

# Automatic detection of depression from text data

## A systematic literature review

Felipe Magami

felipe.magami@usp.br

Escola de Artes, Ciências e Humanidades, USP  
São Paulo, Brazil

Luciano Antonio Digiampietri

digiampietri@usp.br

Escola de Artes, Ciências e Humanidades, USP  
São Paulo, Brazil

### ABSTRACT

Depression is a mental disorder that affects hundreds of millions of people worldwide, with potentially serious consequences if left without treatment. Despite that, many people still suffer from depression without a diagnosis. Recently, the amount of studies related to the automatic detection of depression has improved. The objective of this paper is to identify the methods and techniques used by studies about depression detection through text data, by conducting a systematic review.

### CCS CONCEPTS

• **Information systems** → **Social networks**; • **General and reference** → **Surveys and overviews**.

### KEYWORDS

Detection of depression, Social networks, Sentiment analysis

#### ACM Reference Format:

Felipe Magami and Luciano Antonio Digiampietri. 2020. Automatic detection of depression from text data: A systematic literature review. In *XVI Brazilian Symposium on Information Systems (SBSI'20)*, November 3–6, 2020, São Bernardo do Campo, Brazil. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3411564.3411603>

## 1 INTRODUÇÃO

A depressão é um transtorno caracterizado pela falta de interesse e prazer, tristeza, sentimento de culpa, falta de autoconfiança, cansaço persistente, falta de concentração e distúrbios do sono, entre outros sintomas [51]. Se não tratada, a depressão pode, em casos extremos, levar ao suicídio. Segundo a Organização Mundial de Saúde [55], em 2017, o número de pessoas com depressão no mundo inteiro foi de mais de 300 milhões, número que tende a crescer. No Brasil, os depressivos representam 11,5 milhões, uma das maiores taxas de depressão no mundo e a maior da América Latina. Provisões e serviços para suporte e identificação de transtornos mentais em geral ainda são considerados insuficientes [20], e muitas vezes os próprios depressivos podem não procurar ajuda, não só pelo preconceito que eles possam vir a sofrer, mas também pela vergonha e estigma que eles sentem por si [90]. Os transtornos mentais podem acarretar

em grandes custos, grande parte devido a causas indiretas, como perdas com morte precoce, absenteísmo e presenteísmo [70]. Mais de 700 mil pessoas morrem por ano por suicídio, uma das principais causas de morte para a faixa etária entre os 15 e 29 anos [55].

Existem várias pesquisas com interesse em detectar automaticamente pessoas potencialmente depressivas com o intuito de facilitar esse diagnóstico e promover maior busca por tratamento. Por exemplo, Alghowinem [4] descreve uma abordagem multimodal para detecção de depressão utilizando atributos vocais, movimento dos olhos e balanço da cabeça. Similarmente, outros trabalhos [32, 36, 103] também utilizaram características acústicas de gravações de falas de depressivos para identificação.

Vários trabalhos estudam a relação entre linguagem e transtornos mentais [63]. Rude, Gortner e Pennebaker [74] identificaram que universitários americanos depressivos utilizavam a palavra “eu” e palavras relacionadas a emoções negativas com maior frequência. Já Al-Mosaiwi e Johnstone [2] encontraram palavras “absolutistas”, isto é, palavras com maior grau de polarização, como “absolutamente”, “tudo” e “nunca”, em maior frequência em fóruns de internet relacionados a depressão, ansiedade e ideação suicida do que em outros fóruns.

O presente trabalho apresenta uma revisão sistemática da literatura com o objetivo de identificar os métodos e técnicas utilizados para a tarefa de detecção de depressão a partir de dados textuais.

## 2 METODOLOGIA E CONDUÇÃO

A revisão sistemática da literatura é uma metodologia empregada com o propósito de identificar, avaliar e interpretar pesquisas pertinentes a um tópico, pergunta ou fenômeno de interesse em específico [37]. Em geral, uma revisão sistemática é dividida em três etapas: o planejamento, em que são definidos os propósitos da revisão, seus objetivos e as perguntas a serem respondidas, bem como detalhes acerca das bibliotecas científicas a serem adotadas e os critérios de seleção de trabalhos; a condução, em que, de acordo com as decisões tomadas no passo anterior, os trabalhos são identificados, selecionados e os dados relevantes são extraídos e sintetizados; e, por último, a divulgação, em que os resultados da revisão são documentados e publicados. Devido à limitação de espaço, apenas os pontos julgados mais importantes do processo da revisão sistemática realizada neste trabalho serão apresentados.

Como mencionado, esta revisão teve como objetivo identificar os métodos e técnicas empregados para detecção de depressão a partir de dados de texto. Ela procurou responder às seguintes perguntas:

- (1) Quais são os métodos e técnicas utilizados para a criação (ou cálculo ou extração) de atributos a partir de dados textuais (incluindo características sobre uso de palavras, gramática e representação vetorial)?

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*SBSI'20, November 3–6, 2020, São Bernardo do Campo, Brazil*

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-8873-3/20/11...\$15.00

<https://doi.org/10.1145/3411564.3411603>

- (2) Quais são os modelos de classificação<sup>1</sup> utilizados para detecção de depressão a partir de dados textuais?
- (3) Quais foram os conjuntos de dados textuais utilizados?
- (4) Como a classificação real foi encontrada (isto é, de que forma os depressivos foram identificados)?

Para a etapa de seleção dos trabalhos relevantes, apenas critérios de inclusão foram definidos e o artigo deve obedecer a todos estes critérios para ser selecionado. Os critérios são:

- I1. Serão incluídos trabalhos publicados e disponíveis integralmente em bases de dados científicas e acessíveis pelos autores;
- I2. Serão incluídos trabalhos que já possuam aprovação pela comunidade científica (isto é, foram submetidos a um processo de revisão por pares);
- I3. Serão incluídos trabalhos que abordam métodos e técnicas para detecção de depressão a partir de dados textuais;

As bibliotecas digitais utilizadas para a condução desta revisão foram a IEEE Xplore Digital Library, ACM Digital Library e Scopus. Consideramos que essas bibliotecas sejam abrangentes o suficiente para alcançarmos pelo menos a maioria dos trabalhos disponíveis pertinentes para a nossa revisão. A expressão de busca utilizada foi a expressão “*depress\**” AND (“*language*” OR “*text*”). Para as duas primeiras bibliotecas, o filtro “Metadata Only” foi utilizado. Para a biblioteca Scopus a expressão de busca foi aplicada apenas para o título, resumo e palavras-chave, restringindo-se para as áreas de computação, matemática e engenharia. Não foi utilizado refinamento por ano ou critérios de qualidade. As buscas nas bibliotecas digitais foram realizadas entre o período de 27/04/2019 e 17/09/2019. Não foi realizado nenhum filtro em relação à data de publicação dos artigos. Os trabalhos encontrados foram examinados a partir de seus títulos, palavras-chave e resumos e selecionados a partir dos critérios mencionados anteriormente. A Tabela 1 apresenta a quantidade de trabalhos encontrados com a busca descrita e a quantidade de trabalhos incluídos.

**Tabela 1: Resultados das buscas nas bibliotecas científicas**

Base	Encontrados	Incluídos
IEEE	128	17
ACM	65	12
Scopus	524	47

### 3 RESULTADOS E DISCUSSÃO

Todos os 76 trabalhos incluídos na etapa de condução foram lidos na íntegra e as informações referentes ao objetivo e às perguntas definidas no planejamento foram extraídas. As Tabelas 2 e 3 contêm a lista dos trabalhos incluídos e os dados extraídos.

