SFU SIMON FRASER UNIVERSITY

Ride-Hailing Order Demand Prediction in Vancouver Area

Zoe(Chenhao) Wang, Philip Cho, Logan Born, Zhuo Cen, Yijing Xue

INTRODUCTION

We have developed a demand prediction system for Kabu, a ride-hailing service company in Richmond, BC. We aim to forecast future demand in Greater Vancouver using machine learning. Based on our findings, Kabu will be able to prioritize service to regions that have higher demand, reduce overall costs, and minimize dispatches to low-demand areas.

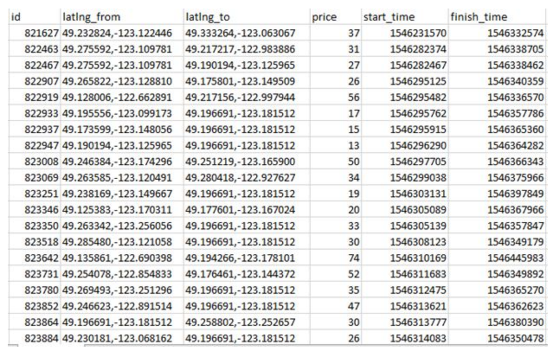
DATA PREPROCESSING

Raw Data:

- >500,000 ride-hailing orders from Kabu inc.
- Daily weather information
- Temporal information (e.g. holidays)

Group data into...

- 1 hour timesteps
- 30 by 50 grid (700m by 700m cells)



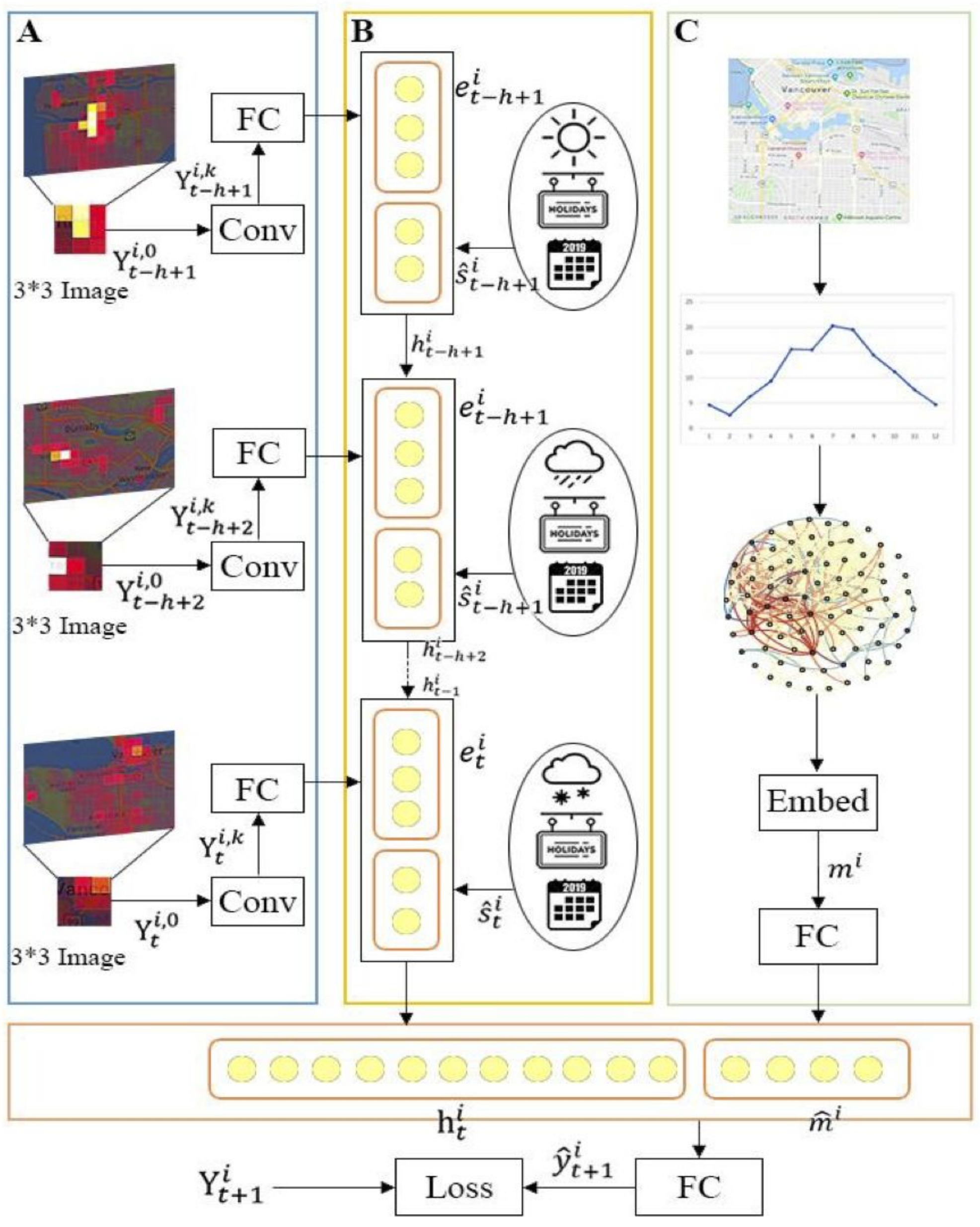
Over 500,000 total orders (2018)

30*50 grid in Great Vancouver Area

Generate heatmap at each timestep

Our training data comprises a full year of ride hailing orders from 2018. These are grouped into discrete chunks of time and locations on a grid to create a demand heatmap. Regions with extremely low demand (<25 orders per year) were filtered out.

