



UNIVERSIDADE
ESTADUAL DE LONDRINA

JOSÉ LUIZ VILLELA MARCONDES MIONI

**PROCESSAMENTO DA LÍNGUA PORTUGUESA NA
DETECÇÃO DE TOXICIDADE NA REDE SOCIAL TWITTER**

Londrina
2023

JOSÉ LUIZ VILLELA MARCONDES MIONI

**PROCESSAMENTO DA LÍNGUA PORTUGUESA NA
DETECÇÃO DE TOXICIDADE NA REDE SOCIAL TWITTER**

Dissertação apresentada ao Curso de Mestrado em
Ciência da Computação do Departamento de
Computação da Universidade Estadual de Londrina,
como requisito parcial para a obtenção do título de
Mestre em Ciência da Computação.

Orientadora: Profa. Dra. Cinthyan Renata Sachs
Camerlengo de Barbosa

Londrina
2023

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

M669p Mioni, José Luiz Villela Marcondes.
Processamento da língua portuguesa na detecção de toxicidade na rede social twitter / José Luiz Villela Marcondes Mioni. - Londrina, 2023.
82 f. : il.

Orientador: Cinthyan Renata Sachs Camerlengo de Barbosa.
Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ciência da Computação, 2023.
Inclui bibliografia.

1. Processamento de Linguagem Natural - Tese. 2. Análise Pragmática - Tese. 3. Cyberbullying - Tese. I. de Barbosa, Cinthyan Renata Sachs Camerlengo. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU 519

JOSÉ LUIZ VILLELA MARCONDES MIONI

**PROCESSAMENTO DA LÍNGUA PORTUGUESA NA
DETECÇÃO DE TOXICIDADE NA REDE SOCIAL TWITTER**

Dissertação apresentada ao Curso de Mestrado em
Ciência da Computação do Departamento de
Computação da Universidade Estadual de Londrina,
como requisito parcial para a obtenção do título de
Mestre em Ciência da Computação.

BANCA EXAMINADORA

Orientadora: Prof. Dra. Cinthyan Renata Sachs
Camerlengo de Barbosa
Universidade Estadual de Londrina – UEL

Prof. Dr. Alan Salvany Felinto
Universidade Estadual de Londrina – UEL

Prof. Dr. Wagner Ferreira Lima
Universidade Estadual de Londrina – UEL

Londrina, 27 de março de 2023.

Este trabalho é dedicado a todos que escolhem apreciar a jornada, ao invés do destino.

AGRADECIMENTOS

Agradeço a Deus por permitir que eu tenha a oportunidade de trabalhar neste projeto e, principalmente, pelo privilégio de atravessar os recentes anos de pandemia com saúde.

Agradeço aos meus pais, José Luiz e Maria Regina, por amorosamente tirarem de si as peças necessárias para construir o meu caminho. Do constante sacrifício que permitiu minha educação às valiosas lições sobre amor, persistência, superação e consistência.

Agradeço às minhas irmãs Beatriz e Bianca por serem a maior inspiração que alguém precisaria para ser melhor.

Agradeço à minha esposa Rafaella por ser uma verdadeira luz de *Èarendil* em minha vida, tanto neste momento de pesquisa quanto em tantos outros.

Agradeço à minha orientadora Cinthyan Renata Sachs Camerlengo de Barbosa pela paciência, confiança e apoio. Aos membros da banca e ao departamento de Computação da UEL pela oportunidade de realizar este trabalho.

*“I imagine the fish out of water is the only fish
to have an inkling of water.”
(John Ronald Reuel Tolkien)*

MIONI, J. L. V. M. **Processamento da língua portuguesa na detecção de toxicidade na rede social Twitter**. 2023. 82 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina, 2023.

RESUMO

O crescimento de conteúdo ofensivo, comportamento tóxico e assédio de indivíduos em ambientes de internet aumentou significativamente à medida que a sociedade se torna cada vez mais digital. Diante desses eventos, diversas comunidades da web como redes sociais, portais e jogos online buscam dar suporte tecnológico na filtragem de conteúdo, proteção individual e verificação de dados. Apesar desses esforços, casos de depressão, ansiedade, outros traumas psicológicos e suicídio ainda são relatados em várias plataformas. Este trabalho descreve uma ferramenta de Processamento da Língua Portuguesa que sinalizará possível cyberbullying e discurso de ódio na rede social twitter por meio de categorização, classificação e concorrência para alcançar um índice binário de toxicidade pragmaticamente, assim como sua análise sobre um corpus alvo compilado a partir de conteúdo em redes sociais. Por meio da análise, um índice de 7,3% de toxicidade foi encontrado no corpus, reforçando a necessidade de atenção ao assunto, assim como expansões das ferramentas e pesquisas futuras na área.

Palavras-chave: processamento de linguagem natural; análise pragmática; cyberbullying.

MIONI, J. L. V. M. **Natural language processing of the portuguese language for toxicity detection on the Twitter social network.** 2023. 82 p. Dissertation (Masters in Science in Computer Science) – State University of Londrina, Londrina, 2023.

ABSTRACT

The growth of offensive content, toxic behavior, and harassment of individuals in internet environments has increased significantly as society becomes more and more digital. Given these events, several web communities such as social networks, portals and online games seek to provide technological support in content filtering, individual protection, and data verification. Despite these efforts, cases of depression, anxiety, other psychological traumas, and suicide are still reported on several platforms. This work describes a Portuguese language processing tool that will signal possible cyberbullying and hate speech in the twitter social network through categorization, classification, and concurrency to pragmatically reach a binary toxicity index, as well as analyze a target corpus compiled from social media content. According to the analysis, a 7.3% toxicity index was found in the corpus, reinforcing the need for attention to the issue as well as expansion of available tools and future research in this field.

Key words: natural language processing; pragmatic analysis; cyberbullying.

LISTA DE ILUSTRAÇÕES

Figura 1	Diferentes fases no processamento de um comando à Alexa	29
Figura 2	Estágios de análise no PLN.....	29
Figura 3	Processo de tokenização da frase alvo	30
Figura 4	Processo de Análise Léxico-morfológica para a frase alvo extraindo termos POS-Tag.....	31
Figura 5	Descrição gráfica de possível análise pragmática.....	33
Figura 6	Mapa de termos em tempo real na comunidade médica, do projeto Vozes da Pandemia.....	39
Figura 7	BERT classificando uma frase em inglês	41
Figura 8	Cyber Guardian determinando toxicidade em um nome de usuário em Língua Inglesa.....	42
Figura 9	Resultado da mesma palavra, servido pela API aberta do Cyber Guardian.....	42
Figura 10	Cyber Guardian interferindo em possível toxicidade durante uma conversa de usuários	43
Figura 11	Arquitetura do Sistema	44
Figura 12	Importando um pipeline pré-treinado no Spark NLP e inserindo uma frase alvo sem toxicidade.....	46
Figura 13	Importando o modelo pré-treinado disponibilizado pelos desenvolvedores do ToLD-BR	46
Figura 14	Separação em tokens da frase alvo sem toxicidade	46
Figura 15	Separação da frase alvo sem toxicidade em tags POS.....	47
Figura 16	Separação da frase alvo sem toxicidade em NERs	48
Figura 17	Importando um pipeline pré-treinado no Spark NLP e inserindo uma frase alvo com toxicidade	49
Figura 18	Separação em tokens da frase alvo com toxicidade.....	49
Figura 19	Separação da frase alvo com toxicidade em tags POS	49
Figura 20	Compilado dos resultados da ferramenta sobre a frase alvo sem toxicidade.....	49
Figura 21	Compilado dos resultados da ferramenta sobre a frase alvo com toxicidade.....	50
Figura 22	Toxicidade positiva para frase oriunda do corpus MasterChef Brasil.....	50

Figura 23 – Recorte do corpus MasterChef Brasil utilizado no trabalho	51
Figura 24 – Índice de toxicidade global do corpus determinado pelo software.....	51

LISTA DE TABELAS

Tabela 1 –	Ranking de cyberbullying por país em 2016 e 2018.....	36
Tabela 2 –	Entidades NER, suas descrições e exemplos	47
Tabela 3 –	Análises feitas via software e análises feitas manualmente.....	52

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
BERT	<i>Bidirectional Encoder Representations from Transformers</i>
GLCL	Gramática Livre de Contexto Lexicalizada
IA	Inteligência Artificial
LN	Linguagem Natural
NER	<i>Named Entity Recognition</i>
PLN	Processamento de Linguagem Natural
POS	<i>Part of Speech</i>
RoBERTa	<i>Robustly-optimized BERT approach</i>
ToLD-BR	<i>Toxic Language Dataset for Brazilian Portuguese</i>

SUMÁRIO

1	INTRODUÇÃO	23
1.1	PROBLEMAS E DESAFIOS	25
1.2	OBJETIVOS	26
1.3	ORGANIZAÇÃO DO TRABALHO	26
2	FUNDAMENTAÇÃO TEÓRICA	28
2.1	PROCESSAMENTO DE LINGUAGEM NATURAL	28
2.2	ANÁLISE PRAGMÁTICA	32
2.3	CYBERBULLYING NO BRASIL.....	35
2.4	CYBERBULLYING E SUA CONEXÃO COM O SUICÍDIO E MONITORAMENTO DE TOXICIDADE.....	36
2.4.1	O caso Megan Maier	37
2.4.2	O caso Hana Kimura	37
2.4.3	O caso Lucas Santos.....	38
2.4.4	Massacre de Columbine	38
3	TRABALHOS E FERRAMENTAS CORRELATAS	39
3.1	TRABALHOS PESQUISADOS.....	39
3.2	FERRAMENTAS CORRELATAS	40
3.2.1	BERT.....	40
3.2.2	RoBERTa	41
3.2.3	Cyber Guardian	42
4	MATERIAIS E MÉTODOS	44
4.1	PREPARAÇÃO E TESTES	45
4.2	CORPUS E RESULTADOS DA ANÁLISE	50
5	CONCLUSÕES E TRABALHOS FUTUROS	53
	REFERÊNCIAS	55
	ANEXOS	64

ANEXO A – CORPUS MASTERCHEF 64

TRABALHOS PUBLICADOS PELO AUTOR 81

1 INTRODUÇÃO

O mundo passa por diversas mudanças tecnológicas à medida que avanços ocorrem em diferentes frentes de desenvolvimento. Demandas de mercado, novas tecnologias, tendências e problemas com necessidade de solução são abordados por meio de novas perspectivas e ferramentas.

Sendo parte fundamental da experiência humana, a capacidade de se comunicar e conectar existe em diversas iniciativas tecnológicas e não tecnológicas ao passo que evoluímos enquanto sociedade. Seja na comunicação entre indivíduos que usam a mesma linguagem ou na interação entre esses utilizando de linguagens artificiais e não só naturais para interagir com diferentes tipos de hardware e software, a habilidade de se expressar e ser compreendido entre emissores e receptores permanece como um alicerce das evoluções tecnológicas e de qualidade de vida.

Um dos motivos da existência do Processamento de Linguagem Natural (PLN) é o intuito de permitir a comunicação entre indivíduos utilizando-se de língua natural e no auxílio de ferramentas tecnológicas na compreensão do discurso nessa por meio de software e outros dispositivos. Usado por mais de cinquenta anos, o PLN tem crescido além do campo da linguística à medida que computadores são cada vez mais utilizados na humanidade.

Segundo Jurafsky e Martin [1], o PLN tem como objetivo extrair textos em Linguagem Natural (LN) e executar tarefas relevantes, permitindo o diálogo entre homem e máquina, melhorando a comunicação humano-computador ou processamento de texto ou fala (discurso).

O PLN abrange diversos usos da habilidade de detectar e processar a palavra escrita e suas informações [2], tais como chatbots, geração de conteúdo [3] e diversos outros softwares utilizados no reconhecimento de palavras e discurso.

Os desafios que envolvem a linguística computacional são, em alguns casos, motivados pela perspectiva científica, onde alguém procura prover uma explicação para um fenômeno linguístico ou psicolinguístico [4] [5]. Em outros casos, a motivação pode ser puramente tecnológica, buscando montar um componente funcional de um sistema de linguagem natural [6].

À medida que a tecnologia avança e nos conectamos, problemas e soluções do mundo real são amplificados por meio de maior alcance e conexão proporcionados pelo

mundo digital. Um dos problemas amplificados pelas maiores capacidades de comunicação permitidas pelo mundo digital é, infelizmente, a potencialização do impacto de um problema antigo já conhecido no mundo: o *bullying*. Sua versão digital, conhecida como *cyberbullying*, infelizmente já foi apontada como um fator agravante ou até mesmo como o principal fator motivador em casos de suicídio de jovens de diferentes idades [7].

Cyberbullying, discurso de ódio e outras atividades maléficas realizadas *online* podem criar ambientes agressivos e, às vezes, perigosos [7], sendo esses relacionados a distúrbios psicológicos e, em casos mais severos, ferimentos autoinduzidos ou até mesmo atos de violência extrema como tiroteios em escolas e suicídio [8].

Porém, diversos fatores relacionados ao dia a dia da comunicação digital dificultam o emprego de diferentes tecnologias no intuito da diminuição do impacto de relações tóxicas na internet. A grande diversidade de dados demográficos de usuário assim como a variedade tecnológica entre plataformas, formatos de agressão e escassez de produções especificamente voltadas à Língua Portuguesa contribuem para a dificuldade da detecção eficaz de ameaças de *cyberbullying*. Entretanto, casos de depressão, ansiedade, outros traumas psicológicos e até mesmo suicídio muitas vezes não são relatados em várias plataformas.

Assim, este trabalho busca atuar na direção de mitigar problemas do mundo real gerados pelo *cyberbullying* ao produzir uma ferramenta que realiza e disponibiliza as diferentes etapas do Processamento de Linguagem Natural, assim como classifica binariamente a toxicidade de conteúdo extraído da rede social Twitter. Também se tem a intenção neste trabalho de trazer uma atenção especial ao tópico do dano psicológico e perda de vida humana realizado pelo *cyberbullying*, assim como contribuir com os recursos disponíveis no campo de estudo para que no futuro a prevenção de eventos trágicos motivados ou potencializados pela violência *online*, como atos de violência em escolas, casos clínicos de doenças psiquiátricas e perda de vida humana por meio do suicídio em casos de perseguição pela internet possam ser mitigados.

Este trabalho se justifica pelo devido aumento de número de conteúdos ofensivos, comportamentos tóxicos e assédio de indivíduos em ambientes da internet em diversas comunidades na web, dentre essas as de jogos e plataformas. Busca-se aqui fornecer suporte tecnológico proteção individual e filtragem de conteúdo tóxico.

1.1 Problemas e desafios

Este trabalho busca aplicar o PLN como ferramenta na solução, ainda que parcial, dos seguintes problemas:

- comportamento persecutório e violento em plataformas digitais;
- traumas e perdas de vida humana causados ou amplificados por comportamentos tóxicos;
- dificuldades relacionadas à ambiguidade do processamento de uso de profanidade e palavras com alto índice de agressão;
- escassez de conteúdos no tema utilizando a Língua Portuguesa, especialmente quando comparados a outros idiomas como, por exemplo, a Língua Inglesa;
- aparente invisibilidade de casos graves enquanto ainda estão apenas no meio digital;
- aumento do uso de meios digitais sem verificações de segurança;
- resultados crescentes da não-educação digital em um ambiente não policiado.

Dentre os principais desafios mapeados por pesquisas e corporações no intuito de frear o progresso do *cyberbullying*, mais bem descritos nas seções posteriores deste trabalho, fazem-se dignos de menção:

- a enorme quantidade de informações disponíveis e inseridas por usuários durante seu dia a dia em redes sociais que torna o policiamento não automatizado praticamente impossível em grandes comunidades digitais, como *instagram* e *facebook*;
- a natureza ilusória do uso de profanidade no português brasileiro, usado no reforço tribal ou incremento performático, assim como em polaridade contrária na exclusão ou expulsão de indivíduos de grupos sociais faz dos esforços automatizados de detecção de comportamento hostis realizáveis, porém não uma solução completamente assertiva
- a combinação dos desafios acima naturalmente cria um esforço, onde o uso de automação guiada ou auxiliada por seres humanos costuma ser a melhor abordagem possível no caminho para a solução de problemas relacionados aos *bullying* em meios digitais;
- vítimas de perseguição e *bullying* nem sempre buscam ajuda específica de guardiões legais, amigos ou autoridades na solução do problema, o que culmina na não divulgação ou análise tardia da situação, muitas vezes elucidando um cenário onde danos já foram realizados.

1.2 Objetivos

O objetivo geral deste trabalho é o Processamento da Língua Portuguesa, principalmente no que diz respeito à análise pragmática (nível de organização linguística no qual o discurso tóxico é encontrado) na rede social Twitter no intuito de apontar comportamento tóxico.

Devido ao número de conteúdos ofensivos, comportamentos tóxicos e assédio de indivíduos em ambientes da internet em diversas comunidades na *web*, entre jogos e plataformas, times de desenvolvimento e segurança buscam fornecer suporte tecnológico, proteção individual e filtragem de conteúdo tóxico. Porém, casos psiquiátricos ou até mesmo de perda de vida humana ainda são uma presença tristemente corriqueira em diversos ambientes sociais digitais, de redes sociais aos jogos *online*.

Os objetivos deste trabalho são:

- contribuir para o estudo de Processamento da Língua Portuguesa em conteúdos inseridos por usuários em plataformas como redes sociais;
- aumentar a percepção na comunidade sobre os perigos e riscos inerentes ao comportamento tóxico *online*;
- uso combinado de softwares de PLN e detecção de toxicidade na análise de conteúdo extraído de plataformas digitais;
- análise de dados reais de toxicidade coletados da web;
- aumento no potencial de prevenção futura de traumas psicológicos, violência e perda de vida humana;
- crescimento de software disponível para detectar *cyberbullying* sobre o tema na Língua Portuguesa, o qual, apesar das crescentes pesquisas nos últimos anos, ainda permanece escasso quando comparado à Língua Inglesa.

1.3 Organização do Trabalho

Este trabalho possui a seguinte estrutura: o Capítulo 2 é destinado ao referencial teórico da pesquisa e abrange conceitos de linguística computacional e *cyberbullying*.

