Task Report

Prediction of Target Location Using RSS Signals

Introduction: Automatic user localization, particularly using WLAN fingerprinting, has gained significant attention due to its applications in various real-world scenarios. In this project, I aim to implement a machine learning algorithm to predict user location based on Received Signal Strength (RSS) signals collected from Access Points (Wi-Fi). I used a RandomForestRegressor ML model. The dataset used for this task consists of 103,584 Wi-Fi fingerprints collected over a period of 25 months in a library building at Universitat Jaume I, Spain.

Objective: The primary objective of this project is to develop and evaluate a predictive model capable of estimating user location (latitude, longitude, and altitude) based on the RSS signals from Wi-Fi access points.

Implementation:

1. **Data Preprocessing:** Before training the model, I pre-processed the dataset. This involves tasks such as data cleaning, handling missing values, feature scaling.
2. **Model Selection:** I explored various machine learning models suitable for regression tasks. Finally, I selected RandomForestRegressor model for the regression task.
3. **Training:** I trained the selected model using the training dataset provided (first.train.xlsx). I employ techniques such as checking Correlations, plotting pair plot.
4. **Testing:** Once trained, I checked the model's performance using the testing datasets provided for each month (test.month.xlsx). I calculated the average localization error (ALE) using the Euclidean formula and Root Mean Squared Error (RMSE) for each month's test data.
5. **Results Reporting:** The results are compiled into an Excel file, containing the average localization error and RMSE computed for each of the twenty-five months.

Challenges:

- **Data Quality:** Ensuring the quality and reliability of the dataset is crucial for accurate model training and evaluation.
- **Model Selection:** Identifying the most suitable model architecture and algorithms for the given task is challenging and require experimentation.
- **Overfitting:** Preventing overfitting while training the model is essential to ensure its generalizability to unseen data. Sometimes we need to drop some attributes from the training dataset.

Conclusion: The successful implementation of an accurate and efficient localization model will have significant implications for various applications, including indoor navigation, asset tracking, and location-based services. By addressing the outlined objectives and challenges, I tried to develop a robust solution that can predict user location based on RSS signals with less error.

Project Timeline:

- Data Preprocessing: 2 days
- Model Selection and Training: 3 days
- Testing and Evaluation: 2 days
- Results Reporting and Documentation: 1 day

Acknowledgments: I acknowledge the resources provided by the Universitat Jaume I, Spain, and IIT Patna for providing the dataset and task specifications.

This report outlines the plan and approach for the prediction of target location using RSS signals and sets the foundation for the successful completion of the project.