modelagem de tarefas de Mineração de Dados. Os artefatos mais comuns produzidos no trabalho de docência nas áreas já citadas são os *slides* de suporte às aulas, as especificações de trabalhos práticos

e a produção de listas de exercícios. Em sua maioria, esses artefatos estão disponíveis publicamente na homepage da autora⁶⁴.

Além disso, vale apresentar neste texto de sistematização de pesquisa, duas produções acadêmicas que tem o potencial de impactar a docência não só da autora como de outras pessoas que se propuserem a usar tais materiais. A primeira produção, resultante de um trabalho de conclusão de curso orientado pela autora, é um tutorial referente aos algoritmos *Fuzzy-c-Means* e *Fuzzy Learning Vector Quantization* [177]⁶⁵. Ambos algoritmos foram estudados em detalhes e os resultados do estudo foram organizados em um artigo que contempla a definição dos problemas de otimização associados aos algoritmos, a demonstração formal e ilustração gráfica e numérica do raciocínio pertinente ao papel dos parâmetros dos algoritmos, e a apresentação de experimentos computacionais para apoiar a validação do comportamento dos algoritmos em conjuntos de dados sintéticos, com características especialmente escolhidas para apoiar o entendimento do assunto.

A segunda produção constitui-se em um livro texto didático, projetado para suportar o aprendizado inicial na área de Mineração de Dados [2]⁶⁶. O livro trata principalmente de três tarefas de Mineração de Dados (predição, agrupamento e regras de associação), é escrito em uma linguagem de fácil leitura, apresentando um contexto de exemplificação único (análise de dados de um restaurante) que se desenvolve por todo o livro, motivando a prática dos algoritmos. Todos os conceitos são apresentados a partir de um embasamento teórico, seguido por exemplos didáticos e pela apresentação de implementação em R⁶⁷.

Agradecimentos

O trabalho de pesquisa da autora está fortemente atrelado ao trabalho de diversas outras pessoas, tanto pelo aspecto interdisciplinar de alguns dos problemas tratados, envolvendo pesquisadores da área de Linguística, Fonoaudiologia e Semiótica, quanto pela sua dinâmica de desenvolvimento de pesquisa colaborativa com colegas da área de Computação e Engenharia. Também colaboraram fortemente para os resultados aqui apresentados os vários alunos de graduação e pós-graduação orientados pela autora, em seus trabalhos de iniciação científica, trabalhos de conclusão de curso e dissertações de mestrado. A todas essas pessoas, pesquisadoras e empreendedoras, muitas das quais coautoras em artigos científicos, é destacado aqui o agradecimento da autora pela inestimável

⁶⁴www.each.usp.br/sarajane

⁶⁵Rocha, T.; **PERES, S. M.**; Bísvaro, H. H.; Madeo, R. C. B.; Boscaroli, C. Tutorial sobre Fuzzy-c-Means e Fuzzy Learning Vector Quantization: Abordagens Híbridas para Tarefas de Agrupamento e Classificação. *Revista de Informática Teórica e Aplicada: RITA*, v. 19, p. 120–163, 2016.

⁶⁶Silva, L. A.; **PERES, S. M.**; Boscaroli, C. *Introdução à Mineração de Dados: Com Aplicações em R*. São Paulo: Elsevier, 2016, 277p.

⁶⁷Ambiente de software livre que possibilita a programação genérica apoiada por facilidades referentes à implementação de funções estatísticas e pela disponibilização de ferramentas para visualização gráfica.

colaboração.

Um agradecimento também deve ser feito para os agentes financiadores de pesquisa que por meio do apoio financeiro para participação em eventos, bolsas de iniciação científica e mestrado, e fomento a projetos, possibilitaram o engrandecimento profissional da autora, suportaram o desenvolvimento da pesquisa e valorizaram os resultados obtidos. Tais agentes incluem: Fapesp, Capes, CNPq, MEC (Ministério da Educação via Programa de Educação Tutorial), e programas especiais da Universidade de São Paulo.

References

- [1] The Economist, “Data, data everywhere: a special report on managing information.” <http://www.economist.com/node/15557421>, 2010.
- [2] L. A. Silva, S. M. Peres, and C. Boscarioli, *Introdução à Mineração de Dados: Com Aplicações em R*. Elsevier, 2016.
- [3] A. Gandomi and M. Hainder, “Beyond the hype: Big data concepts, methods and analytics,” *International Journal of Information Management*, vol. 35, pp. 137–144, 2015.
- [4] R. Goldman, J. McHugh, and J. Windom, “From semistructured data to XML: Migrating the core data model and query language,” in *Proceedings of the 2nd International Workshop on the Web and Databases*, pp. 25–30, 1999.
- [5] M. E. Zorrila, E. Mora, and J. L. Crespo, “Non-structured data management by means of object relational database management systems,” *Systems Analysis Modelling Simulations*, vol. 43, no. 9, pp. 1173–1187, 2003.
- [6] T. Kavanagh, C. Beall, W. Heiny, J. Motycka, S. Pendleton, T. Smallwood, B. Terpening, and K. Traut, “Object oriented database management system,” Nov. 17 1998. US Patent 5,838,965.
- [7] M. Atkinson, D. DeWitt, D. Maier, F. Bancilhon, K. Dittrich, and S. Zdonik, “Building an object-oriented database system,” ch. The Object-oriented Database System Manifesto, pp. 1–20, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1992.
- [8] M. Stonebraker and D. Moore, *Object-Relational DBMSs: The Next Great Wave*. Morgan Kaufmann Publishers, 1996.
- [9] R. Cattell, “Scalable sql and nosql data stores,” *SIGMOD Rec.*, vol. 39, no. 4, pp. 12–27, 2011.

