UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

FILIPE RAMALHO DA SILVA

CRIPTOGRAFIA HOMOMÓRFICA NO APRENDIZADO DE MÁQUINA

CAMPINA GRANDE - PB

2024

FILIPE RAMALHO DA SILVA

CRIPTOGRAFIA HOMOMÓRFICA NO APRENDIZADO DE MÁQUINA

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador : Eanes T. Pereira

CAMPINA GRANDE - PB

2024

FILIPE RAMALHO DA SILVA

CRIPTOGRAFIA HOMOMÓRFICA NO APRENDIZADO DE MÁQUINA

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA:

Eanes T. Pereira

Orientador – UASC/CEEI/UFCG

Andrey Elisio Brito Monteiro

Examinador – UASC/CEEI/UFCG

Francisco Vilar Brasileiro

Professor da Disciplina TCC – UASC/CEEI/UFCG

Trabalho aprovado em: 17 de Maio de 2024.

CAMPINA GRANDE - PB

RESUMO

A criptografia homomórfica representa uma mudança de paradigma no âmbito do processamento seguro de dados, permitindo cálculos em dados criptografados sem a necessidade de descriptografia. Essa propriedade promete enormes avanços para aprimorar a privacidade e segurança em diversos domínios, incluindo computação em nuvem, saúde, finanças e também no aprendizado de máquina. Este TCC adentra nos fundamentos da criptografia homomórfica no aprendizado de máquina, elucidando suas bases matemáticas e explorando suas aplicações práticas. Através de uma revisão da literatura existente e metodologias, esta pesquisa avalia os pontos fortes, fraquezas e desafios potenciais associados. Além disso, investiga as implicações de desempenho e sobrecargas computacionais incorridas por diferentes esquemas de criptografía homomórfica. O estudo também examina casos de uso do mundo real e cenários de implementação para avaliar a viabilidade e eficácia da criptografia homomórfica para processamento seguro de dados e tecnologias de preservação de privacidade.

HOMOMORPHIC ENCRYPTION IN MACHINE LEARNING

ABSTRACT

Homomorphic encryption represents a paradigm shift in the realm of secure data processing, allowing computations on encrypted data without the need for decryption. This capability promises significant advancements in enhancing privacy and security across various domains, including cloud computing, healthcare, finance, and also in machine learning. This Final paper delves into the fundamentals of homomorphic encryption in machine learning, elucidating its mathematical underpinnings and exploring its practical applications. Through a comprehensive review of existing literature and methodologies, this research evaluates the strengths, weaknesses, and potential challenges associated with it. Additionally, it investigates the performance implications and computational overhead incurred by different homomorphic encryption schemes. The study also examines real-world use cases and implementation scenarios to assess the viability and effectiveness of homomorphic encryption for secure data processing and privacy-preserving technologies.

Criptografia homomórfica no aprendizado de máquina

Filipe Ramalho da Silva
Universidade Federal de Campina Grande
filipe.silva@ccc.ufcg.edu.br

Eanes T. Pereira
Universidade Federal de Campina Grande
eanes@computacao.ufcg.edu.br

ABSTRACT

Homomorphic encryption represents a paradigm shift in the realm of secure data processing, allowing computations on encrypted data without the need for decryption. This capability promises significant advancements in enhancing privacy and security across various domains, including cloud computing, healthcare, finance, and also in machine learning. This Final paper delves into the fundamentals of homomorphic encryption in machine learning, elucidating its mathematical underpinnings and exploring its practical applications. Through a review of existing literature and methodologies, this research evaluates the strengths, weaknesses, and potential challenges associated with it. Additionally, it investigates the performance implications and computational overhead incurred by different homomorphic encryption schemes. The study also examines real-world use cases and implementation scenarios to assess the viability and effectiveness of homomorphic encryption for secure data processing and privacy-preserving technologies.

RESUMO

A criptografia homomórfica representa uma quebra de paradigma no âmbito do processamento seguro de dados, permitindo cálculos em dados criptografados sem a necessidade de descriptografia. Essa capacidade promete enormes avanços para aprimorar a privacidade e segurança em diversos domínios, incluindo computação em nuvem, saúde, finanças e também no aprendizado de máquina. Este TCC adentra nos fundamentos da criptografia homomórfica no aprendizado de máquina, elucidando suas bases matemáticas e explorando suas aplicações práticas. Através de uma revisão da literatura existente e metodologias, esta pesquisa avalia os pontos fortes, fraquezas e desafios potenciais associados. Além disso, investiga as implicações de desempenho e sobrecargas computacionais incorridas por diferentes esquemas de criptografia homomórfica. O estudo também examina casos de uso do mundo real e cenários de implementação para avaliar a viabilidade e eficácia da criptografia homomórfica para processamento seguro de dados e tecnologias de preservação de privacidade.

Keywords

Criptografia, criptografia homomórfica, aprendizado de máquina.

1. INTRODUÇÃO

A privacidade e segurança de dados são cruciais no mundo digital contemporâneo, onde a crescente interconexão e digitalização tornam as informações pessoais cada vez mais vulneráveis. Esses temas têm impacto em diversas esferas, desde a proteção dos direitos individuais até a segurança nacional, e são fundamentais para preservar a autonomia, dignidade e liberdade das pessoas. É essencial que se adotem medidas proativas, como práticas robustas de segurança cibernética.

Tradicionalmente, o treinamento de modelos de aprendizado de máquina requer acesso direto aos dados brutos, o que pode levantar preocupações sobre privacidade e segurança, especialmente em setores sensíveis, como saúde e finanças. No entanto, a criptografia homomórfica pode permitir que os dados permaneçam criptografados durante todo o processo de análise, inclusive durante o treinamento do modelo.

Essa combinação de aprendizado de máquina e criptografia homomórfica oferece um potencial significativo para avanços em áreas como medicina personalizada, segurança cibernética e análise financeira. Ao preservar a privacidade dos dados, as organizações podem colher os benefícios do aprendizado de máquina sem comprometer a confidencialidade dos dados dos usuários. No entanto, é importante notar que a criptografia homomórfica ainda está em estágios iniciais de desenvolvimento e pode enfrentar desafios de desempenho e escalabilidade. No entanto, com o avanço da tecnologia, espera-se que essa abordagem inovadora continue a crescer e se tornar uma ferramenta poderosa para análise de dados seguros e privados.

