

# AVICO MNF: UM SOFTWARE SENTINELA PARA O MONITORAMENTO DE NOTÍCIAS FALSAS SOBRE VACINAS DA COVID

Daniel Oliveira de Freitas \*  
Fábio Paulo Basso \*\*  
Gabriel Machado Lunardi \*\*\*

## RESUMO

Durante a pandemia de COVID-19, a proliferação de *fake news* nas mídias sociais e outras plataformas causaram sérios impactos na saúde física e mental das pessoas. Essa desinformação gera desconfiança nas autoridades de saúde e orientações médicas, levando a comportamentos arriscados. Com milhões de casos e centenas de milhares de mortes confirmadas, combater as fake news é crucial. Pesquisadores utilizam técnicas avançadas como mineração de dados, *web scraping*, *web crawling*, aprendizado de máquina e aprendizado profundo para detectar essas informações falsas. Apesar dos avanços, desafios como a necessidade de grandes volumes de dados e a rápida evolução das táticas de desinformação persistem, exigindo abordagens multifacetadas para lidar com o problema. Este trabalho apresenta uma proposta de solução: um software sentinela que permite o monitoramento de notícias falsas acerca das vacinas da COVID-19.

**Palavras-chaves:** data mining, text mining, web mining, fake news.

## ABSTRACT

During the COVID-19 pandemic, the proliferation of fake news on social media and other platforms has had a serious impact on people's physical and mental health. This misinformation has led to distrust of health authorities and medical advice, leading to risky behavior. With millions of confirmed cases and hundreds of thousands of deaths, combating fake news is crucial. Researchers are using advanced techniques such as data mining, web scraping, web crawling, machine learning, and deep learning to detect this false information. Despite these advances, challenges such as the need for large volumes of data and the rapid evolution of disinformation tactics persist, requiring multifaceted approaches to address the problem. This paper presents a proposed solution: a sentinel software that allows monitoring of false news about COVID-19 vaccines.

**Keywords:** data mining, text mining, web mining, fake news.

\* Aluno do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil  
E-mail: [danielodf2.aluno@unipampa.edu.br](mailto:danielodf2.aluno@unipampa.edu.br)

\*\* Orientador, Professor do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil, E-mail: [fabiobasso@unipampa.edu.br](mailto:fabiobasso@unipampa.edu.br)

\*\*\* Co-Orientador, Professor do Departamento de Eletrônica e Computação da Universidade Federal de Santa Maria, Santa Maria, Rio Grande do Sul, Brasil, E-mail: [gabriel.lunardi@uufsm.br](mailto:gabriel.lunardi@uufsm.br)

## 1. INTRODUÇÃO

O crescimento dos dados gerados e armazenados ocorreu de forma exponencial na última década. Além disso, a triste ocorrência de uma pandemia de *Covid-19*, a utilização de aparelhos eletrônicos se tornou um meio principal para a realização de atividades. Atualmente, a Web é indiscutivelmente a maior fonte de informações eletrônicas disponível. No entanto, encontrar informações relevantes pode muitas vezes ser uma experiência frustrante devido à sua natureza dinâmica. Para solucionar esse problema, técnicas de mineração (IGNACZAK et al., 2021) de dados capazes de analisar as informações disponíveis na Web. Por exemplo, a utilização de técnicas de processamento de linguagem natural e mineração de textos tem se mostrado promissora na detecção de desinformações e *fake news* (MISINFORMATION, 2022). Além disso, diferentes aplicações, como a análise de opinião nas mídias sociais, a detecção de conteúdos indesejados e a classificação automática de notícias falsas, têm sido exploradas por meio do uso de inteligência artificial, aprendizado de máquina e *big data* (FERNANDES; MORO; CORTEZ, 2023). A partir desses esforços, é possível avançar na tarefa de buscar informações relevantes na Web enquanto se aumenta a confiabilidade e veracidade das informações que são disponibilizadas (DIAZ-GARCIA; RUIZ; MARTIN-BAUTISTA, 2023).

O termo *fake news* é utilizado para descrever informações falsas ou enganosas que são conscientemente disseminadas, com o objetivo de influenciar a opinião pública e iludir as pessoas (PROCESSING, 2021). A propagação de notícias falsas pode trazer consequências graves, tais como a desinformação, o aumento da polarização política e até mesmo a instabilidade social. Por isso, é importante identificar e refutar informações falsas, visando garantir a integridade das informações e proteger a sociedade (ASUDANI; NAGWANI; SINGH, 2023). A detecção de *fake news* requer o uso de técnicas e ferramentas capazes de identificar informações falsas ou enganosas e distingui-las de informações verdadeiras. Isso pode envolver a verificação das fontes, a análise de evidências e a comparação com informações confiáveis (MINING, 2022). Além disso, a inteligência artificial e o aprendizado de máquina estão sendo aplicados para auxiliar na detecção de *fake news*, por meio da identificação de padrões e tendências nas informações enganosas (RAZALI et al., 2021). No entanto, é importante lembrar que a tecnologia não é uma solução definitiva e que a responsabilidade de diferenciar o verdadeiro do falso ainda cabe ao público em geral. É importante ainda desenvolver habilidades críticas e a capacidade de avaliar a informação de forma independente (FERNANDES; MORO; CORTEZ, 2023).

A Mineração de Dados é um conjunto de técnicas e processos que são voltados para a busca de padrões, relações e exploração de grandes quantidades de informações em banco de dados ou em páginas web amplamente utilizada em diversas áreas, análise de mídias sociais, como marketing, monitoramento de reputação empresarial, análise financeira, entre outras (IGNACZAK et al., 2021). Tornando-se uma ferramenta indispensável para extrair informações valiosas de grandes volumes de dados não estruturados, permitindo que empresas e organizações possam tirar o máximo proveito da informação disponível para tomar decisões mais informadas e eficientes (DIAZ-GARCIA; RUIZ; MARTIN-BAUTISTA, 2023). Ocasionalmente, com o crescimento de informações geradas diariamente na web, a produção dessa grande quantidade de informação causa um problema de organização. Tal problema pode ser solucionado utilizando uma técnica de mineração de dados, que visa estruturar e organizar essas

informações (MINING, 2022). Um dos principais objetivos de quem utiliza uma técnica de mineração de dados é o de analisar e explorar os dados em busca de tendências, relações entre as informações e com isso prever comportamentos futuros, onde seja possível classificar notícias em verdadeiras ou em *fake news* (ALAMOODI et al., 2021).

A *machine learning* e o *deep learning* são subcampos da Inteligência Artificial que se concentram em desenvolver sistemas de computação capazes de aprender por conta própria (SYNTAX, 2021). Isso é alcançado através de algoritmos que são treinados com grandes quantidades de dados, a fim de identificar padrões e tomar decisões com base em novos inputs. A *deep learning*, em particular, é uma abordagem avançada de *machine learning* que usa camadas profundas de redes neurais para modelar complexos relacionamentos entre os dados de treinamento. O objetivo é imitar a capacidade humana de aprendizado e desenvolvimento de conhecimento, tornando-o um dos caminhos mais promissores para a realização da inteligência artificial (ANALYSIS, 2022).

### 1.1. JUSTIFICATIVA

A Associação de Vítimas de Covid-19 (AVICO) é uma organização comunitária com mobilização social que visa minimizar os diferentes impactos da pandemia no Brasil, lutando pela memória das vítimas da Covid-19, assim como a garantia e acesso aos Direitos Humanos. Atualmente, a AVICO possui dificuldades para a utilização de várias funcionalidades automatizadas em o seu sistema.

Dentro desse contexto, a AVICO necessita da aplicação de técnicas e métodos para a identificação automática das *fake news*, tendo em vista a quantidade de informações disponíveis se torna demorada e dificultosa uma classificação manual. Nesse sentido, uma das primeiras etapas para a construção de uma abordagem para a detecção automática de notícias falsas é a aquisição de dados de texto para o treinamento dos algoritmos.

Portanto, a justificativa para o desenvolvimento do presente Trabalho de Conclusão de Curso (TCC) é justamente essa aquisição de dados, isto é, utilizar técnicas de *web mining*. Além disso, a implementação de uma ferramenta de monitoramento para detecção de *fake news* relacionadas à AVICO.

