



25 a 28
setembro
2024
Campus Central UEPG
Ponta Grossa | PR

Explorando as Interseções das Inteligências
Artificiais na Sociedade Atual

Realização:



Apoio:



COMTURPG



PROCESSAMENTO DE LINGUAGEM NATURAL COMO FERRAMENTA PARA ANÁLISE DE TEXTO EM DIFERENTES CONTEXTOS DA ADMINISTRAÇÃO

NATURAL LANGUAGE PROCESSING AS A TOOL FOR TEXT ANALYSIS IN DIFFERENT ADMINISTRATION CONTEXTS

ÁREA TEMÁTICA: MARKETING - MARKETING DE SERVIÇOS, DE RELACIONAMENTO E DE
VENDAS

Bárbara Adelle Dalamaria, Universidade de Passo Fundo, Brasil, 174283@upf.br

Cassiana Maris Lima Cruz, Universidade de Passo Fundo, Brasil, cassiana@upf.br

Resumo

O Processamento de Linguagem Natural (PLN) é caracterizado como um tipo de inteligência artificial que atua com identificação e interpretação da linguagem natural humana. Esta pesquisa objetivou mapear e compreender as possibilidades de aplicação do PLN em diferentes contextos da Administração que necessitam de análise de texto. O procedimento metodológico utilizado foi uma Revisão Sistemática da Literatura guiada pelo Protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta Analysis*) para a seleção dos dados, seguido da análise de conteúdo dos materiais selecionados. Os resultados apontam que os mecanismos de PLN vêm se aprimorando com o passar dos anos e atualmente têm potencial de ser utilizados para Análise de Sentimento, Pesquisas Organizacionais, Análise Preditiva e em Instituições de Ensino. Por fim, foi desenvolvido um *framework* que sintetiza os resultados encontrados.

Palavras-chave: Processamento de Linguagem Natural; Análise de Texto; Administração; Revisão Sistemática da Literatura.

Abstract

Natural Language Processing (NLP) is characterized as a type of artificial intelligence that works with the identification and interpretation of human natural language. This research aimed to map and understand the possibilities of applying PLN in different Administration contexts that require text analysis. The methodological procedure used was a Systematic Literature Review guided by the PRISMA Protocol (Preferred Reporting Items for Systematic Reviews and Meta Analysis) for data selection, followed by content analysis of the selected materials. The results indicate that NLP mechanisms have been improving over the years and currently have the potential to be used for Sentiment Analysis, Organizational Research, Predictive Analysis and in Educational Institutions. Finally, a framework was developed that summarizes the results.

Keywords: *Natural Language Processing; Text Analysis; Administration; Systematic Literature Review.*

1. INTRODUÇÃO

Nas últimas décadas o mundo vem vivenciando uma série de inovações e mudanças a nível global no que diz respeito às novas tecnologias e à transformação digital. Frente a esse cenário, a sociedade e as organizações enfrentam o desafio de incorporar essas tecnologias em seus processos cotidianos de maneiras cada vez mais otimizadas e inteligentes. As demandas mudam, a economia opera sem pausas, a vigilância aumenta e as incertezas começam a tomar conta (Weiss, 2019).

O conceito de Inteligência Artificial (IA) possui diversas concepções e situações de aplicabilidade. Existem dimensões que caracterizam as diversas possibilidades de atuação da IA e todas determinam que essa tecnologia consiga reproduzir práticas cognitivas humanas (Gomes, 2010) nas atividades de pensar (Haugeland, 1985), agir (Kurzweil, 1990), pensar racionalmente (Charniak & Mcdermott, 1985) e agir racionalmente (Poole, Mackworth & Goebel, 1998). Portanto, a IA é um ramo da Ciência da Computação cujos processos objetivam reproduzir sistemas cognitivos de pensamento e ação humanos, podendo realizar tarefas que exigem a inteligência humana para serem efetuadas (Russell & Norvig, 2021; Liu, 2023).

Paralelo a isso, o processamento de linguagem natural (PLN) está vinculado ao processamento da fala e a todas as suas dimensões. Para que uma máquina aprenda a realizar o processamento de linguagem natural é necessário que ela seja treinada através de algoritmos e programação. Portanto, o processamento de linguagem natural é uma prática de *machine learning* de IA que possibilita que a tecnologia simule as conexões realizadas pelas redes neurais do cérebro humano durante o processo de comunicação falada ou escrita e também interprete seus significados (Santos, 2001).

Na área de Administração, o PLN tem potencial para ser ainda mais explorado, especialmente no contexto atual, onde a quantidade de informação disponível é inviável de ser assimilada unicamente por seres humanos (Pupo, 2023). Assim sendo, esta pesquisa tem como objetivo compreender como o sistema de Processamento de Linguagem Natural (PLN) pode ser utilizado na análise de textos na área da Administração.