O treinamento com dados encriptados pode ser aplicado em uma ampla variedade de conjuntos de dados sensíveis, em diversos setores e áreas de aplicação. Aqui estão alguns exemplos de tipos de dados sensíveis nos quais essa abordagem pode ser útil:

Dados de Saúde: informações médicas pessoais, como histórico de saúde, registros de pacientes, resultados de exames laboratoriais e imagens médicas, podem ser criptografados durante o treinamento de modelos de aprendizado de máquina. Isso permite a análise segura desses dados para desenvolver modelos de diagnóstico médico, previsão de doenças e personalização de tratamentos, sem comprometer a privacidade dos pacientes.

Dados Financeiros: dados financeiros confidenciais, incluindo transações bancárias, histórico de crédito, informações fiscais e investimentos, podem ser protegidos por meio de criptografia durante o treinamento de modelos de detecção de fraudes, previsão de riscos financeiros e recomendações de investimento. Isso garante a confidencialidade dos dados financeiros dos clientes e reduz o risco de fraudes e violações de segurança.

Dados Pessoais Identificáveis (PII): informações pessoais identificáveis, como nomes, endereços, números de identificação pessoal (como CPF ou SSN), podem ser encriptadas durante o treinamento de modelos de análise de clientes, segmentação de mercado e personalização de serviços. Isso protege a privacidade dos indivíduos e cumpre regulamentações de proteção de dados, como o GDPR e o CCPA.

Dados Governamentais: dados governamentais sensíveis, como registros de votação, informações de segurança nacional e registros criminais, podem ser protegidos por meio de criptografia durante o treinamento de modelos de análise de segurança, previsão de tendências criminais e detecção de atividades suspeitas. Isso assegura a integridade e a confidencialidade dos dados do governo.

Dados de Pesquisa Científica: dados de pesquisa sensíveis, incluindo resultados de experimentos, dados genéticos e informações sobre ensaios clínicos, podem ser criptografados durante o treinamento de modelos de análise de dados científicos e médicos. Isso permite a colaboração segura entre pesquisadores e instituições, mantendo a confidencialidade dos dados de pesquisa.

Esses são apenas alguns exemplos de dados sensíveis nos quais o treinamento com dados encriptados pode ser aplicado. Em geral, essa abordagem pode ser útil em qualquer cenário no qual a privacidade e a segurança dos dados sejam preocupações fundamentais.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são introduzidos alguns conceitos e tecnologias importantes para entender o contexto do trabalho apresentado.

2.1 Criptografia

A criptografia pode ser entendida como um conjunto de métodos e técnicas para cifrar ou codificar informações legíveis através de um algoritmo, convertendo um texto original em um texto ilegível, sendo possível através do processo inverso recuperar as informações originais (SIMON, 1999)

Pode-se criptografar informações basicamente através de códigos ou de cifras. Os códigos protegem as informações trocando partes da informação por códigos predefinidos. Sendo que todas as pessoas autorizadas a ter acesso à uma determinada informação devem conhecer os códigos utilizados. As cifras são técnicas nas quais a informação é cifrada através da transposição e/ou substituição das letras da mensagem original. Assim, as pessoas autorizadas podem ter acesso às informações originais conhecendo o processo de cifragem.

Figura 1 - Esquema geral de cifragem simétrica

Existem dois tipos principais de criptografia: simétrica e assimétrica. Em que a criptografia simétrica usa apenas uma chave para criptografar e descriptografar, e a assimétrica, proposta por Whitfield Diffie e Martin Hellman publicam,que, em 1976, o artigo ”New Directions in Cryptography”[2], onde introduzem a ideia de criptografia de chave pública, em que apenas a chave privada poderá descriptografar a mensagem, mas qualquer pessoa com a chave pública poderá encriptar-la

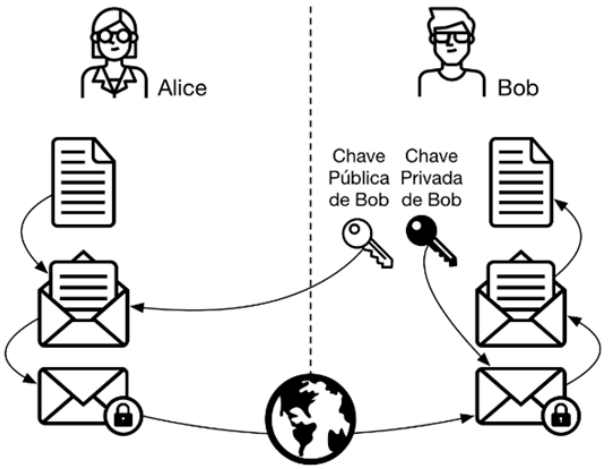

Figura 2 - Diagrama de funcionamento da criptografia assimétrica[3]

A criptografia desempenha um papel crucial na proteção da privacidade e na segurança dos dados em um mundo digital cada vez mais interconectado. Além de garantir a confidencialidade das informações, ela também ajuda a prevenir fraudes, ataques cibernéticos e o acesso não autorizado a sistemas e redes. Com o avanço da tecnologia, novas técnicas e algoritmos de criptografia continuam sendo desenvolvidos para enfrentar os desafios emergentes de segurança e privacidade, mantendo nossas comunicações e dados protegidos.

2.2 Criptografia Homomórfica

A criptografia homomórfica é um sistema de criptografia que permite realizar operações matemáticas nos dados criptografados sem a necessidade de descriptografá-los primeiro. Isso significa que você pode executar operações como adição, multiplicação e outras sobre os dados criptografados e obter um resultado criptografado que, quando descriptografado, corresponde ao resultado da operação aplicada aos dados originais.

1. Adição Homomórfica:

Se E(a) representa a criptografia de (a) e E(b) representa a criptografia de (b), então a adição homomórfica nos dados criptografados poderá ser expressa como:

$$E(a) + E(b) = E(a + b)$$

2. Multiplicação Homomórfica:

Similarmente, se quisermos multiplicar (a) e (b) de forma homomórfica, podemos usar a seguinte expressão

$$E(a) \times E(b) = E(a \times b)$$

Essas propriedades permitem que operações sejam realizadas nos dados criptografados sem revelar a informação subjacente. No entanto, é importante notar que existem diferentes tipos de criptografia homomórfica, como criptografia totalmente homomórfica (FHE) e criptografia parcialmente homomórfica (PHE), cada uma com suas próprias características e aplicações.

