VNIT Housing Price Prediction

Welcome to the VNIT Housing Price Prediction project documentation. This comprehensive guide covers the complete machine learning pipeline including data preprocessing, exploratory data analysis, model optimization, and deployment.

Project Overview

The VNIT Housing Price Prediction project is an end-to-end machine learning solution that predicts house prices using advanced regression techniques. The project implements best practices in data science including:

- **Data Preprocessing & EDA** - Comprehensive data cleaning and exploration
- **Hyperparameter Optimization** - Optuna-based parameter tuning
- **Model Training** - XGBoost with cross-validation
- **Experiment Tracking** - MLflow for reproducibility
- **Performance Monitoring** - Evidently for model health checks
- **Workflow Orchestration** - Prefect for pipeline management

Key Metrics

- **Model R^2 Score**: 0.8990
- **RMSE**: $27,836.04
- **MAE**: $17,279.64
- **Optimization Trials**: 50 (Optuna)
- **Cross-Validation Folds**: 5

Quick Start

1. **Install Dependencies**:

```
pip install -r requirements.txt
```

2. **Run Preprocessing Pipeline**: