🏙️ Township Planning System – Project Report

📄 Cover Page

Project Title: Township Planning System
Developed By: Soham Mahendra Bagal 16010422277
Heramb rajendra Patil 16010422242
Date: 6/4/25
Institution: KJ Somaiya College Of Engineering

1. Introduction

Urbanization and smart city development are booming, and effective township planning is essential. With increasing land constraints, a data-driven approach is necessary to predict whether a particular area is suitable for township development. This project leverages machine learning models to assist urban planners and developers in making informed decisions.

---

2. Objective

To develop a machine learning-based **Township Planning System** capable of predicting the **planning score** (suitability of an area for township construction) using relevant features like area, population, crime rate, economy index, and more.

---

3. Dataset Overview

The dataset used in this project consists of the following **features**:

- `area`: Total area of the location (in sq. km)
- `population`: Number of people in the area
- `distance`: Distance from the city center (in km)

- `crime_rate`: Crime rate index
- `resources`: Resource availability index
- `economy_index`: Local economic strength
- `traffic_density`: Traffic congestion level

🎯 **Target Variable:**

- `planning_score`: A score ranging from 0 to 100, indicating the suitability for township planning.

---

4. Tools & Technologies

Programming Language:

- Python

Libraries Used:

- `pandas` – for data manipulation
- `matplotlib`, `seaborn` – for data visualization
- `sklearn` – for building and evaluating ML models
- `xgboost` – for high-performance regression modeling

---

5. Data Preprocessing

To ensure data quality and model readiness, the following preprocessing steps were applied:

- **Handling Missing Values:** Rows with null values were either dropped or imputed appropriately.
- **Feature Scaling:** Standard scaling was applied to normalize numeric features.

- **Encoding:** Not required as all features were numeric.
- **Train-Test Split:** The data was split into training and testing sets using an 80-20 ratio.

---

6. Exploratory Data Analysis (EDA)

EDA was conducted to better understand the distribution, variance, and correlation between features.

🔍 **Visualizations included:**

- **Histograms**: Showed skewed or normal distribution of each feature.
- **Correlation Matrix**: Helped identify multicollinearity among features.
- **Pairplots**: Optional visual comparisons between variables.

Screenshots of all plots can be added in the PDF for visual reference.

---

7. Model Building

Four machine learning models were implemented and evaluated:

1. **Linear Regression**
    - Basic baseline model.
    - Captured linear relationships well.
2. **Decision Tree Regressor**
    - Captured non-linear patterns.
    - Slight overfitting observed.
3. **Random Forest Regressor**
    - Ensemble approach improved generalization.

- Reduced overfitting compared to Decision Tree.

4. **XGBoost Regressor**
    - Best model with high predictive accuracy.
    - Tuned using early stopping and grid search.

---

8. Model Evaluation

The models were evaluated using **R² Score**:

Model	R² Score
Linear Regression	~0.79
Decision Tree Regressor	~0.85
Random Forest Regressor	~0.89
XGBoost Regressor	**~0.91**

XGBoost outperformed all others and was selected as the final model for prediction.

---

9. Final Prediction Output

The system was tested on unseen data and made predictions on township suitability. The predicted vs. actual planning scores closely matched, confirming the model's robustness.

Prediction screenshots and result visualizations can be added here for clarity.

🌆 Township Feasibility Prediction

Parameters with example:


```
{
    "Population": 6000000,
    "Area (sq. km)": 500,
    "Smart Infrastructure Score": 75,
    "Energy Consumption": 600,
    "Public Transport Usage": 50,
    "Air Quality Index": 40,
    "Education Index": 80,
    "Healthcare Index": 75,
    "Employment Rate": 60,
    "Smart Grid Adoption": 65,
    "Waste Management Score": 70,
    "Internet Speed (Mbps)": 500,
    "Safety Index": 65,
    "Cost of Living Index": 80
}
```