### 1.2. OBJETIVO GERAL E ESPECÍFICO

Este trabalho tem por objetivo extrair conteúdo textual da *web* para o treinamento de algoritmos baseados em classificação de texto para a detecção de conteúdo falso relacionado à AVICO. Para atingir tal objetivo, os seguintes objetivos específicos podem ser elencados:

- Estudar sobre técnicas de *web mining*.
- Estudar sobre técnicas de limpeza e transformação de dados textuais.
- Implementar o suporte ferramental para a classificação de notícias;
- Implantar e testar a aplicação de monitoramento na plataforma da AVICO;

### 1.3. METODOLOGIA

Nesta seção, é apresentada a metodologia pretendida para a execução do Trabalho de Conclusão de Curso. Para tanto, foi desenvolvida a metodologia que apresenta a Figura 1, de forma esquemática, um diagrama BPMN com as atividades do trabalho. No diagrama, destacam-se os seguintes processos: Concepção, Fundamentação Teórica, TCC 1, e TCC 2. A seguir, é apresentada uma descrição detalhada do que foi realizado em cada etapa:

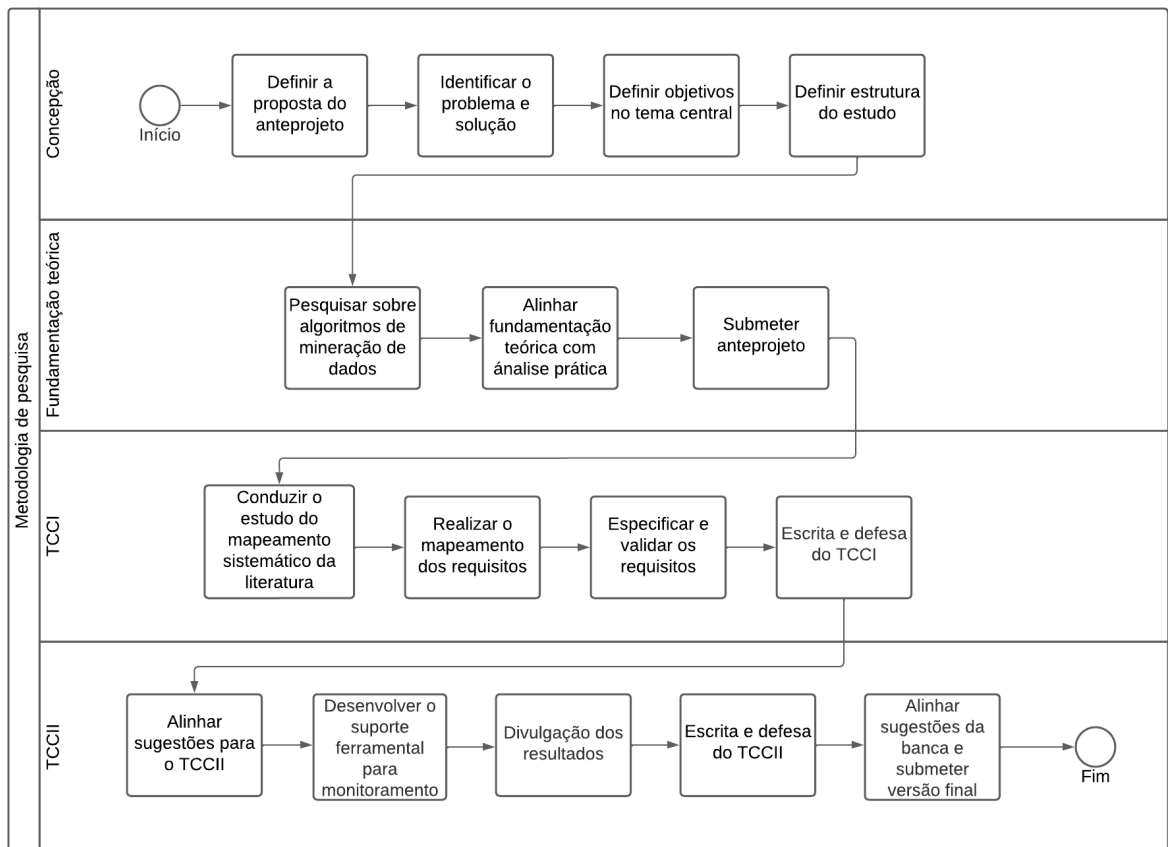


Figura 1 - Desenho da Metodologia de pesquisa. Fonte: Autor

- **Concepção:** É nessa etapa inicial, teve como objetivo principal a definição de qual será o tema do projeto, sua estruturação, assim como uma leitura inicial na parte ferramental relacionada ao tema;
- **Fundamentação teórica:** É a fase em que serão realizadas pesquisas exploratórias em algoritmos de mineração e análise de dados, assim como o sistema que a AVICO fará uso, contribuindo para o conhecimento do que será necessário desenvolver o algoritmo que automatizará a pesquisa por fake news, melhorando em relação à situação que a AVICO possui no momento que é a análise manual dos dados;
- **TCC I:** Durante essa fase, foi realizada a leitura dos artigos relacionados ao tema abordado no trabalho de conclusão de curso, o qual foi posteriormente redigido e

apresentado na defesa. Devido ao limite de espaço do presente artigo, tal estudo não foi incluído neste documento.

- **TCC II:** Por último, na fase do TCC II, foram realizados ajustes na redação. Foi desenvolvido o suporte ferramental, sendo possível acessá-lo por meio de requisições via API, que inclui dois algoritmos de mineração de dados com o objetivo de minerar os links e textos das notícias. Além disso, ocorreu a escrita e a defesa do TCC II.

#### **1.4. CONTRIBUIÇÕES**

As principais contribuições do TCC apresentado estão ligadas à implementação de dois algoritmos de mineração (raspagem) assim como métodos e técnicas para a filtragem de dados. Em consequência disso, a integração com um suporte ferramental visando automatizar e tornar mais eficiente o processo de seleção de *fake news* direcionados a associação AVICO. Portanto, as contribuições para o Trabalho de Conclusão de Curso são:

- Criação de um suporte ferramental integrado para monitoramento de notícias falsas a *fake news*.

#### **1.5. ORGANIZAÇÃO**

O artigo está estruturado da seguinte forma: A Seção 2 apresenta a fundamentação teórica e a motivação do trabalho. Na Seção 3, é apresentada a proposta de desenvolvimento, implementação e demonstração conceitual. A seção 4 detalha a avaliação do software proposto. A seção 5 apresenta a conclusão do trabalho, discutindo as considerações finais e possíveis direções futuras de pesquisa.

## **2. FUNDAMENTAÇÃO TEÓRICA**

A disseminação de notícias falsas pode causar impactos negativos na sociedade e na política. A teoria por trás da detecção de *fake news* se baseia em métodos de processamento de linguagem natural, mineração de dados e aprendizado de máquina. Esses métodos procuram identificar diferenças entre notícias verdadeiras e falsas, utilizando técnicas de análise para alcançar resultados precisos.

A mineração de dados é uma técnica que desempenha um papel fundamental na detecção de notícias falsas. Com ela, é possível coletar e analisar grandes volumes de informações, identificando padrões e características únicas que podem ser usadas para distinguir notícias que tenham conteúdo enganoso. Com essas informações, é possível desenvolver técnicas mais eficazes para combater a propagação de notícias falsas e garantir que a sociedade receba informações precisas e confiáveis.

## Extração dos Dados

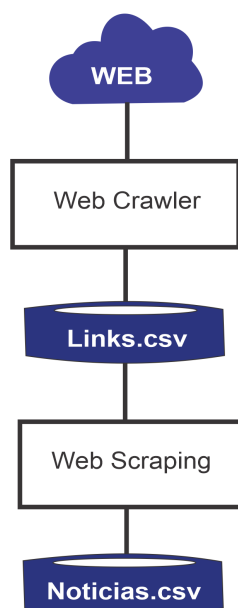


Figura 2 - Fluxo do processo de extração dos dados. Fonte: Autor

A Figura 2 ilustra o fluxo do processo de extração de dados utilizado para identificar e analisar notícias falsas. Esse processo é composto por várias etapas essenciais, conforme detalhado a seguir:

- **Web Crawler:** Esta etapa envolve a utilização de um *web crawler*, um tipo de robô da internet que navega na *web* de forma automatizada, coletando dados de diversas páginas. O *crawler* visita uma página web, baixa seu conteúdo e extrai os hiperlinks contidos nela, seguindo esses links para continuar o processo de coleta de dados. O resultado desse processo é um arquivo *Links.csv* contendo os URLs coletados, que servirão como ponto de partida para a próxima etapa.
- **Web Scraping:** Com os URLs coletados pelo *web crawler*, a próxima etapa é o *web scraping*, que envolve a extração específica de informações relevantes das páginas web. Utilizando algoritmos de *scraping*, os dados são baixados automaticamente de uma página web e informações específicas são extraídas. Essas informações podem ser armazenadas em diversos formatos, como bancos de dados ou arquivos. No contexto deste projeto, os dados extraídos são armazenados em um arquivo *Noticias.csv*.

O *web scraping* e o *web crawling* são técnicas complementares, onde o *web crawling* é responsável por coletar informações genéricas, enquanto o *web scraping* foca em dados específicos. Essas técnicas são fundamentais para o processo de mineração de dados, permitindo a coleta e análise de grandes volumes de informações da web.

Uma etapa que antecedeu ao desenvolvimento de uma solução para monitoramento de fake news para a AVICO Brasil foi um estudo de revisão de literatura, apresentada como contribuição do TCC 1. A motivação para realizar uma revisão sobre mineração de notícias falsas é o aumento significativo da disseminação de informações falsas e enganosas na era digital. Com a facilidade de acesso à informação e

a ampla disponibilidade de plataformas de mídia social, é cada vez mais comum que notícias enganosas sejam criadas e compartilhadas em larga escala. Com base nisso, a condução da revisão teve como objetivo pesquisar técnicas de mineração utilizadas pelos estudos a fim de caracterizar aquelas que apresentam maior potencial.