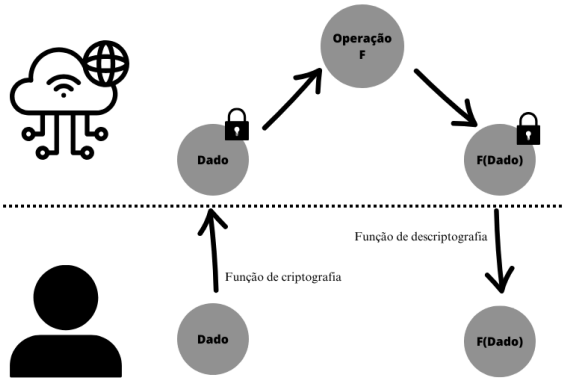

Figura 3 - Diagrama do Funcionamento da criptografia homomórfica

No entanto, é importante ressaltar que, assim como qualquer outra técnica de criptografia, a criptografia homomórfica não é imune a possíveis vulnerabilidades e ataques. Uma vez que a segurança da criptografia homomórfica é baseada na dificuldade de resolver certos problemas matemáticos, como o problema do logaritmo discreto ou a fatoração de inteiros. Se um algoritmo de criptografia subjacente for quebrado ou se novas técnicas criptográficas forem desenvolvidas, os esquemas de criptografia homomórfica podem se tornar vulneráveis.[4]

Ataques de canal lateral exploram informações acessíveis durante o processo de criptografia ou descriptografia, como consumo de energia, tempo de execução ou radiação eletromagnética. Essas informações podem ser utilizadas para inferir dados secretos ou chaves criptográficas, comprometendo assim a segurança da criptografia homomórfica.

Mesmo que as operações sejam realizadas em dados criptografados, em alguns casos, é possível inferir informações sobre os dados originais ou as operações realizadas com base em padrões nos dados criptografados. Isso pode representar uma ameaça à privacidade e à segurança dos dados.

Por isso, é fundamental que os sistemas que a utilizam sejam projetados e implementados de forma cuidadosa, levando em consideração as melhores práticas de segurança da informação. Apesar dos desafios e limitações atuais, a criptografia homomórfica tem o potencial de revolucionar a maneira como lidamos com a privacidade e segurança dos dados em um mundo cada vez mais digitalizado e interconectado.

2.3 Aprendizado de máquina

O aprendizado de máquina é um campo da inteligência artificial que se concentra no desenvolvimento de algoritmos e modelos que permitem aos computadores aprenderem a partir de dados e fazer previsões ou tomar decisões sem serem explicitamente programados para isso. Esse processo de aprendizado é realizado através da análise de padrões nos dados, permitindo que os sistemas automatizem tarefas e aprimorem seu desempenho ao longo do tempo. O aprendizado de máquina é aplicado em uma ampla gama de áreas, incluindo reconhecimento de padrões, processamento de linguagem natural, visão computacional, medicina, finanças e muitas outras.

Existem diversos tipos de aprendizado de máquina, sendo os principais o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado por reforço. No aprendizado supervisionado, o algoritmo é treinado com um conjunto de dados que contém pares de entrada e saída esperada, permitindo que o modelo faça previsões ou classificações em novos dados. No aprendizado não supervisionado, o algoritmo é treinado apenas com os dados de entrada, e o objetivo é encontrar padrões ou estruturas nos dados sem a necessidade de rótulos prévios. Já no aprendizado por reforço, o sistema aprende através de interações com um ambiente, recebendo recompensas ou punições conforme suas ações.

O aprendizado de máquina tem revolucionado diversos setores, possibilitando a automação de tarefas complexas, a personalização de serviços e produtos, a otimização de processos e a descoberta de insights valiosos a partir de grandes volumes de dados. No entanto, é importante ressaltar que o sucesso do aprendizado de máquina depende não apenas da qualidade dos algoritmos e modelos, mas também da disponibilidade de dados de qualidade, da capacidade de interpretação e explicação dos resultados, e da consideração de questões éticas e sociais relacionadas ao seu uso[7].

2.4 Criptografía Homomórfica Total

O FHE, ou Fully Homomorphic Encryption (Criptografia Homomórfica Total)[6], é uma forma avançada de criptografia homomórfica que permite realizar operações de soma, subtração e multiplicação nos dados criptografados, sem a necessidade de descriptografá-los. Esta capacidade de realizar operações arbitrariamente complexas em dados criptografados é o que distingue o FHE de outras formas de criptografia homomórfica. Ou seja, $E(a) \times E(b) = E(a \times b)$ é verdade, bem como $E(a) + E(b) = E(a + b)$.

No entanto, é importante notar que o FHE também apresenta desafios significativos em termos de desempenho e eficiência computacional. As operações realizadas em dados criptografados pelo FHE geralmente são muito mais lentas do que as operações equivalentes em dados não criptografados. Além disso, os algoritmos de FHE podem exigir uma quantidade considerável de recursos computacionais para executar, o que pode limitar sua aplicabilidade em certos cenários.

2.5 Dataset

Um *dataset*, ou conjunto de dados, é uma coleção estruturada de informações, geralmente apresentada em formato digital, que contém observações, variáveis e seus valores correspondentes. Em termos simples, um *dataset* é um conjunto de dados que representa uma parte do mundo real ou de um sistema, organizado de forma que seja possível analisar, manipular e extrair informações dele. Cada linha de um *dataset* pode representar uma entrada individual, como um exemplo, uma observação ou uma transação, enquanto cada coluna representa uma característica específica ou uma variável associada a essas entradas.

Os *datasets* podem variar amplamente em tamanho, complexidade e formato, dependendo do contexto em que são utilizados. Eles são comumente utilizados em diversas áreas, incluindo ciência de dados, aprendizado de máquina, pesquisa acadêmica, análise de negócios e muitos outros campos. A qualidade e a relevância dos dados em um *dataset* são fundamentais para garantir que as análises e os modelos construídos a partir deles sejam precisos e úteis. Portanto, a seleção, a preparação e o gerenciamento cuidadoso de datasets são etapas críticas em qualquer projeto de análise de dados ou de desenvolvimento de modelos de aprendizado de máquina.

2.6 MÉTRICAS DE MODELOS

As métricas utilizadas para avaliar modelos de machine learning desempenham um papel crucial na compreensão do quão bem um modelo está se saindo em uma tarefa específica. Aqui estão algumas das métricas mais comuns:

Acurácia: É uma medida simples que calcula a proporção de previsões corretas em relação ao total de previsões feitas pelo modelo. É adequada quando as classes estão balanceadas, ou seja, têm aproximadamente o mesmo número de amostras. No entanto, pode ser enganosa em conjuntos de dados desbalanceados [14].

