🏨 Text Mining Hotel Reviews: How I Used Sentiment Analysis to Rate Guest Experiences

Have you ever wondered what hidden patterns lie beneath the sea of hotel and restaurant reviews on the internet? That's exactly what sparked my interest—and led me on a journey to apply text mining and sentiment analysis to real-world customer feedback from over 100 hospitality businesses.

📌 The Motivation

Online reviews are powerful. They shape opinions, influence bookings, and define brand reputations. However, manually going through thousands of reviews is both inefficient and prone to bias. That's where Natural Language Processing (NLP) comes in—particularly, sentiment analysis.

I wanted to uncover how guests truly feel about their experiences—positive, negative, or neutral—and whether machine learning could reliably detect and classify those emotions.

🧰 Tools & Technologies Used

- **Programming Language**: Python
- **Libraries**: NLTK, Scikit-learn, Pandas, Matplotlib
- **Models Applied**:
    - Naive Bayes Classifier
    - Random Forest Classifier
- **Preprocessing Techniques**: Tokenization, stop word removal, lemmatization, vectorization (TF-IDF)

🔍 Data Preparation

The dataset consisted of thousands of customer reviews across a wide variety of hotels and restaurants. To prepare this raw text for analysis:

- I cleaned the text by removing punctuation, numbers, and stopwords.
- I converted the text to lowercase and lemmatized each word to its root form.
- Then, I transformed the cleaned text into numerical features using the TF-IDF method, enabling machine learning models to make sense of it.

📊 Exploratory Data Analysis

Before jumping into model training, I explored:

- Most frequently used positive and negative words
- Distribution of review sentiments
- Length and structure of reviews per sentiment class

This helped me identify common patterns—for instance, how shorter reviews tended to skew negative or how specific words like “comfortable,” “friendly,” and “dirty” heavily influenced sentiment scores.

🤖

✅ Conclusion

This project was a practical example of how data science can transform unstructured data into valuable business insights. For anyone looking to break into NLP or hospitality analytics, sentiment analysis on real-world reviews is a great place to start.