

ANALYZING



HOTEL REVIEWS

Introduction

Problem Statement

Hotels receive thousands of reviews, but manually analyzing them is inefficient.

Why Businesses Care:

- Customer feedback helps improve services/experience.
- Better customer experiences/feedback links to more bookings.

Project objective:

- Use NLP to analyze customer reviews, classify sentiment, and provide insights.



Project Overview

Project Goal

- Analyze [hotel reviews](#) to classify sentiments and provide actionable insights.

Key Result

- Model **89%** accuracy.

Business Impact

- Enables these groups of hotels to make data-driven decisions to enhance guest satisfaction.

NLTK Approach

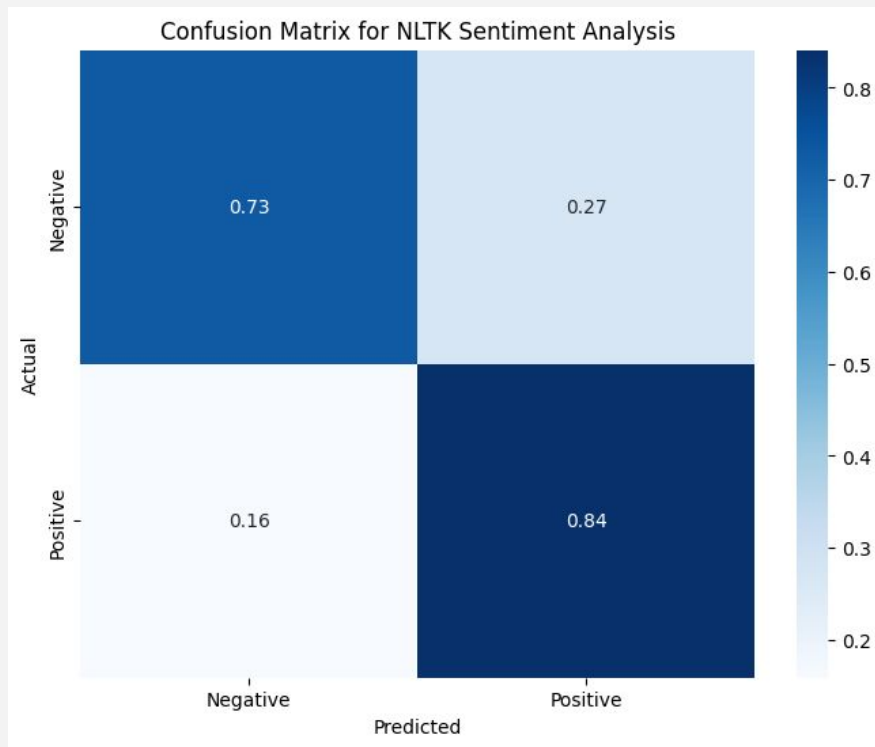
Natural Language Toolkit (NLTK)

- Used NLTK Vader
`SentimentIntensityAnalyzer()` to get a sentiment score and a predicted sentiment.

```
→ Accuracy: 0.7859
Classification Report:
              precision    recall  f1-score   support

   NEGATIVE      0.82      0.73      0.77     38690
   POSITIVE      0.76      0.84      0.80     38690

 accuracy              0.79              0.79     77380
 macro avg              0.79      0.79      0.79     77380
 weighted avg           0.79      0.79      0.79     77380
```



