



---

## *Elementos de Inteligência Artificial e Ciência de Dados*

### Trabalho Prático 2

A avaliação de imóveis é uma prática essencial no mercado imobiliário, influenciando decisões importantes para compradores, vendedores, instituições financeiras e investidores. Este processo visa determinar o valor justo de uma propriedade com base numa série de fatores, incluindo as suas características físicas, localização, condições de mercado e tendências económicas. Apesar de ser um processo complexo, a avaliação de imóveis baseia-se em dados objetivos e mensuráveis, sendo portanto passível de ser automatizado através do uso técnicas de Ciência de Dados, que permitem desenvolver modelos preditivos precisos que estimam o valor de um imóvel com base num conjunto de variáveis explicativas.



Neste trabalho prático pretende-se usar modelos de aprendizagem automática para realizar a avaliação do potencial preço de venda de um imóvel com base num conjunto de características físicas e de localização do imóvel. Para isto, será necessário realizar as seguintes fases:

- **Recolha e Limpeza de Dados:** Esta fase consiste na recolha de dados sobre transações de imóveis que devem ser agregados de forma a construir um conjunto de dados contendo variáveis relacionadas com as características físicas e de localização do imóvel, assim como o preço e data de venda desse imóvel.
- **Limpeza e Preprocessamento de Dados:** Esta fase pretende preparar o conjunto de dados para o uso de métodos de aprendizagem supervisionada, sendo portanto essencial realizar a correção ou remoção de exemplos contendo *outliers* ou valores em falta, assim como realizar a normalização das escalas dos dados.
- **Construção do Protocolo de Treino e Avaliação:** Esta fase tem como objetivo a definição dos dados que serão usados durante a aprendizagem e avaliação, assim como a função de custo e as métricas de avaliação do modelo.
- **Treino e Avaliação de Modelos de Aprendizagem Automática:** Realização do processo de aprendizagem e análise crítica sobre o desempenho do mesmo.

No final, deve ser elaborado um relatório contendo os detalhes do processo de implementação de cada uma das fases, e uma apresentação contendo um sumário sobre os principais resultados obtidos.

## Elementos a Entregar:

No final do projeto, devem ser entregues através do Moodle os seguintes elementos:

- Repositório Git contendo o código Python usado nas diferentes fases do projeto
- Relatório organizado de acordo com as fases do projeto descrevendo as técnicas usadas em cada uma delas, a justificação para o seu uso, e os resultados obtidos. O relatório deve obrigatoriamente fazer referência à função ou script usada para implementar os processos descritos no mesmo, não sendo necessário incluir blocos de código no mesmo.
- Apresentação com um sumário dos principais resultados e conclusões obtidos sobre o desempenho do(s) método(s) implementado(s).

---

## Avaliação:

A classificação final do projeto será obtida através da seguinte fórmula:

$$Nota_{TP2} = 0.9A + 0.1B,$$

sendo  $A$  o repositório e relatório desenvolvido, e  $B$  a apresentação dos resultados obtidos.

O repositório e o relatório serão avaliados de acordo com as técnicas usadas em cada uma das fases, tendo em conta a sua correta aplicação e complexidade. Cada fase contribui para a classificação final de acordo com as seguintes percentagens:

- Recolha e Integração de Dados [5%]
- Limpeza e Preprocessamento de Dados [15%]
- Construção do Protocolo de Treino e Avaliação [20%]
- Treino e Avaliação de Modelos de Aprendizagem Automática [50%]

O desempenho na apresentação dos resultados será medido pela forma como os mesmos são comunicados, em particular se é fácil perceber quão bem o modelo consegue prever o preço de um imóvel.