



**30º** CONGRESSO  
BRASILEIRO  
DE BIBLIOTECONOMIA  
E DOCUMENTAÇÃO



25 a 29 de novembro 2024



**Bibliotecas Fortes:**  
Sociedade Democrática Recife, PE

Eixo 6 - O mundo digital: apropriação e desafios

Modalidade: [Trabalho Completo]

## **Inteligência artificial aplicada à indicação de leitura baseado no perfil de estudantes usuários de biblioteca**

*Artificial Intelligence Applied to Reading Recommendation Based on the Profile of Library Student Users*

**Maria Aparecida Rodrigues de Souza** – Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG)

**Oscar Junior Soares da Silva** – Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG)

**Viviane Margarida Gomes Pacheco** – Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG)

**Wesley Pacheco Calixto** – Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG)

**Rita Rodrigues de Souza** – Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG)

**Resumo:** O objetivo da pesquisa foi contribuir para criação de um sistema de sugestão de leituras de acordo com o perfil dos estudantes usuários do Sistema Integrado de Bibliotecas do Instituto Federal de Goiás (IFG), a partir do caso Câmpus Inhumas, utilizando técnicas de inteligência artificial. Nesta pesquisa, utilizamos um estudo de caso com dados provenientes de relatórios anuais emitidos pela biblioteca Atena do IFG Câmpus Inhumas. A pesquisa é de abordagem descritiva-explicativa, incluindo levantamento bibliográfico e de dados primários emitidos pelo sistema de gerenciamento de acervo Sophia referente aos empréstimos realizados pelos estudantes do Câmpus Inhumas no período de 2021 a 2023. Nesse sentido, buscou-se utilizar a inteligência artificial, especificamente, redes neurais artificiais, para desenvolvimento de um sistema inteligível aplicado a serviços de disseminação seletiva de informação aos estudantes usuários de bibliotecas.

**Palavras-chave:** Recuperação de informação. Sistema de descoberta de dados. Sugestão de leitura. Filtragem de informação. Disseminação seletiva da informação.

**Abstract:** This research aimed to create a reading suggestion system based on the profiles of students using the Integrated Library System of the Federal Institute of Goiás (IFG), focusing on the Inhumas campus. A descriptive-explanatory approach was used, incorporating bibliographic surveys and primary data from the library's Sophia management system. We developed a neural network to optimize reading suggestions.





**Keywords:** KInformation retrieval. Data discovery system. Reading suggestions. Information filtering. Selective dissemination of information.

## 1 INTRODUÇÃO

Em uma sociedade cada vez mais digital e com um vasto acervo literário, técnico-científico disponível à comunidade por meio de buscadores, faz necessário incentivar esse público à leitura por meio de indicação categorizada. Segundo Silva (2021), a quantidade de informações na rede de computadores propiciou pesquisas sobre a recuperação de informações pertinentes ao usuário. Em consideração a lógica de sistemas de recuperação da informação surgiram ferramentas inteligentes de base digital para indicar itens relevantes e de acordo com os interesses dos usuários, com foco na disseminação seletiva da informação.

O artigo de Ludermir (2021) aponta como a inteligência artificial (IA) tem contribuído positivamente para aprimorar processos, a tomada de decisões e a criação de oportunidades. Acompanhar e entender essas tendências da IA é crucial para aproveitar seus benefícios e impulsionar a eficiência e a qualidade dos serviços e, principalmente, disseminar a informação de maneira personalizada. A integração da IA na recomendação de leitura contribui para uma experiência de descoberta mais eficiente e enriquecedora, fomentando a democratização do acesso à informação e incentivando a leitura em diferentes contextos.

Nos estudos de Souza et al. (2015) e Oliveira (2016), identificou-se a possibilidade de incentivar os usuários de uma biblioteca a ler por meio da categorização do perfil psicológico, utilizando-se do processo de mineração de dados e, em seguida, o desenvolvimento preliminar de um banco de dados. Essas duas pesquisas são complementares e corroboram com a importância e a possibilidade de uso das ferramentas de IA a serviço da disseminação seletiva de informação.

Considerando a importância de sugerir leitura aos usuários de uma biblioteca pelo seu perfil, a partir da rotulação dos dados retirados do relatório de empréstimo de materiais bibliográficos, tem-se a possibilidade da idealização de um sistema automatizado de indicação de leitura que melhor se adequa às necessidades e interesses do estudante.

A utilização da mineração de dados para gerar regras a partir da preferência dos leitores no estudo de Souza et al. (2015) obteve um piloto do banco de dados com todo o processo para identificação do perfil psicológico conectado ao socioeconômico, do grupo pesquisado. O problema de reconhecimento de padrões relacionados ao histórico de leituras realizadas por estudantes segundo faixa etária, área do curso, nível de escolaridade, entre outras questões, ainda não foi investigado no contexto da atuação



da Biblioteca Atena, do Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG) Câmpus Inhumas. Logo, considera-se oportuna pesquisa que investigue solução para esse problema e que resulte na indicação de leitura de acordo com os padrões reconhecidos.

