foss asia

Machine Learning for Smart Cities

Truc Viet 'Joe' Le

@vietexob

March 25, 2018

About me:

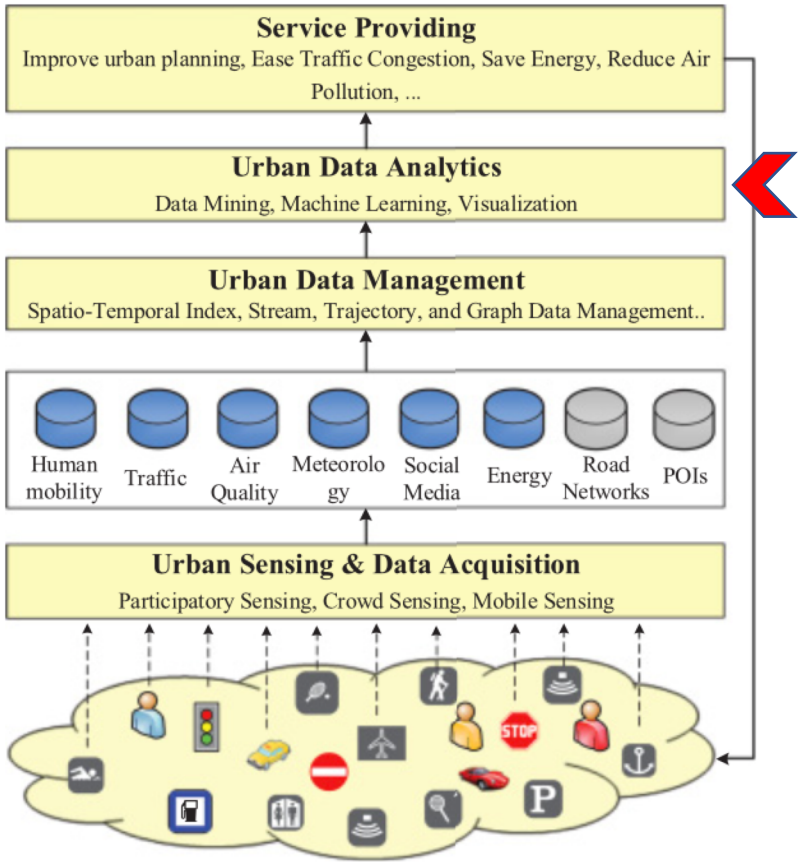
http://trucvietle.me

- Machine Learning (ML) engineer at **SAP** since July 2017
- Received Ph.D. in Information Systems from **Singapore Management University** (SMU) in Feb. 2018
- Thesis work on ML methods for **urban problems** using **spatiotemporal data**
- Avid traveler, adventurer and aspiring writer/photographer

Big Data and the Big Cities

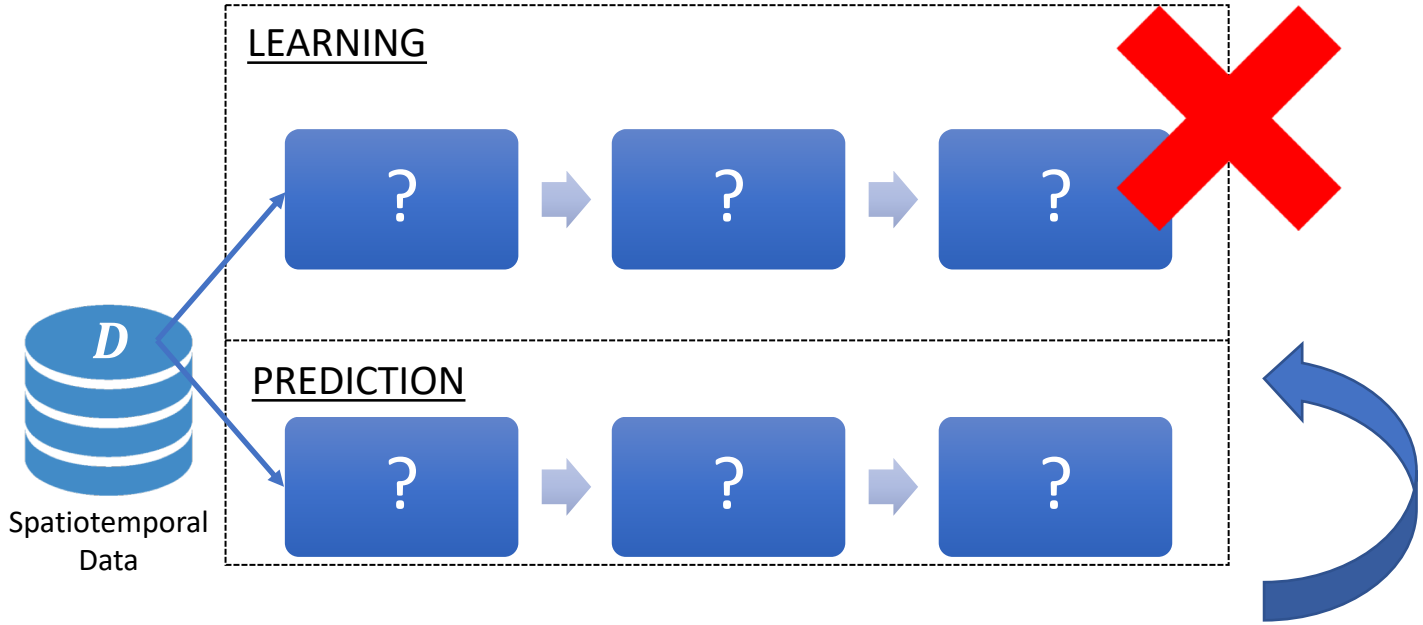
> Solve Big Challenges in Big Cities using Big Data.

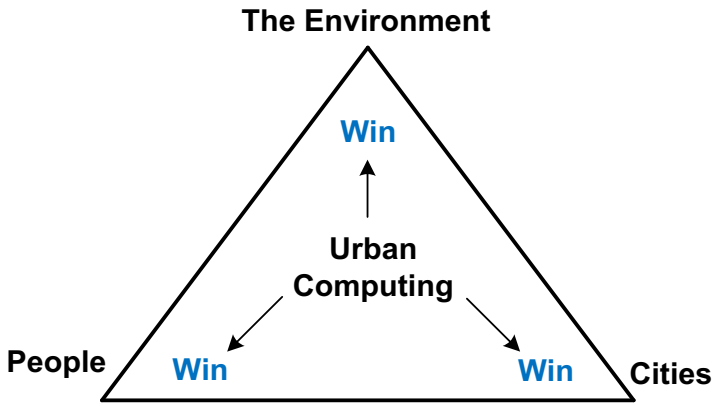
[1]Heilig, G. K. (2014). World urbanization prospects the 2014 revision. *United Nations, Department of Economic and Social Affairs (DESA), Population Division, Population Estimates and Projections Section, New York*.

