UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

Gabriel Cortizo Ferraz

Utilizando aprendizagem de máquina para avaliar políticas de privacidade

RECIFE
2023

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO

Gabriel Cortizo Ferraz

Utilizando aprendizagem de máquina para avaliar políticas de privacidade

Monografia apresentada ao Centro de Informática (CIN) da Universidade Federal de Pernambuco (UFPE), como requisito parcial para conclusão do Curso de Engenharia da Computação, orientada pela professora Jéssyka Flavyanne Ferreira Vilela.

RECIFE
2023

AGRADECIMENTOS

Inicialmente, desejo expressar minha gratidão à minha família e amigos pelo contínuo apoio que me concederam. Em especial, a minha mãe por sempre ressaltar a importância da educação. Além disso, gostaria de estender meus agradecimentos aos professores e professoras do Centro de Informática da Universidade Federal de Pernambuco, que contribuíram para minha formação acadêmica durante os anos de graduação. Em particular, quero expressar minha gratidão à professora Jéssyka Vilela pelo suporte na realização deste trabalho.

RESUMO

Contexto: A partir do século 20, surgiram as primeiras leis de proteção de dados pessoais em reconhecimento ao impacto que o uso desregulado desses dados pode causar aos seus titulares. Em resposta aos grandes vazamentos de dados, as leis foram adaptadas às mudanças dinâmicas do mundo digital. No Brasil, a Lei Geral de Proteção de Dados Pessoais (LGPD) regulamenta a coleta, uso e retenção de dados pessoais por instituições, garantindo segurança jurídica aos usuários que compartilham seus dados. A LGPD também concede aos titulares dos dados o direito de acesso às informações sobre o tratamento de seus dados, que devem ser claras e adequadas em políticas de privacidade. Problema: Embora políticas de privacidade sejam de suma importância para informar ao usuário como seus dados são utilizados e coletados, as políticas de privacidade geralmente possuem textos longos e de difícil compreensão. Objetivo: Desenvolver uma ferramenta para avaliação de políticas de privacidade de acordo com critérios de avaliação definidos na literatura, utilizando aprendizagem de máquina na identificação dos critérios. Método**:** (1) Foram analisadas 23 Políticas de Privacidade para construção da base de dados; (2) Algoritmos de aprendizagem de máquina foram avaliados e comparados para a classificação de um título quanto a um critério; (3) Ferramenta foi analisada manualmente. Resultados**:** O algoritmo de classificação utilizado teve obteve uma eficácia de 81%, enquanto a ferramenta final obteve uma taxa de sucesso de 72.7%. Conclusões: De acordo com a taxa de eficácia da política ser acima de 72.7%, a ferramenta mostrou ser uma alternativa válida na avaliação de títulos de políticas de privacidade.

Palavras-chaves: Política de Privacidade, Privacidade, Aprendizado de máquina, Critérios de avaliação.

ABSTRACT

Context: Starting in the 20th century, the first laws protecting personal data emerged in recognition of the impact that unregulated use of such data can have on its owners. In response to major data breaches, laws have been adapted to the dynamic changes of the digital world. In Brazil, the General Data Protection Law (LGPD) regulates the collection, use and retention of personal data by institutions, providing legal security to users who share their data. The LGPD also grants data owners the right to access information about the treatment of their data, which must be clear and adequate in privacy policies. Problem: Although privacy policies are of utmost importance to inform the user how their data is used and collected, privacy policies often have long and difficult-to-understand texts. Objective: Develop a tool for evaluating privacy policies according to criteria defined in the literature, using machine learning in the identification of criteria. Method: (1) 23 Privacy Policies were analyzed for database construction; (2) Machine learning algorithms were evaluated and compared for the classification of a title according to a criterion. (3) The tool was manually analyzed. Results: The classification algorithm used had an efficacy of 81%, while the final tool achieved a success rate of 72.7%. Conclusions: According to the policy's efficacy rate being above 72.7%, the tool proved to be a valid alternative in evaluating privacy policy titles.

Keywords: Privacy Policy, Privacy, Machine Learning, Quality Rating Criteria.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

1.1 Contexto

No século 20 surgem as primeiras leis que levam em consideração a proteção de dados pessoais [1], um dos pontos iniciais no reconhecimento do impacto que o uso desregulado de dados podem ter nos seus titulares. Diante de um ambiente dinâmico como o mundo digital, as leis se adaptaram aos tempos atuais principalmente em respostas a grandes vazamentos de dados. As leis surgiram em resposta a grandes incidentes de vazamento de dados, como o da empresa de buscas Yahoo[1], onde mais de 1 milhão de usuários tiveram dados como nome, email e senhas expostos entre os anos de 2013 e 2016 [19].

No Brasil, a Lei nº 13.709/2018, conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD), regulamenta o modo como as instituições devem agir durante a coleta, o uso e retenção de dados pessoais de usuários. A promulgação da lei contribui para uma segurança jurídica aos usuários que compartilham dados com terceiros. A lei brasileira, também estabelece o direito do titular dos dados de ter o acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso, essas informações são geralmente agrupadas em um documento disponibilizado pelas empresas chamado de política de privacidade, onde também constam informações sobre como os dados dos usuários são coletados, armazenados e utilizados.

Tendo em vista que políticas de privacidade são fundamentais para a tomada de decisão do usuário para o compartilhamento de dados pessoais, é imprescindível que as políticas atendam a critérios de qualidade, como clareza e transparência para que sejam eficazes [22].

1.2 Motivação e Justificativa

Política de privacidade é um documento que descreve como uma organização coleta, usa, armazena e protege as informações pessoais dos usuários. Esse documento fornece transparência sobre como os dados pessoais são gerenciados e ajuda os usuários a entenderem como suas informações serão tratadas [7].

[1] Yahoo: https://yahoo.com/

Embora políticas de privacidade sejam fundamentais para compreensão do usuário de como seus dados são coletados e tratados, é comum que as mesmas possuam textos longos e de difícil compreensão [13, 18]. No trabalho de 2015, *"The readability of privacy policies"* [18], os autores avaliaram a legibilidade das políticas de privacidade de 16 redes sociais populares para estimar o nível de leitura necessário para compreender as políticas de privacidade de cada rede social. Eles descobriram que a maioria das políticas de privacidade eram lidas e compreendidas, com uma média de leitura de nível universitário. A política de privacidade do Facebook foi a mais difícil de ler, exigindo o nível de leitura de um estudante universitário de pós-graduação.

No trabalho, *"Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade"* [12] , o autor propõe a criação de uma ferramenta para análise de políticas de privacidade por meio de critérios pré-definidos, com o uso de processamento de linguagem natural, contribuindo assim para tornar a leitura mais acessível e compreensível ao usuário. O autor sugere como oportunidades futuras a utilização de aprendizado de máquina na identificação dos critérios utilizados na avaliação. O trabalho atual propõe a construção de uma ferramenta que utilize processamento de linguagem natural e aprendizado de máquina para identificar se uma determinada política de privacidade atende a determinados critérios de avaliação de políticas de privacidade. O trabalho também propõe a comparação dos resultados obtidos com outras ferramentas da literatura.

1.3 Objetivos

O objetivo geral deste trabalho é o desenvolvimento de uma ferramenta para avaliação de políticas de privacidade de acordo com critérios definidos na literatura, especificamente com o uso de aprendizagem de máquina na avaliação dos critérios.

Para atingir esse objetivo geral, destaca-se os seguintes objetivos específicos:

- Avaliar e definir critérios de avaliação de políticas na literatura para serem utilizadas na ferramenta;
- Escolha e treinamento do modelo de aprendizagem para classificação de critérios;
- Definir o fluxo de dados da ferramenta;

- Definição das tecnologias utilizadas na implementação da ferramenta;
- Implementação da ferramenta;
- Comparação da eficácia da ferramenta com outras ferramentas da literatura com propósitos semelhantes.

1.4 Trabalhos Relacionados

Tendo em vista a importância que as políticas de privacidade exercem na segurança jurídica do usuário e ao mesmo tempo as dificuldades de interpretação e clareza que as mesmas possuem [13], a utilização de ferramentas para avaliação de políticas de privacidade é tema já explorado, com alguns exemplos na literatura.

Os autores do trabalho *"Automatic Privacy Policy Analysis and Its Legal Implications"* [21] utilizaram mineração de texto na extração de informações relevantes das políticas de privacidade, como os tipos de dados coletados e a finalidade da coleta de dados. Posteriormente, os dados extraídos foram comparados com as leis de privacidade da União Europeia (GDPR) e da Califórnia (CCPA) para avaliar a conformidade das políticas de privacidade com essas leis. A eficácia do modelo proposto foi avaliada em uma amostra de 50 políticas de privacidade de empresas do setor de tecnologia, e possui uma taxa de acerto de 94% na identificação das principais obrigações impostas pelas leis de privacidade, o que indica sua eficácia na avaliação de conformidade das políticas de privacidade com as leis de privacidade.