O IFG, Câmpus Inhumas, oferece cursos na área da Ciência da Computação, a nível de graduação, bem como tem pesquisadores desta área e de Biblioteconomia tornando o projeto factível de ser executado. As duas características são fatores favoráveis no desenvolvimento da temática do projeto que abarca a inteligência artificial aplicada às questões de disseminação seletiva da informação para se chegar a categorização de perfil de leitor visando a indicação de leitura a usuários com o mesmo perfil. Visto isso, foram realizadas a coleta de dados de locação de utilizadores da biblioteca do curso de Engenharia de Software, entre o período de 2021 a 2023. Esses dados foram de extrema importância, visto que serão necessários para a base de dados para a recomendação de leitura.

O objetivo geral deste estudo é contribuir para a criação de um sistema de sugestão de leituras que atenda ao perfil dos estudantes usuários do Sistema Integrado de Bibliotecas (SIB) do Instituto Federal de Goiás (IFG), especificamente no caso do Câmpus Inhumas, utilizando técnicas de inteligência artificial a fim de desenvolver um sistema de sugestões para usuários da biblioteca do Câmpus.

Os objetivos específicos deste estudo incluem realizar pesquisa bibliográfica acerca das técnicas e metodologias de inteligência artificial utilizadas na classificação de perfis de leitores; obter dados pertinentes dos estudantes usuários da Biblioteca Atena, compreendendo obras emprestadas, gostos literários e técnicos, faixa etária, curso, entre outros; criar algoritmos de aprendizado de máquina que possam analisar os dados coletados e reconhecer padrões e preferências dos usuários; desenvolver um mecanismo de sugestão de leituras personalizadas que empregue os algoritmos elaborados, visando indicar livros, artigos e materiais de leitura pertinentes e ajustados aos interesses individuais de cada estudante; e analisar a efetividade do mecanismo de sugestão através de avaliações dos usuários referentes às recomendações de leitura.

## **2 FUNDAMENTAÇÃO TEÓRICA**

A Ciência da Computação tem como objetivo criar sistemas que possam executar tarefas que requerem habilidades humanas, e isso é conhecido como IA. O termo IA foi criado por Alan Turing, um matemático britânico, em 1956. Essa ciência tem evoluído consideravelmente nos últimos anos. Turing, em suas obras pioneiras, como *Computing Machinery and Intelligence* (1948) e *The Imitation Game* (1950), explorou a possibilidade de as máquinas exibirem comportamento inteligente e serem indistinguíveis dos seres humanos em determinadas tarefas. Uma de suas ideias centrais foi o "Teste de Turing", que propunha um método operacional para verificar se uma máquina poderia ser



considerada inteligente ao enganar um juiz humano quanto à sua natureza. Suas contribuições foram fundamentais para o desenvolvimento teórico e prático da IA, abrindo caminho para pesquisas e avanços subsequentes nesta área fascinante, como o desenvolvimento de algoritmos de aprendizado de máquina e redes neurais que permitem às máquinas aprenderem e tomarem decisões de forma autônoma (Turing, 2009).

Dentre os inúmeros modelos de IA, podemos destacar a Rede Neural. Uma rede neural é um método de IA que ensina computadores a processar dados de uma forma inspirada pelo cérebro humano. É um tipo de processo de machine learning, chamado aprendizado profundo, que usa nós ou neurônios interconectados em uma estrutura em camadas, semelhante ao cérebro humano. A rede neural cria um sistema adaptativo que os computadores usam para aprender com os erros e se aprimorar continuamente. As redes neurais artificiais tentam solucionar problemas complicados, como resumir documentos ou reconhecer rostos com grande precisão. As redes neurais artificiais podem ser categorizadas com base em como os dados fluem do nó de entrada para o nó de saída (AWS, 2024). Geralmente, as arquiteturas de redes neurais podem ser colocadas em três categorias específicas: Redes Neurais Feed-Forward, Redes Recorrentes e Redes Conectadas Simetricamente.

As redes Neurais Feedforward são o tipo mais comum de rede neural em aplicações práticas. A primeira camada é a entrada e a última camada é a saída. Se houver mais de uma camada oculta, nós as chamamos de redes neurais “profundas” (ou Deep Learning). Esses tipos de redes neurais calculam uma série de transformações que alteram as semelhanças entre os casos. As atividades dos neurônios em cada camada são uma função não-linear das atividades na camada anterior.

