

Análise de redes sociais para detecção de transtornos de saúde mental

Ivandr  Paraboni

School of Arts, Sciences and Humanities, University of S o Paulo (USP / EACH)
Av. Arlindo Bettio, 1000 - S o Paulo, Brazil
ivandre@usp.br

1 Introdu o

Este relat rio descreve os resultados do projeto ‘*An lise de linguagem em redes sociais para detec o precoce de transtornos de sa de mental*’. Para este fim,   apresentado um breve resumo da proposta (se o 2), o trabalho realizado (se o 3), os resultados obtidos at  o presente momento (se o 4), e considera es sobre a continuidade da pesquisa realizada (se o 5).

2 Resumo do projeto

O projeto objetivou o estudo e desenvolvimento de recursos computacionais b sicos e modelos de detec o precoce de transtornos de sa de mental sob uma perspectiva de PLN baseada em textos e outras informa es provenientes da rede social Twitter em portugu s, com contribui es te ricas e pr ticas possivelmente in ditas em tr s frentes principais, a saber:

1. Cria o de um recurso lingu stico-computacional de larga escala dedicado ao portugu s, sem similar conhecido, cobrindo tanto o problema de detec o de depress o (que   o mais frequentemente estudado na  rea) como tamb m de transtorno de ansiedade, e incluindo dados textuais e n o-textuais (e.g., relativos aos contatos e intera es na rede social).
2. Reformula o do problema computacional como uma tarefa de classifica o restrita aos dados cronologicamente anteriores ao diagn stico, tornando a aplica o potencialmente mais  til na preven o dos transtornos investigados.
3. Modelos de detec o precoce de depress o/ansiedade nos moldes definidos.

O projeto seguiu uma metodologia tradicional da  rea de PLN de estudo do problema, constru o de conjunto de dados, treinamento e teste de modelos de aprendizado de m quina supervisionado. Em linhas gerais, o projeto pode ser dividido em duas grandes fases, cada uma correspondendo aproximadamente a um ano de pesquisa: a constru o e valida o do recurso lingu stico-computacional b sico (i.e., o c rpus no dom nio do Twitter com dados de usu rios diagnosticados com depress o/ansiedade e um grupo de controle), e o uso da base de dados

coletada para estudo e desenvolvimento de modelos computacionais de detecção destes fenômenos a partir de dados textuais, comportamentais, demográficos e outros.

3 Trabalho realizado

O relato a seguir resume as atividades realizadas nos dois anos de projeto. Em linhas gerais, o primeiro ano do projeto concentrou-se predominantemente na atividade de construção da base de dados, denominada *córpus SetembroBR*, desenvolvimento de sistemas de *baseline* e resultados iniciais. No segundo ano, os modelos textuais desenvolvidos foram aprimorados com o uso de grandes modelos de língua (LLMs) e engenharia de *prompts*, e foram investigados modelos ditos estendidos (i.e., que utilizam dados não textuais como relações de rede social, de forma isolada e em combinação com modelos textuais). Estas atividades são discutidas individualmente a seguir.

3.1 Construção do *córpus SetembroBR*

O problema computacional de detecção de transtornos de saúde mental pode ser visto como um caso específico da tarefa de caracterização autoral [1–3], fazendo igualmente uso de métodos supervisionados com base em um *córpus* previamente rotulado. Assim, a primeira etapa do projeto tratou justamente da construção de um recurso deste tipo.

A construção do *córpus SetembroBR* (atividades 1,2 do cronograma) consistiu da coleta e disponibilização de dados para reuso em pesquisa segundo as diretrizes do plano de gestão de dados do projeto. Este processo, iniciado em [4], é descrito em detalhes em [5]. Uma discussão mais ampla, apresentada em formato adaptado à comunidade de humanidades digitais, foi também apresentada em um resumo publicado em [6].

