
Análise de algoritmos de aprendizado de máquina para mineração de conversas e aperfeiçoamento de chatbots

Analysis of machine learning algorithms for conversation mining and chatbot improvement

Eliane Somavilla^{1*}, Marcilde Sabadin¹, Vinicius Ramos¹, Giovani Gracioli¹

RESUMO

Esse trabalho **objetiva**, por meio da aplicação dos algoritmos *Latent Dirichlet Allocation (LDA)*, *NGram* e *Bag of Words (BOW)*, analisar como o aprendizado de máquina usado na mineração de conversas pode ajudar no aperfeiçoamento de chatbots através da identificação da intenção do cliente ao entrar em contato com uma empresa. A **metodologia** será qualitativa, por meio de pesquisa bibliográfica e do uso da ferramenta Knime, a fim de apoiar a implementação de algoritmos de aprendizado de máquina, em base de dados cedida por uma empresa desenvolvedora de software para a área de *Customer Experience (CX)*. Os **resultados** derivam de uma comparação entre os algoritmos considerando o peso (weight) das palavras e frases chave e validada pelo Analista de Domínio (AD) da empresa. O extrator de tópicos LDA apresenta o melhor resultado sobre as intenções dos usuários e o algoritmo *NGram* apresenta características complementares significativas para as intenções. Ambos algoritmos trazem restrito aprendizado de máquina, e demandam constante monitoramento e acompanhamento do AD, devido a evolução do processamento da linguagem natural. O algoritmo *BoW* apresentou limitações neste estudo.

Palavras-chave: Chatbots; Inteligência Artificial; LDA; Bag of Words; *NGram*;

ABSTRACT

This work aims, through the application of *Latent Dirichlet Allocation (LDA)*, *NGram* and *Bag of Words (BOW)* algorithms, to analyze how machine learning used in conversation mining can help improve chatbots by identifying customer intent when contacting a company. The methodology will be qualitative, through bibliographical research and the use of the Knime tool, in order to support the implementation of machine learning algorithms, in a database provided by a software developer company for the *Customer Experience (CX)* area. . The results derive from a comparison between the algorithms considering the weight (weight) of the key words and phrases and validated by the *Domain Analyst (AD)* of the company. The LDA topic extractor presents the best result on the users' intentions and the *NGram* algorithm presents significant complementary characteristics for the intentions. Both algorithms bring restricted machine learning, and demand constant monitoring and monitoring of AD, due to the evolution of natural language processing. The BOW algorithm showed limitations in this study.

Keywords: Chatbots; Artificial Intelligence; LDA; Bag of Words; *NGram*;

¹ Universidade Federal de Santa Catarina.

*Email: eliane.somavilla@gmail.com

INTRODUÇÃO

A pandemia de Covid-19 tem acelerado os processos de transformação digital nas empresas privadas e entidades públicas. A digitalização das empresas aumenta a importância de canais digitais para contato com clientes e estimula o consumo de produtos tecnológicos compatíveis com Internet das Coisas (IoT), Inteligência Artificial (IA) e Big Data. Ou seja, cada vez mais pessoas interagem com empresas usando mecanismos de comunicação acessíveis de qualquer lugar a qualquer momento.

A atual popularidade de chatbots, “programas computacionais que emulam uma conversa como as estabelecidas entre seres humanos” (COMARELLA; CAFÉ, 2008, p. 55), também designados como um agente de software (robô), permite a interação homem-máquina usando linguagem natural (ARSOVSKI, 2019).

Neste contexto, diversos temas ganharam destaque como: o desenho de conversas e fluxos de diálogos, a análise de interação entre humano e máquina, a estruturação de bases de conhecimentos, o mapeamento de intenções e ações de usuários, além da criação de “chatbots personas”, que é uma representação visual dada ao chatbot para ampliar a empatia humana. De acordo com Reina e Cruz (2020), as discussões e trocas de conhecimento se tornaram fundamentais para apoiar a maturação gradativa de chatbots, processo que foi nomeado de Curadoria de *Chatbots*. Entretanto, apesar da crescente dedicação de esforços profissionais na Curadoria de *Chatbots*, há poucos estudos acadêmicos sobre esta área. Profissionais de diferentes especialidades obtiveram diversos conhecimentos através da prática das novas atividades demandadas, o que gerou, ao longo do tempo, conhecimento especializado (REINA; CRUZ, 2020).

Outro ponto importante refere-se ao fato de que a inteligência artificial (IA), ou inteligência demonstrada por máquinas ao executar tarefas consideradas inteligentes, está ganhando cada vez mais atenção dos líderes de negócios. Como principal ferramenta de IA, *chatbots* têm sido cada vez mais utilizados pelas empresas para apoiar o atendimento ao cliente (WANG; LIN; SHAO, 2020).

Considerando este cenário, é importante avaliar o desempenho do chatbot para que seja possível entregar ao usuário uma ferramenta de atendimento que tenha qualidade. Todavia, faltam estudos sobre quais são os principais indicadores que devem ser avaliados e ferramentas que apoiem na tomada de decisão (KULIGOWSKA, 2015).

Dessa forma, o objetivo deste trabalho é avaliar como o uso de técnicas de aprendizado de máquina pode ajudar a aperfeiçoar chatbots de fluxo fechado, ou seja,

sem inteligência artificial. Para alcançar este objetivo, será contextualizada a importância de chatbots e sua aplicação na área de atendimento ao cliente, etapa fundamental para obtenção de bons resultados relacionados à experiência do cliente (*Customer Experience*). Serão também apresentados dados sobre a demanda do mercado por chatbots com inteligência artificial, os chatbots de fluxo aberto e híbridos.

Após um breve relato dos principais desafios do processamento de linguagem natural (PLN), serão aplicados e comparados os resultados dos algoritmos LDA (extrator de tópicos), NGram e *Bag of Words* (BoW) em uma base de dados privada cedida por uma empresa de tecnologia desenvolvedora de uma plataforma de produtos focados em *Customer Experience*. Os algoritmos escolhidos usam incorporação de palavras (*Word Embedding*) ou clusterização (agrupamento) e podem alcançar bons resultados através da análise de palavras chaves ou frases chaves (KHAN *et. al*, 2022).

Essa aplicação será útil para a curadoria na etapa de cadastro de intenções dos usuários para treinamento de *Chatbots*, além de melhorar a experiência do cliente no contato com *chatbots* de fluxo aberto ou híbrido.