- [10] G. Piatetsky-Shapiro, “Knowledge discovery in real databases: A report on the ijcai-89 workshop,” *AI Magazine*, vol. 11, no. 5, pp. 68–70, 1991.
- [11] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, *Advances in Knowledge Discovery and Data Mining*. The MIT Press, 1996.
- [12] S. Haykin, *Neural Networks and Learning Machines*. Pearson, 3 ed., 2010.
- [13] T. Bäck, D. B. Fogel, and Z. Michalewicz, *Evolutionary Computation 1: Basic Algorithms and Operators*. United Kingdom: Institute of Physics Publishing, 2000.
- [14] T. Bäck, D. B. Fogel, and Z. Michalewicz, *Evolutionary Computation 2: Basic Algorithms and Operators*. United Kingdom: Institute of Physics Publishing, 2000.
- [15] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, 1995.
- [16] T. M. Mitchell, *Machine Learning*. McGraw Hill, 1997.
- [17] M. Turk, “Multimodal interaction: A review.,” *Pattern Recognition Letters*, vol. 36, pp. 189–195, 2014.
- [18] M. Pantic, A. Pentland, A. Nijholt, and T. Huang, “Human computing and machine understanding of human behavior: a survey,” in *Lecture Notes in Artificial Intelligence*, vol. 44512007, (Berlin), pp. 41–71, Springer-Verlag, 2007.
- [19] R. G. Souza, S. M. Peres, and C. A. M. Lima, “Adaptação do jogo da forca para língua de sinais usando luvas coloridas e redes neurais multilayer perceptron,” in *Anais do XIV Simpósio Brasileiro de Jogos e Entretenimento Digital - SBGames*, pp. 283–286, 2015.
- [20] F. A. Freitas, “Reconhecimento automático de expressões faciais gramaticais na língua brasileira de sinais,” Master’s thesis, Universidade de São Paulo, 2015.
- [21] A. Ittoo, L. M. Nguye, and A. van den Bosch, “Text analytics in industry: Challenges desiderata and trends,” *Computers in Industry*, vol. 78, pp. 96–107, 2016.
- [22] L. F. Brunialti, “Fatoração de matrizes no problema de coagrupamento com sobreposição de colunas,” Master’s thesis, Universidade de São Paulo, 2016.
- [23] W. M. P. van der Aalst, *Process Mining – Discovery, Conformance and Enhancement of Business Processes*. Berlin: Springer Publishing Company, Incorporated, 1st ed., 2011.

- [24] W. M. P. v. d. van der Aalst, *Process Mining: Data Science in Action*. Berlin: Springer, 2 ed., 2016.
- [25] S. Baraldi, A. D. Bimbo, and L. Landucci, “Natural interaction on tabletops,” *Multimedia Tools Applications*, vol. 38, pp. 385–405, 2008.
- [26] D. Kim, J. Song, and D. Kim, “Simultaneous gesture segmentation and recognition based on forward spotting accumulative hmms,” *Pattern Recognition*, vol. 40, no. 11, pp. 3012–3026, 2007.
- [27] L. Spano, A. Cisternino, and F. Paterno, “A compositional model for gesture definition,” in *Proceedings of International Conference on Human-Centered Software Engineering*, pp. 34–52, 2012.
- [28] R. C. B. Madeo, S. M. Peres, H. H. Biscaro, D. B. Dias, and C. Boscaroli, “A committee machine implementing the pattern recognition module for fingerspelling applications,” in *Proceedings of the 2010 ACM Symposium on Applied Computing, SAC '10*, (New York, NY, USA), pp. 954–958, ACM, 2010.
- [29] O. Koller, J. Foster, and H. Ney, “Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers,” *Computer Vision and Image Understanding*, vol. 141, pp. 108–125, Dec. 2015.
- [30] G. Lee, F. h. Yeh, and Y.-H. Hslao, “Kinect-based taiwanese sign-language recognition system,” *Multimedia Tools and Applications*, vol. 75, pp. 261–279, Jan. 2016.
- [31] R. A. Seger, M. M. Wanderley, and A. L. Koerich, “Automatic detection of musicians’ ancillary gestures based on video analysis,” *Expert System with Applications*, vol. 41, pp. 2098–2106, 2014.
- [32] A. E. Brown and Y. Sasson, “MAESTRO: Using technology to improve kinesthetic skill learning of music conductors,” in *Proceedings of the International Computer Music Conference*, (Ljubljana, Slovenia), pp. 333–336, Sept. 2012.
- [33] M. G. Jacob and J. P. Wachs, “Context-based hand gesture recognition for the operating room,” *Pattern Recognition Letters*, vol. 36, pp. 196–203, 2014.
- [34] A. Drosou, D. Ioannidis, K. Moustakas, and D. Tzovaras, “Spatiotemporal analysis of human activities for biometric authentication,” *Computer Vision and Image Understanding*, vol. 116, pp. 411–421, 2012.