O *córpus SetembroBR* é uma coleção multimodal de 47 milhões de *tweets* (totalizando 555 milhões de palavras em Português) publicados por mais de 31 mil usuários, e de suas relações de rede social (grafos de amigos, seguidores e menções). O *córpus* contempla dois tipos de *timelines*: conjuntos de *tweets* e informações de rede social públicos de indivíduos que relataram um diagnóstico de depressão/ansiedade, e um conjunto de controle formado por *timelines* de usuários aleatórios. Esses dois conjuntos, denominados *Diagnosticados* e *Controle*, correspondem às classes positiva e negativa do problema de classificação, e são pareados por quantidade de publicações, gênero (masculino/feminino) e datas aproximadas de publicação. Além disso, seguindo-se práticas estabelecidas na área (e.g., [7]), a quantidade de usuários do conjunto *Controle* no *córpus* é sete vezes a quantidade de usuários do conjunto *Diagnosticados*.

No que diz respeito aos usuários *Diagnosticados*, o *córpus* contempla apenas a porção de texto dita ‘útil’ (para fins de classificação) dos dados, ou seja, limitada aos *tweets* anteriores ao momento do diagnóstico feito (segundo o autorrelato publicado) por um especialista da área de saúde. Estes autorrelatos dividem-se

em dois tipos — diagnóstico de depressão e de transtorno de ansiedade — a serem tratados como tarefas computacionais independentes. A Tabela 1 apresenta estatísticas descritivas dos dois conjuntos de dados construídos, adaptada de [5].

Table 1. Estatísticas descritivas do córpus SetembroBR [5].

Estatística	Depressão	Controle	Total	C/D	Ansiedade	Controle	Total	C/A
Todos usuários	1684	11788	13472	7.0	2219	15533	17752	7.0
Sexo feminino	76.7%	76.7%	76.7%	7.0	78.8%	78.8%	78.8%	7.0
Tweets (milhões)	2.43	16.99	19.42	7.0	3.43	23.98	27.41	7.0
Palavras (milhões)	29.32	201.94	231.26	6.9	42.24	281.51	323.75	6.7
Tweets/usuários	1441	1441	1441	1.0	1543	1543	1543	1.0
Palavras/tweets	12.08	11.88	11.91	1.0	12.33	11.74	11.81	1.0
Dias média	492	553	546	1.1	524	540	538	1.0
Dias máximo	3712	4165	4165	1.1	4088	4211	4211	1.0
Amigos média	659	710	704	1.1	678	729	722	1.1
Seguidores média	777	945	924	1.2	810	975	954	1.2
Menções média	125	122	122	1.0	115	114	114	1.0

A Tabela 1 apresenta dados quantitativos de natureza textual (porção superior), temporal (porção central) e não-textual (porção inferior) para dados de indivíduos diagnosticados (D), de controle (C) e suas proporções (C/D e C/A). Os dados textuais representam quantidades de *timelines*, *tweets* e palavras de cada subcórpus (depressão e ansiedade). Os dados temporais representam o intervalo médio e máximo entre o primeiro *tweet* disponível no córpus e o momento em que o diagnóstico de um indivíduo é autorrelatado, servindo assim como um indicativo do potencial de antecipação destes transtornos de saúde mental por meio da análise do histórico de publicações na rede social. Finalmente, os dados não-textuais constituem as redes de amigos, seguidores e menções (ou interações) de cada indivíduo cuja *timeline* está presente no córpus.

O córpus construído é um dos maiores já disponibilizados para este tipo de pesquisa, possivelmente inferior apenas ao córpus em [8] para a língua inglesa (e que, ao contrário do córpus SetembroBR, é puramente textual). O córpus é também um dos poucos a incluir grafos de relações de redes sociais, e possivelmente o único recurso do tipo existente para a língua portuguesa até o presente momento. Além disso, é um dos poucos a modelar não apenas casos de depressão, mas também de transtorno de ansiedade, e a tratar exclusivamente de publicações anteriores ao momento do diagnóstico, enquanto a maioria dos recursos deste tipo geralmente não consideram esta questão. Todos estes pontos são discutidos em detalhes em [5].