No trabalho *"Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade"* [12], o autor propõe a criação de uma ferramenta para avaliar políticas de privacidade brasileiras com a utilização de critérios de avaliação de políticas de privacidade baseadas nos catálogo elencado por Augusto Terra no trabalho *"Catálogo de critérios para avaliação de políticas de privacidade"* [14]. No trabalho de [12], 14 critérios de avaliação foram utilizados na construção de uma ferramenta para identificar a presença ou não desses critérios nos títulos dos textos de políticas de privacidade. O autor utilizou 6 políticas de privacidade na construção da ferramenta que utiliza processamento de linguagem natural para a identificação dos critérios.

Este trabalho explora a adoção de aprendizagem de máquina na identificação de critérios de políticas de privacidade como alternativa ao algoritmo utilizado no

trabalho "*Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade"* [12], comparando se o seu uso possui maior eficácia na identificação de critérios.

1.5 Estrutura do documento

Este trabalho se divide em 6 capítulos. No capítulo 2 é apresentado o referencial teórico, com a familiarização do conceito de privacidade, leis de privacidade e proteção de dados e políticas de privacidade. O terceiro capítulo envolve a metodologia utilizada na construção da ferramenta, na escolha do modelo de classificação, na escolha de tecnologias e na implementação da ferramenta. O capítulo 4 compreende os resultados alcançados com a implementação da ferramenta e os compara com resultados de ferramentas na literatura. O capítulo 5 contempla a conclusão e discute trabalhos futuros para a continuidade da ferramenta.

2 REFERENCIAL TEÓRICO

Nesta seção, são descritos os principais conceitos relacionados neste trabalho: privacidade, dados pessoais, leis de privacidade, proteção de dados pessoais, avaliação de políticas de privacidade, processamento de linguagem natural e algoritmos de aprendizagem de máquina para classificação de textos.

2.1 Privacidade

O conceito de privacidade tem evoluído ao longo dos anos, refletindo mudanças nas práticas sociais, tecnológicas e legais. Inicialmente, o conceito de privacidade era associado ao reconhecimento da inviolabilidade da residência. Posteriormente foi expandido, compreendendo não apenas a propriedade material mas o indivíduo em si, como exemplificado no artigo de 1890 *“The Right to Privacy”* [3] onde os autores defendem o *“direito de ser deixado em paz”*, em resposta ao desenvolvimento tecnológico da época com o surgimento de câmeras fotográficas e da popularização imprensa. A partir desse momento a privacidade foi se tornando um tema mais presente nas discussões jurídicas e políticas

A criação da Organização das Nações Unidas (ONU) foi fundamental para a adoção da privacidade como um direito humano fundamental, a Declaração Universal dos Direitos Humanos da ONU em 1948, estabelece que *"Ninguém será sujeito a interferências em sua vida privada, em sua família, em seu lar ou em sua correspondência, nem a ataques à sua honra e reputação. Todo ser humano tem direito à proteção da lei contra tais interferências ou ataques."*[5].

Com a popularização da internet e com o uso aparelhos eletrônicos, o conceito de privacidade passou a ser associado às informações digitais tendo em vista que as mesmas podem englobar desde dados pessoais, como nome e informações de contato, até informações mais sensíveis, como histórico médico e dados de localização.

2.2 Dados Pessoais

Segundo o Regulamento do Parlamento Europeu e o Conselho da União Européia 2016/679 [4], dados pessoais são informações relativas a uma pessoa singular que possa ser identificada diretamente ou indiretamente, em especial por referência de um identificador, como nome, dados de localização ou um ou mais

elementos específicos da identidade física, econômica ou social dessa pessoa singular. É válido pontuar que dados como cookies e endereços IP também podem ser considerados dados pessoais.

Alguns tipos de dados pessoais podem ser classificados como **sensíveis.** Esse tipo de dado está relacionado a características da personalidade do indivíduo e suas escolhas pessoais, como origem racial ou étnica, convicção religiosa, opinião religiosa, opinião política, filiação a sindicato ou organização de caráter religioso, filosófico,político, genético ou biométrico quando vinculado a uma pessoa natural.

2.3 Leis de privacidade e de proteção de dados pessoais

As leis de privacidade e proteção de dados são interdependentes, pois ambas visam proteger as informações pessoais dos indivíduos. As leis de privacidade garantem o controle sobre o uso desses dados, enquanto as leis de proteção de dados regulam a coleta, processamento e armazenamento dessas informações[6], assim, servindo como uma garantia para o exercício de direitos fundamentais.

Embora existam diferentes legislações entre países e regiões, foi notado um conjunto comum de princípios a serem aplicados na proteção de dados pessoais [1] baseados nas Guidelines da OCDE [2].

Os princípios de acordo com OCDE [2] são:

1. **Princípio da publicidade (ou transparência):** A existência do armazenamento de dados pessoais deve ser de conhecimento público, seja por meio da exigência de autorização prévia para funcionar, da notificação a uma autoridade sobre sua existência.
2. **Princípio da exatidão:** Dados armazenados devem ser fiéis à realidade.
3. **Princípio da finalidade:** Qualquer utilização de dados pessoais deve ser baseada na finalidade comunicada ao titular dos dados antes da coleta de seus dados, os fins devem ser legítimos e específicos.
4. **Princípio do livre acesso:** O indivíduo deve ter acesso a quais informações estão sendo coletadas, podendo obter cópias desses registros.
5. **Princípio da segurança física e lógica:** Dados devem ser protegidos contra riscos de extravio, acesso indevido, destruição e modificação.

2.3.1 LGPD

No Brasil, a Lei nº 13.709/2018, também conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD) foi promulgada para regulamentar o modo como as instituições devem agir durante a coleta, retenção e processamento de dados pessoais, contribuindo para segurança jurídica aos usuários que compartilham dados.

A LGPD se aplica a todos que realizam tratamento de dados pessoais, organizações públicas ou privadas, pessoas físicas ou jurídicas que realizam tratamento de dados, seja utilizando o meio físico ou digital, desde que possa envolver ao menos um dos elementos:

- Ocorrer em território nacional;
- Tenha por objetivo oferta ou fornecimento de bens ou serviços ou tratamento de dados de indivíduos localizados em território nacional;
- Em que os dados sejam coletados em território nacional;

Desse modo, a lei não se restringe à nacionalidade ou cidadania do titular dos dados, nem seu local residencial atual.

É válido pontuar que de acordo com o art 4º da LGDP, a lei não se aplica quando o tratamento de dados é realizado por uma pessoa física, para fins exclusivamente particulares e não econômicos, para fins jornalísticos e artísticos e para tratamentos realizados para fins de segurança pública e defesa nacional.

Quanto às penalidades previstas pela LGPD constam: advertência, com indicação de prazo para adoção de medidas corretivas, multa de até 2% do faturamento anual da organização no Brasil – limitado a R$ 50 milhões por infração, eliminação dos dados pessoais a que se refere a infração, proibição parcial ou total de atividades relacionadas ao tratamento de dados entre outras.

O 11º artigo da LGPD estabelece as hipóteses legais para o tratamento de dados pessoais. Entre as hipótese legais consta o consentimento, onde o titular de dados ou seu responsável legal consente com o tratamento de dados para finalidades específicas informadas previamente. Também é válido pontuar que existem hipóteses legais de tratamento mesmo quando não há consentimento explícito, desde que os dados forem indispensáveis para:

- Cumprimento de obrigação legal ou regulatória pelo controlador;

- Tratamento compartilhado de dados necessários à execução, pela administração pública, de políticas públicas previstas em leis ou regulamentos;
- Realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais sensíveis;
- Exercício regular de direitos, inclusive em contrato e em processo judicial, administrativo e arbitral, este último nos termos da Lei nº 9.307, de 23 de setembro de 1996 (Lei de Arbitragem);
- Proteção da vida ou da incolumidade física do titular ou de terceiro;
- Tutela da saúde, exclusivamente, em procedimento realizado por profissionais de saúde, serviços de saúde ou autoridade sanitária; ou
- Garantia da prevenção à fraude e à segurança do titular, nos processos de identificação e autenticação de cadastro em sistemas eletrônicos, resguardados os direitos mencionados no art. 9º desta Lei e exceto no caso de prevalecer em direitos e liberdades fundamentais do titular que exijam a proteção dos dados pessoais.