- [35] P. Bremner, A. G. Pipe, M. Frase, S. Subramanian, and C. Melhuish, “Beat gesture generation rules for human-robot interaction,” in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, (Toyama, Japan), pp. 1029–1034, Sept. 2009.
- [36] M. Salem, I. W. S. Kopp, K. Rohlfing, and F. Joublin, “Generation and evaluation of communicative robot gesture,” *International Journal of Social Robotics*, vol. 4, no. 2, pp. 201–217, 2012.
- [37] F. S. Borba, *Dicionário UNESP do português contemporâneo*. São Paulo: UNESP, 2005.
- [38] A. Kendon, “An agenda for gesture studies,” *The Semiotic Review of Books*, vol. 7, no. 3, pp. 7–12, 1996.
- [39] D. McNeill, *Hand and mind: What the hands reveal about thought*. University of Chicago Press, 1992.
- [40] N. Frishberg, “Arbitrariness and iconicity: Historical change in american sign language,” *Language*, vol. 51, no. 3, pp. 696–719, 1975.
- [41] R. M. Quadros and L. B. Karnopp, *Língua de sinais brasileira: estudos lingüísticos*. Artmed, 2004.
- [42] W. C. J. Stokoe, “Sign language structure: An outline of the visual communication systems of the american deaf,” *Journal of Deaf Studies and Deaf Education*, vol. 10, no. 1, pp. 3–37, 2005. Originalmente publicado em 1965 em *Studies in Linguistics*, Occasional Paper 8.
- [43] A. Martinet, *Elementos de lingüística geral*. São Paulo: Marings Fontes, 1978.
- [44] T. d. A. Leite, *A segmentação da língua de sinais brasileira (libras): um estudo lingüístico descritivo a partir da conversação espontânea entre surdos*. PhD thesis, Universidade de São Paulo, 2008.
- [45] R. Battison, “Phonological deletion in american sign language,” *Sign Language Studies*, vol. 5, pp. 1–19, 1974.
- [46] D. Aarons, *Aspects of the syntax of american sign language*. PhD thesis, Boston University, 1994.
- [47] S. K. Liddell, *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press, 2003.

- [48] F. C. Capovilla and W. D. Raphael, *Enciclopédia da Língua de Sinais Brasileira, Volume 1: O Mundo do Surdo em Libras. Educação*. São Paulo: EDUSP, 2004.
- [49] H. Pistori and J. J. Neto, *An Experiment on Handshape Sign Recognition Using Adaptive Technology: Preliminary Results*, pp. 464–473. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.
- [50] S. M. Peres, F. C. Flores, D. Veronez, and C. J. M. Olguin, “LIBRAS signals recognition: a study with learning vector quantization and bit signature,” in *Proceeding of the Ninth Brazilian Symposium on Neural Networks- SBRN*, (Ribeirão Preto), pp. 119–124, out. 2006.
- [51] M. N. Neris, A. J. Silva, S. M. Peres, and F. C. Flores, “Self organizing maps and bit signature: A study applied on signal language recognition,” in *Proceedings of the International Joint Conference on Neural Networks - IJCNN, part of the IEEE World Congress on Computational Intelligence - WCCI*, (Hong Kong), pp. 2934–2941, IEEE, Jun. 2008.
- [52] F. C. Flores, S. M. Peres, and F. J. V. Zuben, “Automatic design of W-operators using LVQ - application to morphological image segmentation,” in *Proceeding of the International Joint Conference on Neural Networks*, vol. 1, (Hawaii, USA), pp. 755–760, IEEE, 2002.
- [53] F. Chung and T. Lee, “Fuzzy learning vector quantization,” in *Proceedings of the International Joint Conference on Neural Networks*, (Nagoya), pp. 2793–2742, IEEE, 1993.
- [54] J. C. Bezked, E. C. Tsao, and N. R. Pal, “Fuzzy kohonen clustering networks,” in *Proceedings of the International Conference on Fuzzy Systems*, (San Diego, USA), pp. 1035–1043, IEEE, 1992.
- [55] J. C. Bezked and S. K. Pal, *Fuzzy models for pattern recognition: methods that search for structures in data*. New York: IEEE, 1992.
- [56] D. B. Dias, R. C. B. Madeo, T. Rocha, H. H. BÍscaro, and S. M. Peres, “Hand movement recognition for brazilian sign language: A study using distance-based neural networks,” in *International Joint Conference on Neural Networks - IJCNN*, (Atlanta), pp. 697–704, IEEE, jun. 2009.
- [57] T. Kohonen, *Self Organizing Maps*. Springer, 3 ed., 2000.

