



UNIVERSIDADE
ESTADUAL DE LONDRINA

LAURA FERREIRA ROCHA

DESENVOLVIMENTO DE UM *CHATBOT* INFORMATIVO
SOBRE SAÚDE MENTAL BASEADO EM APRENDIZADO
DE MÁQUINA

LONDRINA

2024

LAURA FERREIRA ROCHA

**DESENVOLVIMENTO DE UM *CHATBOT* INFORMATIVO
SOBRE SAÚDE MENTAL BASEADO EM APRENDIZADO
DE MÁQUINA**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Ciência da Computação da Universidade Estadual de Londrina para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Gilberto Fernandes Junior

LONDRINA

2024

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

R672d Rocha, Laura Ferreira.
Desenvolvimento de um chatbot informativo sobre saúde mental baseado em aprendizado de máquina / Laura Ferreira Rocha. - Londrina, 2024.
59 f. : il.

Orientador: Gilberto Fernandes Junior.
Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Graduação em Ciência da Computação, 2024.
Inclui bibliografia.

1. Chatbot - TCC. 2. Aprendizado de máquina - TCC. 3. Saúde Mental - TCC.
I. Fernandes Junior, Gilberto. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Graduação em Ciência da Computação. III. Título.

CDU 519

LAURA FERREIRA ROCHA

DESENVOLVIMENTO DE UM *CHATBOT* INFORMATIVO
SOBRE SAÚDE MENTAL BASEADO EM APRENDIZADO
DE MÁQUINA

Trabalho de Conclusão de Curso apresentado
ao curso de Bacharelado em Ciência da Com-
putação da Universidade Estadual de Lon-
drina para obtenção do título de Bacharel em
Ciência da Computação.

BANCA EXAMINADORA

Orientador: Prof. Dr. Gilberto Fernandes
Junior
Universidade Estadual de Londrina

Prof. Dr. Bruno Bogaz Zarpelão
Universidade Estadual de Londrina - UEL

Prof. Dr. Arthur Alexandre Artoni
Universidade Estadual de Londrina - UEL

Londrina, 10 de maio de 2024.

*Este trabalho é dedicado aos que reconhecem
que nesta vida não se conquista nada
sozinho.*

AGRADECIMENTOS

Gostaria de expressar minha mais profunda gratidão a todas as pessoas que contribuíram para a realização deste trabalho.

Primeiramente, agradeço ao meu orientador, Prof. Dr. Gilberto Fernandes, que me acolheu como orientanda em meio a uma troca de tema do TCC. Sua expertise e seu apoio foram fundamentais não apenas para a condução deste projeto, mas também para a minha formação acadêmica e pessoal. Sou profundamente grata por sua confiança e suporte em cada passo deste desafiante percurso.

Agradeço também a minha mãe, meu pai, Maria Luiza, Gustavo e Blenda. Esta é a minha família. Vocês são tudo de mais importante que eu tenho. Vocês são meus pilares! O apoio que recebi de vocês nos momentos de dúvida e exaustão foram decisivos para eu conseguisse seguir em frente. Obrigada por estarem ao meu lado, celebrando cada pequeno sucesso e me apoiando nos desafios.

Por fim, agradeço à Professora Dr.^a Silvia Regina de Souza, que abriu as portas do seu projeto para uma colaboração conosco. Espero que a nossa parceria continue frutífera e que possamos colaborar em futuros projetos.

Este trabalho é também um reflexo do apoio que recebi de cada um de vocês. Obrigada por estarem ao meu lado.

*“Uma pessoa que nunca cometeu um erro,
nunca tentou nada de novo.*

— Albert Einstein

ROCHA, L. F.. **Desenvolvimento de um *chatbot* informativo sobre saúde mental baseado em aprendizado de máquina**. 2024. 58f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina, 2024.

RESUMO

A saúde mental é um aspecto fundamental do bem-estar psicológico humano. Ela influencia diretamente a capacidade de lidar com os desafios da vida, manter relacionamentos saudáveis e contribuir positivamente para a sociedade. Apesar disso, pesquisas recentes destacam a preocupante prevalência de problemas de saúde mental entre os profissionais da saúde, uma classe essencial para o funcionamento adequado do sistema de saúde. Diante deste cenário, faz-se necessário o desenvolvimento de iniciativas voltadas para o suporte emocional e promoção do bem-estar psicológico desses profissionais. Por outro lado, o emprego das tecnologias vem aumentando em todos os níveis de atenção à saúde, tanto nos serviços suplementares quanto na rede privada de assistência. Diante disto, surge a tecnologia do *chatbot*, um programa computacional de conversação que simula diálogos humanos. Os *chatbots* são uma ferramenta moderna e cada vez mais utilizada para prestar algum tipo de assistência ao usuário. Portanto, este trabalho tem como objetivo desenvolver e validar um *chatbot* informativo sobre saúde mental voltado para profissionais de saúde. O chatbot proposto abrange diferentes temas de saúde mental, como ansiedade, luto, comunicação efetiva e relacionamento com pacientes, por exemplo. O sistema conta com um classificador de textos implementado com uma rede neural de propagação direta para classificação das intenções dos usuários. O *chatbot* foi alocado em um *website* desenvolvido em Django e hospedado na plataforma Render. A validação do sistema proposto foi realizada por meio da análise do histórico de perguntas dos participantes para o *chatbot* e da análise do questionário de avaliação que os participantes preencheram ao final da experiência. De forma geral, os resultados foram satisfatórios. Em sua primeira versão, o *chatbot* alcançou uma acurácia de 60,10% na classificação correta das intenções. Além disso, 68,18% dos participantes concordaram ou concordaram totalmente com a afirmação de que as respostas do *chatbot* atendiam às suas perguntas.

Palavras-chave: *Chatbot*. Classificador de Textos. Saúde Mental.

ROCHA, L. F.. **Development of an informative chatbot on mental health based on machine learning**. 2024. 58p. Final Project (Bachelor of Science in Computer Science) – State University of Londrina, Londrina, 2024.

ABSTRACT

Mental health is a fundamental aspect of human psychological well-being. It directly influences the ability to handle life's challenges, maintain healthy relationships, and contribute positively to society. Despite this, recent research highlights the concerning prevalence of mental health issues among healthcare professionals, a group essential for the proper functioning of the health system. Given this scenario, it is necessary to develop initiatives aimed at providing emotional support and promoting the psychological well-being of these professionals. On the other hand, the use of technology has been increasing at all levels of healthcare, both in supplementary services and in the private healthcare network. In this regard, the technology of the chatbot, a computational conversational program that simulates human dialogues, emerges. Chatbots are a modern tool increasingly used to provide some type of assistance to users. Therefore, this study aims to develop and validate an informative chatbot about mental health for healthcare professionals. The proposed chatbot covers various mental health topics such as anxiety, grief, effective communication, and patient relationships, for example. The system includes a text classifier implemented with a feedforward neural network to classify user intentions. The chatbot was deployed on a website developed in Django and hosted on the Render platform. The validation of the proposed system was conducted through the analysis of the participants' question history to the chatbot and the analysis of an evaluation questionnaire filled out by the participants at the end of the experience. Generally, the results were satisfactory. In its first version, the chatbot achieved an accuracy of 60.10% in the correct classification of intentions. Additionally, 68.18% of the participants agreed or strongly agreed that the chatbot's responses met their questions.

Keywords: Chatbot. Text Classifier. Mental Health.

LISTA DE ILUSTRAÇÕES

Figura 1 – Principais técnicas de AM.	18
Figura 2 – Processo de aprendizado supervisionado.	18
Figura 3 – Representação esquemática de um neurônio artificial.	20
Figura 4 – Exemplo de uma RNA de propagação direta.	21
Figura 5 – Exemplo de tokenização.	23
Figura 6 – Exemplo de lematização.	24
Figura 7 – Exemplo de remoção de <i>stop words</i>	25
Figura 8 – Exemplo de criação do dicionário de características.	26
Figura 9 – Exemplo de criação do vetor de características de três diferentes documentos.	27
Figura 10 – Arquitetura geral de um <i>chatbot</i> . Traduzido de [1].	30
Figura 11 – Arquitetura geral do <i>chatbot</i> proposto.	34
Figura 12 – Exemplo do processo de criação dos padrões intenção-resposta para treinamento do <i>chatbot</i>	35
Figura 13 – <i>Pipeline</i> de pré-processamento.	37
Figura 14 – Representação esquemática do classificador de textos.	38
Figura 15 – Tela inicial do <i>website</i>	39
Figura 16 – Tela de interação com o <i>chatbot</i>	40
Figura 17 – Exemplo de interação com o <i>chatbot</i> sobre ansiedade.	40
Figura 18 – Tela de avaliação do <i>chatbot</i>	41
Figura 19 – Tela final do <i>website</i>	42
Figura 20 – Quantidade de perguntas feita por cada participante.	45
Figura 21 – Representatividade das perguntas Corretas, Erradas e Indefinidas.	46
Figura 22 – Escolaridade dos participantes da validação do <i>chatbot</i>	48
Figura 23 – Representatividade de participantes que já utilizaram algum <i>chatbot</i> antes.	49

LISTA DE TABELAS

Tabela 1 – Amostra das intenções identificadas na cartilha de saúde mental. . . .	35
Tabela 2 – Caracterização das perguntas do questionário de avaliação do <i>chatbot</i> . .	43
Tabela 3 – Caracterização das perguntas não respondidas ou respondidas incorre- tamente pelo <i>chatbot</i>	46
Tabela 4 – Resultado da avaliação do <i>chatbot</i> nas perguntas em escalas Likert. . .	50

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
BoW	<i>Bag of Words</i>
CT	Classificação de Textos
OMS	Organização Mundial da Saúde
ReLU	<i>Rectified Linear Unit</i>
RNA	Rede Neural Artificial
UEL	Universidade Estadual de Londrina

SUMÁRIO

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Saúde Mental dos Profissionais da Saúde	16
2.1.1	Projeto Manual de Saúde Mental da UEL	16
2.2	Aprendizado de Máquina	17
2.2.1	Técnicas de Aprendizagem	18
2.2.2	Redes Neurais Artificiais	19
2.3	Classificação de Textos	22
2.3.1	Tokenização	22
2.3.2	Normalização de <i>Tokens</i>	23
2.3.3	Remoção de <i>Stop Words</i>	24
2.3.4	<i>Bag of Words</i>	25
2.4	<i>Chatbots</i>	27
2.4.1	Tipos de <i>Chatbots</i>	28
2.4.2	Arquitetura Geral	29
2.5	Trabalhos Correlatos	31
3	DESENVOLVIMENTO	33
3.1	Arquitetura do <i>Chatbot</i> Proposto	33
3.2	Base de Conhecimento	34
3.3	Pré-Processamento	35
3.4	Classificador de Textos	37
3.5	Interfaces Gráficas de Usuário	38
3.6	Métricas de Avaliação	42
4	RESULTADOS E DISCUSSÃO	44
4.1	Classificador de Textos	44
4.2	Histórico de Perguntas	44
4.3	Avaliação dos Participantes	48
5	CONCLUSÃO	52
	REFERÊNCIAS	53

APÊNDICES	57
APÊNDICE A – INTENÇÕES DA BASE DE CONHECI- MENTO	58

1 INTRODUÇÃO

A saúde mental é um aspecto fundamental do bem-estar psicológico humano. Ela influencia diretamente a capacidade de lidar com os desafios da vida, manter relacionamentos saudáveis e contribuir positivamente para a sociedade [2].

Pesquisas recentes destacam a preocupante prevalência de problemas de saúde mental entre os profissionais da saúde, uma classe essencial para o funcionamento adequado do sistema de saúde [3, 4, 5, 6, 7]. A exposição diária a situações de incerteza, pressão emocional e a intensa carga de trabalho são fatores contribuintes para o sofrimento psicológico desses profissionais [6]. Além disso, a sobrecarga emocional, o medo de agressão por parte dos pacientes e a fadiga decorrente dos turnos de trabalho irregulares são questões adicionais que aumentam o risco de desenvolver problemas de saúde mental [7]. Essas situações podem levar a altos níveis de ansiedade, estresse, depressão e até mesmo ao esgotamento profissional, como a Síndrome de *Burnout* [3, 4].

Diante deste cenário, faz-se necessário o desenvolvimento de iniciativas voltadas para o suporte emocional e promoção do bem-estar psicológico dos profissionais da saúde. Tais medidas não apenas são essenciais para garantir o próprio bem-estar desses profissionais, mas também são importantes para manter a eficácia e qualidade no atendimento aos pacientes [5, 6].

O emprego das tecnologias vem aumentando em todos os níveis de atenção à saúde, tanto nos serviços suplementares quanto na rede privada de assistência [8]. A utilização dessas tecnologias torna-se eficaz em facilitar a promoção da saúde em diversos grupos e condições, como a saúde mental. Isso porque o aumento do acesso às tecnologias ajuda a criar ambientes nos quais as informações de saúde estão disponíveis [9]. Diante disto, surge a tecnologia do *chatbot*, um programa computacional de conversação que simula diálogos humanos. Um *chatbot* realiza processamento e classificação de textos para vincular sentenças comuns a respostas predefinidas, servindo para otimizar processos em serviços de saúde complexos [10].

Em vista desse panorama, faz-se necessário investigar o uso dos *chatbots* no enfrentamento de questões de saúde mental. Portanto, este trabalho visa desenvolver um *chatbot* informativo sobre saúde mental voltado para profissionais de saúde. Para atingir esse objetivo geral, tem-se como objetivos específicos: (a) projetar a arquitetura de um *chatbot* informativo; (b) estruturar uma base de conhecimento científica sobre saúde mental para treinar o *chatbot*; (c) desenvolver e disponibilizar o *chatbot* na Web para poder ser facilmente acessado; e (d) validar o sistema proposto.

Este trabalho é fruto de uma parceria com o projeto Manual de Saúde Mental de-

envolvido no Departamento de Psicologia da Universidade Estadual de Londrina (UEL), sob coordenação da Professora Dr.^a Silvia Regina de Souza e a Professora Dr.^a Nádia Kienen. Nesse projeto foi desenvolvida uma cartilha com diversas informações de saúde mental voltada para profissionais da saúde. Alguns tópicos abordados são os conceitos e estratégias para lidar com ansiedade, estresse e luto; como abordar assuntos difíceis com pacientes; o que é comunicação efetiva; como ter uma comunicação efetiva, dentre outros. Essa cartilha foi cuidadosamente desenvolvida por estudantes de psicologia com alto rigor em termos de referencial teórico. Nessa parceria, a cartilha desenvolvida será utilizada para a criação de uma base de conhecimento para treinamento do *chatbot* informativo proposto.

O restante deste trabalho está organizado como segue. No Capítulo 2, é apresentada a fundamentação teórica, onde se desenvolve os conceitos relevantes neste estudo. O Capítulo 3 detalha todos os procedimentos realizados para projetar, desenvolver e avaliar o sistema proposto. No Capítulo 4 são apresentados e discutidos os resultados obtidos. Por fim, o Capítulo 5 apresenta a conclusão deste trabalho e propostas para futuros desenvolvimentos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Saúde Mental dos Profissionais da Saúde

Segundo a Organização Mundial da Saúde (OMS), saúde mental é um estado de bem-estar psicológico que permite às pessoas lidar com os desafios da vida, relacionar-se de forma saudável e contribuir positivamente para sua comunidade. A saúde mental vai além da ausência de transtornos mentais. Ela é experimentada de diferentes formas por cada indivíduo, com graus de dificuldade e sofrimento variados e, eventualmente, resultados muito diferentes. A saúde mental é um direito humano fundamental e é essencial para o desenvolvimento pessoal, comunitário e socioeconômico [2].