As recorrentes são tipos de redes neurais que têm ciclos direcionados em seu grafo de conexão. Isso significa que, às vezes, se pode voltar ao início, seguindo as setas. Eles podem ter uma dinâmica complicada e isso pode torná-los muito difíceis de treinar. Entretanto, estes tipos são mais biologicamente realistas. Atualmente, há muito interesse em encontrar formas eficientes de treinamento de redes recorrentes. As redes neurais recorrentes são uma maneira muito natural de modelar dados sequenciais. Eles são equivalentes a redes muito profundas com uma camada oculta por fatia de tempo; exceto que eles usam os mesmos pesos em cada fatia de tempo e recebem entrada em cada fatia. Eles têm a capacidade de lembrar informações em seu estado oculto por um longo período de tempo, mas é muito difícil treiná-las para usar esse potencial.

Por fim, as Redes Conectadas Simetricamente são como redes recorrentes, mas as conexões entre as unidades são simétricas (elas têm o mesmo peso em ambas as direções). As redes simétricas são muito mais fáceis de analisar do que as redes recorrentes. Elas também são mais restritas no que podem fazer porque obedecem a uma função de energia. As redes conectadas simetricamente sem unidades ocultas são



chamadas de “Redes Hopfield”. As redes conectadas simetricamente com unidades ocultas são chamadas de “Máquinas de Boltzmann” (Data Science Academy, 2024).

No contexto da pesquisa sobre a categorização do tipo de leitor, a IA desempenha um papel crucial. De acordo com Doneda et al. (2018), os sistemas de sugestão baseados em IA utilizam algoritmos avançados e aprendizado de máquina para examinar os costumes e preferências dos usuários. Nos últimos anos, o aprendizado de máquina passou de um campo de interesse de pesquisa acadêmica para um campo aplicado, sendo utilizado constantemente no dia a dia por inúmeros softwares (Paleyes; Urma; Lawrence, 2022).

O Brasil tem a maior extensão territorial da América Latina e a quinta maior do mundo. É também conhecido como um dos países mais diversificados, e uma formação tão diversificada destaca a sua riqueza cultural. Na verdade, os estados brasileiros têm preferências próprias de vocabulário, sotaque, gastronomia, moda, música e literatura. Especialmente na literatura, a leitura de livros é um comportamento cultural moldado por fatores sociais, contextos econômicos e locais. As pessoas escolhem o que ler de acordo com as suas ideias e realizações intelectuais, além da influência que as rodeiam numa perspectiva social. De uma perspectiva econômica, o acesso e o consumo cultural são diretamente afetados pelo estatuto socioeconômico. Então, do ponto de vista local, cada estado ou região de um país possui características que podem interferir nas preferências de leitura (Silva et al. 2021). De acordo com Silva (2021), as estratégias de recomendação de sites que utilizam recomendações podem ser muito interessantes para bibliotecas, por exemplo, como forma de incentivar a leitura na comunidade.

As IAs são capazes de propor leituras relevantes e personalizadas aos interesses individuais, proporcionando uma experiência de descoberta mais eficiente e enriquecedora. A integração da IA na recomendação de leitura pode ter o potencial de fomentar a democratização do acesso à informação, como destacado por Parreira et al. (2021), incentivando a leitura em diferentes contextos e fortalecendo o envolvimento dos leitores.

No que se refere aos aspectos morais da IA, estudiosos como Doneda et al. (2018) ressaltam a relevância de levar em conta os princípios de independência individual e moralidade na elaboração e uso dessas tecnologias. É importante assegurar que os sistemas de inteligência artificial sejam desenvolvidos de forma responsável e clara, considerando questões como a confidencialidade dos usuários, o preconceito algorítmico com dados inseridos e a transparência dos resultados. Preconceito devido a inserção de dados já imbuídos de preconceito, por exemplo, a IA não analisa dados de pessoas de xis cidade.

Ademais, o uso da IA para classificar o perfil do leitor tem implicações significativas na área educacional e no desenvolvimento de estratégias de ensino personalizadas. Parreira et al. (2021) examinam os desafios e vantagens das tecnologias



de IA na educação, destacando a importância de compreender as percepções e avaliações dos educadores em relação a essas ferramentas. A IA pode ajudar na identificação dos interesses e necessidades dos estudantes, possibilitando uma abordagem mais individual e efetiva no processo de aprendizagem.

Diante dessas perspectivas, o uso do alcance sintético para definir a silhueta do leitor revela-se altamente promissor na compreensão da recepção da informação e do envolvimento dos leitores. Conforme apontado por Silva e Mairink (2019), os sistemas de informação que se baseiam no alcance sintético são capazes de analisar de forma sofisticada os hábitos e preferências dos usuários, oferecendo sugestões de interpretação pertinentes e adaptadas aos seus interesses. A incorporação do alcance sintético na definição da silhueta do leitor visa proporcionar uma experiência de interpretação mais personalizada e enriquecedora, como destacado por Ludermitz (2021), além de contribuir para a democratização do acesso à informação.