- [58] R. C. B. Madeo, S. M. Peres, C. A. M. Lima, and C. Boscarioli, “Hybrid architecture for gesture recognition: Integrating fuzzy-connectionist and heuristic classifiers using fuzzy syntactical strategy,” in *Proceedings of International Joint Conference on Neural Networks – IJCNN*, (Brisbane), pp. 1–8, IEEE, Jun. 2012.
- [59] M. V. Lamar, *Hand Gesture Recognition using TCombNET – A Neural Network Model dedicated to Temporal Information Processing*. PhD thesis, Nagoya Institute of Technology, 2001.
- [60] R. Y. Wang and J. Popović, “Real-time hand-tracking with a color glove,” *ACM Transactions on Graphics*, vol. 28, pp. 63:1–63:8, jul. 2009.
- [61] P. Mistry, P. Maes, and L. Chang, “WUW – wear Ur world: A wearable gestural interface,” in *Proceedings of Extended Abstracts on Human Factors in Computing Systems*, (New York), pp. 411–4116, ACM, 2009.
- [62] D. B. Dias, S. M. Peres, and H. H. Bísaro, “Libras movement data set,” 2009.
- [63] M. Lichman, “UCI machine learning repository,” 2013.
- [64] P. Ekman, “Facial expression,” *Nonverbal Behavior and Communication*, pp. 97–116, 1977.
- [65] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Los Altos, California: Malor Books, 2003.
- [66] C. Corneanu, M. Simón, J. Cohn, and S. Guerrero, “Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 8, pp. 1548–1568, 2016.
- [67] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, “Face recognition: A literature survey,” *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.
- [68] A. Mehrabian, “Communication theory,” ch. Communication Without Words, pp. 193–200, Transaction Publishers, 2008.
- [69] T. D. Nguyen and S. Ranganath, “Facial expressions in american sign language: Tracking and recognition,” *Pattern Recognition*, vol. 45, pp. 1877–1891, 2012.
- [70] U. von Agris, M. Knorr, and K. F. Kraiss, “The significance of facial features for automatic sign language recognition,” in *Proceedings of the 8th IEEE International Conference on Automatic Face Gesture Recognition*, (Amsterdam, The Netherlands), pp. 1–6, IEEE, Sep 2008.

- [71] L. Ferreira-Brito, *Por uma gramática de línguas de sinais*. Rio de Janeiro, RJ: Tempo Brasileiro, 1995.
- [72] J. Arroteia, *O papel da marcação não-manual nas sentenças negativas em Línguas de Sinais Brasileira (LSB)*. PhD thesis, Universidade de Estadual de Campinas, 2005.
- [73] I. Ari and L. Akarun, “Facial feature tracking and expression recognition for sign language,” in *Proceedings of the 17th Signal Processing and Communications Applications Conference*, (Antalya, Turkey), pp. 229–232, IEEE, April 2009.
- [74] T. D. Nguyen and S. Ranganath, “Towards recognition of facial expressions in sign language: Tracking facial features under occlusion,” in *15th IEEE International Conference on Image Processing*, pp. 3228–3231, Oct 2008.
- [75] N. Michael, D. Metaxas, and C. Neidle, “Spatial and temporal pyramids for grammatical expression recognition of american sign language,” in *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, (New York, NY, USA), pp. 75–82, ACM, 2009.
- [76] O. Lostakis, P. Papapetrou, and J. Hollmén, “Distance measure for querying sequences of temporal intervals,” in *Proceedings of the 4th International Conference on Pervasive Technologies Related to Assistive Environments*, pp. 40:1–40:8, ACM, 2011.
- [77] M. Uddin, “An ada-random forests based grammatical facial expressions recognition approach,” in *2015 4th International Conference on Informatics, Electronics and Vision*, 2015.
- [78] M. Bhuvan, D. Rao, S. Jain, T. Ashwin, R. Guddetti, and S. Kulgod, “Detection and analysis model for grammatical facial expressions in sign language,” pp. 155–160, 2016.
- [79] P. Campr, M. Hruz, and J. Trojanová, “Collection and preprocessing of czech sign language corpus for sign language recognition,” in *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, (Marrakech-Morocco), 2008.
- [80] Z. Krnoul, M. Hruz, and P. Campr, “Correlation analysis of facial features and sign gestures,” in *IEEE 10th International Conference on Signal Processing*, pp. 732–735, IEEE, 2010.