DESENVOLVIMENTO DO TRABALHO

Referencial Teórico

De acordo com Lemon e Verhoef (2016), a Experiência do Cliente (CX) envolve suas reações cognitivas, emocionais, sensoriais, sociais e espirituais a todas as interações com uma empresa. Além disso, foca em resolver problemas dos clientes, responder a perguntas e moldar as percepções de marca dos usuários.

Dessa forma, o atendimento é apenas uma parte do CX, que corresponde a soma de todos os pontos de contato que um cliente tem com uma empresa, incluindo o atendimento, marketing, produto, vendas e outros departamentos.

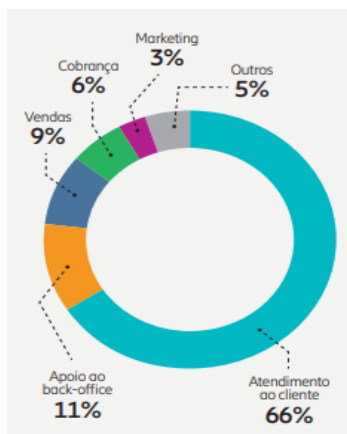
Uma das maiores preocupações da Curadoria de *Chatbots* é a da satisfação do cliente, que é influenciada pela utilidade e facilidade de uso do chatbot durante um atendimento. A humanização dos robôs está aliada à comunicação, trazendo empatia e proximidade entre empresa e cliente. A humanização de chatbots, em um nível adequado, tem se mostrado um fator importante na aceitação dos clientes ao robô como opção de atendimento, resultando em trabalho constante para a Curadoria de *Chatbots* (REBECCHI, 2020).

Para Lemon e Verhoef (2016), é preciso entender quais fatores podem influenciar

as experiências dos clientes, principalmente entre os fatores hedônicos (ou seja, imagens mentais, entretenimento e estética). Tais fatores afetam os conhecimentos do cliente, o que acaba influenciando na satisfação e, logo, na tomada de decisão.

De acordo com a revista *Mobile Time* (2022), plataforma de conteúdo sobre a indústria móvel e o setor de robôs de conversação, a qual produz notícias, pesquisas e eventos relacionados a chatbots, a finalidade com maior demanda por chatbots é o atendimento ao cliente, conforme é possível observar na Figura 1.

Figura 1 - Áreas com maior uso de chatbots



Fonte: Revista Mobile Time (2022)

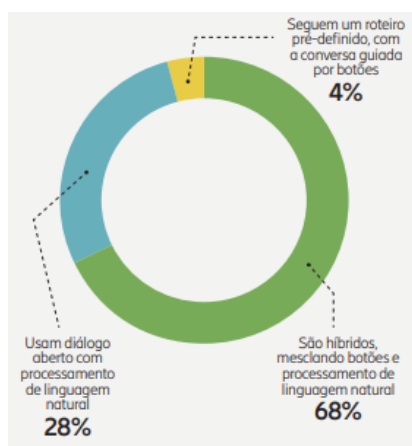
O apoio ao *backoffice*, ou seja, chatbots que apoiam os funcionários de uma empresa, como por exemplo os atendentes de *call center*, aparece como a segunda finalidade mais popular entre os chatbots desenvolvidos no Brasil.

Ainda, de acordo com a revista *Mobile Time* (2022), a maioria dos chatbots da atualidade são híbridos, ou seja, são implementados para atender os clientes através de algoritmos de inteligência artificial capazes de compreender tanto a linguagem natural quanto um fluxo fechado onde o usuário informa sua intenção por meio da escolha de opções listadas em botões, conforme é possível observar na Figura 2.

No intuito de atender de forma ágil os clientes, os *Chatbots* têm utilizado modelos de processamento de linguagem natural (PLN) para apoiar na identificação das intenções de seus usuários ao entrar em contato com a empresa naturalmente (PRZEGALINSKA *et. al*, 2019). As intenções dos usuários são solicitações colhidas em frases escritas durante um diálogo. Normalmente acompanham um verbo, como por exemplo, Pagar, Imprimir, Cancelar, Trocar, entre outros. Tais intenções possuem atributos que detalham a solicitação do usuário como: “Imprimir o Boletim”, “Pagar o Boletim”, ”Trocar Endereço”. Ou seja, os recursos atuais de aprendizado de máquina incluem a geração e compreensão natural de fala e escrita (PRZEGALINSKA *et. al*, 2019). Esse cenário

perpassa por desafios decorrentes do processamento de linguagem natural, especialmente a compreensão do português e seus variados significados.

Figura 2 - ChatBots com fluxo aberto (com IA) x fechado (sem IA) x híbrido



Fonte: Revista Mobile Time (2022)

Processamento de Linguagem Natural

O Processamento de Linguagem Natural (PLN) é uma abordagem computacional para analisar textos de acordo com um conjunto de teorias, regras e tecnologias (LINDEN *et al.*,2003).

Segundo (BRAUN *et al.*,2018), o PLN é um ramo da IA que tem por objetivo interpretar e gerar textos em uma linguagem natural, ou seja, se dedica ao desenvolvimento de máquinas com a capacidade de entender e responder textos e palavras faladas (voz) da mesma forma que os seres humanos. No entanto, entender a posição que o processamento de Língua Portuguesa ocupa nesse cenário e quais as melhores direções a serem tomadas para o seu desenvolvimento pleno (FINGER, 2021), torna-se um desafio.

O autor Gadelha (2019), resume os desafios da PLN e dos *Chatbots* enfrentados pelas empresas em quatro pontos: o domínio de sintaxe da linguagem, que pode variar dependendo do contexto do diálogo, visto que humanos não interagem em uma ordem definida; a interface de conversação, que interfere na experiência do usuário e pode causar insatisfação àqueles que dispõem de pouco tempo e esperam respostas rápidas; o terceiro desafio se refere ao teste dos *chatbots*, que estão em constante evolução, principalmente em seus modelos de linguagem natural e cuja precisão de compreensão deve ser frequentemente verificada; o quarto ponto se refere a necessidade de dados, visto que é vital para uma empresa realizar o enriquecimento de chatbots com dados significativos para compor respostas que representarão a identidade da marca ao seu público-alvo.