Pesquisas recentes apontam para os profissionais da saúde como uma classe alvo de problemas de saúde mental [3, 4, 5, 6, 7]. Um estudo descritivo realizado por Bonazza e Schuh [3] com 85 profissionais de diferentes áreas da saúde identificou que esses profissionais apresentaram níveis de ansiedade acima da média populacional. Além disso, uma revisão da literatura conduzida por Ferrari, França e Magalhães [4] evidenciou uma predisposição dos profissionais da saúde a apresentarem elevado nível de estresse, podendo levar à Síndrome de *Burnout*, uma doença resultante do esgotamento profissional.

A exposição diária a incertezas, tensões e angústias, combinada com a intensa carga emocional dos pacientes, pode contribuir para o sofrimento psíquico dos profissionais da saúde [6]. A sobrecarga emocional, o medo de agressão e o cansaço resultante do trabalho por turnos são fatores adicionais que aumentam o risco de estresse, ansiedade e depressão [7]. A complexidade das decisões e ações requeridas, juntamente com o ambiente estressante, podem gerar alterações comportamentais significativas [3]. Além disso, o esgotamento mental desses profissionais é frequentemente banalizado, muitos trabalhadores não dão a devida relevância a seus problemas de saúde, o que pode ter diversas implicações [7].

Essas evidências científicas apontam para o alto impacto que as demandas às quais os profissionais de saúde são submetidos diariamente podem ter para sua própria saúde mental [3]. Diante desse cenário, é essencial o desenvolvimento de iniciativas de suporte emocional e promoção do bem-estar psicológico desses profissionais. Essas ações também são importantes para garantir o bem-estar e a eficácia no atendimento aos pacientes [5, 6].

2.1.1 Projeto Manual de Saúde Mental da UEL

No departamento de Análise do Comportamento da UEL, a Professora Dr.^a Silvia Regina de Souza e a Professora Dr.^a Nádia Kienen coordenam um projeto de suporte

psicológico aos profissionais da saúde. O projeto conta com a participação de graduandos e mestrandos do curso de Psicologia da UEL e também com professores e graduandos de outros departamentos, como o de Computação e o de Design Gráfico.

O projeto iniciou-se durante a pandemia da Covid-19, que impactou diretamente a saúde mental dos profissionais da saúde. Nesse contexto, os participantes do projeto levantaram uma série de referenciais teórico-científicos para propor uma cartilha destinada a capacitar psicólogos, médicos, enfermeiros e outros profissionais que atuam no contexto da saúde a gerenciarem variáveis ambientais que afetam a própria saúde mental.

A cartilha conta com seis capítulos. O primeiro capítulo apresenta os conceitos de ansiedade, estresse e luto. Ele também conta com o descritivo de uma série de estratégias para lidar com esses sentimentos. O segundo capítulo, por sua vez, fala sobre comunicação, dando dicas para comunicar situações delicadas, manejar situações de conflito e compartilhar sentimentos de forma efetiva. Já o terceiro capítulo é dedicado a apresentar um protocolo para garantir maior segurança na comunicação com o paciente. O capítulo também reflete sobre como o profissional da saúde pode autoavaliar sua comunicação. Os capítulos quatro e cinco exploram as relações de trabalho e relações familiares, respectivamente. Por fim, o capítulo seis conceitua autocuidado e apresenta práticas essenciais nesse contexto.

2.2 Aprendizado de Máquina

O Aprendizado de Máquina (AM) é um campo da inteligência artificial dedicado a desenvolver sistemas capazes de aprender a executar tarefas específicas sem instruções explícitas. Entende-se que um programa aprende se seu desempenho em realizar uma tarefa melhora com base em experiência [11].

Na prática, o objetivo é identificar padrões precisos e confiáveis em dados que podem ser generalizados para novas observações. Para isso, são executados algoritmos que estimam parâmetros de ajuste (ou pesos). Esses parâmetros são, de forma simplista, a medida de influência de cada variável no conjunto de dados. Por fim, é obtido um modelo capaz de gerar algum tipo de previsão quando apresentado a um novo conjunto de dados [12].

Não existe uma fronteira bem definida entre o aprendizado de máquina e a estatística, e nem sempre há um consenso se determinado algoritmo é considerado aprendizado de máquina [13]. Ainda que metodologicamente semelhantes, o aprendizado de máquina se concentra em uma precisão preditiva. Isso significa que o principal foco do aprendizado de máquina é criar modelos que façam previsões precisas em conjuntos de dados não conhecidos, mesmo que isso traga complexidade em extrair relações causais entre as diferentes variáveis dos dados [14].

2.2.1 Técnicas de Aprendizagem

No que diz respeito às técnicas de aprendizagem dos modelos, as principais são: (a) aprendizado supervisionado; (b) aprendizado não supervisionado; e (c) aprendizado por reforço [15]. A Figura 1 mostra uma representação esquemática dessas técnicas com foco no aprendizado supervisionado.

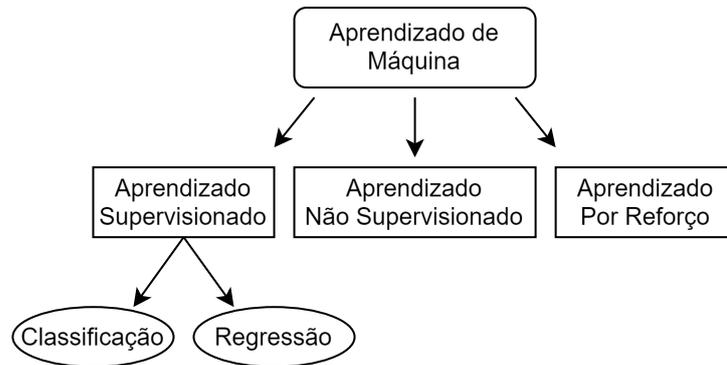


Figura 1 – Principais técnicas de AM.

O aprendizado supervisionado é realizado em dois passos: treinamento e teste [16]. No treinamento, são apresentados ao algoritmo dados organizados como pares de entrada (ou observações) e seus respectivos rótulos conhecidos (a saída esperada) [17]. Conforme processa cada observação, o algoritmo corrige os pesos das variáveis, reduzindo a magnitude de uma função de perda global. Com isso, os padrões presentes nos dados são incorporados no modelo treinado que, em fase de teste, acessa uma entrada não conhecida e deve ser capaz de rotulá-la corretamente (Figura 2) [16]. Dessa forma, a combinação de um algoritmo de aprendizado flexível com um conjunto de dados coerente irá gerar um modelo de alta precisão global [15].

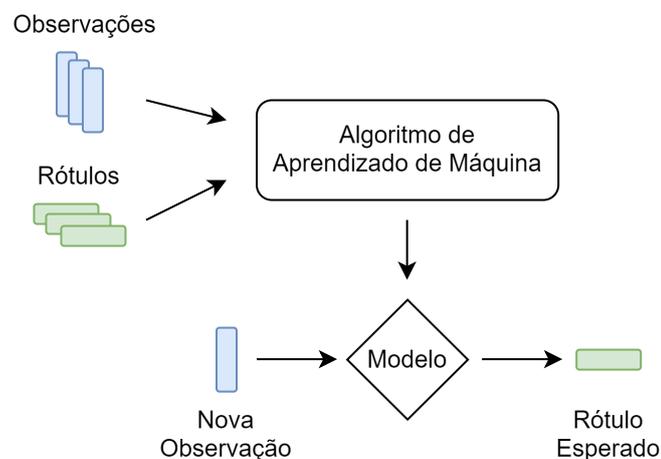


Figura 2 – Processo de aprendizado supervisionado.

Em particular, os modelos supervisionados são utilizados em duas clássicas tarefas estatísticas: classificação (reconhecimento de padrões) e regressão (aproximação de função) [18]. Modelos de regressão mapeiam os dados de entrada para um domínio de valores contínuos, enquanto os classificadores mapeiam os dados para um domínio discreto (de classes predefinidas) [16]. Algumas aplicações de classificadores envolvem a detecção de *spam* em caixas de mensagens, tomada de decisão do pedido de empréstimo pelo devedor e reconhecimento de rostos. Por outro lado, os modelos de regressão podem ser utilizados para predição do preço da energia, predição do desempenho de estudantes em um exame e predição do efeito da variação de temperatura na venda de produtos gelados, por exemplo [15, 19, 20].

O aprendizado não supervisionado, por sua vez, é uma abordagem exploratória, na qual o algoritmo busca identificar relações nos dados sem ter qualquer referência prévia de uma resposta esperada [13]. Uma vez que os dados não são rotulados, essa técnica é utilizada para entender como um conjunto de dados pode ser agrupado de acordo com suas semelhanças (ou diferenças). Algumas aplicações comuns dos modelos de aprendizado não supervisionado são na segmentação de objetos e detecção de semelhança, por exemplo [15].

Por fim, no aprendizado por reforço, o sistema melhora seu desempenho em uma tarefa ao realizá-la por meio de tentativa e erro [15]. A cada tentativa, o sistema recebe uma recompensa (positiva ou negativa) útil para entender se a ação realizada é positiva ou não. Em seguida, ele se adapta e tenta realizar a tarefa novamente com sua nova percepção [13]. O objetivo dessa abordagem é aprender uma sequência de ações úteis, chamada de política, para tomar sempre a melhor decisão, visando a maior recompensa imediata e acumulada. Algumas aplicações do aprendizado por reforço são na melhoria da política para jogar jogos de videogame [15] e na recomendação personalizada de artigos de notícias [21], por exemplo.

2.2.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) constituem uma classe de algoritmos de AM inspirada no comportamento dos neurônios em redes neurais biológicas [22]. Uma RNA consiste em uma estrutura de neurônios interconectados capaz de generalizar interações complexas entre um grupo de variáveis independentes a fim de prever um resultado (variável dependente) [13].

O neurônio (Figura 3) é a unidade básica de processamento de uma RNA. Ele possui n entradas de dados (x_i), cada uma caracterizada por um peso (w_i). O processamento nessa unidade consiste em multiplicar cada entrada por seu respectivo peso e somar um viés opcional (b). O viés ajusta o plano de decisão da tarefa que o neurônio está tentando aprender, ele é uma constante que permite trasladar o resultado do neurônio. Isso torna o

modelo mais flexível, o que permite que ele se adapte de maneira mais eficaz a diferentes padrões de dados [15].

O somatório desse processo é apresentado a uma função de ativação (f_a). Esta, por sua vez, retorna o dado que configura a saída do neurônio, que serve de entrada para os neurônios da camada adjacente [15]. Dependendo do tipo de função de ativação, a saída pode ser dicotômica (“1” quando o somatório excede um determinado limiar e “0” caso contrário) ou contínua [13].

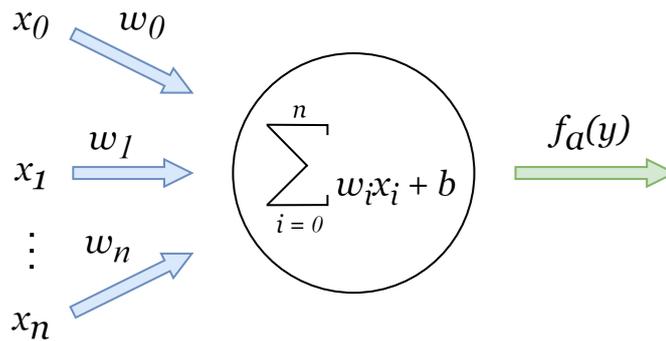


Figura 3 – Representação esquemática de um neurônio artificial.

A função de ativação é um componente fundamental do neurônio. Ela reduz a amplitude da saída do neurônio e introduz a não linearidade nas operações realizadas pela rede. A função de ativação permite que a rede aprenda e modele relações não lineares complexas nos dados. A importância da não linearidade reside no fato de que diversos problemas do mundo real não podem ser adequadamente modelados por relações lineares simples entre as variáveis [22, 23].

A função ReLU (*Rectified Linear Unit*) é uma das funções de ativação mais utilizadas em RNAs. Ela se caracteriza por sua simplicidade e eficácia na introdução de não-linearidades no modelo sem afetar a escala dos dados positivos. A definição da ReLU é dada pela Equação 2.1. Ela produz o próprio valor de entrada se esse valor for positivo e zero caso contrário [24].

$$ReLU(x) = \max(0, x) \quad (2.1)$$

Em problemas de classificação com múltiplas classes, a função *Softmax* é amplamente adotada. Essa função converte os valores de saída do modelo em probabilidades, que facilitam a interpretação dos resultados como chances de pertencimento às classes em questão. A definição matemática da *Softmax* é dada pela Equação 2.2 para um vetor de entradas x . Essa função assegura que a saída esteja entre 0 e 1, com a soma das probabilidades de todas as classes totalizando 1, caracterizando uma distribuição de probabilidade.

A *Softmax* é particularmente útil em tarefas de classificação onde a escolha da classe mais provável é o objetivo da modelagem [25].

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (2.2)$$

A RNA é formada por diferentes camadas de neurônios. Essas camadas são divididas em (a) uma camada de entrada, (b) uma ou mais camadas intermediárias (ou ocultas) e (c) uma camada de saída. Os neurônios das camadas de entrada e saída correspondem, respectivamente, às variáveis independentes e dependentes [15].

A Figura 4 ilustra uma RNA simples, com arquitetura de propagação direta. Em RNAs de propagação direta os dados se deslocam de maneira unidirecional, da camada de entrada à camada de saída [13]. Além disso, cada neurônio de uma camada está conectado a todos (e apenas) os neurônios da camada subsequente [23].

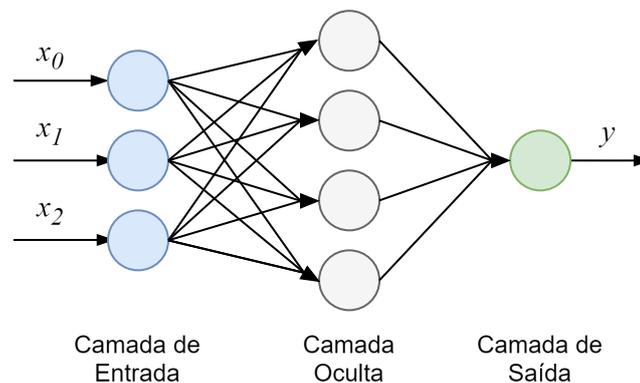


Figura 4 – Exemplo de uma RNA de propagação direta.

Em aplicações de aprendizado supervisionado, inicialmente, define-se o número de camadas e neurônios. Os pesos (w_i) de todas as conexões são inicializados aleatoriamente. Durante o treinamento, a RNA é exposta a um conjunto rotulado de dados. Nesse processo, os dados obtidos na camada de saída são comparados com os rótulos conhecidos utilizando uma função de perda, que mede a diferença entre os valores previstos e verdadeiros. O objetivo é reduzir a função de perda a zero, ou seja, fazer com que a saída prevista pela RNA corresponda o mais fielmente possível à verdade [13].