- [81] O. Aran, I. Ari, A. Guvensan, H. Haberdar, Z. Kurt, I. Turkmen, A. Uyar, and L. Akarun, “A database of non-manual signs in turkish sign language,” in *Proceedings of the 15th Signal Processing and Communications Applications*, (USA), pp. 1–4, IEEE, Jun 2007.
- [82] C. Neidle, “SignstreamTM: A database tool for research on visual-gestural language,” Tech. Rep. 10, Boston University, aug. 2000.
- [83] F. A. Freitas, S. M. Peres, C. A. M. Lima, and F. V. Barbosa, “Grammatical facial expression recognition with machine learning,” in *27th Florida Artificial Intelligence Research Society Conference (FLAIRS)*, pp. 180–185, Palo Alto: The AAAI Press, 2014.
- [84] F. A. Freitas, F. V. Barbosa, and S. M. Peres, “Grammatical facial expressions data set,” 2014.
- [85] “ISGS: International society for gesture studies.” <http://www.gesturestudies.com>, 2013. Online; acessado em 7 de janeiro de 2017.
- [86] A. Kendon, “Gesticulation and speech: Two aspects of the process of utterance,” in *The Relationship of verbal and nonverbal communication* (M. R. Key, ed.), pp. 207–227, Mouton Publishers, 1980.
- [87] D. McNeill, *Gesture and Thought*. University of Chicago Press, 2005.
- [88] S. Kita, I. van Gijn, and H. van der Hulst, “Movement phases in signs and co-speech gestures, and their transcription by human coders,” in *Gesture and Sign Language in Human-Computer Interaction* (I. Wachsmuth and M. Frohlich, eds.), vol. 1371 of *Lecture Notes in Computer Science*, pp. 23–35, Springer Berlin / Heidelberg, 1998.
- [89] R. C. B. Madeo, C. A. M. Lima, and S. M. Peres, “Studies in automated hand gesture analysis: an overview of functional types and gesture phases,” *Language Resources and Evaluation*, pp. 1–33, 2016. Online First.
- [90] F. Maricchiolo, A. Gnisci, and M. Bonaiuto, “Coding hand gestures: a reliable taxonomy and a multi-media support,” in *Cognitive Behavioural Systems*, pp. 405–416, Springer, 2012.
- [91] M. Kipp, *Multimedia Annotation, Querying, and Analysis in Anvil*, pp. 351–367. John Wiley & Sons, Inc., 2012.

- [92] H. Brugman and A. Russel, “Annotating multimedia/multi-modal resources with elan,” in *Proceedings of the 4th International Conference on Language Resources and Evaluation*, (Lisbon, Portugal), pp. 2065–2068, 2004.
- [93] F. Quek, D. McNeill, R. Bryll, S. Duncan, X. Ma, C. Kirbas, K. McCullough, and R. Ansari, “Multimodal human discourse: gesture and speech,” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 9, no. 3, pp. 171–193, 2002.
- [94] M. Gullberg, “Some reasons for studying gesture and second language acquisition (Hommage à. Adam Kendon,” *International Review of Applied Linguistics in Language Teaching*, vol. 44, pp. 103–124, 2006.
- [95] Y.-H. Hsieh, S. C. Hidayati, W.-H. Chen, M.-C. Hu, and K.-L. Hua, “Who’s the best charades player? mining iconic movement of semantic concepts,” in *Proceedings of the 20th Anniversary International Conference on MultiMedia Modeling*, (New York), pp. 231–241, Springer, 2014.
- [96] S. Sathyanarayana, R. K. Satzoda, A. Carini, M. Lee, L. Salamanca, J. Reilly, D. Foster, M. Bartlett, and G. Littlewort, “Towards automated understanding of student-tutor interactions using visual deictic gestures,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition - Workshop*, pp. 480–487, IEEE, 2014.
- [97] M. Salem, K. Rohlfing, S. Kopp, and F. Jaoubli, “A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction,” in *Proceedings of 20th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 247–252, IEEE, 2011.
- [98] R. C. B. Madeo, “Máquinas de vetores suporte e a análise de gestos: incorporando aspectos temporais,” Master’s thesis, Universidade de São Paulo, 2013.
- [99] R. C. B. Madeo, C. A. M. Lima, and S. M. Peres, “A review on temporal reasoning using support vector machines,” in *Proceedings of 19th International Symposium on Temporal Representation and Reasoning*, (Washington, DC, USA), pp. 46–52, IEEE Computer Society, sep. 2012.
- [100] M. Hermans and B. Schrauwen, “Recurrent kernel machines: Computing with infinite echo state networks,” *Neural Computation*, vol. 24, pp. 104–133, 2012.
- [101] Z. Shi and M. Han, “Support vector echo-state machine for chaotic time-series prediction,” *IEEE Transaction on Neural Networks*, vol. 18, no. 2, pp. 359–372, 2007.

