



- Desenvolver um sistema inteligente para identificação automática de defeitos.
- Baseado em 31 variáveis extraídas de imagens.

Entendimento do Problema

- Classificação baseada em características geométricas e estatísticas.
- 7 categorias de defeitos: 6 tipos + 1 categoria 'outros'.

Análise Exploratória dos Dados

- Análise das distribuições e correlações entre variáveis.
- Exemplos: Distribuição da área e matriz de correlação.



- Tratamento de valores inconsistentes.
- Exclusão de colunas não relevantes.
- Normalização dos dados.

Modelagem Preditiva

- Algoritmos utilizados:
- Random Forest
- XGBoost
- LightGBM
- Avaliação com F1-Score macro, precisão e recall.



- Melhor modelo selecionado com base no F1-Score.
- Avaliação final realizada com conjunto de teste.

Conclusão

- Projeto demonstra a aplicação prática da ciência de dados na indústria.
- Próximos passos:
- Implantação do modelo em produção.
- Criação de dashboard com Streamlit.

Links do Projeto

- GitHub: github.com/Carregosa/projeto_final_cdia
- Notebook: drive.google.com/file/d/140TRs9Pa18pOdLsW fyols9SAYbuh5Uym

Distribuição da Área dos Defeitos



Matriz de Correlação entre Variáveis

