

Artificial Intelligence

Talent Week 2025

Sentiment Evaluation

Challenge

bip.xTech

Introduzione

Dato un **database di recensioni** di Tripadvisor di hotel viene richiesto, tramite tecniche di NLP, di effettuare una ***sentiment analysis e ranking***.

Nello specifico, il tuo obiettivo è costruire un modello che permetta di effettuare in modo automatico una **classificazione binaria** delle recensioni tra positive e negative.

Il database

Viene fornito un **dataset** di *training* (**Allegato 1 - data_classification.xlsx**) non spartito in *train* e *test* di circa 350.000 **record** di recensioni testuali suddivise in **2 classi**:

- **Positive** quando la variabile assume il valore 1;
- **Negative** quando la variabile assume il valore 0.

Viene inoltre fornito un altro allegato (**Allegato 2 - data_evaluation.xlsx**) composto da **86.168 entry** di soli testi paragonabili a quelli del dataset di *training*. Questo ulteriore dataset dovrà essere utilizzato per **fare inferenza**.

I task

1. Splittare il dataset di *review* in **train** e **test set**.
2. Addestrare un **modello di classificazione binaria** e misurarne le prestazioni.
3. Elaborare il **grafico ROC** (deve avere sulle ordinate il TP rate e sulle ascisse il FP rate) e calcolare l'**AUC** rispetto al grafico rappresentato.
4. Costruire la **Matrice di Confusione**.
5. Addestrare il modello sul dataset completo di *review* a disposizione.
6. **OUTPUT CONCLUSIVO:** restituire il file **Allegato 2 - data_evaluation.xlsx** arricchito di una colonna denominata “**promotore_pred**” che contenga le predizioni del modello. Il file arricchito deve essere denominato **output_pred_nome_cognome.xlsx** (es. **output_pred_mario_rossi.xlsx**).

Con cosa lavorare

LINGUAGGI E SOFTWARE

Si sviluppi codice utilizzando **Python**, dato l'alto livello di standardizzazione della challenge.

LIBRERIE

Nessuna limitazione sulle librerie che puoi utilizzare. Puoi sbizzarrirti!

Formato della soluzione

Cartella **.zip** denominata **nome_cognome.zip** (es. **mario_rossi**) che contenga all'interno:

1. un'ulteriore cartella **.zip** con all'interno:
 - un file **main_train.py** che, quando viene lanciato, esegua in automatico il *training* di un modello di *machine learning* (ovvero qualsiasi tipologia o combinazione di tecniche AI) che riceve in input il dataset di addestramento (**Allegato 1 - data_classification.xlsx**); per il modello dovranno già essere state fatte analisi di **Grid Search** e perciò gli iperparametri utilizzati dovranno essere fissi;
 - un file **main_eval.py** che, quando viene lanciato, esegua in automatico l'inferenza sul dataset di valutazione (**Allegato 2 - data_evaluation.xlsx**) e produca in output un file denominato **output_pred_nome_cognome.xlsx** (es. **output_pred_mario_rossi.xlsx**), esattamente uguale ed ordinato come l'Allegato 2 - data_evaluation.xlsx, ma con in più una colonna nominata “**promotore_pred**” dove viene salvato il risultato dell'inferenza;

Formato della soluzione

- tutti i file e i dati necessari a runnare il `main_train.py` e il `main_eval.py` in automatico, nonché eventuali notebook per la Grid Search;
2. il **file di output .xlsx** denominato **output_pred_nome_cognome.xlsx** contenente le predizioni del modello;
 3. un file **nome_cognome_model.pickle** (es. **mario_rossi_model.pickle**) oppure **nome_cognome_model.json** (es. **mario_rossi_model.json**) dove viene salvato il modello già addestrato.

Valutazione delle soluzioni

La commissione appurerà la correttezza del formato della soluzione e il suo funzionamento sulla base delle richieste della challenge.

Dopo di che, utilizzerà **output_pred_nome_cognome.xlsx** per la valutazione delle *performance* del modello, prendendo come parametro l'**F1-score (sklearn.metrics.f1_score)**.

Il punteggio ottenuto da tutti i partecipanti verrà **normalizzato e parametrizzato in 30-esimi**. Per i **migliori 5-10 partecipanti** la commissione potrà procedere ad un'analisi attenta del codice e della pulizia e potrà valutare di aggiungere discrezionalmente dei punti **per merito**.

Eventuali irregolarità ravvisate verranno punite con l'**invalidità della prova**.

Punteggio Talent Week

Il **punteggio totale** ottenuto da ciascun candidato alla Talent Week è calcolato come **media ponderata**, espressa in **30-esimi**, dei punteggi conseguiti nelle due prove previste: la Online Challenge e l'**Hard Skills Test**.

I pesi assegnati per il calcolo sono pari a **7 per la Online Challenge** e **3 per l'Hard Skills Test**.

Artificial Intelligence Talent Week 2025

**Hai tempo fino a venerdì 5
dicembre alle ore 10 per
completare la Challenge!**

bip. xTech