- [102] R. C. B. Madeo, S. M. Peres, and C. A. M. Lima, “Overview on support vector machines applied to temporal modeling,” in *Anais do IX Encontro Nacional de Inteligência Artificial*, pp. 1–6, oct. 2012.
- [103] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, UK: Cambridge University Press, 2nd ed., 2004.
- [104] R. C. B. Madeo, S. M. Peres, and C. A. M. Lima, “Gesture phase segmentation using Support Vector Machines,” *Expert Systems with Applications*, vol. 56, pp. 100 – 115, 2016.
- [105] P. K. Wagner, R. C. B. Madeo, S. M. Peres, and C. A. M. Lima, “Segmentação de unidades gestuais com Multilayer Perceptron,” in *Anais do X Encontro Nacional de Inteligência Artificial*, (Fortaleza, Ceará), 2013.
- [106] R. C. B. Madeo, P. K. Wagner, and S. M. Peres, “Gesture phase segmentation data set,” 2014.
- [107] C. A. Montgomery, “Linguistic an information science,” *Journal of the Association for Information Science and Tecnhnology*, vol. 23, pp. 195–219, 1972.
- [108] K. S. Jones and M. Kay, “Linguistic and information science: A postscript,” in *Natural Language in Information Retrieval* (D. E. Walker, H. Karlgren, and M. Kay, eds.), pp. 183–192, 1977.
- [109] O. Kolomiyets and M.-F. Moens, “A survey on question answering technology from an information retrieval perspective,” *Information Sciences*, vol. 181, no. 24, pp. 5412 – 5434, 2011.
- [110] N. Y. Saiyad, H. Prajapati, and V. Dabhi, “A survey of document clustering using semantic approach,” in *Proceedings of International Conference on Electrical, Electronics, and Optimization Techniques*, pp. 2555–2564, 2016.
- [111] C. J. V. Rijsbergen, *Information Retrieval*. London: Butterworths, 2nd ed., 1979.
- [112] N. P. Luhn, “The automatic creation of literature abstracts,” *IBM Journal of Research and Development*, vol. 2, pp. 159–165, 1958.
- [113] R. Jindal, R. Malhotra, and A. Jain, “Techniques for text classification: Literature review and current trends,” *Webology*, vol. 12, 2015.
- [114] G. Salton, A. Wong, and C. S. Yang, “A vector space model for automatic indexing,” *Communications of the ACM*, vol. 18, pp. 613–620, Nov. 1975.

- [115] F. Sebastiani, “Machine learning in automated text categorization,” *ACM Computing Surveys*, vol. 34, pp. 1–47, Mar. 2002.
- [116] M. Porter, “An algorithm for suffix stripping,” *Program*, pp. 130–137, 1980.
- [117] T. Müller, R. Cotterell, A. M. Fraser, and H. Schütze, “Joint lemmatization and morphological tagging with lemming,” in *Proceedings of Conference on Empirical Methods on Natural Language Processing - EMNLP* (L. Màrquez, C. Callison-Burch, J. Su, D. Pighin, and Y. Marton, eds.), pp. 2268–2274, The Association for Computational Linguistics, 2015.
- [118] C. C. Aggarwal and C. Zhai, “A survey of text clustering algorithms,” in *Mining Text Data* (C. C. Aggarwal and C. Zhai, eds.), pp. 77–128, Berlin: Springer, 2012.
- [119] J. C. Bezdek and J. C. Dunn, “Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions,” *IEEE Transactions on Computers*, vol. C-24, pp. 835–838, 1975.
- [120] J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A K-means clustering algorithm,” *Applied Statistics*, pp. 100–108, 1979.
- [121] Z. Huang, “Extensions to the K-means algorithm for clustering large data sets with categorical values,” *Data Mining and Knowledge Discovery*, vol. 2, pp. 283–304, 1998.
- [122] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. 1990: John Wiley & Sons, 1990.
- [123] B. B. C. Silva, S. M. Peres, B. C. Letti, L. F. Brunialti, M. Fantinato, and C. A. M. Lima, “Análise de agrupamento aplicada a textos para construção de perfis de trabalhos de conclusão de curso,” in *Anais do 12o. Congresso Brasileiro de Inteligência Computacional - CBIC 2015*, (Curitiba-PR), pp. 1–6, 2015.
- [124] D. D. Lee and H. S. Seung, “Learning the parts of objects by nonnegative matrix factorization,” *Nature*, pp. 788–791, 1999.
- [125] S. Yoo, J.; Choi, “Orthogonal nonnegative matrix tri-factorizations for co-clustering: multiplicative updates on stiefel manifolds,” *Information Processing and Management*, vol. 46, pp. 559–570, 2010.
- [126] Y. Yan, C. L., and W.-C. Tjhi, “Fuzzy semi-supervised co-clustering for text documents,” *Fuzzy Sets and Systems*, vol. 215, pp. 74–89, 2013.

- [127] C. Laclau and M. Nadif, “Hard and fuzzy diagonal co-clustering for document-term partitioning,” *Neurocomputing*, vol. 193, pp. 133–147, 2016.
- [128] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Advances in Neural Information Processing Systems 13* (T. K. Leen, T. G. Dietterich, and V. Tresp, eds.), pp. 556–562, MIT Press, 2001.
- [129] W.-C. Tjhi and L. Chen, “A heuristic-based fuzzy co-clustering algorithm for categorization of high-dimensional data,” *Fuzzy Sets and Systems*, vol. 159, no. 4, pp. 371 – 389, 2008.
- [130] M. Khoshneshin, M. Ghazizadeh, W. N. Street, and J. W. Ohlmann, “A memetic heuristic for the co-clustering problem,” in *Proceedings of the 8th International Conference on Bioinspired Information and Communications Technologies*, (ICST, Brussels, Belgium, Belgium), pp. 400–403, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014.
- [131] S. C. Madeira and A. L. Oliveira, “Biclustering algorithms for biological data analysis: A survey,” *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 1, pp. 24–45, jan. 2004.
- [132] F. O. França, “A hash-based co-clustering algorithm for categorical data,” *Expert System wirh Applications*, vol. 64, pp. 24–35, 2016.
- [133] G. Getz, E. Levine, and E. Domany, “Coupled two-way clustering analysis of gene microarray data,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 22, pp. 12079–12084, 2000.
- [134] Y. Cheng and G. M. Church, “Biclustering of expression data,” in *Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology*, (California), pp. 93–103, 2000.
- [135] J. A. Hartigan, “Direct clustering of a data matrix,” *Journal of the American Statistical Association*, vol. 67, no. 337, pp. 123–129, 1972.
- [136] P. A. D. de Castro, F. O. de França, H. M. Coelho, and F. J. Von Zuben, “Query expansion using an immune-inspired biclustering algorithm,” *Natural Computing*, vol. 9, no. 3, pp. 579–602, 2000.
- [137] S. Mitra and H. Banka, “Multi-objective evolutionary biclustering of gene expression data,” *Pattern Recognition*, vol. 39, pp. 2464–2477, 2006.