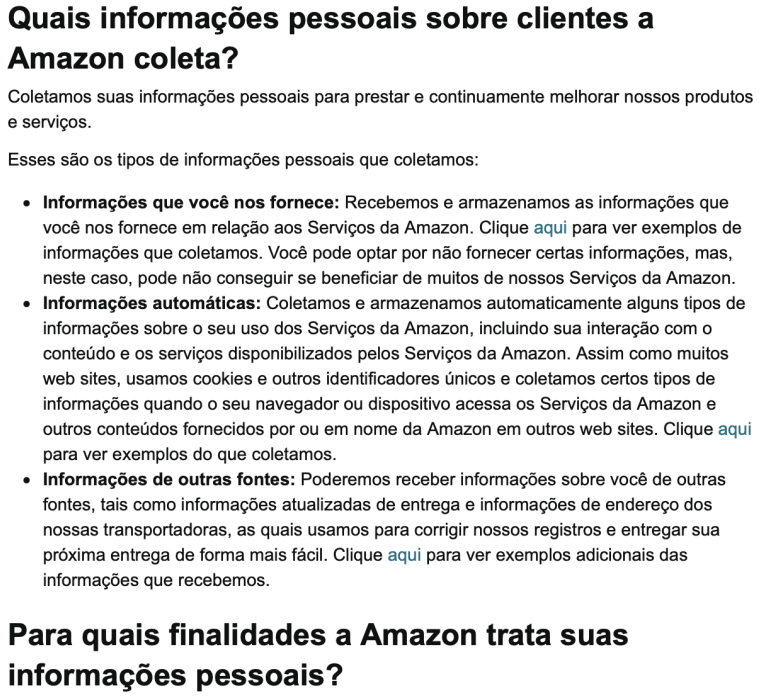
2.4 Políticas de privacidade

Política de privacidade é um documento que descreve como uma organização coleta, usa, armazena e protege as informações pessoais dos usuários [23]. Esse documento fornece transparência sobre como os dados pessoais são gerenciados e ajuda os usuários a entenderem como suas informações serão tratadas [7]. As políticas de privacidade geralmente estão disponíveis nos sites ou aplicativos das empresas, fornecendo informações claras e precisas sobre como suas informações pessoais são tratadas.

Algumas das principais legislações de proteção de dados reforçam a criação e a divulgação de políticas de informações por parte das empresas. O 13° artigo da GDPR estabelece que as empresas devem fornecer informações claras e transparentes aos titulares dos dados sobre como seus dados pessoais serão processados e o motivo de sua coleta, enquanto o 9° artigo da LGPD estabelece que "*O titular tem direito ao acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva*

acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso”.

O trecho da política de privacidade da Amazon na Figura 1, exemplifica algumas das informações esperadas em documentos desse modelo, como informações sobre os dados são coletados do usuários e qual a motivação por trás da coleta.



Quais informações pessoais sobre clientes a Amazon coleta?

Coletamos suas informações pessoais para prestar e continuamente melhorar nossos produtos e serviços.

Esses são os tipos de informações pessoais que coletamos:

- **Informações que você nos fornece:** Recebemos e armazenamos as informações que você nos fornece em relação aos Serviços da Amazon. Clique aqui para ver exemplos de informações que coletamos. Você pode optar por não fornecer certas informações, mas, neste caso, pode não conseguir se beneficiar de muitos de nossos Serviços da Amazon.
- **Informações automáticas:** Coletamos e armazenamos automaticamente alguns tipos de informações sobre o seu uso dos Serviços da Amazon, incluindo sua interação com o conteúdo e os serviços disponibilizados pelos Serviços da Amazon. Assim como muitos web sites, usamos cookies e outros identificadores únicos e coletamos certos tipos de informações quando o seu navegador ou dispositivo acessa os Serviços da Amazon e outros conteúdos fornecidos por ou em nome da Amazon em outros web sites. Clique aqui para ver exemplos do que coletamos.
- **Informações de outras fontes:** Poderemos receber informações sobre você de outras fontes, tais como informações atualizadas de entrega e informações de endereço dos nossas transportadoras, as quais usamos para corrigir nossos registros e entregar sua próxima entrega de forma mais fácil. Clique aqui para ver exemplos adicionais das informações que recebemos.

Para quais finalidades a Amazon trata suas informações pessoais?

Figura 1 - Trecho da política de privacidade da Amazon. Fonte: Amazon, 2023.

2.5 Avaliação de políticas de privacidade

Tendo em vista a importância que as políticas de privacidade possuem e os problemas comuns de legibilidade, complexidade e interpretação que as mesmas possuem [13], estudos foram feitos com o empenho de avaliar o quão adequada a uma Lei de privacidade uma política de privacidade é, possivelmente auxiliando os responsáveis pela escrita das mesmas na construção de um documento mais adequado ao usuário.

Um dos modos de avaliação é através do uso de critérios de avaliação, de modo que a adequação de uma política de privacidade a uma lei vigente seja compatível com a quantidade de critérios de avaliação atendidos pela mesma. Por

exemplo, alguns dos critérios de avaliação de políticas de privacidade são relativos à política de privacidade especificar como os dados de usuário são utilizados por uma empresa ou organização. O Apêndice A deste trabalho, disponibiliza exemplos de critérios de avaliação de políticas de privacidade utilizados neste trabalho.

2.6 Processamento de linguagem natural

O processamento de linguagem natural é a interseção da linguística, ciência da computação e matemática que se preocupa com o desenvolvimento de algoritmos e modelos estatísticos que permitem o reconhecimento, a análise e a geração de linguagem natural por computador [36]. Análise sintática, análise semântica, extração de informação são algumas das técnicas desenvolvidas na área.

É válido pontuar que existe uma sobreposição entre aprendizagem de máquina e processamento de linguagem natural. Algumas aplicações de PLN usam técnicas de AM para melhorar o desempenho de seus modelos, como por exemplo, na classificação de textos e na tradução automática [37, 38].

2.7 Aprendizagem de máquina e suas aplicações no problema classificação de textos

A aprendizagem de máquina é uma área da inteligência artificial que se dedica ao estudo e desenvolvimento de algoritmos que permitem que um sistema automatizado possa aprender a partir de dados, sem que seja explicitamente programado para isso [7]. A aprendizagem de máquina possui diversas aplicações, incluindo reconhecimento de fala, visão computacional e processamento de linguagem natural.

Uma das aplicações da aprendizagem de máquina é no problema de classificação de textos, como por exemplo na análise de sentimentos [8] e na detecção de spams [9]. Esse problema envolve o treinamento de um modelo através de um conjunto de dados rotulados, de modo que novos textos possam ser categorizados de acordo com as categorias pré-definidas.

Support Vector Machine, Árvores de decisão, Random Forest Tree, Logistic Regression e Naive Bayes são alguns dos algoritmos de aprendizado de máquina supervisionado, e serão abordados nas próximas subseções deste capítulo.

2.7.1 Support Vector Machine

O SVM é um algoritmo de aprendizado de máquina supervisionado que busca encontrar o hiperplano de separação ótimo entre duas classes de dados distintas. Esse hiperplano é escolhido de forma a maximizar a margem entre as classes, definida como a distância perpendicular entre o hiperplano e os pontos de dados mais próximos de cada classe [11]. O SVM é capaz de lidar com dados de alta dimensionalidade e pode ser utilizado tanto para problemas de classificação binária como para problemas de classificação multiclasse.

O SVM pode ser usado para classificar textos em uma ampla variedade de aplicativos, como categorização de tópicos e classificação de documentos. Um outro exemplo da técnica é na análise de sentimentos, como exemplificado no trabalho *"Twitter as a Corpus for Sentiment Analysis and Opinion Mining"* [10], onde os autores exploram o uso do SVM para classificação de sentimentos em mensagens curtas de mídia social.

2.7.2 Árvores de decisão

Árvore de decisão é um modelo de classificação que divide recursivamente um conjunto de dados em subconjuntos menores e mais homogêneos. Cada divisão é feita com base em uma das variáveis do conjunto de dados e o resultado é uma árvore binária onde cada nó interno representa uma variável e cada folha representa uma classe [33], formando um conjunto de regras de decisão organizadas em uma estrutura hierárquica.

O processo de construção de uma árvore de decisão começa com a seleção da melhor característica para o primeiro nó interno, com base em critérios como o ganho de informação ou a redução de entropia. A partir desse nó inicial, a árvore é expandida recursivamente, dividindo os dados de treinamento em subconjuntos menores em cada nó interno, com base em novos testes em outras características dos dados [34]. Uma vez construída a árvore, ela pode ser usada para prever a classe ou o valor de saída para novos dados, seguindo o caminho da árvore que corresponde às respostas dos testes sobre as características desses dados.

