



**Ok, il prezzo è giusto!**



**Il crescente problema del consumismo.**

**Introduzione di una IA per decisioni di Smart Shopping.**

**Riduzione di e-waste e consumismo.**

**Conclusioni.**

---

# Il crescente problema del consumismo

Il mondo stà gradualmente evolvendo verso una società sempre più consumer-oriented. Con l'avvento della tecnologia e di internet, è diventato più facile che mai comprare prodotti e servizi con il semplice click di un tasto. Questo ha portato ad un incremento del consumismo, ovvero l'eccessivo shopping di prodotti e servizi.

Il problema del consumismo è che questo porta inevitabilmente ad una produzione maggiore di rifiuti e ad un utilizzo eccessivo delle risorse. Questo ha avuto un impatto negativo sia sull'ambiente che sui risparmi delle persone. Ha portato, statisticamente, anche ad un incremento del debito per quelle le persone che cercano a tutti i costi di stare sempre al passo e di avere l'ultima uscita.



---

# Introduzione di una IA per decisioni di Smart Shopping

Per combattere il problema del consumismo, ho sviluppato una IA che può predire un prezzo “vincente” per introdurre sul mercato uno smartphone usato. Questo tool utilizza una tecnica di regressione per fornire una stima di prezzo che permetta di vendere il dispositivo in un lasso di tempo ragionevole e con la minor perdita possibile in termini economici.

Con l'utilizzo di questo tool, l'utente può prendere decisioni più consapevoli quando si tratta di gettare via uno smartphone. Questo può aiutare a ridurre il numero di rifiuti, poichè le persone sarebbero incoraggiate a vendere, con un prezzo che si è già provato essere vincente senza doversi preoccupare di andare per tentativi al rialzo o al ribasso.





## **Specifica PEAS**



# Performance, Environment, Actuators and Sensors

P. erformance: prestazione delle previsioni

E. nvironment: ambiente operativo

A. ctuators: azioni attraverso le quali l'agente può operare

S. ensors: come l'agente raccoglie le informazioni



## Specifiche dell'ambiente

- Completamente osservabile



## Specifiche dell'ambiente

- Completamente osservabile
- Deterministico





## Specifiche dell'ambiente

- Completamente osservabile
- Deterministico
- Statico



## Specifiche dell'ambiente

- Completamente osservabile
- Deterministico
- Statico
- Discreto



## Specifiche dell'ambiente

- Completamente osservabile
- Deterministico
- Statico
- Discreto
- Episodico



## Specifiche dell'ambiente

- Completamente osservabile
- Deterministico
- Statico
- Discreto
- Episodico
- Singolo agente



# Il Dataset

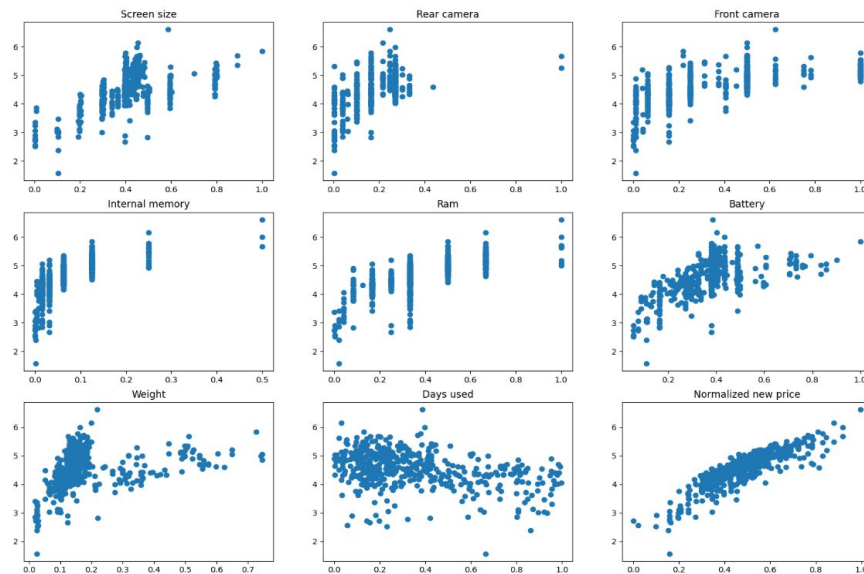
- <https://www.kaggle.com/datasets/ahsan81/used-handheld-device-data>