Considerando estes quatro pontos, podemos compreender a importância de pessoas com conhecimento aprofundado no negócio da empresa para apoiar em estratégias e enfrentar estes desafios.

Aprendizado de máquina interativo

Atualmente, para Wondimu, Buche, Visser (2022, p.12), “o aprendizado de máquina interativo (iML) tem ganhado cada vez mais atenção dos pesquisadores por conta de sua modalidade *"human-in-the-loop"*. Os usuários finais e especialistas de domínio participam do processo de construção do modelo alimentando interativamente parâmetros de treinamento, inspecionando saídas do modelo e fornecendo feedback sobre resultados intermediários.

Uma importante preocupação é quanto a qualidade da base de conhecimento dos chatbots. Para chatbots de fluxo fechado, normalmente baseados em regras, o esforço de construção da base de conhecimento é realizado manualmente por humanos. Diferentemente, *chatbots* que utilizam inteligência artificial podem aprender com base nos exemplos pré cadastrados (ARSOVSKI *et al*, 2019). Consequentemente, podemos entender que construir e manter uma base de conhecimento detalhada é um desafio para a Curadoria de *Chatbots* e contar com uma especialista de domínio (*human-in-the-loop*) para apoiar nesta tarefa pode ser um diferencial.

Para se criar chatbots híbridos e bem treinados para responder ao usuário, é necessário identificar a intenção do usuário ao entrar em contato com uma empresa ou marca por meio de um canal como chat, telefone ou whatsapp. Ao identificar o que o cliente deseja, os chatbots podem ser treinados para detectar a principal intenção dos usuários, viabilizando a criação de um modelo com vários exemplos de como o usuário expressa sua intenção na linguagem natural. Para apoiar na criação deste modelo, há uma gama de algoritmos de inteligência artificial já desenvolvidos e consolidados na área de aprendizado de máquina, além disso, o especialista de domínio pode participar do processo para validar os resultados gerados.

Algoritmos

Conforme a quantidade de dados textuais não estruturados aumenta, a necessidade de processar e extrair conhecimento automaticamente dos textos torna-se cada vez mais importante. (KHAN *et.al*, 2022).

As palavras ou frases mais significativas que descrevem o documento textual em um nível abstrato geram as palavras-chaves ou frases-chaves que são o resultado de uma abordagem de indexação automática de texto que resume os principais temas, conceitos, ideias ou argumentos em um número reduzido de frases e palavras. Esta indexação favorece sistemas que buscam quais palavras-chave são frequentemente usadas para consulta, extração de resumo de um documento de identificação ou agrupamento de tópicos (KHAN *et.al*, 2022).

Uma palavra-chave ou frase-chave pode ser entendida como uma palavra que descreve o conteúdo de um documento, no todo ou em parte. Uma palavra-chave é um unigrama, enquanto uma frase-chave é um *n-gram*, ou seja, mais de uma palavra. Por exemplo, “família” é uma palavra-chave, e "membros da família" é uma frase-chave. Muitas vezes a frase-chave ou *n-gram* é preferencialmente usada em algoritmos por conter mais informações e significados do que uma palavra-chave ou unigrama (KHAN *et.al*, 2022). As frases e palavras-chaves são usadas neste estudo para interpretar a intenção dos usuários.

Além disso, o significado contextual pode variar dependendo do texto e do ambiente. Por exemplo, a palavra “banco” pode significar uma organização bancária, ou pode ser um assento. Portanto, o contexto é essencial para entender o significado de uma palavra. O contexto também pode ser usado para inferir o significado de palavras desconhecidas. Uma das técnicas utilizadas nos algoritmos como o *Bag of Words* é a análise da palavra anterior ou posterior ao termo desejado (KHAN *et.al*, 2022).

Existem várias categorias de métodos usados para encontrar palavras-chave e frases-chave. Neste estudo focaremos em métodos baseados em *cluster* (agrupamento) e métodos baseados em incorporação (*embedding*).

Métodos baseados em *cluster* ou agrupamento

As abordagens baseadas em *cluster* ou agrupamento aplicam algoritmos capazes de agrupar frases ou palavras candidatas em clusters de tópicos, então os tópicos mais representativos de cada cluster são selecionados como frase-chave ou palavra-chave (KHAN *et.al*, 2022).

LDA

O algoritmo LDA (*Latent Dirichlet Allocation*) é um modelo probabilístico muito utilizado para identificar os tópicos ou palavras-chaves de um documento de texto. Para

um *corpus* de documentos, “cada documento é descrito como uma mistura de tópicos latentes e cada tópico é modelado como uma distribuição discreta sobre o vocabulário dos documentos” (LENG *et.al*, 2020).

Considerando a ampla aplicabilidade do algoritmo LDA, de acordo com os estudos dos autores Onan, Bulut e Korukoglu (2016) e as recentes pesquisas realizadas por Khan *et al* (2022), as quais apontam resultados relevantes quanto a identificação de palavras e frases-chaves na análise de diferentes tópicos que compõem um texto, os autores do presente artigo optaram por analisá-lo frente ao tema emergente de interação entre humanos e robôs (*chatbots*).

Métodos baseados em incorporação (*embedding*)

As palavras são representadas numericamente de uma forma que é facilmente processada por algoritmos. Alguns algoritmos usam a semelhança do cosseno na incorporação de palavras e frases candidatas e a comparam com a incorporação do cosseno do documento para fornecer uma classificação adequada (KHAN *et.al*, 2022).

Bag Of Words (BoW) e NGram

Bag of Words (BoW) é uma técnica de vetorização de texto que converte o texto em vetores de comprimento finito. Cada coluna de um vetor representa uma palavra. Os valores em cada célula de uma linha mostram o número de ocorrências de uma palavra em uma frase. A configuração mais comum do BoW usa todas as palavras de um documento para identificar a presença de uma palavra específica. O método mais simples para identificar se uma palavra está presente em um documento é o binário em que o peso é um - se a palavra correspondente está presente no documento – ou zero caso contrário (FELDMAN; SANGER, 2016).

O Algoritmo *NGram* comporta-se similarmente ao BoW, mas usa uma sequência de palavras configurável. Através do uso de *n-grams* pode-se diminuir a probabilidade de erros de entendimento dependendo do contexto (FELDMAN; SANGER, 2016). Neste estudo usamos 2 palavras em cada *n-gram*.