2.7.3 Random Forest Tree

Random Forest é um algoritmo de aprendizado de máquina que utiliza a técnica de bootstrap agregado para construção de uma uma coleção de árvores de decisão não correlacionadas construídas a partir de uma amostra aleatória dos dados de treinamento e variáveis de entrada com o objetivo de reduzir de *overfitting,* quando os o modelo se ajusta aos dados de treinamento, mas não é capaz de generalizar para os novos dados de teste . O algoritmo utiliza a média ou a maioria das previsões das árvores para gerar uma previsão final [7].

Classificação de texto é uma das áreas onde a técnica de Random Forest possui aplicações, como nos exemplos de classificação de spam e categorização de tópicos [35].

2.7.4 Logistic Regression

Logistic Regression ou Regressão Logística é um modelo de aprendizado de máquina supervisionado, utilizado quando a variável de resposta é atribuída a uma categoria. O modelo prevê a probabilidade da resposta pertencer a uma das categorias de acordo com suas variáveis utilizando uma função sigmóide que transforma a variável de resposta linearmente em uma escala de 0 a 1, o que representa a probabilidade da resposta pertencer à categoria de interesse [27].

A regressão logística pode ser estendida para problemas de classificação com múltiplas categorias[29]. Um dos exemplos de implementação de regressão logística para múltiplas categorias é através do treinamento um modelo de regressão logística para cada classe, considerando apenas os exemplos dessa categoria como positivos e todos os demais exemplos como negativos. Para classificar um novo exemplo, é feita a predição de probabilidade de pertencer a cada classe com base em cada modelo treinado, e a classe com a maior probabilidade é selecionada como a classe de predição [28]

2.7.5 Naive Bayes

O modelo Naive Bayes é um algoritmo de classificação probabilístico que utiliza o Teorema de Bayes no cálculo da probabilidade de uma instância pertencer a cada classe possível, com base nas suas características. Ele assume que a presença ou ausência de uma característica é independente da presença ou

ausência de qualquer outra característica, dada a classe de saída. O classificador calcula a probabilidade de cada classe para uma instância e seleciona a classe com a maior probabilidade como a classe predita [30].

O modelo Naive Bayes é amplamente utilizado em problemas de classificação de texto, como exemplificado no trabalho de análise de sentimentos *"Sentiment Analysis of Twitter Data using a Naïve Bayes Classifier"* [32], onde os os autores utilizam o algoritmo Naive Bayes para analisar o sentimento de tweets relacionados a eventos esportivos. Outra aplicação do algoritmo em problemas de classificação de texto é na detecção e na construção de filtros de spams[31].

2.7.6 Válidação cruzada

A validação cruzada é uma técnica para avaliar o desempenho de um modelo de aprendizado de máquina. O conjunto de dados é dividido em K partições disjuntas, com o modelo sendo treinado em K-1 dessas partições e avaliado no restante. O processo é repetido K vezes, com cada uma das K partições sendo usada uma vez para teste e as outras K-1 sendo usadas para treinamento. A média dos resultados obtidos nas K iterações é usada como uma estimativa do desempenho do modelo [40].

A Figura 2 ilustra o processo de validação cruzada com um número de partições igual a 10. É válido ressaltar que a eficácia total é o valor da média obtida com as eficácias das diferentes partições.

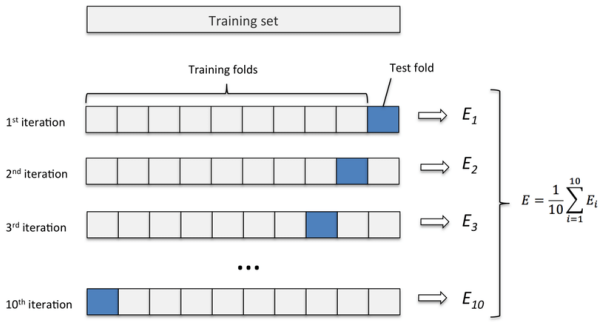


Figura 2 - Ilustração da técnica de validação cruzada utilizando 10 partições. Fonte: Karlrosaen, 2016

3 METODOLOGIA

A metodologia desse trabalho pode ser divida em 6 etapas:

1. Seleção de critérios de avaliação de políticas de privacidade;
2. Criação da base de dados utilizada para o treinamento de modelos;
3. Definição de modelo de aprendizagem de máquina para classificar políticas de privacidade;
4. Definição de tecnologias utilizadas na elaboração da ferramenta;
5. Implementação da ferramenta;
6. Testes na ferramenta.

3.1 Definição de critérios de avaliação de políticas de privacidade

Inicialmente, foram definidos os critérios de avaliação de políticas de privacidade que serão utilizados na construção da ferramenta e na classificação de títulos de políticas de privacidade. No trabalho Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade [12], Rodrigo Ferreira utiliza 14 critérios de avaliação de políticas de privacidade[2] baseados no catálogo de critérios criado por Augusto Terra [14]. Os mesmos 14 critérios foram utilizados no trabalho atual.

3.2 Criação da base de dados para treinamento de modelos

Após a definição dos critérios de avaliação de políticas de privacidade utilizados neste trabalho, títulos de 23 políticas de privacidade[3] escritas na língua portuguesa de médias e grandes empresas foram coletadas em um arquivo em formato csv e posteriormente classificados manualmente de acordo com os critérios definidos na primeira etapa. A base foi populada com políticas de privacidade de empresas de diferentes segmentos, como jornais, redes sociais e e-commerces com o objetivo da captura de textos de políticas de privacidade em diferentes contextos.É válido pontuar que devido a restrição temporal do trabalho atual, um número maior de políticas não pode ser avaliado.

[2] Disponível no Apêndice A
[3] Disponível no Apêndice B

A base de dados é composta por 3 colunas sendo estas:

- **company:** Empresa atribuída ao título da política de privacidade
- **text**: Texto dos títulos de uma das políticas de privacidade avaliadas.
- **class:** Corresponde ao critério de avaliação que a o título foi classificado.

Ao todo 375 títulos foram classificados em 15 diferentes categorias, 14 representando as diferentes categorias de critérios e uma representando a ausência de um critério. A coluna “Número de amostras na base de treinamento" no Apêndice A[4] deste trabalho, representa o número de vezes que um critério foi identificado na base de dados construída nessa etapa. Alguns critérios tiveram maior ocorrência na base de dados do que outros, enquanto o critério referente a *A Política de Privacidade específica claramente como a empresa pode usar os dados coletados?* possuiu 29 exemplos na base de dados o critério de Decisões automatizadas possuiu apenas uma amostra.

A Figura 3 exemplifica parte do conteúdo da base de dados construída nesta etapa.

	company	text	class
0	amazon	Notificação de Privacidade da Amazon	0
1	amazon	Última atualização: 29 de junho de 2022.	0
2	amazon	Controladores de informações pessoais	0
3	amazon	Quais informações pessoais sobre clientes a Am...	1
4	amazon	Para quais finalidades a Amazon trata suas inf...	2
...	...	...	...
370	correios	15. Alterações da Política de Privacidade d...	14
371	correios	16. Jurisdição para resolução de conflitos	0
372	correios	17. Informações Gerais	0
373	correios	18. Definições	0
374	correios	POLÍTICA DE COOKIES DOS CORREIOS	10

Figura 3 - Amostra da base de dados utilizada no treinamento de modelos.

Fonte: Autor, 2023.

[4] Disponível no Apêndice A

3.3 Definição de modelo de aprendizagem de máquina para classificar políticas de privacidade

Nesta etapa foi definido qual o modelo de aprendizagem de máquina deve ser utilizado para classificar um título quanto a um critério de avaliação de políticas de privacidade.

Após a criação e população da base de dados de títulos de políticas de privacidade e seus critérios de qualidade correspondentes, 5 modelos de aprendizagem supervisionada foram comparados quanto à eficácia na classificação de títulos ao seu critério de qualidade correspondente. Os modelos de aprendizagem de máquina avaliados foram avaliados foram os seguintes:

- Naive Bayes
- Random Forest
- Árvore de decisão
- Regressão Logística
- Support Vector Machine

Os modelos selecionados levam em consideração a limitação na quantidade de amostras de classes diferentes na base de dados, tendo em vista a utilização prévia desses algoritmos em trabalhos científicos com um número reduzido de amostras [15, 16, 17]. Por exemplo no trabalho *Identifying PTSD Symptoms using Twitter Data: A Microblogging Platform-based Machine Learning Approach* [17] os autores utilizam SVM na identificação de sintomas de transtorno de estresse pós-traumático em tweets, com um conjunto de apenas 300 amostras rotuladas.