Precisão : Indica a proporção de instâncias positivas corretamente classificadas em relação ao total de instâncias previstas como positivas. É útil quando o foco está na minimização de falsos positivos [14].

Recall: Calcula a proporção de instâncias positivas corretamente classificadas em relação ao total de instâncias que realmente são positivas. É importante em situações em que a detecção de todos os casos positivos é crucial, mesmo que isso aumente o número de falsos positivos [14].

F1-Score: É a média harmônica entre precisão e recall. É útil quando há um desequilíbrio entre as classes, pois dá mais peso às classes menos frequentes [14].

Matriz de Confusão: Fornece uma visão mais detalhada do desempenho do modelo, mostrando o número de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos [14].

Curva ROC: A curva ROC é uma representação gráfica da taxa de verdadeiros positivos em relação à taxa de falsos positivos para diferentes limiares de classificação [14].

2.7 SCIKIT-LEARN

Scikit-learn, frequentemente abreviado como sklearn, é uma das bibliotecas de aprendizado de máquina mais populares e amplamente utilizadas em Python. Ela oferece uma variedade de algoritmos de aprendizado supervisionado e não supervisionado, bem como ferramentas para pré-processamento de dados, seleção de modelos, avaliação de desempenho e muito mais.

Essa biblioteca é conhecida por sua facilidade de uso e sua vasta documentação, o que a torna uma escolha popular entre cientistas de dados, pesquisadores e desenvolvedores. O Scikit-learn é construído sobre outras bibliotecas Python essenciais, como NumPy, SciPy e Matplotlib, o que permite uma integração suave com o ecossistema de ferramentas de análise de dados em Python.

O Scikit-learn é uma excelente escolha para quem está começando a aprender sobre aprendizado de máquina, pois oferece uma ampla gama de algoritmos e funcionalidades essenciais em uma interface simples e consistente. Ele também é frequentemente utilizado em projetos de produção devido à sua eficiência e desempenho confiável.

3. TRABALHOS RELACIONADOS

Existem alguns trabalhos na literatura que focam nas aplicações de criptografia no aprendizado de máquina. Dentre esses trabalhos, podemos destacar:

- ***Homomorphic Encryption for Machine Learning in Medicine and Bioinformatics*** [8] que mostra como a criptografia homomórfica pode ser útil no machine learning para aplicações médicas
- ***Homomorphic Encryption for Secure Computation on Big Data*** [9] O uso da criptografia homomórfica na computação de nuvem e *big Data*
- ***Processing Encrypted Data Using Homomorphic Encryption*** [10] mostra como os cálculos de criptografia homomórfica podem ser úteis também no machine learning
- ***Oblivious Neural Network Computing via Homomorphic Encryption*** [11] Mostra como a computação em redes neurais pode se beneficiar da criptografia homomórfica.

4. Uso da Criptografia na ML

4.1 Descrição do problema

Um desafio fundamental no campo do aprendizado de máquina é a necessidade de preservar a privacidade dos dados durante o processamento. Em diversos contextos, especialmente em aplicações de aprendizado de máquina que envolvem dados sensíveis, é essencial assegurar a proteção da privacidade dos usuários, especialmente quando esses dados podem conter informações pessoais ou confidenciais. Tradicionalmente, ao utilizar algoritmos de aprendizado de máquina, os dados precisam ser descriptografados antes de serem processados, o que levanta preocupações com a privacidade, pois os dados podem ficar expostos a riscos de segurança durante esse processo de descriptografia.

No entanto, ao criptografar dados usando um método de criptografia homomórfica e utilizá-los em um modelo de aprendizado de máquina já treinado, o resultado obtido pode ser o esperado. Entretanto, é importante notar que esse modelo foi treinado com dados não criptografados, o que não é seguro quando se trata de informações sensíveis, e pode não ser preciso caso o modelo tenha sido treinado com dados simulados. Portanto, torna-se necessário encontrar uma maneira de treinar os modelos utilizando dados já criptografados, garantindo assim a segurança e a precisão dos resultados, especialmente em ambientes nos quais a privacidade dos dados é uma preocupação primordial.

4.2 Design da Solução

Para solucionar essa problemática de uso de dados sensíveis no treinamento de modelos de ML, propõe-se adotar o FHE no treinamento de modelos de aprendizado de máquina utilizando dados criptografados. Embora essa abordagem permita o treinamento funcional do modelo com dados mantidos em sigilo, é importante ressaltar que isso acarretará um aumento no custo computacional devido à complexidade dos cálculos envolvidos na manipulação de dados criptografados.

O procedimento proposto envolve a criação de um modelo de teste a partir de um toy *dataset* conhecido, ou seja, um conjunto de dados simplificado utilizado para fins de experimentação. Em seguida, o modelo será treinado utilizando tanto o *dataset* original não criptografado quanto uma versão criptografada dos mesmos dados. Os resultados obtidos através dos dois métodos de treinamento serão comparados, assim como o desempenho dos modelos resultantes.

Essa abordagem visa avaliar a eficácia do treinamento de modelos de ML com dados criptografados, considerando não apenas a precisão dos resultados, mas também o custo computacional e outras métricas relevantes.

4.3 Prova de Conceito

A abordagem da criptografia homomórfica permite que os dados permaneçam seguros e privados, mesmo quando são processados por algoritmos de aprendizado de máquina. Isso significa que as organizações podem colaborar e compartilhar dados de maneira confiável, sem comprometer a privacidade dos usuários. Além disso, a criptografia homomórfica possibilita a realização de análises em dados sensíveis, como registros médicos ou informações financeiras, sem a necessidade de revelar informações confidenciais.

Além da criptografia homomórfica, existem outras formas de garantir segurança no machine learning como: Differential Privacy: Esta abordagem foca em adicionar ruído controlado aos dados durante o treinamento do modelo, de forma a preservar a privacidade dos indivíduos representados nesses dados. Isso impede que o modelo aprenda informações específicas sobre pontos de dados individuais[5].

Encriptação de Dados em Repouso e em Trânsito: Antes mesmo de aplicar algoritmos de Machine Learning, é fundamental garantir que os dados estejam seguros enquanto estão sendo armazenados ou transmitidos. A criptografia dos dados em repouso e em trânsito ajuda a protegê-los contra acessos não autorizados[5].