2. REFERENCIAL TEÓRICO

2.1 Processamento de Linguagem Natural

O Processamento de Linguagem Natural (PLN) surgiu ligeiramente após os computadores, visto que a tradução automática de um idioma para outro foi uma das primeiras práticas submetidas aos computadores. O PLN passou a ser estudado cientificamente com maior frequência no Brasil a partir da década de 1990, quando os aparelhos digitais começaram a ser utilizados de forma abrangente no país (Caseli & Nunes, 2023).

Inicialmente, o PLN se caracterizava pela união de recursos linguístico-computacionais com a finalidade de fazer com que a máquina compreendesse e interpretasse perguntas feitas em linguagem natural, ou seja, a linguagem coloquial com termos, gírias e maneirismos de cada idioma. Com o passar do tempo e o advento da transformação digital as práticas de PLN vêm sendo cada vez mais desenvolvidas (Caseli & Nunes, 2023).

Logo, o PLN é uma prática que permite que máquinas compreendam e processem a linguagem humana, atualmente, o PLN é amplamente utilizado em *chatbots* e assistentes virtuais (Jurafsky & Martin, 2019). O PLN é um dos sistemas de IA que seguem sendo cada vez mais aperfeiçoados para que um dia consigam reproduzir e compreender a linguagem humana em toda sua complexidade (Russell & Norvig, 2021).

2.2 Análise de Texto

A análise de material textual com a implementação de PLN, mineração de textos e outras técnicas vinculadas à IA têm se mostrado eficiente nos processos de avaliação automatizada da

linguagem escrita. Essas práticas podem ser implementadas para pesquisas e avaliações textuais no sentido de otimizar tais processos (Pinho, 2021).

O PLN pode ser utilizado para análise textual através de seus recursos para reconhecimento automático de fala e reconhecimento de emoções e predição de pontuação necessária. O PLN aliado ao *machine learning* constantemente se aprimora com relação à ordem e função de palavras em uma sentença, aos recursos para processamento da sintaxe das frases, à semântica das palavras e à extração de sentido das frases (Santos, 2001; Caseli & Nunes, 2023).

Alguns exemplos de como o PLN é utilizado atualmente são a tradução automática, a correção automática de pontuação e agentes de conversação, como ChatGPT (Caseli & Nunes, 2023). Porém, seu uso ainda pode ser ampliado, como objetivou-se comprovar neste estudo.

3. PROCEDIMENTOS METODOLÓGICOS

Este estudo qualitativo é caracterizado como uma Revisão Sistemática da Literatura (RSL) para compreender as principais contribuições à pesquisa com relação às temáticas abordadas neste estudo e realizar um *framework* com os resultados a fim de auxiliar pesquisas futuras acerca das temáticas. Partindo do princípio da robustez metodológica para a realização de uma revisão como esta, adotou-se o Protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta Analysis*) que abarca as etapas de identificação, triagem e análise dos materiais selecionados (Araújo, Queiroz & Aouar, 2023).

Primeiramente, foi definida a pergunta de pesquisa que conduziria os procedimentos seguintes. Segundamente, foram delimitados os critérios de elegibilidade e as bases de dados de onde o conteúdo seria extraído. As bases foram selecionadas por sua ampla quantidade de artigos e demais materiais em estudos organizacionais. Ademais, foram delimitados os termos de pesquisa a serem utilizados com o objetivo de extrair os materiais de maior assertividade e relevância para responder à questão de pesquisa.

A pesquisa inicial, com aplicação dos critérios de elegibilidade, resultou em 18 artigos na base de dados *Scopus* e 16 artigos na base de dados *Web of Science*. Após isso, foi realizada a triagem dos conteúdos inicialmente encontrados através da leitura dos títulos, resumos e palavras-chave de cada um dos 34 artigos pré-selecionados. Excluindo os 9 artigos repetidos e artigos que não respondiam ao problema de pesquisa, chegou-se a um total de 11 artigos para a análise de conteúdo.

ETAPA	DESCRIÇÃO
Questão de pesquisa	Como o sistema de Processamento de Linguagem Natural (PLN) pode ser utilizado na análise de textos na área da Administração?
Critérios de elegibilidade	Critérios de inclusão: artigos referentes à área de Administração que respondessem à pergunta de pesquisa. Critérios de exclusão: teses, capítulos de livro e publicações fora da área de Administração.
Fontes de conteúdo	<i>Scopus</i> e <i>Web of Science</i>
Termos de pesquisa	" <i>natural language processing</i> " AND " <i>text analysis</i> "
Análise de conteúdo	11 artigos.

Quadro 1 - Procedimento metodológico. Fonte: elaborado pelas autoras (2024).

A análise de conteúdo segue os elementos concernentes à técnica de Bardin (2011), dentro desse escopo, são respeitados os processos de (i) pré-análise; (ii) exploração do material encontrado; (iii) categorização do conteúdo analisado (iv) tratamento dos resultados e; (v) elaboração de

inferências e interpretações. Todas as etapas também foram adotadas no processo de análise de conteúdo dos artigos selecionados. A seguir, são apresentados os resultados.

4. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Inicialmente, são apresentadas as principais contribuições à pesquisa de cada um dos artigos analisados para esta RSL.

AUTOR(ES)	ANO DE PUBLICAÇÃO	PRINCIPAIS CONTRIBUIÇÕES
García-Crespo, Colomo-Palacios, Gómez-Berbís e Ruiz-Mezcua	2010	PLN para compreensão das opiniões dos consumidores sobre as marcas e organizações; assimilação de informações advindas das redes sociais e identificação de sua relação com as emoções de consumo dos usuários.
Antonio, Almeida, Nunes, Batista e Ribeiro	2018	PLN e <i>machine learning</i> no desenvolvimento de modelos com capacidade preditiva; aplicação de processos de PLN em múltiplos idiomas.
Bharti e Kumar	2019	PLN utilizado para análise de sentimentos com a finalidade de compreender sistematicamente informações subjetivas; mapeamento e classificação de opiniões dos usuários.
Barakhnin, Kozhemyakina, Mukhamediev, Borzilova e Yakunin	2019	Configuração de sistemas de PLN; etapas do processo de análise de texto realizado com PLN; PLN para identificar informações de natureza destrutiva.
Miric, Jia e Huang	2022	PLN utilizada para análise textual em pesquisas com o intuito de criar variáveis categóricas; PLN e <i>machine learning</i> para mapeamento de dados textuais estruturados e não estruturados; PLN e <i>machine learning</i> para transformar dados textuais em imagens e representações numéricas.
Greselin, Zitkikis, Vincenzo e Piacenza	2022	Sistemas de PLN utilizados para mapear e compreender riscos operacionais, fornecendo oportunidades de mitigação de riscos futuros.
Karvetski, Meinel, Maxwell, Lu, Mellers e Tetlock	2022	Torneios de previsão geopolítica; PLN para mapear probabilidade de eventos futuros.
Alantari, Currim, Deng e Singh	2022	PLN e análise de sentimentos no contexto de avaliações digitais de consumidores.
Poschmann, Goldenstein, Büchel e Hahn	2023	Utilização de PLN em pesquisas organizacionais como abordagem metodológica devido à grande quantidade de material textual, especialmente, online.
Baharuddin e Naufal	2023	Modelos de PLN para avaliar e classificar questões de provas escolares.
Todd, Bowden e Moshfeghi	2024	PLN e <i>machine learning</i> utilizados para análise de sentimento; identificação de

		elementos do discurso de gerentes e analistas financeiros; impacto dos sentimentos dos funcionários nos processos organizacionais.
--	--	--

Quadro 2 - Principais contribuições à pesquisa. Fonte: elaborado pelas autoras (2024).

Após o mapeamento dos estudos e a análise dos conteúdos, realizou-se a segmentação dos resultados advindos dessa análise. Essas compreensões serão apresentadas em quatro subitens a seguir.

4.1 Processamento de Linguagem Natural como Ferramenta para Análise de Sentimento

Uma prática no desenvolvimento de algoritmos de IA é o desenvolvimento de processos de *machine learning* com o objetivo de realizar o PLN. Trata-se de uma tarefa complexa visto que a criação de práticas universais para a análise de textos ainda é um processo em desenvolvimento. Para que o PLN possa ser realizado com mais eficiência existem alguns elementos que precisam ser considerados, como as especificidades da estrutura da língua e a existência, ou não, de sistemas completos de processamento de texto naquele idioma (Barakhnin et al., 2019).

O PLN identifica e avalia unidades do léxico existentes, busca extrair informações sobre o ambiente informal e a linguagem criativa, esse último processo é um dos mais complexos feitos pelo PLN pois a informalidade e os maneirismos gramaticais podem ser intrincados de modo que um sistema tenha dificuldade em interpretá-los. Com o passar do tempo e o desenvolvimento da IA, esses processos de PLN estão ficando cada vez mais qualificados em compreender a complexidade de recursos léxicos utilizados (Bharti & Kumar, 2019).

Quando se configura um sistema de PLN existem algumas dimensões que devem ser levadas em consideração: (i) processamento de dados; (ii) armazenamento de dados. Na primeira dimensão, os arquivos de texto são coletados e processados com um conjunto de métodos de *machine learning*. Na segunda dimensão, os resultados são armazenados em um banco de dados relacional. Por meio desses processos, é possível visualizar os textos processados e classificados (Barakhnin et al., 2019).

O processo de análise de texto envolvendo PLN possui as seguintes etapas: (i) Inicialização (formação do texto e seu pré-processamento); (ii) Análise Estrutural (classificação de fonética e metrorrítmica de poemas); (iii) Análise Semântica (definição de construções semânticas); (iv) Análise Pragmática (definição de gênero para textos literários); (v) Síntese dos Resultados Obtidos (determinação dos efeitos dos elementos textuais encontrados uns sobre os outros) (Barakhnin et al., 2019).