É essencial considerar, no entanto, os aspectos éticos, a transparência e a responsabilidade na aplicação dessas tecnologias, conforme ressaltado por Garcia (2020). Nesse sentido, durante o desenvolvimento e aplicação do sistema, os dados foram utilizados de forma segura e responsável, não sendo coletadas informações pessoais dos usuários e sim dados acerca da área de conhecimento dos livros utilizados. As informações foram tratadas apenas como entradas para o modelo, garantindo que nenhum indivíduo específico fosse identificado ou exposto. Uma vez que trata de IA em contextos educacionais e de bibliotecas, as questões éticas foram asseguradas observando o comportamento geral de uma coletividade com fim no rigor acadêmico e alinhamento com as expectativas da área de Biblioteconomia e IA derivados de valores éticos universais (Santos, 2017). A consulta frequente e a atualização nessa área são necessárias para desvendar a realidade e a confiabilidade dos sistemas de informação baseados em alcance sintético, garantindo uma experiência de interpretação que seja não apenas personalizada, mas também ética e responsável.

### **3 METODOLOGIA**

A metodologia foi dada por meio de pesquisa descritiva-explicativa. Na pesquisa descritiva, buscou-se um aprofundamento no tema, enquanto a explicativa procurou-se conectar as ideias para compreender causas e efeitos (Diana, 2024). A coleta e o levantamento de dados foram de base qualitativa e quantitativa, por meio de métodos experimentais com os dados disponíveis por meio de planilhas. Os métodos experimentais foram feitos a partir de conversão de dados nominais e numéricos, por exemplo a transformação da cidade em CEP e, posteriormente, em código de município.

A pesquisa foi realizada por meio de levantamento bibliográfico, estudos de caso e de dados primários emitidos pelo sistema de gerenciamento de acervo Sophia referente aos empréstimos realizados pelos estudantes do Câmpus Inhumas, no período de 2021 a 2023.

Para a obtenção dos dados foram analisados os relatórios emitidos pelo sistema

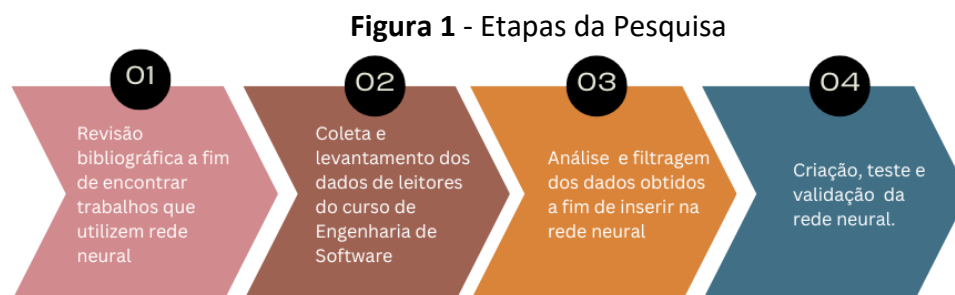
de gerenciamento de acervo, o Sophia, mais especificamente as categorias área do conhecimento conforme Classificação Decimal de Dewey (CDD) e indexação do assunto do material retirado por empréstimo por estudantes de determinado curso.

Vale ressaltar que, de acordo com Parreira, Lehmann e Oliveira (2021), há uma necessidade de levar em conta questões éticas e de responsabilidade ao implantar modelos de aprendizado de máquina, assegurando transparência e capacidade de explicação dos resultados. Tais considerações foram fundamentais para o desenvolvimento de soluções de IA confiáveis e eficazes na seleção de informações pertinentes aos usuários.

Para a criação da rede neural artificial, inicialmente, foi feito uma base de dados em uma planilha a fim de fazer o tratamento desses dados e, posteriormente, a criação da rede neural para sugestões de leitura. O tratamento desses dados passaram por diversas filtragens a fim de conseguirmos os dados mais otimizados possíveis para a rede.

Para a criação da rede neural, utilizamos a linguagem de programação *Python* (Python, 2024), devido à sua extensa coleção de bibliotecas voltadas para o aprendizado de máquina e sua fácil integração com ferramentas de análise de dados. Especificamente, utilizamos algumas bibliotecas como Scikit-learn (Scikit-Learn, 2024), que oferece uma vasta gama de algoritmos e funcionalidades para a construção e treinamento de modelos de aprendizado de máquina. A programação toda foi feita no ambiente de desenvolvimento disponibilizado pelo *Google*, o *Google Collab*, um serviço de nuvem gratuito hospedado pelo próprio Google para incentivar a pesquisa de Aprendizado de Máquina e Inteligência Artificial (SANTOS, 2023).