Considerando que, de acordo com Harris (2015), o *Bag Of Words (BoW)* utiliza a abordagem mais popular para representar dados textuais para alimentar algoritmos de aprendizado de máquina e que, de acordo com Feldman e Sanger (2016), o método binário de avaliação da presença de uma palavra em um texto torna os algoritmos BoW e *Ngram* de simples entendimento e ampla utilização, os autores do presente artigo

optaram por analisá-lo frente ao tema emergente de interação entre humanos e robôs.

A implementação destes três algoritmos ocorreu conforme descrito na Seção Materiais e Métodos, com o apoio da empresa Hi Platform² para obtenção dos dados.

MATERIAIS E MÉTODOS

Para fundamentar a pesquisa, serão utilizados procedimentos técnicos de pesquisa bibliográfica nas publicações científicas como artigos, teses e dissertações disponibilizadas nos últimos 8 anos na base da CAPES. Este conjunto de bibliografia compõe o *corpus* desta proposta de pesquisa (ZANELLA, 2013).

Com abordagem qualitativa, Brenner (1981, p.51), considera que a "metodologia qualitativa" é aquela que "... permite ao pesquisador estar próximo dos dados, desenvolvendo, de uma ou outra forma, os componentes analíticos, conceituais e categóricos da explicação a partir dos próprios dados". Portanto, nesse método, os resultados permitirão conhecer a problemática e permitirão a análise e comparação entre os algoritmos implementados no intuito de identificar o melhor resultado. A base de dados conterá dados anonimizados e cedidos pela empresa Hi Platform² e os algoritmos serão implementados com o apoio da ferramenta Knime³.

Base de Dados

A base de dados usada no presente estudo foi cedida pela empresa Hi Platform² e corresponde a diálogos da ferramenta Hi Chat entre usuários e um cliente anônimo (empresa) do segmento de *e-commerce*. Os dados que compõem a base se referem apenas às mensagens enviadas pelos usuários, *datas* e o código individual de cada diálogo. O código é uma forma de identificar um diálogo específico entre um usuário e a empresa dentro das centenas de atendimentos que podem ser realizados diariamente por meio da ferramenta Hi Chat.

As mensagens dos operadores da empresa não fazem parte do *corpus* de mensagens deste estudo, visto que o objetivo é identificar apenas a intenção informada através do contato do usuário. Ao analisar a base, observou-se que logo no início do

² HI PLATAFORM. Sobre. Disponível em: <<https://www.hiplatform.com/sobre>>. Acesso em: 23 out. 2022.

³ KNIME ANALYTICS PLATFORM. Open for innovation. Switzerland. Disponível em: <<https://www.knime.org/>>. Acesso em: 08 nov. 2022.

diálogo o usuário informa o que deseja, por isso, optou-se por considerar apenas as 3 primeiras mensagens como limite de pesquisa.

Os dados estudados correspondem a 7 dias de atendimento no mês de Novembro de 2022. Quaisquer dados que possam representar a empresa cliente foram anonimizados pela Hi Platform.

Knime Analytics Platform

O *KNIME⁴ Analytics Platform* é um sistema de código livre e aberto, de análise de dados, baseado na plataforma Eclipse (Java) com possibilidade de customizar e implementar novos módulos em um ambiente visual, o que permite sua aplicação em ambientes de produção comercial, ensino e/ou pesquisa.

O *KNIME* permite que os usuários criem visualmente fluxos de dados (*pipelines*), executem seletivamente alguns ou todos os passos da análise e depois inspecionem os resultados e modelos, usando visualizações interativas. A plataforma iniciou com versão 2.1 e atualmente, encontra-se na versão 4.6.3.

Assim, o *pipeline* é construído de forma dinâmica, nó após nó, sendo configurado, executado e inspecionado. O documento vai se formando como uma rede de caixas, interligadas, podendo ter uma ou mais entradas e uma ou mais saídas, como será apresentado neste trabalho na fase de aplicação de algoritmos.

Pré-Processamento

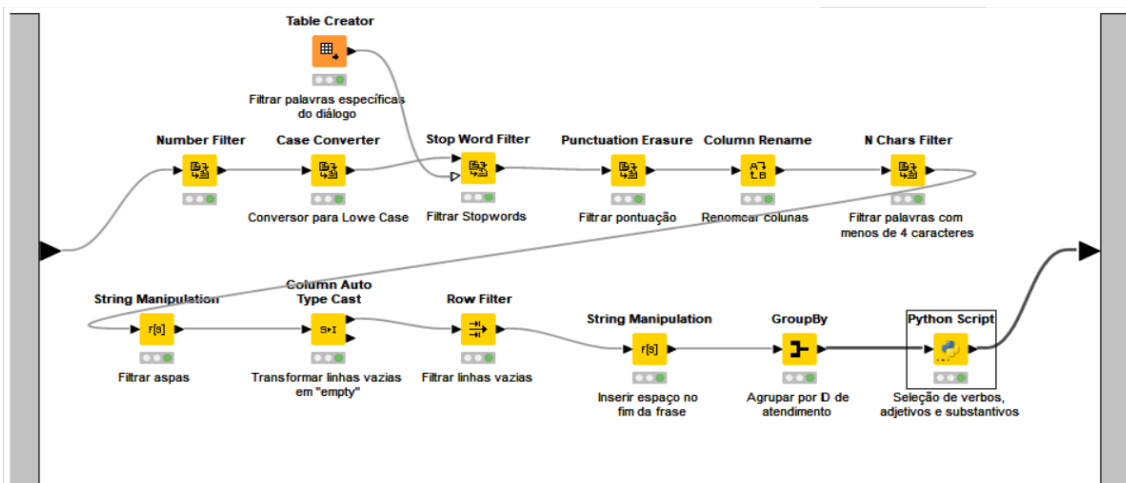
Os dados foram importados no *Knime* por meio de um leitor de arquivo Excel e representam mais de 50 mil mensagens, as quais foram dialogadas/trocadas entre usuários e a empresa em 7 dias de atendimento, utilizando a ferramenta Hi Chat. Inicialmente, cada mensagem da base de dados foi convertida em ‘Documento’ por meio do *node* ‘*String to Document*’. Na sequência realizou-se um pré-processamento de limpeza e normalização, conforme Figura 3, que incluiu:

- Limpeza de caracteres numéricos, pontuação;
- Conversor de caracteres para letras minúsculas;

⁴ KNIME ANALYTICS PLATFORM. Open for innovation. Switzerland. Disponível em: <<https://www.knime.org/>>. Acesso em: 08 nov. 2022.

- Filtragem de palavras com menos de 4 caracteres e de *stopwords*, ou seja, termos comuns em diálogos do idioma português nativamente identificados pelo *Knime* e também através de uma tabela de termos adicionais identificados na análise do diálogo, como mensagens de cumprimentos iniciais, despedidas e agradecimentos.
- Linhas que ficaram vazias no final do processo foram desconsideradas;
- Mensagens foram agrupadas em uma única linha de acordo com o código do atendimento.
- Aplicação de um *script Python* para identificar a classe de palavras e considerar apenas verbos, adjetivos e substantivos.

Figura 3 - Pré-Processamento dos dados



Fonte: elaborado pelos autores (2023)

Para utilizar o *script Python* criamos um ambiente *Python Deep Learning* versão 3.6 na ferramenta *Knime* e, através da plataforma *Anaconda*, instalamos o *plugin spaCy*⁵. O *spaCy* se destaca em tarefas de extração de informações em larga escala e, atualmente, oferece suporte para os diversos idiomas, entre eles, o português. Nos cinco anos desde seu lançamento, o *spaCy* se tornou um padrão da indústria com um enorme ecossistema de componentes. Os pacotes treinados do *spaCy* podem ser instalados como pacotes *Python*. Isso significa que eles são um componente versionado do *Knime*, assim como qualquer outro. No estudo, usamos o pacote *pt_core_news_sm*.

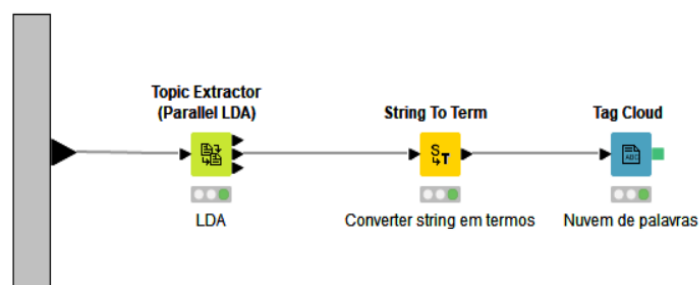
⁵ spaCy. Models & Languages. Disponível em: <<https://spacy.io/usage/models>>. Acesso em: 26 nov. 2022.

Na sequência a *string* gerada no resultado do agrupamento foi convertida em ‘Documento’ para que pudesse ser processado pelos algoritmos utilizados.

Aplicação de Algoritmos

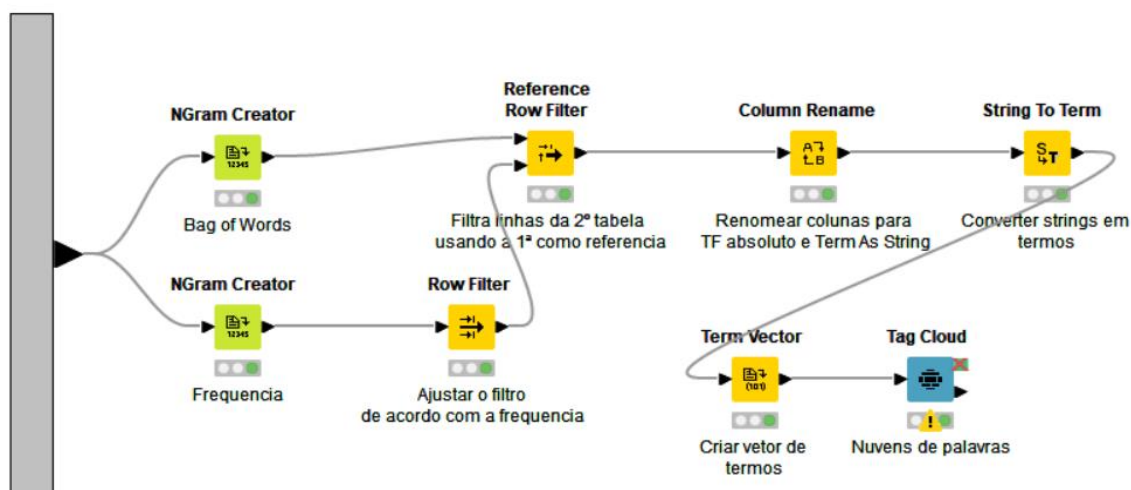
O primeiro algoritmo aplicado foi o extrator de tópicos LDA, cuja configuração foi efetuada para gerar 10 tópicos para cada documento, considerando as 5 principais palavras a serem extraídas por cada tópico. A *string* gerada foi convertida em ‘Termos’ os quais foram apresentados em uma Nuvem de Palavras que apresenta os 50 principais termos identificados, conforme Figura 4.

Figura 4 - Implementação do extrator de tópicos LDA no knime



Fonte: elaborado pelos autores (2023)

Figura 5 - Implementação do algoritmo Ngram no knime



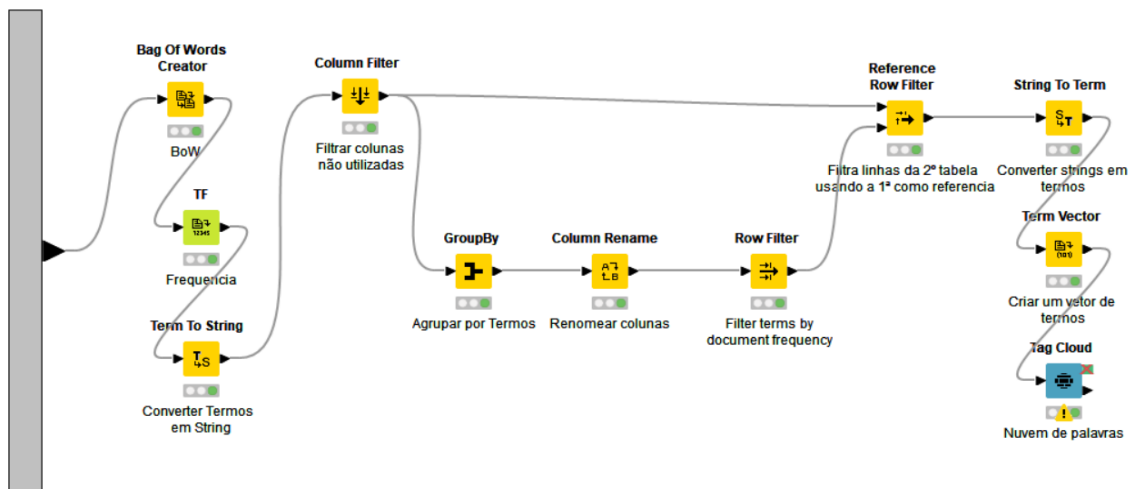
Fonte: elaborado pelos autores (2023)