O PLN pode ser utilizado para análise de sentimento. A análise de sentimento é um estudo computacional das opiniões dos indivíduos, assim como suas atitudes e emoções com relação a uma instituição, ou entidade. Inicialmente, essas técnicas de análise de sentimento eram realizadas de forma intuitiva através de contagem de palavras, com as tecnologias atuais, é possível agrupar um conjunto de sentimentos gerais, assim como seus sinônimos e antônimos. O uso do digital torna essa contagem e as comparações muito mais abrangentes e precisas para a análise. Portanto, as abordagens de PLN e *machine learning* se mostram mais eficientes na classificação e análise de sentimentos (Todd et al., 2024).

A análise de sentimentos, ou mineração de opinião, é uma técnica de análise linguística de texto que faz uso do PLN para sua execução. Os objetivos dessa análise são identificar, extrair e compreender de forma sistemática informações subjetivas acerca dos sentimentos dos indivíduos (Bharti & Kumar, 2019).

O advento da internet e das redes sociais veio acompanhado de mudanças nas relações e no comportamento de consumo dos usuários. A influência das redes sociais afetou a Gestão de

Relacionamento com o Consumidor (*Customer Relationship Management - CRM*) sendo caracterizada como CRM 2.0. Dentro desse escopo, compreender as relações entre os consumidores e suas opiniões sobre as marcas e organizações se tornou essencial para a sobrevivência das empresas (García-Crespo et al., 2010).

Existem plataformas que classificam semântica e possibilitam a análise de texto a partir de métodos de PLN. Os autores apresentam a abordagem SEMO, ou seja, a captação de informações de redes sociais e sua posterior relação com as emoções de consumo dos usuários. A falta de precisão na análise de emoções básicas pode ser fruto de definições incompletas de vocabulário utilizadas no PLN. Por isso, é importante que as práticas utilizadas para mapeamento de emoções possuam o PLN aprimorado com o objetivo de captar todas as nuances necessárias (García-Crespo et al., 2010).

Atualmente, existem diversas plataformas online com recursos para que os usuários adicionem suas opiniões e relatos acerca de suas experiências com as marcas, serviços e produtos. Esses comentários são relevantes para que as organizações por trás das marcas, produtos e serviços possam avaliar sua atuação e seu impacto com o público. Como esses relatos são publicados em domínio público, as opiniões de um usuário podem influenciar no comportamento de outros (Bharti & Kumar, 2019).

Existe uma quantidade expressiva de avaliações digitais de consumidores em formato de texto. Proporcional a esse aumento, também aumentaram as abordagens de *machine learning* para análise de sentimentos baseadas em material textual. Pesquisadores de marketing possuem práticas que buscam avaliar preditivamente a preferência do cliente sobre a concorrência (Alantari et al., 2022).

As redes sociais possuem participação ativa no impacto social no aumento ou diminuição da popularidade e, conseqüentemente, na participação de mercado de uma marca. Dentro desse contexto, a análise de sentimentos pode ser utilizada para mapear e classificar as opiniões dos usuários, servindo como base para a reunião de elementos que atraíam emoções positivas sobre a marca para determinado segmento de usuários (Bharti & Kumar, 2019).

Esses diagnósticos são feitos baseados em estimativas e parâmetros e a análise textual das avaliações submetida a uma ferramenta de PLN tem potencial de ser um componente adicional para esses diagnósticos, tornando seu resultado final mais assertivo. Além disso, os pesquisadores de marketing podem identificar clientes que possuem sentimentos positivos com relação à marca (Alantari et al., 2022).

Por meio de análises dos modelos linguísticos utilizados é possível inferir sobre a influência do gerente sobre os resultados, especialmente no campo financeiro. Identificação de aspectos comuns em discurso identificam, por exemplo, que gerentes tendem a ter uma linguagem mais otimista e menos complexa sobre o respectivo negócio do que seus colegas de trabalho. Por esse motivo, os sentimentos do gerente e do analista financeiro com relação ao mesmo negócio podem ser divergentes (Todd et al., 2024).

Os métodos de análise utilizados para definir os sentimentos estão ficando cada vez mais complexos, especialmente quando se trata das relações entre as variáveis sentimento e mercado. Mesmo assim, os estudos em sua maioria ainda se valem de técnicas anteriores às evoluções e computacionais, como abordagens de dicionário específicas. O impacto dos sentimentos nos processos organizacionais estudado com o auxílio de PLN e *machine learning* é uma temática relevante para estudos futuros, assim como, atualização de procedimentos metodológicos (Todd et al., 2024).

4.2 Processamento de Linguagem Natural em Pesquisas Organizacionais

Documentos baseados em texto são fontes relevantes de informação tanto para pesquisadores como para o processo de análise de negócios. Os pesquisadores utilizam dados textuais não estruturados com o intuito de construir variáveis quantitativas para análise. Inicialmente, esse

processo era feito utilizando palavras-chave específicas e mapeamento dos conceitos teóricos relacionados. Nos tempos atuais, existem ferramentas de *machine learning* que conseguem realizar o processamento de linguagem natural e possibilitam a classificação de variáveis quantitativas em dados textuais não estruturados (Miric et al., 2022).