A Figura 1, representa de forma dinâmica, como foram feitas, as etapas da metodologia da nossa pesquisa englobando do início ao fim.



Fonte: Produzido pelos próprios autores (2024).

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Coleta de dados

A coleta de dados foi realizada com o intermédio da Coordenação da Biblioteca de Inhumas. Os dados foram disponibilizados por meio de relatórios em PDF contendo todas as informações de locação de cada aluno do curso de Engenharia de *Software* no período de 2021 a 2023.

Dentre os dados obtidos através dos relatórios, conseguimos identificar 381 locações de livros entre o período de 2021 a 2023. Dentro dessas locações, utilizamos os seguintes dados: Matrícula (a fim de diferenciar cada locação), ISBN (International Standard Book Number), CDD (Classificação Decimal de Dewey), Tipo Literário englobando Didático,



Literário, Livro técnico, Paradidático e Referência, Gênero (masculino ou feminino), Dias Vividos (Um macro utilizado para pegar os dias de vida do usuário desde o nascimento até a criação da tabela), Cidade (Transformada em código do município segundo o IBGE), Período do Curso, Data de Saída, Dias Permitidos, Dias utilizados e se está em Atraso. Esses dados foram essenciais para identificarmos os perfis de leitores e assim, posteriormente, fazermos uma rede neural a fim de gerar recomendações aos leitores. Em síntese, os dados para a rede neural foram: matrícula, ISBN, CDD, tipo literário, dias vividos, gênero, cidade, período do curso, data de saída, dias permitidos, dias utilizados, verificação de atraso.

#### 4.2 Processamento de dados

Depois do acesso aos dados, foi realizado o tratamento destes em planilhas a fim de entender como utilizar da melhor forma na rede neural. Inicialmente, houve uma organização dos dados, sendo adicionado um requisito a cada seleção.

Para o maior entendimento da rede neural, decidimos transformar todos os dados em números inteiros, pois assim quando houvesse a inserção no software seria mais otimizado. Sendo assim, houve a necessidade da criação de uma legenda para termos uma compreensão do que se trata e fazer a padronização em números. Na Tabela 1, podemos ver os dados com sua respectiva numeração.

**Tabela 1** - Legenda dos dados.

<b>Gênero</b>	M (0), F (1)
<b>Tipo</b>	Didático (0), Literário (1), Livro técnico (2), Paradidático (3), Referência (4)
<b>Cidade</b>	Inhumas (5210000), Aparecida de Goiânia (5201405), Araçu (5201603), Brazabrantes (5203609), Caturai (5205208), Cristalina (5206206), Goianira (5208806), Itauçu (5211404), Jacobina (2917508), Joviânia (5212105), Santa Rosa de Goiás (5219506), Senador Canedo (5220454), Uruaçu (5221601), Uruana de Minas (3170479), Goiania (3170479)
<b>Área</b>	Ciência da Computação, Informação e Trabalhos Gerais (0), Filosofia, Parapsicologia, Ocultismo e Psicologia (1), Religião (2) Ciências Sociais (3), Linguagens (4), Ciências Naturais e Matemática (5), Tecnologia e Ciência Aplicada (6), Artes (7), Literatura e Retórica (8) História, Geografia e Disciplinas Auxiliares (9)
<b>Em Atraso</b>	Não (0), Sim (1)

Fonte: Produzida pelos próprios autores (2024).

As locações identificadas nesse período foram somente referentes aos cursos de Bacharelado em Engenharia de *Software*, abrangendo do 1º período até o 6º período. Com base nesses dados, foi possível elaborar a tabela 1 de dados para a rede.



### 4.3 Validação dos dados

Posteriormente a esse processo, iniciamos a etapa de validação, verificando os dados e como iríamos trabalhar da melhor forma para inserir na rede neural, para compreender o que seria mais propício para utilizarmos na base de dados da rede neural.