<sup>1</sup>Para esta revisão também foram considerados modelos para regressão. Trabalhos que tiveram como resultado predições de pontuação para depressão poderiam ser transformados em classes estabelecendo intervalos para diferentes níveis de depressão, estratégia equivalente aos questionários de diagnóstico de depressão utilizado por profissionais.

### 3.1 Conjuntos de dados

Percebe-se que grande parte dos trabalhos se referem a participações em desafios, especificamente nos desafios eRisk para detecção precoce de risco na internet dos anos 2017 [44], 2018 [45] e 2019 [46]; e Audio/Visual Emotion Challenge and Workshop (AVEC) de 2014 [93], 2016 [92] e 2017 [72]. Para o desafios eRisk dos anos 2017 e 2018, o conjunto de dados primeiramente proposto em [43] foi adotado, em que postagens do Reddit foram coletadas e os usuários depressivos foram localizados a partir de expressões específicas, como “Fui diagnosticado com depressão”, e verificados manualmente. Como o desafio aborda a detecção precoce, a tarefa é dividida em etapas, em que para cada semana um subconjunto das postagens é disponibilizado e os participantes decidem se devem realizar a classificação de um usuário ou se devem esperar até que novas postagens sejam disponibilizadas, de forma que classificações corretas com menos postagens recebem uma pontuação maior. Para o eRisk 2019, foram propostas algumas mudanças: postagens de redes sociais de usuários foram disponibilizadas de forma completa e os participantes tiveram como objetivo prever a pontuação para cada uma das perguntas encontradas no Inventário de Depressão de Beck (BDI) [9].

Os conjuntos de dados utilizados pelo desafio AVEC são de gravações de interações humano-computador. Em particular, o conjunto Distress Analysis Interview Corpus (DAIC) [28], adotado nas tarefas de 2016 e 2017, também foi utilizado pelos estudos em [5, 15, 40]. Por causa da natureza dos dados, alguns trabalhos do AVEC e de conjuntos de áudio e vídeo tendem a focar menos nos atributos textuais. Trabalhos como em [102] e [81] utilizam informações sobre o conteúdo próprio do DAIC, por exemplo, se o participante já foi diagnosticado previamente com depressão ou estresse pós-traumático e se ele sofre de problemas de sono, combinadas com outras características como gestos, movimentação da cabeça e segmentação de voz.

Analisando os artigos não-relacionados a estes desafios, observa-se que a maior parte dos conjuntos são oriundos de dados online, em particular de redes sociais, como Twitter [10, 18, 20, 22, 33, 58, 104], Facebook [19, 35, 99] e LiveJournal [52–54, 105]. Um conjunto interessante é o in-the-Wild Speech Corpus (WSM), coletado em [85] e utilizado também por [15] e [16], de vídeos do YouTube sobre um determinado tema. Embora o foco do estudo não seja só em depressão, vídeos sobre depressão foram coletados pela expressão “vlog de depressão” e verificados manualmente.

### 3.2 Classificação real

A identificação das classes reais, isto é, se o indivíduo ou postagem apresenta ou não conteúdo depressivo, ocorreu de acordo com duas abordagens: uma que inclui o contato direto (pessoalmente, por questionários, etc) com a pessoa que produziu os dados e outra sem nenhum contato.

No caso das anotações sem contato/envolvimento, os usuários ou postagens são identificados com base em um tipo de estratégia, como procurar por sintomas depressivos específicos [23] ou por expressões específicas [43].

Tabela 2: Síntese dos artigos selecionados - parte 1

Ref.	Conjunto de dados	Classificação real	Criação de atributos textuais	Modelos de classificação
[58]	Twitter	Anotação por especialistas	Pos <i>tagging</i> e dicionários	Fórmula de pontuação [34]
[22]	Twitter	Não especificado	Pos <i>tagging</i> e BoW	SVM e NB
[49]	Conjunto em AVEC 2014 [93]	BDI-II	LIWC, n-gramas de palavras, caracteres e PoS e métricas de contagem	SVM
[100]	AVEC 2016 [92]	PHQ-8	Métodos de <i>embedding</i>	SVM
[35]	Facebook	Questionário de saúde mental tailandês	Métricas de contagem, análise de sentimentos e PoS <i>tagging</i>	SVM, florestas randômicas (RF) e redes neurais profundas
[79]	AVEC 2017 (DAIC [28])	PHQ-8	BoW e métodos de <i>embedding</i>	SVM
[7]	Gravações de interações entre adolescentes e seus pais	Diagnóstico clínico	Dicionário de valência e excitação ([97])	SVM
[23]	Diary.com	Anotação manual	Sintomas do questionário de depressão fornecido na plataforma PsychCentral e no Unified Medical Language System (UMLS)	Modelos Bayesianos
[90]	eRisk 2017 (Reddit [43])	Expressões específicas	Presença do pronome "Eu", PoS <i>tagging</i> , pontuações de legibilidade, métricas de contagem, expressões selecionadas, nomes de antidepressivos, LIWC e métodos de <i>embedding</i>	CNN (para <i>embedding</i> ), LR (para outros atributos) e ensemble
[33]	Twitter	Expressões específicas	Dicionários de sentimentos (VSO, SentStrength, WordNet) e de <i>emoicons</i> e PoS <i>tagging</i>	SVM
[82]	Reddit [65]	De acordo com grupo de origem das postagens	LIWC, n-gramas e LDA	LR, SVM, RF, AdaBoost e Perceptron multiclasse (MLP)
[15]	WSM [85] e DAIC [28]	Anotação manual (WSM) e PHQ-8 (DAIC)	BoW	SVM
[52]	LiveJournal	De acordo com grupo de origem das postagens	ANEW, LIWC, LDA e etiquetas de humor do LiveJournal	Lasso (LR)
[14]	Comunidades online	Extração automática com verificação manual de sintomas do questionário ICM-10	N-gramas	SVM, NB e árvores de decisão (DT)
[40]	DAIC [28]	PHQ-8	BoW, modelagem de tópicos manual e <i>data augmentation</i>	Modelos de redes de transformação
[84]	Mensagens de textos e postagens em redes sociais	PHQ-9	Análise de sentimentos com ferramenta TextBlob, avaliação de subjetividade do texto, PoS <i>tagging</i> e métricas de contagem	Algoritmo k-vizinhos mais próximos (KNN), SVM, RF, AdaBoost, XGBoost, LR e NB.
[3]	Mensagens de textos entre pacientes e psicólogos	Anotação manual	N-gramas	KNN, SVM, gradiente descendente estocástico (SGD), NB, DT, RF e AdaBoost
[102]	AVEC 2017 (DAIC [28])	PHQ-8	Métricas de contagem e termos específicos e atributos semânticos relacionados a trechos específicos do conjunto	SVM
[81]	AVEC 2017 (DAIC [28])	PHQ-8	Atributos semânticos relacionados a trechos específicos do conjunto	RF
[99]	Facebook	Anotação manual	Dicionários chineses de emoção e <i>emoicons</i> e métodos de <i>embedding</i>	Pontuação através dos dicionários e CNN
[76]	eRisk 2017 (Reddit [43])	Expressões específicas	BoW, métodos de <i>embedding</i> treinados em n-gramas relacionados a depressão e conceitos médicos	SVM e redes neurais recorrentes (RNN)
[59]	AVEC 2016 (DAIC [28])	PHQ-8	Métricas de contagem e palavras depressivas, ANEW e LIWC	DT e ensembles
[101]	AVEC 2017 (DAIC [28])	PHQ-8	Atributos semânticos relacionados a trechos específicos do conjunto e métodos de <i>embedding</i>	CNN
[98]	AVEC 2016 (DAIC [28])	PHQ-8	Métodos de <i>embedding</i> e atributos semânticos relacionados a trechos específicos do conjunto	CNN
[19]	Facebook	Autodeclaração e PHQ-9	LIWC e palavras relacionadas a perguntas (e.g. "como", "por quê")	SVM
[18]	Twitter	Autodeclaração e CES-D	LIWC, ANEW e n-gramas	SVM
[104]	Twitter	Expressões específicas com anotação manual por sintomas do PHQ-9	LDA com termos sementes a partir de sintomas do PHQ-9 e palavras relacionadas a depressão	Classificação a partir do modelo de tópicos gerado, SVM e NB
[29]	AVEC 2014 [93]	BDI-II	Dicionários próprios de excitação e valência, modelos de distribuição semântica, n-gramas e PoS <i>tagging</i>	SVM
[17]	AVEC 2017 (DAIC [28])	PHQ-8	BoW, PoS e n-gramas indexados e dicionários ANEW, EmoLex, SentieNet e Laswell	Regressores Gaussianos e máquinas de vetores relevantes
[31]	Western Collaborative Study [73]	CES-D	PoS <i>tagging</i> , LIWC, avaliador de densidade de ideias (CPIDR) e frequência de palavras em primeira pessoa	LR, MLP
[95]	eRisk 2017 (Reddit [43])	Expressões específicas	Modelagem de grafos de gramas de palavras e letras	Cálculo de medidas de similaridade entre o documento de interesse e os grafos das classes gerados
[67]	eRisk 2018 (Reddit [43])	Expressões específicas	BoW, PoS <i>tagging</i> , termos de negação, capitalizações, pontuações, <i>emoicons</i> , métricas de contagem, palavras em primeira pessoa, palavras "absolutistas", termos no pretérito e relacionados, sintomas e medicamentos relacionados a depressão, frequência da expressão "depress", n-gramas, análise de sentimentos, termos relacionados a sono, pontuação de legibilidade e métodos de <i>embedding</i>	RF e LR
[56]	Conjunto de áudio em turco	Diagnóstico clínico	Radicals, afixos, PoS, análise de sentimentos, dicionário de psicologia Harvard-III e BoW	NB, LR Bayesiano e SVM
[71]	Redação de universitários sobre seus sentimentos [74]	BDI	LDA e LIWC	Regressão linear
[27]	AVEC 2017 (DAIC [28])	PHQ-8	LIWC e atributos semânticos relacionados a trechos específicos do conjunto	RF, SGD e SVM
[13]	eRisk 2019 [46]	BDI	Framework SS3 [12]	Framework SS3 [12]
[69]	Redações em alemão sobre retrospectiva do ano anterior e expectativas para o ano seguinte	BDI-II	LIWC	LR
