Progetto Analisi: Sovrapprezzo dei laptop

Basata sul dataset <https://www.kaggle.com/datasets/asinow/laptop-price-dataset/data>

Creatore: Tufano Leonardo

Matricola: 05121 17411

Contenuti:

[**1 Introduzione:**](#_8q2hxl6ccun2) **1**

[1.1 Obiettivo del problema](#_9875xin0oytv) 1

[1.2 Obiettivo del progetto](#_y2xseibrk8zz) 2

[**2 Analisi dei dati**](#_n43fv9gr20s) **2**

[2.1 Descrizione dataset](#_v7ipnucapqvm) 2

[2.2 Formato caricamento dati](#_at6r5no5nr84) 4

# 1 Introduzione:

Questo progetto punta alla risoluzione di un problema di classificazione binaria, nel contesto della valutazione degli strumenti tecnologici. In particolare l’obiettivo è determinare se un laptop è sovrapprezzo o meno. Questo progetto può essere utile per la scelta di acquisto sia a livello personale sia a livello aziendale.

## 1.1 Obiettivo del problema

L’obiettivo principale è classificare i laptop in due gruppi:

* 1: sovrapprezzo
* 0: non sovrapprezzo

Per fare ciò viene calcolato prima di tutto la qualità del laptop, con cui poi si potrà andare a vedere il sovrapprezzo, la qualità del laptop è calcolata sia numericamente che categorizzata in tre gruppo:

* 1: qualità alta
* 2: qualità media
* 3: qualità bassa

## 1.2 Obiettivo del progetto

1. Analizzare e processare i dati per garantire la qualità dei dati
2. Creare un modello predittivo utilizzando una tecnica di machine learning supervisionato
3. Valutare le performance del progetto grazie ad alcune metriche come accuracy, F1-score, precision, recall

# 

# 2 Analisi dei dati

## 2.1 Descrizione dataset

Il dataset contiene dati sui laptop compresi la loro grandezza, prezzo, brand, Ram (in Giga), modello di GPU, durata della batteria (in ore), sistema operativo ecc…

La tabella seguente

| Nome Colonna | Definizione |
| --- | --- |
| Brand | Indica il brand di appartenenza del laptop |
| Processor | E’ il processore utilizzato, uno degli elementi usati per la valutazione |
| RAM | La ram in GB del laptop, uno degli elementi usati per la valutazione |
| Storage | La memoria del laptop |
| GPU | Indica la gpu del laptop, uno degli elementi usati per la valutazione |
| Screen size | Indica la grandezza dello schermo misurata in Inch |
| Resolution | Risoluzione dello schermo del laptop |
| Battery Life | Indica la durata della batteria del laptop in ore |
| Weight | Indica il peso del laptop in Chilogrammi |
| Operating System | Indica il sistema operativo |
| Price | Indica il prezzo di vendita del laptop in Euro |

## 

## 2.2 Formato caricamento dati

Il dataset è fornito in formato CSV, un formato facile da importare su Python e gestito grazie ad alcune librerie come pandas. In questo progetto il dataset è caricato nel seguente snippet di codice:

def load\_data(filepath):

"""Carica il dataset da un file CSV."""

return pd.read\_csv(filepath)

La variabile “filepath” è definita nella funzione main:

def main():

"""Funzione principale per eseguire il flusso di elaborazione e salvataggio."""

file\_path = 'Dataset/laptop\_prices.csv' # Percorso dataset

df = load\_data(file\_path)