Embora as práticas de *machine learning* em PLN estejam sendo implementadas, ainda existem diversos pesquisadores que desconhecem suas possibilidades para uso em pesquisas envolvendo dados textuais não estruturados. *Machine learning* trata-se de um conjunto de algoritmos estatísticos com capacidade de “aprender” e melhorar seus processos através de interações, não necessariamente com instruções explícitas (Miric et al., 2022).

Essas técnicas podem ser utilizadas em pesquisas para classificar observações em grupos e criar variáveis categóricas. Ademais, podem ser utilizadas para identificar padrões dominantes em textos ou dados de imagens sem a influência de preconceitos teóricos que podem vir do pesquisador no momento de fazer essa identificação (Miric et al., 2022).

A análise de texto com PLN pode ser utilizada para mapear e compreender riscos operacionais, bem como, auxiliar na elaboração de estratégias para mitigar tais riscos em cenários futuros. O risco operacional é o risco de perda de dados relevantes durante procedimentos internos, falhas humanas ou de sistemas eletrônicos. Instituições financeiras possuem diversas estratégias para gerenciar esses riscos (Greselin et al., 2022).

Diversos sistemas de análise de PLN baseados em *machine learning* são utilizados para analisar corpus textual de notícias com objetivo de identificação de informações de natureza destrutiva (Barakhnin et al., 2019). Técnicas de análise de textos podem ser aplicadas às descrições de eventos de risco operacional por meio de métodos estatísticos. Greselin et al. (2022) promovem uma estrutura analítica de medição que considera as informações analisadas de forma holística, com o objetivo final de obter uma visão do risco operacional em toda sua complexidade.

A perspectiva de que os dados textuais em pesquisa organizacional fornecem uma importante janela para a sociedade, assim como a percepção de que esses dados permitem medir construtos baseados em teorias é tratada por Poschmann et al. (2023). Uma abordagem de espaço vetorial utilizando PLN possibilita a compreensão da multidimensionalidade desses construtos, assim como permite o cálculo de semelhanças entre as características das organizações em suas interações com o ambiente externo e interno. Tendências recentes em pesquisas organizacionais buscam explorar o potencial dos dados textuais utilizando PLN (Poschmann et al., 2023).

O uso da IA também se mostra valioso para investigadores em estratégia e inovação nas organizações. Portanto, a IA proporciona meios interessantes para análise de negócios e compreensão de mercado do ponto de vista estratégico (Miric et al., 2022).

Existem orientações acerca de como os pesquisadores organizacionais podem assegurar as medidas de similaridade relacional. Tais orientações se caracterizam por: (i) definição das dimensões dos construtos baseados nas teorias utilizadas; (ii) seleção de procedimentos de análise formal de textos; (iii) seleção dos textos; e (iv) definição teórica e seleção de métricas de similaridade (Poschmann et al., 2023).

É possível implementar métodos de PLN e *machine learning* para mapear dados de texto estruturados e não estruturados e relacioná-los com conceitos teóricos pré-existentes. Essa prática utilizando PLN e *machine learning* pode ser utilizada para transformar dados textuais e imagens em representações numéricas. Além disso, é possível identificar frequências de relacionamentos entre os dados, fornecendo subsídios valiosos para a formação de teorias (Miric et al., 2022).

Existe um apelo por essa abordagem metodológica em pesquisas organizacionais devido à quantidade expressiva de material textual disponível por meio de diversos canais de comunicação atualmente. A onipresença dos conteúdos textuais como notícias, artigos, relatórios e comunicados, possibilitam o uso do PLN para quantificar e compreender fenômenos

organizacionais em maior escala. Com a evolução das tecnologias, o PLN pode ser usado não apenas para agrupar e classificar, mas também para compreender e analisar esses dados (Poschmann et al., 2023).

4.3 Processamento de Linguagem Natural em Análise Preditiva

A combinação de práticas de PLN e *machine learning* possibilita a previsão das classificações de avaliações online. Recursos quantitativos das avaliações podem ser complementados com análise de sentimentos das informações textuais. Além disso, as ferramentas digitais atuais possibilitam que essas análises e previsões sejam realizadas em diversos idiomas simultaneamente, o que qualifica e amplia as possibilidades de previsão (Antonio et al., 2018).

O estudo de Antonio et al. (2018) propõe a criação de um modelo com capacidade preditiva de classificações de avaliações online de múltiplas fontes de hotéis, possibilitando, inclusive, a identificação de avaliações fraudulentas. O desenvolvimento desse modelo foi feito utilizando o PLN e *machine learning* juntamente com características qualitativas e quantitativas integradas com o intuito de gerar uma classificação mais encorpada. A possibilidade de compreender os componentes quantitativos e qualitativos de uma avaliação oferece a chance de compreender o real valor daquela avaliação. Isso pode auxiliar uma empresa a identificar os pontos fortes e fracos de seu negócio perante os clientes, além de possibilitar personalização de ofertas para os clientes (Antonio et al., 2018).

