

Introduction to Artificial Intelligence and Generative Learning

CPS 769

Segundo Trimestre de 2024

Professores:

Edmundo de Souza e Silva (PESC/COPPE/UFRJ)

Rosa M. Leão (PESC/COPPE/UFRJ)

Participação Especial: Gaspare Bruno (Diretor Inovação, ANLIX)

Lista de Exercícios 2

ATENÇÃO! Faça as listas de forma que TODAS AS RESPOSTAS sejam DEVIDAMENTE COMENTADAS (passos para se chegar a resposta). Ser claro e objetivo facilitará organizar as ideias para as discussões em classe.

Para facilitar escrever a lista de forma clara, é possível traduzir equações escritas a mão para LaTeX:

<https://mathpix.com/>, ver também

https://www.overleaf.com/learn/latex/Questions/Are_there_any_tools_to_help_transcribe_mathematical_formulae_into_LaTeX%3F.

Questão 1

O objetivo deste trabalho é entender como construir um modelo preditivo simples (de Redes Neurais) usando uma ou mais camadas ocultas, e se familiarizar com os códigos em Python. Nesta tarefa, você deverá construir um modelo de rede neural simples para prever a nota de cada aluno em uma turma com base em duas *features*:

- Fração de palestras assistidas
- Número de horas estudadas por semana (até o máximo de 8 horas em uma semana).

Sobre a Rede Neural e programa:

- A sua rede neural deverá ter uma camada oculta com 3 neurônios. Portanto teremos 2 entradas (as 2 *features*) uma camada oculta e, na camada de saída, um neurônio.
- Os dados de entrada são fornecidos em uma planilha .ods (libreoffice), com os dados de 500 estudantes.
- Uma vez lido, o seu dataset deve ser aleatoriamente dividido de forma a que 80% seja para treino do modelo e os restantes 20% para teste.
- No seu programa:
 - Use um *Scaler* padrão do *sklearn* para dimensionar os dados de treinamento. Explique o motivo de escalonar os dados de entrada.
 - Para a camada oculta, use a função de ativação ReLU (*Rectified Linear Unit*).
 - A camada de saída usa a ativação linear padrão (explique o que é).
 - Use o algoritmo de otimização *Adam* (*Adaptive Moment Estimation*). Explique bem resumidamente as vantagens em relação ao *Stochastic Gradient Descent* padrão.
 - Use *mean square error* para a função de perda. Explique bem resumidamente o objetivo da função de perda.

Responda:

1. Como o modelo de rede neural está estruturado? Explique a arquitetura.
2. Explique o papel da função de ativação usada na camada oculta.

3. Treine o modelo com o conjunto de dados fornecido. Qual o erro quadrático médio nos dados de teste?
4. Trace o erro quadrático médio em função das “épocas” (dos passos para a convergência). Descreva a tendência que você observa.
5. Mostre os pesos e *bias* de cada camada após o treinamento, isto é, mostre os parâmetros aprendidos do modelo.
6. Use o modelo treinado para prever as notas a partir dos dados de novos alunos (com um segundo dataset fornecido sem as notas).
Mostre as previsões feitas pelo modelo e explique os resultados.
7. Modifique o programa para adicionar mais camadas e/ou maior número de neurônios ocultos. Como suas modificações afetam o desempenho do modelo?
8. Quais seriam algumas melhorias potenciais ou recursos adicionais que poderiam ser adicionados ao modelo para melhorar sua precisão preditiva?
9. Faria sentido usar a função de ativação sigmoid no modelo? Explique em poucas palavras.