Dentro do conteúdo do primeiro campo mencionado, são abordados tópicos como Processamento de Linguagem Natural (PLN) e suas fases quando aplicado ao Português

Brasileiro, linguagens natural e artificial, estrutura e funcionamento de aplicações de PLN, dispositivos que usam tecnologias dependentes ou conectadas ao PLN. Finalmente, a Análise Pragmática e de sentimento em textos é abordada.

A segunda parte do referencial teórico deste trabalho busca fundamentar em números e exemplos a validade e existência do problema: casos de violência ou suicídio causados ou amplificados devido aos comportamentos de grupo, agressão verbal e validação por meio do uso de profanidade ou discursos agressivos em ambientes *online*, assim como tentativas não automatizadas de prevenção são abordados.

O Capítulo 3 trata de trabalhos correlatos ao PLN e análise pragmática, conectados ou não ao *cyberbullying*, bem como linguística computacional que exploram o uso de profanidade ou discurso tóxico.

O Capítulo 4 aborda os materiais e métodos utilizados na construção da ferramenta de análise de discurso, foco deste trabalho. Tecnologias como Spark NLP, explaindocument-md, Python, BERT, ToLD-BR, Jupyter notebooks, Google Colab e modelos treinados utilizados no trabalho são revistas e definidas. Na sequência, o corpus escolhido como estudo de caso da análise de toxicidade e os resultados encontrados pela ferramenta desenvolvida neste trabalho são descritos.

Finalmente, o Capítulo 5 mostra as conclusões obtidas durante a realização deste trabalho. Os próximos passos e resultados esperados em futuras ampliações deste trabalho também são apontados.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo aborda conceitos necessários para a compreensão dos capítulos seguintes, como a definição de Processamento de Linguagem Natural, as etapas envolvidas durante sua execução, assim como definições e exemplos de análise pragmática. Também são comentados casos de *cyberbullying* e seus impactos psicológicos, essenciais na compreensão do problema-alvo deste trabalho.

2.1 Processamento de Linguagem Natural

Processamento de Linguagem Natural, conhecido como PLN, é o campo da Inteligência Artificial (IA) que pesquisa como os computadores podem ser utilizados para entender e manipular texto ou fala em linguagem natural para fazer coisas úteis. Ou seja, desenvolve modelos computacionais para analisar e gerar interações entre humanos e computadores, as quais ocorrem por meio de Linguagem Natural (LN) [9] .

Convenções que regem o uso da linguagem nas interações sociais apontam várias competências linguísticas sociais, imprescindíveis no ato de comunicação, dentro da pragmática principalmente.

As habilidades de comunicação apontam como pré-requisito para a atribuição de função comunicacional à linguagem, ou seja, possibilitam o uso da linguagem em um processo de interação [1].

Dada a importância da pragmática para a comunicação verbal, os algoritmos por trás das tarefas do PLN são, na realidade, tão complexos quanto o funcionamento da linguagem verbal. A título de exemplo do ponto mencionado, apesar de corriqueiro em certos dispositivos como o Google, a *Alexa* da Amazon é complexa em seu real funcionamento e composto de diversas etapas. Tal complexidade pode ser demonstrada pelo infográfico da Figura 1, apontando as diferentes etapas envolvidas em um simples comando de voz submetido a um dispositivo como esse.

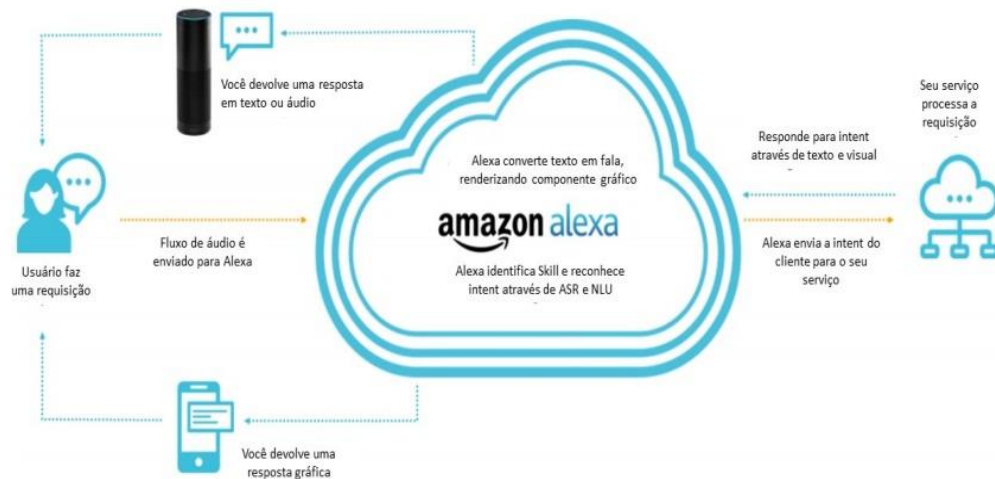


Figura 1 - Diferentes fases no processamento de um comando à Alexa [10].

Miura [11] aponta que o processo de análise de uma sentença na PLN pode resultar em mais de uma interpretação. Essa ambiguidade pode ser determinada em diferentes formas, como: morfológica, sintática ou semântica, sendo a análise pragmática, instrumental na resolução de problemas linguísticos de ambiguidade.

Na Figura 2 temos as etapas que podem ser usadas em um sistema de PLN.

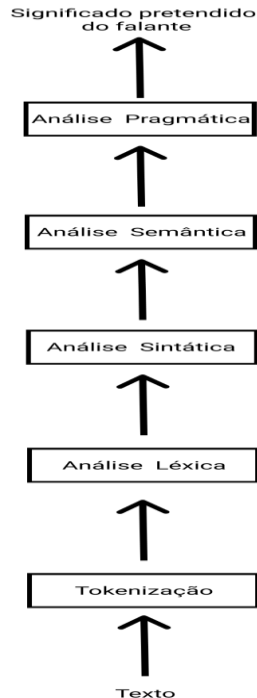


Figura 2 - Estágios de análise no PLN.

A tokenização aborda um processo de separação das palavras. Dessa forma, letras maiúsculas e minúsculas são reconhecidas, assim como palavras compostas, verificação ortográfica e quebra de caracteres.

O processo de Análise Léxica (ou léxico-morfológica) é composto da classificação léxico-morfológica que armazena o seu significado.

A tarefa de *parsing* (análise sintática) consiste da extração de informações sintáticas de uma frase representada por meio de regras gramaticais e/ou árvores sintáticas.

A análise semântica aborda o processo de análise do significado das palavras, interpretando expressões fixadas, enunciados no contexto e sentenças inteiras [12].

Já em relação à análise pragmática, o estudo fundamenta-se em reconhecimento do significado mais situado da frase, sendo essa analisada dentro de um contexto [13].

Para clarear o que significa cada fase de análise, pequenos exemplos serão elucidados.

O processo de tokenização consiste da separação de uma sentença em tokens, partes de uma sentença que futuramente serão processadas pelo analisador léxico-morfológico.

Um exemplo de uma separação de uma sentença em tokens pode ser observado na Figura 3.

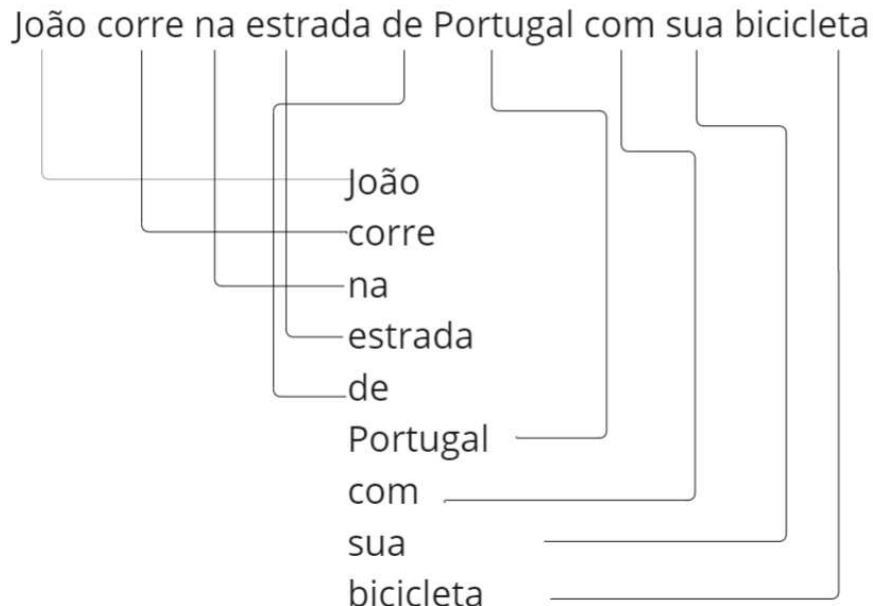


Figura 3 – Processo de tokenização da frase alvo.

O processo de Análise Léxico-morfológica faz a verificação ortográfica e classificação léxico-morfológica, categorizando palavras em diferentes tipos, como substantivo, verbo, artigo, pronome, numeral, preposição, conjunção, interjeição, artigo e

adjetivo, por exemplo. As partes podem ser identificadas segundo sua estrutura e formação em: radical, tema, vogal temática, dentre outras [14]. Um exemplo pode ser visto na Figura 4.

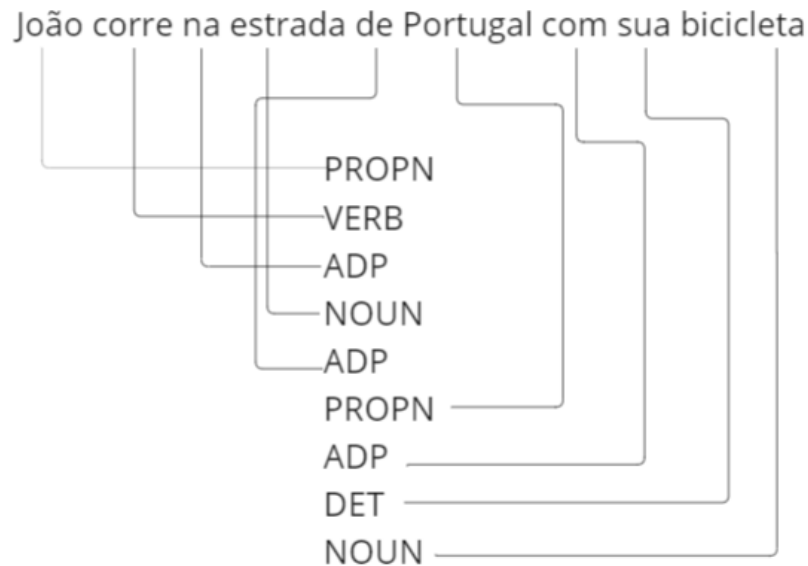


Figura 4 - Processo de Análise Léxico-Morfológica para a frase alvo extraindo termos POS-Tag.

A Análise Sintática aborda a realização de tarefas que elucidam justamente a função sintática de cada um dos tokens, observada nas frases. Dessa forma, a tarefa da Análise Sintática é extrair as informações contidas em uma frase, de acordo com suas regras gramaticais.

A finalidade do *parsing* é analisar e gerar sentenças corretas de acordo com a estrutura de cada palavra [14]. Dessa forma, a análise sintática é aquela onde uma sequência de unidades lexicais, será decomposta para determinar sua descrição estrutural de acordo com uma gramática formal [15], tratando-se de uma fase responsável por organizar o conjunto das palavras e, na sequência, aplicar regras gramaticais à sentença para reconhecer a estrutura e extrair seus significados [16]. Segundo Thanaki [17], a análise sintática é uma área de grande importância para lidar com a sintaxe da LN, visto que algumas sentenças podem possuir erros sintáticos.

Conforme visto em Jurafsky and Martin [1], podemos determinar dois formalismos que desempenham papéis críticos na análise sintática:

- Gramática de Estrutura Frasal. Composta de palavras que se agrupam e formam uma unidade e essa agrupa-se formando constituintes.

- Gramática de Dependência, aplicada justamente no processamento de fala e linguagem. É possível gerar uma estrutura de árvore por meio da derivação da frase de entrada que descreve a formação sintática da sentença analisada [1].

Gramáticas Livres de Contexto Lexicalizadas (GLCLs) também podem ser aplicadas, como a descrita por Barbosa [18]. Estudos de *parsing* para GLCL também têm sido feitos [19].

Já sobre a Análise Semântica, Oliveira e Costa [3] afirmam que a semântica pode ser dividida em dois tipos: a análise semântica léxica e a análise semântica gramatical.

Ao passo de que a análise semântica léxica busca uma representação conceitual para descrever o sentido [3], a análise semântica gramatical busca determinar o sentido da mensagem por meio de uma fórmula lógico-semântica. Esse fator, porém, pode criar alguns tipos de ambiguidade ou diferentes sentidos e significados de uma oração, um dos problemas cuja solução é buscada pela análise pragmática do discurso.

A seguir será abordada a Análise Pragmática em uma seção individual, cujo tema é primordial desta Dissertação.

2.2 Análise pragmática

A análise pragmática refere-se às convenções que regem o uso da linguagem nas interações sociais, ou seja, a um vasto leque de competências linguísticas sociais, imprescindível no ato de comunicação [20]. As habilidades de comunicação (como a capacidade de detectar ironia, contexto, significado de uma mensagem e intuito do emissor) são as que se constituem pré-requisito para a atribuição de função comunicacional à linguagem, isto é, que possibilitam o uso dessa em um processo de interação [21]. A pragmática estuda os atos da fala e os contextos nos quais eles se realizam [22].

Por exemplo, conforme ilustrado na Figura 5, dentro da mesma expressão escrita “que ótimo!” podemos ter dois significados completamente diferentes, positivo ou negativo, de acordo com o contexto em que a frase estiver inserida.

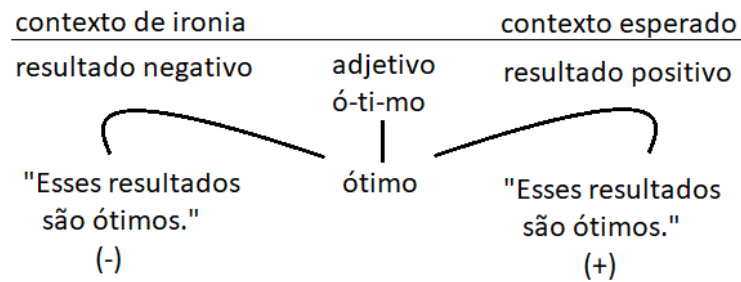


Figura 5 - Descrição gráfica de possível Análise Pragmática.

Pode-se então também determinar a análise pragmática do discurso como o ramo da linguística que analisa todo o contexto comunicativo com a finalidade de perceber as intenções do enunciado. Portanto, a pragmática trata principalmente do uso da linguagem do ponto de vista do falante, considerando suas motivações psicológicas, o impacto que sua fala causa no interlocutor e os fatores que interferem na escolha das palavras, das estruturas sintáticas, do nível de formalidade, do canal de comunicação [23].

Faz-se também importante apontar os três diferentes graus de pragmática, sendo esses de primeiro, segundo e terceiros graus.

A *pragmática de primeiro grau* é o estudo dos símbolos indexicais, ou seja, das expressões sistematicamente ambíguas. Expressões cujo sentido é tal que sua referência varia sistematicamente segundo as circunstâncias de seu uso, ou seja, segundo seu contexto de enunciação [22]. O exemplo já citado na Figura 5 trata-se de uma pragmática de primeiro grau.

A *pragmática do segundo grau* compreende o estudo da maneira como a proposição expressa se liga à frase pronunciada, lá onde, nos casos relevantes, a proposição expressa deve ser distinguida do significado literal da frase [22]. Em Armengaud [22] temos o exemplo clássico do contexto ampliado para evitar ambiguidades nas frases que, sem conter indexais, exprimem, contudo, proposições diferentes segundo o contexto: “– *Que busca Don Juan? – Ele busca uma esposa*”. Entre a interpretação “*Ele busca uma esposa*” e a outra que é “*há uma mulher casada que ele cobiça*” é o contexto, ou seja, na ocorrência concreta, aquilo que cada um sabe de Don Juan, que permite optar pela segunda. Ou seja, há a distinção fundamental entre o sentido literal e o sentido expresso.

Já a *pragmática do terceiro grau* aborda a teoria dos atos de fala. Trata-se de saber o que é que se realiza com a utilização de certas formas linguísticas. Os atos de fala são linguisticamente marcados, mas isso não basta para erradicar as ambiguidades e indicar o

que foi realmente realizado em determinada situação de comunicação [22]. Dessa forma, o uso coloquial da língua muda de formato para a mesma informação, dependendo da situação em questão. Como o caso de um padre ou pastor ao pronunciar “*Eu vos declaro marido e mulher*”, esse não está informando sobre o casamento, mas realizando-o. Caso fosse outra pessoa não autorizada a proferir o mesmo enunciado não teria valor de verdade, pois o verbo não tem força performática. Em uma frase, por exemplo, “*Pode me passar o anel?*” para uma dama de honra não quer saber se essa tem a condição física para realizar a ação, mas está reformulando um pedido que é: “*Passe-me o anel, por favor*”.

Ainda no intuito de exemplos que determinam a fase pragmática do entendimento de discurso entre interlocutores, também se faz importante pontuar exemplos de significado implícito, carregados pela própria linguagem natural em sua interpretação pragmática. Por exemplo, a frase “*É proibido fumar neste recinto*” pode, além da informação óbvia da regra, também transmitir um pedido a um indivíduo de que a regra em si seja cumprida. Ou seja, “*é proibido fumar neste recinto*” pode também significar “*por favor, pare de fumar enquanto estiver neste recinto*”.

Uma vez elucidados exemplos de análise pragmática no dia a dia, se faz necessário o uso de exemplos e contextos conectados à toxicidade e *cyberbullying*. Por exemplo, o adjetivo “*feio*” passa a, além de ter um contexto descritivo, também ter um contexto de agressividade. No entanto, a definição pragmática mostra-se ainda mais desafiadora em casos como esse, em que a palavra “*feio*”, apesar de carregar uma entrega agressiva de significado, não necessariamente ataca o interlocutor. Por exemplo, as frases “*Você é feio*”, “*É feio falar mal dos outros*” e “*O acidente foi feio*” trazem significados pragmáticos diversos usando as mesmas palavras.