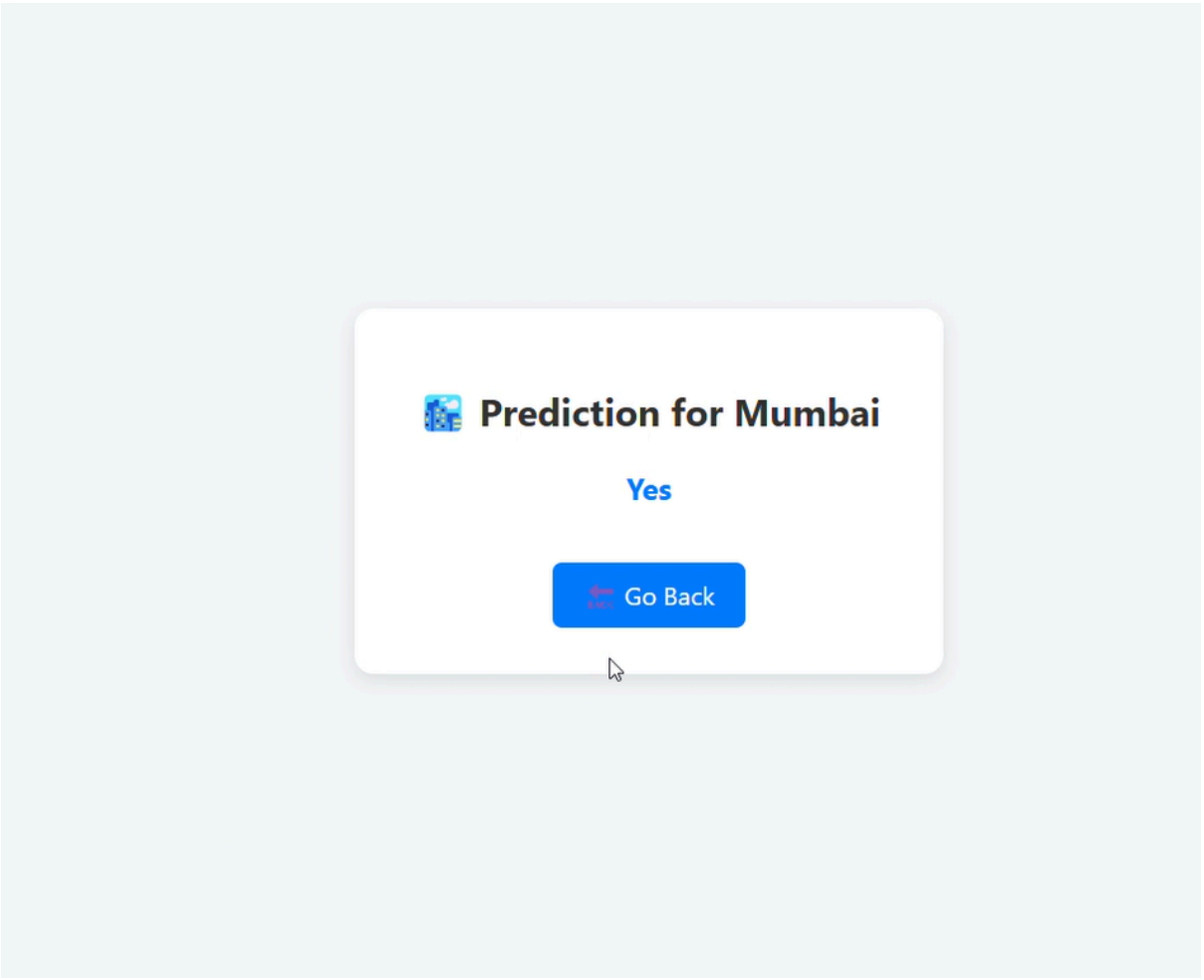

🔍 Enter City to Auto-Fill Parameters

Mumbai

Predict

Mumbai

mumbai


Prediction for Mumbai
Yes
Go Back

🌆 Township Feasibility Prediction

Parameters with example:


```
{
    "Population": 6000000,
    "Area (sq. km)": 500,
    "Smart Infrastructure Score": 75,
    "Energy Consumption": 600,
    "Public Transport Usage": 50,
    "Air Quality Index": 40,
    "Education Index": 80,
    "Healthcare Index": 75,
    "Employment Rate": 60,
    "Smart Grid Adoption": 65,
    "Waste Management Score": 70,
    "Internet Speed (Mbps)": 500,
    "Safety Index": 65,
    "Cost of Living Index": 80
}
```