Antes do treinamento dos modelos, os textos dos títulos da base de dados foram pré-processados, a fim de corrigir inconsistência e remover informações irrelevantes que podem influenciar negativamente a qualidade da análise. As operações de limpeza de dados utilizadas no pré-processamento foram, respectivamente:

1. Reescrita do texto utilizando letras minúsculas;
2. Lematização do texto;
3. Remoção de *stopwords*, palavras que ocorrem com frequência em um idioma e geralmente são filtradas durante o pré-processamento de dados em NLP e análise de texto [20], exemplo: “as”, “os” e “uns”.
4. Remoção de caracteres especiais, pontuação, números, e-mails e urls;

Após o pré-processamento inicial dos textos, os mesmos precisam ser reescritos em um formato de representação numérica, que possa ser utilizado pelo modelo em treinamento, no trabalho foi utilizado a técnica de *Frequency-Inverse Document Frequency*, para avaliar a relevância de uma palavra em um documento, em relação a um conjunto de documentos. Os dados pré-processados nesta etapa são em seguida utilizados para o treinamento dos modelos.

É válido pontuar que durante o treinamento de modelos, uma das etapas fundamentais é a de ajuste de *hiperparâmetros* dos modelos. *Hiperparâmetros* são parâmetros geralmente definidos antes do treinamento do modelo e que afetam o processo de aprendizado [28], influenciando diretamente a eficácia do modelo.

Para avaliar a eficácia de diferentes modelos e de diferentes configurações de hiperparâmetros foi utilizada a técnica de validação cruzada com configurações comuns em trabalhos científicos de aplicação de aprendizagem de máquina em problemas de classificação de textos [24, 25, 26], com número de divisão igual a 5. A Tabela 1 indica ilustra a eficácia dos diferentes algoritmos avaliados obtidos com o uso da validação cruzada.

O modelo que possui a maior eficácia foi o baseado em Support Vector Machine com 81% de acerto de eficácia, e por isso foi adotado na construção da ferramenta.

Modelo de aprendizado de máquina	**Eficácia**
Decision Tree Classifier	64%
Naive Bayes	69%
Random Forest	76%
Regressão Logística	79%
Support Vector Machine	81%

Tabela 1 - Comparação de eficácia de diferentes modelos de aprendizagem de máquina quanto a classificação de títulos em seus respectivos critérios de avaliação de políticas de privacidade.

3.4 Definição de tecnologias utilizadas na elaboração da ferramenta

Python[5] foi a linguagem de programação escolhida para implementação da ferramenta, tendo em vista a diversidade de bibliotecas de desenvolvimento web, aprendizado de máquina e processamento de linguagem natural, o que permitiria o uso de uma mesma linguagem em todos os módulos da ferramenta.

A ferramenta faz o uso da biblioteca sklearn[6] para a execução do modelo de aprendizagem de máquina, da biblioteca *nltk* no pré-processamento dos dados de entrada do modelo e do framework *flask*, utilizado na construção de aplicações web.

Quanto a construção da interface gráfica, o framework front-end Bootstrap[7] foi escolhido pela possibilidade de acelerar a prototipação da ferramenta

3.5 Implementação da ferramenta

Após a definição das tecnologias utilizadas na elaboração da ferramenta, o fluxo da aplicação foi definido, e a implementação da ferramenta foi iniciada.

Primeiramente, o usuário deve ser capaz de fazer o upload de uma política de privacidade em um arquivo de formato PDF, a escolha pela adoção do arquivo PDF se deve a falta de padronização na implementação das políticas de privacidade, o que dificulta a extração de informações do texto da política, em alguns casos avaliados sites de diferentes empresas utilizavam diferentes frameworks na construção das páginas web, utilizando diferentes tags HTML para estilizar títulos.

Após o upload do arquivo PDF, o texto da política é processado e os títulos identificados são extraídos, essa identificação de títulos ocorre, principalmente, pelo critério do tamanho da fonte no documento. Após a extração dos títulos, os mesmos são pré-processados com as mesmas operações utilizadas na seção 3.3 e posteriormente classificados pelo modelo de aprendizagem de máquina em algum dos critérios. Por fim, os resultados na etapa de predição bem como os títulos detectados na política de privacidade são formatados e retornados para o usuário. A Figura 4 ilustra o fluxo de fluxo de dados da aplicação desenvolvida

[5] Python: https://www.python.org/
[6] Sklearn: https://scikit-learn.org/stable/
[7] Bootstrap: https://getbootstrap.com/

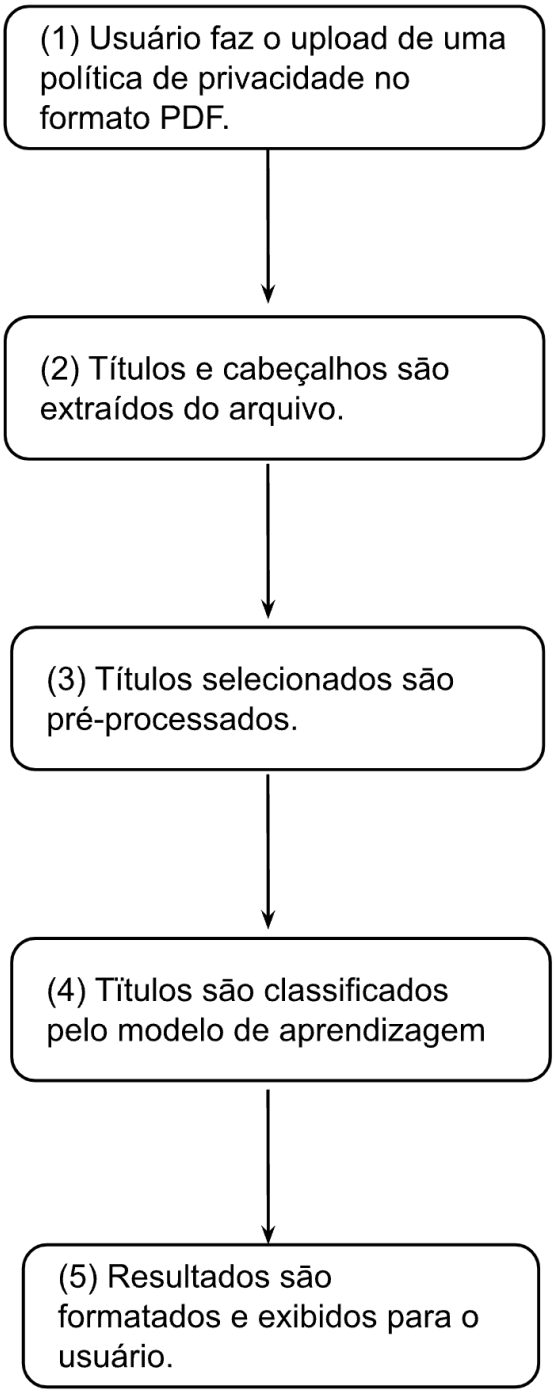


Figura 4 - Diagrama do fluxo utilizado na construção da aplicação

Fonte: Autor, 2023

3.6 Testes na ferramenta

Nesta seção são apresentados os testes de eficácia da ferramenta. Na avaliação foram utilizadas 3 métricas distintas, levando em consideração apenas a

capacidade da ferramenta na detecção de títulos, apenas a classificação dos títulos detectados e a eficácia total da ferramenta.

Na seção 3.3, o modelo de aprendizado de máquina baseado em Support Vector Machine, teve o melhor desempenho entre os demais modelos utilizados neste trabalho, com acurácia de 81%. Tendo em vista que a ferramenta possui duas funcionalidades, a de detecção cabeçalhos e a classificação de títulos quanto aos critérios de privacidade, os testes finais da ferramenta avaliaram 4 políticas de privacidade[8] que não foram utilizadas na construção da base de treinamento.

Ao longo da avaliação, a ferramenta detectou 37 critérios de avaliação de políticas de privacidade entre os títulos das políticas, 5 deles critérios pertencentes a uma classe diferente da atribuída pela ferramenta. É válido mencionar que 7 títulos que deveriam ter sido classificados como critérios não foram.

Considerando o total de critérios entre os títulos como 44, sendo 32 o número de critérios corretamente classificados pela ferramenta, a taxa de eficácia da mesma foi de 72.7%.

[8] Disponível no Apêndice C

4. RESULTADOS

Nesta seção, são apresentados os resultados obtidos na construção da ferramenta, os passos para a utilização da mesma e comparação com ferramentas e trabalhos já existentes.