Il dataset trovato non fornisce, sfortunatamente, tutte le informazioni necessarie a risolvere il problema presentato. Manca infatti, per dirne una, la quantità di tempo dall'immissione del dispositivo in un sito di vendita alla data d'acquisto.

Questa è una demo, ma il modello potrebbe essere adattato qualora si riuscissero a trovare dati migliori.

# Data Understanding

Per capire quale modello potrà approssimare meglio i nostri dati si guarda la distribuzione di questi su un grafico.



I dati sembrano seguire una distribuzione lineare.



# Test di Pearson

Una ulteriore verifica potrebbe essere quella del test statistico di Pearson.

- +1 indica una correlazione lineare **positiva** perfetta
- -1 indica una correlazione lineare **negativa** perfetta
- 0 correlazione nulla



# Test di Pearson

Una ulteriore verifica potrebbe essere quella del test statistico di Pearson.

- +1 indica una correlazione lineare **positiva** perfetta
- -1 indica una correlazione lineare **negativa** perfetta
- 0 correlazione nulla

Le features risultano essere tutte linearmente e positivamente correlate tra loro, ad eccezione di quella che descrive i giorni di utilizzo che segue una correlazione sempre lineare ma negativa.





# Data Preparation

202 Entry Mancanti -> Data Imputation con strategia **Most-Frequent**

Variabili Categoricali -> Scalate con **OneHotEncoding**

Trattandosi di un Dataset costruito a scopo didattico non sono state necessarie ulteriori trasformazioni come **Data Balancing** o **Feature Selection**.



# Modelling

Una volta appurata la natura lineare dei dati, si può procedere a scegliere il migliore algoritmo di regressione

- Elastic Net
- Decision Tree Regressor
- Linear Regression



# Elastic Net

- Una combinazione della regressione **Lasso e Ridge**
- Meno suscettibile all'overfitting rispetto alla regressione Lasso
- In grado di gestire un gran numero di variabili indipendenti, anche se alcune di esse sono correlate



# Decision Tree Regressor

Utilizza un albero di decisione per effettuare previsioni sulla base dei valori di alcune caratteristiche. L'albero di decisione è costruito mediante un processo di suddivisione ricorsiva dei dati in base al valore delle caratteristiche. Ogni nodo interno dell'albero rappresenta una caratteristica e ogni foglia rappresenta un valore predetto.



# Linear Regression

È un metodo statistico usato per la relazione tra una variabile indipendente ( $x$ ) e una o più variabili dipendenti ( $y$ ). Viene utilizzato per costruire una linea di regressione, che rappresenta il modello matematico che descrive la relazione tra  $x$  e  $y$ .

Ovviamente risulta essere il miglior modo di approssimare i nostri dati.



# Validazione

La cross-validation è una tecnica utilizzata per valutare le performance di un modello di Machine Learning su un insieme di dati di test indipendente.

Consiste nel dividere i dati di input in più parti (“fold”), addestrare il modello su una parte di essi e testarlo sull’altra parte. Per questo progetto si è scelto di optare per una 10-fold cross validation.



# Valutazione dei modelli

Una volta addestrati i modelli, sono necessarie delle metriche oggettive per verificarne la bontà, quelle utilizzate sono:

- MAE
- MSE
- $R^2$



# Valutazione dei modelli

Una volta addestrati i modelli, sono necessarie delle metriche oggettive per verificarne la bontà, quelle utilizzate sono:

- MAE
- MSE
- $R^2$

Come facilmente intuibile, i risultati migliori si ottengono con la regressione lineare.





**Grazie per l'attenzione.**