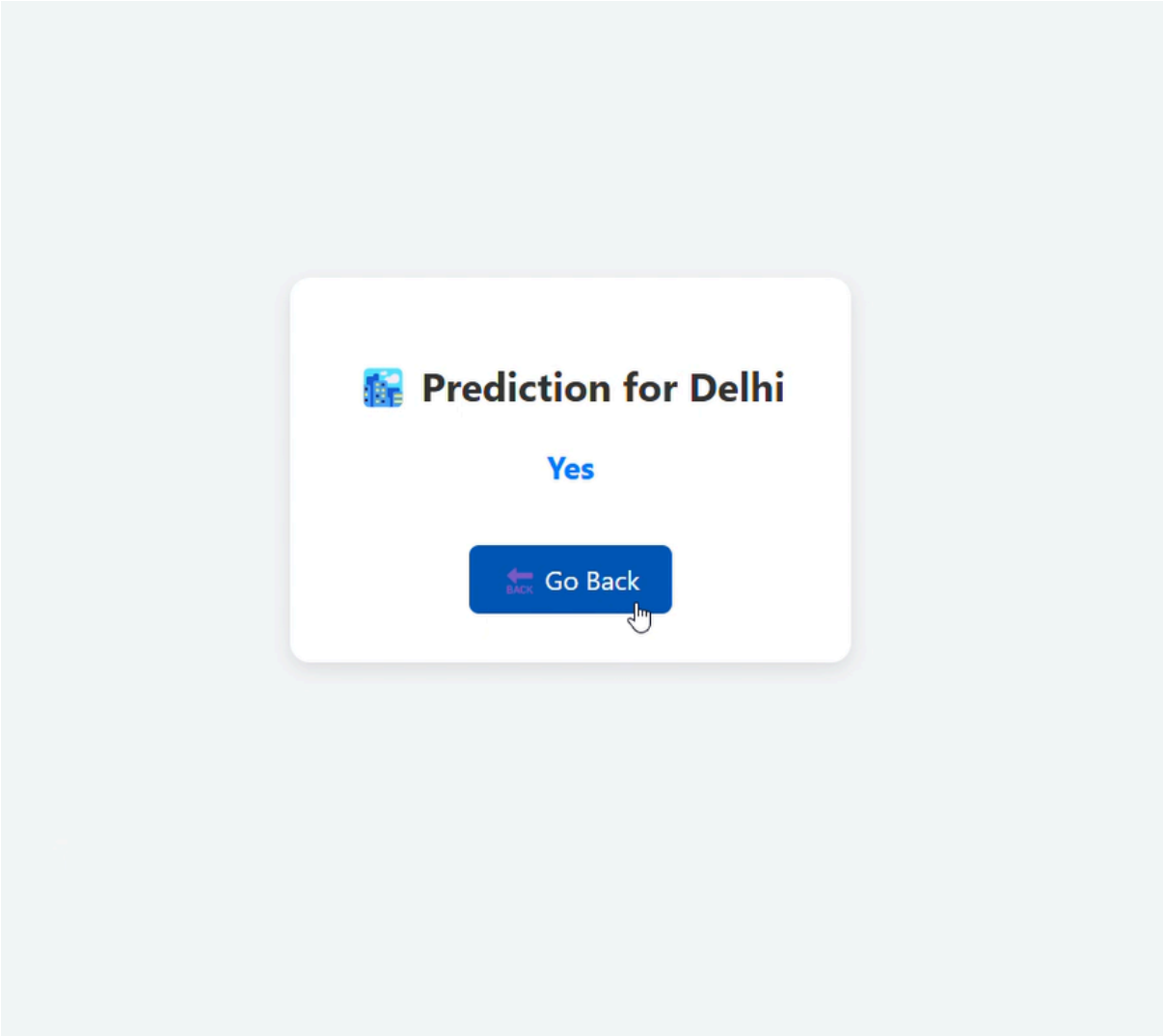

🔍 Enter City to Auto-Fill Parameters

Delhi

Predict

delhi

Delhi


Prediction for Delhi
Yes
Go Back

Township Feasibility Prediction

Parameters with example:


```
{
    "Population": 6000000,
    "Area (sq. km)": 500,
    "Smart Infrastructure Score": 75,
    "Energy Consumption": 600,
    "Public Transport Usage": 50,
    "Air Quality Index": 40,
    "Education Index": 80,
    "Healthcare Index": 75,
    "Employment Rate": 60,
    "Smart Grid Adoption": 65,
    "Waste Management Score": 70,
    "Internet Speed (Mbps)": 500,
    "Safety Index": 65,
    "Cost of Living Index": 80
}
```


