Predicting bike demand using deep learning techniques

Mehizel Ali

Introduction

In this project, we developed three models to predict bike demand for the next 60 minutes at each station. The work was divided into the following parts:

1. **Exploratory Data Analysis (EDA):** We provided analytical results to answer key business questions:
   - Identifying the top 5 stations with the highest bike demand compared to the lowest 5.
   - Analyzing the most frequent ride routes between stations.
   - Identifying demand trends over time.
2. **Model Development:** We created three models:
   - LSTM-based Convolutional Model
   - LSTM with Attention Mechanism
   - Graph Convolutional Network (GCN) *(The GCN was particularly challenging to implement.)*
3. **Evaluation and Techniques:**
   - **Metric:** Mean Absolute Error (MAE)
   - **Optimization Technique:** ADAM optimizer
   - **Challenge:** We were unable to include weather-related features due to limited access to data from open-source APIs.

1 First part: EDA

Business Recommendations

Note: visualization provided in last page.

- **Move Underused Stations:**
  - **Action:** Find stations that are not used much and move them to busier areas.
  - **Why:** Moving stations to high-traffic areas will help increase their usage.
- **Target Morning and Evening Riders:**
  - **Action:** Focus on people who use bikes in the morning and evening. Offer special deals or discounts during these times.
  - **Why:** Morning and evening riders are more regular. Giving them offers can build loyalty.
- **Encourage Night Riders:**
  - **Action:** Offer discounts or deals for people who use bikes at night. Place stations near nightlife areas or hotels.
  - **Why:** Getting more people to ride bikes at night can help balance out the quieter hours.
- **Plan for Winter Demand Drop:**
  - **Action:** When demand drops in early December due to cold weather, offer special promotions to keep people using the service.
  - **Why:** Adjusting to the winter slowdown can keep ridership steady and prevent losing customers.

2 Second part: modeling

Methodology

1. Data Preparation

We restructured the data into a cross-table format with lag features to capture both spatial and temporal patterns.

2. Models

(a) LSTM-based Convolutional Model:

- Two ConvLSTM layers: 64 and 32 units, respectively.
- ReLU activation function.
- A two-layer MLP: 64 units → 1 unit.
- Batch normalization to address vanishing/exploding gradients.

(b) Attention-based LSTM:

- An attention mechanism to capture spatial features.
- A single LSTM layer with 64 units to extract temporal dependencies.

(c) Graph Convolutional Network (GCN):

- **Graph Structure:**
  - Nodes represent stations.
  - Edges represent rides between stations. An edge exists if there's at least one ride between two stations.
  - The graph is undirected (A ↔ B for a ride from Station A to Station B).
- **Example Representation:**
  - Input data row:

    ```
    start_station_id: 123, end_station_id: 456, rides_: 20
    start_lat: 40.7, start_lng: -74.0, end_lat: 40.8, end_lng: -73.9
    ```

  - Preprocessed graph:
    * Node 0 (Station 123) → Features: [40.7, -74.0, 40.8, -73.9, 20].
    * Edge index: [[0], [1]] (connection between Node 0 and Node 1).

Results

- **ConvLSTM:** Lowest train loss: **1.16**.