3.2 Sistemas de baseline e validação inicial

A publicação do córpus em [5] incluiu a descrição de uma série de sistemas de *baseline* (atividades 3.4) construídos com uso de redes do tipo BiLSTM e do modelo de língua BERT e outros existentes na literatura da área para predição

de depressão e transtorno de ansiedade a partir de dados textuais e não textuais. Estes resultados iniciais foram então tomados por referência para refinamentos e melhorias nas etapas subsequentes do projeto.

Os métodos em questão fazem uso de um modelo BERT treinado especificamente para o domínio do Twitter em português, denominado BERTabaporu, detalhado em [9]. Este modelo foi treinado em uma massa de 238 mil *tweets*, totalizando cerca de 2,9 bilhões de *tokens*, e que é um subproduto da etapa de coleta de dados do cópulo SetembroBR. Os resultados deste modelo foram comparados aos obtidos pelo tradicional BERTimbau [10], treinado em cerca de 2,7 bilhões de *tokens*, para as tarefas de predição de transtornos de saúde mental e outras em [9]. O modelo BERTabaporu propriamente dito está disponível para reuso no repositório *Hugging Face*¹.

Finalmente, embora a ênfase da pesquisa tenha sido o uso de métodos orientados a dados baseados em aprendizado profundo, foi constatada a necessidade de comparar estes métodos com uma abordagem mais tradicional, baseada em engenharia de características, como forma de justificar as presentes decisões de projeto. Assim, foi conduzida uma avaliação de uma estratégia fortemente baseada em indicadores linguísticos de depressão, dicionários de sentimentos, sintomas e outros, e sua comparação com modelos baseados em BERT e aprendizado profundo. Estas características são relacionadas a seguir, e detalhes são discutidos em [11].

1. *DA_Mentions*: proporção de tweets contendo os termos ‘depressão’ ou ‘ansiedade’.
2. *Meds_01*: proporção de tweets contendo termos de natureza médica em geral.
3. *Symptoms* [1..9]: proporção de palavras associadas a sintomas de depressão.
4. *LIWC* [1..64]: proporção de palavras associadas a diferentes categorias LIWC [12].
5. *Abs*: proporção de termos absolutos no texto [13].
6. *Night_01*: proporção de tweets publicados depois das 21h e antes das 6h da manhã, com o intuito de capturar eventual atividade que possa estar relacionada a distúrbios de sono [14].
7. *1p Verbs*: proporção de pronomes pessoais de primeira pessoa, frequentemente associados à depressão [15], obtidos por *PoS tagging*.
8. *MeMimComigo*: proporção de palavras que correspondem a pronomes me/mim/comigo.
9. *Gender*: o gênero linguístico (0=masculino, 1=feminino) mais frequentemente adotado pelo indivíduo ao falar de si mesmo.

Como alternativa à engenharia de características, modelos BERT do tipo introduzido em [5] foram combinados em uma arquitetura de comitê de máquina dinâmico de mistura de especialistas baseada em [16]. Esta arquitetura é constituída de três modelos especialistas com uma rede *gating* ponderadora, em que cada modelo especialista, assim como sua rede *gating*, é constituído de um modelo de rede recorrente do tipo LSTM que recebe como entrada uma representação

¹ <https://huggingface.co/pablocosta/bertabaporu-large-uncased>

em nível de usuário construída pelo modelo de representação distribuída BERT [17]. Os resultados obtidos, que confirmam a ampla superioridade dos modelos de aprendizado profundo baseados em mistura de especialistas BERT, é apresentado em [11].

3.3 Modelos textuais e estendidos

Foram realizados também diversos estudos mais detalhados para construção de modelos baseados em dados textuais e estendidos com dados não-textuais. Alguns dos resultados obtidos nestas duas frentes são resumidos a seguir.

