Project Proposal

Student Names:

1. Abhishek Konduri
2. Nitish Joshi
3. Rohan Kamat
4. Varun Jain

Team Name: Group 2

Project Title: AirBnb Data Analysis – Predictions and Recommendations

Project Description: In this project, we will be using the Seattle Airbnb Dataset. This dataset contains information about the listings and review details of homestays in Seattle, WA. Below is the list of objectives we propose to achieve/implement during the entire duration of the project.

- Perform natural language processing to predict the review scores based on features like descriptions, summaries, property type, room type, etc.
- Perform analysis based on the different prices of listings.
- Perform analysis and draw insights to identify regions which yield maximum returns after making an investment in the property of that region.
- Analysis to identify the seasonal pattern of prices.
- Sentiment analysis of the user reviews.
- Predicting the best time to plan a trip to get the deal in price.
- Build a recommender system to provide a great user experience to the user as well as predict the listing prices based on features like location, property, room types, etc.

These are some of the objectives that we hope to achieve during our entire project implementation phase.

Dataset link: https://www.kaggle.com/airbnb/seattle

Proposed Methodology/Techniques:

- Data Exploration
- Data Cleansing (noise, duplicate data, missing values, outliers, wrong data)
- Data Preprocessing

Proposed Algorithms for implementation:

We plan to use sklearn regression model for predicting the prices of new properties, predicting regions which can give maximum profit post investment, etc.

Apart from that, we plan to use and experiment using the below mentioned algorithms to test the accuracy of our models and develop efficient models.

- Decision Trees
- Linear Regression
- Random Forest Classifier
- SGD Classifier
- Logistic Regression

Tools:

- Python 3.0
- Jupyter Notebook
- sklearn