### 3. PROPOSTA E DESENVOLVIMENTO

A proposta de implantação de um sistema de monitoramento para detecção de notícias falsas no sistema AVICO envolveu a implementação de um conjunto de algoritmos e técnicas que permitam identificar e filtrar possíveis notícias falsas ou enganosas. Além disso, foi importante considerar aspectos éticos, como a transparência no processo de detecção, a privacidade dos usuários e a divulgação adequada dos resultados obtidos pelo sistema. Vale ressaltar que a detecção de *fake news* é um desafio em constante evolução, e nenhum sistema é totalmente infalível. Portanto, é essencial fornecer aos usuários orientações claras sobre a importância de verificar fontes e ser crítico em relação às informações encontradas na web.

Com isso, foram criadas duas *APIs*: uma para executar o algoritmo *scrapy*, responsável por realizar a raspagem dos links de notícias contidos na página, e outra para executar o algoritmo *crawler*, responsável por extrair o título usando a tag `<h1>` e os textos usando a tag `<p>` de cada link coletado pelo *scrapy*. Essas *APIs* foram integradas em duas ferramentas: um *chatbot*, acessível a qualquer visitante do site da AVICO, e um sistema de monitoramento destinado à parte administrativa, ativada com o cadastro de novas pesquisas com o intervalo de tempo determinados pelo usuário e enviando uma resposta por *e-mail* em uma tabela contendo o link da notícia e sua classificação.

Esse sistema proporciona suporte aos verificadores, buscando otimizar o tempo de resposta, evitar sobrecargas no sistema e garantir uma navegação fluida. É importante ressaltar que o desenvolvimento dessas ferramentas não substitui completamente a necessidade de avaliação humana, dado que a detecção de *fake news* é um desafio em constante evolução, e nenhum sistema é infalível. O desenvolvimento desse sistema foi dividido em etapas, conforme descrito na seção abaixo:

#### 3.1. REQUISITOS

- O sistema de monitoramento deve ser capaz de coletar dados de diferentes fontes, como notícias de sites, redes sociais ou outras fontes de informação relevantes para a detecção de *fake news*.
- Os dados coletados devem passar por etapas de pré-processamento, incluindo limpeza, *tokenização*, remoção de *stopwords* e normalização, para prepará-los para a análise.
- O sistema de monitoramento deve ser capaz de extrair características relevantes dos textos das notícias, como palavras-chave, padrões linguísticos, entre outros, para ajudar na detecção.
- A interface deve incluir funcionalidades para inserir notícias para análise e visualizar os resultados de detecção de *fake news* de forma clara e compreensível.
- O usuário deve acessar o suporte ferramental por uma interface intuitiva que permita interagir com o sistema facilmente.

- O suporte ferramental deve classificar automaticamente as notícias como verdadeiras ou falsas com base nas características extraídas.

### 3.2. IMPLEMENTAÇÃO E DEMONSTRAÇÃO CONCEITUAL

As APIs de monitoramento de notícias falsas estão integradas à página web com a ferramenta de varredura, para analisar e avaliar a veracidade das notícias, garantindo uma abordagem abrangente na verificação de *fake news* em diferentes contextos de interação online.

Uma interface de usuário intuitiva e amigável é fundamental para permitir que os usuários interagem facilmente com o sistema. As interfaces foram desenvolvidas com o intuito de possibilitar a visualização dos resultados da detecção de *fake news* de forma clara e oferecer indicações de confiabilidade para as notícias apresentadas.

A tela de listar é a tela principal da ferramenta de monitoramento automática para detecção de *fake news* como mostra a Figura 3. Nesta tela, os usuários têm a capacidade de gerenciar de forma fácil e simples as funcionalidades da ferramenta, incluindo o cadastro, a visualização de uma lista das pesquisas já cadastradas, edição e exclusão de itens cadastrados. Ela proporciona uma visão geral de todos os dados relevantes para o processo de monitoramento, permitindo uma gestão eficiente e organizada.

Para acessar a tela de listar, os usuários podem seguir os seguintes passos:

- A partir da página inicial da aplicação, navegue até a barra de navegação localizada na parte superior da tela.
- Na barra de navegação, localize a opção "Administração" e clique sobre ela. É importante ter acesso como administrador pois se trata de uma ferramenta de uso administrativo.
- No menu suspenso que será exibido, procure pela opção "Detecção de *fake news*" e selecione-a.

O usuário será redirecionado para a tela principal da ferramenta de monitoramento.

Na tela de cadastro mostrado na Figura 4, os usuários podem inserir novas informações para iniciar o processo de monitoramento de *fake news*. Nesta tela, há três campos principais a serem preenchidos:

- O primeiro campo destina-se à inserção do link da página de notícias ou da própria notícia específica que deseja ser verificada. Os usuários podem fornecer o URL diretamente neste campo para que a ferramenta possa acessar e analisar o conteúdo.
- O segundo campo permite que os usuários selecionem a frequência com que desejam que o monitoramento seja executado. Há três opções de escolha entre os períodos diário, semanal e mensal, dependendo da necessidade e da frequência desejada para o monitoramento das notícias.



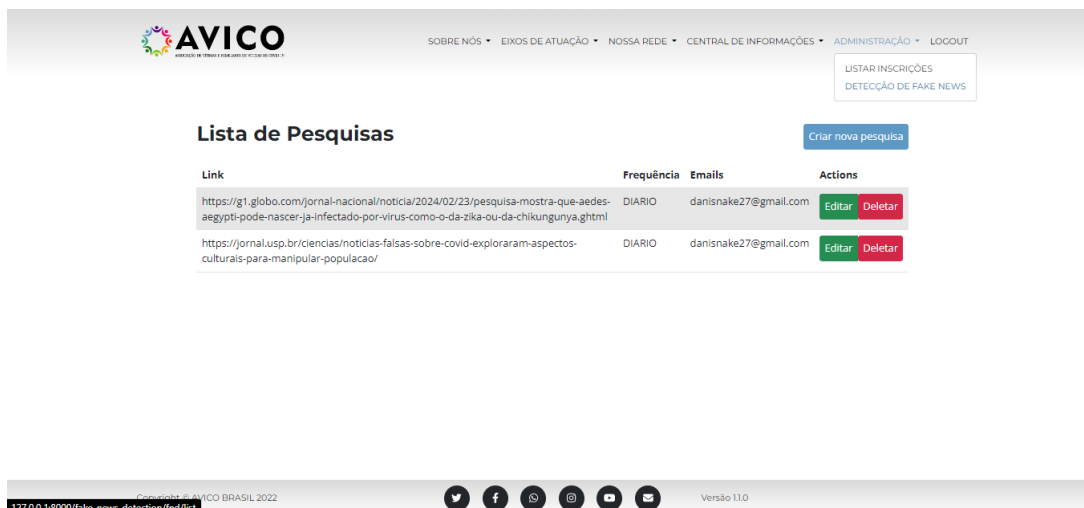


Figura 3 - Tela de listar da ferramenta de varredura.

- O terceiro campo é destinado à inserção do *e-mail* do usuário, onde ele deseja receber os resultados das classificações das notícias. Uma vez que o monitoramento é concluído, os resultados são enviados para o endereço de *e-mail*, fornecendo aos usuários uma maneira conveniente de acessar as informações sobre a veracidade das notícias verificadas.

Essa tela de cadastro fornece uma interface intuitiva e fácil de usar, permitindo que os usuários configurem facilmente as preferências de monitoramento e recebam os resultados de classificação das notícias de acordo com suas necessidades e especificações.

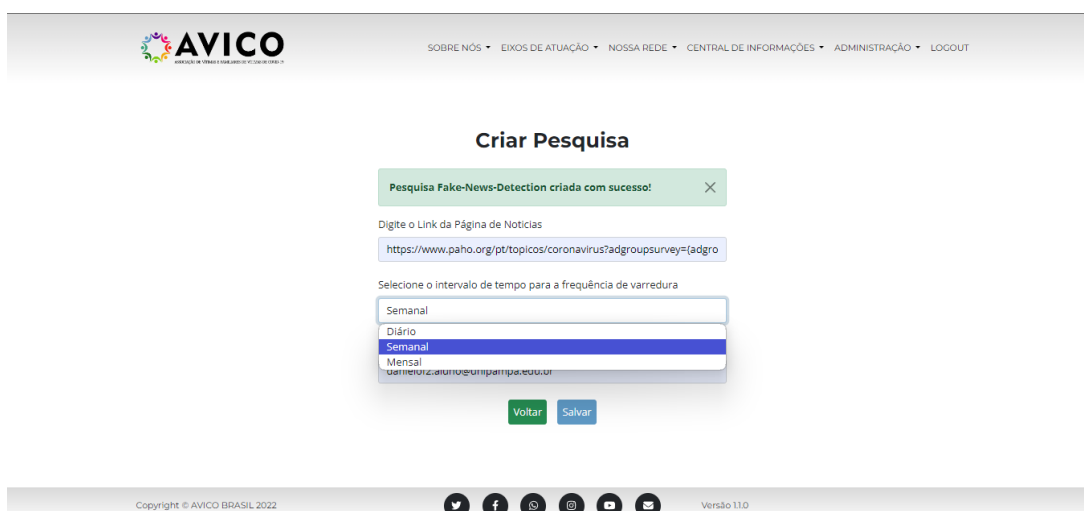


Figura 4 - Tela de cadastro da ferramenta de varredura.