Outro desafio, ainda em questão do uso de toxicidade, é a natureza ilusória do uso de profanidade no diálogo entre interlocutores ou em seus grupos, principalmente em ambientes cognitivos. Profanidades na Língua Portuguesa, como a palavra “*porra*” podem significar diferentes contextos e reações a diferentes acontecimentos. Por exemplo, a exata mesma palavra pode ser exaltada em comemoração, repúdio ou até mesmo sinalizar confusão, fazendo-se ainda mais importante, senão imprescindível, o uso do contexto na definição do sentido do uso de profanidade seja na linguagem natural ou na utilização do PLN de ferramentas como BERT [24] ou Spark [25], as quais serão abordadas futuramente.

Ao observar a ideia da análise pragmática por meio de um modelo de linguagem genérico, nota-se a ideia geral de predizer a próxima palavra em uma sentença.

Torna-se assim necessário entender o contexto com um modelo treinado para assim determiná-la com acurácia.

A título de exemplo, tomemos a sentença “*Qual comida __ sua favorita?*”. O resultado esperado de um modelo treinado é, não necessariamente a determinação da palavra “*e*” mas sim a determinação do verbo “*ser*”, devido ao entendimento do espaço vazio se localizar entre um substantivo e um pronome possessivo.

Basicamente, o modelo corre as palavras, normalmente da esquerda para à direita, tentando prever a palavra de acordo com o contexto e uma probabilidade de acerto. Esse conceito é mudado em um modelo de linguagem como o BERT (*Bidirectional Encoder Representations from Transformers*) [24], onde o mesmo pode também abordar a frase da direita para à esquerda, funcionalidade crítica à arquitetura de *Transformer* [26] (que permite o processamento de uma palavra em relação a todas as palavras em uma sentença e não necessariamente processando uma palavra por vez) [27].

Sendo assim, considerando os diferentes casos e exemplos da pragmática, assim como a natureza ilusória do uso da profanidade citadas, pode-se entender que o *cyberbullying* e sua detecção são um problema que reside no domínio da pragmática.

2.3 *Cyberbullying* no Brasil

O suicídio é um problema de saúde pública que afeta, diariamente, populações do mundo inteiro. No Brasil, os dados são preocupantes e segundo um estudo realizado em 2017 pelo Sistema de Informações de Mortalidade (SIM) [28], um órgão que faz parte do Ministério da Saúde, a taxa de suicídio entre os jovens vem aumentando desde 2002. Esses dados que apresentam um crescimento de 10% no número de mortes mostram que o problema atinge, principalmente, indivíduos na faixa etária de 15 a 29 anos [29].

Infelizmente, casos de *cyberbullying* responsáveis por traumas psicológicos severos [30], [31] inclusive a perda de vidas devido ao suicídio, podem ser encontrados no mundo todo [32] [33].

Tal ponto se faz ainda mais importante ao observamos que a nação brasileira obteve o segundo lugar no ranking mundial de casos de *bullying* nos meios digitais (perdendo apenas para a Índia) durante pesquisa global [34] onde, conforme a pesquisa, 29% dos pais ou responsáveis brasileiros consultados relataram que os filhos já foram vítimas de violência *online*. A média determinada pela pesquisa mencionada é de 17%, colocando o Brasil em

mais de 50% que a média mundial. O crescimento e números de resultado relacionados ao *cyberbullying* no mundo podem ser observados na Tabela 1.

Tabela 1 – Ranking de *cyberbullying* por país em 2016 e 2018.

Ranking	País	% de <i>cyberbullying</i> em 2018	% de <i>cyberbullying</i> em 2016
1°	Índia	37%	32%
2°	Brasil	29%	19%
3°	Estados Unidos	27%	34%
4°	Bélgica	26%	13%
5°	África do Sul	25%	25%
6°	Malásia	23%	-
7°	Suécia	23%	20%
8°	Canadá	20%	17%
9°	Turquia	19%	14%
10°	Arábia Saudita	19%	17%

A divergência entre o relato de alunos e o apresentado pelos pais se faz mais clara quando observamos uma pesquisa realizada menos de um ano após a pesquisa mencionada, onde ao invés de pais, foram entrevistados alunos [35]. Nesse caso, o número diminui para 13.6% de estudantes que afirmam já terem sofrido algum tipo de violência *online*. Já segundo uma pesquisa realizada por um *site* estadunidense, 37% dos adolescentes entre 12 e 17 anos já sofreram algum tipo de *bullying online* [36].

Esse número torna-se ainda mais preocupante quando consideramos o aumento do uso de alunos e indivíduos das redes sociais após o distanciamento social nas escolas gerado pela pandemia de COVID-19 [37].

Também se faz importante analisar a correlação de profanidade em texto [38], assim como mapear e incrementar o corpus de acordo com dados linguísticos [39] e verificar alguma técnica para separação de dialetos dentro da linguagem [40].

2.4 *Cyberbullying* e sua conexão com o suicídio e monitoramento de toxicidade

Diferentes casos de situações psiquiátricas e perda de vida humana associados ao *bullying* e *cyberbullying* já foram documentados em diversos veículos de notícia. A seguir, casos de relevância mundial são abordados.

2.4.1 O caso Megan Maier

O suicídio de Megan T. Maier, uma jovem de 13 anos, devido ao *cyberbullying* [41] [42] e perseguição no *myspace* tomou proporções globais em 2008 [43]. Megan foi perseguida e agredida mentalmente por diversos usuários, entre eles uma mulher adulta, mãe de uma de suas colegas de classe que usou uma conta falsa para interagir com a adolescente ao longo de meses.

Nesse caso específico, Megan buscou o auxílio dos pais, porém os mesmos não diagnosticaram imediatamente a severidade da situação, uma vez que o *cyberbullying* ainda não era um termo largamente conhecido. Após seu falecimento, a fundação *Megan Maier* foi criada para aumentar o conhecimento do assunto por pais, filhos e professores [44].

Observa-se que o monitoramento da rede *myspace* poderia ter contribuído para a prevenção de perda de vida humana.

2.4.2 O caso Hana Kimura

Um dos casos mais recentes conectados ao *cyberbullying* foi o suicídio da profissional de luta livre de 22 anos, Hana Kimura [45] após uma sequência de centenas de mensagens negativas em suas redes sociais depois de uma participação em um programa de TV. Esse caso aponta que, apesar da detecção do conteúdo tóxico que foi informado por usuários e até mesmo pela própria Hana, em alguns casos a detecção tardia não consegue evitar a prevenção da perda de vida humana. Sendo assim, nota-se que em alguns casos a inibição do conteúdo antes mesmo de sua chegada ao usuário final torna-se essencial na prevenção do dano psicológico e físico.

O caso teve tamanha proporção que o governo japonês aumentou a severidade das penalidades do *cyberbullying* para até um ano de prisão (previamente, o crime era apenas passível de multa). Em casos passados de suicídio, o governo japonês também tentou uma abordagem de voluntários treinados na detecção manual de *cyberbullying*. No entanto, apesar da identificação de alguns casos e prevenção, a crescente quantidade de conteúdo e informação nas redes fez com que os esforços manuais não conseguissem acompanhar a quantidade de informação gerada pelos usuários de maneira eficiente [46].

2.4.3 O caso Lucas Santos

Lucas Santos, 16 anos, se suicidou após insistentes comentários homofóbicos após a postagem de um vídeo *meme* na rede social *tiktok* [47]. Segundo sua mãe, o adolescente já tinha apresentado sinais que foram ampliados e engatilhados pela onda de comentários na rede social.

Atualmente, a mãe do jovem faz parte de uma campanha que busca legislar como crime a inserção de comentários de ódio e perseguição na internet [48].

2.4.4 Massacre de Columbine

Talvez um dos casos mais relevantes de violência estudantil divulgados no mundo [49], tendo sido adaptado em diversos formatos de mídia [50] [51], foi o massacre de Columbine nos Estados Unidos que tem suas principais raízes no *bullying* físico (o que faz sentido lógico ao se considerar que o mesmo ocorreu em 1999).

No entanto, a análise mais funda de evidências do caso aponta que o planejamento do atentado vinha sendo feito de maneira *online* [52], onde um dos realizadores dos disparos não só apontou o que gostaria de fazer, como citou nominalmente uma dos estudantes que pretendia ferir, ao longo de 12 páginas de descrições e profanidade criadas em seu *site* [52]. Dessa forma, observa-se que mesmo em casos onde a violência é iniciada em um meio não digital, a detecção de discurso de violência aumentaria as chances da prevenção da perda de vida humana.

Ao observar os casos relatados, nota-se a possível perda de oportunidade de prevenção, fosse o quadro detectado precocemente por meio do PLN e automação.

3 TRABALHOS E FERRAMENTAS CORRELATAS

A fim de entender resultados possíveis de serem atingidos com uso do PLN, assim como exemplos de ferramentas correlatas tanto ao Processamento de Linguagem Natural quanto ao combate ao *cyberbullying*, alguns exemplos são necessários.

Este capítulo aborda os resultados obtidos no mapeamento de dados relacionados à pandemia, assim como na previsão de acontecimentos usando apenas como ferramenta o PLN. Na sequência, tecnologias importantes para a execução deste trabalho e outras tentativas de solução do problema existentes são exemplificadas.

3.1 Trabalhos Pesquisados

Dentre alguns exemplos, escolhidos para elucidar os diferentes impactos e aplicações da PLN a nível nacional, notam-se contribuições de nível nacional, como a previsão dos vencedores de um famoso *reality show* brasileiro, antes do fim da votação por uma empresa brasileira de PLN [53] apenas analisando o tráfego de postagens e comentários na internet em uma rede social.

A mesma empresa realizou, também com PLN, o mapeamento em tempo real de termos e palavras-chave sobre o cenário atual do Covid-19, novamente apenas utilizando conteúdo disponível nas redes [54]. Tal resultado pode ser visualizado de maneira gráfica na Figura 6.



Figura 6 - Mapa de termos em tempo real na comunidade médica, do projeto Vozes da Pandemia [54].

Nota-se também o crescimento de trabalhos [55] [56], também no estudo de casos de *cyberbullying* em outras instâncias de redes sociais [57], como professores da rede pública de ensino [58], além de estudos de detecção de *cyberbullying* em outros idiomas, como para a Língua Polonesa [59] e a Língua Inglesa [60].

3.2 Ferramentas Correlatas

Abordagens tecnológicas utilizadas no PLN, assim como ferramentas desenvolvidas especificamente com o intuito de evitar o contato com a toxicidade *online* são descritas a seguir.

3.2.1 BERT

Entre ferramentas correlatas ao Processamento de Linguagem Natural, faz-se importante a menção do BERT (*Bidirectional Encoder Representations from Transformers*), a qual é uma técnica de Aprendizado de Máquina (AM) de código aberto para pré-treinamento de PLN [24] que possibilita o treinamento de diferentes modelos de Inteligência Artificial em algumas horas aplicando o trabalho de apenas uma placa de processamento gráfico utilizando *Transformers* [61]. Esse é um modelo de aprendizado profundo de máquina, introduzido em 2017 [26], usado primariamente no campo de PLN.

Um exemplo de classificação realizado pelo BERT na frase em inglês "*The animal didn't cross the street because it was too wide*" (*O animal não cruzou a rua, pois ela era muito larga*) pode ser observado na Figura 7.

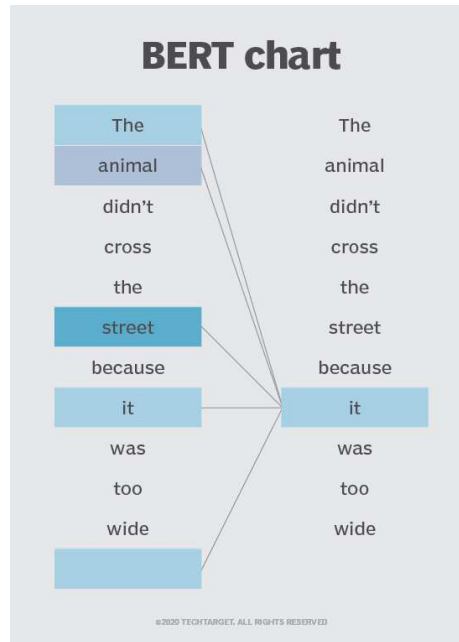


Figura 7 – BERT classificando uma frase em inglês [62].

BERT, mais especificamente um modelo de BERT pré-treinado para detecção de toxicidade chamado ToID-BR (*Toxic Language Dataset for Brazilian Portuguese*) [63] é o maior conjunto de dados disponível para análise de dados tóxicos em mídias sociais para a Língua Portuguesa [63] e o primeiro conjunto de dados com informações demográficas sobre anotadores [63]. Essa é uma das ferramentas utilizadas na montagem deste trabalho durante a pesquisa realizada e é abordada no Capítulo 4.

3.2.2 RoBERTa

Também relevante, a ferramenta RoBERTa ou *Robustly Optimized BERT approach*, fornecida pelo *Facebook*, é um modelo que adota diversas técnicas utilizadas pelo modelo BERT [64].

Essa usa a estratégia de *masking* (já trabalhada de diferentes maneiras em outros momentos por outras tecnologias [65]) da ferramenta correlata usada pelo Google, onde o sistema tenta prever seções intencionais escondidas do texto usando exemplos. Hiperparâmetros-chave foram modificados em sua produção, como um tempo muito maior de treinamento, maior magnitude de dados de análise e a remoção de seu objetivo de pré-treinamento da próxima sentença na frase, permitindo melhores resultados que o modelo base [64].

3.2.3 Cyber Guardian

Outro exemplo de tecnologia de PLN, nesse caso já orientada ao impedimento de atividades tóxicas que possam gerar *cyberbullying*, nota-se a existência do *Cyber Guardian* fornecido pela empresa Samurai AI [66]. O modelo de tecnologia disponibilizado proporciona três métodos interessantes de detecção, assim como prevenção de violência da internet.

O *Cyber Guardian* verifica e valida a criação de nomes de usuário durante a criação de contas em um *site*. Dessa forma, torna-se mais complexo e desestimula-se o intuito da criação de contas de usuário cujo intuito essencial é distribuir conteúdo agressivo *online*. A validação de um nome ofensivo em inglês (a palavra *bitch*) é descrita nas figuras abaixo, tanto em seu resultado de *frontend* na Figura 8 [67] como em seu resultado de *backend* retornado por sua API aberta na Figura 9 [67].

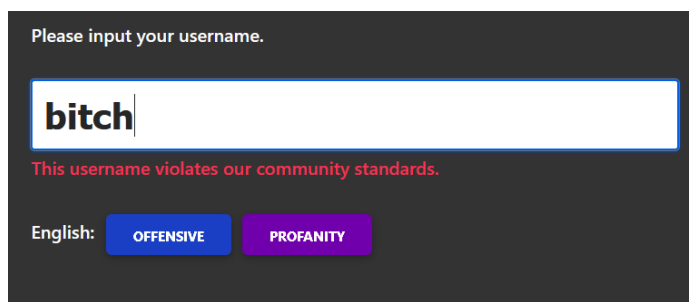


Figura 8 – *Cyber Guardian* determinando toxicidade em um nome usuário em Língua Inglesa [66].

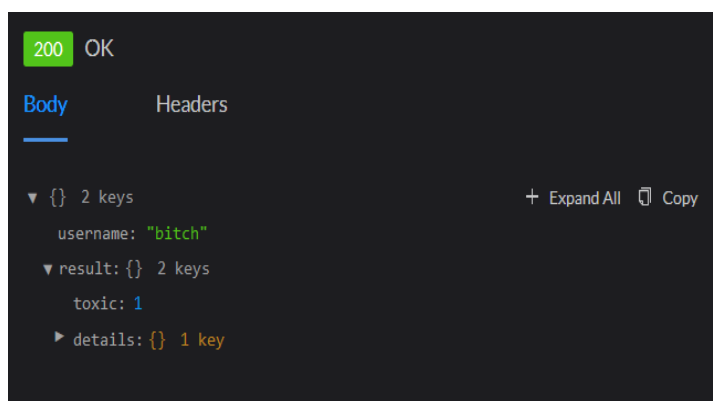


Figura 9 – Resultado da mesma palavra, servido pela API aberta do *Cyber Guardian* [67].

O agente inteligente também monitora e analisa em tempo real as mensagens enviadas por usuários na plataforma. Em momentos necessários, de acordo com o algoritmo do *Cyber Guardian*, o agente surge na conversa apontando uma possível infração às políticas de comunidade, conforme observado na Figura 10.

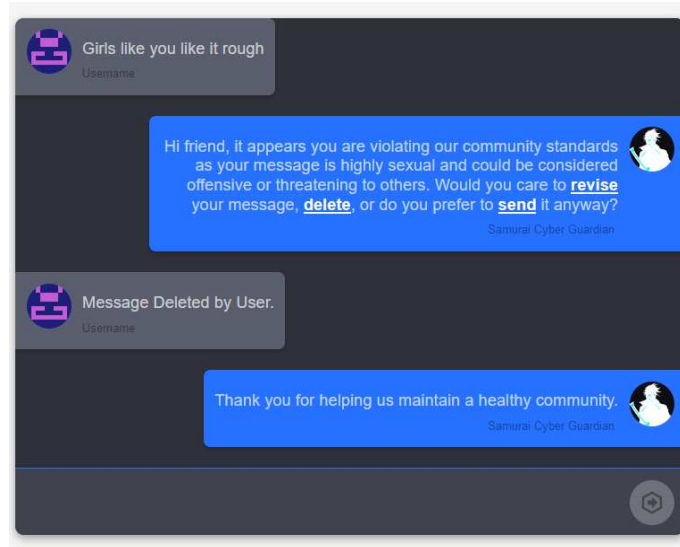


Figura 10 – *Cyber Guardian* interferindo em possível toxicidade durante uma conversa de usuários.

Também digna de nota, talvez sua função de mais impacto, o *Cyber Guardian* pode impedir o envio de uma mensagem, fazendo com que a mesma jamais seja lida pela vítima. Essa abordagem é importante, uma vez que ela permite a prevenção do dano ao invés de apenas sinalizar um usuário com possível comportamento violento. Sendo assim, o resultado é que o *cyberbullying* não chega a existir.

De uma maneira geral, a ferramenta pode conter diferentes tipos de atividades tóxicas na Língua Inglesa. Segundo a Samurai AI [66], a mera existência do *Cyber Guardian* e sua instalação em uma plataforma, muda a postura de até 78% dos usuários impactados [66]. Ao analisar as ferramentas mencionadas, é possível observar sua relação e abordagens em comum no processamento de discurso no intuito de diagnosticar algum tipo de intenção (seja em uma frase de busca ou discurso tóxico), indicador esse essencial no trabalho com PLN e na detecção de toxicidade.

4 MATERIAIS E MÉTODOS

No intuito de demonstrar a metodologia e métodos utilizados durante o trabalho, faz-se necessária a elucidação dos recursos tecnológicos utilizados, assim como os passos necessários para a preparação e execução do software.

Na Figura 11 temos a Arquitetura do Sistema para detecção de toxicidade na rede social twitter, a qual será mais bem explicitada nas seções seguintes.

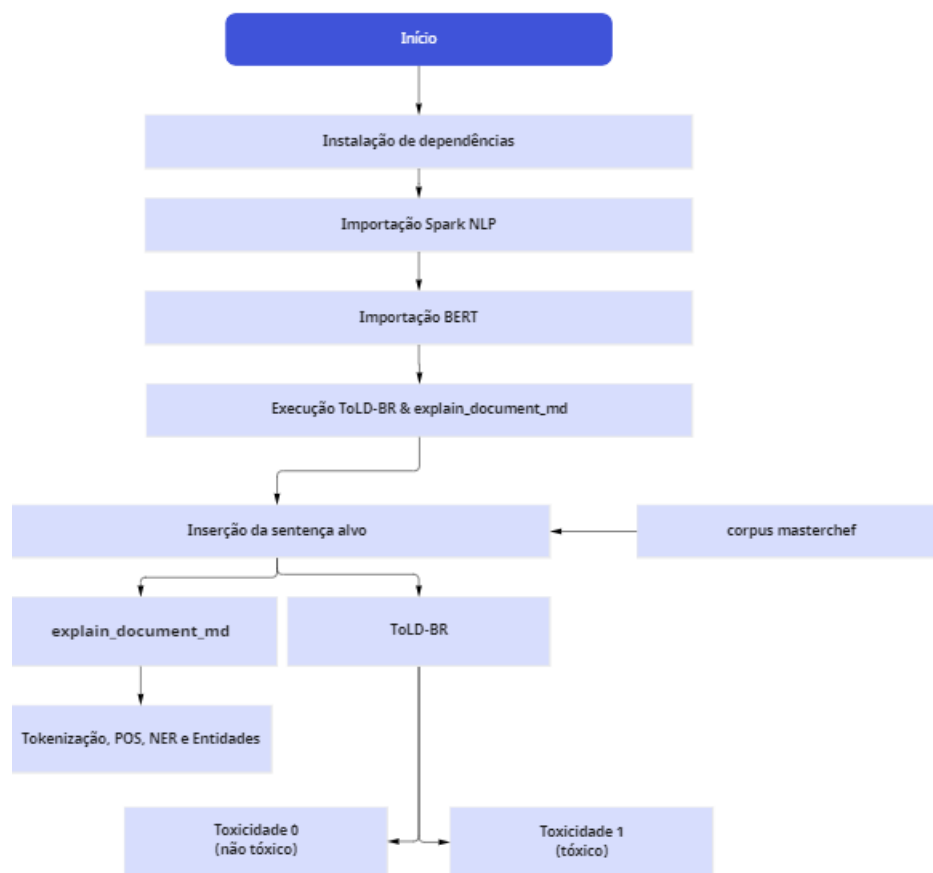


Figura 11 – Arquitetura do Sistema.

Etapas de instalação de dependências, importação e execução das tecnologias, inserção de sentença alvo e sua classificação são determinadas durante este capítulo. Comandos e tecnologias necessários para executar os testes feitos no corpus, assim como resultados de exemplos atingíveis pela ferramenta também são demonstrados.

A princípio, uma sentença de exemplo é usada na explicação dos mecanismos envolvidos na catalogação e classificação de conteúdo. Uma vez abordados os principais tópicos, as explicações e resultados obtidos terão como objeto de estudo, itens do corpus usado no trabalho.

Por fim, os resultados obtidos durante a análise do corpus, assim como análises percentuais de acurácia e explicações sobre resultados divergentes são abordados.

4.1 Preparação e testes

A ferramenta descrita neste trabalho foi elaborada para classificar frases e detectar possíveis sentenças tóxicas, podendo futuramente apontar possíveis suspeitos de abuso, perseguição e outras atividades tóxicas *online*, por meio da análise de sentenças em sessões de comentário, postagens e qualquer outro discurso escrito que esteja disponível *online* e em Língua Portuguesa. As etapas, corpus e recursos que compõem a ferramenta deste trabalho, assim como o corpus utilizado para análises preliminares são mais bem descritos a seguir.

Spark NLP [68], software estado da arte, largamente usado na indústria em diferentes atividades de Processamento de Linguagem Natural combinado ao ToLD-BR [63], um modelo pré treinado da ferramenta BERT [24] com grau de de 74% de acerto na classificação de uma sentença como tóxica, disponibilizado sob a licença *Creative Commons* na plataforma *GitHub* são os recursos essenciais na execução deste trabalho. Inicialmente, faz-se necessário a instalação e preparação do Spark NLP.

Utilizando um *pipeline* (um *pipeline* de AP é a construção de ponta a ponta que orquestra o fluxo de dados e a saída de um modelo de aprendizado de máquina. Inclui entrada de dados brutos, recursos, saídas, o modelo de AP e parâmetros do modelo e saídas de previsão [69]) de código aberto *explain document* [68], disponível no Spark NLP e na plataforma *GitHub*, podemos inserir uma frase alvo para ser analisada pela aplicação, como observada na Figura 12.

```
from sparknlp.pretrained import PretrainedPipeline
from sparknlp.base import *
```

```
pipeline = PretrainedPipeline('explain_document_md', lang = 'pt')
```

```
text = 'João corre na estrada de Portugal com sua bicicleta'
result = pipeline.annotate(text)
```

Figura 12 – Importando um *pipeline* pré-treinado no Spark NLP e inserindo uma frase alvo sem toxicidade.

Estando o Spark NLP já disponível e funcionando, faz-se necessário carregar o modelo pré-treinado do ToLD-BR para classificação binária de toxicidade, conforme demonstrado na imagem extraída do caderno Jupyter [70] da aplicação (editor executado no navegador que permite a manipulação de código Python) demonstrada na Figura 13.

```
from simpletransformers.classification import ClassificationModel
from google.colab import drive, files
import gdown
!pip install --upgrade --no-cache-dir gdown
!gdown "1HQSKxn2dhtc8JJH49e4NMxChQrIi2nje&confirm=t"
output = "toxic_bert_model.zip"
os.environ['modelpath'] = "./toxic_bert_model.zip"
!unzip "$modelpath" -d .
```

Figura 13 – Importando o modelo pré-treinado disponibilizado pelos desenvolvedores do ToLD-BR [63].

Uma vez que os dados estejam prontos para análise na aplicação, é possível realizar tarefas como tokenização, separação em diferentes partes de discurso (ou *Parts of Speech Tagging*) e indicação de Entidades Nomeadas (*Named Entity Recognition* - NERs).

A aplicação então desconstrói as sentenças em verbetes, separando “João corre na estrada de Portugal” nos diferentes tokens: 'João', 'corre', 'na', 'estrada', 'de', 'Portugal', 'com', 'sua' e 'bicicleta'. O resultado do processo de tokenização disponível é ilustrado na Figura 14.

```
[ ] result['token']
['João', 'corre', 'na', 'estrada', 'de', 'Portugal', 'com', 'sua', 'bicicleta']
```

Figura 14 - Separação em tokens da frase alvo sem toxicidade.

Após o processo de tokenização, são dados os primeiros passos da realização do processo de POS - *Part of Speech tags* (classificar todas as sentenças do texto por categorias gramaticais e a relação de dependência entre palavras), conforme ilustrado na Figura 15.

```
[ ] result['pos']
['PROPN', 'VERB', 'ADP', 'NOUN', 'ADP', 'PROPN', 'ADP', 'DET', 'NOUN']
```

Figura 15 - Separação da frase alvo sem toxicidade em *tags* POS.

Spark NLP separa a frase alvo entre substantivos (do inglês *noun*, representado na ferramenta como *NOUN*), substantivos próprios (do inglês *proper noun*, representado por *PROPN*, verbos (do inglês *verb*, representado por *VERB*), adposições (do inglês *adposition*, representado por *ADP*) e determinante (do inglês *determiner*, representado por *DET*) ao se usar POS [71].

É destacado por Ferreira [72] que ao nível de palavras, as principais *parts of speech tags* fornecem informações significativas sobre uma palavra em seu contexto. Além disso, elementos como a subjetividade, polaridade ou até mesmo a presença de sentimentos na escrita de um determinado autor podem ser classificados pela quantidade de adjetivos presentes no texto [73].

Ainda dentro da aplicação, é possível determinar NERs existentes dentro da frase alvo. Conforme demonstrado por Tripathi [74] na Tabela 2 temos NERs que podem ser classificadas da seguinte forma e foram usadas neste trabalho.

Tabela 2 – Entidades NER, suas descrições e exemplos [74].

Tipo de NER	Descrição	Exemplo
PERSON	Pessoas, reais ou fictícias	João
NORP	Grupos de nacionalidades, religiões ou entidades políticas	Partido Democrático Trabalhista
FAC	Construções, estradas, e outros pontos geográficos	Aeroporto de Guarulhos
ORG	Empreas, instituições, etc	Microsoft
GPE	Países, cidades, estados	França
LOC	Localizações não petencentes ao GPE	Europa
PRODUCT	Objetos, veículos, etc	Formula 1

EVENT	Conflitos, eventos esportivos ou históricos	Olimpíadas
WORK_OF_ART	Títulos de livros, canções, filmes	O Senhor dos Anéis
LAW	Itens que constituem leis	Lei Maria da Penha
LANGUAGE	Linguagens nomeadas	Português Brasileiro
DATE	Datas ou períodos históricos	20 de Julho de 2020
TIME	Unidades de tempo menores que 24h	Dez minutos
PERCENT	Porcentagem	Cinquenta por cento, 50%
MONEY	Valores monetários	Quinze centavos
QUANTITY	Medidas diversas	10 quilômetros
ORDINAL	Posicionamento numérico	Primeiro
CARDINAL	Numerias não cobertos por outra NER	2, quarenta, 3

O projeto de separação em entidades referenciado na Tabela 2 pode ser aplicado à frase alvo utilizada neste trabalho. O resultado dessa análise é ilustrado na Figura 16.

```
[ ] result['ner']
    ['B-PER', 'O', 'O', 'O', 'O', 'B-LOC', 'O', 'O', 'O']

[ ] result['entities']
    ['João', 'Portugal']
```

Figura 16 - Separação da frase alvo sem toxicidade em NERs.

As duas NERs encontradas na frase alvo são “João”, uma entidade de pessoa (do inglês *person*) refletida em B-PER e a entidade de local “Portugal”, apontada como B-LOC.

Para elucidar a execução da ferramenta com uma frase com certo grau de toxicidade, segue a frase alvo “*João é feio por ter vindo de Portugal*” na Figura 17.

```
text = 'João é feio por ter vindo de Portugal'
result = pipeline.annotate(text)
```

Figura 17 – Importando um *pipeline* pré-treinado no Spark NLP e inserindo uma frase alvo com toxicidade.

O processo de tokenização disponível é ilustrado na Figura 18.

```
result['token']
['João', 'é', 'feio', 'por', 'ter', 'vindo', 'de', 'Portugal']
```

Figura 18 - Separação em tokens da frase alvo com toxicidade.

Com a ferramenta é possível separar a sentença “*João é feio por ter vindo de Portugal*” nos diferentes tokens 'João', 'é', 'feio', 'por', 'ter', 'vindo', 'de' e 'Portugal'.

A seguir faz o processo de POS Tag conforme ilustrado na Figura 19.

```
result['pos']
['PROPN', 'AUX', 'ADJ', 'ADP', 'AUX', 'VERB', 'ADP', 'PROPN']
```

Figura 19 - Separação da frase alvo com toxicidade em POS *tags*.

Na sequência, a análise preditiva do ToID-BR é ativada, classificando binariamente a sentença entre tóxica (1) e não-tóxica (0). O compilado do resultado da sentença do primeiro exemplo, naturalmente não tóxica, considerando a frase alvo, é observado na Figura 20.

```
100% ██████████ 1/1 [00:00<00:00, 5.35it/s]
100% ██████████ 1/1 [00:00<00:00, 1.21it/s]
Sentença: ['João corre na estrada de Portugal com sua bicicleta']
Lematizador: ['João', 'correr', 'o', 'estrada', 'de', 'Portugal', 'com', 'sua', 'bicicleta']
Part of Speech (POS): ['PROPN', 'VERB', 'ADP', 'NOUN', 'ADP', 'PROPN', 'ADP', 'DET', 'NOUN']
Tokenizador: ['João', 'corre', 'na', 'estrada', 'de', 'Portugal', 'com', 'sua', 'bicicleta']
Named Entity Recognition (NER): ['B-PER', 'O', 'O', 'O', 'O', 'B-LOC', 'O', 'O', 'O']
Entidades: ['João', 'Portugal']
Toxicidade: [0]
```

Figura 20 – Compilado dos resultados da ferramenta sobre a frase alvo sem toxicidade.

Agora mostraremos a segunda frase mostrada que já tem toxicidade.

```
100% ██████████ 1/1 [00:00<00:00, 4.34it/s]
100% ██████████ 1/1 [00:00<00:00, 2.06it/s]
Sentença: ['João é feio por ter vindo de Portugal']
Lematizador: ['João', 'ser', 'feio', 'por', 'ter', 'vir', 'de', 'Portugal']
Part of Speech (POS): ['PROPN', 'AUX', 'ADJ', 'ADP', 'AUX', 'VERB', 'ADP', 'PROPN']
Tokenizador: ['João', 'é', 'feio', 'por', 'ter', 'vindo', 'de', 'Portugal']
Named Entity Recognition (NER): ['B-PER', 'O', 'O', 'O', 'O', 'O', 'O', 'B-LOC']
Entidades: ['João', 'Portugal']
Toxicidade: [1]
```

Figura 21 - Compilado dos resultados da ferramenta sobre a frase alvo com toxicidade.

A seguir será exibido o comportamento da ferramenta com frases reais do twitter no domínio do MasterChef para terem noção como é esse Corpus. A frase é claramente tóxica, já oriunda do corpus que será descrito na próxima seção usado para testes deste trabalho. “*Maldita, eu te odeio sua escrota #MasterChefBR*”, onde a toxicidade é apontada na Figura 22.

```
100% ██████████ 1/1 [00:00<00:00, 4.51it/s]
100% ██████████ 1/1 [00:00<00:00, 1.64it/s]
Sentença: ['Maldita,eu te odeio sua escrota #masterchefbr']
Lematizador: ['Maldita,eu', 'te', 'odiar', 'sua', 'escrota', '#masterchefbr']
Part of Speech (POS): ['PROPN', 'VERB', 'VERB', 'DET', 'NOUN', 'ADJ']
Tokenizador: ['Maldita,eu', 'te', 'odeio', 'sua', 'escrota', '#masterchefbr']
Named Entity Recognition (NER): ['B-PER', 'O', 'O', 'O', 'O', 'O']
Entidades: ['Maldita,eu']
Toxicidade: [1]
```

Figura 22 – Toxicidade positiva para frase oriunda do corpus MasterChef Brasil.

4.2 Corpus e resultados da análise

Os testes realizados neste trabalho foram realizados em um corpus, catalogado por Silva, Teixeira e Caseli [75] [76], selecionado por conter conteúdo real gerado por usuários na internet durante um momento de emoção, composto de 450 *tweets* capturados enquanto o programa competitivo MasterChef Brasil era exibido no dia 02 de maio de 2017 no canal de TV Rede Bandeirantes. Os textos foram rotulados manualmente por dois anotadores [75], com experiência em redes sociais e que acompanham o *reality* de culinária do referido programa.

Previamente separado durante segmentos do programa, o corpus usado neste trabalho é uma combinação dos 450 *tweets* do corpus em um único bloco de informação e está presente no Anexo A. Uma breve amostra do conteúdo do corpus pode ser observada na Figura 23.

```

69 #masterchefbr aderize eu te venero aaaaa
70 não pode respirar nessa cozinha hoje que geral tá com nervos à flor da pele! #masterchefbr
71 a cara da miriam hahaha raivosa #masterchefbr
72 aderize com sangue nos olhos pra sambar na cara da miriam #masterchefbr
73 miriam colando na prova descaradamente. #masterchefbr
74 miriam é o otariano do masterchef #masterchefbr
75 #masterchefbr miriam eu te odeio
76 rt @bicmuller: agora pare pegue no bumbum #masterchefbr https://t.co/gxiba0nczu
77 entrando macio eh sempre bom ne #masterchefbr
78 miriam copiando a aderize chocado com a vilã da edição. #masterchefbr
79 #masterchefbr paola pisa menos demonia ei te amoo
80 a marian cozinhando parece eu colando nas provas! imitou o prato da colega! hahahahah #masterchefbr

```

Figura 23 – Recorte do corpus MasterChef Brasil utilizado no trabalho.

Na sequência, todos os 450 itens do corpus foram formatados (conversão para letras minúsculas e remoção de *links*) e analisados no trabalho, o qual apontou um total de 33 itens de toxicidade positiva, o que corresponde a 7,3% do total. Apesar de um número relativamente pequeno é um dado que se torna alarmante ao se cogitar a considerável quantidade de conteúdos tóxicos sendo gerada pelos usuários em redes sociais, assim como o número de indivíduos expostos a esses dados diariamente.

