

UNIVERSIDADE POSITIVO
CURSO DE CIÊNCIA DA COMPUTAÇÃO

BRUNA AMARAL CHAPELIN

LANNA DOS ANJOS COSTA

LUÍSA DE SOUZA KOERICH

INTELIGÊNCIA ARTIFICIAL APLICADA A DADOS REAIS

CURITIBA

2025

RESUMO

O seguinte projeto propõe a utilização de uma rede neural do tipo Multilayer Perceptron (MLP) para prever aspectos sociais de mulheres residentes nas capitais brasileiras. A abordagem consiste em relacionar indicadores socioeconômicos, como taxa de alfabetização e renda média, à probabilidade de aquisição e manutenção de uma moradia. Para isso, o modelo é alimentado com dados numéricos normalizados e organizados, destinados ao processo de aprendizado supervisionado da rede.

A escolha do tema fundamenta-se em sua relevância social e econômica. Compreender como esses indicadores se relacionam permite identificar possíveis origens de problemas estruturais, tais como baixos níveis de alfabetização, dificuldades de acesso à moradia e limitações de poder aquisitivo nessa parcela da população. A análise dessas relações é essencial para subsidiar ações que visem à mitigação dessas desigualdades.

Assim, o objetivo do projeto é treinar um modelo MLP capaz de aprender os padrões existentes entre diferentes índices e, a partir deles, prever resultados que auxiliem na formulação de estratégias e soluções direcionadas. Os resultados esperados incluem a avaliação dos erros de treinamento por meio de vetores de perda, e a visualização de métricas como precisão e evolução da função de custo ao longo das iterações.

SUMÁRIO

1 DESCRIÇÃO DO DATASET	4
2 METODOLOGIA	4
3 RESULTADOS OBTIDOS	5
4 DISCUSSÃO	5
5 CONCLUSÃO	6

1 DESCRIÇÃO DO DATASET

- Tema: Indicadores Sociais das Mulheres Brasileiras
- Fonte: IBGE – Censo Demográfico 1991 e 2000
- Período: 1991 – 2000
- Quantidade de amostras e atributos: 20 amostras de treino e 7 amostras de teste para cada tabela (5 tabelas), apenas do ano 2000. Foram utilizados 8 atributos.

2 METODOLOGIA

A metodologia aplicada neste estudo consiste no desenvolvimento e treinamento de uma Rede Neural Artificial do tipo Multilayer Perceptron (MLP), implementada em python, utilizando a biblioteca PyTorch. O modelo foi projetado com o objetivo de realizar classificação binária, a partir de um conjunto de indicadores socioeconômicos estruturados e normalizados previamente.

• Arquitetura da Rede

A arquitetura é composta por uma camada de entrada, três camadas ocultas e uma camada de saída. A classe `ClassificadorMLP` define essa estrutura da seguinte forma:

- Camada de entrada: recebe `tam_entrada` atributos provenientes do dataset.
- Camada Oculta 1: 64 neurônios totalmente conectados.
- Camada Oculta 2: 32 neurônios totalmente conectados.
- Camada Oculta 3: 16 neurônios totalmente conectados.
- Camada de Saída: 2 neurônios, correspondentes às classes alvo do problema de classificação.

As conexões entre camadas são implementadas por `nn.Linear()`.

• Funções de Ativação

Além da arquitetura principal utilizando ReLU, foram realizadas avaliações experimentais com diferentes funções de ativação, com o objetivo de identificar qual delas proporcionaria melhor desempenho no problema de classificação. Para isso, foi utilizada uma versão alternativa do modelo (`ClassificadorMLP_Exp`), que permite a parametrização da função de ativação utilizada nas camadas ocultas.

Nesse experimento comparativo, foram testadas três funções de ativação: Sigmoid, Tanh e LeakyReLU. Cada uma delas foi aplicada mantendo-se constantes

todos os demais parâmetros da rede e os mesmos dados de treino e teste, resultando assim em uma avaliação justa de desempenho. Cada variante foi treinada por 100 épocas, registrando-se a evolução da perda e a acurácia final obtida por cada modelo.

Além disso, utilizou-se a técnica Dropout, com taxa de 0,3, aplicada após cada camada oculta, com o objetivo de reduzir overfitting e melhorar a capacidade de generalização do modelo.