Os erros mensurados pela função de perda são distribuídos retroativamente pela rede, da saída para a entrada. Assim, é atribuída uma contribuição de erro a cada neurônio nas camadas ocultas e de entrada. Esse processo é denominado retropropagação. Além disso, a fim de reduzir a função de perda, os pesos de conexão são atualizados, processo conhecido como ajuste de peso. Essa otimização se repete por um determinado número de iterações (ou épocas) com diferentes lotes de dados (ou *batches*) até que a RNA atenda a uma regra de parada predefinida ou um limiar de taxa de erro [13, 22].

2.3 Classificação de Textos

A Classificação de Textos (CT) é uma tarefa de aprendizado supervisionado que associa textos em linguagem natural a uma ou mais categorias predefinidas. Ela é utilizada em diferentes domínios de pesquisa, como extração de informações, recuperação de informações e mineração de texto, por exemplo. Existem dois tipos principais de CT, a baseada em tópicos e a estilística. A primeira busca agrupar textos com base no assunto principal ou no tópico que abordam. Um exemplo é determinar se um artigo de notícia trata de finanças, estilo de vida, tecnologia ou esportes. Por outro lado, a classificação estilística se concentra nos aspectos linguísticos e estilísticos dos textos, como gênero literário, tom e estilo de escrita. Um exemplo de aplicação dessa tarefa é a classificação de textos em gêneros literários, como ação, comédia, fantasia e ficção científica [26].

No contexto da CT, são muito utilizados os termos documento, *corpus* e dicionário. Uma coleção de textos livres é definida como um documento e uma coleção de documentos é definida como um *corpus*. As características de um documento são as próprias palavras, e uma lista de todas as possíveis palavras únicas no *corpus* é conhecida como dicionário [26].

Na preparação de um *corpus* para ser apresentado ao algoritmo de AM são realizadas etapas de pré-processamento e, em seguida, aplica-se alguma técnica de representação das palavras restantes como características do documento. Por fim, essa representação é utilizada para treinamento e teste dos algoritmos de AM para classificar textos não conhecidos [27].

Aplicar etapas de pré-processamento no conjunto de dados pode melhorar sua qualidade, em especial para a tarefa de CT. Esse processo auxilia na eliminação de ruídos nos dados e, em alguns casos, a aplicação desses métodos pode melhorar a qualidade do conjunto de dados para tarefas de CT [26]. A seguir são detalhadas as principais etapas de pré-processamento e, por fim, é explicado o modelo de representação numérica de textos utilizado neste trabalho.

2.3.1 Tokenização

A tokenização consiste em dividir um texto em componentes individuais significativos, conhecidos como *tokens*. Um tipo comum de tokenização é a identificação das palavras individuais em uma sentença. Nessa representação, cada palavra existe por si só, sem o contexto fornecido por suas vizinhas, resultando na perda de parte do significado [28]. A Figura 5 apresenta um exemplo de sentença antes e depois da tokenização.

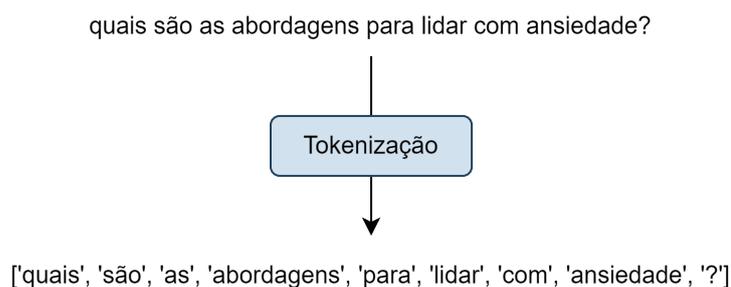


Figura 5 – Exemplo de tokenização.

2.3.2 Normalização de *Tokens*

A normalização de *tokens* é a redução das palavras a uma forma raiz, de modo que variações da mesma palavra sejam reconhecidas como uma única entidade. Existem duas abordagens comuns para normalização de *tokens*, *stemming* e lematização. No *stemming*, os marcadores de nuances de significado, como pluralidade e tempo verbal, são modificados ou removidos. Um exemplo de algoritmo de *stemming* amplamente utilizado é o algoritmo de *stemming* de Porter [28]. Ao aplicar esse algoritmo às palavras “quais”, “abordagens” e “ansiedade”, por exemplo, obtemos os resultados “quai”, “abordagen” e “ansiedad”, respectivamente. Embora gere palavras que não existam na língua portuguesa, o uso de um radical comum para múltiplas formas da mesma palavra simplifica as técnicas de processamento de textos posteriores [27].

Outro método de normalização de *tokens* é a lematização, que utiliza um dicionário de variações conhecidas das palavras em vez de uma abordagem baseada em regras [28]. Ao aplicar a lematização às palavras “quais”, “abordagens” e “ansiedade”, obtemos os resultados “qual”, “abordagem” e “ansiedade”, respectivamente. A Figura 6 apresenta um exemplo de sentença que, inicialmente, passa pela tokenização e, em seguida, sua saída é lematizada.

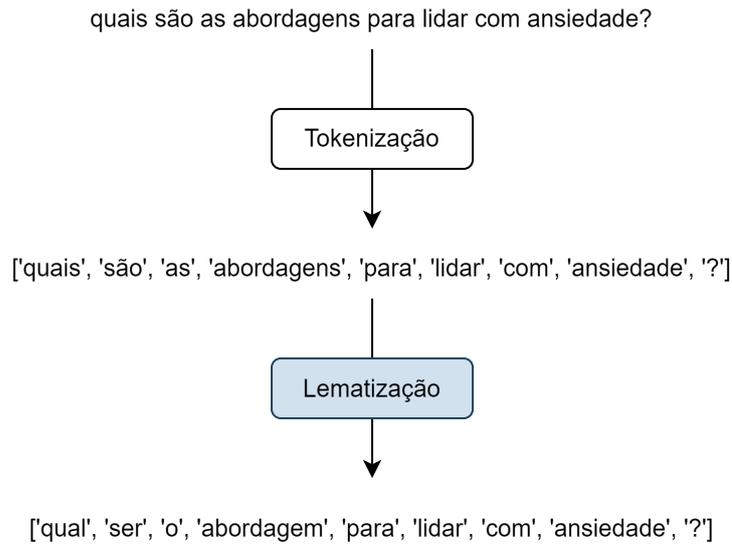


Figura 6 – Exemplo de lematização.

2.3.3 Remoção de *Stop Words*

Stop words são palavras que ocorrem com maior frequência na fala e na escrita dentro do idioma-alvo. Elas são comumente usadas na construção de frases e geralmente não carregam nenhum significado especial, como “o” e “é”, por exemplo. Por essa razão, a remoção dessas palavras irá reduzir significativamente o tamanho do dicionário de palavras únicas, o que é um resultado desejável. Não existe uma única lista de *stop words* universalmente aceita, e variações existem em diferentes dicionários de *stop words* [28].

Ao remover todas as *stop words* de um documento, algumas significações podem ser perdidas, e isso poderia alterar a funcionalidade de uma aplicação. Portanto, é necessário cautela ao remover *stop words* [27]. A Figura 7 complementa o exemplo anterior utilizando a saída da lematização como entrada para o processo de remoção de *stop words*.

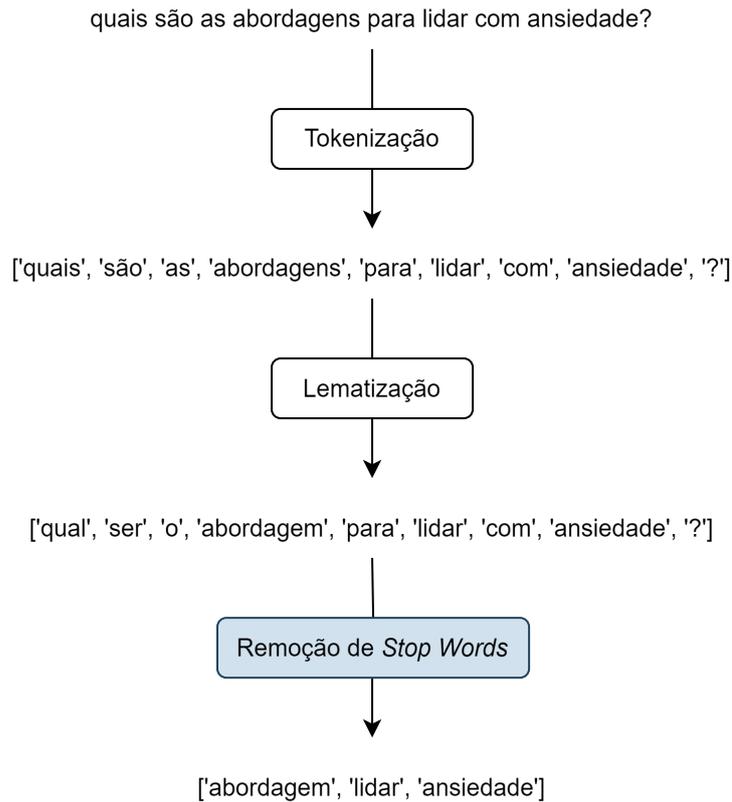


Figura 7 – Exemplo de remoção de *stop words*.

2.3.4 *Bag of Words*

O modelo tradicional para classificação baseada em tópicos é a representação de *Bag of Words* (BoW), que associa um texto a um vetor indicando a ocorrência de cada palavra escolhida no *corpus* de treinamento [26]. Esse é um dos métodos mais comuns para categorização de textos e objetos [29].

Na construção de um modelo BoW, é primeiramente definido o dicionário de características. Um dicionário é definido como uma lista de termos permitidos, sendo composto pelas palavras únicas disponíveis após aplicar as etapas de pré-processamento no *corpus* dos documentos [27]. A Figura 8 ilustra o fluxo de construção de um dicionário de características de um *corpus* composto por três documentos.

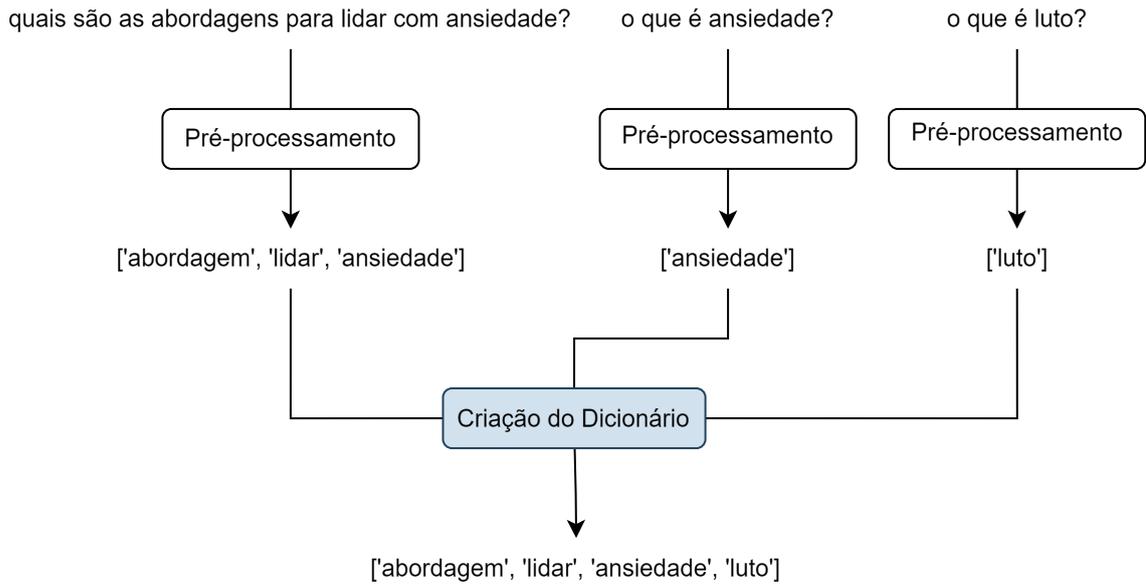


Figura 8 – Exemplo de criação do dicionário de características.

Após a construção do dicionário, o próximo passo é realizar a extração de características de documentos individuais em um vetor de características. Para cada documento do *corpus*, verifica-se se esse documento possui a primeira palavra do dicionário. Se sim, coloca-se algum valor na célula correspondente do vetor de características. Na abordagem mais simples, o valor colocado é baseado na presença do termo, onde atribuímos um “0” ou “1” quando um determinado termo está ausente ou presente, respectivamente. Outros métodos são a contagem de termos e a frequência do termo, por exemplo. Esse processo é então repetido para todas as outras palavras do dicionário, finalizando a representação do primeiro documento em um BoW. Então, todos os documentos são representados em BoW para poderem ser processados pelo algoritmo de AM [27]. A Figura 9 ilustra um exemplo da criação do vetor de características de três diferentes documentos.

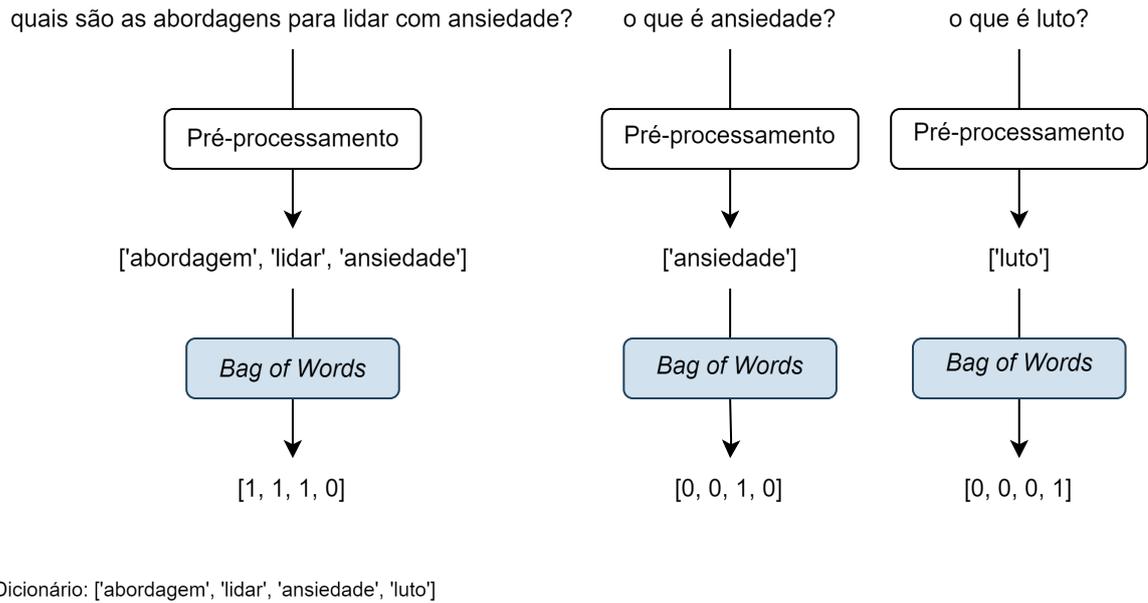


Figura 9 – Exemplo de criação do vetor de características de três diferentes documentos.

