ANA502 Revised Proposal: Image classification for Oral cancer detection from clinical images

KOHEI NISHITANI

1. Introduction and Motivation

The number of international tourists visiting Japan has dramatically increased thus making the country a worldwide destination for travelers. The substantial economic expansion has created a serious deficit of available lodging particularly in locations with high tourist interests and urban centers. Airbnb and similar platforms have provided essential substitute accommodations by enabling users to book rooms in various sized facilities spread throughout Japan.

The increasing demand promotes new investment possibilities and innovative solutions in the short-term rental business. The proposed analysis focuses on Airbnb booking trends within Japan's specific area to identify marketplace requirements for customer demand and generate mission-critical business intelligence. The research aims to support strategic choices through data analytics and machine learning methods for shortened rental market investors and owners alongside hospitality entrepreneurs looking to excel in this business environment.

As part of my CPT, I’m exploring new business opportunities in the hospitality and property management space. This project aims to analyze Airbnb booking occupancy in a specific area in Japan, providing actionable insights that could support investment or operational decisions.

2. Research Question

What are the key factors that influence booking occupancy rates on Airbnb in Japan, and how can they inform strategic business opportunities?

3. Dataset Selection

The dataset used for this project includes Airbnb listings in a specific region in Japan, using [web scraping](https://apify.com/tri_angle/airbnb-scraper).

4. Methodology

4.1 Data Cleaning & Preparation

* Handle missing values, especially in review or availability columns.
* Feature engineering: average occupancy, distance from city center, extract qualitative info from meta data etc.

4.2 Exploratory Data Analysis (EDA)

* Visualize relationships between occupancy and features like price, location, season.
* Identify high-performing listings and commonalities.

4.3 Predictive Modeling

* Use regression models (e.g., Random Forest, XGBoost, Linear Regression) to predict occupancy rates.
* Cluster listings based on performance or pricing using k-means or DBSCAN.

4.4 Business Insights

* Recommend pricing strategies based on season or listing type.
* Identify underserved areas or segments for new business opportunities.

5. Expected and Anticipated Outcomes

* A model or framework to predict occupancy rates based on image impression, not listed written information
* Identification of high-opportunity zones for new Airbnb listings or investment
* Strategic insights for business owners to optimize operations and marketing
* Limitations may include data completeness or lack of booking confirmation records (vs. availability).

6. Conclusion & Next Steps

* This project aims to combine data analytics and business intelligence to uncover actionable trends in Airbnb occupancy in Japan. The outcomes can serve both hosts and real estate stakeholders by guiding decisions around pricing, listing optimization, and expansion.