4.1 Ferramenta

A ferramenta é composta por quatro páginas distintas:

- Tela inicial
- Tela de resultados
- Critério de avaliação de políticas de privacidade
- Sobre a ferramenta

Para exemplificar o uso da ferramenta, a política de privacidade do empresa Magazine Luiza[9] (https://especiais.magazineluiza.com.br/politica-de-privacidade/) foi utilizada. A figura 5 ilustra parcialmente o texto da política da Magazine Luiza

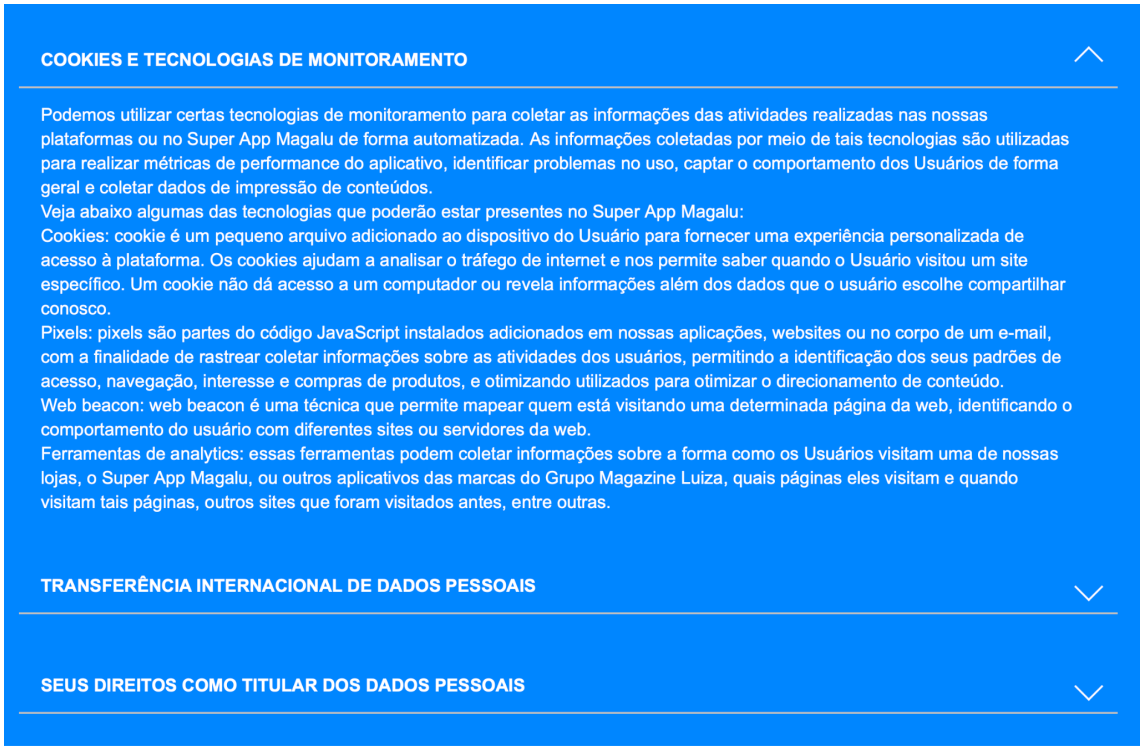


Figura 5 - Trecho da política de privacidade do Magazine Luiza.
Fonte: Magazine Luiza, 2023.

[9] Magazine Luiza: https://www.magazineluiza.com.br/

Na tela inicial da ferramenta o usuário pode anexar a política de privacidade para ser avaliada, como é demonstrado na Figura 6.

Figura 6 - Tela inicial da ferramenta com o arquivo PDF selecionado.
Fonte: Autor, 2023.

Após a submissão do arquivo de política de privacidade, o usuário é redirecionado para a página de resultados, onde os títulos detectados pela ferramenta e os critérios de avaliação de políticas de privacidade foram encontrados. As Figuras 7 e 8 detalham os resultados obtidos na avaliação da política.

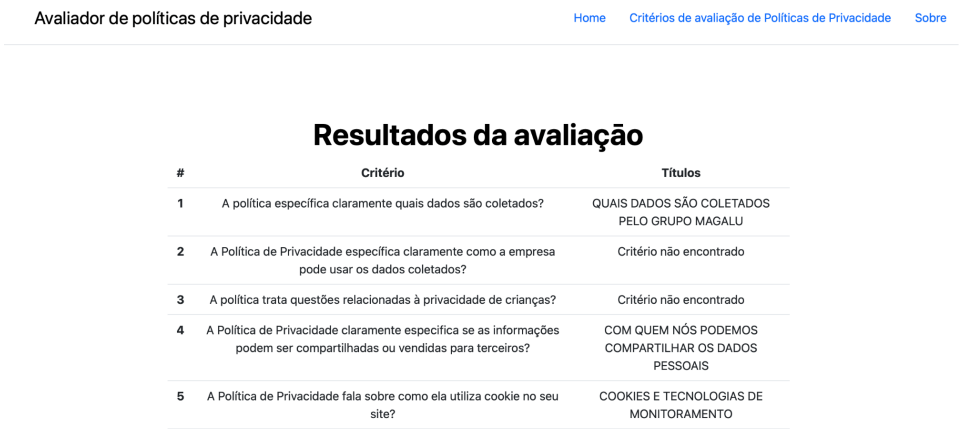


Figura 7 - Tela de resultados da ferramenta para a política de privacidade selecionada.
Fonte: Autor, 2023.

Títulos detectados no arquivo

#	Título
1	QUAIS DADOS SÃO COLETADOS PELO GRUPO MAGALU
2	COMO NÓS UTILIZAMOS OS SEUS DADOS PESSOAIS
3	COM QUEM NÓS PODEMOS COMPARTILHAR OS DADOS PESSOAIS
4	ARMAZENAMENTO E SEGURANÇA DOS DADOS PESSOAIS
5	COOKIES E TECNOLOGIAS DE MONITORAMENTO
6	TRANSFERÊNCIA INTERNACIONAL DE DADOS PESSOAIS
7	SEUS DIREITOS COMO TITULAR DOS DADOS PESSOAIS
8	RETENÇÃO E EXCLUSÃO DOS SEUS DADOS PESSOAIS
9	ALTERAÇÕES DESTA POLÍTICA DE PRIVACIDADE
10	FALE CONOSCO

Figura 8 - Tela de resultados da ferramenta para a política de privacidade selecionada.

Fonte: Autor, 2023

A ferramenta também possui uma tela de informação sobre os critérios de avaliação de políticas de privacidade (Figura 9), explicando o conceito de critérios de avaliação de políticas de privacidade e enumerando os critérios utilizados na implementação da ferramenta.

Figura 9 - Tela de Critérios de avaliação de políticas de privacidade.

Fonte: Autor, 2023.

A última tela da ferramenta possui informações sobre a motivação para a criação da ferramenta, e algumas das tecnologias utilizadas no processo. A figura abaixo representa a tela sobre a ferramenta.

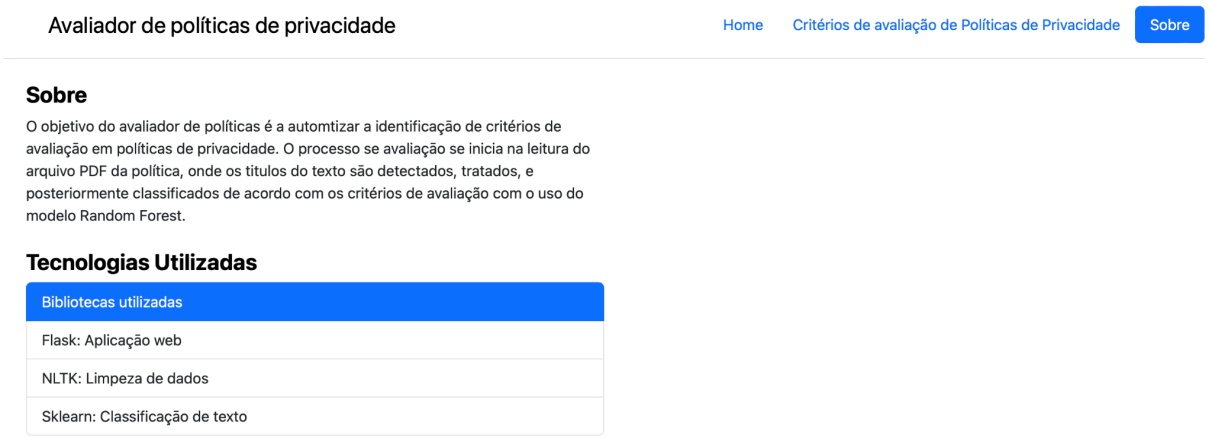


Figura 10 - Tela sobre a ferramenta.
Fonte: Autor, 2023.

O código fonte da ferramenta, o modelo de apredizagem e a base de dados utilizada no treinamento dos modelos estão disponibilizados no repositório: https://github.com/GabrielCortizo/avaliador-de-politicas-de-privacidade

4.2 Comparação com ferramentas existentes

A Tabela 2 faz o comparativo entre a ferramenta desenvolvida nesse trabalho e a ferramenta desenvolvida no trabalho de Rodrigo Ferreira, utilizando processamento de linguagem natural para avaliar Políticas de Privacidade [12], tendo em vista que ambas ferramentas propõem a detecção de critérios de avaliação de políticas de privacidade de maneira automatizada.