[6]	eRisk 2017 (Reddit [43])	Expressões específicas	N-gramas, PoS <i>tagging</i> , dicionários (SentieNet, medicamentos, doenças relacionadas e drogas) e métricas de contagem	Recuperação de informação com algoritmo baseado em KNN, ensemble de SVM, ensemble de RF, ensemble dos métodos de classificação e de recuperação de informação.
[62]	Sina Weibo	Expressões específicas	Dicionário de emoções, dicionário de emoções para linguagem de internet e dicionário de <i>emoicons</i>	SVM, NB, KNN, DT, ensemble

Tabela 3: Síntese dos artigos selecionados - parte 2

Ref.	Conjunto de dados	Classificação real	Criação de atributos textuais	Modelos de classificação
[107]	Redações de pessoas com dores crônicas	CES-D	LIWC	LR
[34]	Psycho-Babble	Anotação de especialistas	Dicionários de sintomas e sinônimos e de palavras frequentes, pronomes e palavras de negação	Pontuação dada por fórmula própria
[106]	Blog chines "Tree Hole"	Não especificado	Métricas de contagem, <i>emoicons</i> , palavras em primeira pessoa, padrões chineses nas sentenças e métodos de <i>embedding</i>	SVM e CNN
[96]	eRisk 2018 (Reddit [43])	Expressões específicas	BoW	CNN
[57]	eRisk 2018 (Reddit [43])	Expressões específicas	BoW	SVM
[12]	eRisk 2017 (Reddit [43])	Expressões específicas	Framework SS3	Framework SS3
[61]	eRisk 2017 (Reddit [43])	Expressões específicas	BoW, análise semântica concisa, n-gramas e LIWC	NB, RF e DT
[88]	eRisk 2017 (Reddit [43])	Expressões específicas	BoW, métodos de <i>embedding</i> , pronomes pessoais e possessivos, verbos no pretérito, métricas de contagem e índices de legibilidade	LR e RNN
[26]	eRisk 2018 (Reddit [43])	Expressões específicas	Análise semântica concisa e abordagem similar ao Framework SS3 [12]	LR, SVM e abordagem similar ao Framework SS3 [12]
[1]	eRisk 2019 [46] e conjunto próprio de universitários de psicologia para treinamento (aprendizado por transferência)	BDI	Modelos de pré-treinamento generativo (GPT) e LIWC	SVM
[43]	Reddit	Expressões específicas	BoW	LR
[77]	AVEC 2017 (DAIC [28])	PHQ-8	Métricas de contagem, dicionário de palavras relacionadas à depressão e dicionário AFINN	RF
[66]	eRisk 2018 (Reddit [43])	Expressões específicas	Métodos de <i>embedding</i> , sentenças positivas e negativas, inversão Bayesiana	MLP, SVM e média móvel
[78]	eRisk 2017 (Reddit [43])	Expressões específicas	BoW, métodos de <i>embedding</i> , PoS <i>tagging</i> e métricas de contagem	SVM e RF
[89]	eRisk 2018 (Reddit [43])	Expressões específicas	BoW, métodos de <i>embedding</i> , pronomes pessoais e possessivos, verbos no pretérito, métricas de contagem, índices de legibilidade, dicionário de medicamentos, frases específicas e LIWC	LR e RNN
[80]	Transcrição do áudio de casos no MentalHelp.net	Questionários BDI-II e DSM-5	Dicionários baseados no SentiWordNet, PoS, sintomas, palavras de intensidade e de fatores relevantes como família e histórico médico	Pontuação dada por construção de rede semântica
[94]	eRisk 2019 [46]	BDI	Dicionários Multi Perspective Question Answering, WordNet, de desejos sexuais e de distúrbios alimentares, <i>tokens</i> selecionados por informação mútua e medidas de similaridade semântica dadas por vetores de <i>embeddings</i>	LR, linear de pontuações e ensemble
[105]	LiveJournal	Postagens com a expressão regular "depress" no título	Dicionários ANEW, MPQA e SentiWordNet	NB e SVM
[5]	DAIC ([28])	PHQ-8	Métodos de <i>embedding</i>	LR e RNN
[30]	Blogs japoneses	Expressões específicas	N-gramas	NB, SVM e LR
[75]	eRisk 2017 (Reddit [43])	Expressões específicas	Dicionário de palavras relacionadas à expressão regular "depress" do Yahoo! Respostas e conceitos do UMLS	SVM, RNN e ensemble
[24]	eRisk 2017	Expressões específicas	N-gramas	NB
[50]	Blogs japoneses	Categoria das postagens anotadas pelos autores dos blogs	Palavras-chave do Self-Rating Depression Scale e modalidade gramatical e semântica dos termos em japonês	SVM
[42]	AVEC 2016 (DAIC [28])	PHQ-8	Métodos de <i>embedding</i>	SVM e LR
[86]	eRisk 2019 [46]	BDI	Análise de sentimentos com TextBlob, pronomes pessoais, palavras absolutistas, dicionário de antidepressivos e distúrbios mentais e palavras-chave específicas	Pontuação própria
[60]	eRisk 2018 (Reddit [43])	Expressões específicas	BoW, conceitos do UMLS e métodos de <i>embedding</i>	SVM, RF, AdaBoost, LR e RNN
[91]	Forum PPT	Anotação por especialistas	Dicionários de eventos negativos, emoções negativas, sintomas e pensamentos negativos e PoS <i>tagging</i>	Fórmula de pontuação própria
[21]	Reddit [43]	Expressões específicas	Dicionários LIWC, Wariner, NRC e SenticNet 5, índices de legibilidade, frequência do pronome "eu" e métodos de <i>embedding</i>	RF
[68]	eRisk 2018 (Reddit [43])	Expressões específicas	Dicionário LIWC e de antidepressivos, palavras absolutistas e n-gramas	LR e RF
[25]	Reddit	Grupo de origem e expressões específicas	Dicionário LIWC, palavras absolutistas e métricas de contagem	SVM, MLP e LR
[16]	WSM [85]	Anotação manual	BoW e redes neurais recorrentes de tensores para características de sentimento	LR e SVM
[85]	WSM e DAIC [28]	Anotação manual (WSM) e PHQ-8 (DAIC)	BoW e redes neurais recorrentes de tensores para características de sentimento	LR e SVM
[54]	LiveJournal	De acordo com grupo de origem das postagens	LDA e LIWC	LR, SVM e NB
[48]	AVEC 2014 [93]	BDI-II	BoW e PoS <i>tagging</i>	SVM
[10]	Twitter	Anotação manual	Similaridade de grafos de sentimentos, auxiliado por WordNet	SVM, KNN e DT
[53]	LiveJournal	De acordo com grupo de origem das postagens	LDA e LIWC	LR
[20]	Twitter	CES-D e BDI	Dicionários LIWC, ANEW, de termos depressivos retirados do Yahoo! Respostas e de antidepressivos	SVM