Durante a fase de treinamento, o algoritmo de aprendizado de máquina recebe os vetores de características de cada documento, bem como seus respectivos rótulos conhecidos (classes de textos ao qual pertencem). Assim, constrói um modelo capaz de classificar novos documentos dentro dos rótulos existentes [27].

2.4 *Chatbots*

Desde que os primeiros computadores começaram a operar, muitas pessoas já se perguntavam se eles teriam a capacidade de pensar e se um dia seriam comparáveis aos seres humanos. Explorando essa temática, em 1950, Alan Turing introduziu o conceito do Teste de Turing, que questionava a capacidade das máquinas de pensar. Nesse teste, uma pessoa, um computador e um interrogador humano (juiz) são isolados em salas separadas e se comunicam apenas por texto. O juiz deveria analisar o conteúdo das conversas e determinar qual participante é o computador e qual é o humano. A questão central levantada por Turing era se a máquina poderia imitar o pensamento humano o suficiente para enganar o juiz. O Teste de Turing deu então origem à ideia de *chatbots* [30].

Chatbots são sistemas capazes de conduzir uma conversa com um indivíduo, seja por texto ou fala. Eles também são conhecidos como robôs de conversação, *bots* inteligentes, assistentes digitais, entre outros. As respostas de um *chatbot* são geradas com base em análise da entrada do usuário e acessando informações relevantes armazenadas [31]. Os *chatbots* são uma ferramenta moderna e cada vez mais utilizada para prestar algum tipo de assistência ao usuário [32].

Um dos primeiros exemplos foi o ELIZA, criado em 1966, projetado para simular uma terapeuta fazendo perguntas com base nas respostas do usuário. Embora suas habilidades de conversação fossem limitadas, ELIZA era capaz de intrigar as pessoas naquela época em que a interação com computadores era novidade, incentivando o desenvolvimento de outros *chatbots* [1].

Com o passar dos anos, avanços significativos nos campos de processamento de linguagem natural e AM impulsionaram o surgimento de *chatbots* cada vez mais sofisticados. A introdução de RNAs e algoritmos de aprendizado profundo melhorou significativamente a qualidade e a fluência das interações com *chatbots*. Com a popularização da computação em nuvem, *big data* e o acesso a conjuntos de dados em grande escala, o desenvolvimento de *chatbots* foi acelerado [33]. Grandes empresas como Facebook, Google, Microsoft, Apple e OpenAI contribuíram significativamente para esse avanço, introduzindo plataformas de *chatbot* como ChatGPT [34], Siri [35], Alexa [36], Bard [37] e Google Gemini [38], por exemplo.

Os *chatbots* podem ser programados de diversas formas, oferecendo flexibilidade e adaptabilidade. *Chatbots* voltados para tarefas específicas, como automação de reservas, assistência ao cliente e interação com marcas, aumentam a eficiência e produtividade. Por outro lado, os *chatbots* de respostas a perguntas fornecem informações precisas e relevantes aos usuários [32]. Assim, os *chatbots* se tornam uma ferramenta útil para aprimorar a experiência do usuário de variadas formas e em diferentes áreas do conhecimento [33].

2.4.1 Tipos de *Chatbots*

Os *chatbots* podem ser classificados de diversas formas, cada uma referente aos diferentes aspectos de seu design e funcionalidade. Em relação ao domínio de conhecimento, *chatbots* de domínio aberto podem engajar em conversas sobre uma ampla gama de tópicos, enquanto os de domínio fechado são especializados em áreas específicas e respondem a perguntas mais complexas [31], mas podem ter dificuldades com perguntas não relacionadas [1].

O objetivo de cada *chatbot* também pode variar. *Chatbots* informativos são projetados para fornecer informações de um banco de dados ou inventário preexistente [31]. *Chatbots* conversacionais envolvem os usuários em diálogos o mais naturais possível. *Chatbots* baseados em tarefas realizam apenas funções específicas onde as ações são determinadas, como reservar voos [1, 31].

Os *chatbots* também podem ser classificados de acordo com sua plataforma de desenvolvimento. Plataformas de código aberto permitem maior intervenção do desenvolvedor. Enquanto isso, plataformas fechadas podem oferecer recursos avançados, mas limitar o acesso a seu funcionamento interno [1].

O método de processamento de entrada e geração de resposta oferece outra forma de classificar os *chatbots*. Modelos baseados em regras seguem regras predefinidas para seleção de resposta, modelos baseados em recuperação buscam respostas em um banco de dados e modelos generativos geram respostas com base em algoritmos de aprendizado de máquina [1].

Chatbots baseados em regras são frequentemente utilizados por empresas devido à simplicidade de desenvolvimento e eficácia em ambientes controlados [39]. Esses *chatbots* operam com um conjunto predefinido de regras e padrões de conversa, o que os torna aptos a responder apenas a perguntas específicas para as quais foram programados. A organização desses *chatbots* segundo padrões de conversa limita sua capacidade de lidar com perguntas fora desse escopo [40].

Por outro lado, *chatbots* baseados em recuperação oferecem mais flexibilidade em suas respostas, adaptando-se a uma gama mais ampla de interações ao consultar e analisar recursos disponíveis. Esses sistemas são particularmente adequados para domínios fechados, onde o foco está em resolver problemas recorrentes e específicos, utilizando um repertório de respostas predefinidas para selecionar a mais apropriada com base na entrada do usuário [41].

Finalmente, os *chatbots* generativos representam uma abordagem mais avançada. Eles utilizam modelos de linguagem para criar respostas fluidas e coerentes para cada interação, processando e entendendo a linguagem natural de maneira eficaz [39]. *Chatbots* generativos são treinados em grandes volumes de dados, o que lhes permite aprender padrões e relacionamentos complexos entre entidades, facilitando uma interação mais natural e humana com os usuários [42]. O ChatGPT é um exemplo de *chatbot* generativo recente que conquistou uma popularidade mundial. Este *chatbot* alcançou um milhão de usuários em apenas 5 dias [43].

Embora os *chatbots* generativos sejam robustos para a geração e análise de conteúdo, eles carecem de uma verdadeira noção de verdade ou realidade. Eles operam baseando-se em palavras que comumente ocorrem juntas nos dados de treinamento, e por meio de métodos como ajuste fino supervisionado, modelo de recompensa e otimização de política. No entanto, os dados de treinamento são limitados tanto temporalmente quanto pela sua natureza retrospectiva, o que pode introduzir riscos epistemológicos. Essa limitação significa que, embora os *chatbots* possam produzir conteúdo que soa coerente, ele pode ser inexato ou fabricado, fenômeno esse conhecido como alucinação [44].

2.4.2 Arquitetura Geral

Desenvolver um *chatbot* envolve uma série de técnicas e etapas. Definir o que o *chatbot* irá oferecer e em quais categorias se enquadra auxilia os desenvolvedores na escolha dos algoritmos, plataformas e ferramentas para essa construção. Além disso, essas

definições também são importantes para que os usuários finais entendam o que devem esperar. Para projetar um *chatbot*, é necessária uma representação precisa do conhecimento, uma estratégia de geração de respostas e uma resposta neutra predefinida para responder quando a pergunta do usuário não é compreendida [1]. A Figura 10 mostra a arquitetura geral de um *chatbot*.

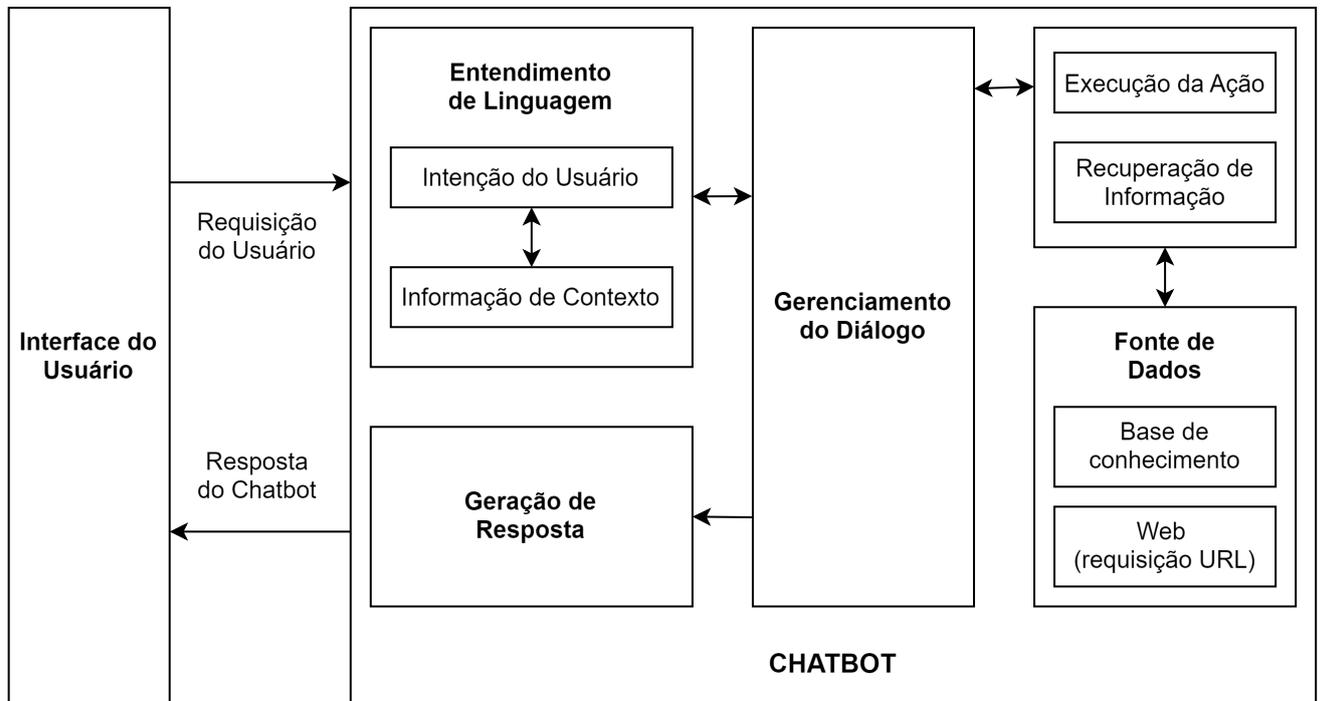


Figura 10 – Arquitetura geral de um *chatbot*. Traduzido de [1].

O processo inicia com uma pergunta do usuário ao *chatbot* por meio de um aplicativo de mensagem (ex: Facebook, Slack, WhatsApp) ou por um aplicativo com funcionalidade de entrada de texto/voz. Então, o Componente de Entendimento de Linguagem, analisa essa pergunta para determinar a intenção do usuário e os detalhes relacionados [1]. A intenção representa a tarefa que o usuário pretende que o *chatbot* realize por meio de sua frase [32]. Os detalhes relacionados, por sua vez, são as informações relevantes para a realização da tarefa, em casos de *chatbots* que armazenam dados de contexto [1].

Uma vez interpretada a solicitação, o *chatbot* decide o próximo passo a tomar, que pode variar entre agir baseado na informação recebida, guardar o entendimento para ações futuras, solicitar mais dados contextuais ou pedir esclarecimentos ao usuário. Após compreender a solicitação, o *chatbot* executa as ações necessárias ou busca as informações requeridas em suas fontes de dados, como uma base de conhecimento interna ou por APIs externas. Então, o Componente de Geração de Resposta, formula uma resposta em linguagem próxima à humana, fundamentada na intenção e contexto identificados

na análise da mensagem do usuário. As respostas são geradas por um dos três métodos descritos: baseado em regras, baseado em recuperação ou generativos. O Componente de Gestão de Diálogo do *chatbot* gerencia e atualiza o contexto da conversa, incluindo a intenção atual, as entidades identificadas ou faltantes, além de solicitar informações adicionais, processar pedidos de esclarecimento e/ou realizar perguntas complementares [1].

Existem várias opções comerciais e de código aberto para o desenvolvimento de *chatbots*. Eles podem ser desenvolvidos usando linguagens de programação (como Java, Clojure, Python, C++, PHP, Ruby, e Lisp) ou por meio de plataformas especializadas (como Google's DialogFlow, Facebook's wit.ai, Microsoft LUIS, IBM Watson Conversation, Amazon Lex e SAP Conversation AI) [1].

2.5 Trabalhos Correlatos

No departamento de computação da Universidade Federal de Ouro Preto, foi desenvolvido o *chatbot* Helena para auxiliar os estudantes em questões universitárias. Neste estudo, foram apresentados a arquitetura, a implementação e o funcionamento do *chatbot*. Seu objetivo é auxiliar os alunos de graduação de computação a sanar dúvidas recorrentes sobre trâmites universitários, com a intenção de reduzir o tempo gasto pelo colegiado do curso respondendo a *e-mails* repetitivos [45].

O *chatbot* foi implementado utilizando a plataforma *Watson Assistant da IBM*, integrada com um *website* e um *dashboard* para coletar o *feedback* dos usuários. Foi utilizada a plataforma Heroku para hospedagem do *website*. A interação do *chatbot* com os usuários foi realizada por meio do Facebook Messenger, seguindo as diretrizes do *Watson da IBM* para treinamento de intenções a partir de abordagens dos usuários [45].

O *chatbot* Helena foi testado para medir sua acurácia e eficácia, sendo conduzidos experimentos para avaliar seu desempenho. Os testes foram realizados sem um roteiro definido, possibilitando interações espontâneas com os usuários. Os dados durante o teste demonstraram que o *chatbot* Helena obteve resultados satisfatórios em termos de experiência de uso, acurácia e eficiência, indicando um nível de maturidade do projeto que o torna viável para disponibilização ao público. Além disso, foram propostas melhorias futuras, como a criação de um sistema próprio para processamento de linguagem natural e a expansão da base de treino para oferecer uma experiência ainda melhor ao conversar com o *chatbot* [45].

Outro *chatbot* informativo é o Dr. Insulina, desenvolvido para fornecer informações sobre diabetes tipo 2. Ele busca disponibilizar informações relevantes para pacientes e pessoas interessadas em aprender mais sobre essa doença. O Dr Insulina foi implementado utilizando a abordagem baseada em regras por meio do Botfather, uma ferramenta para

criar e gerenciar *bots* do Telegram. Para criar um *chatbot* no Telegram basta traçar a ordem da conversa para que o *bot* siga essa lógica, construí-lo no editor visual e testá-lo imediatamente para confirmar que funciona. Podem ser feitas as alterações necessárias e criados cenários adicionais para as conversas no Telegram, a fim de tornar o *chatbot* mais funcional [46].

O *chatbot* desenvolvido foi submetido a um teste de usabilidade para avaliar sua precisão, funcionalidade, facilidade de uso, clareza das informações e satisfação dos usuários. A avaliação foi realizada com 18 estudantes do curso de inteligência artificial da *Benemérita Universidad Autónoma de Puebla*, utilizando métricas padronizadas de usabilidade e de satisfação. Os resultados obtidos indicaram que o *chatbot* obteve uma qualificação qualitativa boa, com uma média de avaliação variando de 7,77 a 8,66. Esse resultado demonstra que o *chatbot* foi bem recebido pelos usuários e atendeu às expectativas estabelecidas durante o desenvolvimento [46].

