



**BDM 3014-INTRODUCTION TO ARTIFICIAL
INTELLIGENCE**

FINAL PROJECT REPORT

**PROJECT TITLE:
GERMANY APARTMENT RENTAL
PREDICTION**

**GUIDED BY
BHAVIK GANDHI**

GROUP 2

**ALWIN SCARIA
DONA BIJU
MOHAMMAD AFTAB
JERIN THOMAS
SNEHA SUJATHA
SANIKA SABEESH
TANVI TAKKAR**

MECE TABLE

Type of activity	
Data Collection	<p>Jerin and Dona: Selected the dataset from Kaggle.</p> <p>Dona: Choose the features and specific regions of dataset to concentrate on.</p> <p>Sanika and Sneha: Collect information from a range of sources, including publications, case studies, documentations and the like.</p> <p>Jerin: Work together as a team and talk about any difficulties that may have arisen.</p> <p>All group members: Study on dataset's attributes, evaluating their significance</p>
Project Proposal	<p>All group members: Developed the project's statement, target customer service, and possibilities and the project's dataflow and wireframes were designed.</p>
Git Repository and Project board	<p>Jerin: Created Git Repository, Project Board</p> <p>All group members: Created Task, added the appropriate description and created branches for uploading their work. Pull requests to merge each person's contribution on different branches were reviewed and then merged to main branch after approval.</p>
Dataset loading and import libraries	<p>Sanika: Loaded the dataset and imported the libraries and display basic data information.</p>
Data Cleaning	<p>Dona and Tanvi: Look for missing values, find the duplicates, checking datatypes, identifying unique values. Moreover, handling missing values using imputation method.</p>
Data Visualization	<p>Tanvi, Sneha, Sanika and Jerin: Plotted different graphs using different libraries like matplotlib, Plotly, seaborn and Streamlit.</p>
Data Preprocessing	<p>Dona: Performed outlier detection and handling.</p> <p>Dona, Sneha, Sanika, Tanvi, Alwin: Checked for correlation among the independent as well as the dependent feature</p> <p>Dona and Aftab: Performed data encoding using different encoding technique</p>

Feature Engineering	<p>Dona, Alwin, Tanvi, Sneha, and Sanika: Drop unwanted features and selected the relevant features.</p> <p>Aftab and Jerin: Performed standardization methods such as MinMaxScaler, Standard Scaler and Robust Scaler</p>
Model Building	<p>Dona: Split the data into train and test.</p> <p>Jerin: Selected different model and train the model.</p> <p>Aftab and Alwin: Evaluated the model by using metrics</p>
Model Tunning & Development	<p>Sneha: Performed hyper parameter tuning in random forest model using GridSearchCV to improve the performance.</p> <p>Sanika: Performed Stacking ensemble technique on Decision Tree model and Random Forest model.</p> <p>Tanvi: Performed Cross Validation using KFold on Random Forest model.</p> <p>Aftab: Performed Model Interpretation by checking the Cumulative Importance and generated dependence plots using SHAP library.</p> <p>Dona: Evaluated the model</p>
Model Deployment	<p>Alwin: Deployment of the model using flask and ensured his requirements ,changes and updates are completed by the team</p>
Model Demonstration	<p>Jerin: Demonstration of the model using Streamlit.</p>
Report and Presentation PowerPoint Drafting	<p>All team members: Worked together to create presentation,report and ensure a platform for content style and formatting uniformity.</p>
Report Merging	<p>All team members: Consolidated and final editing was completed.</p>

PROJECT TABLE

TASK	TASK DESCRIPTION	
Stacked Ensemble	Model 1: LinearRegression	Accuracy 61%
	Model 2: Ridge and Lasso	61%
	Model 2: Polynomial and Ridge	63%
	Model 3: Decision Tree	72%
	Model 4: Random Forest	84%
Interpretation	Local interpretation: SHAP	Done
	Global interpretation: Partial Dependence Plots(PDPs)	Incomplete
	Individual Conditional Expectation(ICE)Plots	Done
	Feature Importance Analysis	Done

Model Tuning	Issue with Linear Regression model: <ul style="list-style-type: none"> Low Accuracy. MSE and RMSE are highly indicating significant deviations between predicted and actual values. 	Solution: Applied regularization techniques like Ridge regression or Lasso regression to reduce overfitting and improve generalization.
	Issue with Ridge and Lasso Model: <ul style="list-style-type: none"> The model's R-squared value was still showing the same performance. MSE and RMSE are still highly indicating significant deviations between predicted and actual values. 	Solution: Performed Polynomial and Ridge Regression model to check the non-linearity relationship, capture complex problems and reduce overfitting.
	Issue with Ridge and Polynomial Model: <ul style="list-style-type: none"> The model accuracy increased from 61% to 63%. And the MSE, MAE, RMSE showed less deviation. A better prediction was expected with less deviations in the various evaluation methods. 	Solution: Experiment with different model hyperparameters using techniques like grid search.
	Issue with Decision Tree Model: The decision tree model indicates a moderately high prediction. Since Decision Tree may become more complex as tree grows deeper and is also sensitive to outliers in the data	Solution: Performed Random Forest model which showed strong performance out of all the models and minimized the errors.
	Issue with Random Forest Model: <ul style="list-style-type: none"> Because of the complexity in the dataset, it was challenging to understand the decision-making process. 	Solution: Hyperparameter Tuning Ensemble Method Cross-Validation Model Interpretation
Deployment and Demo	Deployment- Done Demo - Done	Tech used: Flask Tech used: Streamlit
GitHub Repository Link	Code check-ins are done by each member each week. All deployment from dev to test code is reviewed by a separate member with comments etc. Notebooks in a separate folder in dev and py files moving through to test/master	Done Done

GIT REPOSITORY LINK:

https://github.com/JerinT/rental_prediction_system/blob/main/apartmentRentalPrediction.ipynb

PROJECT BOARD LINK:

<https://github.com/users/Jerin-T/projects/1/views/1>

STREAMLIT VISUALISATION LINK:

<https://app-egiqrixlbqjhmgupcsetcn.streamlit.app/>