Trabalho	Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade [12]	Trabalho atual
Ano	2022	2023

Paradigmas utilizados	Processamento de linguagem natural.	Processamento de linguagem natural e modelo de aprendizado supervisionado Support Vector Machine.
Número de políticas utilizadas na construção da ferramenta	7	23
Número de critérios avaliados	14	14
Eficácia em um conjunto de políticas não utilizados na construção da ferramenta(%)	48%	72.7%

Tabela 2 - Comparação entre o trabalho proposto e trabalhos relacionados.

Tendo em vista que o desempenho de um algoritmo de aprendizado de máquina está diretamente relacionado ao número de exemplos e a qualidade de dados utilizados para treinar o modelo [39], um número maior de políticas foi utilizado no trabalho atual.

Quanto à avaliação da eficácia, tanto o *"Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade"* [12] como o trabalho atual utilizaram o cálculo da proporção de previsões corretas em relação ao número total de previsões, divergindo quanto ao número de políticas utilizadas no teste que foram respectivamente de duas e quatro políticas.

4.3 Limitações

Essa seção enumera algumas das limitações da ferramenta desenvolvida no trabalho atual.

O suporte da ferramenta limitado à políticas de privacidade em formato PDF afeta diretamente a experiência do usuário, tendo em vista que as políticas de privacidade salvo exceções são disponibilizadas em páginas web, sendo necessário que o usuário faça o download do conteúdo da página web para cada política e que o converta em formato PDF para utilizar a ferramenta.

A ferramenta também assume que o texto da política de privacidade em avaliação seja dividido em títulos, e que os títulos possuam uma formatação de texto diferente dos parágrafos, o que inviabilizaria a identificação de critérios caso documento de política não seguisse essas duas limitações.

5 CONCLUSÕES E TRABALHOS FUTUROS

O trabalho propôs a implementação de uma ferramenta utilizando aprendizado de máquina na avaliação da qualidade de políticas de privacidade por meio de critérios de avaliação de propostas por Augusto Terra [14]. É válido mencionar que a ferramenta desenvolvida no trabalho atual obteve eficácia na identificação de critérios de 72.7% na avaliação de políticas não utilizadas no treinamento, sugerindo que a utilização de aprendizagem de máquina melhorou significativamente a eficácia na identificação de critérios em relação ao ferramenta desenvolvida no trabalho *"Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade"* [12].

O motivo no desenvolvimento da ferramenta é a avaliação automatizada de uma política de privacidade, tendo em vista que os textos das mesmas podem ser longos e de difícil compreensão [13].

A ferramenta foi desenvolvida para a identificação de critérios de avaliação no texto de uma Política de Privacidade. Posteriormente disponibilizada online junto com o seu código fonte. O código e a base de dados utilizada no treinamento e na avaliação do modelo também foram disponibilizados.

Os principais desafios encontrados ao longo do desenvolvimento se devem a falta de padronização no documento de políticas de privacidade, os diferentes tipos de formato, alguns possuindo imagens e tabelas dificultam a extração de campos do texto, como títulos e parágrafos.

Outro ponto de dificuldade foi a quantidade de amostras por critérios na base de dados, tendo em vista a necessidade de avaliação manual e que alguns critérios não frequentes em políticas limitam a eficácia da ferramenta na classificação dos critérios de privacidade.

Algumas sugestões de trabalhos futuros envolvem:

- Aumento do número de critérios de avaliação de políticas de privacidade utilizados na ferramenta a fim de melhorar a eficácia da ferramenta.
- Teste de usabilidade da ferramenta.
- Testes de eficácia da ferramenta.

- A possibilidade de leitura de textos de políticas a partir de links *URL* e a habilitação da ferramenta na utilização não apenas dos títulos das políticas mas também do texto completo.

6 REFERÊNCIAS

[1] Danilo Doneda (2011). **A proteção dos dados pessoais como um direito fundamental.** Universidade do Oeste de Santa Catarina. Disponível em: https://periodicos.unoesc.edu.br/espacojuridico/article/view/1315/658. Acesso em 20/04/2023.

[2] OECD. (1980). **OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data**. OECD. Disponível em: www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html. Acesso em 20/04/2023

[3] Warren, Samuel; Brandeis, Louis (1890)**: The Right to Privacy.** Harvard Law Review. IV (5): pp. 193–220

[4] Parlamento Europeu e do Conselho da União Europeia(2016), **Regulamento (UE) 2016/679.** Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32016R0679. Acesso em 20/04/2023.

[5] Assembleia Geral das Nações Unidas. (1948). **Declaração Universal dos Direitos Humanos. Artigo 12**. Assembleia Geral das Nações Unidas. Disponível em: https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/por.pdf. Acesso em 20/04/2023.

[6] Solove, D. J. (2011). **Information Privacy Law** (2nd ed.). New York: Wolters Kluwer Law & Business.

[7] Hastie, T., Tibshirani, R., & Friedman, J. (2001). **The Elements of Statistical Learning: Data Mining, Inference, and Prediction** (1st ed.). New York: Springer.

[8] Khalid, H., & Mehmood, A. (2019). **Sentiment Analysis of Product Reviews using Machine Learning Techniques**. In 2019 2nd International Conference on Intelligent Sustainable Systems (ICISS) (pp. 324-327). IEEE. DOI: 10.1109/ISS1.2019.8713536.

[9] Carrascosa, C., Castillo, G., & Martínez, J. M. (2016). **A survey of email spam filtering techniques. Journal of Network and Computer Applications**, 71, pp. 19-30. DOI: 10.1016/j.jnca.2016.05.006.

[10] Pak, S., & Paroubek, C. (2010). **Twitter as a Corpus for Sentiment Analysis and Opinion Mining**. In Proceedings of the Seventh Conference on International

Language Resources and Evaluation (LREC'10) (pp. 1320-1326). European Language Resources Association (ELRA).

[11] Cortes, C. (1995). **Support-vector networks. Machine Learning**, 20(3), pp. 273-297. DOI: 10.1007/BF00994018.

[12] Rodrigo Ferreira Oliveira de Paula(2022). **Utilizando processamento de linguagem natural para avaliar Políticas de Privacidade.** UNIVERSIDADE FEDERAL DE PERNAMBUCO

[13] Singh, Ravi Inder, Manasa Sumeeth, and James Miller. "**Evaluating the readability of privacy policies in mobile environments**." International Journal of Mobile Human Computer Interaction (IJMHCI) 3.1 (2011), pp. 55-78.

[14] Terra, Augusto Henriques.(2021) **Catálogo de critérios para avaliação de políticas de privacidade**. UNIVERSIDADE FEDERAL DE PERNAMBUCO.

[15] Mohanty, S. S., Patra, M. R., & Panda, G. (2014). **Emotion recognition from speech signal using support vector machine**. International Journal of Computer Applications, 101(10), pp. 26-31.

[16] Almazaydeh, L., Rantz, M. J., Skubic, M., Miller, S. J., & Pakhomov, S. (2015). **Predicting sleep states in the elderly using wireless sensors and support vector machines**. Journal of Medical Systems, 39(8), 91. DOI: 10.1007/s10916-015-0292-3.

[17] Sinha, A., Sharma, A., & Jha, S. (2018). **Identifying PTSD Symptoms using Twitter Data: A Microblogging Platform-based Machine Learning Approach**. In 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1097-1101). IEEE.

[18] Zimmerman, J., Toubia, O., & Schwartz, H. A. (2015). **The readability of privacy policies**. Computers in Human Behavior, 52, pp. 479-487.

[19] **Yahoo Says 1 Billion User Accounts Were Hacked**. The New York Times, 14/12/2016, Disponível em: https://www.nytimes.com/2016/12/14/technology/yahoo-hack.html#:~:text=In%20the%20hacking%20disclosed%20Wednesday,attack%20occurred%20in%20August%202013. Acesso em 20/04/2023.

[20] Bird, S., Klein, E., & Loper, E. (2009). **Natural Language Processing with Python** (1st ed.). O'Reilly Media, Inc.

[21] Li, F., Li, Q., Yang, W., & Zhang, S. (2021). **Automatic Privacy Policy Analysis and Its Legal Implications**. IEEE Transactions on Information Forensics and Security, 16, 1267-1278. DOI: 10.1109/TIFS.2020.3037241.

[22] Solove, D. J. (2015). **Understanding Privacy** (1st ed.). Harvard University Press.