O segundo algoritmo aplicado foi o NGram. O primeiro *node* cria um *NGram* composto de 2 palavras e gera como saída um ‘saco de palavras com *n-grams*’. O segundo *node* contabiliza a frequência de cada *n-grams* em cada documento e também em todo o *corpus* de documentos. Na sequência, um filtro é aplicado para que sejam considerados

apenas *n-grams* com frequência maior que 5. Os resultados dos dois *nodes* de *Ngram* (*NGgram Bag of Words* e frequência) são novamente filtrados comparando os termos gerados. A *string* gerada é convertida em um ‘Vetor de Termos’ os quais foram apresentados em uma Nuvem de Palavras com os 50 principais termos identificados compostos por 2 *n-grams*, conforme Figura 5.

O terceiro algoritmo aplicado foi o *Bag Of Words* (BoW). O primeiro *node* aplica o algoritmo conforme a configuração nativa do *Knime*. O segundo *node* contabiliza a frequência de cada termo em cada documento e, após a conversão para *string*, os termos são agrupados e contabilizados novamente. O valor da frequência é calculado dividindo a frequência absoluta de um termo em um documento pelo número de todos os termos deste documento. Após, um filtro é aplicado para que sejam considerados termos com frequência maior que 5. O resultado é novamente filtrado comparando com os termos gerados originalmente pelo algoritmo BoW. A *string* gerada é convertida em um ‘Vetor de Termos’ os quais foram apresentados em uma Nuvem de Palavras com os 50 principais termos identificados, conforme Figura 6.

Figura 6 - Implementação do algoritmo Bag Of Words (BoW) no knime



Fonte: elaborado pelos autores (2023)

RESULTADOS E DISCUSSÃO

Após análises realizadas, apresentamos o resultado dos algoritmos estudados no intuito de apontar soluções de IA baseada em PLN para aperfeiçoamento de *chatbots*.

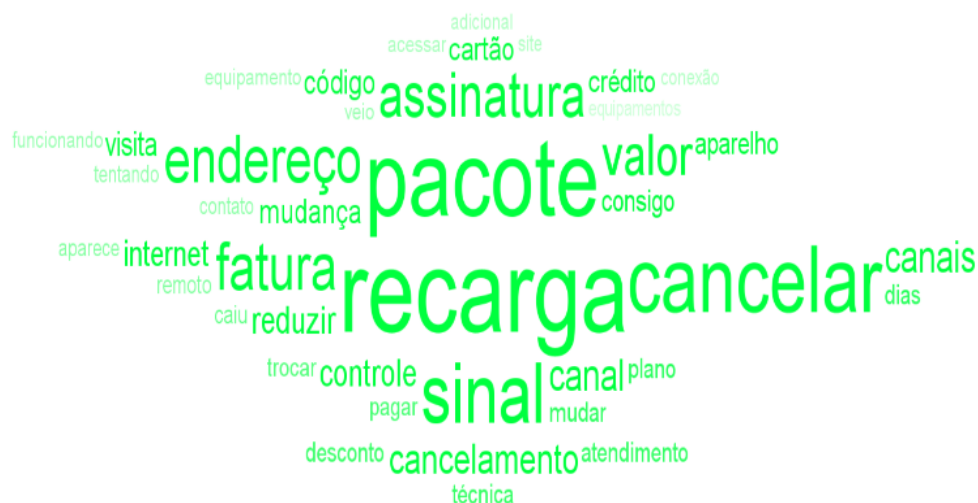
LDA

O algoritmo extrator de tópicos apresentou o melhor resultado, segundo análises realizadas e validadas com o administrador da Hi Platform. Foram testadas configurações iniciadas com 20 tópicos e 10 principais termos, após, são decrescidas a quantidade de tópicos em uma escala de 5 em 5 até atingir um valor agregado representativo com 10 tópicos e 5 principais termos ou palavras chave. Algumas configurações apresentaram um resultado com baixo valor semântico, como no caso de 20 tópicos e 10 termos.

Observou-se que as 5 palavras mais representativas de um tópico são suficientes para entender a intenção de um cliente. As outras 5 palavras apresentadas na configuração com 10 termos não são significativas, ou seja, não alteram a interpretação sobre a intenção.

Por meio do *script Python* gerado com o pacote treinado *pt_core_news_sm* e instalado com apoio do *pluin scpaCy*. apenas as palavras pertencentes às classes de verbos, adjetivos e substantivos foram consideradas. Todavia, o resultado gerado utiliza o português de Portugal, por isso algumas classificações podem ser diferentes do português do Brasil. Ademais, alguns nomes próprios ou escritas de forma equivocada ou com baixo valor semântico podem ser identificados como pertencentes às classes de verbos, adjetivos ou substantivos e apresentadas na Nuvem de Palavras quando deveriam ser ignoradas. As palavras chaves mais representativas estão na Figura 7.

Figura 7 - Nuvem de Palavras - LDA - 10 *Topic* e 5 *Words*



Fonte: elaborado pelos autores (2023)

Conforme a Figura 8, quanto mais representativo o termo ou palavra chave, maior o peso (*weight*) calculado pelo algoritmo LDA. Ao ordenar os termos por peso, observa-se uma ênfase maior nos termos 'Pacote', 'Cancelar' e 'Recarga' com pesos calculados

entre 3200 e 4400. Observamos que a diminuição de tópicos influencia na quantidade de repetição de termos e o peso (*Weight*) aumenta à medida que ocorre o agrupamento de tópicos, o que demonstra que o algoritmo está conseguindo compreender tópicos similares .