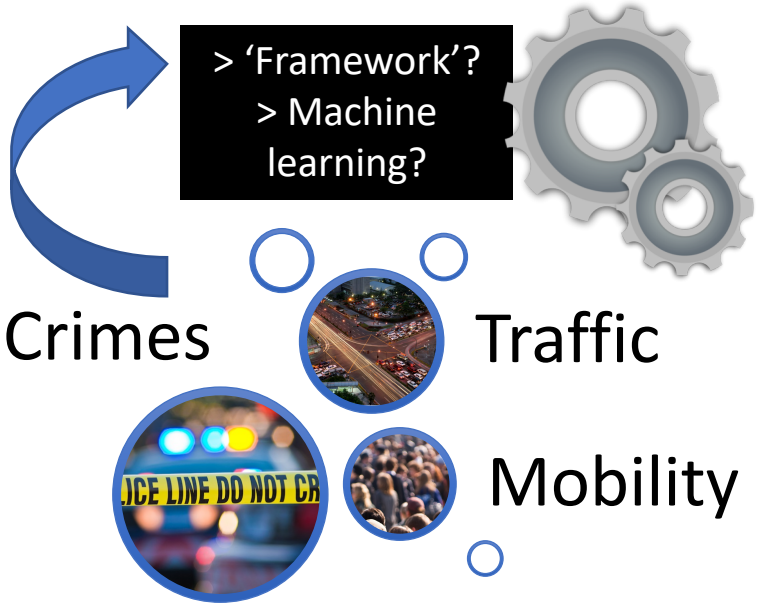


General Framework for Urban Computing Research
Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(3), 38.

"...unlocks the power of big data collected in urban spaces to solve major issues cities face today."







What is this talk about?

Static
(e.g., POIs)

Dynamic (e.g.,
sensors, CCTV)