Alguns trabalhos separam grupos depressivos de não depressivos de acordo com sua origem, como em [52], em que as postagens são categorizadas de acordo com o grupo do LiveJournal de origem (depressão, bipolaridade, separação, suicídio e auto-mutilação para o grupo clínico e animais de estimação, moda, comida, parentalidade e tecnologia para o grupo de controle), ou em [50], em que postagens em blogs já estavam categorizadas de acordo com tags colocadas pelos próprios autores.

Para os conjuntos anotados com envolvimento dos autores dos textos a estratégia mais utilizada é a aplicação de questionários autoavaliativos, frequentemente aplicados por psicólogos para auxílio ao diagnóstico. Os mais frequentes foram o Patient Health Questionnaire com oito [39] e nove [38] itens (PHQ-8 e PHQ-9), o mesmo utilizado no DAIC; o Beck Depression Inventory (BDI [9] e BDI-II [8]); e o Center for Epidemiologic Studies - Depression (CES-D).

Segundo [43], o motivo de se utilizar anotações automáticas ou manuais, sem envolvimento com os autores dos textos depressivos, é que justamente esse contato pode ser laborioso. Porém, isso pode limitar a qualidade e volume dos dados, ainda que isso possa degradar a qualidade das classificações corretas (não se pode garantir, por exemplo, que uma postagem sem conteúdo depressivo aparente não tenha sido escrita por uma pessoa com depressão). Para os conjuntos de dados de vídeo, com exceção do WSM, por ser de vídeos do Youtube, a aplicação de questionários é frequente, possivelmente por este contato direto com os participantes dos vídeos ser normalmente inevitável. Trabalhos como [19, 20, 84] utilizaram uma abordagem de *crowdsourcing* em que, ao mesmo tempo em que são recrutados participantes dispostos a disponibilizar suas publicações em redes sociais para comporem o *corpus* do estudo, os *crowdworkers* devem responder aos questionários de depressão. Ainda assim, no trabalho em [19] para detecção de mulheres com depressão pós-parto no Facebook, participantes que foram diagnosticadas como tendo depressão pelo questionário PHQ-9 mas que declararam não ter diagnóstico prévio de depressão foram descartadas para evitar essa discrepância. É importante notar que apenas os participantes dos trabalhos de transcrições de áudio de pessoas de uma clínica psiquiátrica [56] e de gravações de interações entre adolescentes e seus pais [7] foram avaliados clinicamente para serem diagnosticados como depressivos ou não. Isso se deve, provavelmente, pelas dificuldades mencionadas anteriormente.

### 3.3 Criação de atributos

Diversos métodos de criação/extração de características textuais foram observados. Técnicas habitualmente utilizadas em outras tarefas de processamento de língua natural e mineração de texto, como *Bag-of-Words* (BoW), *n*-gramas e *Part-of-Speech tagging* (PoS *tagging*) foram frequentes. O uso de PoS *tagging* para depressão é particularmente interessante, já que alguns estudos já verificaram singularidades no uso da língua por depressivos, como maior uso de pronomes em primeira pessoa [74] e verbos no pretérito [87]. [34] realiza classificação a partir da pontuação dada por uma fórmula baseada em palavras de frequência (como “sempre”, “nunca” e “geralmente”), sintomas depressivos, pronomes e negações. Tal abordagem também foi utilizada por [58]. Dicionários voltados para a aplicação em computação, como o Linguistic Inquiry and Word Count (LIWC) [83], com categorias relacionadas a PoS, análise de

sentimentos e interesses pessoais, e o Affective Norms for English Words (ANEW) [11], para valência e excitação afetiva das palavras também foram empregados. Realizar uma análise de emoções nos textos pode ser útil levando em conta o impacto emocional que a depressão provoca. Alguns trabalhos também utilizaram dicionários de criação própria, como os dicionários de [34] para termos de frequência e sintomas. O uso de modelagem de tópicos a partir da técnica Latent Dirichlet Allocation [52–54, 71, 82, 104] não foi tão comum quanto o das outras técnicas.

Métodos de *word embedding*, que ganharam popularidade após a introdução do *word2vec* [47], também são bastante frequentes em trabalhos mais recentes. Os modelos de *embedding* adotados são variados; [79] faz uso de modelos do *word2vec* treinados em dados do Google News, já [100] utiliza *embedding* de parágrafo, enquanto [42] emprega vetores pré treinados do GloVe [64].