As avaliações online são de extrema importância para a indústria hoteleira, especialmente no que diz respeito ao impacto na reputação social do hotel. A pesquisa demonstra a importância do desenvolvimento de modelos preditivos para a especulação da receita. Além disso, embora complexa, a aplicação de sistemas de PLN e análise de texto em múltiplos idiomas é um recurso valioso para organizações de diversos segmentos. Esse sistema de classificação auxilia na identificação de avaliações fraudulentas e possibilita que os hoteleiros analisem tais avaliações e, se necessário, as denunciem à plataforma onde foram feitas (Antonio et al., 2018).

Karvetski et al. (2022) adaptaram métodos de PLN e análise textual para com o objetivo de encontrar estratégias de raciocínio distintas de análise, tais como: estilos cognitivos (complexidades dialéticas) e uso de comparação e taxas para informar previsões. Existem torneios de previsão geopolítica que buscam identificar as melhores ferramentas para julgar a probabilidade de eventos no mundo real.

Os principais previsores se valem de maior complexidade dialética e maior quantidade de classes de comparação em seus raciocínios. Práticas de treinamento de equipes podem aumentar a precisão das tarefas dos previsores por meio de treinamento de perspectiva, ou recrutamento de indivíduos que são tolerantes com pontos de vista divergentes, característica essencial para um bom previsor (Karvetski et al., 2022).

4.4 Processamento de Linguagem Natural na Educação

Modelos de PLN podem ser utilizados para avaliar e classificar questões de exames escolares. Os autores buscaram classificar questões de exames do ensino fundamental indonésio com o intuito de avaliar o nível de pensamento cognitivo. Existem modelos de PLN que podem ser projetados e configurados para análise de textos de países específicos, além disso, os modelos podem ser pré-treinados para identificar modelos de linguagem e codificar representações textuais não rotuladas (Baharuddin & Naufal, 2023).

É importante que as ferramentas de PLN sejam projetadas, configuradas e treinadas para as ações específicas de classificação e compreensão textual que elas precisarão realizar. Quanto mais qualificados forem esses treinamentos e configurações, mais qualificada será a análise posterior. Os modelos desenvolvidos e pré-treinados no estudo foram um sucesso. Esses modelos de análise são ferramentas relevantes para educadores, auxiliando na categorização e classificação de questões de exames (Baharuddin & Naufal, 2023).

5. CONCEPÇÃO DE FRAMEWORK DOS CONTEÚDOS ANALISADOS

O PLN é uma prática de IA que tem potencial para identificar e compreender conteúdo textual de modo a levar em consideração as características da linguagem natural através de unidades léxicas e construção de sentido. O PLN vêm sendo cada vez mais aprimorado com o passar dos anos e os processos de *machine learning* proporcionados pela IA também auxiliam essa evolução.

As práticas de PLN são utilizadas, principalmente, para pesquisas envolvendo análise de sentimentos relacionados às emoções de consumo e avaliações de consumidores em plataformas digitais. Tais pesquisas podem ser úteis para personalizar atendimentos e mapear o grau de satisfação geral dos clientes (García-Crespo et al., 2010; Bharti & Kumar, 2019; Barakhnin et al., 2019; Alantari et al., 2022; Todd et al., 2024).

O PLN também pode ser utilizado em pesquisas envolvendo práticas de inovação e gestão das organizações. Dados textuais também fornecem importantes informações sobre o ambiente externo à organização quando interpretados de forma correta e a grande quantidade de dados disponíveis na internet atualmente pode ser melhor analisada com processos efetuados por máquinas (Greselin et al., 2022; Poschmann et al., 2023).

Com relação à pesquisa acadêmica, o PLN auxilia ao explorar a análise de texto partindo de uma perspectiva holística através de dados quantitativos e qualitativos considerando o fato de ser um processo valioso quando se considera a análise de textos gravados ou escritos de maneira informal e grande quantidade de informações. Análises textuais são um importante aspecto a ser estudado visto que a quantidade de informações textuais está cada vez mais vasta, principalmente em plataformas online (Miric et al., 2022; Poschmann et al., 2023).

Previsões também podem ser feitas por meio de práticas de PLN através de mecanismos de *machine learning* que se valem de comparações e cálculos baseados nas informações disponíveis. Além disso, a análise preditiva proporcionada por processos de PLN e *machine learning* pode auxiliar no prognóstico de futuras avaliações online e na identificação de pontos de melhoria com eficiência (Antonio et al., 2018; Karvetzki et al., 2022).

Ademais, o PLN também pode ser utilizado para aferir o grau de complexidade de questões de provas no âmbito educacional (Baharuddin & Naufal, 2023). Ainda existem alguns desafios envolvendo o PLN no que diz respeito à complexidade da linguagem, visto que, além das regras do idioma, existe uma série de termos coloquiais ditos e escritos que podem confundir o processamento. Processos de *machine learning* vêm tentando contornar tais desafios ao se aprimorar cada vez mais.

