

Relatório do Trabalho Computacional 1

Gabriel Ferreira Lima¹

¹Departamento de ciência da computação - IFCE, campus maracanaú

Abstract. *This article describes the activity carried out in Computational Work 2 of the Applied Computational Intelligence course. The task required the implementation of an AI model using Grid-Search and it allowed the use of libraries that implement models such as sklearn.*

Resumo. *Esse artigo descreve a atividade realizada no Trabalho Computacional 2 da disciplina de Inteligência Computacional Aplicada. Foi requerida a implementação de um modelo de IA que utilize Grid-Search e foi permitido o uso de bibliotecas que implementam os modelos como o sklearn.*

1. Introdução

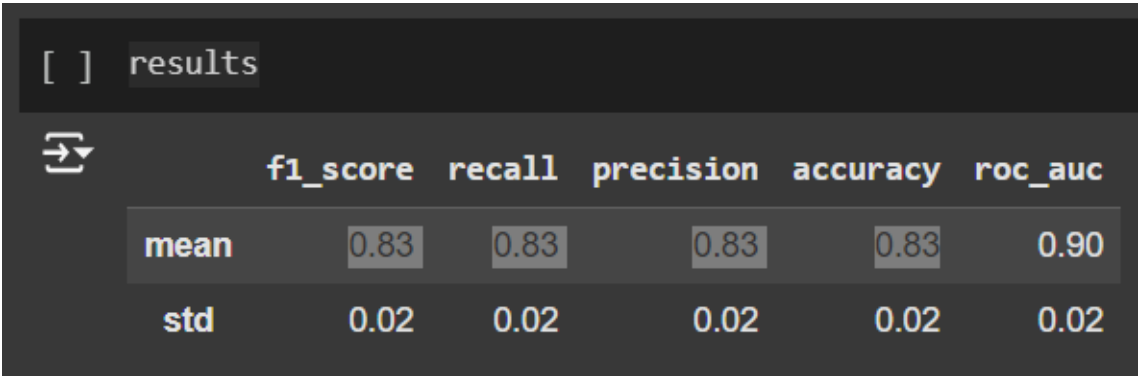
A atividade consistia em criar um modelo que realize otimização com o grid-search e preveja um conjunto de dados sobre diagnóstico de doença de alzheimers. Para esse trabalho foi utilizado o modelo svc (C-Support Vector Classification) do Scikit-Learn. O histograma de hyperparamêmetros analisados com o grid-search foi:

'C': [0.1, 1, 10, 100], 'kernel': ['linear', 'rbf', 'sigmoid']

O modelo foi treinado e testado 10 vezes, cada iteração coletando os dados de F1-Score, recall, precisão, acurácia e roc-auc. Ao final das 10 iterações os resultados foram agrupados, calculando a média e o desvio padrão.

2. Apresentação das Métricas

O F1-Score, recall, precisão e acurácia retornaram como valor 0.83 enquanto o roc-auc retornou com valor 0.90. Todos os valores tiveram um desvio padrão de aproximadamente 0.02, ou seja, não há severa variação dos resultados.



	f1_score	recall	precision	accuracy	roc_auc
mean	0.83	0.83	0.83	0.83	0.90
std	0.02	0.02	0.02	0.02	0.02

3. Discussão dos Resultados

Os resultados obtidos mostram que o modelo é útil para uma avaliação auxiliar, onde um médico poderia usar os resultados do modelo como base para identificar casos. Entretanto, as taxas de erro apresentaram-se altas demais para o contexto da medicina, que exige taxas de erro baixíssimas

Julgo que meu modelo apresentou resultados insatisfatórios devido às limitações na atuação do grid-search pelo demasiado tempo exigido pelo GridSearchCV e pela falta de uma limpeza de dados mais minuciosa.

4. Interpretação no Contexto da Saúde

Tal modelo nunca poderia ser usado isoladamente para casos de resultado negativo devido à gravidade de um potencial falso negativo. Ademais, não aconselho seu uso para o caso de resultado positivo pois a taxa de erro do modelo ainda é grande demais.

Entretanto, esse modelo pode muito bem ser usado para avaliação precoce e tem o potencial de salvar muitas pessoas por causa disso.