|  |  |
| --- | --- |
| 高层建筑风压预测项目  （外立面平均风压系数和风压系数标准差的预测）  项目概述  本项目是开发用于预测高层建筑风荷载影响的机器学习模型的初始阶段。 目标是预测：   * 平均风压系数（Cp） * 风压系数标准差（StdDev）   即在不同风向和干扰建筑位置变化下，建筑外立面特定点的风压行为。  所采用的基础模型结构为多层感知机（MLP），包含两个隐藏层，且隐藏层单元数相同。该结构因其简单性、可解释性以及对非线性特征建模的表达能力而被选中，非常适合作为开发初期的基线架构。  模型训练数据来源于风洞实验，实验对象为具有相同高度与尺寸的建筑模型。  数据集大小：   * 总计：153,216 行 * 预处理后：   + 训练集：121,114 行   + Holdout 测试集：30,279 行   对于用于预测风压系数标准差（StdDev）的模型，训练集和测试集未进行异常值剔除。所有原始值均保留，以保持数据的真实波动性，避免扭曲风压波动的物理特性。此决策旨在提高模型在真实工程数据上的泛化能力。  仓库内容概览  本仓库包含两套独立训练的机器学习模型，均基于 MLP 架构，目标是预测高层建筑外立面的风压行为：  1. 模型 1 — 平均风压系数（Cp）  预测在不同流动条件下，建筑外立面某点的平均风压系数。  2. 模型 2 — 风压系数标准差（StdDev）  预测风压的波动性，即风压系数的标准差，反映了该点的湍流强度。  两套模型使用相同的输入特征空间，但分别独立训练和验证。  每个模型归档中包含：   * 模型配置（超参数与架构） * 训练与验证日志 * Holdout 测试集性能指标（MAE, MSE, R²） * 训练后的模型权重 * 指标可视化图表 * 多组数据稳定性测试结果   数据集说明  数据来源为模拟建筑群风流场的风洞实验。 每一行数据代表一次在特定风向和干扰建筑配置下，立面某点的风压测量。  输入特征：   * X\_int, Y\_int: 干扰建筑的坐标 * X\_fac, Y\_fac: 主楼立面测点的坐标 * Ang: 风向角（单位：度）   预测目标：   * Cp: 平均风压系数（模型 1） * StdDev: 风压系数的标准差（模型 2）   数据集大小：   * 总行数：153,216 * 训练集：121,114 * Holdout 测试集：30,279   模型结果  两套模型在 holdout 测试集上均表现出色，以下是最终的性能指标与 50 次稳定性测试结果： | **Wind Pressure Prediction for High-Rise Buildings**  *(Prediction of Mean Pressure Coefficient and Standard Deviation of Pressure Coefficient on Facades)*  **Project Overview**  This project represents the initial stage of developing a machine learning model for predicting wind effects on high-rise buildings. The goal is to forecast:   * **Mean pressure coefficient (Cp)** * **Standard deviation of pressure coefficient (StdDev)**   in specified points on a building’s facade under different wind directions and varying positions of interfering buildings.  The baseline model architecture is a **Multi-Layer Perceptron (MLP)** with two hidden layers of equal size. This structure was selected due to its simplicity, interpretability, and sufficient expressive power for the initial development phase. Model training was performed using the results of wind tunnel experiments on physical building models of identical height and dimensions.  **Dataset size:** Total: 153,216 rows After preprocessing:   * Training set: 121,114 rows * Holdout set: 30,279 rows   For the model predicting the **standard deviation of pressure coefficient (StdDev)**, the training and test sets were **not filtered for outliers**. All values were used in their original form to preserve realistic variability and avoid distorting the nature of pressure fluctuations. This decision was made intentionally to improve the model’s generalization ability when working with raw engineering data.  **Repository Summary**  This repository includes two independently trained machine learning models based on the MLP architecture, designed to predict wind pressure characteristics on high-rise building facades:   1. **Model 1: Mean Pressure Coefficient (Cp)** Predicts the average pressure coefficient at a given facade point under different flow conditions. 2. **Model 2: Standard Deviation of Pressure Coefficient (StdDev)** Predicts the variability of pressure using the standard deviation of the coefficient, representing turbulence intensity at the measurement point.   Both models use the **same set of input features** but were trained and validated **independently**.  Each model archive includes:   * Model configuration (hyperparameters, architecture) * Training and evaluation logs * Performance metrics on the holdout set (MAE, MSE, R²) * Model weights * Visualization of metrics * Stability analysis charts for different subsets   **Dataset Description**  The dataset was obtained from wind tunnel tests simulating wind flow around building groups. Each row represents a single pressure measurement at a specific facade point under a given wind direction and position of an interfering building.  **Input features:**   * X\_int, Y\_int: coordinates of the interfering building * X\_fac, Y\_fac: coordinates of the pressure measurement point on the main facade * Ang: wind direction (degrees)   **Target variables:**   * Cp: Mean pressure coefficient (for Model 1) * StdDev: Standard deviation of pressure coefficient (for Model 2)   **Dataset size:**   * Total: 153,216 rows * Training set: 121,114 rows * Holdout set: 30,279 rows |
| 🔹 模型 1 — 平均风压系数（Cp）  Holdout 测试集指标：   * R²: 0.9784 * MAE: 0.0518 * MSE: 0.005328     稳定性测试（50 次）：   * R² 平均值: 0.9651 ± 0.0089 * MAE 平均值: 0.0644 ± 0.0082 * MSE 平均值: 0.008654 ± 0.002215 | **Results**  Both models demonstrated high prediction accuracy on the holdout set. Below are the final evaluation metrics and stability test results (across 50 training set variations).  **1) Model 1 — Cp (Mean Pressure Coefficient)**  **Holdout Metrics:**   * R²: 0.9784 * MAE: 0.0518 * MSE: 0.005328     **Stability (50 runs):**   * R² mean: 0.9651 ± 0.0089 * MAE mean: 0.0644 ± 0.0082 * MSE mean: 0.008654 ± 0.002215 |
| 🔹 模型 2 — 风压系数标准差（StdDev）  Holdout 测试集指标：   * R²: 0.847 * MAE: 0.0241 * MSE: 0.001076     稳定性测试（50 次）：   * R² 平均值: 0.7071 ± 0.1372 * MAE 平均值: 0.0316 ± 0.0072 * MSE 平均值: 0.002034 ± 0.000954     访问链接  最终模型、训练笔记本与可视化图表请见：   * GitHub 项目仓库： <https://github.com/fzakharov129/wind_pressure_mlp> * Kaggle 页面： <https://www.kaggle.com/datasets/fedorzakharov331/wind-pressure-mlp> | **2) Model 2 — StdDev (Standard Deviation of Pressure Coefficient)**  **Holdout Metrics:**   * R²: 0.847 * MAE: 0.0241 * MSE: 0.001076     **Stability (50 runs):**   * R² mean: 0.7071 ± 0.1372 * MAE mean: 0.0316 ± 0.0072 * MSE mean: 0.002034 ± 0.000954     Final versions of the trained models and metric visualizations are available at:   * **GitHub Repository:** <https://github.com/fzakharov129/wind_pressure_mlp> * **Kaggle Page:** <https://www.kaggle.com/datasets/fedorzakharov331/wind-pressure-mlp> |