Também se faz importante apontar que, apesar de competitivo, o programa MasterChef Brasil não é necessariamente um programa que incentivaria esse tipo de comportamento nas redes sociais, o que traz preocupação sobre resultados futuros em momentos mais delicados do país, seja em situações políticas, competitivas ou em outro cenário de atrito. Uma representação visual dos dados obtidos é determinada no gráfico da Figura 24 que mostra 7,3% de conteúdos tóxicos e 92,7% de não tóxicos.

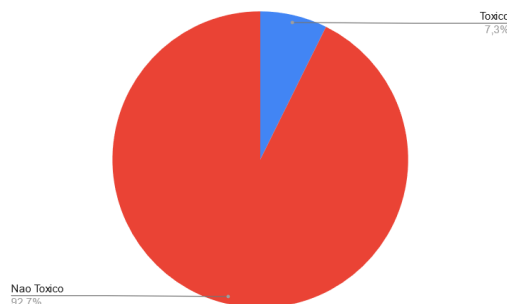


Figura 24 – Índice de toxicidade global do corpus determinado pelo software.

Após a realização da revisão do corpus pelo software, uma revisão manual foi realizada por meio da leitura individual e análise pragmática de cada um dos *tweets* no corpus, a fim de determinar de maneira manual e binária os conteúdos do corpus. Os dados obtidos estão listados na Tabela 3.

Tabela 3 – Análises feitas via software e análises feitas manualmente.

Tipo de Análise	Itens tóxicos	Itens não tóxicos
Análise de Software	33	417
Análise Manual	36	414

Faz-se então necessária a avaliação do resultado divergente. A explicação do problema e da suposta não acurácia total dos resultados encontra-se não só na Análise Pragmática como na própria habilidade humana de entendimento de contexto e na bagagem cultural usada por indivíduos brasileiros em seu dia a dia.

É possível citar o seguinte exemplo de sentença erroneamente classificada como tóxica pelo software. Durante o programa *MasterChef*, os participantes têm uma quantidade limitada de tempo para concluir o preparo de diferentes pratos de diversos tipos de culinária. A gestão desse tempo é realizada pela apresentadora do referido programa. Em um certo momento, para sinalizar o fim do tempo disponível, a apresentadora levantou as mãos, realizando um sinal de parada para os participantes. Dois usuários na rede social de *microblogs* associaram o ato a letra da música “*Pega no Bumbum*” do conjunto brasileiro “*É o Tchan do Brasil*”, nos itens listados no corpus como “*AGORA PARE PEGUE NO BUMBUM #masterchefBR*” e “*RT: AGORA PARE PEGUE NO BUMBUM #masterchefBR*”. Sem a bagagem cultural brasileira de um ser humano em seu modelo, o software então catalogou erroneamente ambas as sentenças como tóxicas.

Situações como a acima mencionada, apontam o caráter determinante não só da Análise Pragmática na detecção de conteúdo tóxico, como também a ainda latente necessidade de revisão manual dos conteúdos eventualmente apontados como perigosos por um software que analisa redes sociais.

5 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho buscou listar e definir as diferentes etapas do Processamento de Linguagem Natural da Língua Portuguesa, assim como possibilitar sua disponibilização por meio de uma ferramenta que pudesse fornecê-las a partir de uma sentença escrita (usando o Spark NLP e o *pipeline explain_document_md*) por usuários reais (coletados pelo corpus MasterChef) além de também determinar possível toxicidade (com o uso do modelo ToLD-BR).

Sendo assim, a ferramenta obtida ao final do trabalho é de grande valia no campo de estudo da PLN e também da detecção de conteúdo tóxico, uma vez que tal combinação de ferramentas, de acordo com o levantamento realizado durante a etapa de pesquisa, ainda não se encontrava disponível em trabalhos da área na Língua Portuguesa.

A ferramenta mencionada pode ser manipulada, utilizada e modificada pelo seu link do Google Colaboratory [77], o qual permite a execução de código Python por meio do navegador.

O resultado acima foi obtido pela combinação de mecanismos já existentes de lematização, tokenização e afins combinados com um modelo pré-treinado previamente disponibilizado de catálogo binário de toxicidade na Língua Portuguesa aliado ao conteúdo da Análise Pragmática.

Um *pipeline* pré-treinado foi utilizado no reconhecimento e atribuição de palavras da Língua Portuguesa em diferentes sentenças presentes em um determinado corpus.

Os resultados obtidos durante a análise do corpus podem contribuir para os estudos referentes ao *cyberbullying* e reforçar a importância desse tema, ao apontar a existência de toxicidade no diálogo e comportamento de indivíduos na internet até mesmo em momentos corriqueiros como o ato de comentar um programa de televisão nas redes sociais. Também é de resultado do trabalho a contribuição para a divulgação e conscientização do crescente e urgente problema conhecido como *cyberbullying*.

Observou-se, ao longo de todo o trabalho, que a Análise Pragmática mostra-se parte muito importante da análise de toxicidade em sentenças, sendo não só o grande fator diferencial no atingimento da análise correta, como também a diferença entre um software que apenas procura por palavras-chave e um programa que de fato pode realizar um entendimento minimamente mais próximo do atingido por um indivíduo usando de linguagem natural.

No entanto, a partir dos resultados observados, nota-se que a própria pragmática ainda é justamente o fator decisivo entre a compreensão parcial e a acurácia total de um software de PLN que busca determinar toxicidade, fazendo-se necessário ainda mais estudos nesse campo no objetivo de, em algum momento futuro, atingir-se um patamar similar ao de ser humano no entendimento do texto escrito.

Diante dos problemas descritos, a elaboração deste trabalho permite um passo na direção de contribuir com a sociedade e com o relacionamento entre diferentes pessoas no mundo *online*, utilizando técnicas de PLN na identificação de discurso violento e perseguição na internet.

Possíveis trabalhos futuros abaixo estão elencados:

- a expansão deste trabalho para uma ferramenta web que poderia ser usada em trabalhos correlatos como um recurso de análise pragmática;
- a criação de um módulo para não apenas apontar e determinar toxicidade da frase de maneira binária, como também fornecer o grau de gravidade dessa coletada dentro de uma escala. Essa expansão permitirá adicionar maior assertividade ao projeto, permitindo a priorização de casos mais graves;
- a formatação do projeto como um todo em um *plugin* de código aberto para redes sociais que seria capaz de adicionar usuários possivelmente mal intencionados ao painel de administração de grupos de moderação, permitindo assim a revisão de conteúdos ou a simples remoção de usuários agressivos em ambientes *online*.

REFERÊNCIAS

- [1] JURAFSKY, D.; MARTIN, J. H. *Speech and Language Processing: an introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 2^a. ed. New Jersey, USA: Prentice Hall, v. I, 2008. 988p
- [2] AVRON, B.; FEIGENBAUM, E. A. *Handbook of Artificial Intelligence*. 1^a. ed. Boston: Addison Wesley, 1986. 424p
- [3] OLIVEIRA, H. G.; COSTA, D.; A., P. Automatic Generation of Internet Memes from Portuguese News Headlines. *12th International Conference on the Computational Processing of the Portuguese Language*. Tomar, Portugal, Springer. 2016. p. 340-347.
- [4] SILVA, C. B. B.; PARABONI, I. Personality Recognition from Facebook Text. *13th International Conference on the Computational Processing of the Portuguese Language*. Canela, Springer. 2018. p. 83-92.
- [5] SANTOS, A. S.; BARROS J., J. D.; CAMARGO, H. A. Annotation of a Corpus of Tweets for Sentiment Analysis. *13th International Conference on the Computational Processing of the Portuguese Language*. Canela, Springer. 2018. p. 294-302.
- [6] ACL - ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *What is computational linguistics?* 2022. Disponível em: <<https://www.aclweb.org/portal/>>. Acesso em: 22 jun. 2020.
- [7] PORTELA, G. Cyberbullying e casos de suicídio aumentam entre jovens. *Agência Focruz de Notícias*, 2014. Disponível em: <agencia.fiocruz.br/cyberbullying-e-casos-de-suic%C3%ADdio-aumentam-entre-jovens>. Acesso em: 24 jun. 2020.
- [8] BAUMAN, S.; RUSSEL, B.; WALKER, J. Associations among bullying, cyberbullying, and suicide in high school students. *Journal of Adolescence*, v. 36, n. 2, april, 2013 p. 341-350.
- [9] PEREIRA, S. L. *Processamento de Linguagem Natural*. Instituto Matemática e Estatística - USP. Disponível em: <www.ime.usp.br/~slago/IA-pln.pdf>. Acesso em: 19 maio 2020.
- [10] KITAMURA, C. *O que é Alexa? (Ou Quem é Alexa)?* 2019. Disponível em: <<https://celsokitamura.com.br/o-que-e-alexa/>>. Acesso em: 03 jun. 2022.
- [11] MIURA, N. K. *Geração incremental de parsers dependentes de contexto para o português brasileiro*. Escola Politécnica. Universidade de São Paulo. Tese (Doutorado em Engenharia

Elétrica). São Paulo. 2019. 132p

- [12] GODDARD, C.; SCHALLEY, C. A. *Semantic analysis: A Practical Introduction*. 2ª. ed. Oxford, Londres: Oxford University Press, 2012. 512 p.
- [13] MULLER, D. N. *Processamento de Linguagem Natural*. Instituto de Informática. Universidade Federal do Rio Grande do Sul. 2003. Disponível em: <<https://www.inf.ufrgs.br/~danielnm/docs/pln.pdf>>. Acesso em: 30 mar. 2021.
- [14] DOMINGUES, M. L. C. S. *Abordagem para o Desenvolvimento de um Etiquetador de Alta Acurácia para o Português do Brasil*. Instituto de Tecnologia. Universidade Federal do Pará. Tese (Doutorado em Engenharia Elétrica). Belém, 2011. 140p.
- [15] LJUNGLÖF, P.; WIRÉN, M. *Syntactic parsing*. 2ª. ed. Boca Raton: Chapman & Hall/CRC, 2010. 34p.
- [16] FARIA, C. R. *Ferramenta Carolina para Identificação de Pragmas e Doenças na Cultura da Soja utilizando Processamento de Linguagem Natural*. Departamento de Computação. Universidade Estadual de Londrina. Dissertação (Mestrado em Ciência da Computação). Londrina, 2021. 87p.
- [17] THANAKI, J. *Python Natural Language Processing: Advanced Machine Learning and Deep Learning Techniques for Natural Language Processing*. 1ª. ed. Birmigham: Packt Publishing, 2017. 798 p.
- [18] BARBOSA, C. R. S. C. *Gramática para consultas radiológicas em Língua Portuguesa*. Instituto de Informática da Universidade Federal do Rio Grande do Sul. Dissertação (Mestrado em Ciência da Computação). Porto Alegre. 1998. 143p.
- [19] BARBOSA, C. R. S. C. *Técnicas de Parsing para Gramática Livre de Contexto Lexicalizada da Língua Portuguesa*. Instituto Tecnológico de Aeronáutica. Tese (Doutorado em Engenharia Eletrônica e Computação). São José dos Campos,- 2004. 160p.
- [20] WILKINSON, K. M. Profiles of Language and Communication Skills in Autism. *Mental Retardation and Developmental Disabilities Research Reviews*, v. 4, dec., 1998. p. 73-79.
- [21] LOPES-HERRERA, S. A. O uso da linguagem no autismo de alto funcionamento e na síndrome de Asperger: uma perspectiva pragmática na intervenção fonoaudiológica. *Cadernos de Comunicação e Linguagem*, v. 1, jan., 2009. p. 87-106.
- [22] ARMENGAUD, F. *A Pragmática. Tradução Marcos Marcionilo*. São Paulo: Parábola Editorial, 2006. 159p.

- [23] CLUBE DO PORTUGUÊS. *Pragmática – definição e campos de estudo*. 2021. Disponível em: <<https://www.clubedoportugues.com.br/pragmatica/>>. Acesso em: dez 2022.
- [24] DEVLIN, J. J., CHANG, M-W; LEE, K.; TOUTANOVA, K. *Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2018. Disponível em: <<https://arxiv.org/abs/1810.04805>>. Acesso em: 03 mar 2021.
- [25] JOHN SNOW LABS. *Spark NLP*, 2018. Disponível em: <<https://nlp.johnsnowlabs.com>>. Acesso em: 25 fev. 2022.
- [26] USZKOREIT, J. Transformer: a Novel Neural Network Architecture for Language Understanding. *Google Blog - AI*. 2017. Disponível em: <<https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>>. Acesso em: 03 maio 2020.
- [27] TECH TARGET. *BERT Language Model*. 2022. Disponível em: <<https://www.techtarget.com/searchenterpriseai/definitionBERT-language-model>>. Acesso em: 04 jan. 2023.
- [28] SECRETARIA DE VIGILÂNCIA EM SAÚDE - MINISTÉRIO DA SAÚDE. *Suicídio. Saber, Agir e Prevenir*. 2017. Disponível em: <<http://portal.arquivos2.saude.gov.br/images/pdf/2017/setembro/21/2017-025-Perfil-epidemiologico-das-tentativas-e-obitos-por-suicidio-no-Brasil-e-a-rede-de-atencao-a-saude.pdf>>. Acesso em: 24 jun. 2020.
- [29] HOSPITAL SANTA MONICA. Cyberbullying e suicídio: como influenciam crianças e adolescentes? 2018. Disponível em: <<https://hospitalsantamonica.com.br/cyberbullying-e-suicidio-como-influenciam-criancas-e-adolescentes/>>. Acesso em: 24 jun. 2020.
- [30] HINDUJA, S.; PATCHIN, J. W. Bullying, Cyberbullying, and Suicide. *Archives of Suicide Research*, v. 14, n. 3, jul. 2010. p. 206-221.
- [31] CRAIG, W.; BONIEL-NISSIM M.; KING, N.; WALSH, S. D.; BOER, M. DONNELLY, P. D.; HAREL-FISCH Y.; MALINOWSKA-CIESLIK, M.; MATOS, M. G.; COSMA, A.; EIJNDEN, R. V. D.; VIENO A.; E. F. J. E.; MOLCHO, M.; BJERELD Y.; PICKETT W. Social media use and cyber-bullying: A cross-national analysis of young people in 42 countries. *Journal of Adolescent Health*, v. 66, n. 6, jun., 2020 p. 100-108.
- [32] BERNE, S.; FRISÉN, A.; BERNE, J. *Cyberbullying in Childhood and Adolescence: Assessment, Negative Consequences and Prevention Strategies*. In: Lunneblad, J. (eds) *Policing Schools: School Violence and the Juridification of Youth. Young People and Learning Processes in School and Everyday Life*, vol 2. Springer International Publishing. set. 2019. p. 141–152.

- [33] PINTO, M. *Menina de 15 anos suicida-se por sofrer de Cyberbullying*. Peopleware, 2012. Disponível em: <<https://pplware.sapo.pt/informacao/menina-de-15-anos-suicida-se-por-sofrer-de-cyberbullying/>>. Acesso em: 24 jun. 2020.
- [34] EXAME S.A. *Brasil fica em segundo lugar em raking global de ofensas na internet*. Brasil fica em segundo lugar em raking global de ofensas na internet, 2018. Disponível em: <<https://exame.com/brasil/brasil-fica-em-segundo-lugar-em-ranking-global-de-ofensas-na-internet/>>. Acesso em: 04 jan. 2023.
- [35] TOKARNIA, M. *IBGE: um em cada dez estudantes já foi ofendido nas redes sociais*. 2019. Disponível em: <<https://agenciabrasil.ebc.com.br/geral/noticia/2021-09/ibge-um-em-cada-dez-estudantes-ja-foi-ofendido-nas-redes-sociais>>. Acesso em: 04 dez. 2022.
- [36] DOSOMETHING.ORG. *11 facts about cyberbullying*, 2022. Disponível em: <<https://www.dosomething.org/us/facts/11-facts-about-cyber-bullying#fn1>>. Acesso em: 04 jan. 2023.
- [37] AGENCIA BRASIL. *Celular principal ferramenta de estudo e trabalho na pandemia*. 2020. Disponível em: <<https://agenciabrasil.ebc.com.br/economia/noticia/2020-11/celular-e-principal-ferramenta-de-estudo-e-trabalho-na-pandemia>>. Acesso em: 03 dez. 2022.
- [38] LABOREIRO, G.; OLIVEIRA, E. What we can learn from looking at profanity. *11th Computational Processing of the Portuguese Language*. In: BAPTISTA, J., J., Mamede, N., Candeias, S., Paraboni, I., Pardo, T.A.S., Volpe Nunes, M.G. (eds) Sao Carlos: Springer International Publishing. 2014. p. 108-113.
- [39] ALMEIDA, J. J. *Dicionário de Calão e Expressões Idiomáticas*. 1ª. ed. Lisboa: Guerra e Paz, 2019. 261p.
- [40] WICHMANN, S. How to distinguish languages and dialects. *Computational Linguistics*. v.45, n. 4, dez., 2019. p. 823–831.
- [41] THE GUARDIAN. *Death of 13-year-old prompts cyberbullying test case*. 2008. Disponível em: <<https://www.theguardian.com/world/2008/jun/17/usa.news>>. Acesso em: 01 jan. 2023.
- [42] THE NEW YORK TIMES. *Verdict in MySpace Suicide Case*. 2008. Disponível em: <<https://www.nytimes.com/2008/11/27/us/27myspace.html>>. Acesso em: 04 out. 2022.
- [43] ABC NEWS. *Parents: Cyber Bullying Led to Teen's Suicide*. 2009. Disponível em: <<https://abcnews.go.com/GMA/story?id=3882520&page=1>>. Acesso em: 01 jan. 2023.
- [44] MEGAN MAIER FOUNDATION. *We work to create a world where bullying and cyberbullying no longer exist. Will you join us and Be the Change?* 2022. Disponível em:

- <<https://www.meganmeierfoundation.org/>>. Acesso em: 01 jan. 2023.
- [45] NEW YORK POST. *Japan makes cyberbullying punishable by up to year in prison after reality TV star's suicide*. 2022. Disponível em: <<https://nypost.com/2022/06/15/japan-makes-cyberbullying-punishable-after-reality-tv-star-hana-kimuras-suicide/>>. Acesso em: 01 jan. 2023.
- [46] PTASZYNSKI, M. A Decade of Automatic Cyberbullying Detection: How far are we from eradicating Internet harassment? *Youtube Inc.* 2021. Disponível em: <<https://www.youtube.com/watch?v=cU03mWBdKeE>>. Acesso em: 05 out. 2022.
- [47] MONTEIRO, P. Pai de Lucas Santos, que se suicidou após comentários em vídeo, deseja: “Abrir os olhos dos pais”. *Yahoo vida e estilo*. 2021. Disponível em: <<https://br.vida-estilo.yahoo.com/pai-lamenta-suicidio-filho-lucas-santos-apos-comentarios-video-tiktok-222031969.html>>. Acesso em: 04 dez. 2022.
- [48] BAZILIO, A. L. Os alertas deixados pelo suicídio de Lucas, um adolescente vítima do ódio e da LGBTfobia no TikTok. *Carta Capital*, ago. 2021. Disponível em: <<https://www.cartacapital.com.br/diversidade/os-alertas-deixados-pelo-suicidio-de-lucas-um-adolescente-vitima-do-odio-e-da-lgbtfobia-no-tiktok>>. Acesso em: 20 ago. 2022.
- [49] CNN EDITORIAL RESEARCH. *US Violence in School Fast Facts*. 2022. Disponível em: <<http://edition.cnn.com/2013/09/19/us/u-s-school-violencefacts/index.html?iref=allsearch>>. Acesso em: 10 jan. 2023.
- [50] INTERNET MOVIE DATABASE. *Fusillade de Columbine: Dans la tête des tueurs*. *Internet Movie*. 2007. Disponível em: <<https://www.imdb.com/title/tt2926246/>>. Acesso em: 20 dez. 2022.
- [51] INTERNET MOVIE DATABASE. *Elefant*. 2003. Disponível em: <<https://www.imdb.com/title/tt0363589/>>. Acesso em: 20 dez. 2022.
- [52] CNN. *Affidavit: Columbine shooter posted threat on Web*. 2001. Disponível em: <<http://edition.cnn.com/2001/LAW/04/10/columbine.evidence/index.html>>. Acesso em: 14 nov. 2022.
- [53] STILINGUE SA. *LinkedIn*. Disponível em: <www.linkedin.com/posts/stilinguebbb20-activity-6660693056481701888-wraK>. Acesso em: 20 mar. 2022.
- [54] VOZES da Pandemia. *Stilingue SA*. Disponível em: <<https://covid.stilingue.com.br/>>. Acesso em: 21 jun. 2020.
- [55] YADAV, J.; KUMAR, D.; CHAUHAN, D. Cyberbullying Detection using Pre-Trained

- BERT Model. 2020 International Conference on Electronics and Sustainable Communication Systems. Coimbatore, India. *Proceedings IEEE*. 2020. p. 1096-1100.
- [56] GRANEY-WARD, C.; ISSAC B.; KETSBAIA L.; JACOB S. M. Detection of Cyberbullying Through BERT and Weighted Ensemble of Classifiers. *TechRxiv*, 2022. Disponível em: https://www.techrxiv.org/articles/preprint/Detection_of_Cyberbullying_Through_BERT_and_Weighted_Ensemble_of_Classifiers/17705009 Acesso em: 10 dez. 2022.
- [57] SAINJU, K. D.; NITI MISHRA, A.; KUFFOUR, L. Y. Bullying discourse on Twitter: An examination of bully-related tweets using supervised machine learning. *Computers in Human Behavior*, v. 120, n. 106735, ago., 2021.
- [58] ALMEIDA, R. J. D. A. *Estudo da Ocorrência de Cyberbullying contra Professores na Rede Social Twitter por meio de um Algoritmo de Classificação Bayesiano*. 2021. Disponível em: <https://www.redalyc.org/articulo.oa?id=577163628009>. Acesso em: 13 jan. 2023.
- [59] SKOK.AI. *Training BERT for Cyberbullying Detection - HF Trainer Baseline*. 2020. Disponível em: <https://skok.ai/2020/10/27/Training-BERT-for-Cyberbullying-Detection-Part-1.html>. Acesso em: 10 jan. 2023.
- [60] CUOGHI, L. *Detecting Bullying Tweets with PyTorch Bidirectional LSTM and BERT Transformer*. 2022. Disponível em: <https://www.kaggle.com/code/ludovicocuoghi/detecting-bullying-tweets-pytorch-lstm-bert>. Acesso em: 10 dez. 2022.
- [61] WANG, C.; LI, M.; ANDSMOLA, A. J. *Language models with transformers*. 2019. Disponível em: <https://arxiv.org/abs/1904.09408>. Acesso em: 24 jun. 2020.
- [62] LACERDA, L. BERT: o novo algoritmo do Google que promete revolucionar as SERPs. *Rock Content*. 2020. Disponível em: <https://rockcontent.com/br/blog/bert/>. Acesso em: 2 out. 2021.
- [63] LEITE, J. A.; SILVA, D. ; BONTCHEVA, K. SCARTON, C. Toxic Language Detection in Social Media for Brazilian Portuguese: New Dataset and Multilingual Analysis. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. Suzhou, China. ACL. 2020. p. 914-924.
- [64] LIU, Y.; OTT M.; NAMAN, G.; DU, J.; JOSHI, M.; CHEN, D.; LEVY, O.; LEWIS, M.; ZETTLEMOYER, L.; STOYANOV, V. *RoBERTa: A robustly optimized BERT pretraining approach*. 2019. Disponível em: <https://arxiv.org/abs/1907.11692>. Acesso em: 24 jun.

2020.

- [65] SALAZAR, J.; LIANG, D.; NGUYEN, T. Q.; KIRCHHOFF, K. Masked Language Model Scoring. *Annual Meeting of the Association for Computational Linguistics*. 2020. Disponível em: <<https://www.amazon.science/publications/masked-language-model-scoring>>. Acesso em: 12 abr. 2021.
- [66] INC, S. L. *Samurai Labs AI*. Disponível em: <<https://www.samurailabs.ai/#cyberguardian>>. Acesso em: 04 jan. 2023.
- [67] SAMURAI LABS INC. CyberGuardian API. *Rapid API*, 2022. Disponível em: <<https://rapidapi.com/samurai-labs-samurai-labs-default/api/username-guardian1>>. Acesso em: 01 dez. 2023.
- [68] JOHN SNOW LABS INC. *Spark NLP: State of the Art Natural Language Processing*. Disponível em: <<https://nlp.johnsnowlabs.com/>>. Acesso em: 25 fev. 2022.
- [69] C3.AI. *Machine Learning. Pipeling*. Disponível em: <<https://c3.ai/glossary/machine-learning/machine-learning-pipeline/>>. Acesso em: 25 fev. 2022.
- [70] JUPYTER INC. *Jupyter project*. Disponível em: <<https://jupyter.org>>. Acesso em: 01 jan. 2023.
- [71] AL-RFOU, R. *Part of Speech Tagging. Polyglot*. 2015. Disponível em: <<https://polyglot.readthedocs.io/en/latest/POS.html>>. Acesso em: 25 fev.2022
- [72] FERREIRA, R. C. B. *Uma Abordagem Semiautomática para Identificação de Elementos de Processo de Negócio em Texto em Linguagem Natural*. Instituto de Informática. Universidade Federal do Rio Grande do Sul. Dissertação (Mestrado em Ciência da Computação). Porto Alegre. 2017. 103p.
- [73] SAKURAI, G. Y. *Processamento de Linguagem Natural - Detecção de Fake News*. Departamento de Computação. Universidade Estadual de Londrina. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação). Londrina. 2019. 37p.
- [74] TRIPATHI, A. Named Entity Recognition NER using spaCy | NLP. *Towards Data Science*. 2020. Disponível em: <<https://towardsdatascience.com/named-entity-recognition-ner-using-spacy-nlp-part-4-28da2ece57c6>>. Acesso em: 12 abr. 2022.
- [75] SILVA C. F.; TEIXEIRA C. A. C.; CASELI, H. M. Classificação de tweets por relevância para concepção de um modelo de aprendizado de máquina para uso em aplicações de TV Social. *XIV Encontro Nacional de Inteligência Artificial e Computacional*. Uberlandia. Springer. 2017. p. 685-696

- [76] SILVA C. F.; TEIXEIRA C. A. C.; CASELI, H. M. *Classificação de tweets MasterchefBR - Recursos*. 2017. Disponível em <<https://sites.google.com/view/classificacao-tweetsmasterchef/recursos>>. Acesso em 21 out. 2022
- [77] MIONI, J. L. V. M. *Análise de PLN e índice binário de toxicidade*. 2023. Disponível em: <https://colab.research.google.com/drive/1txbknH0MwRMwwXklIDa4hEiS0j7oy3Cw#scrollTo=BZUG9U_Z3egf>. Acesso em: 21 jan. 2023.

ANEXO A – CORPUS MASTERCHEF

um beijo pro ratatouille #MasterChefBR
A líder do time azul é igual a muitos líderes nas empresas que são arrogantes e insuportáveis #MasterChefBR
RT @MChefRoger: olha aqui querida de peixe eu entendo!!! Vraaaaaa @MChefFabrizio @masterchefbr #MasterChefBR https://t.co/97DHyH6wFk
RT @Candidatavip: o crush visualizou e esta escrevendo #MasterchefBR https://t.co/sYwUelKgcF
Não quero que nem a miriam e nem o fabrizio saia se não meu barraco esta perdido #MasterChefBR #MasterChefBR https://t.co/GsOXYZqucB
RT @thiago_p: A mulher não conseguiu cozinhar uma mandioca e joga a culpa no outro que ficou sozinho com o peixe. hahahahaha #MasterChefBR
Cozinhando sozinha #MasterChefBR https://t.co/et7H4vxXC
RT @edyenecaroline: Para alarmes com a @PaolaCarosella dizendo: "acorda,acorda!" eu digo SIM! #MasterChefBR
RT @misilmg: 45 minutos dá tempo de, no máximo, abrir uma lata de sardinha! #MasterChefBR
Tá vendo como ele surge? #MasterChefBR https://t.co/7rRyynEUoj
RT @missbannwart: Conselhos para a vida com Erick Jacquin #MasterChefBR https://t.co/7w1FuOsO59
RT @cleytu: qnd vc ta deitado quase dormindo e de repente tem a sensação de q tá caindo #MasterChefBR https://t.co/067HSq5AH1
RT @laarissamarcal: Qualquer semelhança é mera coincidência #MasterChefBR https://t.co/OD0JHclQyp
RT @masterchefbr: Sua vó começa a chorar com a novela e quer que você se emocione também! #MasterChefBR https://t.co/jMTQXzZJ9g
ads com a tartaruga que eu achei que era da Tortuguíta, mas é de apartamento... hahahah #MasterChefBR https://t.co/nbDz27wyoa
RT @AmericooJr: O crush mudou o status para solteiro!!! #MasterChefBR https://t.co/o67i0LT616
RT @NonLacerda: Quando tem treta no #MasterChefBR 😊😊👏👏 https://t.co/PD4oaykvJr
Isto que esta passando no Jornal da Globo é uma das coisas mais asquerosas que eu vi hoje. E tem gente limpando peixe no #MasterChefBR
RT @cleytu: eu assistindo as tretas no #MasterChefBR https://t.co/lt0nCfCe8Y
RT @Ketiigomes: Não acredito que uma senhorinha é a embuste da edição #MasterChefBR https://t.co/OKOg4Zu8nW

Cheguei no #masterchefBR e já quero sentar na janelinha
"O salmão é a vaca do mar" #MasterChefBR Eu sou a vaca do Twitter.
a gente torceu pela miriam e foi ela que tirou a marzilia eu odeio o mundo #MasterChefBR
RT @twispn: aqui em casa ja montei meu prato com uma redução de biscoito e um pão desconstruido beijo @PaolaCarosella 🍷 #MasterChefBR
RT @BlueLarry_: #MasterChefBR quando vejo meus favs subindo pro mezanino https://t.co/1NQJ3WAd8d
RT @cleytu: QUANTA TRETA #MasterChefBR https://t.co/b6LmwmxBmj
Gente, o que é DIL que eles tanto falam? #MasterChefBR
#MasterChefBR https://t.co/FoU1CmzgFI
Qual o nome do peixe que tá sendo preparado hoje no #MasterChefBR? https://t.co/yjKhvXhe0
O pessoal desse #MasterchefBR gosta de usar esse tal de DIU né
A prova eliminatória de hoje tem uma pitada de tensão a mais! Estamos adorando? #MasterChefBR https://t.co/5Y5T0gku9M
Miriam ela vende Mary Kay #MasterChefBR
#MasterChefBR eu amo o cabelo da nana
Receiteira sou eu. No "Tudo Gostoso" da vida eu sou aquela que procura qualquer prato com "Fácil e Rápido" no título! #MasterChefBR
#MasterChefBR bem cobras no mezanino
Nathalia Dill #MasterChefBR
O @MChefLeonardo é a enciclopedia do MC ☐☺. #masterchefbr
Quem é a mais chata do programa levanta a mão. #MasterChefBR https://t.co/vUdm3UOYWn
Minha barriga te ama #MasterChefBR
#MasterChefBR MEZANINO ME REPRESENTA
RT @miguelevmusic: "eu sou receitera" it's the new "eu sou rolezera" #MasterChefBR
Quando as pessoas me dão apelidos carinhosos: #MasterChefBR https://t.co/aVqCJkpo2e
Quem a gente tá tetestando no #masterchefBR?
yuko apagada da edição de hj #MasterChefBR
Acho que sou o único aqui que não gosta de ver treta no MasterChef #MasterChefBR
Nível de confiança atualizado. #MasterChefBR https://t.co/1bNIFNwEP
Não quero que a Mirian saia, mas também não quero que ela ganhe essa prova! kkkkk #MasterChefBR
Se eu alimentasse a fome que o #MasterChefBR sempre me provoca acho que estaria com uns 300 Kg.

Nível de incentivo: "EMPRATA AGORA ESSA PORRA! #MasterChefBR
Por favor, não passem do ponto da truta #MasterChefBR https://t.co/tdn2ESS0p4
Empratar, filetar... Verbos que são a cara do #MasterchefBr
que afliçãoooooo #MasterChefBR
Os dez segundos da Ana Paula são 20 segundos, né. #MasterChefBR
Mirian e Fabrizio brigam no #MasterchefBR e internautas vão à loucura - https://t.co/xAqFWi6fc1 https://t.co/foHw4yg0Xw
Depois do climão vem o drama #MasterChefBR
Já podemos lançar um dicionário. #MasterChefBR https://t.co/XYEyql9MwY
Nossa, aguei aqui #MasterChefBR
Minha única regra pra comer peixe é: se ele olhar pra mim eu não como. #MasterChefBR
Quando você dá aquele unfollow libertador! #MasterChefBR https://t.co/Pgoty0qSVB
RT @masterchefbr: Já podemos lançar um dicionário. #MasterChefBR https://t.co/XYEyql9MwY
#MasterChefBR PISA MENOS ADELIZE
#MasterChefBR PISA MENOS EU TE IMPLORO
RT @_CrazyGenious: @MChefDouglas QUE NAJAAAA AHAHAHAHAHAH ADORO #MasterChefBR
AGORA PARE PEGUE NO BUMBUM #masterchefBR https://t.co/GXlba0NCZU
A Aderlize é muito boa. Vai de mansinho e se destaca. #MasterChefBR
Quando eu vejo aquela @ de sunga #MasterChefBR https://t.co/YC3MfZsVXW
É tão bonito os chefs elogiando um prato, falando com carinho #MasterChefBR
#MasterChefBR ADERLIZE EU TE VENERO AAAAA
Não pode respirar nessa cozinha hoje que geral tá com nervos à flor da pele! #MasterChefBR
A cara da Miriam HAHAHA raivosa #MasterChefBR
Aderlize com sangue nos olhos pra sambar na cara da Miriam #MasterChefBR
Miriam colando na prova descaradamente. #MasterChefBR
miriam é o otariano do masterchef #MasterChefBR
#MasterChefBR MIRIAM EU TE ODEIO
RT @bicmuller: AGORA PARE PEGUE NO BUMBUM #masterchefBR https://t.co/GXlba0NCZU
ENTRANDO MACIO EH SEMPRE BOM NE #MasterChefBR
miriam copiando a aderlize chocado com a vilã da edição. #MasterChefBR
#MasterChefBR PAOLA PISA MENOS DEMONIA EI TE AMOO
A Marian cozinhando parece eu colando nas provas! Imitou o prato da colega! hahahahah #MasterChefBR

Miriam servindo os pratos para os amigos de jogo: #MasterChefBR https://t.co/lrMGlgfRry
Brasil inteiro gorando o prato da Miriam #MasterChefBR
Só eu que torço pra Mirian se ferrar na prova kkkk #MasterChefBR
Miriam deu uma de Otariano na colega? #MasterChefBR
#MasterChefBR fez o mesmo prato e fez pior.
eu vendo a mirian se dar mal copiando o prato da amiga #MasterChefBR https://t.co/NumxL7bQv5
ufa acabou a contagem muito dificil isso #MasterChefBR https://t.co/MO24Ed092z
Pela primeira vez vimos a cola numa prova de culinária. #MasterChefBR
será q a band dá ticket refeição pros jurados do #masterchefbr ou não precisa
#MasterChefBR MIRIAM ACUSADA DE PLÁGIO
E ESSE KIBE NO #MasterChefBR TA ASSADO OU CRU?
#MasterChefBR foi decapitado atoa kkkkk
Essa contagem regressiva da Ana Paula Padrão é mais mentirosa que a Miriam. #MasterChefBR
Quando sinto cheiro de gasolina na rua. #MasterChefBR https://t.co/kwIVeo9k8s
Indo stalkear / Não gostei do que vi. #MasterChefBR https://t.co/Z3y5Nwslro
MIRIAM KIBANDO OUTRA PARTICIPANTE, É O OTARIANO DO PROGRAMA. #MasterChefBR
#MasterChefBR "É uma treta rei da França"
ana paula fazendo as tretas acontecerem que rainha. #MasterChefBR
batom preto. me faz perder a fome. #MasterchefBR
- me empresta a sua receita?? - empresto mirian só não vai fazer igual - pode deixar - #MasterChefBR https://t.co/DaGgY5aYx7
Mas é obvio que não vão tirar a Mirian rainha das tretas dessa edição hoje #MasterChefBR
"O peixe já tá morto e ainda decapitado" 😏😏 #MasterChefBr
Que batom trevososo! #MasterChefBR
E essa moça com uma jabuticaba na boca? #MasterChefBR
Ana Paula Padrão só botando lenha na fogueira GOSTO ASSIM #MasterChefBR
Quero tretaaa.. Alias, Trutaaa ❤️ #MasterChefBR
RT @JJhonwespinoza: Minhas Laricae são mais gourmets que alguns pratos @EuFumoUm #MasterchefBr
qnd seu amigo te chama pra ir na casa dele comer #MasterChefBR https://t.co/a23WuJz0NB
Jacquin é mítico. hahahahaha #MasterChefBR
Adivinha quem vai aproveitar o intervalo do masterchef pra assaltar a geladeira? () euzinha (x) minha amiga, claro! #MasterChefBR

Esse batom da taise lacrou total #MasterChefBR
Paola Carosella é criticada após comentário no "MasterChef": "Seu papel é avaliar, e não ofender!". #MasterChefBR https://t.co/25U38pW1oZ
Paola Carosella é criticada após comentário no "MasterChef": "Seu papel é avaliar, e não ofender!". #MasterChefBR https://t.co/MqByHJ473D
Espero que não saia nem Mirian, nem Fabirizio e que semana que vem os dois tretem de novo #MasterChefBR https://t.co/3NbQmnhBZu
Qual a sua relação com o regime? #MasterChefBR https://t.co/xN8VSWLru0
RT @GiuVigiani: ai a pessoa te chama pra sair e ainda diz que paga #MasterChefBR https://t.co/XhTypsGfnB
Quem mais tá lendo o nome dos pratos como TRETA em vez de TRUTA levanta a mão! #MasterChefBR
#MasterChefBR so eu que acho a Paola super sexy experimentando?
Esse prato do Fabrício tá com uma apresentação horrorosa! Espero que esteja saboroso! #MasterChefBR
Aquele quiabo ali não tem nada a ver!! #MasterChefBr
Não sabia que peixe dava tanta treta assim #MasterChefBr
Quem nasceu pra ser #PowerCoupleBrasil nunca será #MasterChefBR https://t.co/7bgb56lwAe
#MasterChefBR https://t.co/Con8d6efmd
Que prato deprê! Parece um despacho light! #MasterChefBR https://t.co/ylqLcpoUwW
RT @possentifer: Você não aderiu à greve? #MasterChefBR https://t.co/XLoz7EN0w2
adorei a participação da Paola Carrosella cozinhando nesse filme #Masterchefbr https://t.co/PZmbuyDDOT
Jacquin kkkkkkkkkkkkkkkkkkkkkkkkkkkkkkk #MasterChefBR
RT @andresinkos: Que esses peixes tejem melhor que o da Miriam #MasterChefBR https://t.co/5EcynKSsYF
Fabrizio e sua cara de quem vai cair no choro #MasterChefBR
RT @mrssexie: Taise tem um senso de superioridade, né? #MasterChefBR
RT @pimentabih_: Eu quando fico dando dicas de dieta, mas na sigo elas. #MasterChefBR https://t.co/GI3oalbx7V
Fabrizio colocando defeito no próprio prato sou eu apresentando trabalho na faculdade #MasterChefBR
#MasterChefBR MIRIAM DE Q VC TÁ RINDO FOFA?
Miriam amando o tombo do Fabrizio #MasterChefBR https://t.co/pbif4HoPJY
Paola elogia com delicadeza e critica com classe, ao mesmo tempo que dá um esculacho.

#MasterChefBR
KKKKKKKK Mirian rindo do rei dos peixes kkkk #MasterChefBR
Paola obrigada por essa aula #MasterChefBR
#MasterChefBR KKKKKKKSKSJEJWKS EU TO PASSANDO MAL
A Caroline é maravilhosa. Seria amigo fácil. hahahahahaha #MasterChefBR
A Mirian rindo da cara do Fabrizio. Putz vacilou feio pra ela #MasterChefBR
Nois na #QuartaComTsunamiSDV ↩️💕 #MasterChefBR https://t.co/mGNoJuz8at
Paola apavorada com o prato da Caroline! hahahaha #MasterChefBR
não entendi o que aconteceu entre paola e caroline mas amei. #MasterChefBR
Paola melhor pessoa hahaha #MasterChefBR
Gente eu amo a Paola #MasterChefBR
A @MChefCaroline é uma figuraça! 😊 #MasterChefBR
Peixe na água? #MasterChefBR
paola e caroline melhores pessoas. #MasterChefBR https://t.co/7Xq5LZaT8H
Imagina sair de role com a Caroline depois dela tomar uns 3 cerveja e uns 5 shots de catuaba? #MasterChefBR https://t.co/ybvANNxXId
#MasterChefBR eita Carol.. pqp
#MasterChefBR VAI YUKO
Acho que Caroline vaza #MasterChefBR
Eu não gosto dessa Carol, tenta ser legal e engraçada, mas é irritante. #MasterChefBR
#MasterChefBR comida de prisão, eu já teria me retirado da cozinha
Certeza que quem vai rodar é a Caroline! Impossível algo pior! #MasterChefBR
Quando a gente encontra um bebezinho fofinho ai ti coisinha mais linda! #MasterChefBR https://t.co/4zjeVKwpCK
MIRIAM VAI SE LASCAR #MasterChefBR
YUKO EU TE AMO AA QUERO GUARDAR NUM POTINHO #MasterChefBR
RT @JuninhoClemente: Ai meodeos yuko #MasterChefBR
Yoku dá vontade de apertar! #MasterChefBR
A Caroline dá risada até quando tá toda lascada, melhor pessoa #MasterChefBR
#MasterChefBR KKKKKKKKKKKKKK SOCORRO
HAHAHAHAHAHAHA Youko #MasterChefBR
Yuko conquistando meu coração. #MasterChefBR

Jiang incorporou em outra asiática. #MasterChefBR
Yuko dá vontade de guardar numa caixinha #MasterChefBR
YUKO RAINHAA! HAHAAA #MasterChefBR
#MasterChefBR KSKEKWKAKWKEKEEEEKEKKSMEJEJS EU TO PASSANDO MAL C A APRESENTAÇÃO DA YUKO
RT @MChefFabrizio: Mas péra, @MChefMirian cadê a mandioca que era a sua única obrigação?! #MasterChefBr #teamFabrizio https://t.co/JZw1ji8w...
nao quero ver ngm dali saindo antes da chatonilda da miriam #MasterChefBR
Yuko e Jiang num mesmo programa, POR FAVOR! #MasterChefBR
RT @keloverdaturma: Quando eu vejo alguém dá minha sala na rua #MasterChefBR https://t.co/p5m6D0WZga
Uma ou duas namoradas? #MasterChefBR https://t.co/Xvk70Pc49v
RT @MattyBala: AAhhhhh vou chorar.... Estou orgulhoso de você @MChefYuko #MasterChefBR https://t.co/HrqMxc4Fkl
Esse prato da Yuko é quase um X-TUDO! #MasterChefBR
PROTEJAM A YUKO A TODOS OS CUSTOS #MasterChefBR
Eu simplesmente ADORO essa YUKO! Maravilhosa 😊😊 #MasterChefBR
NOSSA QUANTA GENTE JA PODE ELIMINAR A METADE #MasterChefBR
quando a miriam começa a falar #MasterChefBR https://t.co/7FuGQuussY
Numa noite de ELIMINAÇÃO "esquecível" é até bom. Só ESQUECEREM de te colocar entre os 3 piores #MasterChefBR
#MasterChefBr amo a Japinha! Humildade , carisma, alegria ... torcendo por ela ! Depois da semana passada , ela merece !
RT @julianojrg: PASEI PARA DIZER: YUKO RAINHA DA EDIÇÃO #MasterChefBR
Yuko ainda curte um menage. RAINHA. HAHAAHAAH #MasterChefBR
"é um prato bem esquecível" MORTOOOO #masterchefBR https://t.co/eqHgHjv3sa
Que miséria de prato! #MasterChefBR
RT @mxriet: caroline sai hoje puts #MasterChefBR
Apenas nesse caso. #MasterChefBR https://t.co/AwufEgpHnQ
RT @desantana_lucas: Yuko e Paola melhor dupla #MasterChefBR
SIM, daqueles que você vai comer e o arroz cai todo pros lados! #MasterChefBR https://t.co/H02tuk4fHH
#MasterChefBR FRESCURA MAL FEITA
#MasterChefBR TOMARA Q SAIA

Essa Mirian tem os defeitos das maioria daquele que se acham líderes, mas não passam de chefes nas empresas #MasterChefBR
Leonardo nao merece sair. Ele é um dos melhores. #MasterChefBR
PATADAS E PATADAS no Professor Aloprado! #MasterChefBR
eita!!!! a semi-precious ... *schadenfreude #MasterchefBR
RT @aryale7: #MasterChefBR Jacquin é hilário.. #ClássicoDePrisão o peixe hahahaha
RT @paiqueroopaul: Yuko e Michelle rainhas, Mirian nadinha! #MasterChefBR
Esperando acabar o #MasterChefBR https://t.co/3tGABbpDso
Humildade faz bem em qualquer profissão #MasterChefBR
Can I get an amen here? #MasterChefBR https://t.co/jP0nOch2F1
Marian colou na prova e passou! #MasterChefBR
TIREM QUALQUER UM MENOS FABRIZIO E MIRIAN EU IMPLORO #masterchefBR https://t.co/l45bM8AZ3s
Aderlize mereceu demais a vitória! #MasterChefBR
#MasterChefBR MIRIAM mt chata
MALDITA, EU TE ODEIO SUA ESCROTA #MasterChefBR
Miriam, pq tão chata? #MasterChefBR
Uia... e aquele #chupa hein+ hahaha #MasterChefBR
Parabéns, @MChefAderlize! Preparou a melhor truta da noite e arrasou nessa treta! #MasterChefBR https://t.co/zyF3XTvdoK
#MasterChefBR MIRIAM N, MIRIAM N
CHUPA ESSA MIRIAM SUA EMBUSTE #MasterChefBR
SAMBA ADERLIZE NA CARA DA MIRIAN, FILMA A CARA DELA A MIRIAN DA MARYKAY #masterchefBR
Tá facil saber quem sai né #MasterChefBR
Caroline né, que prato pavoroso foi aquele? #MasterChefBR
#MasterChefBR NÃO ACREDITO QUE A MICHELE VAI SAIR, MEU DEUS... EU NÃO ACREDITO
"Uma empresa funciona quando os colaboradores colaboram e os líderes lideram" deveria fazer uma placa dessa frase #MasterChefBR
Quem vai deixar a competição na noite de hoje? #MasterChefBR https://t.co/Wltx1uqdYn
Vem falar sobre o #MasterChefBR de hoje com a gente! https://t.co/jrtJNX1xh3
Chupa Mirian nem foi a melhor #MasterChefBR
Yuko e Jiang juntas cozinhando. COMO NÃO AMAR ESSAS MELHORES PESSOAS????? #MasterChefBR https://t.co/uRakXFJ6Zt

A pior foi a Carol acho que ela roda. #MasterChefBR
Quem ceis acham que sai? #MasterChefBR
O Leonardo provavelmente não vai sair mas vai ser bom pra ele deixar um pouco de ser egocêntrico #MasterChefBR
NÃO TIREM A CAROLINE EU IMPLORO #masterchefBR https://t.co/VCMViRVbd9
Truta assada com molho de maracujá e emulsão de batata baroa da @MChefAderlize! #MasterChefBR... https://t.co/XqReaC7jVc
Quando alguém começa a falar mal da minha série favorita perto de mim. #MasterChefBR https://t.co/4pESVtpstK
Carol sai hoje do #MasterChefBr Pena, eu gosto dela. Uma fofa. Mas torço pro Leonardo. E acho Fabricio tão antipático qt Mirian.
RT @masterchefbr: Parabéns, @MChefAderlize! Preparou a melhor truta da noite e arrasou nessa treta! #MasterChefBR https://t.co/zyF3XTvdok
a caroline vai sair to sentindo #MasterChefBR
Leonardo entre os piores não imaginava. Mas é isso aí, o #MasterChefBR é imprevisível.
Receita da @MChefYuko: Truta assada com purê de aspargos com ervilha, ratatouille e caldo de missô #MasterChefBR... https://t.co/CAZRNXd9uS
Quando me SDV na #QuartaDeTremuraSDV ↩️❤️ #MasterChefBR https://t.co/E8YT8pwbD6
Trocaria Michele por Miriam #MasterChefBR
Melhor programa! 📺 @fogaca_henrique @PaolaCarosella e @erickjacquin 😊👏👏👏👏👏 #MasterChef #MasterChefBR
O povo sempre escolhe a Yuko por último nas provas de equipe mas quase sempre ela tá entre os melhores pratos, podem perceber #MasterChefBR
A @MChefMirian ensinou a receita da truta na manteiga com purê de mandioquinha e molho de maracujá #MasterChefBR... https://t.co/cL36evuKCR
RT @cleytu: quando a Yuko fala da vontade de entrar na tv e apertar as bochechas dela #MasterChefBR
RT @cleytu: qnd seu amigo te chama pra ir na casa dele comer #MasterChefBR https://t.co/a23WuJz0NB
RT @cleytu: qnd vc ta deitado quase dormindo e de repente tem a sensação de q tá caindo #MasterChefBR https://t.co/067HSq5AH1
RT @cleytu: QUANTA TRETA #MasterChefBR https://t.co/b6LmwmxBmj
#MasterChefBR AAAAAA EU TE AMO MICHELE
Fala comigo bb. #MasterChefBR https://t.co/MCbxix8CB7A
#MasterChefBR AH NÃO CARA

E agora, quem continua, quem sai? A sua torcida vai pro @MChefLeonardo ou pra @MChefCaroline? #MasterChefBR https://t.co/JnaQ3npy5k
Ficar na berlinda assim é terrível... #MasterChefBr
Caroline foi eliminada com sua "Comida de Prisão"! #MasterChefBR
sheldon eliminada #MasterChefBR
#MasterChefBR AH NÃO CARA, VAI SE FUDE. pq esse lixo ficou?
Carol eliminada NÃOOO #MasterChefBR https://t.co/me5K3bEeke
Caroline é simpática e é boa. Tem gente lá que não é tão bo aquanto ela. Uma pena. #MasterChefBR
É... eu imaginava que a Caroline saísse... Uma pena #MasterChefBR
Caroline caras e bocas eliminada AAAAAAAAAAAAAA #MasterChefBR
O Leo não merecia sair mesmo, mas a Caroline era demais. Fará falta. #MasterChefBR
Não acredito que a rainha dos memes saiu #MasterChefBR
#masterchefbr https://t.co/Hu6WbR1jmF
O #MasterChefBR perde uma fábrica de memes sem a Caroline. Não era uma das melhores mesmo, mas era ousada e engraçada. Espontânea.
Mas tem repescagem #MasterChefBR
A @MChefCaroline não acertou no peixe e precisou deixar a nossa cozinha. Foi uma honra ter uma figura tão especial... https://t.co/0f1zOjUedN
#MasterChefBR eu to bem triste..
"Amanhã vou dormir até 4 da tardeeeeeeeeeeeeeee!!!!!!!" (Caroline) HAHAHAHAHAH #Prioridades #MasterChefBR
tinha que ser a mirian não a carolzinha ela e muito humilde e cozinha com amor #MasterChefBR
A Carol é um amorzinho tao lindo que chega da vontade de agarrar e nao soltar mais #MasterChefBR
Primeiro dia de férias feels like! #MasterChefBR https://t.co/NLK3arSFSD
Não acredito que a filha da Tulla Luana saiu! #MasterChefBR https://t.co/phtjmwshmy
"amanha eu vou dormir ate 16h da tarde" AAAAAAAAAAAAAAAA que invejinha gospel de vc ícone carol #MasterChefBR
e agora? de que memes sobreviveremos? #MasterChefBR
"Quebrei alguns estereótipos", afirma @MChefCaroline ao deixar o #MasterChefBR https://t.co/WOOD48jl2r https://t.co/5V05mml6fi
Vou sentir falta dos comentários da @MChefCaroline... 😊😌 #MasterChefBR
RT @portaldaband: "Quebrei alguns estereótipos", afirma @MChefCaroline ao deixar o #MasterChefBR https://t.co/WOOD48jl2r https://t.co/5V05mml6fi
Todo o sucesso do mundo pra vc, Caroline! Vc é uma querida! #MasterChefBR https://t.co/7LYxXjINXf

caroline era tão amorzinho entendo a eliminação mas fiquei triste. #MasterChefBR https://t.co/yuUI8oKQRu
Os olhos da Caroline é um mix dos olhos desse meme #MasterChefBR https://t.co/2HNpgL9ERI
Queria Mirian saindo, e me sai a Caroline. Ô inferno. #MasterChefBR
ta bom gente ate semana que vem na nossa corrente de pensamentos positivos pra Mirian sair #MasterChefBR https://t.co/1cQtjwUc0h
gente a senhora é a vilã dessa edição #MasterChefBR
Miriam engole esse sapo ai... #MasterChefBR melhor voltar pro consultório em...
PROXIMA PROVA EM EQUIPE, BOTA TODOS CONTRA A MIRIAN, AÍ SIM QUERO VER SE ESSA EMBUSTE NUM SAI #MasterChefBR
RT @masterchefbr: Os pidunchos. They are everywhere! #MasterChefBR https://t.co/L4Vmlo7iJ9
Pequi é muito bom, recomendo. #MasterChefBR
Ave... Oq essa Miriam é chata... #MasterChefBR
Que lavada! Toma Mirian #MasterChefBR
a mirian tá achando q é a M do 007 faz de tudo nessa vida #MasterChefBR
Janeiro: "2017 é o meu ano" Maio: #MasterChefBR https://t.co/fA1D5uL55b
@masterchefbr A equipe do pidunchos está aumentando... #MasterchefBR
ixi, colega... o que mais tem na cozinha é mexer na panela dos outros, hahah e o mais: PEDIR A OPINIÃO DA EQUIPE É IMPORTANTE SIM #MasterchefBR
Adoro todos do time azul, mas a Mirian perdeu sozinha essa prova. #MasterChefBR
#MasterChefBR MAIS CHATO Q A MIRIAM, SÓ A ANA REPETINDO OS PRÊMIOS TODA TERÇA
Time Vermelho SUPREMO. #FATALITYCOMBOPLUS #MasterChefBr https://t.co/SQOZbnJBgP
RT @lilis_hill: Qdo uma das zamiga desencana!#MasterChefBR https://t.co/tw6kGEApsF
Equipe azul indo pro mezanino #MasterChefBR https://t.co/OJsKwLGTcr
Partiu prova de eliminaçãoããooo... Flw! Vlw! #MasterChefBR https://t.co/h4wTMSq62S
brigaaaaaaa #MasterChefBR
Climão #MasterChefBR
Enquanto isso no #MasterChefBR, o @MChefFabrizio está usando uma camiseta com estampa do Chaveco, personagem do Che... https://t.co/lKMuAWQdW8
Gente, maior climão #MasterChefBR
#MasterChefBR TRETAS, TRETAS.... SOCORRO
Essa Miriam é o canal exposto no molar dos participantes do #MasterchefBR.
SEGURA ESSE DEBOCHE DO FABRIZIO #MasterChefBR
RT @Televizona: CADE MIRIAN TU NÃO ERA A FODONA MANA? AH OK! PERDEU A PROVA DE