O trabalho em [67] utiliza a maioria das abordagens de extração de atributos textuais encontradas. Vários modelos foram criados para a participação no eRisk 2018, a partir de duas abordagens: uma baseada em características, que inclui BoW, frequência de PoS, termos de negação, capitalizações, pontuação, *emoticons*, métricas de contagem, como quantidade de postagens e comentários, palavras em primeira pessoa, palavras “absolutistas”, como mencionado anteriormente, expressões temporais remetendo ao passado, sintomas e medicamentos relacionados à depressão, frequência da expressão “depress\*”, *n*-gramas selecionados por relevância, análise de sentimentos, termos relacionados ao sono, análise de legibilidade do texto e métodos de *embedding doc2vec* [41], treinados utilizando modelos de floresta randômica (RF) e regressores logísticos (LR).

### 3.4 Classificadores

Embora tenha sido observada uma grande variedade de modelos de classificação e regressão, os mais utilizados foram máquina de vetores de suporte (SVM) e LR. Métodos de árvores e modelos baseados em algoritmos Bayesianos como Naive Bayes (NB) também foram bastante utilizados. Redes neurais convolutivas (CNN) também foram frequentes quando vetores de *embedding* foram utilizados. Alguns trabalhos não utilizaram modelos de aprendizado clássicos, utilizando, por exemplo, uma função de pontuação [34] ou a abordagem SS3 descrita em [12] e também utilizada por [13, 26].

Na abordagem SS3, o texto é dividido em subníveis compostos de partes de subníveis inferiores e, para cada parte de um subconjunto, é calculado um valor de confiança de “pertencimento” daquela parte em cada classe, como um valor indicando que a palavra “apple” (“maçã”) tem certo pertencimento às classes “comida” e “tecnologia”, mas pouco pertencimento à classe “viagem”. Este valor leva em consideração a frequência da palavra na classe, a partir de textos coletados anteriormente, e ponderada por termos que tentam diminuir este valor de acordo com o valor de confiança para outras classes, de forma que as palavras não tenham valores altos para muitas classes e *stop words* tenham valor próximo de zero. As partes são então combinadas em seus níveis até que um valor para cada classe seja dado para o documento a ser classificado.

## 4 CONSIDERAÇÕES FINAIS

Neste trabalho foi realizada uma revisão sistemática da literatura para identificar trabalhos relacionados à detecção de depressão a

partir de dados textuais. Mais especificamente, objetivou-se identificar os métodos e técnicas computacionais empregados, bem como os conjuntos de dados utilizados. Espera-se que este trabalho possa incentivar o desenvolvimento de mais ferramentas que automatizem parte do processo de diagnóstico de depressivos, auxiliando e complementado assim o processo realizado hoje em dia.

Observou-se que competições de detecção de depressão incentivaram o desenvolvimento desta área. Uma grande parcela dos trabalhos encontrados está relacionada aos desafios propostos pelo eRisk e pelo AVEC. Fora do escopo dessas competições, os dados utilizados são predominantemente originários de fóruns e redes sociais online, com quase nenhuma disponibilização pública dos dados. A língua predominante dos conjuntos utilizados é a inglesa, embora outras línguas, como a chinesa, japonesa e alemã, tenham sido observadas. Não foi encontrado nenhum trabalho utilizando textos em português.

Os métodos de classificação real são variados. A utilização de anotação manual e automática a partir de estratégias como busca por expressões específicas é popular, mas é mais vulnerável a erros. Uma estratégia mais rigorosa, como uma avaliação clínica dos autores dos textos, é muitas vezes laboriosa e dificulta a coleta de uma amostra com dados de um grande número de indivíduos. Questionários de depressão podem ser um bom meio-termo, mas mesmo assim podem ser falíveis. Levando isso em conta, uma solução seria a criação e disponibilização de um grande conjunto de dados com um método de classificação real dos dados confiável. Não só isso encorajaria mais trabalhos relacionados à área, como também estabeleceria uma forma de comparação de desempenho para os diferentes métodos de processamento de texto e classificação encontrados fora do escopo dos desafios mencionados.

Para conjuntos de transcrição de áudio, é comum a aplicação de questionários para a identificação do nível de depressão de cada participante, mas, para conjuntos de redes sociais e fóruns, outros métodos são frequentes, como a anotação manual por especialistas, expressões específicas e o uso do próprio tópico do fórum de origem.

As técnicas para transformação dos textos em atributos para aplicação em métodos de classificação ou regressão também são diversas, mas observa-se o grande uso de estratégias convencionais da área de processamento de língua natural, como BoW, análise de sentimentos, uso de dicionários e *word embeddings*. Nota-se, contudo, que é frequente o uso de características textuais mais específicas para depressão, como o uso de dicionários de antidepressivos e conceitos médicos. O uso de características relacionadas a pronomes pessoais e palavras “absolutistas” também são comuns, justificado por observações empíricas de estudos relacionando depressão e linguística, como mencionado anteriormente.

Os métodos de classificação e regressão mais utilizados foram SVM, LR, modelos Bayesianos e métodos de árvores. Além disso, para trabalhos que utilizaram vetorização por *embeddings*, CNNs foram bastante adotadas.

## ACKNOWLEDGMENTS

O presente trabalho foi parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## REFERÊNCIAS

- [1] P. Abed-Esfahani, D. Howard, M. Maslej, S. Patel, V. Mann, S. Goegan, and L. French. 2019. Transfer learning for depression: Early detection and severity prediction from social media postings. *CEUR Workshop Proceedings* 2380.
- [2] Mohammed Al-Mosaiwi and Tom Johnstone. 2018. 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.
- [3] N. S. Alghamdi. 2019. Monitoring Mental Health Using Smart Devices with Text Analytical Tool. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, Paris, França, 2046–2051.
- [4] S. Alghowinem. 2013. From joyous to clinically depressed: Mood detection using multimodal analysis of a person's appearance and speech. In *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*. Geneva, 648–653.
- [5] T. Alhanai, M. Ghassemi, and J. Glass. 2018. Detecting depression with audio/text sequence modeling of interviews. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Vol. 2018-September. International Speech Communication Association, 1716–1720.
- [6] H. Almeida, A. Briand, and M.-J. Meurs. 2017. Detecting early risk of depression from social media user-generated content. In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [7] M. Asgari, I. Shafra, and L. B. Sheeber. 2014. Inferring clinical depression from speech and spoken utterances. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, Reims, France, 1–5.
- [8] A.T. Beck, R.A. Steer, and G.K. Brown. 1996. *Manual for the Beck Depression Inventory-II*. Psychological Corporation, San Antonio, EUA.
- [9] Aaron T Beck, Calvin H Ward, Mock Mendelson, Jeremiah Mock, and John Erbaugh. 1961. An inventory for measuring depression. *Archives of general psychiatry* 4, 6 (1961), 561–571.
- [10] M. Birjali, A. Beni-Hssane, and M. Erritali. 2017. A method proposed for estimating depressed feeling tendencies of social media users utilizing their data. *Advances in Intelligent Systems and Computing* 552 (2017), 413–420.
- [11] Margaret M Bradley and Peter J Lang. 1999. *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report. Citeseer.
- [12] S.G. Burdisso, M. Errecalde, and M. Montes-y Gómez. 2019. A text classification framework for simple and effective early depression detection over social media streams. *Expert Systems with Applications* 133 (2019), 182–197.
- [13] S.G. Burdisso, M. Errecalde, and M. Montes-Y-Gómez. 2019. UNSL at Erisk 2019: A Unified Approach for Anorexia, Self-harm and Depression Detection in Social Media. In *CEUR Workshop Proceedings*, Vol. 2380. CEUR-WS.
- [14] T. Chomutare, E. Årsand, and G. Hartvigsen. 2015. Mining Symptoms of Severe Mood Disorders in Large Internet Communities. In *2015 IEEE 28th Int. Symposium on Computer-Based Medical Systems*. IEEE, Sao Carlos, Brazil, 214–219.
- [15] J. Correia, B. Raj, and I. Trancoso. 2018. Querying Depression Vlogs. In *2018 IEEE Spoken Language Technology Workshop (SLT)*. IEEE, Athens, Greece, 987–993.
- [16] J. Correia, B. Raj, I. Trancoso, and F. Teixeira. 2018. Mining multimodal repositories for speech affecting diseases. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Vol. 2018-September. International Speech Communication Association, 2963–2967.
- [17] Ting Dang, Brian Stasak, Zhaocheng Huang, Sadari Jayawardena, Mia Atcheson, Munawar Hayat, Phu Le, Vidhyasaharan Sethu, Roland Goecke, and Julien Epps. 2017. Investigating Word Affect Features and Fusion of Probabilistic Predictions Incorporating Uncertainty in AVEC 2017. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 27–35.
- [18] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social Media As a Measurement Tool of Depression in Populations. In *Proceedings of the 5th Annual ACM Web Science Conference* (Paris, France) (WebSci '13). ACM, New York, NY, USA, 47–56.
- [19] Munmun De Choudhury, Scott Counts, Eric J. Horvitz, and Aaron Hoff. 2014. Characterizing and Predicting Postpartum Depression from Shared Facebook Data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & #38; Social Computing* (Baltimore, Maryland, USA) (CSCW '14). ACM, New York, NY, USA, 626–638.
- [20] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz. 2013. Predicting depression via social media. In *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*. AAAI press, Cambridge, MA, 128–137.
- [21] F. Delahunty, I.D. Wood, and M. Arcan. 2018. First insights on a passive major depressive disorder prediction system with incorporated conversational chatbot. In *CEUR Workshop Proceedings*, Vol. 2259. CEUR-WS, 327–338.
- [22] M. Deshpande and V. Rao. 2017. Depression detection using emotion artificial intelligence. In *2017 International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, Palladam, India, 858–862.
- [23] Y. Fang et al. 2014. A mental disorder early warning approach by observing depression symptom in social diary. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, San Diego, CA, USA, 2060–2065.

