

Configuration Manual

MSc Research Project
MSc in Data Analytics

Jaswanth Chandrashekaraiyah

Student ID: 23285265

School of Computing
National College of Ireland

Supervisor: Harshani Nagahamulla

CONFIGURATION MANUAL

Jaswanth Chandrashekaraiiah

Student ID: 23285265

1. System Requirements

- **Hardware:** Minimum 16 GB RAM (32 GB recommended), CPU with 4+ cores; GPU with CUDA support (optional, speeds up model inference).
- **Storage:** At least 10 GB free for datasets, model files, and cached images.
- **Operating System:** Windows 11 / Ubuntu 20.04+
- **Python Version:** Python 3.9 or above □ **Libraries** (from requirements.txt):

- pandas
- NumPy
- h3
- scikit-learn
- matplotlib
- requests
- Pillow
- folium
- tqdm
- Flask
- torch
- torchvision

2. Steps to Execute

Step 1: Install dependencies

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

Step 2: Ensure the following files are in the project root directory:

- `app.py` - Flask backend application
- `map.html` - Interactive map frontend
- `best_hybrid_model.pt` - Trained CNN-LSTM hybrid model

- Dataset NPZ file (`cnn_lstm_data.npz`) containing sequences, hex IDs, and optional crime type mappings.
- `street_view_images/` - Folder with Google Street View images for each H3 hex ID.

Step 3: Run the Flask application python

`app.py`

The app will load the model, sequence data, and image features at startup.

Output will confirm successful loading:

[INFO] Loading model and data ...

[INFO] Model and data loaded successfully

Step 4: Open the map interface

- Visit `http://localhost:5000/map` in a web browser.
- Click on any location in the Los Angeles area to view predicted risk level, common crime type, and model confidence score.

3. Model & Configuration Details

Model Architecture

- CNN branch processes 64×64 RGB images of H3 hexes.
- LSTM branch processes 12-month crime count sequences.
- Fully connected layers combine spatial and temporal features to classify risk as Low, Medium, or High

Prediction Flow

- User clicks a location on the map (`map.html`).
- Latitude and longitude are sent to `/predict` via a POST request.
- Backend converts coordinates to H3 hex ID (level 8) using the **h3** library.
- Sequence vector and street view image for that hex are loaded.
- CNN-LSTM hybrid model predicts risk class and probability.
- Response returns `risk_level`, `common_crime`, and `confidence`.

Color Coding (for map visualization):

- **Low Risk** - Green #33cc33
- **Medium Risk** - Yellow #ffcc00
- **High Risk** - Red #ff3333