Uso de LLMs e engenharia de prompts No que diz respeito à investigação de modelos puramente textuais, foi conduzido um experimento de avaliação do uso da ferramenta ChatGPT² para predição de depressão com base nos dados do cópuz SetembroBR. Os resultados obtidos nesta modalidade de classificação dita *zero-shot* sugerem vantagens e desvantagens desta estratégia em relação aos modelos supervisionados já desenvolvidos (baseados em BERT e BiLSTMs), e são detalhados em [18].

Ainda com relação à tarefa de detecção de depressão a partir de dados textuais, foi investigada também uma abordagem baseada em *prompts* chamada Prompt.Bow que usa o modelo GPT 3.5 para enriquecer um classificador de texto padrão. Essas duas etapas - *prompt* e classificação – são descritas brevemente a seguir.

Primeiramente, o modelo de língua GPT 3.5. foi usado para avaliar uma amostra aleatória de *tweets* do cópuz de acordo com a relevância para a previsão de depressão. Como resultado, estes dados foram categorizados como sendo de alta (1), média (2) ou baixa (3) relevância para a saúde mental. Esta rotulação automática gerada pelo modelo GPT foi então utilizada para treinar um classificador de postagens de alta, média e baixa relevância de acordo com o modelo de língua T5, e este classificador foi então aplicado à rotulação do conjunto completo de 19,42 milhões de postagens do subcópuz de depressão. Este conjunto de rótulos em nível de postagem, provisoriamente chamada de SetembroBRGPT, representa um recurso linguístico-computacional adicional do projeto, e foi parcialmente disponibilizada (no momento apenas para o subconjunto Depressão) para reuso juntamente com o cópuz original (rotulado ao nível de usuário).

A tabela 2 mostra a distribuição dos rótulos na parte de treinamento do cópuz SetembroGPT.

Como esperado, a maioria das publicações (65,2% na classe Diagnosticado e 71,3% na classe Controle) são consideradas de baixa relevância para a predição de saúde mental de acordo com o modelo GPT. Por outro lado, publicações de alta relevância são relativamente raras (5,4% na classe Diagnosticados e 3,0% na classe Controle).

Tão ou mais importante do que esta distribuição, entretanto, é a observação de que ao consultar um LLM sobre a relevância de uma publicação para a saúde

² <https://chat.openai.com/chat>

Table 2. Distribuição dos dados de treino do córpus SetembroGPT de acordo com a relevância para detecção de depressão.

Relevância	Diagnosed class				Control class			
	Tweets	%	Tokens	%	Tweets	%	Tokens	%
alta	39k	2.0%	674k	5.4%	153k	1.1%	2,372k	3.0%
média	180k	9.3%	3,663k	29.4%	1,031k	7.6%	20,566k	25.7%
baixa	1,727k	88.7%	8,122k	65.2%	12,439k	91.3%	56,971k	71.3%

mental estamos nos concentrando em grande parte na semântica, isto é, em sintomas e outros sinais clínicos bem conhecidos de depressão, mas isto não quer dizer que outros fatores possam ou devam ser ignorados. O *prompt* utilizado permite ao modelo identificar uma vasta gama de publicações que podem sugerir, por exemplo, distúrbios alimentares ou uso de linguagem negativa, ambos sabidamente relacionados à depressão [19], mas não considera explicitamente indicadores linguísticos de depressão mais refinados, como o uso de pronomes de primeira pessoa [15, 20], termos absolutos [13] e outros. Estes indicadores, que podem estar presentes em qualquer publicação de baixa ou alta relevância para a depressão, são também importantes preditores do estado de saúde mental, sugerem que nenhuma publicação deveria em princípio ser descartada apenas com base apenas nos resultados do LLM.

Como um meio de manter os dados de treino integralmente disponíveis para o classificador e, ao mesmo tempo, distinguindo mensagens mais e menos relevantes (que mostram claramente diferentes distribuições entre classes na Tabela 2), o modelo proposto Prompt.Bow usou a categoria rótulos fornecidos pelo LLM para dividir os dados de treinamento em subconjuntos de alta, média e baixa relevância, a partir dos quais foram criados três vetores individuais do tipo *bag-of-words* usados na classificação, aos quais é acrescido ainda um vetor de contagens de n-gramas [21] para preservar a ordem destas postagens. Esta abordagem e seus resultados são detalhados em [22].

Uso de dados não textuais No que diz respeito ao uso de dados não-textuais, foram conduzidos diversos experimentos com modelos baseados em características estruturais da rede social representadas por relações de amigos, seguidores e menções do Twitter. Modelos do tipo *ensemble* que combinam estas três fontes de conhecimento são apresentados em [23], o qual modela relações de rede social com uso de *embeddings* produzidos pela ferramenta Node2Vec [24]. O artigo em questão recebeu uma solicitação de correções mínimas em sua mais recente rodada de revisão (novembro/2023), havendo assim razoável expectativa de que possa vir a ser aceito em breve.

Em um estudo análogo, porém tomando emprestado técnicas mais tradicionais de PLN (em especial, a representação do tipo *bag-of-words*), foi conduzido um experimento de predição de depressão e ansiedade a partir de redes de amigos, seguidores e menções utilizando uma representação denominada *bag-of-users*. Detalhes desta abordagem e seus resultados são discutidos em [25].

Nestes estudos, observa-se que tanto no uso de *embeddings* de rede social como na abordagem *bag-of-users*, resultados preliminares sugerem que as informações de interação, ou seja, os usuários com quem um indivíduo estabelece diálogo na rede social, são as mais relevantes para predição de estados de saúde mental.

Em linhas gerais, os resultados destes experimentos sugerem que (1) dados textuais (ou seja, o conteúdo das postagens) são altamente preditivos de depressão/ansiedade mas que (2) dados não-textuais (amigos, seguidores e menções) também apresentam contribuições importantes. Assim, uma parte importante do problema de predição destes transtornos a partir da rede social é a questão de como combinar (1) e (2). Esta questão foi abordada de forma preliminar com a combinação de modelos textuais baseados em BERT e modelos não-textuais utilizando arquiteturas de *ensemble* em [11], e é discutida também em [23].

4 Resultados

Nas seções a seguir enumeramos os resultados obtidos em termos de recursos linguísticos-computacionais, publicações realizadas, e artigos em submissão. Além disso, destacamos ainda a repercussão que o projeto tem obtido junto à imprensa, além de resultados do reaproveitamento dos dados coletados em um segundo projeto, e que constitui um subproduto da presente pesquisa.

4.1 Recursos desenvolvidos

O projeto gerou até o momento o *cópus* SetembroBR propriamente dito [5] e o modelo de língua BERTabaporu [9], ambos voltados ao domínio Twitter em Português. Ambos os recursos estão publicamente disponíveis para reuso a partir dos repositórios divulgados nas respectivas publicações, assim como o código de implementação dos modelos de *baseline* apresentados nos diversos experimentos já realizados [5, 23, 11].

Além disso, considerando os avanços recentes na área de grandes modelos de língua (LLMs), o *cópus* foi também rotulado de forma automática ao nível de postagem utilizando o modelo GPT-3.5. com o objetivo de sinalizar postagens com diferentes graus de relevância para a tarefa de detecção de transtornos de saúde mental. Conforme discutido na seção anterior, estes mais de 19 milhões de rótulos constituem um recurso valioso para pesquisa na área, e já foram disponibilizados para o subconjunto Depressão como forma de complementar o *cópus* já publicado. O subconjunto Ansiedade deverá ser contemplado futuramente, a depender dos recursos computacionais disponíveis.

4.2 Artigos publicados

Até o momento, o projeto gerou três publicações completas em periódicos [5, 11, 23] e cinco artigos em conferências apresentados com recursos do projeto [6, 9, 18, 25, 22].

4.3 Outros resultados

A construção do *córpus* SetembroBR exigiu a coleta inicial de uma grande massa de dados a partir da qual os dados de interesse (ou seja, *timelines* de indivíduos com relatos de depressão ou ansiedade) foram selecionados. O restante dessa massa de dados, de ordem de grandeza várias vezes superior ao efetivamente aproveitado no *córpus*, constitui entretanto um recurso valioso já aproveitado em outras iniciativas de pesquisa.

Além do já descrito modelo BERTabaporu, construído a partir destas e de outras fontes de dados da rede social Twitter, a base de dados existente foi aproveitada para a construção de um segundo *córpus*, denominado UstanceBR, para detecção de posicionamentos ou atitudes (do inglês, *stance*, cf. [26, 27]) de usuários da rede social, com aplicações em áreas como atribuição autoral [28], detecção de discurso de ódio [29], análise de sentimentos [30] e outras.

O *córpus* UstanceBR, descrito em [31], contém assim publicações a respeito de determinados tópicos de interesse (presidentes brasileiros, medidas de tratamento da Covid-19 e outras) e, mesmo sem ter relação direta com o tema de saúde mental, pode ser considerado um subproduto do projeto principal, pois entendemos que sua realização não teria sido possível sem a execução das atividades de coleta de dados deste primeiro ano do projeto SetembroBR. Resultados parciais desta iniciativa foram apresentados em [32].

5 Continuidade da pesquisa

Embora os dois objetivos principais - construção do *córpus* e desenvolvimento de modelos multimodais - tenham em nossa visão sido cumpridos, em um cronograma de 2 anos não é possível esgotar as oportunidades de pesquisa proporcionadas por um recurso complexo e de larga escala como o construído, o qual se prestará ainda a vários anos de pesquisa subsequente no âmbito dos projetos de mestrado e doutorado ainda em andamento. A seguir enumeramos algumas das iniciativas que continuarão sendo desenvolvidas após o término da concessão de recursos.

- Refinamento de modelos textuais baseados em BERT e GPT, enfocando tanto o uso de arquiteturas do tipo *ensemble* e misturas de especialistas já introduzidas em [11], como no uso de métodos mais adequados para tratamento de *timelines* compostas de dados esparsos.
- Experimentos com modelos baseados em relações de interação, a exemplo do já abordado em [23], e sua possível combinação com modelos textuais em soluções ditas multimodais para o problema.
- Detecção precoce de risco de depressão/ansiedade nos moldes da tarefa definida na série de desafios computacionais eRisk [7].
- Avaliação humana, comparando-se as estimativas feitas pelo modelo final com as de especialistas da área de psicologia.
- Análise da tarefa de predição com uso de dados anteriores versus dados posteriores ao diagnóstico, com o objetivo de validar uma das premissas do

projeto que pretende se diferenciar de estudos existentes ao levar em conta a noção de momento do diagnóstico. Para essa análise, foram coletados dados suplementares que não faziam parte do cópulo original (ou seja, publicações posteriores ao momento do diagnóstico de depressão/ansiedade) para investigar o desempenho dos modelos desenvolvidos com base em diferentes porções dos dados de entrada.

- Estudo de outros fenômenos linguísticos que podem ter impacto sobre a produção da língua por indivíduos com e sem depressão em complemento aos indicadores já estabelecidos em [15], como o uso de informação redundante [33, 34] e outros.
- Desenvolvimento e disponibilização pública dos rótulos SetembroBRGPT para o subconjunto Ansiedade como um recurso adicional ao cópulo construído, e possivelmente outras formas de rotulação automática ao nível de postagem, como indicativos de sintomas e outros.