- [24] A.A. Fariás-Anzaldúa, M. Montes-Y-Gómez, A. Pastor López-Monroy, and L.C. González-Gurrola. 2017. UACH-INAOE participation at eRisk2017. In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [25] I. Fatima, B.U.D. Abbasi, S. Khan, M. Al-Saeed, H.F. Ahmad, and R. Mumtaz. 2019. Prediction of postpartum depression using machine learning techniques from social media text. *Expert Systems* 36, 4 (2019).
- [26] D.G. Funez, M.J. Garcíarena Ucelay, M.P. Villegas, S.G. Burdisso, L.C. Cagnina, M. Montes-Y-Gómez, and M.L. Errecalde. 2018. UNSL's participation at eRisk 2018 Lab. *CEUR Workshop Proceedings* 2125.
- [27] Yuan Gong and Christian Poellabauer. 2017. Topic Modeling Based Multi-modal Depression Detection. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 69–76.
- [28] Jonathan Gratch, Ron Artstein, Gale Lucas, Giota Stratou, Stefan Scherer, Angela Nazarian, Rachel Wood, Jill Boberg, David DeVault, Stacy Marsella, David Traum, Skip Rizzo, and Louis-Philippe Morency. 2014. The Distress Analysis Interview Corpus of human and computer interviews. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. European Languages Resources Association (ELRA), Reykjavik, Iceland, 3123–3128.
- [29] Rahul Gupta, Nikolaos Malandrakis, Bo Xiao, Tanaya Guha, Maarten Van Segbroeck, Matthew Black, Alexandros Potamianos, and Shrikanth Narayanan. 2014. Multimodal Prediction of Affective Dimensions and Depression in Human-Computer Interactions. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (Orlando, Florida, USA) (AVEC '14). ACM, New York, NY, USA, 33–40.
- [30] M. Hiraga. 2017. Predicting depression for Japanese blog text. In *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Student Research Workshop*. Association for Computational Linguistics (ACL), 107–113.
- [31] W.L. Jarrold, B. Peintner, E. Yeh, R. Krasnow, H.S. Javitz, and G.E. Swan. 2010. Language analytics for assessing brain health: Cognitive impairment, depression and pre-symptomatic Alzheimer's disease. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6334 LNAI (2010), 299–307.
- [32] H. Jiang, B. Hu, Z. Liu, G. Wang, L. Zhang, X. Li, and H. Kang. 2018. Detecting Depression Using an Ensemble Logistic Regression Model Based on Multiple Speech Features. *Computational and mathematical methods in medicine* 2018 (2018), 6508319.
- [33] K. Kang, C. Yoon, and E. Y. Kim. 2016. Identifying depressive users in Twitter using multimodal analysis. In *2016 International Conference on Big Data and Smart Computing (BigComp)*. IEEE, Hong Kong, China, 231–238.
- [34] C. Karmen, R.C. Hsiung, and T. Wetter. 2015. Screening internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. *Computer Methods and Programs in Biomedicine* 120, 1 (2015), 27–36.
- [35] K. Katchapakin et al. 2018. Facebook Social Media for Depression Detection in the Thai Community. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, Nakhonpathom, Thailand, 1–6.
- [36] G. Kiss, M.G. Tulics, D. Sztahó, A. Esposito, and K. Vicsi. 2016. Language independent detection possibilities of depression by speech. *Smart Innovation, Systems and Technologies* 48 (2016), 103–114.
- [37] Barbara Kitchenham. 2004. Procedures for performing systematic reviews. *Keele, UK, Keele University* 33, 2004 (2004), 1–26.
- [38] Kurt Kroenke and Robert L. Spitzer. 2002. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals* 32, 9 (2002), 509–515.
- [39] Kurt Kroenke, Tara W. Strine, Robert L. Spitzer, Janet B.W. Williams, Joyce T. Berry, and Ali H. Mokdad. 2009. The PHQ-8 as a measure of current depression in the general population. *Journal of affective disorders* 114, 1–3 (2009), 163–173.
- [40] G. Lam, H. Dongyan, and W. Lin. 2019. Context-aware Deep Learning for Multi-modal Depression Detection. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Brighton, UK, 3946–3950.
- [41] Quoc Le and Tomas Mikolov. 2014. Distributed Representations of Sentences and Documents. In *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32* (Beijing, China) (ICML '14). JMLR.org, II–1188–II–1196.
- [42] P. Lopez-Otero, L. Docio-Fernandez, A. Abad, and C. Garcia-Mateo. 2017. Depression detection using automatic transcriptions of de-identified speech. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Vol. 2017-August. International Speech Communication Association, 3157–3161.
- [43] D.E. Losada and F. Crestani. 2016. A test collection for research on depression and language use. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9822 LNCS (2016), 28–39.
- [44] David E Losada, Fabio Crestani, and Javier Parapar. 2017. CLEF 2017 eRisk Overview: Early Risk Prediction on the Internet: Experimental Foundations.. In *CLEF (Working Notes)*.
- [45] David E Losada, Fabio Crestani, and Javier Parapar. 2018. Overview of eRisk: early risk prediction on the internet. In *International Conference of the Cross-Language Evaluation Forum for European Languages*. Springer, 343–361.
- [46] David E Losada, Fabio Crestani, and Javier Parapar. 2019. Overview of eRisk 2019: Early Risk Prediction on the Internet. In *International Conference of the Cross-Language Evaluation Forum for European Languages*. Springer, 340–357.
- [47] T. Mikolov et al. 2013. Efficient Estimation of Word Representations in Vector Space. In *1st International Conference on Learning Representations*. Grimm Brothers Co., Scottsdale, EUA.
- [48] M.R. Morales, S. Scherer, and R. Levitan. 2017. OpenMM: An open-source multimodal feature extraction tool. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Vol. 2017-August. International Speech Communication Association, 3354–3358.
- [49] M. R. Morales and R. Levitan. 2016. Speech vs. text: A comparative analysis of features for depression detection systems. In *2016 IEEE Spoken Language Technology Workshop (SLT)*. IEEE, San Diego, USA, 136–143.
- [50] T. Nakamura, K. Kubo, Y. Usuda, and E. Aramaki. 2014. Defining patients with depressive disorder by using textual information. In *AAAI Spring Symposium - Technical Report*, Vol. SS-14-01. AI Access Foundation, Palo Alto, CA, 39–44.
- [51] NATIONAL INSTITUTE OF MENTAL HEALTH. 2015. Depression. *Government Printing Office* (2015).
- [52] T. Nguyen et al. 2014. Affective and Content Analysis of Online Depression Communities. *IEEE Transactions on Affective Computing* 5, 3 (July 2014), 217–226.
- [53] T. Nguyen, B. O'Dea, M. Larsen, D. Phung, S. Venkatesh, and H. Christensen. 2015. Differentiating sub-groups of online depression-related communities using textual cues. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9419 (2015), 216–224.
- [54] T. Nguyen, B. O'Dea, M. Larsen, D. Phung, S. Venkatesh, and H. Christensen. 2017. Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimedia Tools and Applications* 76, 8 (2017), 10653–10676.
- [55] ORGANIZAÇÃO MUNDIAL DE SAÚDE. 2017. Depression and other common mental disorders: global health estimates.
- [56] Z. Orhan, M. Mercan, and M.K. Gökçöl. 2020. A new digital mental health system infrastructure for diagnosis of psychiatric disorders and patient follow-up by text analysis in turkish. In *IFMBE Proceedings*, Vol. 73. Springer Verlag, 395–402.
- [57] R.M. Ortega-Mendoza, A.P. López-Monroy, A. Franco-Arcega, and M. Montes-Y-Gómez. 2018. PEIMEX at eRisk2018: Emphasizing personal information for depression and anorexia detection. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.
- [58] I. Oyong, E. Utami, and E. T. Luthfi. 2018. Natural Language Processing and Lexical Approach for Depression Symptoms Screening of Indonesian Twitter User. In *2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE)*. IEEE, Kuta, Indonesia, 359–364.
- [59] Anastasia Pampouchidou, Olympia Simantiraki, Amir Fazlollahi, Matthew Pedaditis, Dimitris Manousos, Alexandros Roniotis, Georgios Giannakakis, Fabrice Meriaudeau, Panagiotis Simos, Kostas Marias, Fan Yang, and Manolis Tsiknakis. 2016. Depression Assessment by Fusing High and Low Level Features from Audio, Video, and Text. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge* (Amsterdam, The Netherlands) (AVEC '16). ACM, New York, NY, USA, 27–34.
- [60] S. Paul, J.S. Kalyani, and T. Basu. 2018. Early detection of signs of anorexia and depression over social media using effective machine learning frameworks. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.
- [61] Ma. Paula Villegas, D.G. Funez, Ma. José Garcíarena Ucelay, L.C. Cagnina, and M.L. Errecalde. 2017. LIDIC - UNSL's participation at eRisk 2017: Pilot task on early detection of depression: Notebook for eRisk at CLEF 2017. In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [62] Z. Peng, Q. Hu, and J. Dang. 2019. Multi-kernel SVM based depression recognition using social media data. *International Journal of Machine Learning and Cybernetics* 10, 1 (2019), 43–57.
- [63] James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. 2003. Psychological aspects of natural language use: our words, our selves. *Annual review of psychology* 54, 1 (2003), 547–577.
- [64] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation. In *Empirical Methods in Natural Language Processing (EMNLP)*. ACL, Doha, Qatar, 1532–1543.
- [65] Inna Pirina and Çağrı Çöltekin. 2018. Identifying Depression on Reddit: The Effect of Training Data. 9–12.
- [66] W. Ragheb, B. Moulahi, J. Azé, S. Bringay, and M. Servajean. 2018. Temporal mood variation: At the CLEF eRisk-2018 tasks for early risk detection on the internet. *CEUR Workshop Proceedings* 2125.
- [67] F. Ramiaandrisoa, J. Mothe, F. Benamara, and V. Moriceau. 2018. IRIT at e-Risk 2018. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.

