|  |  |  |  |
| --- | --- | --- | --- |
| **Processo Seletivo** | 00122/2024 - Pesquisador I - Inteligência Artificial | **Etapa** | Estudo de caso |
| **Entidade** | SENAI | **Data** | 11/02/2024 |

Dados a serem preenchidos pelo Candidato(a):

|  |  |  |  |
| --- | --- | --- | --- |
| **Nome Completo** | **Jonas Krause** | | |
| **E-mail** | [**profjonaskrause@gmail.com**](mailto:profjonaskrause@gmail.com) | **CPF** | 006.393.169-99 |

Relatótrio da etapa prática do processo seletivo 00122/2024 - Pesquisador I - Inteligência Artificial.

O código e arquivos estão disponíveis no GitHub (<https://github.com/jonaskrause/SENAI-ProvaPratica>)

A estrutura do projeto está dividida em código chamado “code\_ProvaPratica.py”, e pastas “00-relatorios”, “01-results”, e “02-modelos”.

Para a análise individual dos Dados\_1.npy:

* KNN melhor acurácia com validação cruzada: 0.382025

Para a análise individual dos Dados\_2.npy:

* Random Forest melhor acurácia com validação cruzada: 0.371

Para a análise individual dos Dados\_3.npy:

* KNN melhor acurácia com validação cruzada: 0.41265

Para a análise individual dos Dados\_5.npy:

* Random Forest melhor acurácia com validação cruzada: 0.2135

Para a análise dos Dados\_1, Dados\_2, e Dados\_3 agrupados:

* KNN melhor acurácia com validação cruzada: 0.423975

O arquivo “Dados\_4.npy” possiu somente a constante 50, mostrando que o sensor não fez uma leitura representativa, por isso, esses dados não estão sendo considerados nos experimentos.

O análise do arquivo “Dados\_5.npy” apresentou resultados poucos significativos. Por isso, o experimento final considerou o agrupamento dos Dados\_1, 2, e 3.

Assim, concluo que o conjunto de dados “Dados\_3.npy” possui uma representatividade maior para a discriminação das classes no arquivo “Classes.npy”. Além disso, o conjunto de dados agrupados (Dados\_1, Dados\_2, e Dados\_3) apresentou o melhor treinamento para os modelos de Aprendizado de Máquina e o modelo de KNN alcançou o maior valor de acurácia quando utilizando esse grupo de dados.

**Solução**: KNN utilizando Dados\_1, 2, e 3 resultou em uma acurácia de 0.423975 (ver pasta de resultados).