References

1. Rangel, F., Rosso, P., Verhoeven, B., Daelemans, W., Potthast, M., Stein, B.: Overview of the 4th Author Profiling Task at PAN 2016: Cross-Genre Evaluations. In: CLEF 2016 Evaluation Labs and Workshop, Notebook papers, Évora, Portugal, CEUR-WS.org (2016) 750–784
2. Pavan, M.C., dos Santos, V.G., Lan, A.G.J., ao Trevisan Martins, J., dos Santos, W.R., Deutsch, C., da Costa, P.B., Hsieh, F.C., Paraboni, I.: Morality classification in natural language text. *IEEE transactions on Affective Computing* **14**(1) (2023) 857–863
3. Flores, A.M., Pavan, M.C., Paraboni, I.: User profiling and satisfaction inference in public information access services. *Journal of Intelligent Information Systems* **58**(1) (2022) 67–89
4. dos Santos, W.R., Funabashi, A.M.M., Paraboni, I.: Searching Brazilian Twitter for signs of mental health issues. In: 12th International Conference on Language Resources and Evaluation (LREC-2020), Marseille, France, ELRA (2020) 6113–6119
5. dos Santos, W.R., de Oliveira, R.L., Paraboni, I.: SetembroBR: a social media corpus for depression and anxiety disorder prediction. *Language Resources and Evaluation* **58**(1) (2024) 273–300
6. de Oliveira, R.L., Paraboni, I.: Predição de estado de saúde mental a partir de conteúdo textual e relacionamentos em redes sociais. In: III Congresso Internacional em Humanidades Digitais (HDRio2023), Rio de Janeiro (2023)
7. Losada, D.E., Crestani, F., Parapar, J.: Overview of eRisk 2019 Early Risk Prediction on the Internet. In: *Lecture Notes in Computer Science* vol 11696. (2019)
8. Cohan, A., Desmet, B., Yates, A., Soldaini, L., MacAvaney, S., v Goharian: SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. In: COLING-2018, Santa Fe, USA, Association for Computational Linguistics (2018) 1485–1497
9. da Costa, P.B., Pavan, M.C., dos Santos, W.R., da Silva, S.C., Paraboni, I.: BERTabaporu: assessing a genre-specific language model for Portuguese NLP. In: *Recent Advances in Natural Language Processing (RANLP-2023)*, Varna, Bulgaria (2023) 217–223