Na tela de editar mostrado na Figura 5, os usuários podem fazer ajustes e modificações nas informações previamente cadastradas. Assim como na tela de cadastro, esta também possui os mesmos campos principais: como o link da página de notícias, a frequência de monitoramento e o endereço de *e-mail*.

Após salvar as ações de cadastro ou edição, o sistema exibirá uma mensagem de sucesso informando ao usuário que a operação foi realizada com êxito. Es-

As mensagens geralmente incluem uma confirmação clara, como "Cadastrado com sucesso" ou "Editado com sucesso", para garantir que os usuários saibam que suas ações foram concluídas com sucesso. Essas mensagens fornecem *feedback* imediato aos usuários, confirmando que suas alterações foram salvas corretamente e ajudando a garantir uma experiência de usuário satisfatória.

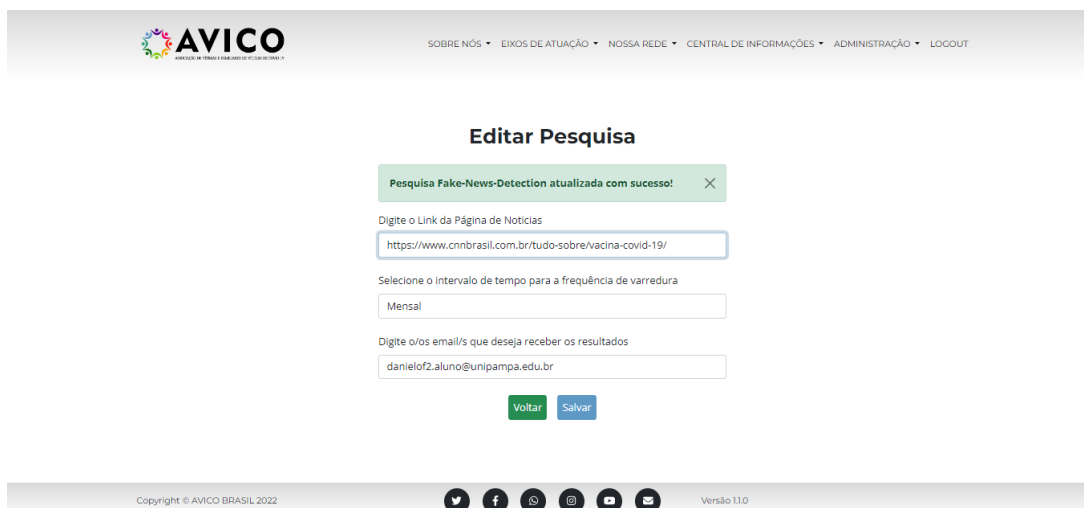


Figura 5 - Tela de editar da ferramenta de varredura.

Na tela de listar, há opção de deletar cada monitoramento cadastrado individualmente. Isso pode ser feito selecionando a opção de "deletar" localizada na coluna de ações ao lado de cada entrada. Após clicar em "deletar", o sistema exibe uma mensagem de confirmação para garantir que o usuário realmente deseja prosseguir com a exclusão como mostra a Figura 6. Após a confirmação, o sistema deleta o cadastro e exibe uma mensagem de sucesso, informando ao usuário que o cadastro de monitoramento foi deletado com êxito mostrado na Figura 7. Além disso, a lista é atualizada automaticamente para refletir as alterações feitas. Essa funcionalidade oferece aos usuários um meio conveniente e eficiente de gerenciar seus monitoramentos cadastrados na ferramenta.

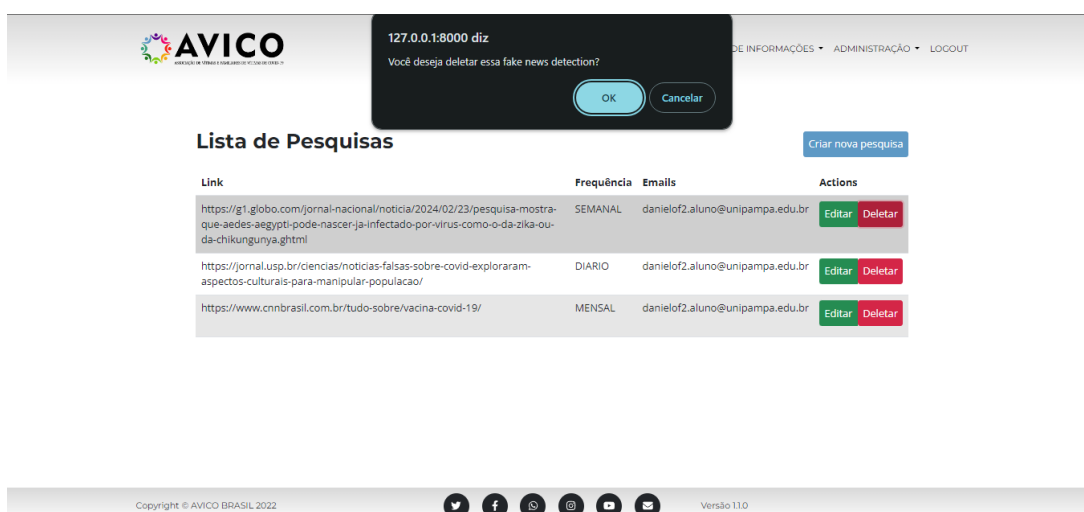


Figura 6 - Aviso de confirmação para exclusão.

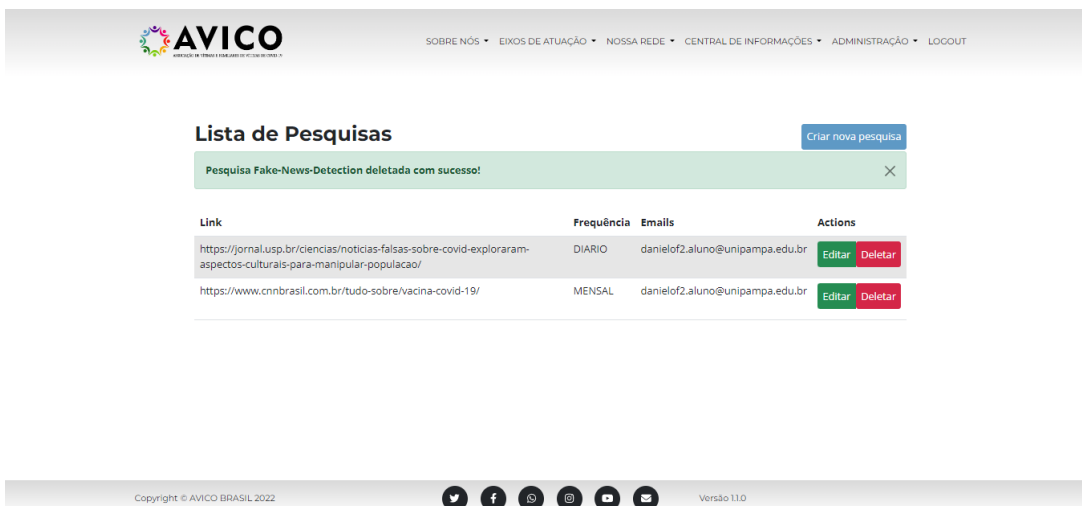


Figura 7 - Mensagem de excluído com sucesso.

- **Atualização e Monitoramento:** A ferramenta é continuamente atualizada com novos dados para manter sua eficácia constante. Após cada monitoramento, os resultados são compilados e enviados automaticamente por *e-mail* para o endereço cadastrado. Esses *e-mails* incluem detalhes abrangentes, como os links das notícias analisadas e suas respectivas classificações mostrados na Figura 8, oferecendo aos usuários uma maneira fácil e conveniente de acessar as informações verificadas. Essa função de monitoramento automatizado garante que os usuários estejam sempre atualizados e bem informados sobre as últimas verificações.