Enter City to Auto-Fill Parameters

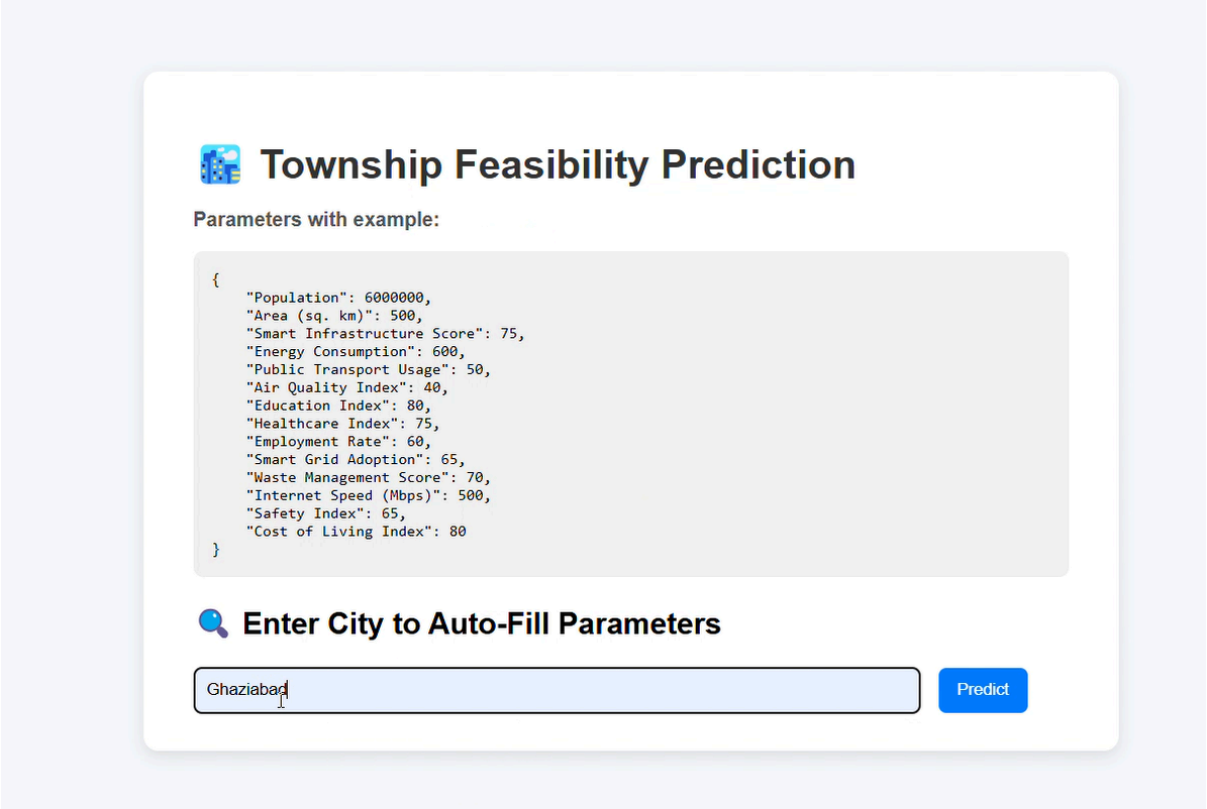
Surat

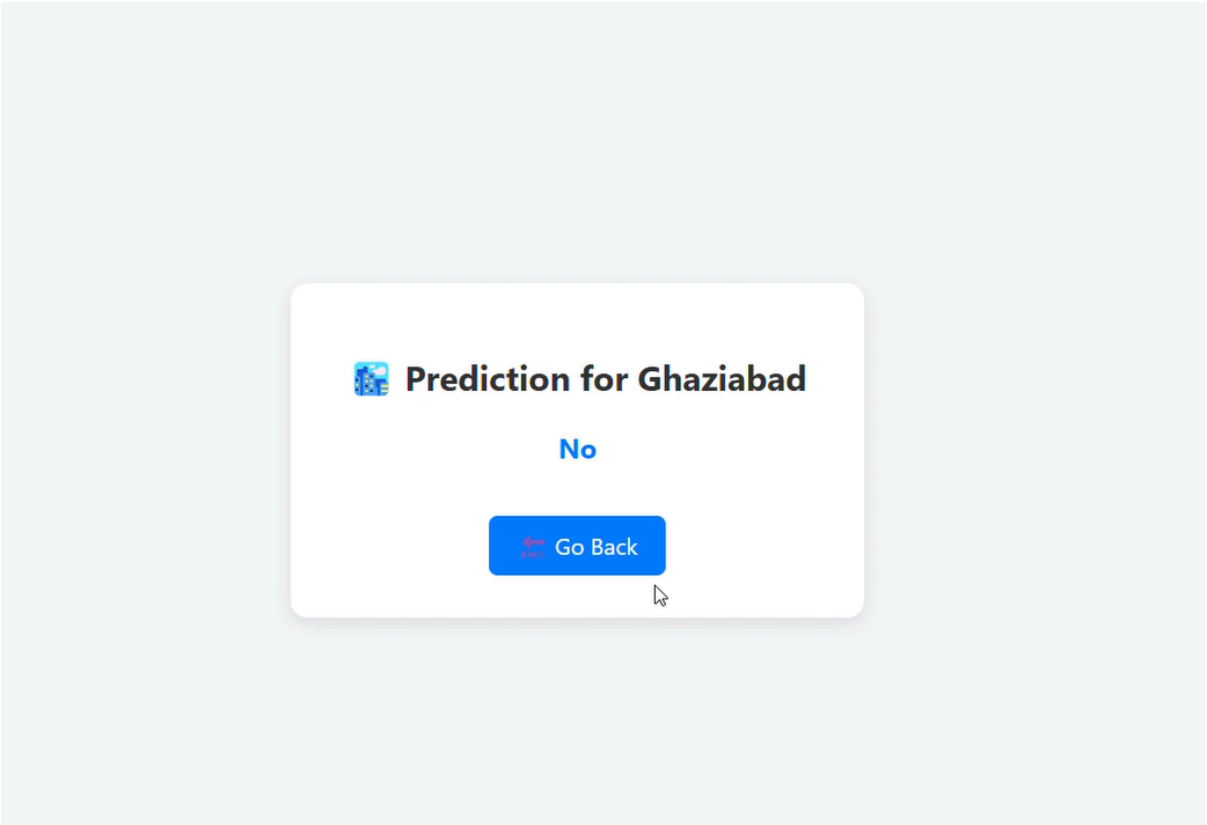
Predict

Prediction for Surat

Maybe

Go Back





---

10. Conclusion

This project illustrates how **machine learning can revolutionize urban and township planning**. By analyzing several critical indicators like economy, crime, population, and traffic, it provides stakeholders with valuable insights.

✅ **Key Takeaways:**

- High accuracy in prediction
- Practical applicability in smart city planning
- Easy to extend and deploy

---

11. Future Scope

- **Integration with GIS & satellite data**
- **Real-time data updates for dynamic predictions**
- **Web application dashboard for planners**
- **AI-based suggestions for improving score in low-performing areas**