10. Souza, F., Nogueira, R., Lotufo, R.: BERTimbau: pretrained BERT models for Brazilian Portuguese. In: 9th Brazilian Conference on Intelligent Systems (BRACIS) - LNCS 12319, Cham, Springer (2020)
11. dos Santos, W.R., Yoon, S., Paraboni, I.: Mental health prediction from social media text using mixture of experts. *IEEE Latin America Transactions* **21**(6) (2023) 723–729
12. Pennebaker, J.W., Francis, M.E., Booth, R.J.: *Inquiry and Word Count: LIWC*. Lawrence Erlbaum, Mahwah, NJ (2001)
13. Al-Mosaiwi, M., Johnstone, T.: In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clinical Psychological Science* **6**(4) (2018) 529–542
14. Choudhury, M.D., Gamon, M., Counts, S., Horvitz, E.: Predicting depression via social media. In: International AAAI Conference on Web and Social Media (ICWSM), AAAI (2013)
15. Trifu, R., Nemes, B., Bodea-Hategan, C., Cozman, D.: Linguistic indicators of language in major depressive disorder (MDD). An evidence based research. *Journal of Evidence-Based Psychotherapies* **17** (03 2017) 105–128
16. Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E.: Adaptive mixtures of local experts. *Neural computation* **3**(1) (1991) 79–87
17. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT-2019, Minneapolis, USA, Association for Computational Linguistics (2019) 4171–4186
18. dos Santos, W.R., Paraboni, I.: Predição de transtorno depressivo em redes sociais: BERT supervisionado ou ChatGPT zero-shot? In: XIV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL-2023), Porto Alegre, Brasil, SBC (2023) 11–21
19. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders 5th edition*. American Psychiatric Association, Arlington, VA (2013)
20. Paraboni, I., de Lima, V.L.S.: Possessive pronominal anaphor resolution in Portuguese written texts. In: Proceedings of the 17th international conference on Computational linguistics-Volume 2, Association for Computational Linguistics (1998) 1010–1014
21. Pereira, D.B., Paraboni, I.: A language modelling tool for statistical NLP. In: 5th Workshop on Information and Human Language Technology (TIL-2007). Anais do XXVII Congresso da SBC, Rio de Janeiro, Sociedade Brasileira de Computação (2007) 1679–1688
22. dos Santos, W.R., Paraboni, I.: Prompt-based mental health screening from social media text. In: XIII Brazilian Workshop on Social Network Analysis and Mining (BraSNAM-2024), Brasília, SBC (2024)
23. de Oliveira, R.L., ao Trevisan Martins, J., Paraboni, I.: Mental health prediction from social media connections. *New Review of Hypermedia and Multimedia* **29**(3–4) (2024) 225–244
24. Grover, A., Leskovec, J.: node2vec: Scalable Feature Learning for Networks. In: KDD16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, USA, Association for Computing Machinery (2016) 855–864
25. de Oliveira, R.L., Paraboni, I.: A Bag-of-Users approach to mental health prediction from social media data. In: 16th International Conference on Computational Processing of Portuguese (PROPOR 2024), Santiago de Compostela, Spain (2024)

26. dos Santos, W.R., Paraboni, I.: Moral Stance Recognition and Polarity Classification from Twitter and Elicited Text. In: *Recent Advances in Natural Language Processing*, Varna, Bulgaria (2019) 1069–1075
27. Pavan, M.C., dos Santos, W.R., Paraboni, I.: Twitter Moral Stance Classification using Long Short-Term Memory Networks. In: *9th Brazilian Conference on Intelligent Systems (BRACIS)*. LNAI 12319, Springer (2020) 636–647
28. Custódio, J.E., Paraboni, I.: Stacked authorship attribution of digital texts. *Expert Systems with Applications* **176** (2021) 114866
29. da Silva, S.C., Ferreira, T.C., Ramos, R.M.S., Paraboni, I.: Data driven and psycholinguistics motivated approaches to hate speech detection. *Computación y Sistemas* **24**(3) (2020) 1179–1188
30. Lin, C., Hu, P., Su, H., Li, S., Mei, J., Zhou, J., Leung, H. In: *SenseMood: Depression Detection on Social Media*. Association for Computing Machinery, New York, USA (2020) 407–411
31. Pereira, C., Pavan, M., Yoon, S., Ramos, R., Costa, P., Cavalheiro, L., Paraboni, I.: UstanceBR: a multimodal language resource for stance prediction. arXiv:2312.06374 (2023)
32. Pavan, M.C., Paraboni, I.: Cross-target stance classification as domain adaptation. In Pichardo Lagunas, O., Martínez-Miranda, J., Martínez Seis, B., eds.: *Advances in Computational Intelligence - MICAI 2022 - Lecture Notes in Artificial Intelligence* vol 13612, Cham, Springer Nature Switzerland (2022) 15–25
33. Paraboni, I.: *Generating references in hierarchical domains: the case of Document Deixis*. PhD thesis, University of Brighton (2003)
34. Teixeira, C.V.M., Paraboni, I., da Silva, A.S.R., Yamasaki, A.K.: Generating relational descriptions involving mutual disambiguation. In: *Computational Linguistics and Intelligent Text Processing (CICLing-2014)*, Lecture Notes in Computer Science 8403, Kathmandu, Nepal, Springer (2014) 492–502