Por fim, destaca-se o estudo de Carvalho [32], que desenvolveu um *chatbot* que fornece informações sobre a COVID-19. O *chatbot* implementado centraliza e permite o acesso a diferentes tipos de informação, como sintomas da doença ou número de vacinas administradas diariamente. [32].

O *chatbot* foi desenvolvido usando os serviços cognitivos da *Azure LUIS* e *QnA Maker*. O *Azure LUIS* é uma plataforma de comunicação empregada para reconhecer a intenção da pergunta do usuário. Além disso, esta ferramenta também tem a capacidade de extrair dados da frase fornecida. Por outro lado, o *QnA Maker* é um serviço de processamento de linguagem natural baseado na nuvem que possibilita o desenvolvimento de uma interface de conversação natural sobre o conjunto de dados. Ele é empregado para determinar a resposta mais adequada para qualquer consulta na base de conhecimento personalizada. Foi implementado também um sistema que permite a monitorização e análise do desempenho do *chatbot* em tempo real [32].

O *chatbot* foi testado pela exposição a pessoas externas ao desenvolvimento e pelas métricas de desempenho coletadas. Inicialmente, os resultados do teste demonstraram um desempenho não satisfatório do *chatbot*, pois ele disponibilizou respostas corretas ao utilizador em apenas 33% das vezes. Contudo, a coleta de dados de desempenho e de satisfação dos usuários permitiu a definição e implementação de modificações que elevaram a percentagem de respostas corretas de 33% para 67%. Esses resultados demonstraram que o *chatbot* foi eficaz em fornecer informações sobre a COVID-19 e que as alterações efetuadas com base nas informações recolhidas durante os testes tiveram o efeito desejado, ou seja, a melhoria do desempenho do *chatbot* [32].

3 DESENVOLVIMENTO

Uma vez que o objetivo deste trabalho é desenvolver um *chatbot* informativo sobre saúde mental, sua arquitetura foi projetada tendo em vista a necessidade de promover informações precisas e confiáveis sobre o tema. Para alcançar uma abordagem que garanta a entrega de informações científicas e cuidadosamente preparadas, optou-se por desenvolver um *chatbot* informativo de domínio fechado, com modelo baseado em recuperação.

Toda a arquitetura do *chatbot* proposto se encontra no *back-end* de um *website* desenvolvido em Django. O *website* conta com quatro telas, a tela inicial, a tela para interação com o *chatbot*, a tela para preenchimento do formulário de avaliação do *chatbot* e a tela de conclusão do processo. A seguir, serão apresentados a arquitetura geral do *chatbot* proposto, o processo de criação da base de conhecimento, as etapas de pré-processamento realizadas, o desenvolvimento do classificador de textos, as interfaces do *website* e, por fim, as métricas de avaliação do *chatbot*.

3.1 Arquitetura do *Chatbot* Proposto

A arquitetura geral do *chatbot* proposto (Figura 11) passa por três principais módulos. O processo se inicia com uma pergunta do usuário ao *chatbot*, realizada por uma interface de entrada de texto. A pergunta é então encaminhada como uma requisição para o módulo de Recuperação da Informação. Este módulo, por sua vez, realiza as seguintes etapas de pré-processamento, respectivamente: tokenização, lematização, retirada de acentos e remoção de *stop words*.

Após o pré-processamento, o módulo de Recuperação da Informação aplica a representação de BoW na pergunta pré-processada e a encaminha para o classificador de textos. O classificador de textos recebe uma representação BoW e prediz a intenção desse vetor de características por meio de um modelo de aprendizado de máquina previamente treinado. A intenção predita é retornada ao módulo de Recuperação de Informação que a utiliza para consultar na base de conhecimento qual a resposta pré-definida para essa intenção. Finalmente, o módulo de Recuperação da Informação envia a resposta para a interface do usuário.

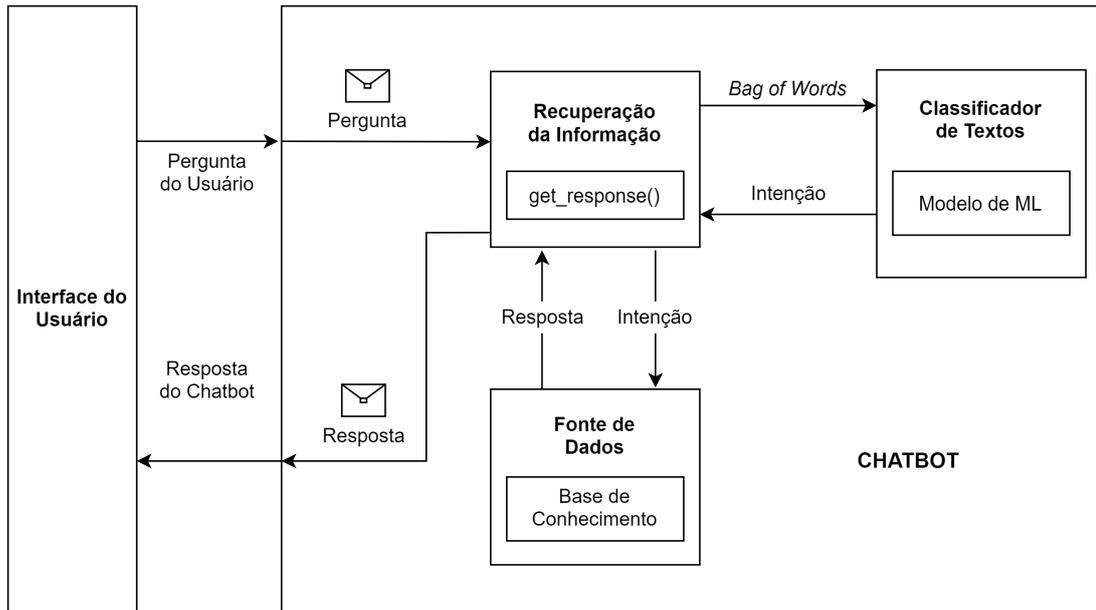


Figura 11 – Arquitetura geral do *chatbot* proposto.

3.2 Base de Conhecimento

Inicialmente, foi construída a base de conhecimento que seria utilizada para treinar o classificador de textos do *chatbot*. Para isso, foi utilizada a cartilha “Manual da Saúde Mental” desenvolvida pelos participantes do projeto de saúde mental da UEL, apresentado no Capítulo 1. A cartilha, que conta com seis capítulos, foi subdividida em tópicos manualmente. A ideia é quebrar os capítulos em unidades de conhecimento menores (os tópicos) para transformá-las em possíveis respostas do *chatbot*. Para cada tópico identificado, foi definido um rótulo (intenção) e foram criadas possíveis perguntas do usuário para o tema daquele tópico. A Figura 12 ilustra o processo de criação da base de conhecimento.

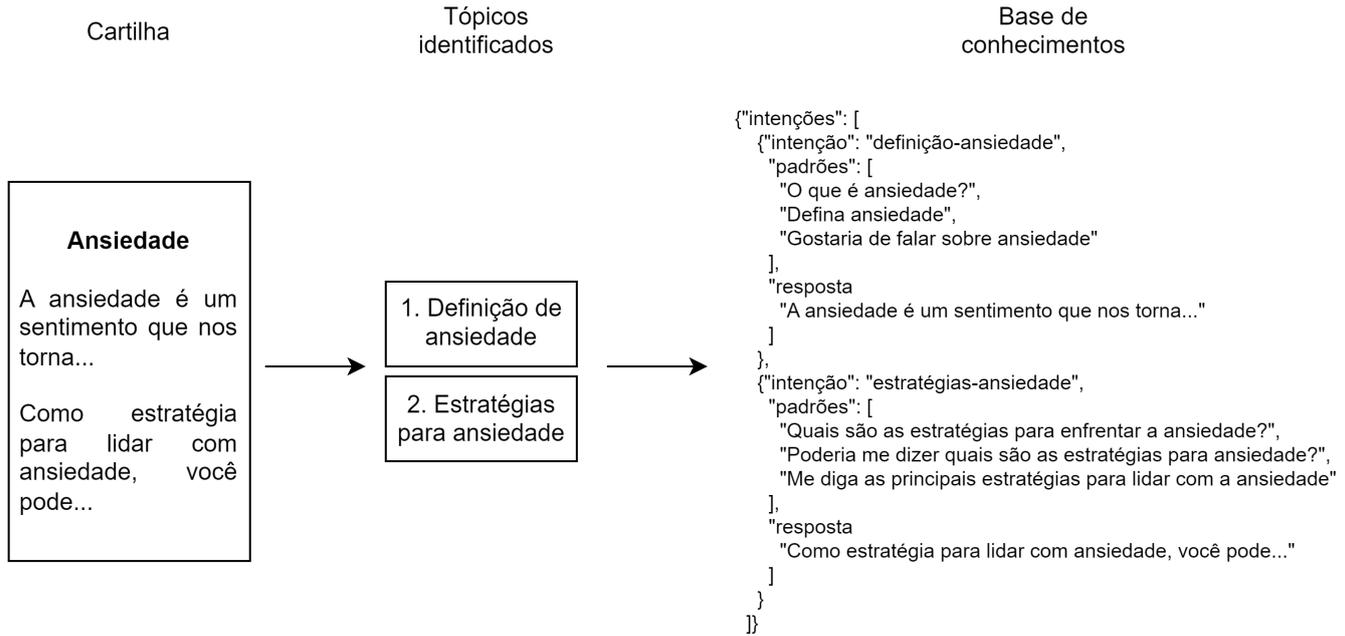


Figura 12 – Exemplo do processo de criação dos padrões intenção-resposta para treinamento do *chatbot*.

Ao todo foram identificados 38 intenções. A Tabela 1 apresenta uma amostra das dez primeiras intenções identificadas. A tabela completa encontra-se no Apêndice A.

Tabela 1 – Amostra das intenções identificadas na cartilha de saúde mental.

Número	Intenção
1	saudacoes
2	definicao-ansiedade
3	identificar-ansiedade
4	lidar-ansiedade
5	definicao-estresse
6	identificar-estresse
7	lidar-estresse-e-estrategias-estresse-e-ansiedade
8	definicao-luto
9	identificar-luto
10	lidar-luto

3.3 Pré-Processamento

O pré-processamento da base de conhecimento é a etapa que prepara os dados para o treinamento de um modelo de AM. Conforme detalhado na seção anterior, o *dataset* desenvolvido neste trabalho consiste em um arquivo JSON, que contém várias intenções

associadas a diferentes padrões de frases. Inicialmente, cada padrão de frase dentro das intenções é submetido a um processo de tokenização utilizando a função `word_tokenize()` da biblioteca NLTK¹ do Python. Essa função segmenta as frases em palavras individuais.

Após a tokenização, é aplicada a lematização em cada *token* para reduzi-los à sua forma canônica. Esse processo é realizado utilizando o atributo *lemma* da biblioteca spaCy² em Python. Esse processo é importante para agrupar variações da mesma palavra, permitindo que o modelo aprenda com mais eficiência sem ser confundido por diferentes flexões.

Uma vez lematizados, os *tokens* lematizados passam por um processo de normalização utilizando a função `unidecode()` da biblioteca padrão do Python. Essa função remove acentos e caracteres especiais, garantindo a uniformidade dos dados. Isso foi feito para que o *chatbot* entendesse as perguntas do usuário mesmo que ele não inserisse os acentos nas palavras ou inserisse os acentos incorretamente.

Posteriormente, foram eliminadas as *stop words*, palavras que ocorrem com frequência na construção de sentenças na língua-alvo, mas que geralmente não carregam um significado especial. Essa remoção é útil para reduzir a dimensionalidade dos dados. As *stop words* removidas foram todas aquelas definidas na biblioteca spaCy e na NLTK [27].

Após o tratamento inicial, cada frase é transformada em um vetor de características por meio da representação de BoW. Nesse processo, implementado manualmente, cada frase é representada como um vetor onde cada posição corresponde à presença de uma palavra no dicionário extraído dos padrões. Cada elemento do vetor representa uma dimensão que o modelo de aprendizado de máquina utilizará para aprender as associações entre padrões de texto e os rótulos de intenções correspondentes. O formato do vetor de características segue o modelo apresentado na Figura 9 da Subseção 2.3.4.

As intenções também são processadas para serem utilizadas no treinamento. Cada intenção é substituída por um índice numérico correspondente, utilizado para a classificação durante o aprendizado do modelo.

A Figura 13 apresenta de forma esquemática o processo de pré-processamento das perguntas e intenções. As saídas desse *pipeline* serão as entradas do modelo de AM do classificador de textos.

¹ <https://www.nltk.org>

² <https://spacy.io>

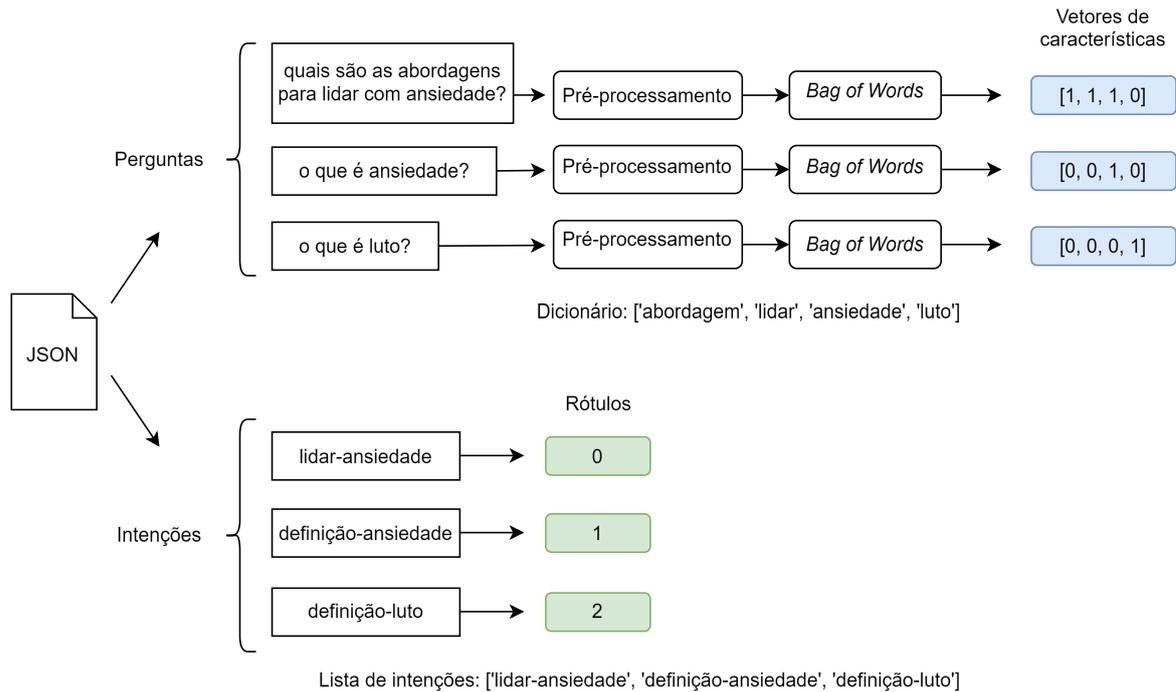


Figura 13 – Pipeline de pré-processamento.