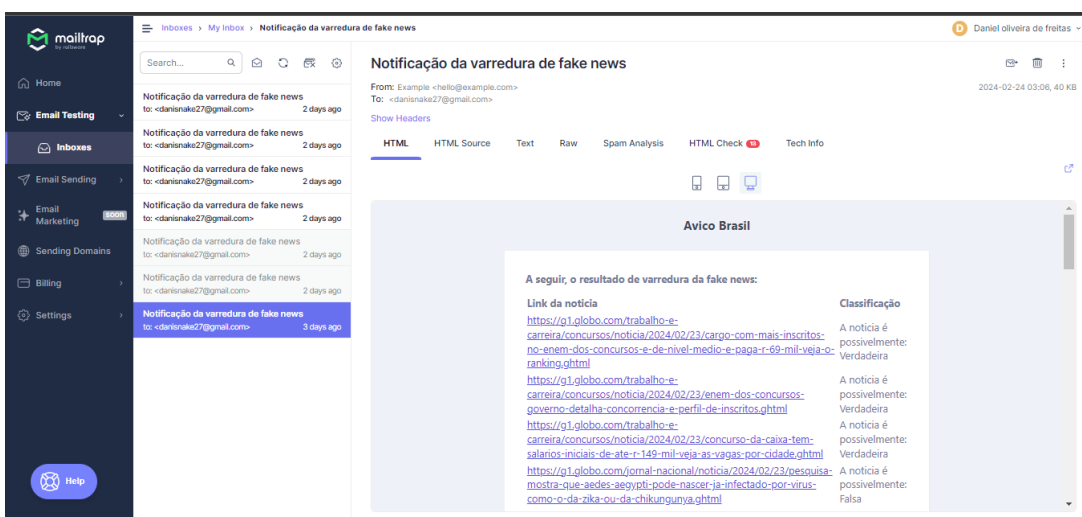


Figura 8 - Tela de *emails* com os resultados do monitoramento.

## 4. AVALIAÇÃO

### 4.1. PROTOCOLO DE AVALIAÇÃO

Este estudo foi elaborado com o propósito de avaliar a eficácia, precisão e facilidade de uso da ferramenta de monitoramento de notícias falsas no que tange

à determinação da veracidade das notícias. O objetivo primordial é oferecer aos usuários um meio confiável para verificar a credibilidade das informações apresentadas. Inicialmente, os participantes receberam o convite contendo o *link* de acesso ao formulário. Essa abordagem facilitou a participação dos respondentes, garantindo que possam contribuir com suas opiniões e experiências de forma conveniente e direta.

Na página inicial do *survey*, contem uma breve descrição sobre o propósito da pesquisa e a natureza da ferramenta de monitoramento de notícias. Nessa introdução, destacamos a importância de avaliar a eficácia, precisão e usabilidade da ferramenta na determinação da veracidade das informações. Além disso, fornecemos uma visão geral da ferramenta, enfatizando sua capacidade de fornecer aos usuários uma maneira confiável de verificar a credibilidade das notícias em um cenário onde a disseminação rápida e vasta de informações imprecisas é cada vez mais prevalente. Essa descrição inicial foi projetada para contextualizar os participantes e incentivá-los a se envolverem ativamente na avaliação, entendendo a importância de seu *feedback* para o aprimoramento contínuo da ferramenta.

Na segunda página do *survey*, é apresentada a Declaração de Consentimento Livre e Esclarecido. Nesta declaração, é explicado o que a participação na avaliação envolverá, incluindo responder questionários ou entrevistas para fornecer *feedback* sobre a ferramenta em questão. Além disso, os participantes são informados de que seus dados serão utilizados exclusivamente para os propósitos da avaliação e que suas identidades serão protegidas de acordo com as leis de proteção de dados aplicáveis.

A declaração enfatizará a natureza voluntária da participação, garantindo aos participantes o direito de retirar seu consentimento a qualquer momento, sem qualquer penalidade ou consequência. Os participantes foram convidados a confirmar que leram e compreenderam os termos do consentimento, concordando em participar da avaliação da ferramenta de monitoramento notícias falsas.

Ao marcar a opção indicada, os participantes autorizarão de forma livre e informada sua participação no estudo, demonstrando seu consentimento para prosseguir com a avaliação. Essa etapa foi essencial para garantir que a participação dos indivíduos seja voluntária e baseada em uma compreensão clara dos termos e objetivos da pesquisa.

Na terceira página do *survey*, os participantes foram direcionados ao Questionário de Perfil do Usuário, que teve como objetivo coletar informações relevantes sobre os usuários. Este questionário foi desenvolvido para obter uma compreensão mais profunda do perfil demográfico e comportamental dos participantes, o que ajudou a contextualizar e interpretar os dados coletados durante a avaliação da ferramenta de monitoramento de notícias falsas.

Na penúltima parte do *survey*, os participantes foram apresentados ao vídeo tutorial para utilização da ferramenta de monitoramento do sistema AVICO. Esse vídeo tutorial foi desenvolvido com o objetivo de fornecer aos participantes uma demonstração prática e passo a passo de como utilizar a ferramenta para monitoramento de notícias falsas.

O vídeo tutorial aborda os principais aspectos da utilização da ferramenta, incluindo como acessá-la, inserir notícias para verificação, interpretar os resultados

e utilizar os recursos disponíveis para analisar a credibilidade das informações. Além disso, o tutorial inclui dicas e melhores práticas para maximizar a eficácia da ferramenta.

A inclusão do vídeo tutorial tem como objetivo facilitar a compreensão e o uso da ferramenta pelos participantes, fornecendo uma experiência prática e visualmente explicativa. Isso poderá aumentar a eficácia do *survey*, garantindo que os participantes estejam bem preparados para interagir com a ferramenta e fornecer *feedback* relevante e informado sobre sua usabilidade e funcionalidade.

Portanto, o vídeo tutorial representa uma parte crucial do *survey*, capacitando os participantes a utilizar plenamente a ferramenta de monitoramento do sistema AVICO e contribuir de forma significativa para a avaliação de sua eficácia na determinação da veracidade de notícias.

A avaliação da ferramenta de monitoramento é uma parte crucial do processo, pois permite aos participantes fornecerem *feedback* detalhado sobre suas experiências e percepções da ferramenta. Utilizando uma escala de avaliação *likert* de 1 a 5, os participantes são convidados a expressar seu nível de concordância com uma série de afirmações relacionadas ao desempenho e funcionalidades da ferramenta.

As perguntas abordam aspectos essenciais, como a clareza do *feedback* fornecido pela ferramenta, a correspondência dos resultados com a veracidade das notícias, a flexibilidade dos usuários para ajustar a varredura de notícias, a consistência da interface, medidas para evitar erros, clareza das opções disponíveis, opções para usuários avançados ou personalização de configurações, organização da interface, disponibilidade de ajuda e documentação.

Além das respostas às perguntas, os participantes também são incentivados a compartilhar quaisquer comentários adicionais que possam ter sobre a ferramenta. Esses comentários adicionais podem incluir sugestões de melhoria, pontos fortes da ferramenta ou qualquer outra observação relevante que possa ajudar na compreensão global da eficácia e usabilidade.

Essa avaliação foi essencial para identificar áreas de melhoria e aprimoramento da ferramenta, garantindo que ela atenda adequadamente às necessidades e expectativas dos usuários. O *feedback* coletado nesta etapa pode orientar futuras atualizações e iterações da ferramenta, contribuindo para sua contínua evolução e eficácia na detecção de *fake news* e na promoção da verificação de informações confiáveis.

#### **4.1.1. RESULTADOS**

Enquanto existem pontos fortes como a consistência da interface e a clareza das opções, há várias áreas críticas que necessitam de aprimoramento urgente, especialmente em termos de precisão dos resultados e suporte ao usuário. A ferramenta tem um bom funcionamento, mas a confiabilidade na precisão dos resultados e o suporte para personalização e resolução de problemas são cruciais para sua eficácia e aceitação pelo usuário. Implementações futuras devem focar na melhoria da precisão e na expansão das funcionalidades da ferramenta.

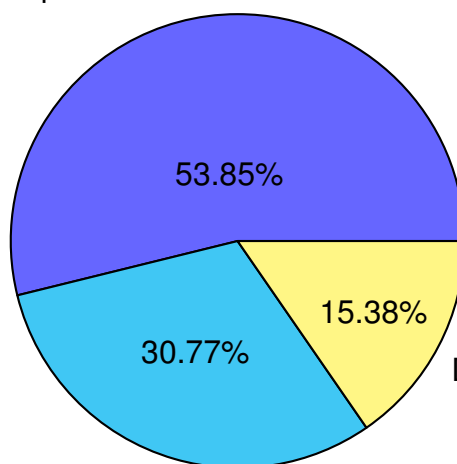
Os resultados da avaliação da ferramenta monitoramento de notícias falsas desenvolvida para a AVICO revelam variados níveis de satisfação e eficácia perce-

bida entre os usuários. A análise detalhada de cada aspecto avaliado oferece *insights* sobre as áreas de força e as que necessitam de melhorias, como segue.

**QUESTÃO 1 - A FERRAMENTA FORNECE *feedback* CLARO SOBRE O ESTADO DA VARREDURA OU ANÁLISE DE NOTÍCIAS EM TEMPO REAL?**

A maioria dos usuários 53.85% afirma que a ferramenta fornece um *feedback* claro durante o monitoramento ou análise de notícias em tempo real. No entanto, 30.77% apenas concordam parcialmente, indicando que a clareza pode ser inconsistente.