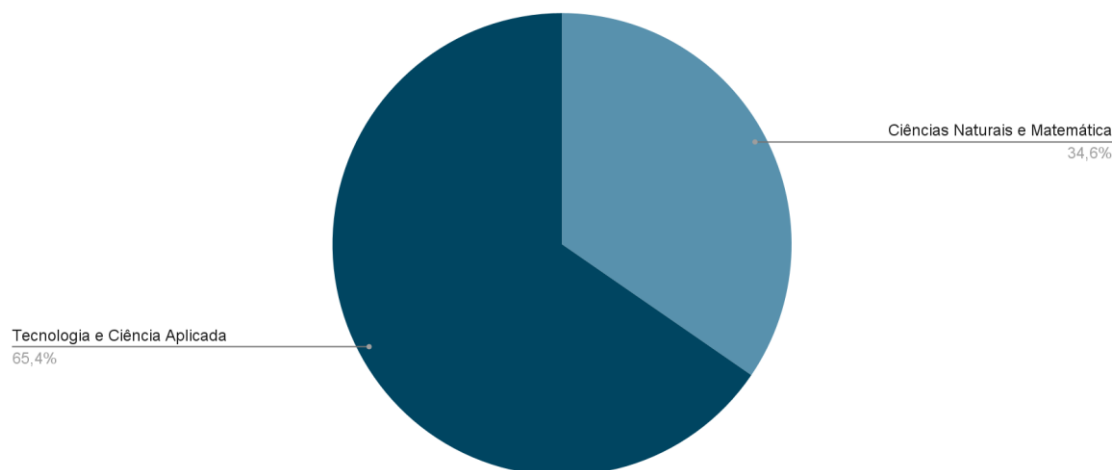
Após essa validação, conseguimos realizar diversos testes, sendo testes lógicos a fim de verificar a melhor forma possível, para a rede neural; conseguindo assim refinar as 381 locações em 81 locações, evitando assim deixarmos muitos dados dispersos ou repetidos para a IA.

Essa filtragem foi essencial para definirmos os perfis dos leitores a fim de assim realizarmos indicações com base nos dados. Com isso podemos ter resultados mais precisos, deixando assim uma rede cada vez mais assertiva nos seus resultados.

Vale ressaltar que como a base de dados foi feita especificamente de um único curso, os perfis para aplicação da rede neural ficaram limitados a somente 2 tipos predominantes. Nesse caso, também há a possibilidade de alguns erros de processamento de determinados perfis se houver uma variação muito extensa. Como podemos ver no gráfico a seguir, dentre os dados inseridos conseguimos classificar da seguinte forma.

Gráfico 1. Perfis de leitura

Perfis Literários



Fonte: Produzida pelos próprios autores (2024).

### 4.4 Rede neural

#### 4.4.1 Treinamento

Para a implementação da rede neural *Multilayer Perceptron* (MLP) utilizamos a linguagem de programação *Python*, devido à sua extensa coleção de bibliotecas voltadas para o aprendizado de máquina e sua fácil integração com ferramentas de análise de dados. Especificamente, utilizamos a biblioteca *Scikit-learn*, que oferece uma vasta gama de algoritmos e funcionalidades para a construção e treinamento de modelos de aprendizado de máquina. A divisão dos dados foi dada da seguinte forma: 80% para



treinamento e 20% para testes e validação.

Para a inserção dos dados, transformamos a tabela de dados em um arquivo .csv para ter uma melhor inserção dos dados, aumentando assim o desempenho da rede.

#### 4.4.2 Testes

A MLP foi configurada com a seguinte estrutura (Data Science Academy, 2024):

- A. Camada de Entrada: A camada de entrada consiste em nove neurônios, correspondendo às nove características de entrada do nosso conjunto de dados, que incluem: matrícula, ISBN, CDD, tipo de livro, gênero do usuário, dias vividos, cidade, período do curso, atraso na devolução;
- B. Camadas Ocultas: Experimentamos diferentes configurações para as camadas ocultas, variando o número de camadas de uma a três e o número de neurônios por camada de 5 a 15. As funções de ativação testadas foram *ReLU* (Rectified Linear Unit) e *Sigmoid* (Probabilidade). O objetivo foi identificar a arquitetura que proporcionasse o melhor desempenho em termos de acurácia;
- C. Camada de Saída: Para cada rótulo a ser previsto (gênero literário preferido e constância de leitura), utilizamos uma camada de saída apropriada. No caso da previsão de categorias múltiplas, aplicamos a função de ativação *softmax* para converter as saídas em probabilidades de classe.

O treinamento do modelo foi realizado utilizando o otimizador Adam (*Adaptive moment estimation*) uma ferramenta conhecida por sua eficiência em problemas de alta dimensionalidade e grande quantidade de dados (Tato, Nkambou, 2018). A taxa de aprendizado inicial foi configurada para 0.001, e o número máximo de interações foi definido como 500.

#### 4.4.3 Validação da rede neural

Para evitar o *overfitting* (em português ajuste excessivo) e garantir a generalização do modelo, dividimos os dados em conjuntos de treinamento e teste. *Overfitting* é um comportamento indesejável de aprendizado de máquina que ocorre quando o modelo de aprendizado de máquina fornece previsões precisas para dados de treinamento, mas não para novos dados. Quando os cientistas de dados usam modelos de aprendizado de máquina para fazer previsões, eles primeiro treinam o modelo em um conjunto de dados conhecido. Em seguida, com base nessas informações, o modelo tenta prever resultados para novos conjuntos de dados. Um modelo de sobreajuste pode fornecer previsões imprecisas e não pode ter um bom desempenho para todos os tipos de novos dados (AWS, 2024).