3.4 Classificador de Textos

O classificador de textos foi implementado utilizando uma RN de propagação direta para a compreensão e classificação de intenções do *chatbot*. A RN foi implementada utilizando a biblioteca PyTorch³ do Python. Essa biblioteca disponibiliza um amplo conjunto de ferramentas e funcionalidades específicas para a construção e treinamento de modelos de aprendizado de máquina de maneira eficiente e intuitiva.

A estrutura da rede inclui três camadas lineares. A camada de entrada possui 606 neurônios, que corresponde ao tamanho do dicionário de palavras (utilizado para a representação do BoW). A camada oculta possui 10 neurônios. Por fim, a camada de saída possui trinta e oito neurônios, correspondente ao número de intenções identificadas.

A função de ativação ReLU é aplicada nas duas primeiras camadas para introduzir a não-linearidade. Na última camada, foi aplicada a função *Softmax*, dado o contexto de classificação de múltiplas classes, fornecendo uma probabilidade para cada classe como saída.

O treinamento do modelo ocorre ao longo de 100 épocas com um tamanho de lote de 5 amostras. Cada passagem de treinamento inclui uma fase de propagação para frente, onde os dados são passados através da rede para obter a saída, seguida por uma fase de retropropagação, onde o gradiente da perda é calculado e usado para atualizar os

³ <https://pytorch.org>

pesos da rede. Este processo é repetido iterativamente para minimizar a função de perda, melhorando assim a precisão das previsões do modelo. Ao final do treinamento, o estado do modelo é salvo, permitindo que o *chatbot* classifique apropriadamente às entradas dos usuários com base no aprendizado realizado. A Figura 14 apresenta de forma simplificada o processo de treinamento da RNA e o posterior uso do modelo gerado para prever a intenção de uma pergunta.

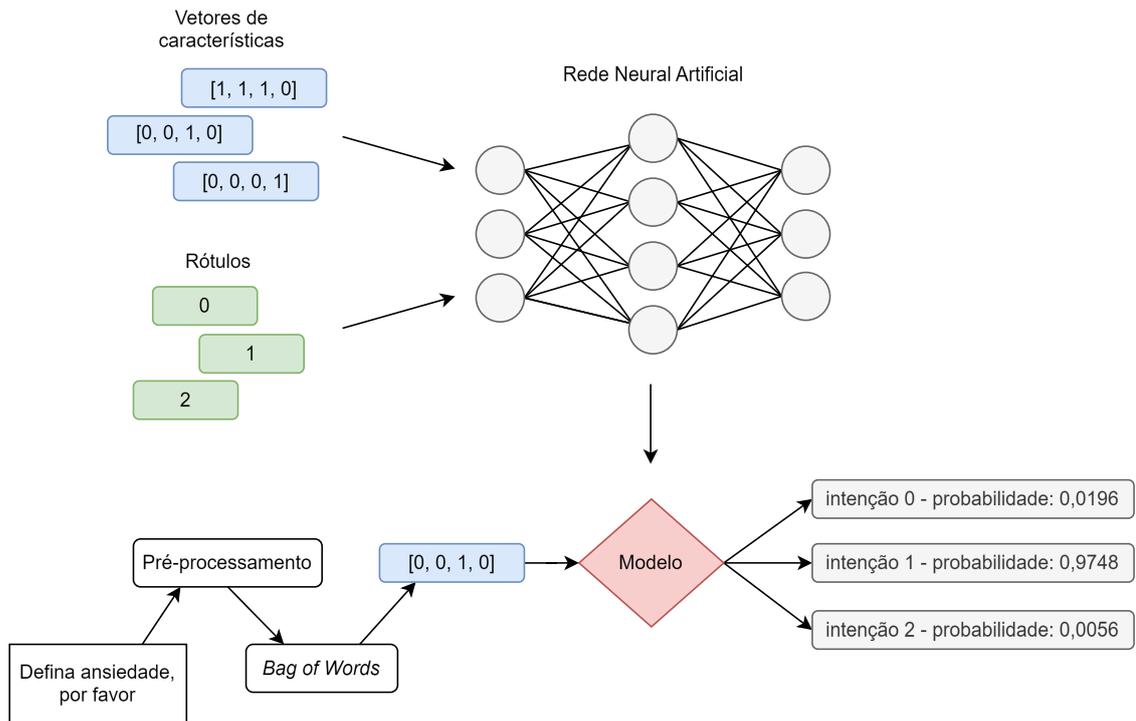


Figura 14 – Representação esquemática do classificador de textos.

Foi definido como limite de pertencimento às intenções o valor 0,55. Ou seja, caso a pergunta do usuário não atinja 0,55 de probabilidade de pertencimento a nenhuma intenção predefinida, será retornado a mensagem de erro, na qual o *chatbot* informa que não entendeu a pergunta.

3.5 Interfaces Gráficas de Usuário

Como plataforma de distribuição e avaliação do *chatbot* proposto, foi desenvolvido um *website* com o *framework* Django hospedado na plataforma Render. Os dados coletados durante o uso desse serviço, foram armazenados em uma instância do banco de dados PostgreSQL, também hospedada na plataforma Render.

A Figura 15 mostra a tela inicial do *website*. Esta tela conta com algumas explicações sobre o objetivo do *chatbot* e a necessidade de avaliá-lo. Além disso, são apresentadas algumas orientações úteis para guiar o usuário no processo. Os usuários foram orientados

a fazer pelo menos dez perguntas ao *chatbot* e também a tentar abranger diferentes tópicos em suas perguntas. Cabe ressaltar que essas orientações são dadas exclusivamente para a fase de testes.

Home

Bem-vindo ao Alkin:
Um Chatbot Informativo sobre Saúde Mental!

Caro Profissional da Saúde,

Este projeto foi desenvolvido com o objetivo de fornecer informações cientificamente embasadas sobre sentimentos, comunicação efetiva, interação com pacientes, entre outros temas relacionados à saúde mental. Você é nosso convidado a participar da avaliação deste chatbot. Estamos muito felizes em ter você aqui!

Por que avaliar este Chatbot?

A avaliação do chatbot é uma etapa importante para garantir a qualidade e a eficácia do serviço prestado. Seu feedback será utilizado exclusivamente para validar a efetividade desta ferramenta.

Como funciona a avaliação?

Ao fim desta introdução, você deve clicar no botão "Iniciar Teste", que te direcionará para interagir com o chatbot. Durante a interação, orientamos que você:

- Faça pelo menos 10 perguntas sobre os temas que deseja saber mais.
- Tente abranger diferentes tópicos em suas perguntas.
- Enquanto estiver na tela do chat NÃO atualize a tela.

Após esse processo, incentivamos você a compartilhar sua experiência clicando no botão "Avaliar Chatbot!" que estará disponível na barra superior do site.

A avaliação é rápida e simples. Você será convidado a responder algumas perguntas sociodemográficas e sobre sua experiência com o chatbot. Suas respostas serão usadas exclusivamente para fins de validação deste serviço.

Vamos lá?!

Iniciar Teste

Figura 15 – Tela inicial do *website*.

Ao clicar no botão “Iniciar Teste” o usuário é direcionado à principal tela do sistema, a tela do *chatbot* (Figura 16). Logo que ingressa nessa tela, o *chatbot* mostra a mensagem de intenção “saudacoes”. Esta mensagem é a abertura da conversa, ela tem o intuito de apresentar o personagem com quem o usuário irá interagir (o pinguim Alkin). Na mensagem de saudações também são citados os temas macros que o *chatbot* aborda. Isso é feito com o intuito de guiar o usuário nas perguntas, de forma a informá-lo sobre quais temas o *chatbot* pode responder.

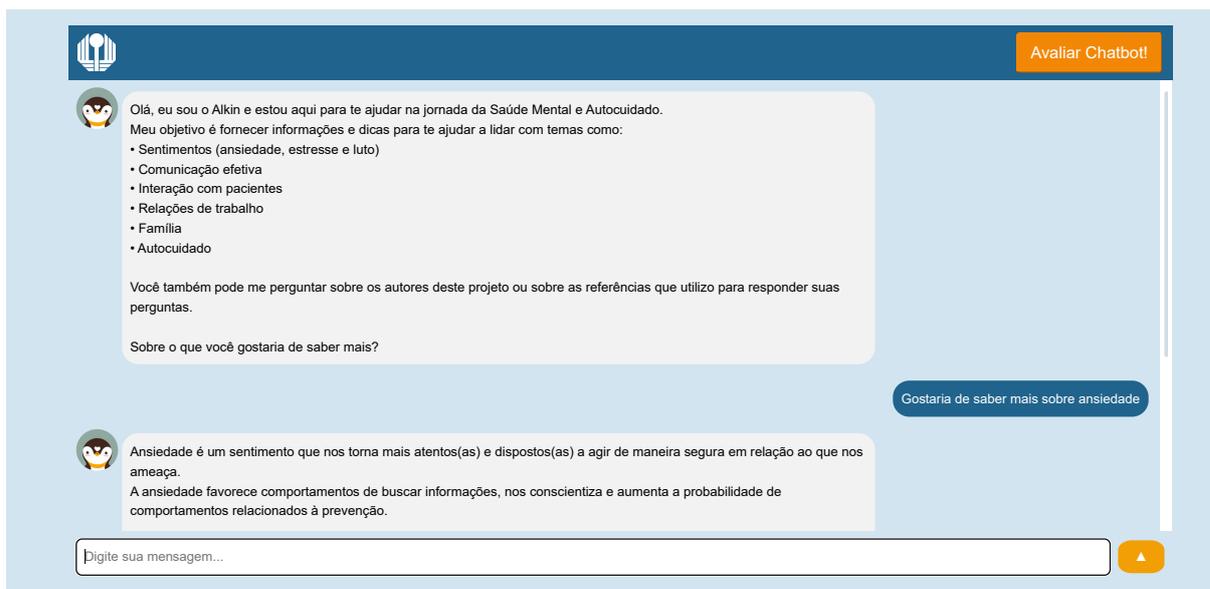


Figura 16 – Tela de interação com o *chatbot*.

Na mensagem de saudações, os temas citados são referentes a cada um dos seis capítulos da cartilha de saúde mental. Conforme o usuário pergunta sobre um destes temas, ele também é informado sobre outras possíveis perguntas que pode fazer ao *chatbot*. Por exemplo, ao perguntar sobre ansiedade, o *chatbot* retorna a resposta da intenção “definicao-ansiedade”. Essa resposta contém o conceito de ansiedade e também contém a indicação de que o usuário pode perguntar mais sobre como lidar com a ansiedade ou sobre estratégias para ansiedade, por exemplo (Figura 17).

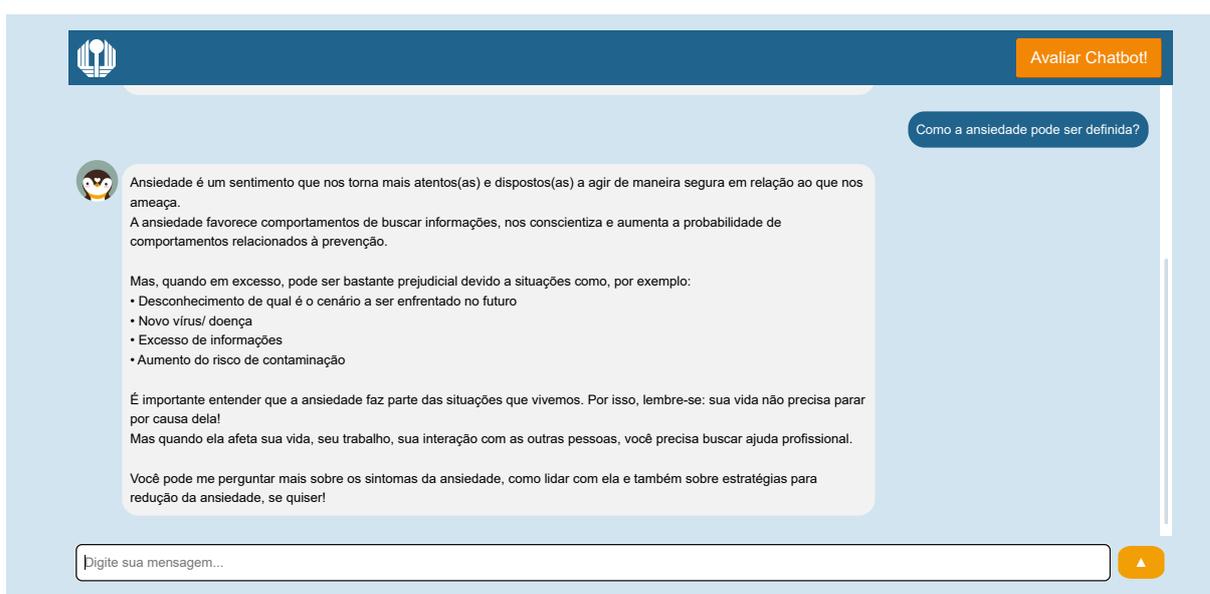


Figura 17 – Exemplo de interação com o *chatbot* sobre ansiedade.

Quando desejar avaliar o *chatbot*, o usuário clica no botão “Avaliar *Chatbot*” posicionado no canto superior direito da tela. A tela de avaliação do *chatbot* (Figura 18) conta com algumas perguntas sociodemográficas e também com perguntas específicas sobre a experiência com o *chatbot*. Em relação às perguntas sociodemográficas, o nome do usuário foi coletado apenas para garantir que o mesmo usuário não iria avaliar o sistema mais de uma vez, mesmo que por diferentes dispositivos. Logo após a coleta de dados, essa informação foi excluída para fins de segurança dos dados. Foram perguntados também a idade, grau de escolaridade dos participantes e se foi a primeira experiência deles com algum *chatbot*. No que diz respeito às perguntas de avaliação do *chatbot*, elas serão detalhadas na seção a seguir juntamente com as outras métricas de avaliação.

Questionário de Avaliação da Experiência do Usuário

Informações Sociodemográficas

Nome Completo

Idade

Você é graduando em psicologia ou psicólogo formado?

Escolaridade

Foi a primeira vez que utilizou um chatbot?

Uso do Chatbot

Classifique as sentenças abaixo em uma escala de 1 a 5 (discordo totalmente a concordo totalmente)

O chatbot entende o que está sendo perguntado.

A velocidade de interação com o chatbot é coerente.

As respostas obtidas frequentemente respondem às questões perguntadas.

De forma geral, estou satisfeito com o desempenho do chatbot.

O que você achou dessa experiência?

Enviar Avaliação

Figura 18 – Tela de avaliação do *chatbot*.

Por fim, após finalizar a avaliação, o usuário é redirecionado para a tela final do *website* (Figura 19). Cabe ressaltar que a paleta de cores utilizada no *website* bem como as artes do pinguim Alkin foram desenvolvidas dentro do projeto Manual da Saúde Mental.



Figura 19 – Tela final do *website*.

3.6 Métricas de Avaliação

Durante a interação com o *chatbot*, foram coletadas todas as perguntas feitas e as respectivas intenções identificadas pelo *chatbot*. Posteriormente, esses pares de pergunta e intenção são manualmente avaliadas para identificar se o *chatbot* classificou a intenção correta para cada pergunta. Essa informação é de extrema importância para mensurar a acurácia do *chatbot* na identificação das intenções dos usuários.

Com o objetivo de investigar a experiência dos usuários, foi aplicado o questionário de avaliação do *chatbot* ao final da interação. Este questionário é de extrema importância para identificar possíveis modificações necessárias nesta primeira versão do *chatbot*.