Federated Learning: Neste método, o treinamento de modelos é distribuído entre várias partes, mantendo os dados localmente em cada parte. A atualização do modelo é feita de forma colaborativa, agregando os resultados dos modelos locais sem a necessidade de compartilhar os dados brutos[5].

Multi-Party Computation (MPC): Esta técnica permite que várias partes realizem cálculos em conjunto sem revelar seus próprios inputs. Cada parte mantém seus próprios dados privados enquanto colabora para calcular uma função de interesse[5].

Como os dados permanecem criptografados durante todo o processo de análise, a FHE oferece um nível extremamente alto de privacidade. Não há exposição dos dados brutos em nenhum momento, garantindo que as informações sensíveis permaneçam confidenciais.

A também FHE permite uma ampla gama de operações a serem realizadas nos dados criptografados, incluindo adição, multiplicação, comparação e outras operações comuns em machine learning. Isso oferece uma grande flexibilidade no tipo de análises que podem ser realizadas enquanto os dados permanecem seguros.

Para demonstrar a eficácia do Fully Homomorphic Encryption (FHE) no treinamento de modelos de aprendizado de máquina, é essencial começar criando um modelo funcional usando um conjunto de dados simples. Isso nos permitirá comparar seu desempenho com um método de treinamento sem criptografia.

O conjunto de dados Iris do SKLearn é uma excelente escolha para esse propósito, pois consiste em uma pequena base de dados que descreve flores com base em características como largura e comprimento da pétala. Usando esse conjunto de dados, desenvolvemos um modelo de classificação simples para distinguir entre os dois tipos de flores com base em seus tamanhos de pétala. Aqui está a divisão alcançada:

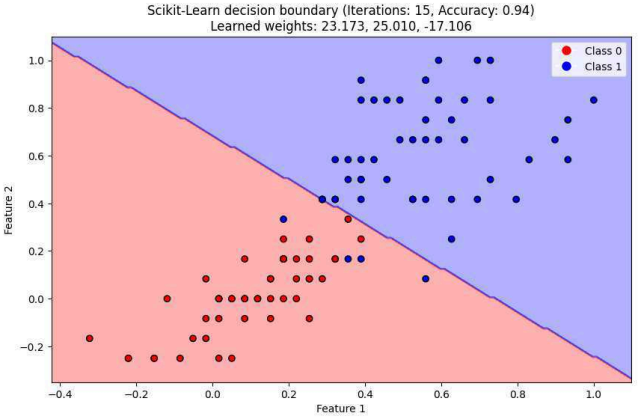

Figura 4 - Classificação do SKLearn

Foi observado que após 15 iterações, alcançamos uma Acurácia notável de 94%, precisão de 88%, recall de 89% e F1-Score de 93% . O tempo médio de execução por iteração foi de cerca de 34 milissegundos, utilizando o interpretador disponibilizado pelo Google Colab. No total, o tempo de treinamento foi de aproximadamente meio segundo, o que demonstra uma eficiência considerável no processo de treinamento do modelo.

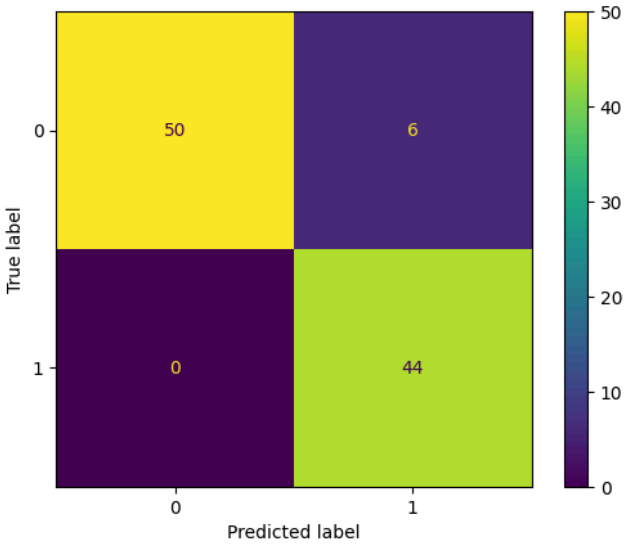

Figura 5 - Matriz de confusão do modelo treinado com dados do iris dataSet sem criptografia

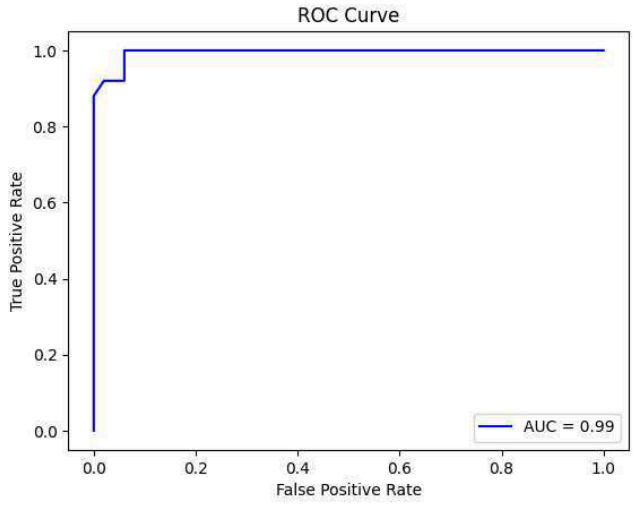

Figura 6 - Gráfico de curva ROC do modelo treinado com dados do iris dataSet sem criptografia

Com o mesmo conjunto de dados, porém criptografado utilizando o esquema BFV (Brakerski-Gentry-Vaikuntanathan) [13] de FHE , um modelo, utilizando o mesmo algoritmo SGD, foi desenvolvido para distinguir as mesmas duas flores, resultando na seguinte divisão:

