

ENG1456 - Redes Neurais - Trabalho 1

Classificação

Aluno: Matheus Carneiro Nogueira

Professora: Marley Velasco

Sumário

1	Apresentação do dataset	2
2	Compreensão do problema e análise de variáveis	2
2.1	Observe a base de dados do problema. Existem variáveis que podem ser eliminadas do dataset? Justifique.	2
2.2	Implemente técnicas de visualização de dados e seleção de variáveis para extrair características importantes sobre a base de dados. Explique a motivação destas técnicas e o que é possível inferir dos resultados obtidos. . . .	3
3	Treinamento do modelo de Rede Neural	3
3.1	Com as configurações do modelo MLP previamente definidas no script, faça o treinamento da Rede Neural sem normalizar os atributos numéricos. Comente o resultado obtido, baseado nas métricas de avaliação disponíveis (acurácia, precision, recall, F1-Score, Matriz de confusão, etc.)	3
3.2	Agora normalize os dados de entrada e treine novamente o modelo MLP. Avalie os resultados obtidos e comente o efeito da normalização no treinamento da Rede Neural.	3
4	Mudança de configurações do modelo	3
4.1	Insira o conjunto de validação para o treinamento do modelo. Avalie o resultado obtido.	3
4.2	Modifique o tempo de treinamento (épocas) da Rede Neural. Escolha dois valores distintos (e.g. 1 e 1000 épocas) e avalie os resultados.	3
4.3	Modifique a taxa de aprendizado da Rede Neural. Escolha dois valores distintos (e.g. 0,001 e 0,1) e avalie os resultados.	3
4.4	Modifique a quantidade de neurônios na camada escondida da Rede Neural. Escolha dois valores distintos (e.g. 2 e 70 neurônios) e avalie os resultados.	3
5	Teste Livre	3
5.1	Faça novos testes para avaliar o desempenho da Rede Neural no problema designado. Use a técnica K-Fold (com $K = 10$) para analisar o resultado obtido.	3
5.2	Faça análises e novas implementações que você julgue importante para o seu trabalho. Não esqueça de explicar a motivação da análise realizada. . .	3

Resumo

Este documento consiste no relatório do trabalho 1 do módulo de Redes Neurais da disciplina ENG1456 da PUC-Rio. Nele será explicada a implementação de modelos de Redes Neurais MLP para a classificação do dataset Ionosphere, disponibilizado pela professora da disciplina e disponível em [1]. As seções do relatório são definidas de acordo com as perguntas principais que constam no arquivo Guia de Atividades.

1 Apresentação do dataset

Com o intuito de criar um modelo de rede neural para classificar o dataset *Ionosphere*, é essencial que, antes de implementar a rede, estudemos o dataset em si. Como descrito no artigo [2] e pelas informações disponibilizadas em [1], este dataset consiste no sistema de radares da região de Goose Bay, Labrador, o qual possui 16 antenas de alta frequência. O alvo dessas antenas foram elétrons livres na Ionosfera e, bons retornos do radar consistem em retornos que indicam alguma estrutura na ionosfera, enquanto más retornos são aqueles cujo sinal passa livre pela Ionosfera e não retorna. É, justamente, esta classificação que a Rede Neural desenvolvida neste trabalho almeja realizar, separar as entradas em "good" ou "bad". Com isso, notamos que a rede pode possuir uma saída única, que mapeia **completar relacao entre 1 e 0 e good e bad**

2 Compreensão do problema e análise de variáveis

2.1 Observe a base de dados do problema. Existem variáveis que podem ser eliminadas do dataset? Justifique.

Ao analisar as colunas do dataset, percebe-se que as duas primeiras colunas, *info_1* e *info_2* aparentam ser colunas inteiramente de 1 e 0, respectivamente. Para ter certeza, uma vez que observar apenas o head e tail não é suficiente, as colunas inteiras foram analisadas e confirmou-se que a segunda coluna, *info_1* de fato é inteiramente de 0. Sendo assim, é recomendável que ela seja excluída do dataset pois, além de não trazer nenhuma informação útil à classificação dos retornos de rádio.

- 2.2 Implemente técnicas de visualização de dados e seleção de variáveis para extrair características importantes sobre a base de dados. Explique a motivação destas técnicas e o que é possível inferir dos resultados obtidos.

3 Treinamento do modelo de Rede Neural

- 3.1 Com as configurações do modelo MLP previamente definidas no script, faça o treinamento da Rede Neural sem normalizar os atributos numéricos. Comente o resultado obtido, baseado nas métricas de avaliação disponíveis (acurácia, precision, recall, F1-Score, Matriz de confusão, etc.)
- 3.2 Agora normalize os dados de entrada e treine novamente o modelo MLP. Avalie os resultados obtidos e comente o efeito da normalização no treinamento da Rede Neural.

4 Mudança de configurações do modelo

- 4.1 Insira o conjunto de validação para o treinamento do modelo. Avalie o resultado obtido.
- 4.2 Modifique o tempo de treinamento (épocas) da Rede Neural. Escolha dois valores distintos (e.g. 1 e 1000 épocas) e avalie os resultados.
- 4.3 Modifique a taxa de aprendizado da Rede Neural. Escolha dois valores distintos (e.g. 0,001 e 0,1) e avalie os resultados.
- 4.4 Modifique a quantidade de neurônios na camada escondida da Rede Neural. Escolha dois valores distintos (e.g. 2 e 70 neurônios) e avalie os resultados.

5 Teste Livre

- 5.1 Faça novos testes para avaliar o desempenho da Rede Neural no problema designado. Use a técnica K-Fold (com $K = 10$) para analisar o resultado obtido.
- 5.2 Faça análises e novas implementações que você julgue importante para o seu trabalho. Não esqueça de explicar a motivação da análise realizada.

Referências

- [1] UCI Machine Learning Repository,
<https://archive.ics.uci.edu/ml/datasets/Ionosphere>

- [2] Sigillito, V. G., Wing, S. P., Hutton, L. V., & Baker, K. B. (1989). *Classification of radar returns from the ionosphere using neural networks*. Johns Hopkins APL Technical Digest, 10, 262-266.