A Tabela 2 apresenta as perguntas que constituem o formulário, com como as suas respectivas escalas de avaliação. As perguntas do questionário seguiram, na sua maioria, uma escala de Likert [47]. Nessas perguntas os usuários foram instruídos a classificar as sentenças em uma escala de 1 a 5, onde 1 significa “discordo totalmente” e 5 significa “concordo totalmente”.

Tabela 2 – Caracterização das perguntas do questionário de avaliação do *chatbot*.

Identificador	Pergunta	Escala de resposta
1	Foi a primeira vez que utilizou um <i>chatbot</i> ?	“Sim” ou “Não”
2	O <i>chatbot</i> entende o que está sendo perguntado.	Escala de Likert
3	A velocidade de interação com o <i>chatbot</i> é coerente.	Escala de Likert
4	As respostas obtidas frequentemente respondem às questões perguntadas.	Escala de Likert
5	De forma geral, estou satisfeito com o desempenho do <i>chatbot</i> .	Escala de Likert
6	O que você achou dessa experiência?	Campo aberto

Em relação ao público-alvo para avaliação do *chatbot*, foi escolhida um grupo particular de profissionais da área da saúde, os psicólogos. Nesta primeira avaliação, estudantes de graduação em psicologia também foram convidados a testar o sistema.

4 RESULTADOS E DISCUSSÃO

4.1 Classificador de Textos

A base de conhecimento gerada para treinamento do classificador de textos, possui 38 diferentes intenções e 996 padrões de perguntas distribuídas entre elas. Foi gerado um dicionário de 606 palavras distintas.

Neste trabalho, a RN implementada apresentou uma perda final de apenas 0,0001. Este valor indica que o modelo conseguiu aprender com uma alta precisão as relações e padrões dos dados de treinamento para classificar as intenções da base de conhecimento. Um valor de perda tão baixo sugere que o modelo está bem ajustado aos dados, minimizando o erro entre as previsões e as respostas reais. No entanto, é importante também considerar a possibilidade de *overfitting*, onde o modelo se ajusta perfeitamente aos dados de treinamento, mas falha em generalizar para novos dados não vistos [15]. Por essa razão, é necessário analisar esse valor de perda juntamente com as demais métricas de desempenho apresentadas a seguir.

4.2 Histórico de Perguntas

Todas as perguntas feitas pelos participantes desta pesquisa foram armazenadas pelo *website* em uma instância do PostgreSQL, bem como as intenções preditas pelo classificador de textos do *chatbot*. Ao todo, foram feitas 188 perguntas ao *chatbot* pelos 22 participantes da avaliação. Desde logo, foi possível observar que nem todos os participantes seguiram a orientação de realizar pelo menos 10 perguntas. Isso porque, 12 participantes fizeram menos de 10 perguntas ao *chatbot*. Destes, a metade fez entre 4 a 9 perguntas, mas o resto fez entre 1 a 3 perguntas apenas. A média de perguntas por participante foi de 8,54 perguntas. Apesar disso, houve participantes que fizeram mais do que 10 perguntas, chegando até a 21 perguntas. A Seção 4.2 detalha a quantidade de perguntas feitas por cada participante.

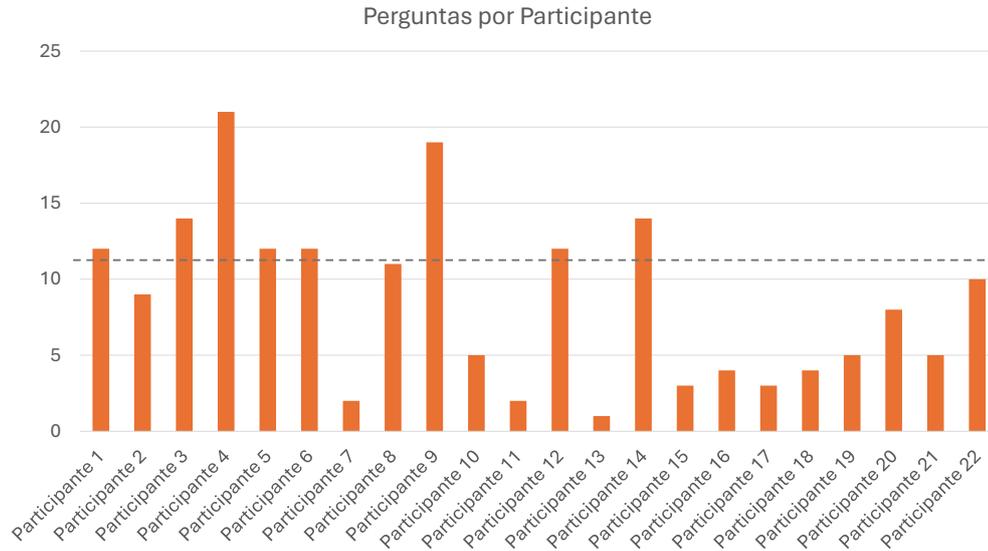


Figura 20 – Quantidade de perguntas feita por cada participante.

Todas as 188 perguntas foram manualmente classificadas como “correta”, “errada” ou “indefinida”. Considera-se “correta” a pergunta do usuário cuja intenção foi corretamente classificada pelo *chatbot*. A pergunta “errada” é aquela que o *chatbot* atribuiu incorretamente a intenção (por exemplo, o usuário pergunta o conceito de ansiedade e o *chatbot* responde o conceito de estresse). Por fim, a pergunta “indefinida” é aquela que o *chatbot* não conseguiu responder, pois nenhuma intenção atingiu a probabilidade mínima utilizada.

Constatou-se que das 188 perguntas feitas ao *chatbot*, 113 foram classificadas como corretas, 47 foram classificadas como erradas e 28 foram classificadas como indefinidas (Figura 4.2). Isso significa que 60,10% das perguntas foram corretamente respondidas, 14,89% receberam a resposta “Desculpe, não entendi a pergunta...” e 25,00% das perguntas receberam uma resposta de outra intenção. Essa acurácia de 60,10% na primeira versão do *chatbot* é bastante satisfatória comparada a outros *chatbots* informativos, como o de Carvalho [32], que na primeira versão de seu *chatbot*, a acurácia foi de 36%.

Quantidade de Respostas Corretas, Erradas e Indefinidas

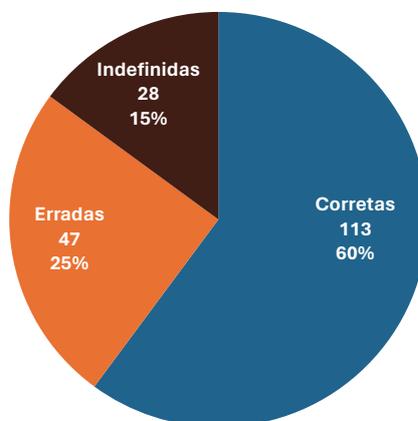


Figura 21 – Representatividade das perguntas Corretas, Erradas e Indefinidas.

Ao analisar mais detalhadamente as perguntas erradas e indefinidas, foi possível notar alguns pontos importantes. Uma possível perspectiva de análise dessas perguntas passa pela classificação do assunto em “fora” ou “dentro” do escopo de treinamento do *chatbot* (Tabela 3).

Tabela 3 – Caracterização das perguntas não respondidas ou respondidas incorretamente pelo *chatbot*.

Tema da Pergunta	Erradas	Indefinidas
Dentro do escopo	22 (46,80%)	12 (42,85%)
Fora do escopo	25 (53,19%)	16 (57,14%)
TOTAL	47 (100%)	28 (100%)

De todas as 28 perguntas indefinidas, 57,14% delas eram perguntas sobre temas não contemplados na cartilha, ou seja, perguntas que o *chatbot* não havia sido preparado para responder. Os temas dessas perguntas passavam por autoestima, depressão, transtornos alimentares, autocontrole e transtorno bipolar, por exemplo. Também houve perguntas sobre o desenvolvimento deste trabalho em si, como “de onde surgiu a ideia desse projeto?” e “qual o objetivo desse projeto de um chat?”. Esses temas não estão contemplados na cartilha de saúde mental utilizada para a criação da base de conhecimento do *chatbot*. Dessa forma, já era esperado que perguntas como essas não obteriam resposta. Apesar

disso, é importante destacar o interesse dos participantes em saber mais sobre esses temas. Ademais, basta introduzir um novo tipo de intenção na base de conhecimento do *chatbot* que dispõe a resposta à pergunta em questão para que essa questão seja resolvida.

Em relação às perguntas erradas fora do escopo da base de conhecimento, estas correspondem a 53,19% das perguntas erradas. Nesses casos, mesmo que a pergunta não correspondesse a nenhuma intenção da base de conhecimentos, o classificador de textos tentou atribuir uma resposta a ela. Isso se deve ao fato de que mesmo após o pré-processamento, algumas palavras da pergunta existiam no dicionário de palavras do classificador. Essas palavras foram suficientes para que o modelo atribuísse alguma probabilidade maior que o limite estabelecido (de 0,55). Isso pode ter acontecido pelo fato de o limite estabelecido para classificação ser baixo, o que permitiria que perguntas indefinidas fossem classificadas erroneamente.

Outra possível razão das perguntas fora do escopo classificadas incorretamente é a ausência de um peso maior para as palavras-chave dos temas da base de conhecimento (como ansiedade, luto e estresse, por exemplo). A representação BoW utilizada armazena apenas a frequência da palavra na pergunta, entretanto, armazenar a frequência não se mostrou muito eficiente nesta proposta. Isso porque as perguntas geralmente foram curtas e sem repetição de palavras-chave do tema. Isso faz com que, na pergunta “como lidar com ansiedade”, seja atribuído o mesmo peso na palavra “lidar” e na palavra ansiedade “ansiedade”. Quando a pergunta é “como lidar com depressão?”, o mesmo acontece com “lidar” e “depressão” e, apesar de essa pergunta não estar na base de conhecimento e de “depressão” não estar no dicionário de palavras, “lidar” está. Mesmo apenas uma palavra (“lidar”) codificada em BoW irá gerar alguma probabilidade pelo modelo.

Em relação às perguntas indefinidas que estavam dentro do escopo deste trabalho, estas correspondem a 42,85% de todas as perguntas indefinidas. Observou-se nesses casos a ocorrência de erros ortográficos e também de perguntas de contexto. Apesar de o *chatbot* lidar com a falta ou uso incorreto de acentos (tratado no pré-processamento), quando houve erros ortográficos, ele não foi capaz de identificar a intenção corretamente. Um exemplo disso foi na ocorrência de “mindfulness” ao invés de “mindfullness”. Nesse caso, a palavra “mindfulness” não foi identificada no dicionário, então a classificação aconteceu com base nas outras palavras da pergunta.

Por fim, as perguntas erradas que estavam dentro do escopo correspondem a 46,80% das perguntas erradas. Foram identificados casos semelhantes aos das perguntas indefinidas que estavam dentro do escopo, como os erros ortográficos e perguntas de contexto. Um caso específico de erro ortográfico que vale a pena destrinchar é o de “auto cuidado”, pois ele gera um efeito totalmente diferente do outro erro citado anteriormente. Ao receber “auto cuidado” ao invés de “autocuidado”, a tokenização quebrou “auto cuidado” em duas palavras distintas. Isso impede a correta representação do BoW

e, conseqüentemente, gera uma classificação errada da intenção do usuário.

Sobre as perguntas de contexto, citadas nas duas últimas classes, percebeu-se que alguns usuários formularam perguntas referentes a mensagens anteriores. Por exemplo, perguntavam sobre o conceito de ansiedade, o *chatbot* respondia corretamente, depois perguntavam “como lidar com ela?”, referindo-se à ansiedade. Entretanto, nesses casos o *chatbot* falhou em responder, pois não armazena o contexto do diálogo. É interessante notar que os três participantes que fizeram perguntas de contexto já haviam utilizado algum *chatbot* antes. O armazenamento de contexto nos principais *chatbots* comerciais pode justificar essa expectativa por parte dos participantes.

Apesar da dificuldade em lidar com erros ortográficos, o *chatbot* teve um bom desempenho em lidar com erros de acentuação. Alguns exemplos de erros ortográficos que foram facilmente normalizados no pré-processamento foram “familia”, “estrategias” e “comunicação”.

4.3 Avaliação dos Participantes

O questionário de avaliação foi uma importante fonte de informação sobre a opinião dos participantes e como estes classificaram a sua interação com o *chatbot*. Ele possibilita a identificação de problemas que são invisíveis às outras análises mais técnicas.

Ao todo 22 pessoas realizaram a avaliação do *chatbot* desenvolvido neste trabalho. Destas, 10 (45,45%) são estudantes de graduação em psicologia. A Seção 4.3 mostra a participação dos estudantes em cada nível de escolaridade.



Figura 22 – Escolaridade dos participantes da validação do *chatbot*.

Dos 22 participantes, 12 deles (55%) relataram ter sido sua primeira experiência

com um *chatbot*. Este indicador é consideravelmente maior do que o esperado, uma vez que os *chatbots* são uma tecnologia emergente que tem sido utilizado em diversos setores. Uma hipótese para isso é no contexto acadêmico de psicologia, esse tipo de tecnologia não seja tão utilizado. Outra hipótese é que a forma como foi perguntado ao usuário (“Foi a primeira vez que utilizou um *chatbot*?”) não tenha ficado clara.

Foi a primeira vez que utilizou um *chatbot*?

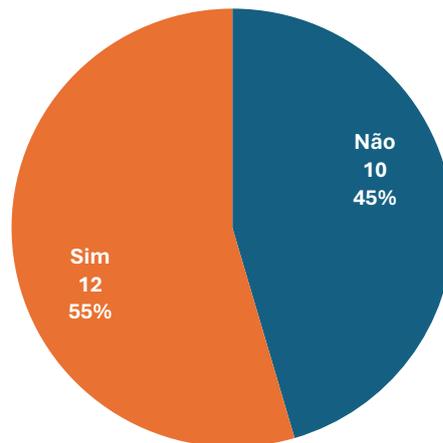


Figura 23 – Representatividade de participantes que já utilizaram algum *chatbot* antes.

Em seguida, foram feitas quatro perguntas a serem respondidas na escala de Likert. Os participantes deveriam analisar as sentenças e classificá-las de 1 (discordo totalmente) a 5 (concordo totalmente). A Tabela 4 apresenta a compilação dos resultados dessas quatro perguntas.

Tabela 4 – Resultado da avaliação do *chatbot* nas perguntas em escalas Likert.

Critério	1	2	3	4	5
O <i>chatbot</i> entende o que está sendo perguntado.	0 (0,00%)	5 (22,73%)	4 (18,18%)	7 (31,82%)	6 (27,27%)
A velocidade de interação com o <i>chatbot</i> é coerente.	0 (0,00%)	2 (9,09%)	0 (0,00%)	7 (31,82%)	13 (59,09%)
As respostas obtidas frequentemente respondem às questões perguntadas.	0 (0,00%)	4 (18,18%)	3 (13,64%)	8 (36,36%)	7 (31,82%)
De forma geral, estou satisfeito com o desempenho do <i>chatbot</i> .	0 (0,00%)	3 (13,64%)	4 (18,18%)	5 (22,73%)	10 (45,45%)

A primeira sentença (“O *chatbot* entende o que está sendo perguntado.”) teve o objetivo de entender a perspectiva do participante sobre a compreensão das perguntas por parte do *chatbot*. Nesta sentença as respostas ficaram bem distribuídas entre as classes 2, 3, 4 e 5. Quarto usuários classificaram a sentença como neutra e 5 como “discordo”, isso provavelmente se deve às perguntas fora do tema e com erros ortográficos mínimos, que o *chatbot* não foi capaz de resolver.