- [138] B. Long, Z. Zhang, and P. S. Yu, “Co-clustering by block value decomposition,” in *Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, (New York, NY, USA), pp. 635–640, ACM, 2005.
- [139] C. Ding, T. Li, W. Peng, and H. Park, “Orthogonal nonnegative matrix tri-factorizations for clustering,” in *Proceedings of ACM SIGKDD Conferences on Knowledge Discovery and Data Mining*, (Philadelphia, USS), pp. 126–135, ACM, 2006.
- [140] C. Ding, X. He, and H. D. Simon, “On the equivalence of nonnegative matrix factorization and spectral clustering,” in *Proceedings of SIAM International Conference on Data Mining*, (Newport Beach, California), pp. 606–610, 2005.
- [141] H. Wang, F. Nie, H. Huang, and F. Makedon, “Fast nonnegative matrix tri-factorization for large-scale data co-clustering,” in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 2, (Barcelona, Spain), pp. 1553–1558, AAAI Press, 2011.
- [142] W. D. Rand, “Objective criteria for the evaluation of clustering methods,” *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 846–850, 1971.
- [143] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, *Recommender Systems Handbook*. Berlin: Springer, 2011.
- [144] L. F. Brunialti, S. M. Peres, V. Freire, and C. A. M. Lima, “Aprendizado de máquina em sistemas de recomendação baseados em conteúdo textual: Uma revisão sistemática,” in *Anais do XI Simpósio Brasileiro de Sistemas de Informação*, (Goiânia, GO), pp. 203–210, 2015.
- [145] M. Kaminskis and D. Bridge, “Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-accuracy objectives in recommender systems,” *ACM Transactions on Interactive Intelligent Systems*, vol. 7, no. 1, pp. 2:1–2:42, 2016. Issue-in-progress.
- [146] M. Weske, *Business Process Management: Concepts, Languages, Architectures*. Berlin: Springer, 2nd ed., 2012.
- [147] M. Dumas, M. Rosa, J. Mendling, and H. Reijers, *Fundamentals of Business Process Management*. Berlin: Springer, 2013.
- [148] A. Lindsay, D. Downs, and K. Lunn, “Business processes-attempts to find a definition,” *Information and Software Technology*, vol. 45, pp. 1015–1019, 2003.

- [149] R. Agrawal, D. Gunopulos, and F. Leymann, “Mining process models from workflow logs,” in *Proceedings of the 6th International Conference on Extending Database Technology* (H.-J. Schek, G. Alonso, F. Saltor, and I. Ramos, eds.), (Berlin), pp. 467–483, Springer, March 1998.
- [150] A. R. C. Maita, “Um estudo da aplicação de técnicas de inteligência computacional e de aprendizado de máquina em mineração de processos de negócio,” Master’s thesis, Universidade de São Paulo, 2016.
- [151] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, “Top 10 algorithms in data mining,” *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.
- [152] A. R. C. Maita, L. C. Martins, C. R. L. Paz, S. M. Peres, and M. Fantinato, “Process mining through artificial neural networks and support vector machines: A systematic literature review,” *Business Process Management Journal*, pp. 1391–1415, 2016.
- [153] W. P. van der Aalst, T. Weijters, and L. Maruster, “Workflow mining: Discovering process models from event logs,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, pp. 1128–1142, 2004.
- [154] A. R. C. Maita, M. Fantinato, S. M. Peres, L. H. Thom, and P. C. K. Hung, “Mining unstructured processes: An exploratory study on a distance learning domain,” in *Proceedings of 30th International Joint Conference on Neural Networks*, (Anchorage, AK, USA), pp. 1–8, IEEE, 2017. Accepted to.
- [155] M. P. Soledade Jr, R. S. Freitas, S. M. Peres, M. Fantinato, R. Steinbeck, and U. Araújo, “Experimenting with design thinking in requirements for a learning management system,” in *Proceeding of Simpósio Brasileiro de Sistemas de Informação*, (João Pessoa), pp. 182–193, 2013.
- [156] D. Yue, X. Wu, H. Wang, and J. Bai, “A review of process mining algorithms,” in *Proceedings of International Conference on Business Management and Electronic Information*, vol. 5, (Guangzhou, China), pp. 181–185, IEEE, May 2011.
- [157] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.
- [158] A. L. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” *Artificial Intelligence*, vol. 97, pp. 245–271, 1997.