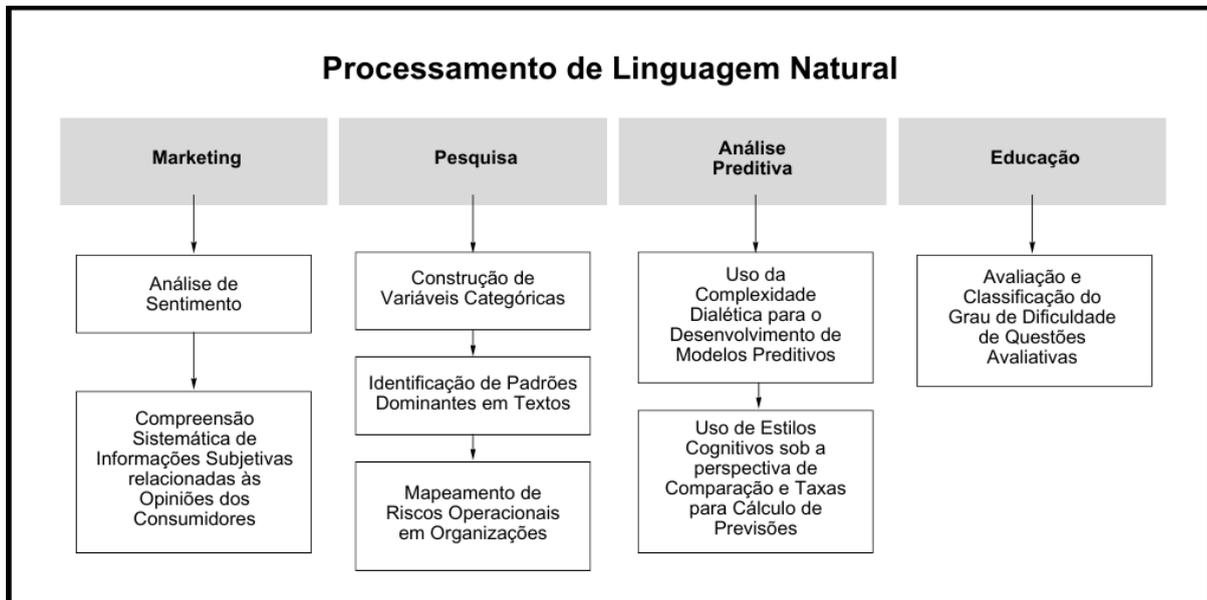


Figura 1 - *Framework* resultante da análise de conteúdo. Fonte: elaborado pelas autoras (2024).

6. CONSIDERAÇÕES FINAIS

O Processamento de Linguagem Natural é um sistema de IA com capacidade de compreender e interpretar a linguagem humana, tal capacidade pode ser utilizada em diferentes contextos que exijam essa compreensão.

Tendo isso em vista, este estudo qualitativo, caracterizado como uma Revisão Sistemática da Literatura guiada pelo Protocolo PRISMA, objetivou compreender as principais situações onde o PLN pode ser utilizado para análise de texto na área de Administração.

Os resultados demonstram que o PLN pode ser utilizado com eficiência para análise de sentimento, especialmente, em avaliações e comentários de consumidores e clientes online. Além disso, pesquisas organizacionais com enfoque tanto no público e conteúdo internos, como externos, também podem se beneficiar com sistemas de PLN.

Processos de análise preditiva com intuito de vislumbrar e prospectar situações e comportamentos futuros podem ser potencializados com o uso de sistemas de PLN. Bem como, classificação da dificuldade de questões e conteúdos escolares e universitários.

Este estudo contemplou apenas artigos da área de Administração de periódicos em bases de dados internacionais. Para pesquisas futuras, sugere-se que sejam realizadas pesquisas quantitativas envolvendo as temáticas de PLN e análise de texto, além de estudos de caso considerando o *framework* elaborado nesta pesquisa.

REFERÊNCIAS

- Alantari, H. J., Currim, I. S., Deng, Y., & Singh, S. (2022). An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*, 39, 1-19. <https://doi.org/10.1016/j.ijresmar.2021.10.011>
- Antonio, N., Almeida, A. M., Nunes, L., Batista, F., & Ribeiro, R. (2018). Hotel online reviews: creating a multi-source aggregated index. *International Journal of Contemporary Hospitality Management*, 30(12), 3574-3591. <https://doi.org/10.1108/IJCHM-05-2017-0302>
- Araújo, F. N. M., Queiroz, M. V. A. B., & Aouar, W. A. E. (2023). Telemedicina, telessaúde e digitalização: revisão sistemática utilizando o protocolo prisma (2020-2022). *Revista de Administração em Saúde (Online)*, 23(91). <http://dx.doi.org/10.23973/ras.91.350>