Concordo completamente



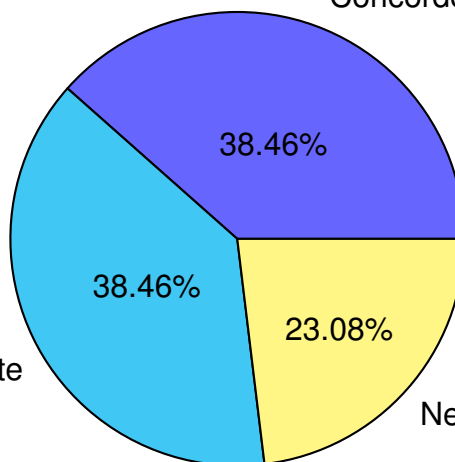
Discordo parcialmente

Concordo parcialmente

**QUESTÃO 2 - OS RESULTADOS DA FERRAMENTA CORRESPONDEM ADEQUADAMENTE À VERACIDADE DAS NOTÍCIAS ESCANEADAS?**

Apenas 38.46% dos usuários concordam completamente que os resultados da ferramenta correspondem à veracidade das notícias escaneadas, com outro 38.46% concordando parcialmente e 23.08% permanecendo neutros, sugerindo uma necessidade de melhoria na precisão da ferramenta.

Concordo completamente

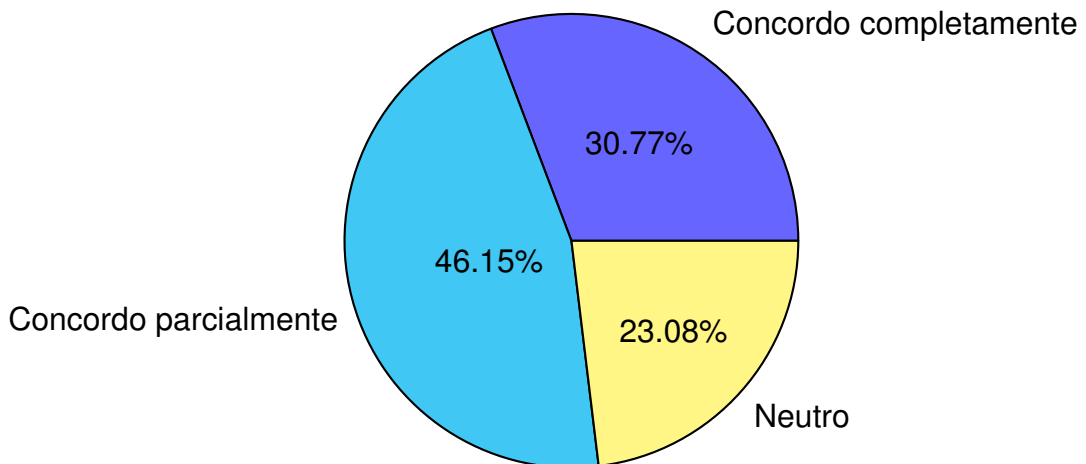


Concordo parcialmente

Neutro

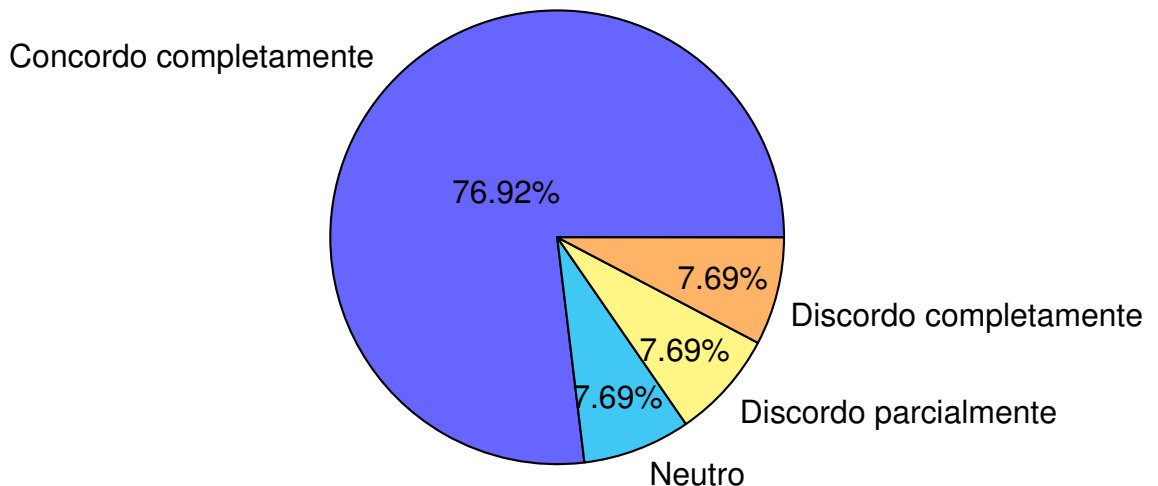
**QUESTÃO 3 - OS USUÁRIOS TÊM A LIBERDADE DE INTERROMPER OU AJUSTAR O MONITORAMENTO DE NOTÍCIAS CONFORME DESEJADO?**

A liberdade para manipular o monitoramento é moderadamente recebida, com 30.77% concordando completamente e 46.15% apenas parcialmente, mostrando que, enquanto a flexibilidade existe, ela pode não ser totalmente efetiva ou conhecida pelos usuários.



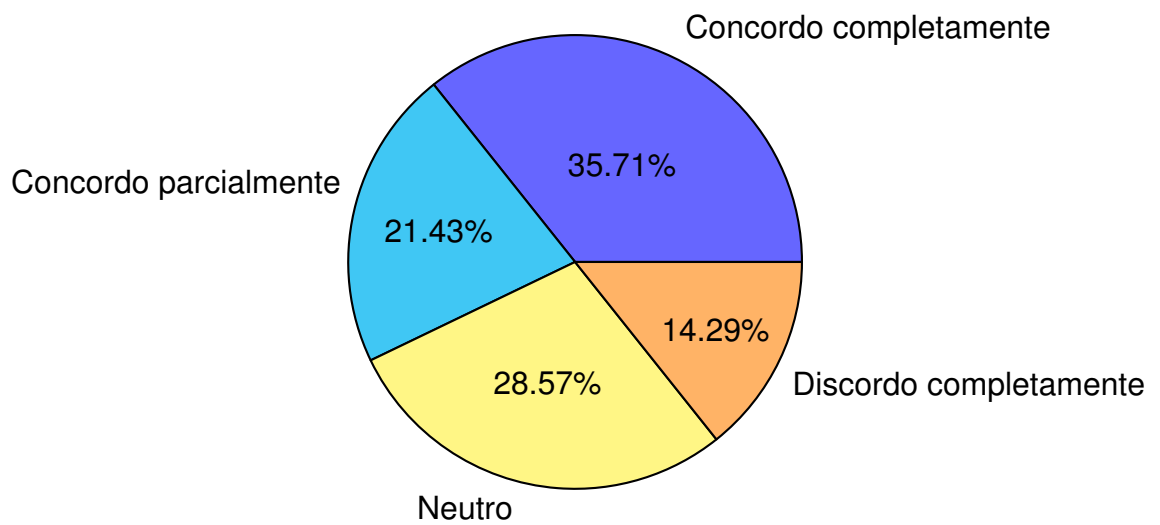
**QUESTÃO 4 - A FERRAMENTA SEGUE UM PADRÃO CONSISTENTE EM SUA APRESENTAÇÃO DE RESULTADOS E INTERFACE?**

Este é um ponto forte, com 76.92% dos usuários concordando completamente com a consistência da interface e da apresentação de resultados. No entanto, 15.38% expressam algum grau de discordância, o que pode indicar problemas isolados de inconsistência.



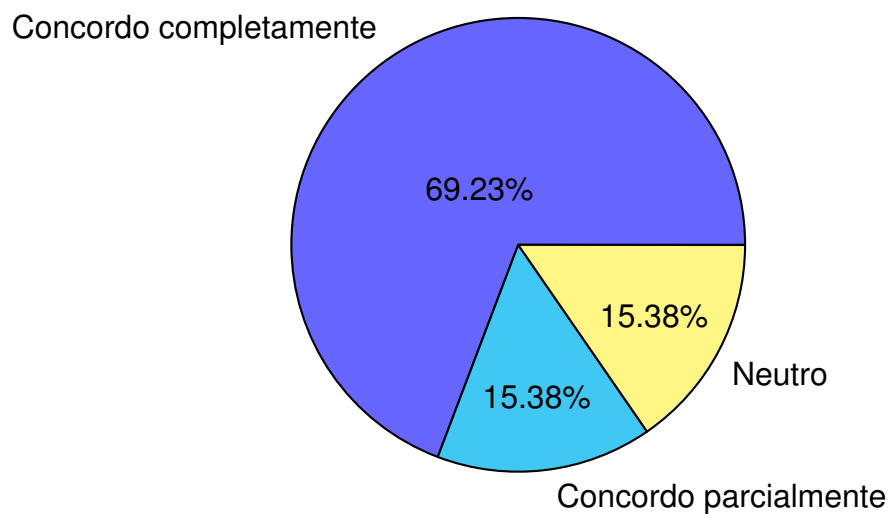
**QUESTÃO 5 - A FERRAMENTA OFERECE MEDIDAS PARA EVITAR OU CORRIGIR ERROS DURANTE O PROCESSO DE MONITORAMENTO?**

A opinião sobre a capacidade da ferramenta de evitar ou corrigir erros é mista, com 35.71% de concordância completa, mas 14.29% discordam completamente, revelando uma área crítica que requer atenção para melhorar a confiabilidade.



**QUESTÃO 6 - AS OPÇÕES DISPONÍVEIS SÃO CLARAMENTE APRESENTADAS?**

A clareza das opções disponíveis é bem vista, com 69.23% concordando completamente, sugerindo que a ferramenta é relativamente intuitiva e fácil de usar.

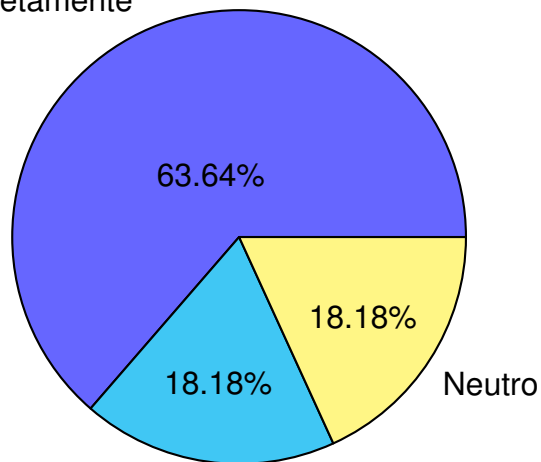


**QUESTÃO 7 - AS AÇÕES DISPONÍVEIS SÃO CLARAMENTE APRESENTADAS?**

Similarmente, 63.64% dos usuários concordam completamente que as ações disponíveis são claras, indicando uma interface amigável.



Concordo completamente

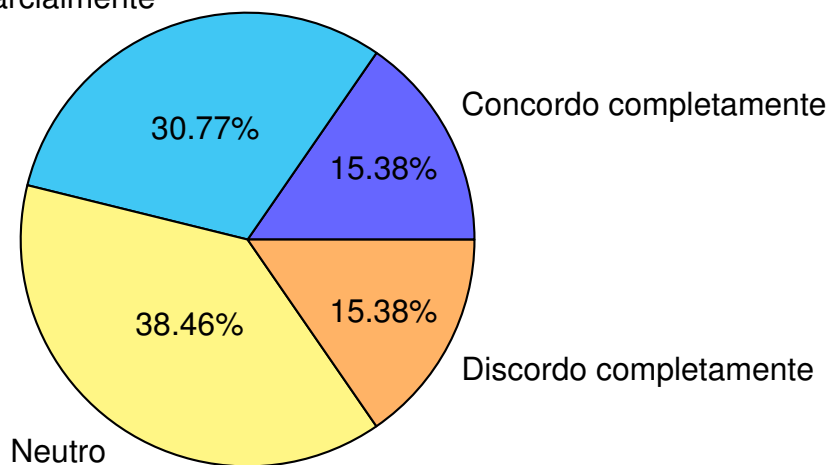


Concordo parcialmente

**QUESTÃO 8 - A FERRAMENTA OFERECE OPÇÕES PARA USUÁRIOS AVANÇADOS OU PERSONALIZAÇÃO DE CONFIGURAÇÕES?**

A personalização e opções avançadas são áreas fracas, com apenas 15.38% de concordância completa e 15.38% de discordância completa, destacando uma disparidade significativa na percepção do usuário.

Concordo parcialmente



Concordo completamente

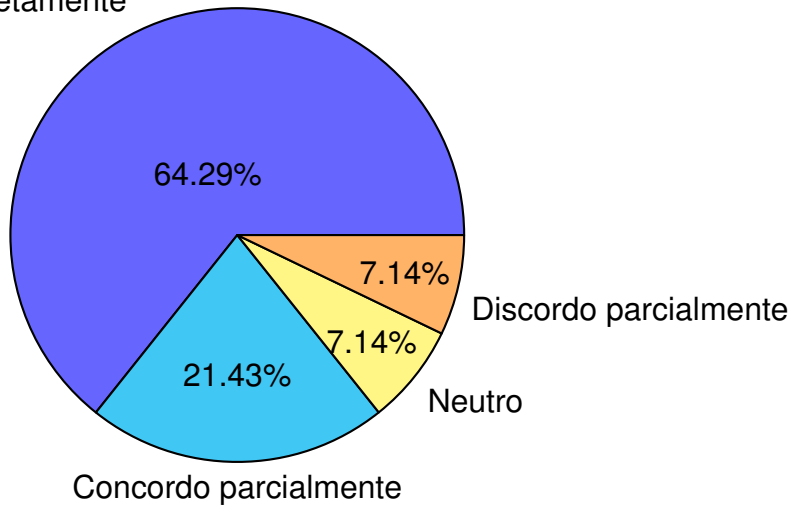
Discordo completamente

Neutro

**QUESTÃO 9 - A INTERFACE DA FERRAMENTA É LIMPA, ORGANIZADA E LIVRE DE ELEMENTOS DESNECESSÁRIOS?**

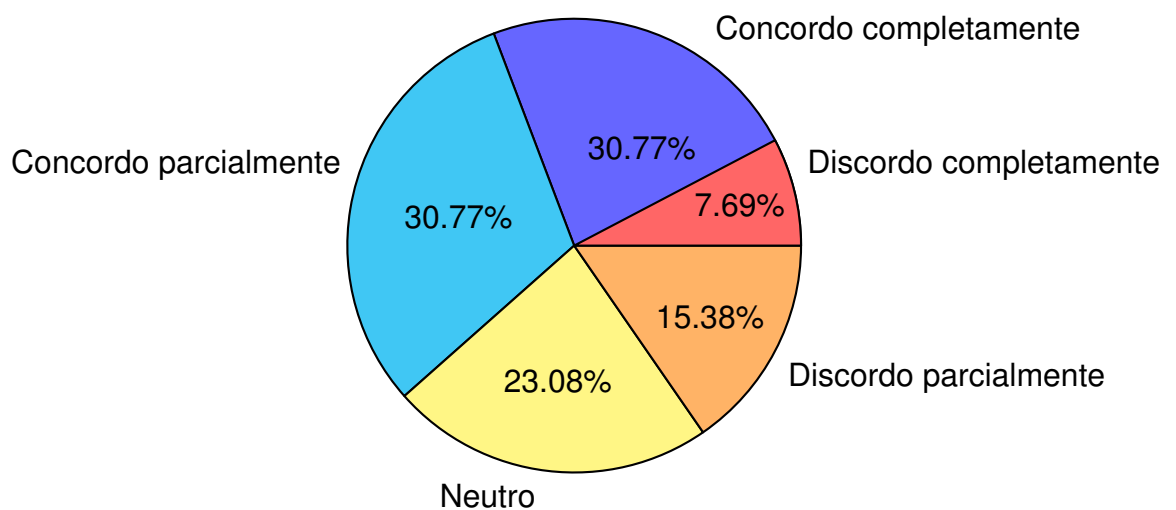
A maioria vê a interface como organizada e livre de elementos desnecessários, com 64.29% de concordância completa, mas ainda há margem para melhorias.

Concordo completamente



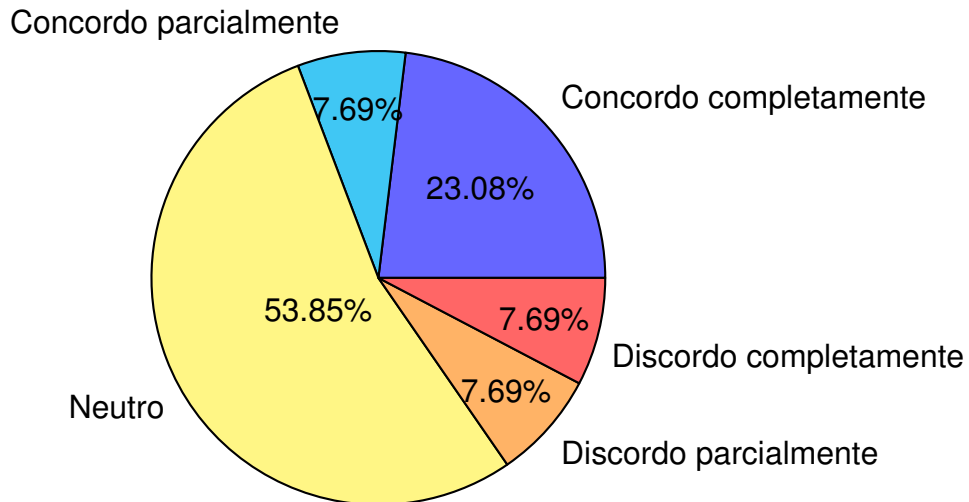
**QUESTÃO 10 - A FERRAMENTA FORNECE AJUDA CLARA E INSTRUÇÕES PARA RESOLVER PROBLEMAS OU INTERPRETAR RESULTADOS?**

As opiniões sobre o suporte oferecido são variadas, com 30.77% concordando completamente e 15.38% discordando completamente, refletindo a necessidade de melhorar a clareza e a acessibilidade do suporte.



**QUESTÃO 11 - EXISTE DOCUMENTAÇÃO OU SUPORTE DISPONÍVEL PARA AJUDAR OS USUÁRIOS A ENTENDER E UTILIZAR A FERRAMENTA DE FORMA EFICAZ?**

A documentação ou suporte é uma área problemática, com 53.85% dos usuários permanecendo neutros e 15.38% discordando completamente, indicando que os recursos de apoio podem ser insuficientes ou inadequados.



## 5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O avanço exponencial na geração e armazenamento de dados ao longo da última década, aliado ao aumento do uso de dispositivos eletrônicos durante a pandemia de COVID-19, trouxe à tona a importância de técnicas avançadas para a análise de informações na Web. A mineração de dados, junto com o processamento de linguagem natural e o aprendizado de máquina, mostrou-se fundamental para a detecção de notícias falsas e outras desinformações. Esse contexto destaca a necessidade de ferramentas que auxiliem na organização e verificação das vastas quantidades de informações disponíveis online. A propagação de notícias falsas representa um sério desafio para a sociedade moderna, podendo levar à desinformação, polarização política e instabilidade social. Identificar e refutar essas informações enganosas é crucial para manter a integridade das informações públicas e proteger a sociedade. As técnicas de verificação de fatos, que incluem a análise de fontes, a verificação de evidências e a comparação com dados confiáveis, são essenciais nesse processo. Além disso, a inteligência artificial e o aprendizado de máquina contribuem significativamente ao identificar padrões que caracterizam as notícias falsas.

O principal objetivo deste trabalho de conclusão de curso 1 foi investigar quais abordagens e métodos para representar o estado atual do domínio de pesquisa relevante e transmitir aos pesquisadores um conhecimento atualizado. Nele, foram discutidos diferentes tipos de *fake news* e pesquisadas quais as técnicas existentes para detectá-las, focando no pré-processamento, métodos de extração de recursos e desempenho da classificação. A mineração de dados desempenha um papel central na detecção de notícias falsas, permitindo a coleta e análise de grandes volumes de informações para identificar padrões específicos que distinguem o conteúdo verdadeiro do enganoso. Com a crescente quantidade de dados gerados na web, a aplicação de técnicas de mineração torna-se indispensável para organizar e estruturar essas informações. Isso possibilita que organizações e empresas tomem decisões mais informadas e eficazes, baseadas em dados confiáveis.

Já o presente trabalho foca em apresentar os resultados do trabalho de conclusão de curso 2. O aprendizado de máquina e o *deep learning*, subcampos da

inteligência artificial, são fundamentais para o desenvolvimento de sistemas capazes de aprender com grandes quantidades de dados. Esses sistemas podem identificar padrões complexos e tomar decisões baseadas em novos dados, imitando a capacidade humana de aprendizado. A aplicação dessas tecnologias na detecção de notícias falsas mostra-se promissora, oferecendo soluções cada vez mais precisas para um problema em constante evolução.

Por fim, a implementação de técnicas automatizadas de detecção de notícias falsas, como as desenvolvidas no contexto da AVICO, é crucial para garantir a confiabilidade das informações. Ferramentas de *web scraping* e *web crawling*, integradas em *APIs* e sistemas de monitoramento, ajudam a identificar e filtrar informações enganosas de maneira eficiente. A ferramenta oferece uma abordagem automatizada para monitorar e analisar conteúdos suspeitos, contribuindo para a mitigação dos impactos negativos causados pelas notícias falsas. Através de atualizações contínuas e do envio de resultados por *e-mail*, os usuários têm acesso conveniente às informações verificadas, permitindo uma resposta rápida e eficaz diante de potenciais ameaças à integridade da informação online.

Além disso, foi avaliada a ferramenta de monitoramento de notícias falsas da AVICO. A ferramenta demonstrou ter uma base sólida com uma interface consistente e opções claras, contribuindo significativamente para a identificação e combate à disseminação de informações falsas. No entanto, a análise do uso da ferramenta revelou a necessidade de melhorias em áreas críticas como a precisão dos resultados e a personalização das funções. Apenas uma parte dos usuários considerou os resultados completamente confiáveis, indicando que a precisão precisa ser aprimorada. Além disso, o suporte ao usuário e a documentação necessitam de melhorias para garantir uma utilização eficaz da ferramenta. Antes da implantação completa, é recomendável focar nessas áreas para aumentar a confiabilidade e a satisfação dos usuários. A integração contínua de atualizações e o envio de resultados por *e-mail* oferecem um acesso conveniente às informações verificadas, permitindo uma resposta rápida e eficaz a potenciais ameaças à integridade da informação online.

## REFERÊNCIAS

ALAMOODI, A. et al. Sentiment analysis and its applications in fighting covid-19 and infectious diseases: A systematic review. **Expert Systems with Applications**, 2021, v. 167, p. 114155, 2021. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417420308988>>.

ANALYSIS, C.-. vaccine; literature review; machine learning; public perception; sentiment. Role of artificial intelligence for analysis of covid-19 vaccination-related tweets: Opportunities, challenges, and future trends. **Mathematics**, 2022, v. 10, 2022. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137805826&doi=10.3390%2fmath10173199&partnerID=40&md5=b2a1e5359eccc3820011941ce499aa>>.

ASUDANI, D. S.; NAGWANI, N. K.; SINGH, P. Impact of word embedding models on text analytics in deep learning environment: a review. **Artificial Intelligence Review**, 2023, 2023. ISSN 0269-2821, 1573-7462. Disponível em: <<https://link.springer.com/10.1007/s10462-023-10419-1>>.

DIAZ-GARCIA, J. A.; RUIZ, M. D.; MARTIN-BAUTISTA, M. J. A survey on the use of association rules mining techniques in textual social media. **Artificial Intelligence Review**, 2023, v. 56, p. 1175–1200, 2023. ISSN 0269-2821, 1573-7462. Disponível em: <<https://link.springer.com/10.1007/s10462-022-10196-3>>.

FERNANDES, E.; MORO, S.; CORTEZ, P. Data science, machine learning and big data in digital journalism: A survey of state-of-the-art, challenges and opportunities. **Expert Systems with Applications**, 2023, p. 119795, 2023. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417423002968>>.

IGNACZAK, L. et al. Text mining in cybersecurity: A systematic literature review. **ACM Comput. Surv.**, 2021, v. 54, 2021. ISSN 0360-0300. Disponível em: <<https://doi-org.ez96.periodicos.capes.gov.br/10.1145/3462477>>.

MINING, C. D. mining; Deep learning; Machine learning; Natural language processing; Social media analysis; Spam content; T. A systematic literature review on spam content detection and classification. **PeerJ Computer Science**, 2022, v. 8, 2022. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126372646&doi=10.7717%2fPEERJ-CS.830&partnerID=40&md5=986b3b16cff26fc9ad16ef489b3b328b>>.

MISINFORMATION, C. C.-. D. learning; Machine learning;. Applications of machine learning for covid-19 misinformation: a systematic review. **Social Network Analysis and Mining**, 2022, v. 12, 2022. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135183706&doi=10.1007%2fs13278-022-00921-9&partnerID=40&md5=1f98c943b3ecc560d967dd3caf20da49>>.

PROCESSING deep learning; fake news; machine learning; Natural language. A comprehensive review on fake news detection with deep learning. **IEEE Access**, 2021, v. 9, p. 156151–156170, 2021. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120045642&doi=10.1109%2fACCESS.2021.3129329&partnerID=40&md5=9ac0830a44fae41640d2ec7e23668004>>.

RAZALI, N. A. M. et al. Opinion mining for national security: techniques, domain applications, challenges and research opportunities. **Journal of Big Data**, 2021, v. 8, p. 150, 2021. ISSN 2196-1115. Disponível em: <<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00536-5>>.

SYNTAX, A. D. tools; Fake news; Linguistics; Machine learning; Semantics; Social media;. Approaches to identify fake news: A systematic literature review. **Lecture Notes in Networks and Systems**, 2021, v. 136, p. 13–22, 2021. Disponível em: <[https://www.scopus.com/inward/record.uri?eid=2-s2.0-85087556918&doi=10.1007%2f978-3-030-49264-9\\_2&partnerID=40&md5=05e066d15e29fd28fea3499aa438e13a](https://www.scopus.com/inward/record.uri?eid=2-s2.0-85087556918&doi=10.1007%2f978-3-030-49264-9_2&partnerID=40&md5=05e066d15e29fd28fea3499aa438e13a)>.