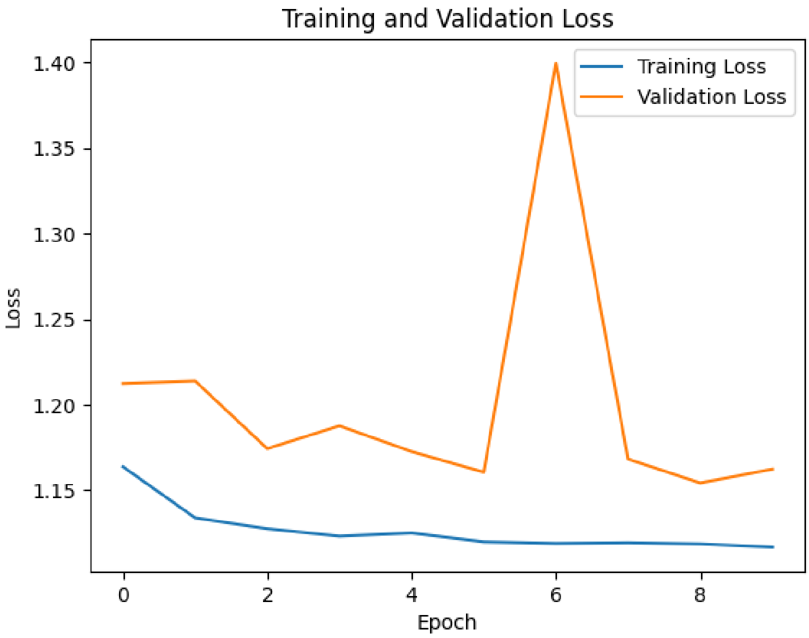

Figure 1: Losses convLSTM

- **Attention LSTM:** Train MAE: **1.57**.

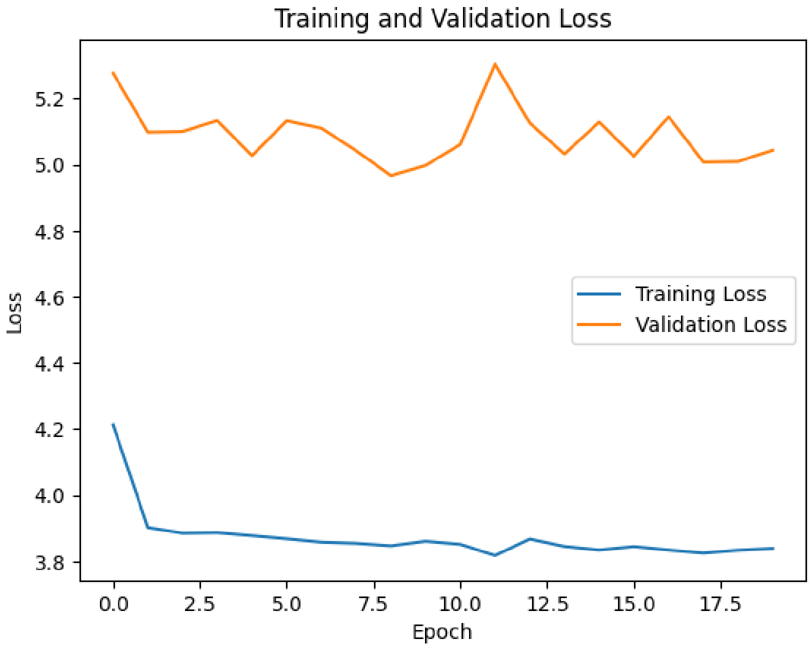

Figure 2: Losses attention

- **Graph Convolutional Network:** Train loss: **3.48**.

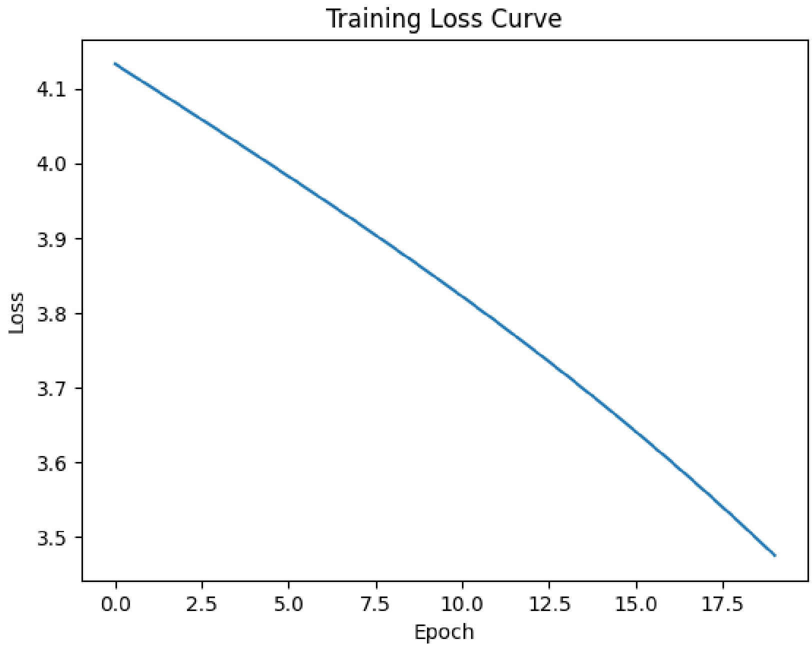

Figure 3: Losses Graph Convolutional Network

Discussion

The convolutional LSTM achieved the lowest error among all models, making it the most effective in this project. However, the Graph Neural Network showed promising scaling behavior—its training loss decreased linearly with the number of epochs (neural scaling laws), indicating strong potential for larger datasets.

