



By:

Rana Alturki

Reem Bin Zeraiban

Introduction:

As part of the Vision 2030 initiative, where the Kingdom aims to attract 100 million visitors annually, which requires a significant increase in the stock of hotel rooms, as the national transformation programs within the framework of developing the tourism sector and national heritage are keen to make the Kingdom a regional and global tourist destination, and the goal of this project is Building unsupervised machine learning models for Arabic Natural Language Processing (NLP) that determine whether a text revision is a positive revision or a negative revision. This project will help hotels to automatically identify and group text review category to improve their services.

Problem Statement:

The Saudi Ministry of Tourism (MT) is a governmental body concerned with tourism. It is focused on encouraging and supporting domestic and external tourism through sponsoring and conducting tourism events across the country and overcoming obstacles to growth of tourism. DS Company will analyze whether the text review is a positive review or a negative review. This project will help hotels to automatically identify and group text review category to improve their services.

Data Description:

We will be using the " To analyze user reviews of hotels " Dataset from Kaggle: the dataset has 2000 numbers of rows and 4 feature.

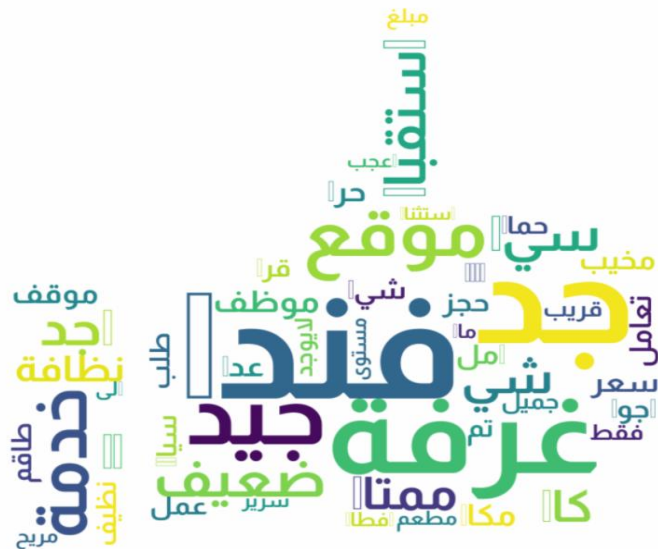
Tools:

Technologies: python, Jupyter notebook

Libraries: pandas, numpy, matplotlib, seaborn, LogisticRegression , WordCloud, Stopword, textblob, tqdm, arabic_resaper, nltk, farasa, sklearn, pyLDAvis.

Word cloud:

- What are the most words positive repeated?



- **What are the most words negative repeated?**



TOPIC MODELING (LDA):

Topic 0	Service	عدم , جيد , استقبال , موقع , ضعيف , جد , حجز , موظف , وجد , لا يوجد , ان , غرفة , فندق , سيارة , موقف
Topic 1	Employee	قرب , موقع , عدم , سيء , مكان , خدمة , وجد , جيد , كان , نظافة , غرفة , استقبال , فندق , مخيب , أمل
Topic 2	Order	نظافة , راحة , فندق , لا شيء , جيد , طاقم , هدوء , عمل , نظافة , موقع , راءع , استثناء , فاي , واي , شيء
Topic 3	Description	أعجب , موقع , كان , سيء , مطعم , شيء , لا يوجد , استقبال , غرفة , فندق , طاقم , مكان , جد , ممتاز , خدمة
Topic 4	Reservation	وجود , موقف , عدم , مكان , سيء , خدمة , وجد , حرم , موقع , نظافة , جيد , ضعيف , غرفة , فندق , جد
Topic 5	Parking	ان , عدم , شيء , كان , أعجب , طلب , موظف , خدمة , جد , استقبال , جيد , سيء , شيء , غرفة , فندق

Recommendation System:

Similar user:

User 10 most similar user 3.

Similar hotel:

The hotel (العين روتانا) most similar hotel (أداجيو البستان).

Recommendation:

Because the user liked Hotel (اكاس ان للشقق الفندقية) , We recommend (أبها كراون للأجنحة) and (فندق ميلينيوم العقيق المدينة).

Conclusion:

Machine Learning to analyze hotel reviews Because the user liked Hotel (اكاس ان للشقق الفندقية)

We recommend (أبها كراون للأجنحة الفندقية) and (فندق ميلينيوم العقيق المدينة).