DistilBERT Approach

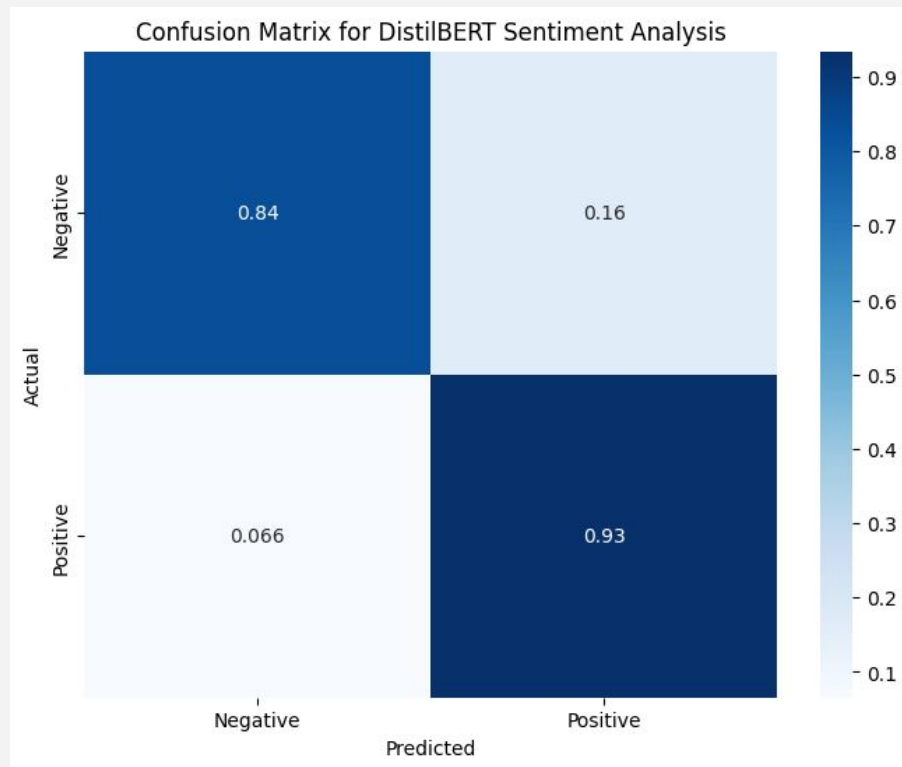
HuggingFace Distilled version of BERT Model

- Used model
distilbert-base-uncased-finetuned-sst-2-english to get a sentiment score and a predicted sentiment.
- Performed better than NLTK.

```
🔗 Accuracy: 0.8848
Classification Report:
              precision    recall  f1-score   support

   NEGATIVE       0.93      0.84      0.88     38690
   POSITIVE       0.85      0.93      0.89     38690

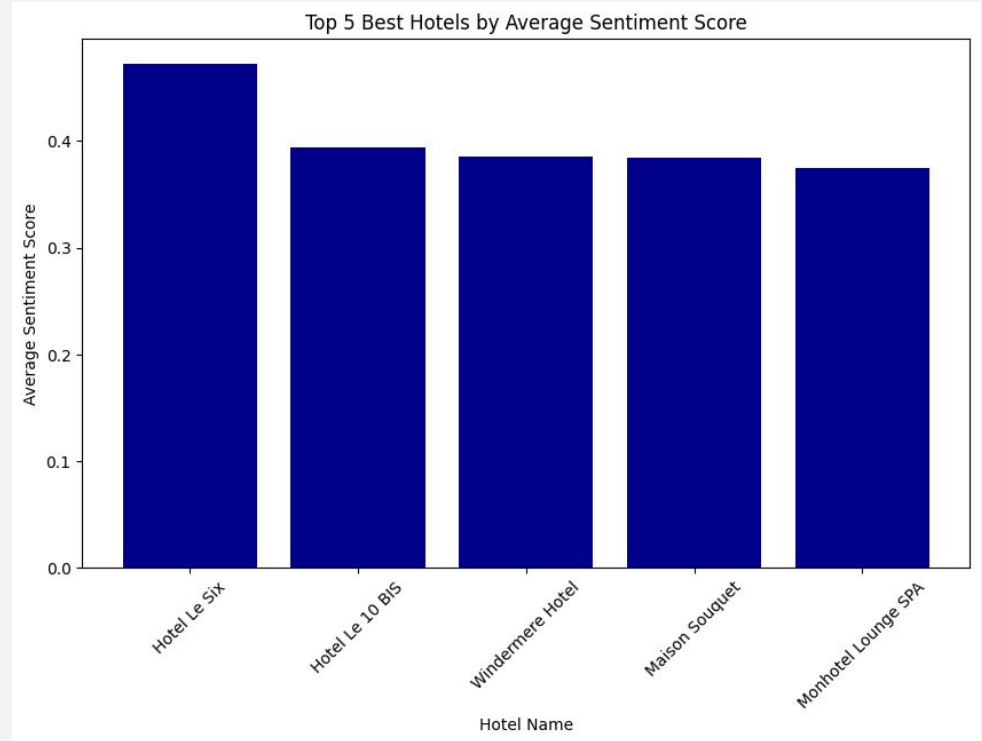
 accuracy              0.88      0.88      0.88     77380
 macro avg              0.89      0.88      0.88     77380
 weighted avg           0.89      0.88      0.88     77380
```