- Baharuddin, F., & Naufal, M. F. (2023). Fine-Tuning IndoBERT for Indonesian Exam Question Classification Based on Bloom's Taxonomy. *Journal of Information Systems Engineering and Business Intelligence*, 9(2). <http://dx.doi.org/10.20473/jisebi.9.2.253-263>
- Barakhnin, V. B., Kozhemyakina, O. Y., Mukhamediev, R. I., Borzilova, Y. S., & Yakunin, K. O. (2019). The design of the structure of the software system for processing text document corpus. *Business Informatics*, 13(4). <https://doi.org/10.17323/1998-0663.2019.4.60.72>
- Bardin, L. (2011). *Análise de conteúdo*. São Paulo, SP: Edições.
- Bharti, N., & Kumar, S. (2019). Sentiment Analysis of Impact of Social Platforms on the Market Share of a Company. *International Journal of Recent Technology and Engineering*, 8(2). <https://doi.org/10.35940/ijrte.B1594.078219>
- Caseli, H. M., & Nunes, M. G. V. (2023). *Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português*. São Carlos, SP: Brasileiras em Processamento de Linguagem Natural (BPLN). Recuperado de: brasileiraspln.com/livro-pln
- Charniak, E., & Mcdermott, D. (1985). *A Bayesian Model of Plan Recognition*. Massachusetts, New England: Addison-Wesley.
- García-Crespo, A., Colomo-Palacios, R., Gómez-Berbrís, J. M., & Ruiz-Mezcua, B. (2010). SEMO: a framework for customer social networks analysis based on semantics. *Journal of Information Technology*, 25, 178-188. <https://doi.org/10.1057/jit.2010.1>
- Gomes, D. S. (2010). Inteligência Artificial: Conceitos e Aplicações. *Revista Olhar Científico*, 1(2). Recuperado de: https://www.professores.uff.br/screspo/wp-content/uploads/sites/127/2017/09/ia_intro.pdf
- Greselin, F., Zitikis, R., Vicenzo, D., & Piacenza, F. (2022). A Text Analysis for Operational Risk Loss Descriptions. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4286208>
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Massachusetts, New England: The MIT Press.
- Jurafsky, D., & Martin, J. H. (2019). *Speech and language processing* (3rd ed.). London, United Kingdom: Pearson.
- Karvetski, C. W., Meinel, C., Maxwell, D. T., Lu, Y., Mellers, B. A., & Tetlock, P. E. (2022). What do forecasting rationales reveal about thinking patterns of top geopolitical forecasters? *International Journal of Forecasting*, 38, 688-704. <https://doi.org/10.1016/j.ijforecast.2021.09.003>
- Kurzweil, R. (1990). *The Age of Spiritual Machines*. Massachusetts, New England: The MIT Press.
- Liu, H. The role of artificial intelligence in robotics. *Robotics Business Review*, 2020. Disponível em: <https://www.roboticsbusinessreview.com/opinion/the-role-of-artificial-intelligence-in-robotics/> Acesso em: 03/07/2024.
- Miric, M., Jia, N., & Huang, K. (2022). Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents. *Strategic Management Journal*, 44, 491-519. <https://doi.org/10.1002/smj.3441>
- Pinho, C. M. A. (2021). *Análise de textos com aplicação de técnicas de inteligência artificial: estudo comparativo para classificação de fuga ao tema em redações*. (Dissertação de Mestrado, Programa de Pós-Graduação em Informática e Gestão do Conhecimento, Universidade Nove de Julho, São Paulo, SP). Recuperado de <http://bibliotecatede.uninove.br/handle/tede/3092>
- Poole, D., Mackworth, A. K., & Goebel, R. (1998). *Computational Intelligence: A Logical Approach*. Oxford, Ohio: Oxford University.
- Poschmann, P., Goldenstein, J., Büchel, S., & Hahn U. (2023). A Vector Space Approach for Measuring Relationality and Multidimensionality of Meaning in Large Text Collections. *Organizational Research Methods*, 1-31. <https://doi.org/10.1177/10944281231213068>

- Pupo, A. S. (2023). *Sinais fracos e o planejamento organizacional: contribuições para o planejamento estratégico por meio do processamento de linguagem natural* (Tese de Doutorado, Programa de Pós-Graduação em Administração, Universidade de São Paulo, São Paulo, SP). Recuperado de <https://www.teses.usp.br/teses/disponiveis/12/12139/tde-20122023-205208/>
- Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). London, United Kingdom: Pearson.
- Santos, D. (2001). *Processamento da linguagem natural: uma apresentação através das aplicações*. In: Ranchhod (org.) 2001. Tratamento das Línguas por Computador. Uma introdução à Linguística Computacional e suas aplicações. Lisboa, Portugal: Caminho.
- Todd, A., Bowden, J., & Moshfeghi, Y. (2024). Text-based sentiment analysis in finance: Synthesising the existing literature and exploring future directions. *Intelligent Systems in Accounting, Finance and Management*, 31. <https://doi.org/10.1002/isaf.1549>
- Weiss, M. C. (2019). Sociedade sensoriada: a sociedade da transformação digital. *Estudos Avançados*, 33(95). <https://doi.org/10.1590/s0103-4014.2019.3395.0013>