Future work: must focus on how to represent spatio temporal data using for example grad based spatial data or others technique it will very powerful to model making them extract new features and better accuracy , try to use another models like KAT kolmogorv arnold transformers, Graph kolmogrov arnold network.

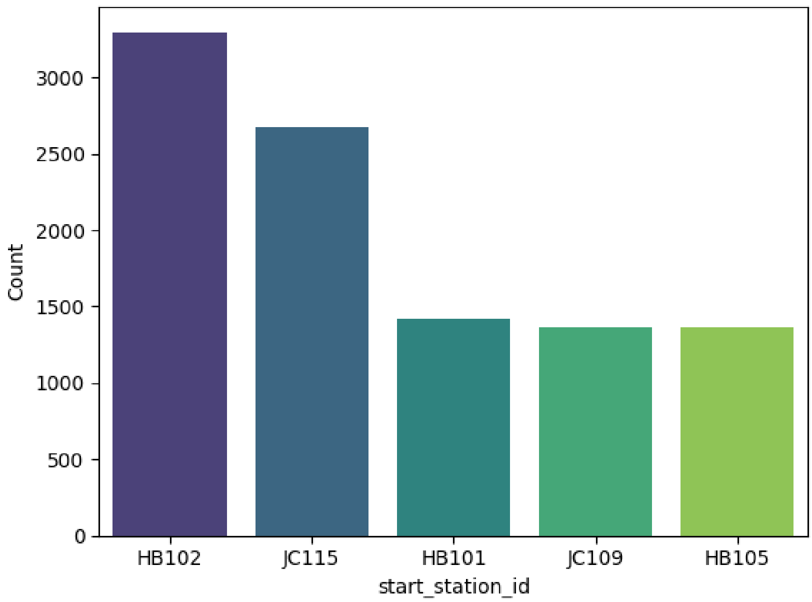

Figure 4: Top 5 active stations over this period that have higher demand

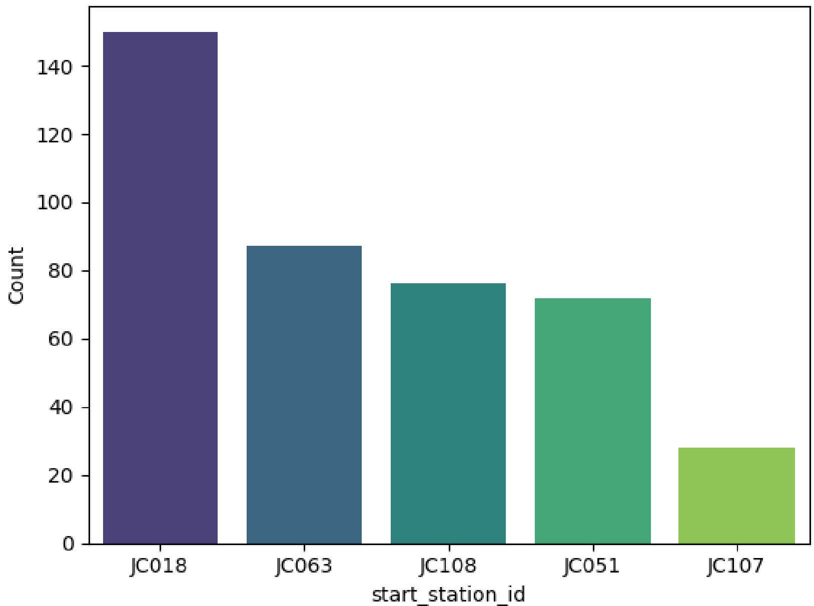

Figure 5: Lowest 5 demand per station

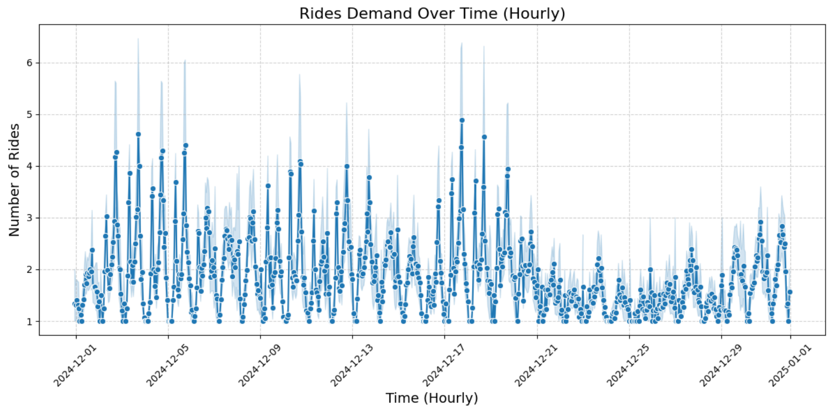

Figure 6: Evolution of bike demand over stations during this period