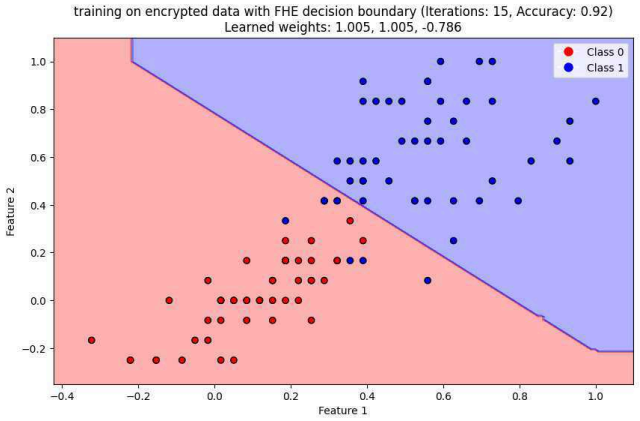

Figura 7 - Classificação com dados encriptados

Com as mesmas 15 iterações, alcançamos uma acurácia de 92%, Precisão de 92%, recall de 92% e F1-Score de 94% uma diferença leve em relação ao modelo não criptografado. No entanto, o desempenho é substancialmente prejudicado devido ao custo computacional envolvido na geração dos dados criptografados e de suas chaves correspondentes, além da complexidade adicional das operações durante o treinamento. Como resultado, o tempo total necessário para este modelo foi de 292.9 segundos, quase 600 vezes mais lento do que a versão não criptografada

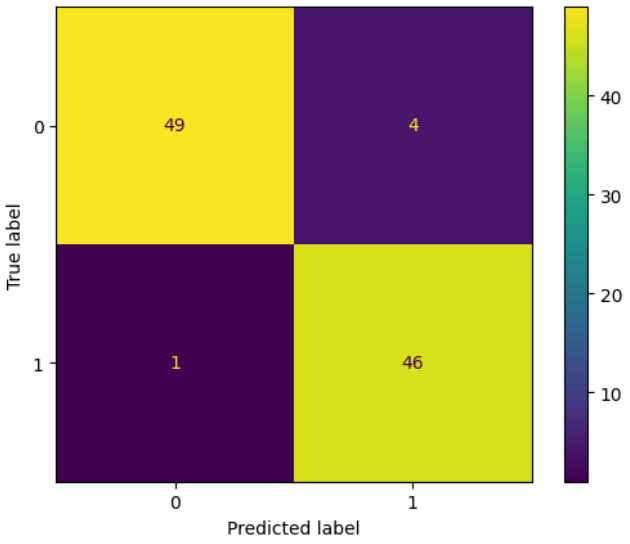

Figura 8 - Matriz de confusão do modelo treinado com dados encriptados

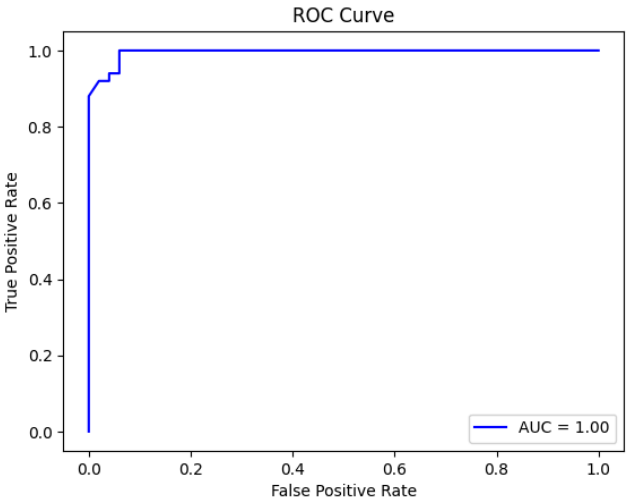

Figura 9 - Curva ROC do modelo treinado com dados encriptados

Ao analisar a matriz de confusão do modelo e a sua curva ROC, podemos perceber uma marcante semelhança com o modelo anterior. Isso sugere uma consistência nos resultados obtidos, reforçando a validade e a confiabilidade do novo modelo proposto.

4.4 Uso em base de dados mais complexas

Com o uso de *datasets* maiores e com mais variáveis, a complexidade computacional aumenta consideravelmente, resultando em um número significativamente maior de cálculos necessários para encontrar os padrões e relações entre os dados. Esse aumento na quantidade de informações a serem processadas amplifica ainda mais a diferença de tempo ao empregar dados criptografados durante o treinamento do modelo. A criptografia adiciona uma camada adicional de processamento, exigindo operações matemáticas intensivas para garantir a segurança dos dados enquanto são manipulados pelo algoritmo de aprendizado

de máquina. Assim, o tempo necessário para realizar tarefas como treinamento e validação do modelo pode se estender consideravelmente, especialmente em comparação com o processamento de dados não criptografados. Essa realidade destaca a importância de desenvolver e otimizar algoritmos de criptografia eficientes para minimizar o impacto no desempenho dos modelos de machine learning quando lidam com conjuntos de dados sensíveis.

Utilizando outro dataset do SKLearn, o *Breast_Cancer*, que relaciona cerca de 30 características para tentar prever um câncer de mama de câncer de mama. Ao importarmos o *dataset* e executar o treinamento utilizando o algoritmo de classificação SGD com os dados não criptografados obtemos uma acurácia de 94%, precisão de 91%, Recall de 86% e F1-Score de 95% após 15 épocas, com a duração total do treinamento de 1,8 seg.