Trajectories

Human Mobility

POIs,
Trajectory
Prediction

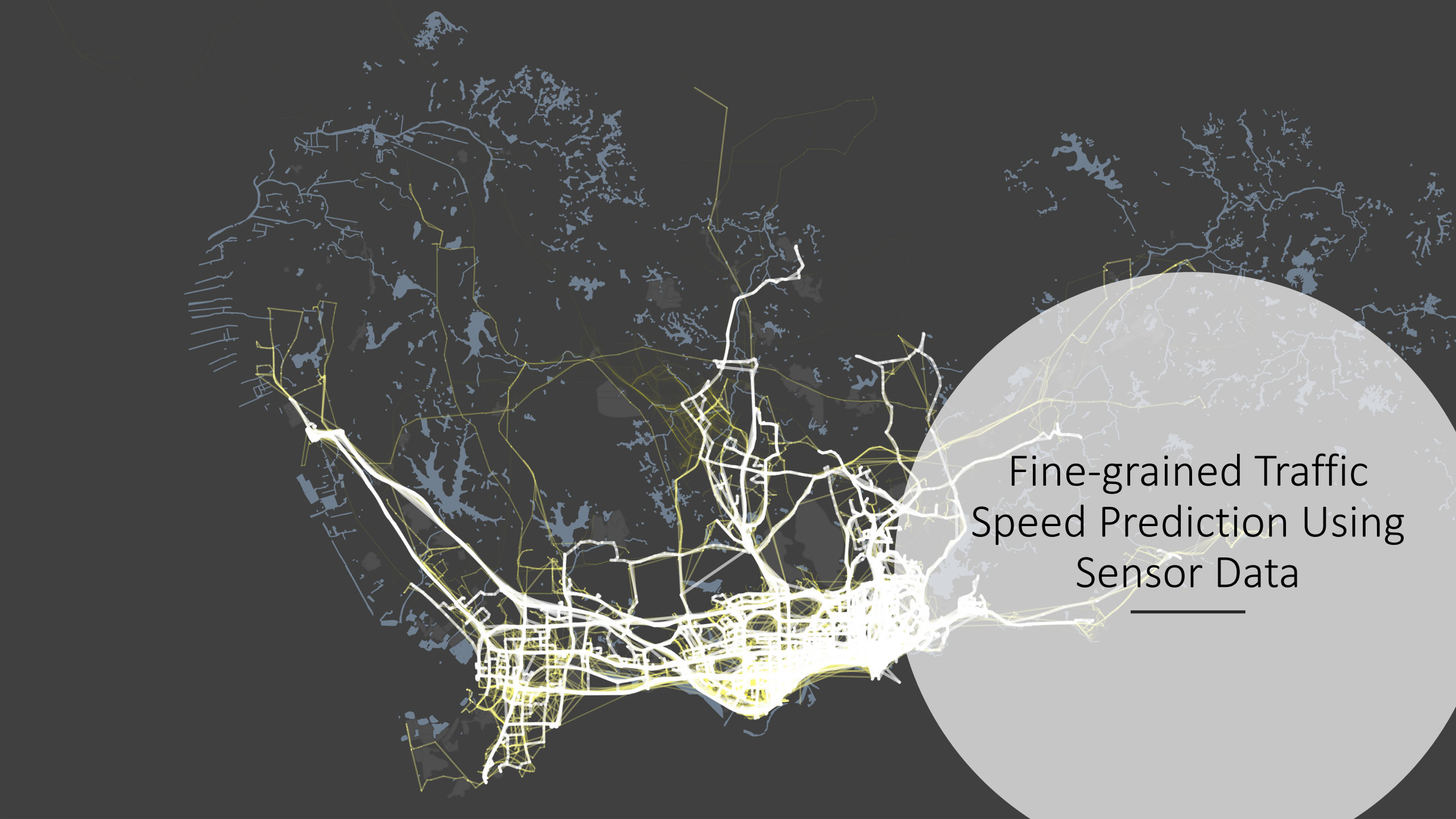
Traffic Speed

Real-time
Speed
Prediction

Crime Incidents

Distribution
of Incidents
over Urban
Areas

Spatiotemporal Data Properties and Problems Addressed



Fine-grained Traffic Speed Prediction Using Sensor Data

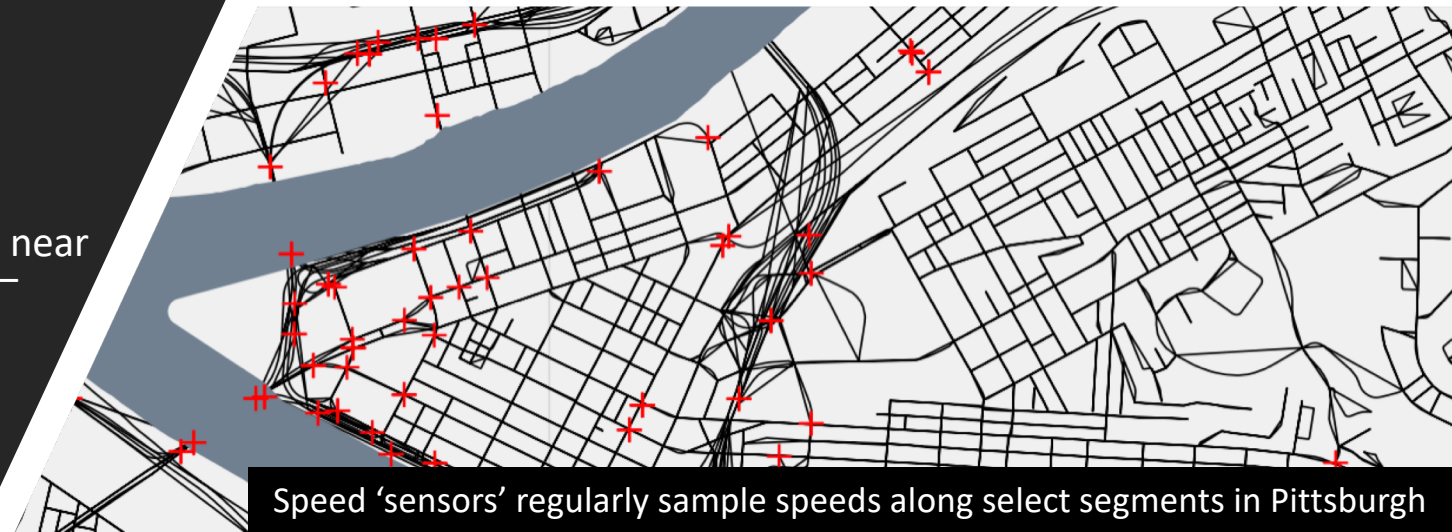
Traffic Data and Problem

- Speed reading every 5 minutes on some road segments in US cities (Pittsburgh and Washington, D.C.)

- Spatially infer speed values for the whole network (unobserved locations)
- Temporally infer speeds at future time steps
  - Fine-grained inferences → Needs accuracy and efficiency for real-time use cases

- **Main idea**: Efficient clustering of spatiotemporally correlated sensors using **Gaussian process**
  - "Everything is related to everything else, but near things are more related than distant things" – Tobler's first law of geography (Tobler, 1970)

Spatially inferred speed distribution for the whole road network

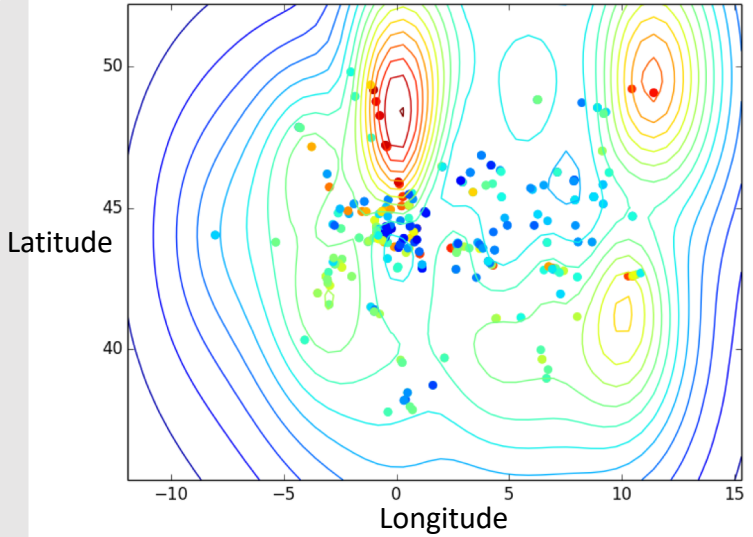


Speed 'sensors' regularly sample speeds along select segments in Pittsburgh

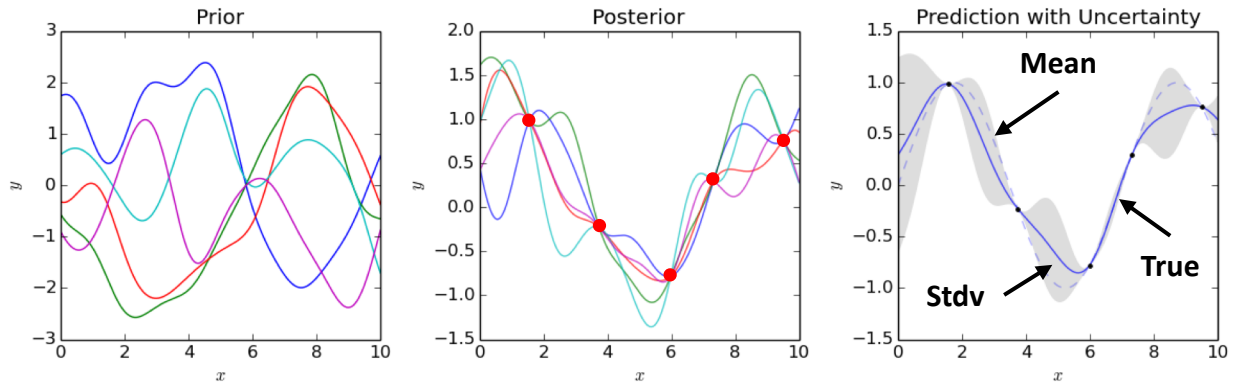
What is a Gaussian Process (GP) and Why?

- GP is a non-linear regression technique that encapsulates ‘closeness’ in space and time via *kernel functions* – which can further incorporate *other* features

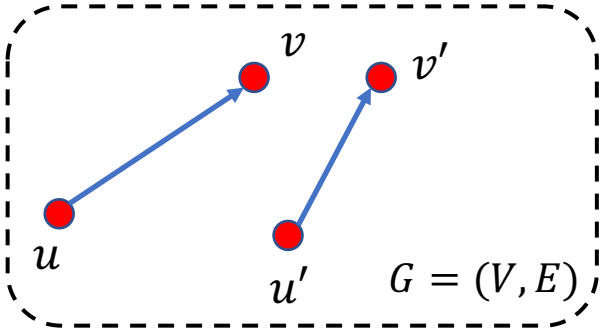
- *GP defines a prior over functions, which can be converted into a posterior over functions once we have seen some data, which can then be used for Bayesian regression.*