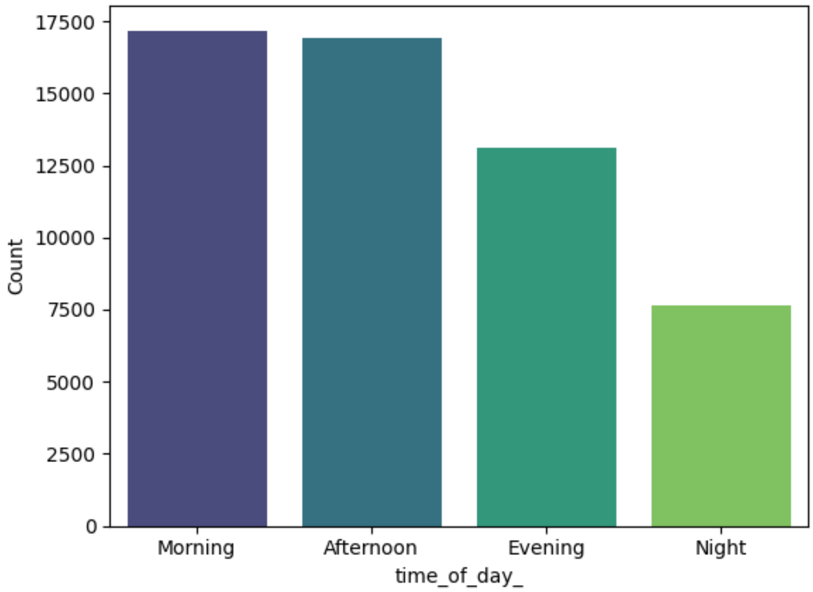

Figure 7: Distribution of bike demand over time of day

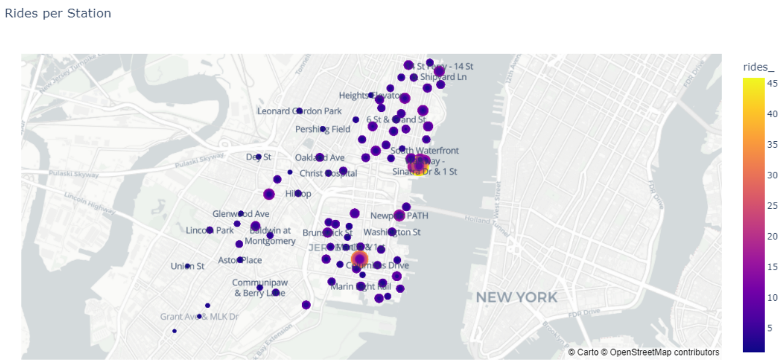

Figure 8: Rides demand per station

	start_station_name	end_station_name	ab_freq_
3258	McGinley Square	Bergen Ave & Sip Ave	317
2034	Grove St PATH	Marin Light Rail	296
2028	Grove St PATH	Liberty Light Rail	285
2844	Liberty Light Rail	Grove St PATH	274
2407	Hoboken Terminal - Hudson St & Hudson Pl	Hoboken Ave at Monmouth St	269
492	8 St & Washington St	Hoboken Terminal - River St & Hudson Pl	268
838	Bergen Ave & Sip Ave	McGinley Square	257

Figure 9: Most frequest destination (a,b)