METHODOLOGY



Model Architecture adapted from [1]**:**

- **Spatial view:** a CNN captures spatial dependencies between nearby regions. We use 3 layers with 8 3×3 filters and 11 output dimensions.
- **Temporal view:** an LSTM layer concatenates the spatial view with context features at corresponding times. We use a sequence length of 6.
- **Semantic view:** a 32 dimensional graph embedding captures similarity between distant regions.

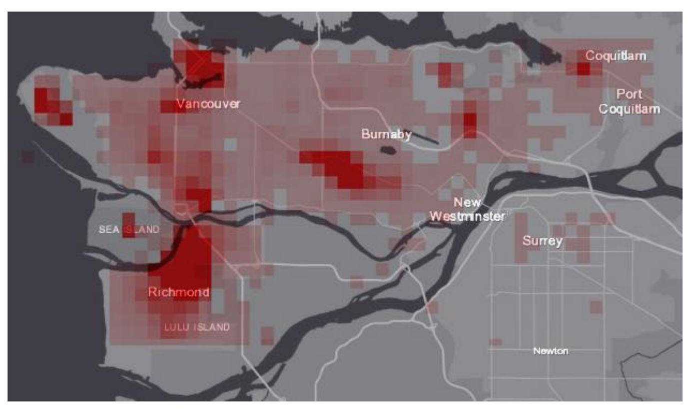
Experimental Design:

- Due to extreme sparsity in our data, we introduce an alternative to mean absolute percentage error (MAPE) loss. **Mean squared logarithmic error (MSLE)** is more suitable for sparse data because it encourages overestimation [2].
- A **hierarchical model** reduces skew in the data. We use a heuristic classifier based on average demand to distinguish between timesteps with zero vs nonzero demand. The regression is trained on the examples predicted to have nonzero demand.
- We compare four models:
  1. Baseline Model: Original architecture described in [1]
  2. Baseline+MSLE: Original architecture optimized with mean squared logarithmic error (MSLE) loss
  3. Hierarchical model: Hierarchical model optimized with the original loss function, mean absolute percentage error (MAPE)
  4. Hierarchical +MSLE: Hierarchical model optimized with MSLE loss

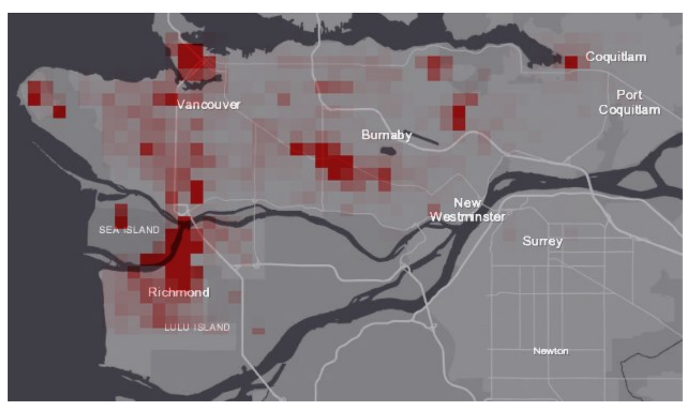
⟵ **For the figure on the left, A. Spatial View B. Temporal View C. Semantic View**

RESULTS

baseline

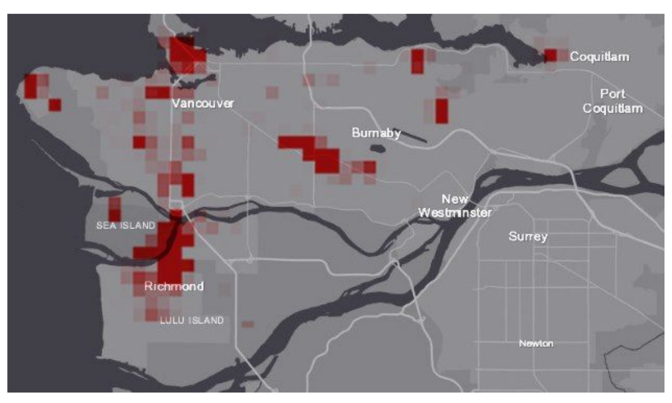


baseline+MSLE



actual demand

hierarchical



hierarchical+MSLE

Model	MAPE	MSLE	RMSE
baseline	0.1114	0.0620	0.5890
baseline+MSLE	0.1399	0.0398	0.4501
hierarchical	0.1101	0.0597	0.5744
hierarchical+MSLE	**0.0499**	**0.0239**	**0.3128**

- Training with MSLE improves accuracy for some metrics but not all
  - MSLE encourages overprediction; MAPE encourages underprediction
  - MSLE increases accuracy when demand is nonzero and reduces accuracy at other times
- A hierarchical model improves accuracy on all metrics
  - Reduces skewness of the training data
- **A hierarchical model trained with MSLE loss gives an optimal performance**
  - MSLE improves accuracy when demand is nonzero
  - Hierarchical structure improves performance when demand is zero

CONCLUSION

We have successfully adapted a state-of-the-art demand prediction model to improve its accuracy on regions with sparse demand. We show that hierarchical models trained with mean squared logarithmic error (MSLE) loss provide the best result in terms of MAPE, MSLE, and RMSE.

FUTURE WORK

Jointly train a neural classifier along with the regression model. We predict this will further increase performance relative to the current hierarchical model which uses a heuristic classifier to predict when demand will be zero.

SELECTED REFERENCES

[1] Huaxiu Yao, Fei Wu, Jintao Ke, Xianfeng Tang, Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye, and Zhenhui Li. Deep multi-view spatial-temporal network for taxi demand prediction, 2018.

[2] Chris Tofallis. A better measure of relative prediction accuracy for model selection and model estimation. Journal of the Operational Research Society, 66(8):1352–1362, Aug 2015.