Traffic speed sample in Pittsburgh



Spatiotemporal GP Kernel for Road Networks



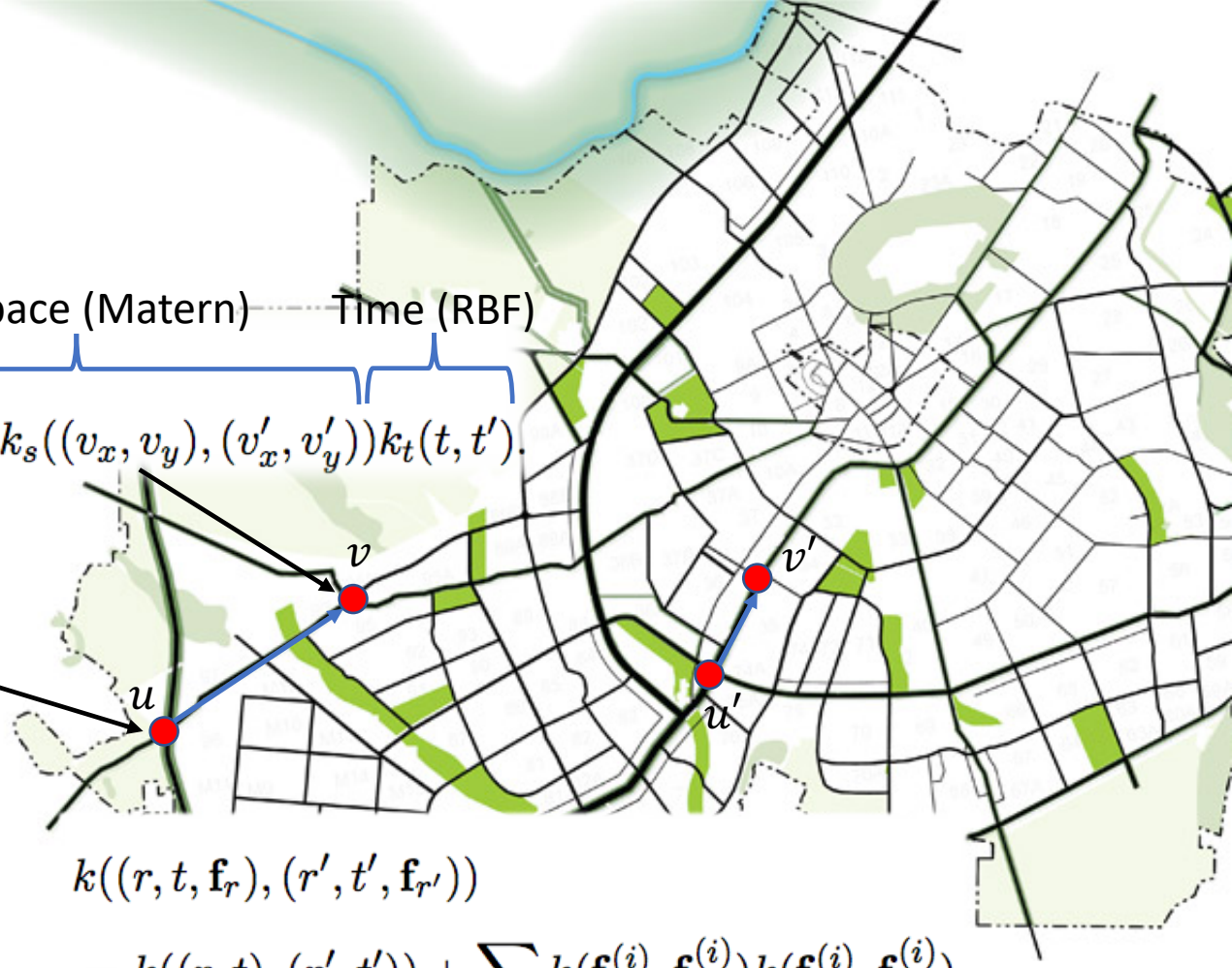
$$k((u, v), (u', v')) = k(u, u')k(v, v'),$$

Yu, K., & Chu, W. (2008). Gaussian process models for link analysis and transfer learning. In *Advances in Neural Information Processing Systems* (pp. 1657-1664).

Space (Matern) Time (RBF)

$$k((r, t), (r', t')) = k_s((u_x, u_y), (u'_x, u'_y))k_s((v_x, v_y), (v'_x, v'_y))k_t(t, t').$$