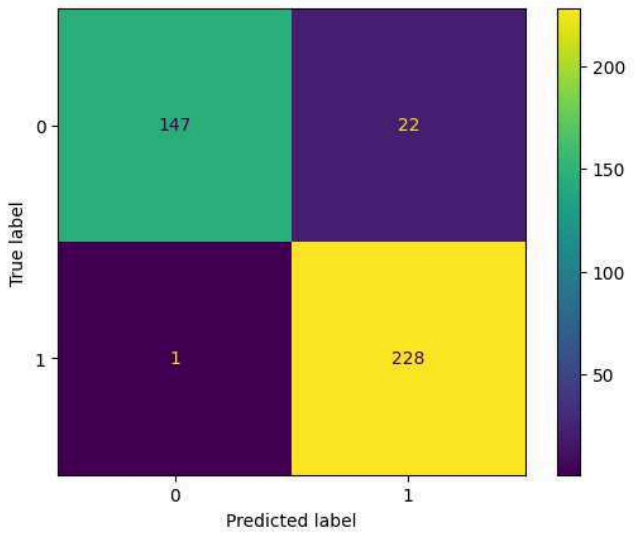

Figura 10 - Matriz de confusão do modelo treinado com o *dataset Breast_cancer*

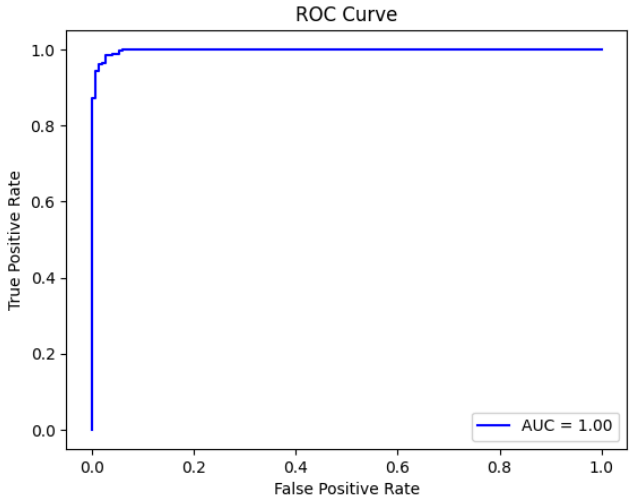

Figura 11 - Curva ROC do modelo treinado com dados do *dataset Breast_Cancer*

Com os dados criptografados, temos o tempo de criptografia e criação de chave semelhante ao exemplo anterior, porém, o tempo das iterações foi aumentado para cerca de 3 minutos e meio, tornando o processo de treinamento de 15 épocas com uma base de dados não tão grande um processo com duração maior que 1h utilizando o google collab. além disso, as métricas também tiveram uma queda significativa, com 86% de acurácia, e F1-Score de 88%.

Na matriz de confusão abaixo podemos ver que não houveram falsos positivos para o conjunto de testes, o que elevou a precisão e o recall do modelo para 100%, porém esses dados não refletem na realidade, uma vez que em outros conjuntos de teste podem ocorrer falsos positivos.

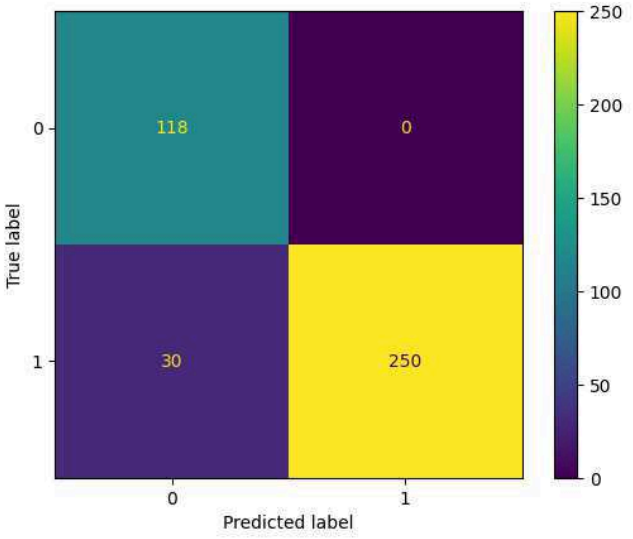

Figura 12 - Matriz de confusão do modelo treinado com o *dataset Breast_cancer* Encriptado

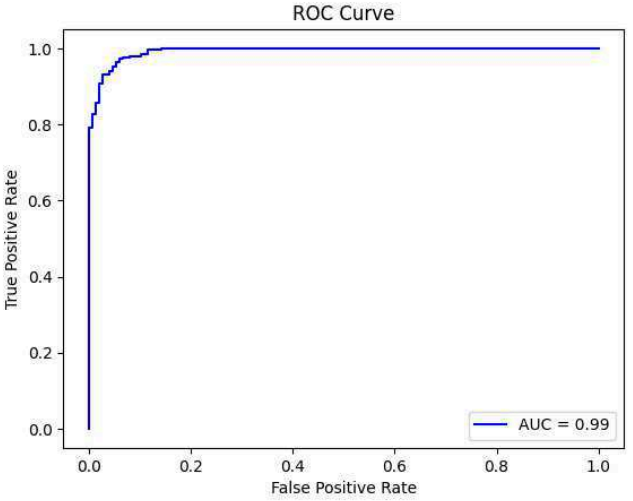

Figura 13 - Curva ROC do modelo treinado com dados do *dataset Breast_Cancer* Encriptado

Ao analisarmos a curva ROC, podemos ver claramente que o modelo perdeu acurácia, uma vez que a curva se distanciou mais de um modelo perfeito.

O uso de conjuntos de dados sensíveis, como os que envolvem informações médicas, como o tamanho do seio, entre

outros, é um exemplo mais próximo do cenário esperado para a aplicação da tecnologia de criptografia homomórfica. Nesses casos, é crucial proteger a privacidade e a confidencialidade dos dados dos pacientes, mesmo durante o processo de treinamento do modelo de aprendizado de máquina.

Essa aplicação da tecnologia de criptografia homomórfica é fundamental para garantir a conformidade com regulamentos de privacidade de dados, como o GDPR e a HIPAA [12], e para construir a confiança dos pacientes no uso de tecnologias de análise de dados em saúde.

5. CONCLUSÃO

Em um contexto onde a proteção da privacidade dos dados é uma prioridade crucial, a criptografia homomórfica surge como uma solução promissora para permitir o treinamento de modelos de aprendizado de máquina sem comprometer a confidencialidade das informações sensíveis. Ao criptografar os dados durante todo o processo de treinamento, essa abordagem oferece uma camada adicional de segurança, garantindo que os dados permaneçam protegidos mesmo durante as operações de processamento e análise.

No entanto, ao avaliar a implementação do FHE no treinamento de modelos de ML, é importante considerar não apenas a segurança dos dados, mas também o desempenho computacional e a precisão dos resultados. Como observado em nossos experimentos com o conjunto de dados Iris do SKLearn e o *dataset* Breast_Cancer, embora o treinamento com dados criptografados possa oferecer uma solução segura, ele também acarreta um aumento significativo no tempo de processamento e uma leve redução na precisão do modelo.

Essa diferença de desempenho é particularmente evidente em conjuntos de dados maiores e mais complexos, onde o custo computacional envolvido na manipulação de dados criptografados se torna substancial. Portanto, é necessário um equilíbrio entre segurança e eficiência, adaptando a abordagem de criptografia homomórfica conforme a necessidade específica do cenário de uso.

Apesar dos desafios computacionais, o potencial da criptografia homomórfica no treinamento de modelos de aprendizado de máquina em dados sensíveis é inegável. Essa tecnologia oferece uma oportunidade única de alavancar os benefícios do aprendizado de máquina em ambientes nos quais a privacidade dos dados é uma preocupação primordial, como na área da saúde, finanças, governo e pesquisa científica.