3 RESULTADOS OBTIDOS

Após os experimentos com diferentes funções de ativação, cada modelo foi treinado por 100 épocas, mantendo constantes todos os demais parâmetros de treinamento, como taxa de aprendizado (0.001), estrutura da rede e conjunto de dados. Durante o processo foram coletadas métricas de perda e acurácia, calculadas a cada época e armazenadas em vetores (perda_hist e acc_hist), permitindo análise visual posterior.

Os resultados finais são apresentados ao término do código, por meio de um resumo automático que imprime a acurácia final obtida por cada função de ativação. Além das métricas numéricas, o código gera gráficos que apresentam a evolução da perda e da acurácia ao longo das épocas, possibilitando visualizar o comportamento do aprendizado, identificar estabilidade ou oscilações e observar a taxa de convergência para cada função de ativação.

No final, são comparadas as acurácias finais de cada modelo, possibilitando identificar qual função de ativação apresentou o melhor desempenho no problema específico de classificação entre as classes definidas no dataset. Esse processo seleciona a escolha da função de ativação mais adequada à arquitetura.

4 DISCUSSÃO

Os resultados obtidos demonstraram que o modelo foi capaz de aprender padrões presentes nos indicadores socioeconômicos utilizados e produzir classificações com desempenho crescente ao longo das épocas, conforme demonstrado pela evolução das métricas de perda e acurácia registradas durante o treinamento. Entretanto, a análise dos diferentes experimentos com funções de ativação (Sigmoid, Tanh e LeakyReLU) revelou que o comportamento do modelo varia significativamente de acordo com a configuração adotada, indicando que a escolha da função de ativação exerce impacto direto na eficiência do aprendizado e na capacidade de generalização do modelo.

Durante a comparação, foi possível identificar diferenças perceptíveis nas curvas de convergência e nas acurácias finais, evidenciando que algumas funções apresentaram convergência mais estável ou menor taxa de perda.

Apesar do desempenho satisfatório, o modelo apresenta algumas limitações que devem ser consideradas. O tamanho reduzido da base de dados restrita às capitais brasileiras pode limitar a capacidade de generalização, especialmente para cenários mais amplos da realidade socioeconômica do país. Também é importante destacar que indicadores socioeconômicos possuem natureza complexa e podem sofrer influência de variáveis externas que não estão representadas no conjunto de atributos utilizado. Assim, mesmo com boa acurácia, o modelo pode estar capturando correlações indiretas e não necessariamente relações causais, o que pode exigir uma interpretação cuidadosa dos resultados.

5 CONCLUSÃO

O projeto demonstrou, de forma prática, a capacidade de Redes Neurais Artificiais do tipo Multilayer Perceptron (MLP) na análise de indicadores socioeconômicos. A experimentação revelou que escolhas arquiteturais, como a seleção de funções de ativação, impactam diretamente o desempenho final do modelo, salientando a necessidade de ajustes finos para a obtenção de resultados robustos.

Por meio de testes com diferentes funções de ativação, constatou-se que decisões aparentemente simples na arquitetura da rede exercem influência significativa em seu rendimento, o que reforça a importância de uma calibração cuidadosa. Os experimentos realizados comprovaram a capacidade do MLP de aprender padrões relevantes nos dados, evidenciada pela evolução consistente da acurácia e pela redução da função de perda ao longo das épocas de treinamento. Contudo, limitações inerentes ao estudo tornaram-se evidentes, destacando-se principalmente o tamanho reduzido da base de dados e a ausência de variáveis capazes de representar a total complexidade dos fatores socioeconômicos em análise. Tais restrições indicam que, embora o modelo tenha apresentado resultados promissores, suas previsões devem ser interpretadas com cautela e não podem ser consideradas como relações estritamente causais.

Entre as lições aprendidas, destaca-se a importância da preparação adequada dos dados, da escolha cuidadosa da função de ativação e do uso de técnicas de regularização, como o dropout, para mitigar problemas de overfitting.

Como potenciais melhorias para trabalhos futuros, recomenda-se: a ampliação do dataset; a inclusão de novos atributos que capturem nuances socioeconômicas adicionais; e a experimentação com arquiteturas alternativas, como Redes Neurais

Profundas (DNNs) ou modelos baseados em árvore de decisão (Random Forest, Gradient Boosting), para fins de análise comparativa.

Em síntese, o projeto alcançou seu objetivo central de treinar e avaliar um classificador baseado em MLP, oferecendo contribuições tanto no âmbito técnico-metodológico quanto no analítico para o campo de estudo.