Utilizamos a biblioteca *StandardScaler* (Scikit-Learn, 2024) para padronizar as características dos dados, melhorando assim a convergência durante o treinamento. A partir do perfil será indicado o gênero que o usuário tem preferência, conseqüentemente a indicação de produções do gênero



## 5 CONSIDERAÇÕES FINAIS

A pesquisa conduzida sobre a aplicação da IA para a indicação de leitura com base no perfil dos estudantes usuários da biblioteca do IFG, Câmpus Inhumas, resultou em avanços significativos no desenvolvimento de sistemas de recomendação personalizados. Um resultado positivo é o mapeamento de perfis dos leitores, o que pode gerar inúmeras outras possibilidades além da indicação de livros.

Este estudo se destacou ao empregar algoritmos de aprendizado de máquina, especificamente uma rede neural MLP, para analisar dados históricos de empréstimos e gerar sugestões de leitura adaptadas aos interesses individuais dos usuários.

A implementação da rede neural utilizando a linguagem Python e a biblioteca Scikit-learn provou ser eficaz e acessível, por ser uma biblioteca de fácil acesso e resolver o problema em questão. A MLP configurada com várias camadas ocultas e diferentes funções de ativação permitiu experimentos que identificaram a melhor arquitetura para o modelo. O uso do otimizador e a padronização dos dados com a biblioteca garantiram a eficiência do treinamento e a generalização dos resultados.

Os resultados mostraram que a rede neural MLP pode efetivamente prever o gênero literário preferido e a constância de leitura dos usuários, alcançando uma acurácia, termo utilizado para medir a precisão da rede neural, satisfatória. Isso não apenas valida a viabilidade técnica do sistema de recomendação, mas também destaca o potencial da IA em transformar a experiência de leitura dos estudantes, incentivando a leitura e democratizando o acesso à informação.

Devido à pequena quantidade de dados para base, esses resultados podem apresentar uma taxa de erro se expostos a mais quantidade de dados para a previsão. Com isso, para uma exposição maior, haveria uma necessidade de uma base de dados maior. Com isso, os resultados estão de acordo com o esperado da pesquisa.

A pesquisa contribui significativamente para o campo da Biblioteconomia e Ciência da Computação, demonstrando como técnicas de IA podem ser aplicadas para melhorar os serviços de disseminação seletiva de informação. Além disso, este estudo fornece uma base sólida para futuras pesquisas e desenvolvimento de sistemas inteligentes de recomendação em bibliotecas e outras instituições educacionais.

Em conclusão, a pesquisa atingiu seu objetivo de criar um modelo de indicação de leitura baseado no perfil dos usuários, oferecendo uma ferramenta para bibliotecas promoverem a leitura de forma personalizada. Devido a sugestão que é feita, as bibliotecas podem se apropriar disso para desenvolver inúmeras questões. Até na parte de aquisição de livros novos poderia levar isso em conta. O resultado deste estudo indica a importância da integração da IA em processos educacionais e culturais, abrindo caminho para inovações contínuas na maneira como consumimos e interagimos com a informação. Contudo, recomenda-se posterior a esse estudo, analisar e avaliar por meio de entrevistas junto aos usuários da biblioteca, os pontos dos quais contribuirá mais qualitativamente para a criação da rede neural.

## REFERÊNCIAS

AMAZON SAGEMAKER (AWS). **O que é uma rede neural?** Disponível em: <https://aws.amazon.com/pt/what-is/neural-network/>. Acesso em: 15 abr. 2024.



DATA SCIENCE ACADEMY. Capítulo 9: a arquitetura das redes neurais. In: \_\_\_\_\_. Deep Learning Book. 2024. Disponível em: <https://www.deeplearningbook.com.br/a-arquitetura-das-redes-neurais/>. Acesso em: 07 jun. 2024.

DIANA, Juliana. **Pesquisa descritiva, exploratória e explicativa**. Disponível em: <https://www.diferenca.com/pesquisa-descritiva-exploratoria-e-explicativa/>. Acesso em: 25 jun. 2024.

DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; AN, Norberto Nuno Martin Becerra Gomes de. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Revista de Ciência Jurídica**, v. 23, n. 4, 2018. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257/>. Acesso em: 15 jun. 2024.

GARCIA, Ana Cristina. Ética e inteligência artificial. **Computação Brasil**, n. 43, p. 14-22, 2020. Disponível em: <https://sol.sbc.org.br/journals/index.php/comp-br/article/view/1791>. Acesso em: 15 jun. 2024.

LUDERMIR, Teresa Bernarda. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, v. 35, p. 85-94, 2021. Disponível em: <https://www.scielo.br/j/ea/a/wXBdv8yHBV9xHz8qG5RCgZd/?lang=pt&format=html>. Acesso em: 15 jun. 2024.