Já em relação à sentença “A velocidade de interação com o *chatbot* é coerente.”, o objetivo era mensurar a satisfação do usuário em relação à velocidade de “escrita” do *chatbot*. Ao responder à pergunta, ao invés de simplesmente apresentar a mensagem na tela, foi implementado um efeito para que as palavras aparecessem aos poucos, como se o personagem estivesse realmente digitando a resposta. Características como essa são implementadas em outros *chatbots* para soarem mais acolhedores. Foi constatado que 31,82% concordaram e 59,09% dos participantes concordaram totalmente com a afirmação. Isso indica que a velocidade de digitação do *chatbot* foi confortável para a maioria dos participantes. No entanto, os dois participantes que discordaram da sentença também escreveram no campo de feedback que gostariam que a resposta fosse escrita mais rapidamente na tela.

No que diz respeito a sentença “As respostas obtidas frequentemente respondem às questões perguntadas.”, os resultados foram satisfatórios. A maioria dos participantes (68,18%) concordaram com a afirmação (marcaram 4 ou 5). Isso reflete o desempenho do sistema na classificação das respostas, com assertividade em 60% das perguntas realizadas.

Na última sentença, os participantes avaliam a sentença “De forma geral, estou satisfeito com o desempenho do *chatbot*.”. Nessa afirmação, 22,73% dos participantes concordaram e 45,45% dos participantes concordaram totalmente. Enquanto isso, 13,64% dos participantes discordaram e 18,18% se mantiveram neutros. Esses resultados levam a

concluir que, em geral, os participantes classificaram a experiência como satisfatória.

Para terminar a análise dos formulários, foram analisados os comentários submetidos. De modo geral, as respostas vão ao encontro da satisfação geral dos participantes. Um participante disse que a interação com o *chatbot* torna a experiência do usuário mais agradável por entregar o conteúdo buscado de forma pontual, uma vez que o conteúdo é denso e complexo. Alguns participantes citaram os momentos em que o *chatbot* não entendeu a pergunta, destacando que ele se saia melhor quando as perguntas feitas eram semelhantes às sugeridas pelo próprio *chatbot*. Houve também um feedback em relação ao tamanho das respostas, nessa contribuição, o participante sugeriu que as informações fossem apresentadas de forma mais sucinta. Por fim, alguns participantes destacaram que gostariam que o repertório do *chatbot* fosse aumentado, para que pudesse responder perguntas de outros temas de saúde mental também.

5 CONCLUSÃO

Neste trabalho, foi desenvolvido e avaliado um *chatbot* informativo sobre saúde mental. A base de conhecimento utilizada para treinar o *chatbot* foi desenvolvida com base na cartilha Manual da Saúde Mental. A cartilha conta com diferentes temas como sentimentos, comunicação efetiva, família e interação com pacientes, por exemplo. Dessa cartilha foram extraídas possíveis intenções do usuário e a respectiva resposta que o *chatbot* retornaria. A base de conhecimento foi submetida a uma etapa de pré-processamento, na qual foram realizadas a tokenização, a lematização, a remoção de acentos e a remoção de *stop words*. Então, foi utilizado o método de BoW para representação numérica dos dados. Estes dados foram submetidos a uma rede neural de propagação direta para criação de um classificador de textos. Por fim, o *chatbot* faz uso do classificador de textos para determinar a intenção do usuário com base em sua pergunta.

Os resultados demonstraram que, embora o *chatbot* tenha alcançado uma acurácia de 60,10% na classificação correta das intenções, ainda existem áreas de melhoria. A capacidade de generalização para perguntas fora do escopo inicialmente treinado e na manipulação de erros ortográficos são pontos a serem explorados em trabalhos futuros.

O desempenho geral do *chatbot* foi considerado satisfatório pela maior parte dos participantes, conforme indicado pela avaliação positiva de 68,18% dos participantes que concordaram ou concordaram totalmente com a afirmação de que as respostas do *chatbot* atendiam às suas perguntas. Este *feedback* positivo, somado à taxa de perda baixa no treinamento do modelo, sugere que o sistema é tecnicamente robusto, mas ainda precisa de ajustes para melhorar sua utilidade prática e relevância.

As contribuições deste estudo incluem a identificação da necessidade de ajustar a tolerância do *chatbot* a variações linguísticas e erros comuns de digitação, além da importância de expandir o escopo de conhecimento do sistema para abranger mais temas no domínio da saúde mental. Além disso, observou-se que a introdução de um sistema de memória de diálogo poderia potencialmente melhorar a interação dos usuários com o *chatbot*, permitindo uma experiência mais contextualizada e personalizada.

REFERÊNCIAS

- [1] ADAMOPOULOU, E.; MOUSSIADES, L. An overview of chatbot technology. In: SPRINGER. *IFIP international conference on artificial intelligence applications and innovations*. [S.l.], 2020. p. 373–383.
- [2] WHO. *Mental health*. [s.n.], 2022. Disponível em: <<https://www.who.int/en/news-room/fact-sheets/detail/mental-health-strengthening-our-response>>.
- [3] BONAZZA, D. S. S.; SCHUH, C. Avaliação dos níveis de ansiedade em profissionais da saúde: Plantonistas diurnos e noturnos. *Connection Line-Revista Eletrônica Do Univag*, n. 22, 2020. Disponível em: <<https://www.periodicos.univag.com.br/index.php/CONNECTIONLINE/article/view/1518/1646>>.
- [4] FERRARI, R.; FRANÇA, F. M. de; MAGALHÃES, J. Avaliação da síndrome de burnout em profissionais de saúde. *Revista Gestão & Saúde*, v. 3, n. 3, p. 868–883, 2012.
- [5] MONTEIRO, D. T.; MENDES, J. M. R.; BECK, C. L. C. Health professionals' mental health: a look at their suffering. *Trends in Psychology*, SciELO Brasil, v. 27, p. 993–1006, 2019.
- [6] ESPERIDIÃO, E.; SAIDEL, M. G. B.; RODRIGUES, J. *Saúde mental: foco nos profissionais de saúde*. [S.l.]: SciELO Brasil, 2020. e73supl01 p.
- [7] FERNANDES, M. A. et al. Transtornos mentais associados ao trabalho em profissionais de enfermagem: uma revisão integrativa brasileira. *Revista brasileira de medicina do trabalho*, Revista Brasileira de Medicina do Trabalho, v. 16, n. 2, p. 218–224, 2018.
- [8] NOVAES, H. M. D.; SOÁREZ, P. C. D. A avaliação das tecnologias em saúde: origem, desenvolvimento e desafios atuais. panorama internacional e brasil. *Cadernos de Saúde Pública*, SciELO Public Health, v. 36, p. e00006820, 2020.
- [9] MARENCO, L. L. et al. Mobile technologies in healthcare: reflections on development, application, legal aspects, and ethicstechnologías sanitarias móviles: reflexiones sobre desarrollo, aplicación, legislación y ética. *Revista Panamericana de Salud Publica= Pan American Journal of Public Health*, v. 46, p. e37–e37, 2022.
- [10] FERREIRA, D. da S. et al. Uso do chatbot no enfrentamento da ansiedade: Uma revisão integrativa. *Revista Enfermagem Atual In Derme*, v. 98, n. 2, p. e024297–e024297, 2024.
- [11] MITCHELL, T. Machine learning. *McGraw Hill*, 1997.
- [12] BZDOK, D.; KRZYWINSKI, M.; ALTMAN, N. Machine learning: a primer. *Nature methods*, NIH Public Access, v. 14, n. 12, p. 1119, 2017.
- [13] BI, Q. et al. What is machine learning? a primer for the epidemiologist. *American journal of epidemiology*, Oxford University Press, v. 188, n. 12, p. 2222–2239, 2019.

- [14] BREIMAN, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, Institute of Mathematical Statistics, v. 16, n. 3, p. 199–231, 2001.
- [15] BONACCORSO, G. *Machine learning algorithms*. [S.l.]: Packt Publishing Ltd, 2017.
- [16] NASTESKI, V. An overview of the supervised machine learning methods. *Horizons*, v. 4, p. 51–62, 2017.
- [17] KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E. Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review*, Springer, v. 26, p. 159–190, 2006.
- [18] SALMAN, R.; KECCMAN, V. Regression as classification. In: IEEE. *2012 Proceedings of IEEE Southeastcon*. [S.l.], 2012. p. 1–6.
- [19] RONG, S.; BAO-WEN, Z. The research of regression model in machine learning field. In: EDP SCIENCES. *MATEC Web of Conferences*. [S.l.], 2018. v. 176, p. 01033.
- [20] SOOFI, A. A.; AWAN, A. Classification techniques in machine learning: applications and issues. *Journal of Basic & Applied Sciences*, Set Publishers, v. 13, n. 1, p. 459–465, 2017.
- [21] BARTO, A. G.; THOMAS, P. S.; SUTTON, R. S. Some recent applications of reinforcement learning. In: *Proceedings of the Eighteenth Yale Workshop on Adaptive and Learning Systems*. [S.l.: s.n.], 2017.
- [22] RAUBER, T. W. Redes neurais artificiais. *Universidade Federal do Espírito Santo*, v. 29, 2005.
- [23] LAUDANI, A. et al. On training efficiency and computational costs of a feed forward neural network: A review. *Computational intelligence and neuroscience*, Hindawi Limited London, UK, United Kingdom, v. 2015, p. 83–83, 2015.
- [24] RIZZO, I. V.; CANATO, R. L. C. Inteligência artificial: funções de ativação. *Prospectus (ISSN: 2674-8576)*, v. 2, n. 2, 2020.
- [25] PARKHI, O.; VEDALDI, A.; ZISSERMAN, A. Deep face recognition. In: BRITISH MACHINE VISION ASSOCIATION. *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*. [S.l.], 2015.
- [26] HACOHEN-KERNER, Y.; MILLER, D.; YIGAL, Y. The influence of preprocessing on text classification using a bag-of-words representation. *PloS one*, Public Library of Science San Francisco, CA USA, v. 15, n. 5, p. e0232525, 2020.
- [27] JULURU, K. et al. Bag-of-words technique in natural language processing: a primer for radiologists. *Radiographics*, Radiological Society of North America, v. 41, n. 5, p. 1420–1426, 2021.
- [28] KOWSARI, K. et al. Text classification algorithms: A survey. *Information*, Multidisciplinary Digital Publishing Institute, v. 10, n. 4, p. 150, 2019.

- [29] QADER, W. A.; AMEEN, M. M.; AHMED, B. I. An overview of bag of words; importance, implementation, applications, and challenges. In: IEEE. *2019 international engineering conference (IEC)*. [S.l.], 2019. p. 200–204.
- [30] ONODY, R. N. Teste de turing e inteligência artificial. 2021.
- [31] XU, L. et al. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR cancer*, JMIR Publications Inc., Toronto, Canada, v. 7, n. 4, p. e27850, 2021.
- [32] CARVALHO, R. M. S. *Chatbot de informação sobre o COVID-19*. Dissertação (Mestrado), 2021.
- [33] CASHEEKAR, A. et al. A contemporary review on chatbots, ai-powered virtual conversational agents, chatgpt: Applications, open challenges and future research directions. *Computer Science Review*, Elsevier, v. 52, p. 100632, 2024.
- [34] OPENAI. *ChatGPT*. 2024. Disponível em: <<https://openai.com/chatgpt>>.
- [35] INC., A. *Siri*. 2024. Disponível em: <<https://www.apple.com/siri/>>.
- [36] AMAZON.COM, I. *Amazon Alexa*. 2024. Disponível em: <<https://www.amazon.com/alexa>>.
- [37] BARD.AI. *Bard*. 2024. Disponível em: <<https://www.bard.ai>>.
- [38] LLC, G. *Google Gen=Mini*. 2024. Disponível em: <<https://gemini.google.com/app>>.
- [39] SILVA, B. F. G. P. d. *Fatores determinantes na utilização dos chatbots no setor da aviação*. Tese (Doutorado) — Instituto Politécnico de Lisboa, Escola Superior de Comunicação Social, 2023.
- [40] RAMESH, K. et al. A survey of design techniques for conversational agents. In: SPRINGER. *International conference on information, communication and computing technology*. [S.l.], 2017. p. 336–350.
- [41] SEREICIKAS, R. T. et al. C2-p2: Uma abordagem baseada em chatbots para navegação de coleções de padrões. 2022.
- [42] CORREA, J. d. S. et al. Lhia: Um agente conversacional inteligente para educação em aleitamento materno e recrutamento de mães doadoras. Universidade Federal do Maranhão, 2023.
- [43] MCKINSEY. *What is generative AI?* 2023. McKinsey & Company. Consultado em 10 de março de 2023. Disponível em: <<https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>>.
- [44] HANNIGAN, T. R.; MCCARTHY, I. P.; SPICER, A. Beware of botshit: How to manage the epistemic risks of generative chatbots. *Business Horizons*, Elsevier, 2024.
- [45] MONTEIRO, G. S. Helena: um chatbot para auxílio dos discentes do decom em trâmites universitários. 2021.

- [46] MORALES, G. H. et al. Chatbot informativo sobre diabetes tipo 2. *Recursos digitales para apoyar a la docencia y el aprendizaje en diferentes niveles educativos*, p. 35.
- [47] GANTE, Á. G. Canto de et al. Escala de likert: Una alternativa para elaborar e interpretar un instrumento de percepción social. *Revista de la alta tecnología y sociedad*, v. 12, n. 1, 2020.

Apêndices

APÊNDICE A – INTENÇÕES DA BASE DE CONHECIMENTO

Identificador	Intenção
1	saudacoes
2	definicao-ansiedade
3	identificar-ansiedade
4	lidar-ansiedade
5	definicao-estresse
6	identificar-estresse
7	lidar-estresse-e-estrategias-estresse-e-ansiedade
8	definicao-luto
9	identificar-luto
10	lidar-luto
11	estrategias-luto
12	respiracao-diafragmatica
13	mindfulness
14	meditacao
15	relaxamento-muscular-progressivo
16	definicao-comunicacao-efetiva
17	comportamentos-para-comunicacao-efetiva
18	estrategias-comunicacao-efetiva
19	estrategias-cominucar-algo-importante
20	estrategias-manejar-conflitos
21	estrategias-cominucar-sentimentos
22	estrategias-ansiedade-estresse-comunicacao
23	interacao-pacientes
24	como-comunicar-pacientes
25	autoavaliacao-comunicacao
26	comunicacao-empatica
27	relacoes-trabalho
28	estrategias-relacoes-trabalho
29	familia
30	dia-a-dia-familia
31	relacionamento-parceiro
32	cuidado-filhos
33	relacionamento-idosos
34	definicao-autocuidado
35	praticas-autocuidado
36	autores-projeto
37	bibliografia-projeto
38	agradecimento