Figura 8 - Lista de termos ou palavras chaves - LDA

Row ID	S Topic id	S Term	D Weight	Row ID	S Topic id	S Term	D Weight
Row0	topic_0	pacote	3,296	Row25	topic_5	recarga	4,334
Row1	topic_0	reduzir	917	Row26	topic_5	sinal	1,302
Row2	topic_0	canais	557	Row27	topic_5	cartão	703
Row3	topic_0	valor	468	Row28	topic_5	crédito	638
Row4	topic_0	trocar	410	Row29	topic_5	dias	500
Row5	topic_1	consigo	590	Row30	topic_6	internet	858
Row6	topic_1	contato	300	Row31	topic_6	aparelho	628
Row7	topic_1	tentando	292	Row32	topic_6	funcionando	279
Row8	topic_1	acessar	246	Row33	topic_6	equipamento	270
Row9	topic_1	site	226	Row34	topic_6	sinal	227
Row10	topic_2	endereço	2,106	Row35	topic_7	pacote	640
Row11	topic_2	mudança	766	Row36	topic_7	canal	573
Row12	topic_2	visita	748	Row37	topic_7	plano	266
Row13	topic_2	mudar	490	Row38	topic_7	canais	239
Row14	topic_2	técnica	466	Row39	topic_7	adicional	232
Row15	topic_3	sinal	1,385	Row40	topic_8	fatura	2,110
Row16	topic_3	código	742	Row41	topic_8	valor	1,273
Row17	topic_3	canal	597	Row42	topic_8	desconto	458
Row18	topic_3	canais	383	Row43	topic_8	pagar	436
Row19	topic_3	aparece	318	Row44	topic_8	veio	276
Row20	topic_4	cancelar	3,276	Row45	topic_9	controle	932
Row21	topic_4	assinatura	1,944	Row46	topic_9	atendimento	479
Row22	topic_4	cancelamento	1,098	Row47	topic_9	caiu	395
Row23	topic_4	plano	232	Row48	topic_9	remoto	328
Row24	topic_4	equipamentos	165	Row49	topic_9	conexão	215

Fonte: elaborado pelos autores (2023)

NGram

O algoritmo *NGram* foi configurado para gerar *ngrams*, ou frases chaves com 2 termos. De acordo com análises realizadas e validadas com o AD da Hi Platform, o *ngram* fornece maior informação sobre a intenção do cliente. Ou seja, fornece informações que vão além do seu objetivo de contato complementando a intenção com uma entidade ou respectivo objeto de contato, que também pode ser entendido como característica da sua intenção. Por exemplo: Comprar Recarga, Colocar Recarga, Manter Pacote, Manter Plano, Prosseguir Cancelamento. A nuvem de palavras ou frases chaves geradas pelo *NGram* é apresentada na Figura 9.

Testamos também o algoritmo *NGram* com 3 termos para gerar frases chaves, porém, não foram observados resultados significativos.

Imagem 9 - Nuvem de Frases Chaves - NGram



Fonte: elaborado pelos autores (2023)

BoW

O algoritmo BoW, na configuração padrão do *Knime*, não apresentou palavras chaves significativas de acordo com análises realizadas e validadas com o AD da Hi Platform. A normalização realizada pelo algoritmo apresentou termos com semântica de difícil compreensão, considerando o conhecimento dos analistas sobre o negócio deste cliente sob estudo .

Outros Testes

Nos algoritmos utilizados testamos *nodes* nativos do *Knime* para aplicar *tags* no intuito de identificar a classe de palavras (substantivos, adjetivos e verbos) para melhorar a identificação de intenções e possibilitar a identificação de entidades. Ou seja, o objeto relacionado à intenção do usuário ou as características de sua intenção. Por exemplo: Trocar endereço ou Trocar Produto, Mudar tamanho. Todavia, os *nodes* disponibilizados pelo Knime para realizar esta classificação ainda não estão preparados para o idioma Português.

Também foi testada a lematização dos termos no pré-processamento, ou seja, redução das palavras ao seu radical. Todavia, observou-se que tal processamento dificultaria a compreensão do resultado exposto nas Nuvens de Palavras geradas e não demonstraria as diversas maneiras do cliente indicar uma mesma intenção.

De acordo com análise realizada e validada com o AD da Hi Platform, a nuvem de palavras do algoritmo LDA demonstra de forma clara as principais intenções do cliente, pois apresenta termos com valor agregado para o negócio com resultado mais assertivo do que as configurações anteriores e do que os demais algoritmos analisados.

Para dar continuidade às ações que promoverão o aperfeiçoamento de *chatbots*, é de extrema relevância, que o AD acompanhe o processo desde o seu início e que haja investimento contínuo para obtenção de conhecimento na ferramenta *Knime*.

Trabalhos Futuros

Futuramente, esse estudo poderá ser aplicado em outros canais, como na base de conversas realizadas por meio de mídias sociais e aplicativos de mensagens, como Whatsapp, Facebook e Instagram a fim de ampliar os insights desse tema para propiciar um banco de dados crescente e contínuo.

Observou-se também que, devido à evolução natural do idioma português, as formas do usuário se expressar por meio da escrita podem mudar, sendo necessário o constante monitoramento sobre os resultados e atualização do fluxo gerado.

O resultado apresentado pelo extrator de tópicos LDA poderia ser refinado através do *script* criado para identificar a relação de dependência sintática, ou seja, a relação entre duas palavras em uma frase com uma palavra sendo a governadora e a outra sendo a dependente da relação. Dessa forma, seria possível reduzir a quantidade de nomes próprios ou escritas de forma equivocada ou com baixo valor semântico que podem ser identificados como pertencentes às classes de verbos, adjetivos ou substantivos e apresentadas na Nuvem de Palavras quando deveriam ser ignoradas devido a sua relação sintática.

A evolução da ferramenta *Knime* também pode influenciar neste processo. Por exemplo, novos *nodes* podem ser desenvolvidos e disponibilizados para a comunidade. Existem grupos de estudos que objetivam desenvolver *nodes* compatíveis com o idioma Português e que futuramente podem ser usados para detalhar melhor os resultados obtidos através da identificação da classe de palavras, como adjetivos, substantivos e verbos, bem como nomes próprios de pessoas e lugares.

Durante os estudos, foi analisada a possibilidade de aplicação do algoritmo BERT (*Bidirectional Encoder Representations from Transformers*), todavia, por necessitar de estudos avançados, a sua utilização não foi possível. O BERT é um modelo de

incorporação de palavras de contexto que sugere uma palavra no contexto ou um vetor de palavras dependente do contexto. O BERT apresenta resultados promissores em muitas áreas, como classificação de sentimentos e Reconhecimento de Entidade Nomeada (NER) (KHAN et.al, 2022).