OLIVEIRA, Rhuandry de. **Plataforma web para manipulação de questionários**. Inhumas, 2016. Monografia (Graduação) Instituto Federal de Goiás, Câmpus Inhumas, Curso de Bacharelado em Informática, 2016.

PALEYES, Andrei; URMA, Raoul-Gabriel; LAWRENCE, Neil D. Challenges in deploying machine learning: a survey of case studies. **ACM Computing Surveys**, v. 55, n. 6, p. 1-29, 2022.

PYTHON. Disponível em: <https://www.python.org>. Acesso em: 23 maio 2024

PARREIRA, Artur; LEHMANN, Lúcia; OLIVEIRA, Mariana. O desafio das tecnologias de inteligência artificial na Educação: percepção e avaliação dos professores. **Ensaio: avaliação e políticas públicas em educação**, v. 29, p. 975-999, 2021. Disponível em: <https://www.scielo.br/j/ensaio/a/nM9Rk8swvtDvwWNrKCZtjGn/?format=html>. Acesso em: 15 jun. 2024.

SANTOS, Luiz Henrique Lopes dos. Sobre a integridade ética da pesquisa. **Cienc. Cult.**, São Paulo, v. 69 n. 3, jul./set. 2017. <http://dx.doi.org/10.21800/2317-66602017000300002>. Acesso em: 12 jun. 2024.

SANTOS, Thiago. **Google Colab: o que é, tutorial de como usar e criar códigos**. 2023. Alura. Disponível em: <https://www.alura.com.br/artigos/google-colab-o-que-e-e-como-usar>. Acesso em: 15 jun. 2024.

SCIKIT-LEARN. **Machine learning in Python**. Disponível em: <https://scikit-learn.org/stable/> Acesso em: 23 maio 2024.



SILVA, J. A. S. DA; MAIRINK, C. H. P. Inteligência artificial. LIBERTAS: **Revista de Ciências Sociais Aplicadas**, v. 9, n. 2, p. 64-85, 13 dez. 2019. Disponível em:

<http://famigvirtual.com.br/famig-libertas/index.php/libertas/article/view/247>. Acesso em: 15 jun. 2024.

SILVA, **Juliene Lobato da. Sistemas de recomendação como ferramenta para indicações de livros: um estudo de caso na plataforma Goodreads.** 2021. Disponível em:

<https://rosario.ufma.br/jspui/bitstream/123456789/5293/1/JULIENELOBATODASILVA.pdf>. Acesso em: 15 jun. 2024.

SILVA, Mariana O.; SCOFIELD, Clarisse; OLIVEIRA, Gabriel P.; SEUFITELLI, Danilo B.; MORO, Mirella M.. Exploring Brazilian Cultural Identity Through Reading Preferences. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING (BRASNAM), 10. , 2021, Evento Online. **Anais ...** Porto Alegre: Sociedade Brasileira de Computação, 2021 . p. 115-126. ISSN 2595-6094. DOI: <https://doi.org/10.5753/brasnam.2021.16130>. Acesso em: 10 jun. 2024.

SOUZA, M. A. R. et al. Leitura categorizada para os/as discentes de licenciatura em Química do IFG Câmpus Inhumas. **Revista Tecnologia Educacional**, Edição Especial do V Congresso Internacional de Tecnologia Educacional da ABT, v. 3. n. 211, p. 108-111, out./dez. 2015. Disponível em:

[https://www.academia.edu/23440575/LEITURA\\_CATEGORIZADA\\_PARA\\_OS\\_AS\\_DISCENTES\\_DE\\_LICENCIATURA\\_EM\\_QU%C3%8DMICA\\_DO\\_IFG\\_C%C3%82MPUS\\_INHUMAS](https://www.academia.edu/23440575/LEITURA_CATEGORIZADA_PARA_OS_AS_DISCENTES_DE_LICENCIATURA_EM_QU%C3%8DMICA_DO_IFG_C%C3%82MPUS_INHUMAS). Acesso em: 15 jun. 2024.

TURING, A. M. **Computing Machinery and Intelligence**. In: Epstein, R.; Roberts, G.; Beber, G. (eds) Parsing the Turing Test. Dordrecht: Springer, 2009. Disponível em:

[https://doi.org/10.1007/978-1-4020-6710-5\\_3](https://doi.org/10.1007/978-1-4020-6710-5_3). Acesso em: 15 jun. 2024.

TATO, Anger; Nkambou, Roger. Improving adam optimizer. 2018. **Workshop track - ICLR 2018**. Disponível em: <https://openreview.net/forum?id=HJfpZq1DM>. Acesso em: 30 maio 2024.