(lon, lat) of node u

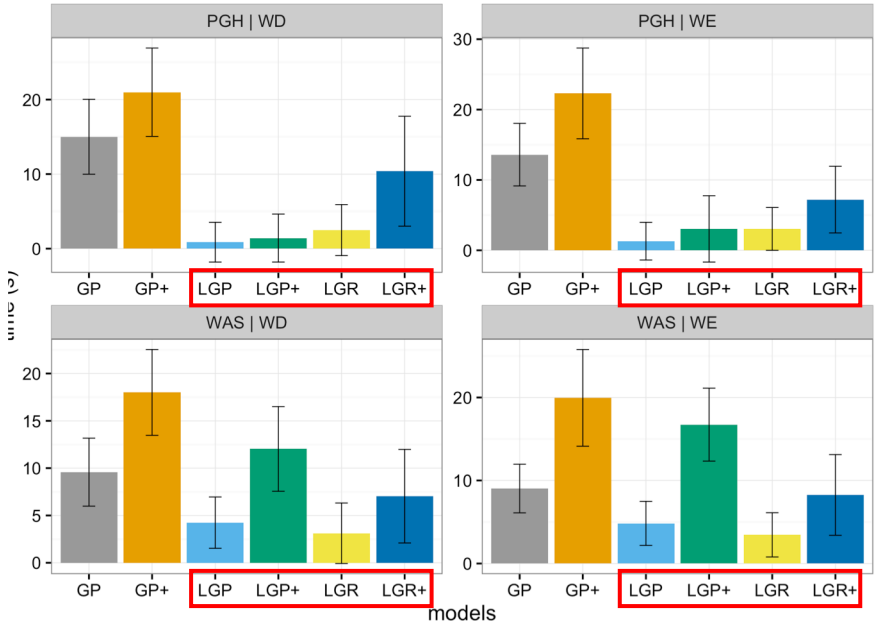


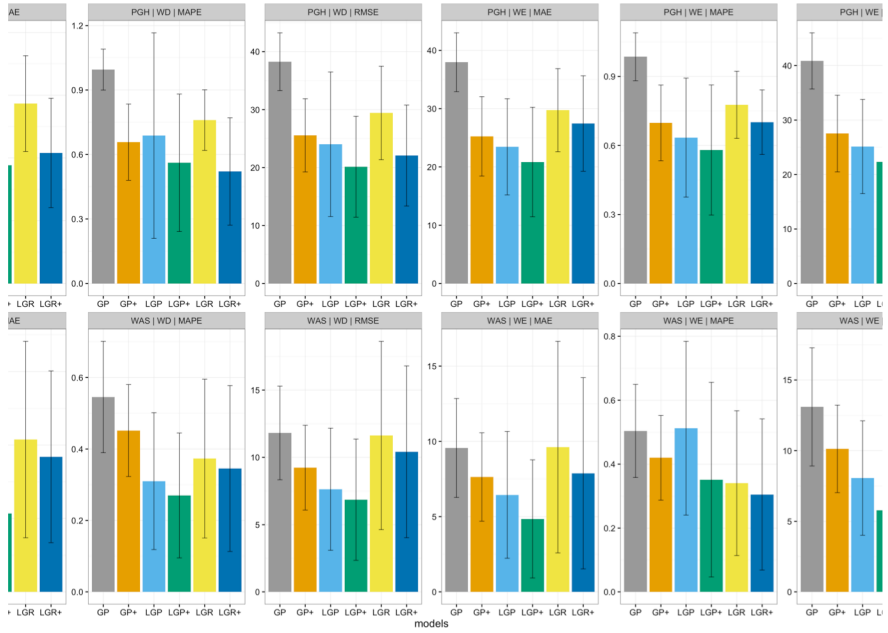
$$k((r, t, \mathbf{f}_r), (r', t', \mathbf{f}_{r'})) = k((r, t), (r', t')) + \sum_i k(\mathbf{f}_u^{(i)}, \mathbf{f}_{u'}^{(i)})k(\mathbf{f}_v^{(i)}, \mathbf{f}_{v'}^{(i)}) + \sum_j k(\mathbf{f}_{(u,v)}^{(j)}, \mathbf{f}_{(u',v')}^{(j)}),$$

Node-wise features

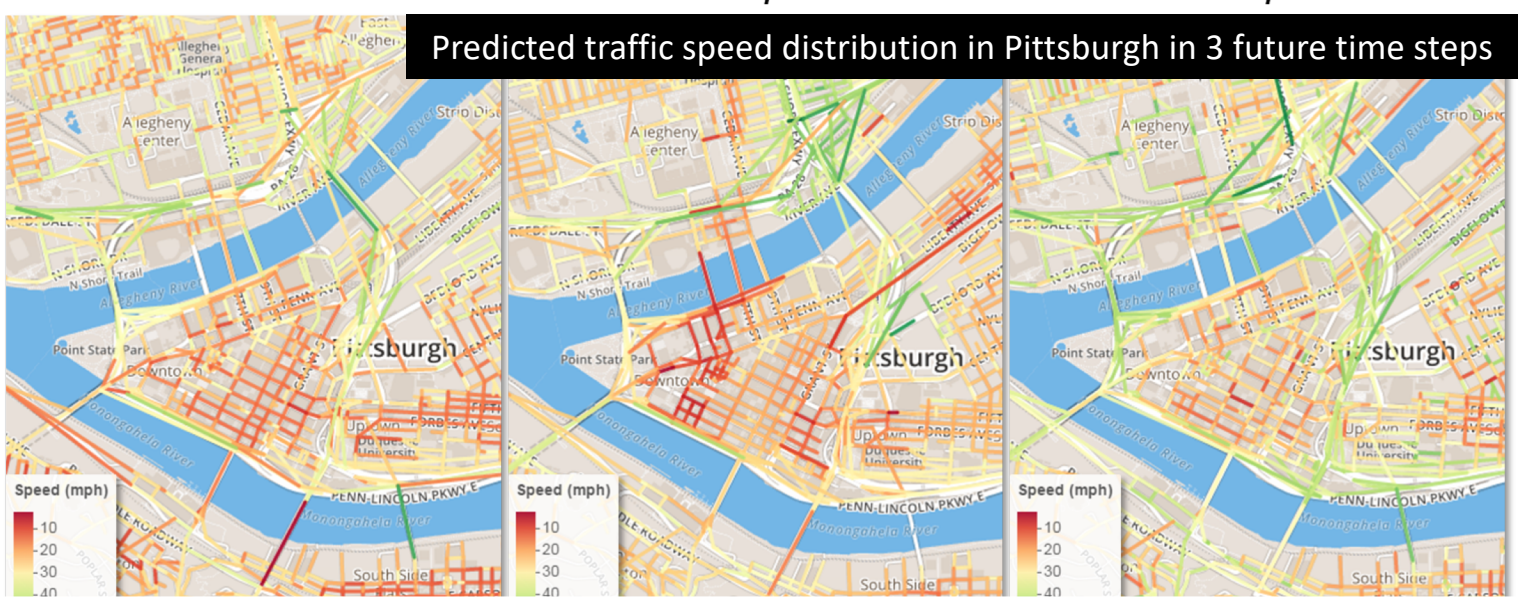
Edge-wise features

Feature	Description
Longitude, latitude	Longitude and latitude coordinates of the two endpoints (nodes) of a segment.
Segment length	Length (in miles) of a segment.
Number of lanes	The number of lanes a segment has in each direction.
Direction	Direction of a segment: northbound, southbound, eastbound, or westbound.
Degree	Degree of two end nodes of an edge (segment).
Betweenness	Edge betweenness centrality of a segment.
One-way	Is this segment one-way?
Road type	One of the 10 defined types: avenue, boulevard, bridge, lane, place, ramp, road,