[23] Herold, R. (2015). **The Privacy Papers: Managing Technology, Consumer, Employee, and Legislative Actions** (1st ed.). Auerbach Publications.

[24] Khan, R., & Baharudin, B. (2016). **A comparative study of classification algorithms for sentiment analysis on Twitter**. Journal of Telecommunication, Electronic and Computer Engineering, 8(8), 81-85. DOI: 10.15282/jtece.08.2016.14.

[25] Tharwat, A., Gaber, T., & Ibrahim, A. (2019). **Classifying texts with support vector machines: An in-depth guide for practitioners**. Knowledge-Based Systems, 163, 473-496. DOI: 10.1016/j.knosys.2018.09.025.

[26] Vergara, J. R., & Flores, J. L. (2018). **Comparing machine learning algorithms for authorship attribution of Spanish texts**. Journal of Information Science, 44(6), 778-788. DOI: 10.1177/0165551517715746.

[27] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). **An Introduction to Statistical Learning: with Applications in R** (Springer Texts in Statistics, Version 1.1). Springer.

[28] Géron, A. (2019). **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems** (2nd ed.). O'Reilly Media.

[29] Agarwal, A., Biadsy, F., & Martin, J. H. (2011). **Sentiment analysis of twitter data. In Proceedings of the workshop on languages in social media**. Association for Computational Linguistics.

[30] Alpaydin, E. (2010). **Introduction to Machine Learning** (2nd ed.). Cambridge, MA: MIT Press.

[31] Chen, T., & Huang, C. (2012). **A study on spam filtering using Naive Bayes algorithm**. Procedia Computer Science, 15, pp. 191-198.

[32] Silva, T., & De Lima, D. F. (2014). **Sentiment Analysis of Twitter Data using a Naïve Bayes Classifier**. Journal of Information and Data Management, 5(2), pp. 107-118.

[33] Han, J., Kamber, M., & Pei, J. (2011). **Data Mining: Concepts and Techniques** (3rd ed.). Morgan Kaufmann.

[34] Alpaydin, E. (2010). **Introduction to Machine Learning** (2nd ed.). Cambridge, MA: MIT Press.

[35] Chang, M. W., & Zhu, W. Y. (2013). **Using Random Forest to Learn Imbalanced Data for Text Classification**. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013). IEEE.

[36] Manning, C. D., & Schütze, H. (1999). **Foundations of Statistical Natural Language Processing**. MIT Press. (1st ed.).

[37] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N. & Polosukhin, I. (2017). **Attention is all you need**. In Advances in neural information processing systems (pp. 5998-6008).

[38] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). **BERT: Pre-training of deep bidirectional transformers for language understanding**. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019).

[39] Chollet, F. (2018). **Deep Learning with Python** (1st ed.). Shelter Island, NY: Manning Publications.

[40] Bishop, C. M. (2006). **Pattern recognition and machine learning**. Springer.

APÊNDICE A - Lista de Critérios de Avaliação de Política de Privacidade utilizados na base de dados para o treinamento dos modelos

Critério	**Descrição**	**Número de amostras na base de treinamento**
A política específica claramente quais dados são coletados?	É importante que a Política de Privacidade detalhe claramente quais dados serão coletados pela aplicação. Os dados coletados se dividem em categorias bem definidas e a Política deve indicar essas áreas.	28
A Política de Privacidade específica claramente como a empresa pode usar os dados coletados?	A Política deve indicar qual o propósito da coleta de informações dos usuários. É necessário afirmar, por exemplo, se os dados estão sendo coletados para contactar o usuário, melhorar os serviços fornecidos, análise e monitoramento durante o uso da aplicação, personalizar a experiência, publicidade direcionada, entre outras ações.	29
A política trata questões relacionadas à privacidade de crianças?	É necessário que a Política explique claramente como se dá questões relacionadas à privacidade com crianças que acessam a aplicação.	11
A Política específica claramente como os dados são coletados?	A Política precisa expressar com clareza quais ferramentas a aplicação utiliza para coletar dados.	7

A Política de Privacidade claramente especifica se as informações podem ser compartilhadas ou vendidas para terceiros?	Caso envolva terceiros, é necessário descrever que tipo de informações são compartilhadas, quem são os terceiros e como os terceiros podem ser classificados, além de estar anexada a Política de Privacidade dessa empresa terceira. É necessário afirmar também caso não haja o compartilhamento com outras organizações.	25
Decisões Automatizadas	Aqui o critério verifica se a política discute se existem recursos tecnlógicos que realizam decisões automatizadas para fim de melhorar o serviço prestado pela empresa	1
A Política de Privacidade claramente especifica quais são as medidas adotadas pela aplicação para garantir a confidencialidade, a integridade e a qualidade dos dados?	Este critério busca avaliar se a aplicação possui algum método para garantir a confidencialidade e integridade dos dados do usuário. Por exemplo, se o armazenamento dos dados é criptografado ou alguma máscara de IP é utilizada.	22
A política explica claramente o que acontece com os dados do usuário caso ele exclua a conta?	É importante que esteja descrito na política o que acontece caso o usuário se desvincule da aplicação.	0
A Política de Privacidade claramente especifica os	As leis de privacidade apresentam direitos que os usuários possuem.	20

direitos do usuário?	É uma boa prática que a política descreva esses direitos em relação a seus dados pessoais.	
A Política de Privacidade fala sobre como ela utiliza cookie no seu site?	Este critério busca avaliar se o site fala sobre os tipos de cookies utilizando pelo website	19
A Política de Privacidade claramente informa dados para contato com a empresa?	Idealmente deve haver o contato da área da empresa que trate de questões de privacidade dos dados de seus usuários.	16
A Política de Privacidade claramente específica como os dados são armazenados?	Ao informar como os dados são armazenados a empresa passa uma maior credibilidade para seus usuários.	15
A Política de Privacidade fala sobre transferir dados do usuário em nível internacional?	O critério idealmente deve falar sobre como transferir os dados do cliente para outras regiões fora do Brasil.	8
Como as alterações nas políticas são tratadas?	Após uma eventual alteração na Política de Privacidade, os usuários precisam ser informados e notificados sobre isso.	17

APÊNDICE B - Políticas de Privacidade utilizadas na construção da base de dados de treinamento dos modelos

Empresa	Link da política
Amazon	https://www.amazon.com.br/gp/help/customer/display.html?nodeId=201283950
Caixa	https://www.caixa.gov.br/privacidade/aviso-de-privacidade/Paginas/default.aspx
Tiktok	https://www.tiktok.com/legal/page/row/privacy-policy/pt-BR
Whatsapp	https://www.whatsapp.com/legal/privacy-policy/?locale=pt_BR
Twitter	https://twitter.com/pt/privacy/previous/version_15
Spotify	https://www.spotify.com/br-pt/legal/privacy-policy/
Skype	https://privacy.microsoft.com/pt-pt/privacystatement/
Nestlé	https://www.nestle.com.br/politica-de-privacidade
Mercado Livre	https://www.mercadolivre.com.br/privacidade
Magalu	https://especiais.magazineluiza.com.br/politica-de-privacidade/
Globo	https://privacidade.globo.com/privacy-policy/
Crunchyroll	https://www.crunchyroll.com/pt-br/privacy/index.html

UOL	https://sobreuol.noticias.uol.com.br/normas-de-seguranca-e-privacidade
Steam	https://store.steampowered.com/privacy_agreement/?l=portuguese
Aliexpress	https://terms.alicdn.com/legal-agreement/terms/suit_bu1_aliexpress/suit_bu1_aliexpress201912031136_53144.html
Shopee	https://help.shopee.com.br/portal/article/77068
OLX	https://ajuda.olx.com.br/s/article/politica-de-privacidade
Bet365	https://help.bet365.com/pt/privacy-policy
Terra	https://www.terra.com.br/privacidade/
TIM	https://www.tim.com.br/sobre-a-tim/institucional/seguranca/politica-de-privacidade
Abril	https://abril.com.br/politica-de-privacidade/
IG	https://institucional.ig.com.br/2020-11-05/politica-de-privacidade-ig.html
Correios	https://www.correios.com.br/falecomoscorreios/politica-de-privacidade-e-cookies

APÊNDICE C - Políticas de Privacidade utilizadas no teste da ferramenta

Empresa	**Link da política**
Gov.br	https://www.gov.br/governodigital/pt-br/governanca-de-dados/conecta-gov.br/termos-de-uso-e-de-politica-de-privacidade/termos-de-uso-versao-1.1
Epic Games	https://www.epicgames.com/site/pt-BR/privacypolicy
Estadão	https://www.estadao.com.br/dados/io/contratos/politica-de-privacidade-v10.pdf
Casas Bahia	https://casasbahia.negociafacil.com.br/privacidade/politica-de-privacidade