CONSIDERAÇÕES FINAIS

Como contribuição acadêmica, este estudo trouxe à comunidade uma visão estruturada sobre a percepção dos benefícios por meio do uso de *chatbots* com IA no atendimento ao cliente. Até o momento nenhum estudo sobre o tema foi identificado dentro do contexto apresentado, uma vez que a literatura acadêmica se volta mais para o uso de IA e seus algoritmos.

Também se identificou contribuição relevante quanto à fundamentação teórica levantada, abordando questões sobre o uso de inteligência artificial nas organizações e sua relação com os processos de atendimento ao cliente. Como contribuição prática, os resultados apresentados pelo estudo podem ser utilizados como balizadores para a adoção do uso de inteligência artificial voltado ao atendimento dos clientes pelas organizações, permitindo uma visão comparativa sobre os benefícios desejados e possivelmente alcançáveis da adoção dessa tecnologia. Permite também aos executivos das organizações uma visão ampla sobre grandes projetos de uso de chatbots, seus benefícios e os impactos percebidos nas empresas.

O estudo apresentado também reforça a ideia de que de nada adianta um chatbot de alta performance e confiança se não for aceito pelos usuários. Sendo assim, é necessário entender porque os usuários aceitam ou rejeitam o robô. Um sistema que frustra os usuários não pode ser considerado eficaz mesmo que use tecnologia de ponta. Para garantir a qualidade, o chatbot deve ser desenvolvido e centrado no usuário (REBECCHI, 2020).

REFERÊNCIAS

ARSOVSKI, S.; OSIPYAN, H.; MUNIRU, I. O.; CHEOK, A. D. Automatic knowledge extraction of any Chatbot from conversation. **Expert Systems with Applications**. v.137, p.343-348, 2019.

BRAUN, D.; MENDEZ, A. H.; MATTHES, F.; LANGEN, M. 2017. Evaluating Natural Language Understanding Services for Conversational Question Answering Systems. In

Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, pages 174–185, Saarbrücken, Germany. Association for Computational Linguistics. <<https://DOI:10.18653/v1/W17-5522>>. Acesso em: 31 out. 2022.

BRENNER, M. Problems in collecting social data: a review for the information researcher. **Social Science Information Studies**, v.1, nr.3, 139-51, April, 1981. <https://revista.ibict.br/article/download>

COMARELLA, R. L.; CAFÉ, L. M. A. **CHATTERBOT: conceito, características, tipologia e construção.** conceito, características, tipologia e construção. 2008.

FELDMAN, R.; SANGER, J. **The Text Mining Handbook: Advances Approaches in Analysing Unstructured Data.** 2. ed. New York, UK: Cambridge University Press, 2016. 423 p.

GADELHA, I. B. L. **O Uso de Chatbots no Atendimento de Clientes de Revenda por Catálogo.** Dissertação (Mestrado em Computação Aplicada – Núcleo de Desenvolvimento da Amazônia em Engenharia da Universidade Federal do Pará - Tucuruí, 2019.

HARRIS, Z. S. Distributional structure. **Word**. v. 10, n. 2-3, p. 146–162, 2015.

KHAN, M. Q. et al. Impact analysis of keyword extraction using contextual word embedding. **Peerj Computer Science**. 2022. Disponível em: <<https://peerj.com/articles/cs-967/>>. Acesso em 24 out 2022.

KIETZMANN, J.; PASCHEN, J.; TREEN, E. Artificial Intelligence in Advertising. **Journal Of Advertising Research**, [s.l.], v. 58, n. 3, p.263-267, set. 2018. WARC Limited. <<http://dx.doi.org/10.2501/jar-2018-035>>. Acesso em 24 nov 2022.

KULIGOWSKA, K. Commercial Chatbot: Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents. **Professionals Center for Business Research**. v.2, p.1-16, 2015.

LEMON, K. N., VERHOEF, P. C. Understanding customer experience throughout the customer journey. **Journal of Marketing**, v.80, nr. 6, p. 69-96. 2016. Disponível em: <[doi:10.1509/jm.15.0420](https://doi.org/10.1509/jm.15.0420)>. Acesso em: 31 out. 2022.

LENG, Yan et al. LDA-based data augmentation algorithm for acoustic scene classification. **Science Direct**. v.195, Mai 2020.

LINDEN, G.; SMITH, B.; YORK, J. Amazon.com Recommendations Item-to-Item Collaborative Filtering. **IEEE Internet Computing**, n. February, 2003. Disponível em: <<https://doi.org/10.1109/MIC.2003.1167344>>. Acesso em: 08 set. 2022.

MOBILE TIME. **Mapa do Ecossistema Brasileiro de Bots 2022.** Disponível em: <<https://www.mobiletime.com.br/pesquisas/mapa-do-ecossistema-brasileiro-de-bots-2022/>>. Acesso em: 23 out. 2022.

ONAN, A.; BULUT, H.; KORUKOGLU, S. An improved ant algorithm with LDA-based representation for text document clustering. **Journal of Information Science**. v.43, Mar 2016.

PRZEGALINSKA A.; CIECHANOWSKI L.; STROZ A.; GLOOR P.; MAZUREK g.. In bot we trust: A new methodology of chatbot performance measures. **Business Horizons**, v. 62, 2019.

REBECCHI, A. B. **ROBÔ HUMANO: Estudo sobre a humanização no atendimento com chatbot**. Dissertação (Mestrado Profissional em Comportamento do Consumidor – Escola Superior de Propaganda e Marketing ESPM - São Paulo, 2020..

REINA, D. R.; CRUZ, C. Curadoria de chatbots: conceptualização, estratégias e indicadores de desempenho. **Revista Ibérica de Sistemas e Tecnologias de Informação**, São Paulo, p.1-14, 2020.

WONDIMU, N. A.; BUCCHE, C.; VISSER, U.. **Interactive Machine Learning: A State of the Art Review**. Cornell University. Jul. 2022.

WANG, X.; LIN, X.; SHAO, B. How does artificial intelligence create business agility? Evidence from chatbots. **International Journal of Information Management**, v.66, 2022.

ZANELLA, L. C. H. **Metodologia de Pesquisa**. 2. ed. Florianópolis, SC: UFSC, 2013. 134 p.