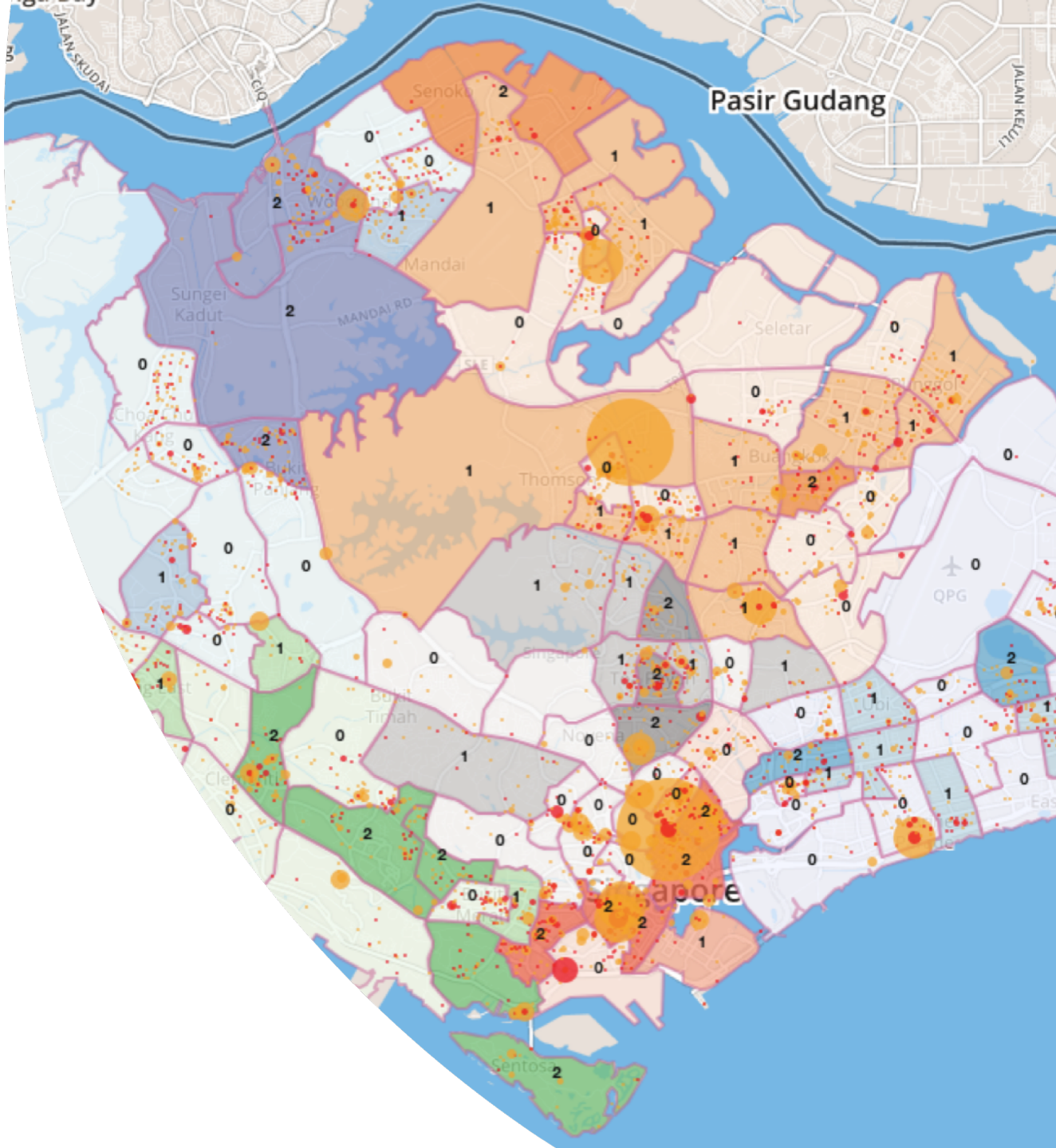
Mean errors measured in MAE, MAPE and RMSE



Experimental Results: Baselines and Comparison

Le, T. V., Oentaryo, R., Liu, S., & Lau, H. C. (2017). Local Gaussian Processes for Efficient Fine-Grained Traffic Speed Prediction. *IEEE Transactions on Big Data*, *3*(2), 194-207.

Incident Prediction for Urban Law Enforcement



Urban Law Enforcement

- Densely populated urban areas increasingly puts pressure on law enforcement agency's manpower trying to meet ever-rising demands
- Large databases of crime incidents provide fine-grained details:
  - Spatiotemporal (where and when)
  - Context (textual description, urgency classification, type and police response)
- It is possible to make high-precision predictions of incident occurrences using ML

Crime Data and Context

- Real-world data provided by a national law enforcement agency over multiple years
- Contains more than 500,000 reported incidents (e.g., from emergency calls)
- Each incident has: location, timestamp, type, urgency classification, dispatch and response information (incl. response time)
- Metadata containing neighborhood/sector boundaries and police deployment information

GP Kernel and Features

The spatiotemporal **kernel function** between bins i and j:

$$k((x_i, y_i, t_i), (x_j, y_j, t_j)) = k_s((x_i, y_i), (x_j, y_j))k_t(t_i, t_j),$$

Matern/RBF kernel: Spatial × Temporal

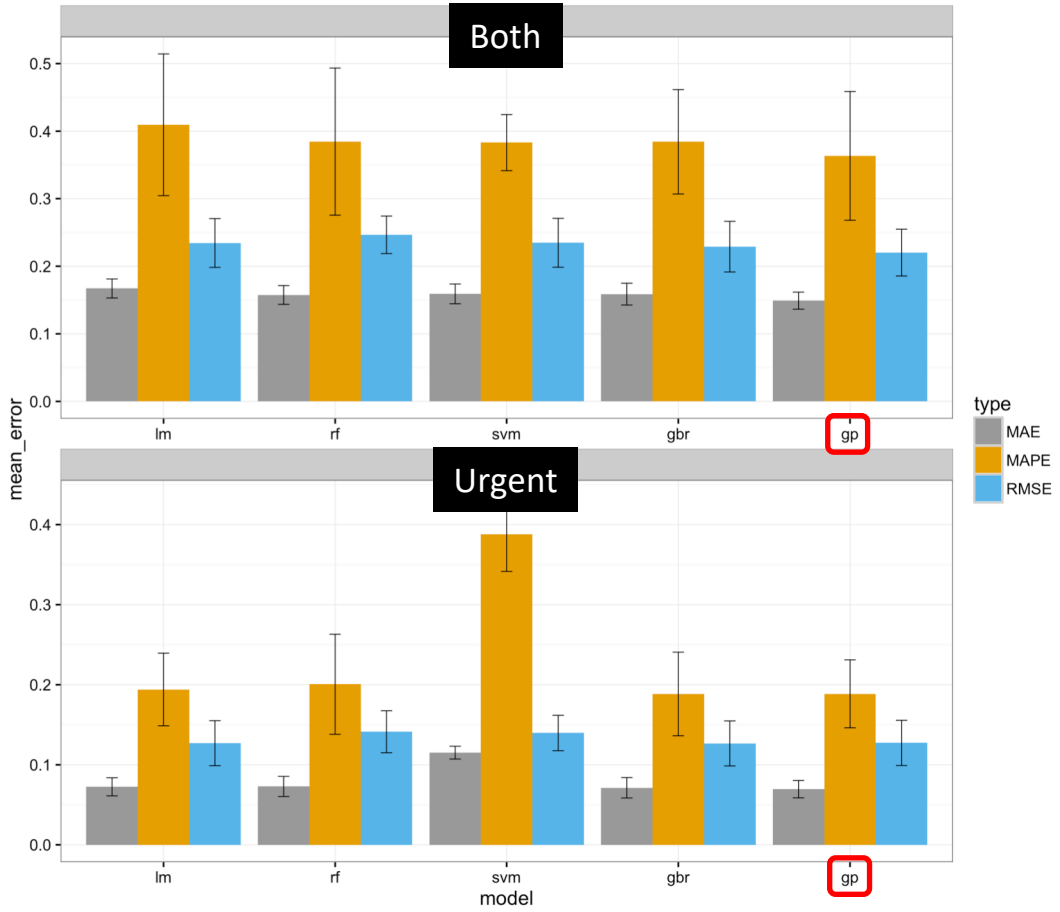
$$k((x_i, y_i, t_i, \mathbf{f}_i), (x_j, x_j, t_j, \mathbf{f}_j))$$

$$= k((x_i, y_i, t_i), (x_j, y_j, t_j)) + \sum_f k(\mathbf{f}_i, \mathbf{f}_j),$$

