





CURSO: <BÁSICO EM MACHINE LEARNING>

● **Atividade 04 (ATIV-04)**

- Tipo: Somativa.
- Tema: Projeto de classificação/segmentação de imagens.
- Conteúdo: Módulo 3.
- Participantes: Grupo.
- Etapa
 - ATIV-04-ET-01
 - Apresentar informações do dataset.
 - ATIV-04-ET-02
 - Apresentar métodos da literatura e propor ideias para métodos próprios.
 - ATIV-04-ET-03
 - Apresentar métodos da literatura e métodos próprios.
- Avaliação do aluno.
 - Objetivo: Avaliar desempenho do aluno em relação ao projeto de classificação/segmentação de imagens.
 - Nota: 0 a 3 supercrítico, 4 a 6 crítico, 5 a 7 razoável e 8 a 10 bom.
 - Critérios avaliados: realização de tarefas individuais e em grupo, organização do projeto e presença em reuniões com docentes/membros do projeto.
- Informações complementares:
 - A atividade é composta por um projeto que utiliza conhecimentos adquiridos durante o módulo 3.
 - Orientações e exemplos para realizar as etapas
 -  [Avanti Bootcamp ML - Etapa 1] Flowers_Classification.i...
 -  [Avanti Bootcamp ML - Etapa 2] Flowers_Classification.i...
- **AO CONCLUIR A ATIVIDADE: ENVIAR APENAS O LINK DO REPOSITÓRIO GITHUB (ESPECIFICAR A BRANCH) PÚBLICO.**

DESCRIÇÃO DA ATIVIDADE

ATIV-04-ET-01

- Objetivo
 - Apresentar informações do dataset.
- Informações esperadas
 - Integridade dos Arquivos
 - Verifique se todas as imagens listadas no arquivo de informações realmente existem no diretório de imagens e vice-versa.
 - Verifique se todas as imagens estão no mesmo formato, ex: JPEG, PNG, etc.
 - Consistência dos Metadados
 - Verifique se há valores ausentes nos metadados e como esses casos são tratados.
 - Verifique valores inconsistentes, por exemplo: dimensões de imagens fora do esperado.
 - Qualidade das Imagens
 - Identifique imagens corrompidas que não podem ser abertas ou processadas.
 - Distribuição das Classes
 - Verifique a distribuição das classes para identificar possíveis desequilíbrios que possam afetar a modelagem
 - Duplicatas
 - Identifique imagens duplicadas que possam enviesar os resultados.
 - Verifique duplicatas no arquivo de informações.
- Informações complementares
 - Elaborar uma apresentação para mostrar os resultados.
 - Apresentar o dataset de forma detalhada.
 - Verifique quais itens em **Informações esperadas** podem ser aplicados no dataset.

ATIV-04-ET-02

- Objetivo
 - Apresentar métodos da literatura e propor ideias para métodos próprios.
- Informações esperadas
 - Métodos da literatura



- Todos devem executar um ou mais códigos (exemplos de códigos podem ser encontrados em **Related Notebooks** no endereço do dataset) que resolvem o problema do projeto.
- Métodos próprios
 - Apresentar ideias que podem melhorar os métodos encontrados e testados da literatura.
 - Ideias para melhorar uma técnica
 - Otimização de Hiperparâmetros - alterar ou reduzir os hiperparâmetros.
 - Elaborar uma nova arquitetura mais leve para que o treino e o teste sejam rápidos.
 - Reduzir número de imagens no treinamento e manter o desempenho.
- Informações complementares
 - Elaborar uma apresentação para mostrar os métodos da literatura e próprios.

ATIV-04-ET-03

- Objetivo
 - Apresentar métodos da literatura e métodos próprios.
- Informações esperadas
 - Metodologia
 - dataset.
 - Métodos da literatura e os próprios de forma detalhada.
 - Métricas para avaliar os métodos.
 - Resultados
 - Comparar os resultados literatura e os próprios.
 - Apresentar tabelas/figuras para ilustrar métricas de avaliação por classe
 - Conclusões
 - Pequeno compilado dos resultados.
 - Contribuições do estudo.
 - Possibilidades de novos estudos.
- Informações complementares
 - Elaborar uma apresentação para mostrar os métodos e resultados.



LISTA DE PROJETOS

projeto 01

Assunto: Classificação de animais

link:

<https://www.kaggle.com/datasets/anthonytherrien/image-classification-64-classes-animal>

projeto 02

Assunto: Classificação de frutas e vegetais

link: <https://www.kaggle.com/datasets/kritikseth/fruit-and-vegetable-image-recognition>

projeto 03

Assunto: Classificação de resíduos recicláveis

link:

<https://www.kaggle.com/datasets/alistairking/recyclable-and-household-waste-classification>

projeto 04

Assunto: Segmentação de pólipos intestinais

link: <https://www.kaggle.com/datasets/balraj98/cvcclinicdb>

projeto 05

Assunto: Segmentação de lesões mamárias

link: <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>

projeto 06

Assunto: Segmentação de lesões de COVID-19 em regiões pulmonares

link:

<https://www.kaggle.com/datasets/maedemaftouni/covid19-ct-scan-lesion-segmentation-dataset>