À medida que a pesquisa e o desenvolvimento continuam avançando, espera-se que novas técnicas e otimizações na implementação da criptografia homomórfica levem a melhorias significativas no desempenho e na escalabilidade dessa abordagem. Isso abrirá portas para uma ampla gama de aplicações em que a segurança dos dados é fundamental, promovendo assim a inovação e o progresso em áreas críticas da sociedade.

Certamente, é crucial destacar que, uma vez treinado, a diferença na complexidade computacional para gerar resultados em um modelo treinado com criptografia homomórfica é praticamente nula quando comparada à de um modelo convencional. Isso torna a tecnologia bastante viável em termos de aplicação prática. No entanto, é importante reconhecer que o custo inicial para implementar essa tecnologia é significativamente alto.

Dessa forma, torna-se imperativo realizar estudos de caso detalhados para avaliar os benefícios e desafios específicos de cada cenário de uso. Enquanto o poder computacional não for suficiente para tornar as questões de desempenho insignificantes ou para surgirem novos métodos que otimizem a criptografia homomórfica, é essencial pesar cuidadosamente os prós e os contras ao decidir implementar um sistema que utilize dados criptografados no treinamento de modelos de aprendizado de máquina.

Em muitos casos, é verdade que modelos treinados com dados não criptografados podem fornecer resultados satisfatórios, mesmo quando os dados de entrada são criptografados durante a inferência. No entanto, é importante considerar os riscos associados à privacidade e à segurança dos dados durante todo o ciclo de vida do modelo, desde o treinamento até a inferência e o armazenamento dos resultados.

Portanto, embora a criptografia homomórfica ofereça uma solução promissora para proteger a privacidade dos dados durante o treinamento de modelos de aprendizado de máquina, é fundamental realizar uma análise abrangente dos custos e benefícios envolvidos, levando em consideração o contexto específico de cada aplicação. Somente dessa forma podemos tomar decisões informadas e garantir a integridade e a confidencialidade dos dados em um mundo cada vez mais orientado por dados e preocupado com a privacidade.

Em suma, o uso da criptografia homomórfica no treinamento de modelos de aprendizado de máquina representa um passo importante em direção a um futuro onde a privacidade dos dados é protegida sem comprometer o avanço da ciência e da tecnologia. Essa abordagem oferece uma maneira segura e eficaz de explorar os benefícios do aprendizado de máquina em ambientes sensíveis, capacitando organizações e pesquisadores a obter insights valiosos enquanto mantêm a confidencialidade dos dados dos usuários.

6. AGRADECIMENTOS

Gostaria de agradecer a todos que acompanharam essa longa Jornada da graduação, em especial a minha mãe, Patrícia, que sempre me apoiou e fez de tudo para que seu filho tivesse a melhor educação possível.

A Gabriel, Antônio, Beatriz e Laís, por não saírem do meu lado mesmo nos piores momentos, e pelas milhares de boas lembranças.

Também gostaria de agradecer aos professores Dalton, Reinaldo, Andrey e Edmar, por todos os conhecimentos adquiridos e oportunidades durante a graduação, também ao professor Eanes por tornar possível a confecção deste trabalho, e também pela compreensão de todos principalmente durante os últimos períodos.

Gostaria também de agradecer a todos os meus colegas de trabalho da Ivory e da Nelogica, pelos momentos compartilhados e trabalhos feitos durante esse período da graduação.

7. REFERÊNCIAS

[1] Criptografia e Segurança, https://www.booki.pt/userfiles/files/loja/preview/9789897232107.pdf

[2] New Directions in Cryptography, https://ee.stanford.edu/~hellman/publications/24.pdf

[3] Criptografia Simétrica e Assimétrica, https://www.estrategiaconcursos.com.br/blog/criptografia-simetrica-assimetrica-seguranca-informacao/

[4] Silvério, Flávia & Dahab, Ricardo. (2019). Avaliação de ferramentas de análise de vulnerabilidades para criptografia. Revista dos Trabalhos de Iniciação Científica da UNICAMP. 10.20396/revpibic262018450.

[5] Nunes, Ronnie & Gonçalves, Alexandre & Barcelos, Bartholomeo. (2023). MACHINE LEARNING NA SEGURANÇA PÚBLICA: UMA ANÁLISE DE POSSÍVEIS PROBLEMAS MECÂNICOS EM VIATURAS POLICIAIS. 10.48090/ciki.v1i1.1290.

[6] Ogburn, Monique & Turner, Claude & Dahal, Pushkar. (2013). Homomorphic Encryption. Procedia Computer Science. 20. 502-509. 10.1016/j.procs.2013.09.310.

[7] Lunkes, Aline & Borges, Fábio. (2022). Sobre a aplicação de homomorfismo na criptografia. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics. 9. 10.5540/03.2022.009.01.0306.

[8] Homomorphic Encryption for Machine Learning in Medicine and Bioinformatics, https://eprints.whiterose.ac.uk/151333/7/main.pdf

[9] Hallman, Roger & Diallo, Mamadou & August, Michael & Graves, Christopher. (2018). Homomorphic Encryption for Secure Computation on Big Data. 10.5220/0006823203400347.

[10] Processing Encrypted Data Using Homomorphic Encryption, https://www.esat.kuleuven.be/cosic/publications/article-2880.pdf

[11] Orlandi, Claudio & Piva, Alessandro & Barni, Mauro. (2007). Oblivious Neural Network Computing via Homomorphic Encryption. EURASIP Journal on Information Security. 2007. 10.1155/2007/37343.

[12] Hublet, François & Basin, David & Krstić, Srđan. (2024). Enforcing the GDPR. 10.1007/978-3-031-51476-0_20.

[13] Suma, M. & Perumal, Madhumathy. (2022). Brakerski‐Gentry‐Vaikuntanathan fully homomorphic encryption cryptography for privacy preserved data access in cloud assisted Internet of Things services using glow‐worm swarm optimization. Transactions on Emerging Telecommunications Technologies. 33. 10.1002/ett.4641.

[14] Naidu, Gireen & Zuva, Tranos & Sibanda, Elias. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. 10.1007/978-3-031-35314-7_2.

SOBRE OS AUTORES:

Filipe Ramalho é um graduando do Curso de Ciência da Computação pela UFCG que durante o curso fez parte do programa de monitoria, PET e projetos de pesquisa e é desenvolvedor profissional a 3 anos, trabalhando agora na Nelogica.

Eanes T. Pereira é Professor-Doutor pela UFCG, com vários artigos publicados