- [68] D. Ramírez-Cifuentes and A. Freire. 2018. UPF's participation at the CLEF eRisk 2018: Early risk prediction on the internet. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.
- [69] E.-M. Rathner, J. Djamali, Y. Terhorst, B. Schuller, N. Cummins, G. Salamon, C. Hunger-Schoppe, and H. Baumeister. 2018. How did you like 2017? Detection of language markers of depression and narcissism in personal narratives. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Vol. 2018-September. International Speech Communication Association, 3388–3392.
- [70] Denise Razzouk. 2016. Capital mental, custos indiretos e saúde mental. In *Saúde mental e trabalho*, Quirino C. Razzouk D, Lima M (Ed.). Conselho Regional de Medicina do Estado de São Paulo, São Paulo, Brasil, Chapter 3, 63–71.
- [71] P. Resnik, A. Garron, and R. Resnik. 2013. Using topic modeling to improve prediction of neuroticism and depression in college students. In *EMNLP 2013 - 2013 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*. Association for Computational Linguistics (ACL), 1348–1353.
- [72] Fabien Ringeval, Björn Schuller, Michel Valstar, Jonathan Gratch, Roddy Cowie, Stefan Scherer, Sharon Mozgai, Nicholas Cummins, Maximilian Schmitt, and Maja Pantic. 2017. AVEC 2017: Real-life Depression, and Affect Recognition Workshop and Challenge. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 3–9.
- [73] Ray H. Rosenman, Meyer Friedman, Reuben Straus, Moses Wurm, Robert Kostichuk, Wilfrid Hahn, and Nicholas T. Werthessen. 1964. A Predictive Study of Coronary Heart Disease: The Western Collaborative Group Study. *JAMA* 189, 1 (07 1964), 15–22. <https://doi.org/10.1001/jama.1964.03070010021004>
- [74] Stephanie Rude, Eva-Maria Gortner, and James Pennebaker. 2004. Language Use of Depressed and Depression-Vulnerable College Students. *Cognition & Emotion* 18 (2004), 1121–1133.
- [75] F. Sadeque, D. Xu, and S. Bethard. 2017. Uarizona at the CLEF eRisk 2017 pilot task: Linear and recurrent models for early depression detection. In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [76] Farig Sadeque, Dongfang Xu, and Steven Bethard. 2018. Measuring the Latency of Depression Detection in Social Media. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (Marina Del Rey, CA, USA) (WSDM '18). ACM, New York, NY, USA, 495–503.
- [77] A. Samareh, Y. Jin, Z. Wang, X. Chang, and S. Huang. 2018. Detect depression from communication: how computer vision, signal processing, and sentiment analysis join forces. *IJSE Transactions on Healthcare Systems Engineering* 8, 3 (2018), 196–208.
- [78] M. Stankevich, V. Isakov, D. Devyatkin, and I. Smirnov. 2018. Feature engineering for depression detection in social media. In *ICPRAM 2018 - Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*, Vol. 2018-January. SciTePress, 426–431.
- [79] E. A. Stepanov et al. 2018. Depression Severity Estimation from Multiple Modalities. In *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE, Ostrava, Czech Republic, 1–6.
- [80] R. Sugandhi and A. Mahajan. 2017. A semantic network approach to affect analysis: A case study on depression. In *Proceedings - 1st International Conference on Intelligent Systems and Information Management, ICISIM 2017*, Vol. 2017-January. Institute of Electrical and Electronics Engineers Inc., 255–266.
- [81] Bo Sun, Yinghui Zhang, Jun He, Lejun Yu, Qihua Xu, Dongliang Li, and Zhaoying Wang. 2017. A Random Forest Regression Method With Selected-Text Feature For Depression Assessment. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 61–68.
- [82] M. M. Tadesse et al. 2019. Detection of Depression-Related Posts in Reddit Social Media Forum. *IEEE Access* 7 (2019), 44883–44893.
- [83] Yla R. Tausczik and James W. Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29, 1 (2010), 24–54.
- [84] M. Tlachac, E. Toto, and E. Rundensteiner. 2019. You're Making Me Depressed: Leveraging Texts from Contact Subsets to Predict Depression. In *2019 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*. IEEE, Chicago, EUA, 1–4.
- [85] I. Trancoso, J. Correia, F. Teixeira, B. Raj, and A. Abad. 2018. Speech analytics for medical applications. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11107 LNAI (2018), 26–37.
- [86] A. Trifan and J.L. Oliveira. 2019. Bioinfo@uAVR at ERIsk 2019: Delving into social media texts for the early detection of mental and food disorders. In *CEUR Workshop Proceedings*, Vol. 2380. CEUR-WS.
- [87] Raluca Trifu, Bogdan Nemes, Carolina Bodea-Hațegan, and Doina Cozman. 2017. Linguistic indicators of language in major depressive disorder (MDD): An evidence based research. *Journal of Evidence-Based Psychotherapies* 17 (03 2017), 105–128. <https://doi.org/10.24193/jebp.2017.1.7>
- [88] M. Troztek, S. Koitka, and C.M. Friedrich. 2017. Linguistic metadata augmented classifiers at the CLEF 2017 task for early detection of depression: FHDO Biomedical Computer Science Group (BCSG). In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [89] M. Troztek, S. Koitka, and C.M. Friedrich. 2018. Word embeddings and linguistic metadata at the CLEF 2018 tasks for early detection of depression and anorexia. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.
- [90] M. Troztek, S. Koitka, and C. M. Friedrich. 2019. Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences. *IEEE Transactions on Knowledge and Data Engineering* (2019), 1–1.
- [91] C. Tung and W. Lu. 2016. Analyzing depression tendency of web posts using an event-driven depression tendency warning model. *Artificial Intelligence in Medicine* 66 (2016), 53–62.
- [92] Michel Valstar, Jonathan Gratch, Björn Schuller, Fabien Ringeval, Denis Lalanne, Mercedes Torres Torres, Stefan Scherer, Giota Stratou, Roddy Cowie, and Maja Pantic. 2016. AVEC 2016: Depression, mood, and emotion recognition workshop and challenge. In *Proceedings of the 6th international workshop on audio/visual emotion challenge*. ACM, 3–10.
- [93] Michel Valstar, Björn Schuller, Kirsty Smith, Timur Almaev, Florian Eyben, Jarek Krajewski, Roddy Cowie, and Maja Pantic. 2014. AVEC 2014: 3D Dimensional Affect and Depression Recognition Challenge. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (Orlando, Florida, USA) (AVEC '14). ACM, New York, NY, USA, 3–10.
- [94] P. Van Rijen, D. Teodoro, N. Naderi, L. Mottin, J. Knafou, M. Jeffryes, and P. Ruch. 2019. A data-driven approach for measuring the severity of the signs of depression using reddit posts. In *CEUR Workshop Proceedings*, Vol. 2380. CEUR-WS.
- [95] E. Villatoro-Tello, G. Ramírez-De-la rosa, and H. Jiménez-Salazar. 2017. UAM's participation at CLEF eRisk 2017 task: Towards modelling depressed bloggers. In *CEUR Workshop Proceedings*, Vol. 1866. CEUR-WS.
- [96] Y.-T. Wang, H.-H. Huang, and H.-H. Chen. 2018. A neural network approach to early risk detection of depression and anorexia on social media text. In *CEUR Workshop Proceedings*, Vol. 2125. CEUR-WS.
- [97] Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45, 4 (01 Dec 2013), 1191–1207.
- [98] James R. Williamson, Elizabeth Godoy, Miriam Cha, Adrienne Schwarzentruher, Pooya Khorrami, Youngjune Gwon, Hsiang-Tsung Kung, Charlie Dagli, and Thomas F. Quatieri. 2016. Detecting Depression Using Vocal, Facial and Semantic Communication Cues. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge* (Amsterdam, The Netherlands) (AVEC '16). ACM, New York, NY, USA, 11–18.
- [99] PinHua Wu, JiaLing Koh, and Arbee L. P. Chen. 2019. Event Detection for Exploring Emotional Upheavals of Depressive People. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing* (Limassol, Cyprus) (SAC '19). ACM, New York, NY, USA, 2086–2095.
- [100] L. Yang, D. Jiang, and H. Sahli. 2018. Integrating Deep and Shallow Models for Multi-Modal Depression Analysis – Hybrid Architectures. *IEEE Transactions on Affective Computing* (2018), 1–1.
- [101] Le Yang, Dongmei Jiang, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke, and Hichem Sahli. 2017. Multimodal Measurement of Depression Using Deep Learning Models. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 53–59.
- [102] Le Yang, Hichem Sahli, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke, and Dongmei Jiang. 2017. Hybrid Depression Classification and Estimation from Audio Video and Text Information. In *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge* (Mountain View, California, USA) (AVEC '17). ACM, New York, NY, USA, 45–51.
- [103] T. Yang et al. 2016. Detection of mood disorder using speech emotion profiles and LSTM. In *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. IEEE, Tianjin, China, 1–5.
- [104] Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (Sydney, Australia) (ASONAM '17). ACM, New York, NY, USA, 1191–1198.
- [105] N.F.A. Yusof, C. Lin, and F. Guerin. 2018. Assessing the effectiveness of affective lexicons for depression classification. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10859 LNCS (2018), 65–69.
- [106] X. Zhao, S. Lin, and Z. Huang. 2018. Text classification of micro-blog's "tree hole" based on convolutional neural network. In *ACM International Conference Proceeding Series*. Association for Computing Machinery.
- [107] K.S. Ziemer and G. Korkmaz. 2017. Using text to predict psychological and physical health: A comparison of human raters and computerized text analysis. *Computers in Human Behavior* 76 (2017), 122–127.