LAVADA NÉ!!!! #MasterChefBR https://t.co/T1hJqaRjf3
A mulher não conseguiu cozinhar uma mandioca e joga a culpa no outro que ficou sozinho com o peixe. hahahahaha #MasterChefBR
CARA NA POEIRAAAAAAAAAAAAAAAA DE NOVO MIRIAAAAAAN #MasterChefBR
RT @agentecorneta: tem TRETA no #MasterChefBR e a gente está como? https://t.co/djx7WaiiXu
A acareação da Miriam #MasterChefBR
miriam completamente sapateada pelo seu grupo. #MasterChefBR
Mirian VS Fabrizio Pow Pow Pow #MasterChefBR https://t.co/L9jUE083qd
mirian tem que sair já tá achando que é casos de família o programa só briga e quer mandar em tudo #MasterChefBR
AMANDO ESSA TRETAAAAAA #MasterChefBR https://t.co/GqQWmQJFSq
QUANTA TRETA #MasterChefBR https://t.co/b6LmwmxBmj
TRETAAAAA! Escondam as facas!!! hahahaha #MasterChefBR
#MasterChefBR MEU DEUS EU AMO ESSE PROGRAMA
RT @ForumChaves: Enquanto isso no #MasterChefBR, o @MChefFabrizio está usando uma camiseta com estampa do Chaveco, personagem do Chespirito...
Mirian SAPATEADA pelo grupo! AMOOOOOOO #MasterchefBR https://t.co/jZEqqSDLYr
PISA MAIS NESSA EMBUSTE FABRIZIO #MasterChefBR
programa tá demais hoje #MasterChefBR
Débora tem dentes ao quadrado!! #MasterchefBR
A cada dia que passa o mezanino se torna mais desejado! #MasterChefBR
Ai, Leonardo nem pra entrar na treta serve #MasterChefBR https://t.co/EGObAzu6bY
Mas a Ana Paula também fica cutucando o povo pra espizinhar a Miriam #MasterChefBR
ana paula padrão muito rainha consegue de forma sutil e com classe instaurar a maior torta de climão você quer tiago leifert? #MasterChefBR
#MasterChefBR até as treta daqui são melhores que o BBB
A MULHER FALA PRA NAO FAZER E AGORA PAGA DE SANTA? AH VAI SE FUDER #MasterChefBR
Esse Leonardo pula fora neh das posições no jogo neh !! Estamos de olho #MasterChefBR https://t.co/cdHEXBkfwr
Vão cozinhar um tiranossauro rex #MasterChefBR
E essa torta de climão 😊 #MasterChefBR https://t.co/YNkYNUEIQQ
RT @edswaglima: IIIIIIIIIIT`S TIME! UFC Masterchef!!! Fight! #masterchefbr
EQUIPE VERMELHA #MasterChefBR arrasaram. Depois dessa, vms morfar @masterchefbr!!!!

https://t.co/g3Oamwgdxp
RT @lookingforlara: essa Débora é inteligente demais bicho, mulherão da porra #masterchefbr
EU QUERO A MIRIAN FORA #MasterChefBR
RT @masterchefbr: Não teve como pisar menos. #MasterChefBR https://t.co/gDtLftm0oY
E depois da TRETA, vamos cozinhar uma TRUTA! #MasterChefBR
A Mariam é insuportável!! Agora imagina ela como sua sogra.... DEUSMELIVRE #MasterChefBr
É mano, a prova hj ta embaçada! Certo TRUTA!! #MASTERCHEFBR
RT @LoohValadares: Agora é a hora do Fabrizio mostrar que sabe fazer um bom peixe #MasterChefBR
"Meus trutas, essa prova é muita treta!" @fogaca_henrique #MasterChefBR https://t.co/IsVnybQ6C6
Quando vejo o pote de sorvete na geladeira, mas descubro que tem feijão dentro #MasterChefBR https://t.co/Nrwx3sCHyz
#MasterChefBR É AGORA QUE O FABRIZIO PISA NA MIRIAM
O PRATO DEVIA SER TORTA DE CLIMÃO #MasterChefBR
Quero esse meme da Vivian escolhendo a equipe dela #MasterChefBR
RT @NonLacerda: EQUIPE VERMELHA #MasterChefBR arrasaram. Depois dessa, vms morfar @masterchefbr!!!! https://t.co/g3Oamwgdxp
eu lembrando da atual situação do país #MasterChefBR https://t.co/xZ2JVReUnR
RT @amaurit: 100 trutas, 1000 tretas! #MasterChefBR https://t.co/VD7dCxFVNN
A cozinha da @AzulinhasAereas não para, então bora cozinhar! O @Fogaca_Henrique foi conferir o menu #MasterChefBR e... https://t.co/tWjVbz72Ja
#MasterChefBR ezza cara da Paola
Meus treta essa prova é mto truta #MasterChefBR
Mil Trutas, Mil Tretas #MasterChefBR
RT @cleytu: O PRATO DEVIA SER TORTA DE CLIMÃO #MasterChefBR
RT @JonassOficial: #MasterChefBR Quando eu como peixe por muito tempo - Só pode ser maldição dos peixes
#MasterChefBR ALGUÉM TIRA O FABRIZIO DO LADO DA MIRIAM PFV
Chegamos para o #MasterChefBR <3
45 minutos #MasterChefBR https://t.co/bDXZufH7LD
a yuko é tão fofinha q ela parece aquele biscoito do shrek #MasterChefBR https://t.co/lw7Vc4dzh1
RT @marthinetto: eu lembrando da atual situação do país #MasterChefBR https://t.co/xZ2JVReUnR
se me dão 3 minutos de mercado eu não consigo nem entrar #MasterChefBR
Só lembro de uma lavada dessa no 7 x 1 da copa do mundo! Agora no #MasterChefBR 58 x 2

@masterchefbr
Quando começam a falar mal de mim. #MasterChefBR https://t.co/BBG8kWIeRU
RT @masterchefbr: Os pidunchos. They are everywhere! #MasterChefBR https://t.co/L4Vml07iJ9
Quando eu chego no MasterChef e descubro que perdi a treta: #MasterChefBR https://t.co/0d0jQvoNKY
RT @NathyVilla: Treta no #MasterChefBR e a @PaolaCarosella não tá envolvida! Inédito!
TRUTA vs TRETAS #MasterChefBR
Quando eu beijo a boca da @ #MasterchefBR https://t.co/XkFGCIsdms
Eu tenho a maior curiosidade de saber a logística pra esses depoimentos, são antes ou depois das provas?... https://t.co/gqwQybE7Gt
Não vale colar na prova! #MasterChefBR https://t.co/aqPh9AFyqF
#MasterChefBR https://t.co/MRS0h6I5BQ
.@masterchefbr @PaolaCarosella @erickjacquin @TwitterBrasil Em Bonito (MS), você também pode nadar com os peixes em... https://t.co/vw3bgMZJuH
os bichinhos unidos! hahaha - "Que frio, mores" #MasterChefBR https://t.co/bxQ4z6MYBO
RT @rickirodrigues: Inventando a receita de hoje #MasterChefBR https://t.co/OIAITzb0P6
Vamos na festa? - Vai ter comida? - Sim. - #MasterChefBR https://t.co/tkO3gdWxwC
E eu achando que truta era só uma gíria de Paulista. #MasterChefBR
Ela quer infernizar a vida do Fabrizio? #MasterChefBR
Que beleza se a Miriam saísse fazendo peixe em hahahahah #MasterChefBR
que sujeita irritante essa myriam #MasterchefBR
TAO COM SANGUE NO ZOIO #MasterChefBR
eu assistindo as tretas no #MasterChefBR https://t.co/lt0nCfCe8Y
@MChefMirian e @MChefFabrizio ainda saem no tapa 🤔🌟até o final desse episódio... 😊 #MasterChefBr
Finalmente vou aprender a cozinhar um belo peixe :) #MasterChefBR
MOSTRA PRA MIRIAM COMO É QUE SE FAZ UM PEIXE, FABRIZIO #MasterChefBR
RT @falarafha: mais uma terça assistindo #MasterChefBR e com fome querendo jantar de novo https://t.co/qqARQTsfLi
JA QUERO PAY PER VIEW DESSA TRETA ENTRE FABRIZIO E MIRIAN #MasterChefBR
a vontade que eu tô de ver a miriam sendo eliminada não está no gibi #MasterChefBR
Jurado do "MasterChef Brasil" choca fãs com foto antiga: "Irreconhecível" #MasterChefBR https://t.co/sQtq5V3XKI
Jurado do "MasterChef Brasil" choca fãs com foto antiga: "Irreconhecível" #MasterChefBR

https://t.co/HPLTIRIQRb
Quando tem treta no #MasterChefBR 😊😊👏👏 https://t.co/PD4oaykvJr
O pessoal do mezanino parece quinta série bicho #MasterChefBR
#MasterChefBR https://t.co/hzuKtM8tA3
RT @masterchefbr: Sobrou uma bronquinha da @PaolaCarosella até pro @MChef_Raul ao vivo! #MasterChefBR https://t.co/hUZBealV5r
#MasterChefBR FORA MIRIAM, FORA MIRIAM
RT @NonLacerda: Quando tem treta no #MasterChefBR 😊😊👏👏 https://t.co/PD4oaykvJr
Jurado do "MasterChef Brasil" choca fãs com foto antiga: "Irreconhecível" #MasterChefBR https://t.co/CFpuP6cgck
O prato de hoje será um peixe ao sugo" #MasterChefBR https://t.co/LcMHCUP3ax
Adoro peixes, adoro truta, mas a prova de eliminação eu passo! #MasterChefBr 🍷🔪🌟
#MasterchefBR: Mara Alcamim participa em prova de pratos do Centro-Oeste - https://t.co/FzU2PEzDui https://t.co/xtqHTamQG5
A truta não namora piranha #MasterChefBR
#MasterChefBR a cafa da Paola de nojo foi a melhor KKKKKKKKKKKK EJ TO PASSANDO MAL
MARAVILHOSO #MasterChefBR
RT @AmericooJr: O crush mudou o status para solteiro!!! #MasterChefBR https://t.co/o67i0LT616
E os jurados colocando lenha na fogueira... 🔥🔥🔥 #MasterChefBr
"Preciso falar com você e tem que ser pessoalmente" #MasterChefBR https://t.co/yEdDrckslh
Se a Miriam não sair hoje a gente apela pro STF. #MasterChefBR
Foca no PIPARUCU! #MasterChefBR https://t.co/0lytteh7SS
A única receita de peixe que eu sei fazer é abrir a lata do atum ou da sardinha. #MasterChefBR
Fico preocupado quando vejo a Yuko na prova de eliminação VAI PASTEL DE FLANGO #MasterChefBR
.@masterchefbr @PaolaCarosella @erickjacquin @TwitterBrasil Você pode conhecer várias receitas com alimentos region... https://t.co/ybKJ15upZA
Batata sautéé com alecrim é muito bom! #MasterChefBR
As pessoas não aprendem que em prova de eliminação não se arrisca, se vai no certo. pq vc nao precisa ganhar, só passar #masterchefbr
quando a pessoa q eu gosto posta foto no insta #MasterChefBR https://t.co/TTIBEOlJkG
Sua vó começa a chorar com a novela e quer que você se emocione também! #MasterChefBR https://t.co/jMTQXzZJ9g
eu quando a miriam começar a falar: #MasterChefBR https://t.co/ZTEp6vcWbP

RT @AlexandreMedina: Corre truta! #MasterChefBR https://t.co/o5BaNyZNfW
Vamos lá!! Torcendo pro @MChefFabrizio , @MChefTaise e a @MChefYuko ! Caprichem neste peixinho aí, gente! 🤞🤞🤞🤞 #MasterChefBR #MasterChef
RT @edyenecaroline: Para alarmes com a @PaolaCarosella dizendo: "acorda,acorda!" eu digo SIM! #MasterChefBR
RT @masterchefbr: Sua vó começa a chorar com a novela e quer que você se emocione também! #MasterChefBR https://t.co/jMTQXzZJ9g
RT @masterchefbr: A cozinha da @AzulinhasAereas não para, então bora cozinhar! O @Fogaca_Henrique foi conferir o menu #MasterChefBR e já bo...
"O salmão é a vaca do mar.." Jacquín. 😊 sensacional 🤩🤩 #MasterChefBr
SINHANA #MasterChefBR
RT @404jrg: VAI FABRIZIO! #MasterChefBR
RT @ROBERTACARLOS: #MasterchefBR Aderlize e demente viu fia
RT @TVShow: Jurado do "MasterChef Brasil" choca fãs com foto antiga: "Irreconhecível" #MasterChefBR https://t.co/sQtq5V3XKI
qnd vc ta deitado quase dormindo e de repente tem a sensação de q tá caindo #MasterChefBR https://t.co/067HSq5AH1
O povo fica inventando moda numa receita tão simples. #MasterChefBR
#MasterChefBR Sobre minha cama: https://t.co/7MSNCp62rx
Vem falar sobre o #MasterChefBR de hoje com a gente! https://t.co/EbfE9ff0oR
Soltei spoiler do #MasterChefBR no quarto e sai correndo 🏃😅
Gentemmm, acho que vou perder o fim da treta das trutas. Buenas noches a todos. Caprichem nos memes! #MasterChefBR
#MasterChefBR Sobre feriado prolongado: https://t.co/ZwXpqeT9RI

TRABALHOS PUBLICADOS PELO AUTOR

1. Mioni, J. L. V. M.; Barbosa, C. R. S. C. Ferramentas para o Aprendizado de Linguagens Formais e Autômatos. **XXIII Simpósio Brasileiro de Games e Entretenimento Digital (SBGames)**. Natal-RN, Brasil, 2022. (Qualis 2022, B1)
2. Mioni, J. L. V. M.; Barbosa, C. R. S. C. Análise Pragmática no Desenvolvimento de Software de Detecção de *Cyberbullying*. **XII Seminário de Estudos sobre Linguagem e Significação (SELISIGNO) – Processamento de Linguagem Natural (PLN) e Suas Implicações para o Entendimento da Língua Materna**, Londrina-PR, Brasil, 2022.
3. Mioni, J. L. V. M.; Barbosa, C. R. S. C. Análise Léxico-Morfológica no desenvolvimento de software de detecção de *cyberbullying*. **III Fórum da Pós-Graduação em Ciência da Computação do Paraná (ForPPGC-PR)**, Cascavel-PR, Brasil, 2021.
4. Gobbo, M. R. M.; Barbosa, C. R. S. C.; Morandini, M; Mafort, F.; Mioni, J. L. V. M. ACA game for individuals with Autism Spectrum Disorder. **Entertainment Computing**. Elsevier, Ago, 2021, v. 38, p. 40-54.
5. Campos, V. V.; Brancher, J. D.; Farias, F. P.; Mioni, J. L. V. M.; Brahim, P. Review and comparison of works on heterogeneous data and semantic analysis in Big Data. **Semina: Ciências Exatas e Tecnológicas.**, Out, 2021, v. 42. p. 113-128.
6. Mioni, J. L. V. M.; Barbosa, C. R. S. C. Processamento de Linguagem Natural na Detecção de Comportamentos Agressivos e Situações de Perigo na Internet. **II Fórum da Pós-Graduação em Ciência da Computação do Paraná (ForPPGC-PR)**, Cascavel-PR, Brasil, 2020.
7. Pereira, Y. H.; Barbosa, C. R. S. C.; Mioni, J. L. V. M.; Artoni, A. A.; Brancher, J. D. Interface em Linguagem Natural para uma Sublinguagem de Câncer de Pele. **XI Computer on the Beach (COBt)**. Florianópolis-SC, Brasil, 2020 p. 469-474. (Qualis 2017-2020, B4)
8. Gobbo, M. R. M.; Barbosa, C. R. S. C.; Morandini, M; Mafort, F.; Mioni, J. L. V. M. Jogo ACA para indivíduos com transtorno do espectro autista. **XXII Simpósio Brasileiro de Games e Entretenimento Digital (SBGames)**. Foz do Iguaçu – PR, Brasil, 2018. p. 1114-1121. (Qualis 2017-2020, B2)
9. Gobbo, M. R. M.; Barbosa, C. R. S. C.; ; Mioni, J. L. V. M; Mafort F. ACA – Learning With Alternative Communication. **XIII International Conference on the**

Computational Processing of the Portuguese Language (PROPOR). Canela – RS, Brazil. Springer. 2018. p. 83-92. (Qualis 2017-2020, B3)

10. Gobbo, M. R. M.; Barbosa, C. R. S. C; Morandini, M; Mafort F.; Mioni, J. L. V. M.. Ferramenta para Alfabetização de Crianças com TEA. **XXIII Conferência Internacional sobre Informática na Educação (TISE)**. Brasília-DF, Brasil. 2018, p.80-88 (Qualis 2017-2020, B5)