Top 5 Hotels

Based on Sentiment from DistilBERT

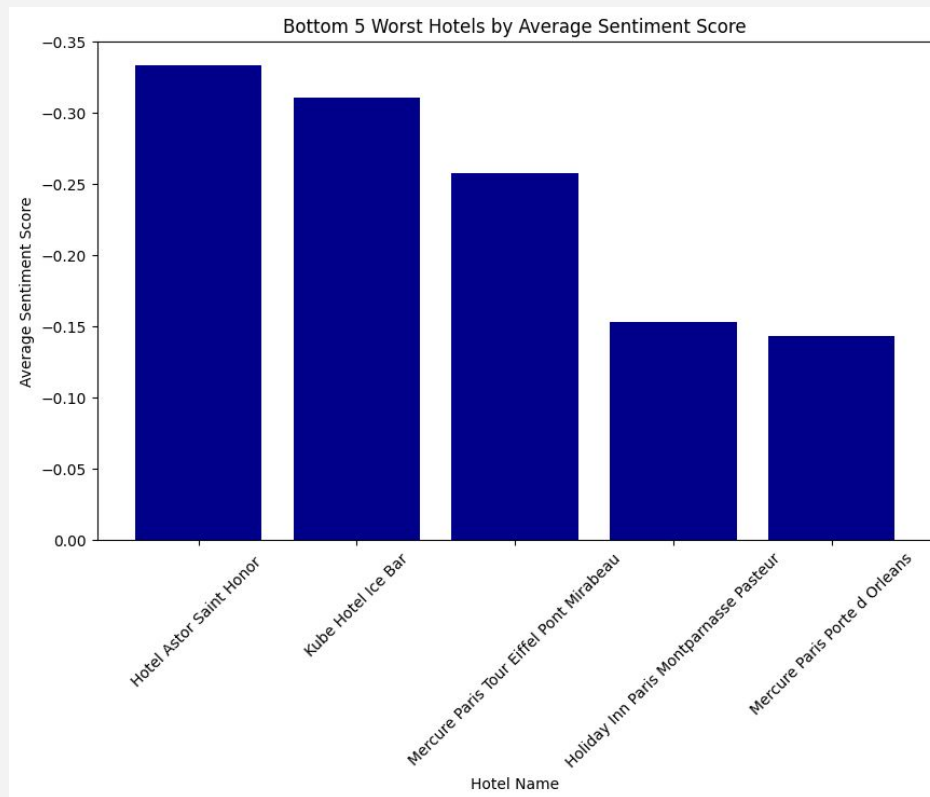
Based on this dataset, Hotel Le Six has the best sentiment score.



Bottom 5 Hotels

Based on Sentiment from DistilBERT

Based on this dataset, Hotel Astor Saint Honor has the worst sentiment score.

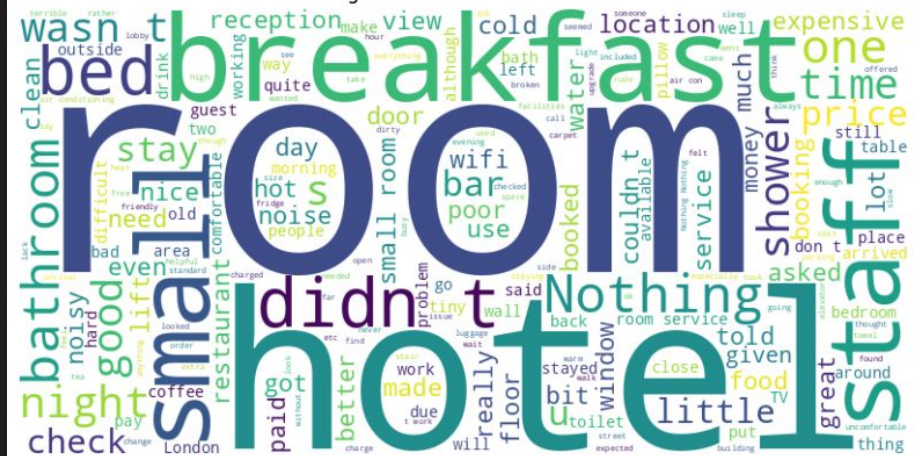


Common Words for Reviews

Positive Reviews Word Cloud



Negative Reviews Word Cloud



→ Questions?