- [159] R. Kohavi and G. H. John, “Wrappers for feature subset selection,” *Artificial Intelligence*, vol. 97, pp. 273–324, 1997.
- [160] V. N. Flores, L. B. Barbisan, M. J. B. Finatto, and M. Teixeira, *Dicionário de Linguística da Enunciação*. São Paulo, SP: Editora Contexto, 2009.
- [161] J.-J. Franckel, D. Paillard, and E. Sunier, *Langages*, ch. Aspects de la théorie d’Antoine Culioli, pp. 52–63. Larousse, 1998.
- [162] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in *Proceedings of the 26th International Conference on Neural Information Processing Systems*, (USA), pp. 3111–3119, Curran Associates Inc., 2013.
- [163] C. A. M. Lima, A. L. V. Coelho, R. C. B. Madeo, and S. M. Peres, “Classification of electromyography signals using relevance vector machines and fractal dimension,” *Neural Computing & Applications (Print)*, vol. 27, pp. 791–804, 2016.
- [164] F. G. S. Teodoro, D. M. M. da Costa, S. M. PERES, and C. A. M. Lima, “Ecg biometric recognition based on fiducial features using support vector machines,” in *X Workshop de Visão Computacional - WVC*, (Uberlândia, MG), pp. 327–332, 2014.
- [165] C. A. M. Lima, R. C. B. Madeo, S. M. Peres, and M. Eisenkraft, *An Investigation on Support Vector Machines and Wavelet Transform in Eletroencephalogram Signal Classification*, pp. 95–113. CRC Press, 2013.
- [166] F. F. S. Teodoro, S. M. Peres, and C. A. M. Lima, “Feature selection for biometric recognition based on electrocardiogram signals,” in *International Joint Conference on Neural Networks*, (Alaska, USA), IEEE, 2017. Accepted to.
- [167] F. G. S. Teodoro, D. M. M. da Costa, S. M. PERES, and C. A. M. Lima, “Seleção de características via algoritmos evolucionários para reconhecimento biométrico baseado em sinais de ecg,” in *12o Congresso Brasileiro de Inteligência Computacional*, (Curitiba, PR), pp. 1–6, 2015.
- [168] B. M. Passos, H. ; Duru, E. L. Oliveira, S. M. Peres, and C. A. M. Lima, “Symbolic representations of time series applied to biometric recognition based on ecg signals,” in *International Joint Conference on Neural Networks*, (Alaska, USA), IEEE, 2017. Accepted to.

- [169] H. Passos, D. M. M. da Costa, S. M. Peres, and C. A. M. Lima, “Patterns detection in ecg signal applied to biometric recognition,” in *XI Workshop de Visão Computacional - WVC*, (São Carlos, SP), 2015.
- [170] D. M. M. da Costa, S. M. Passos, H. ; Peres, and C. A. M. Lima, “Iris recognition using support vector machine and least squares support vector,” in *Workshop de Visão Computacional*, (São Carlos, SP), 2015.
- [171] M. U. Schineider, R. C. B. Madeo, S. M. Peres, and C. A. M. Lima, “Committe machine for iris segmentation,” in *Encontro Nacional de Inteligência Artificial*, 2012.
- [172] D. M. M. da Costa, S. M. Peres, C. A. M. Lima, and P. Mustaro, “Face recognition using support vector machine and multiscale directional image representation methods: A comparative study,” in *International Joint Conference on Neural Networks*, (Killarney, Ireland), pp. 1–8, IEEE, 2015.
- [173] D. M. M. da Costa, F. G. S. Teodoro, S. M. Peres, and C. A. M. Lima, “Face recognition systems based on wavelet transform and svm ensembles,” in *Workshop de Visão Computacional - WVC*, (Uberlândia, MG), pp. 321–326, 2014.
- [174] D. M. M. da Costa, H. Passos, S. M. Peres, and C. A. M. Lima, “Um estudo comparativo das estratégias de fusão no nível de características para sistemas biométricos multimodais baseados em face e íris,” in *Simpósio Brasileiro de Sistemas de Informação*, (Goiania, GO), pp. 219–226, 2015.
- [175] D. M. M. da Costa, S. M. Peres, and C. A. M. Lima, “Floresta de caminhos Ótimos aplicado ao reconhecimento biométrico multimodal,” in *Congresso Brasileiro de Inteligência Computacional*, (Curitiba, PR), pp. 1–6, 2015.
- [176] E. L. Oliveira, C. A. M. Lima, and S. M. Peres, “Fusão de face e forma de andar para reconhecimento biométrico: Revisão sistemática da literatura,” in *Simpósio Brasileiro de Sistemas de Informação*, pp. 108–113, 2016.
- [177] T. Rocha, S. M. Peres, H. H. Bísaro, R. C. B. Madeo, and C. Boscaroli, “Tutorial sobre fuzzy-c-means e fuzzy learning vector quantization: Abordagens híbridas para tarefas de agrupamento e classificação,” *Revista de Informática Teórica e Aplicada: RITA*, vol. 19, pp. 120–163, 2016.