

# Interpretabilidade dos Modelos de Machine Learning

## Proposta de Bolsa/Estágio na Altice Labs



ID do Projeto	2024/25_N27
Departamento	SSO53
Proponente	Maria Manuel Castro
Data de validade	31-Dec-2024

### IDI - Áreas Chave

*Artificial Intelligence & Machine Learning*

### Tema / Título

Interpretabilidade dos Modelos de Machine Learning

### Contexto

A interpretabilidade dos modelos de machine learning é fundamental para a tomada de decisões informadas em muitos setores, incluindo telecomunicações, saúde e finanças. No entanto, muitos modelos, especialmente os mais complexos, são frequentemente considerados "caixas pretas", o que dificulta a compreensão dos resultados por parte dos usuários finais. Reconhecendo a necessidade de transparência e compreensão dos modelos, pretende-se que sejam exploradas e melhoradas as atuais técnicas de interpretabilidade que usamos na AlticeLabs nos nossos modelos de classificação, clustering e previsão de valores.

### Objetivos do Projeto

- Revisão do estado da arte na interpretabilidade dos modelos de machine learning;
- Avaliação das ferramentas existentes em termos de suas capacidades e limitações;
- Desenvolvimento ou adaptação de uma ferramenta que melhore a interpretabilidade dos modelos utilizados pela AlticeLabs, superando as limitações identificadas;
- Criação de diretrizes e melhores práticas para a implementação e uso da ferramenta de interpretabilidade em projetos internos;«.

### Aspetos Inovadores

- Desenvolvimento de métodos ou ferramentas que abordem especificamente os desafios de interpretabilidade;
- Integração de técnicas de visualização de dados avançadas para melhorar a compreensão dos modelos;
- Proposição de uma abordagem adaptável que possa ser facilmente personalizada para diferentes tipos de modelos e casos de uso dentro da AlticeLabs.

### Ferramentas a utilizar

- Linguagens de programação: Python;
- Bibliotecas de Machine Learning: scikit-learn, TensorFlow, PyTorch, etc;
- Ferramentas de interpretabilidade: SHAP, LIME, etc;
- Software de visualização de dados: Matplotlib, Seaborn, Plotly, etc.

### Referências Bibliográficas

- A definir.

### Atividades

- Elaboração do estado da arte;
- Avaliação das Ferramentas Existentes;

- Desenvolvimento de nova Ferramenta/Abordagem;
- Teste e Validação;
- Documentação e elaboração de Diretrizes;
- Coleta de feedback e realização de ajustes;
- Elaboração do relatório final.

## Competências Chave Requeridas

- Entendimento sólido dos princípios de machine learning, incluindo classificação, clustering e regressão;
- Linguagem de programação: Python;
- Capacidade Analítica:
- Assertividade na discussão de conceitos e tecnologias.

## Orientador (nome e e-mail)

Maria Manuel Castro - [maria-m-castro@alticelabs.com](mailto:maria-m-castro@alticelabs.com)

Para concorrer podes enviar a tua candidatura, envia e-mail para o Programa GENIUS: [genius@inova-ria.pt](mailto:genius@inova-ria.pt)