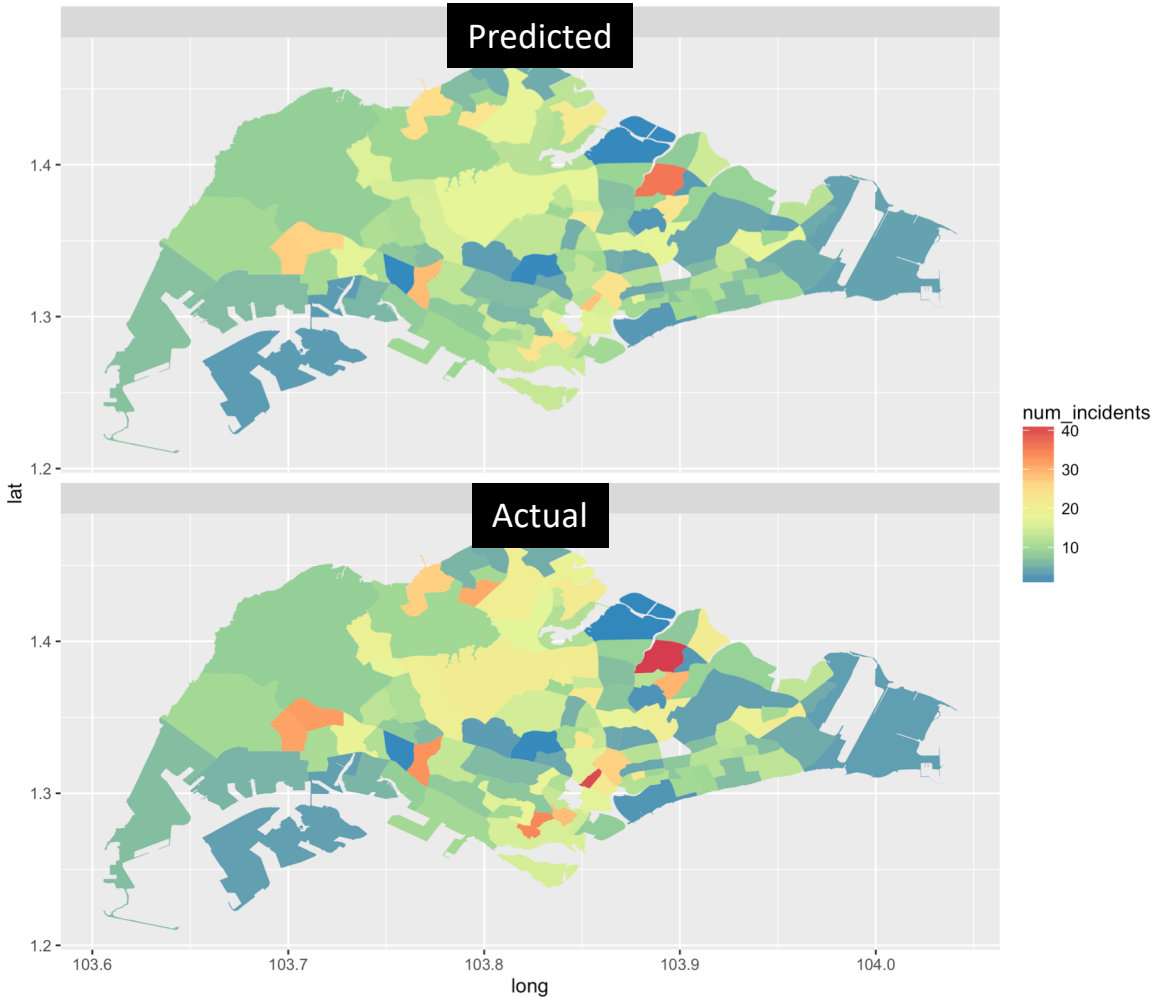
Linear kernel: Additive features

Feature	Description
`location`	Longitude and latitude coordinates of the incident
`hours`	The integer hours of the incident's occurrence time (0–23)
`is_weekend`	Binary variable whether the incident occurs on the weekend
`neighborhood`	Categorical variable specifying the incident's neighborhood
`sector`	Categorical variable specifying the incident's sector

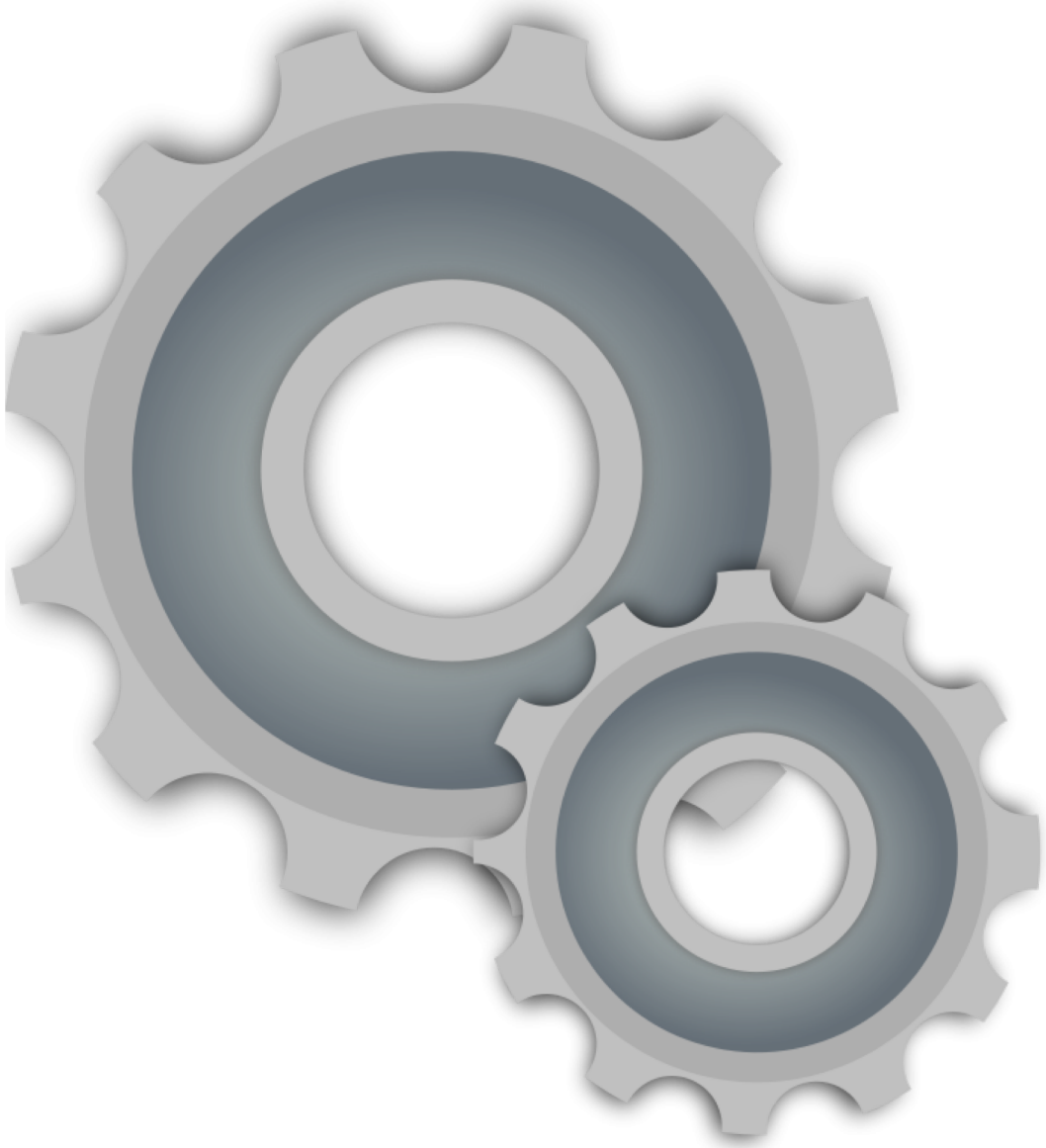
Experimental Results: Baselines and Comparison



- Linear regression (LM)
- Random forest (RF)
- Support vector machine (SVM)
- Gradient boosting regression (GBR)

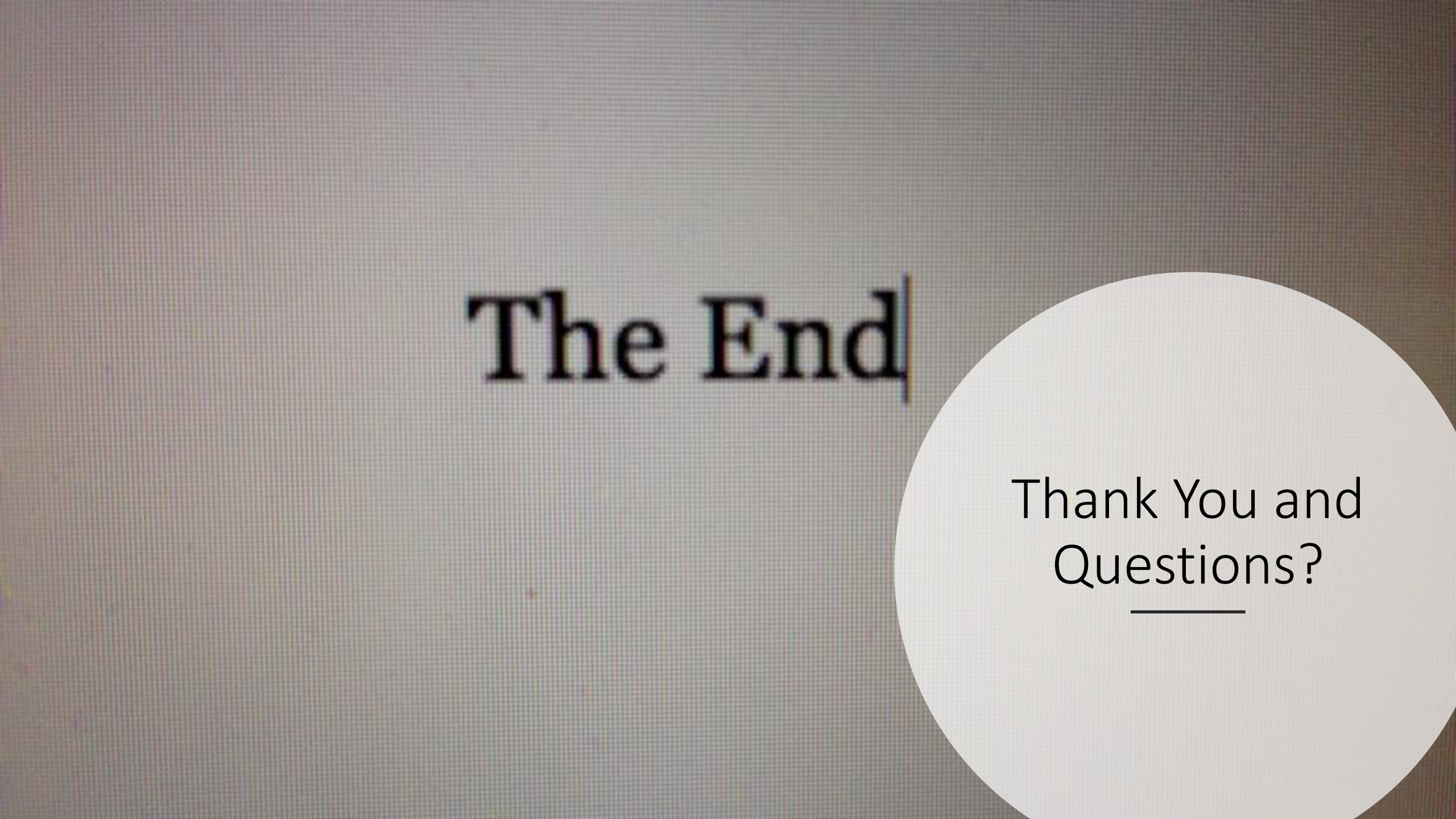


Comparing the **predicted** and **actual** number of incidents (**both** for weekday + weekend) for one particular test week



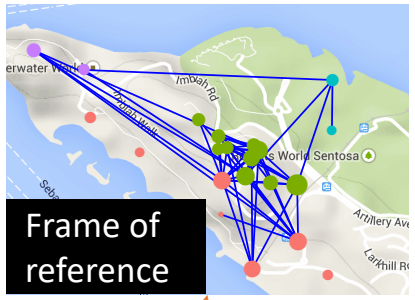
Urban Computing: Takeaways

- 3 B: Big Data, Big Challenges, Big Cities
- 3 M: Data Management, Data Mining, Machine Learning
- 3 W: Win-Win-Win for People, City, and the Environment
- 3 BMW



Thank You and Questions?

Appendix: The Framework



(HMMs)

(Kernel functions)

> Road networks
> Urban area boundaries

> Trajectory clustering
> Matrix factorization
> Incident types

LEARNING

① Spatiotemporal Clustering* → ② Environment Modeling → ③ Machine Learning Model M_j

S → $\mathbf{f}$

$\mathbf{H}_j$

$j \in \{1, \dots, K\}$

- Revealed Preference (RP)
- Reinforcement Leaning (RL)
- Gaussian Process (GP)

PREDICTION

$i \in T$ → $\mathbf{f}_i$ → ① Classification* → k → ② Environment Mapping → $\mathbf{H}_k$ → ③ M_k's Parameters Θ_k → ④ SOLVE

- Knapsack
- Decision Models
- GP Regression

> Logistic regression
> k-NN

D

Spatiotemporal Data