Deep Learning for Traffic Flow Prediction

Presented by Karis Matchaparn
Advisor Prof. Yu-Ling Hsueh

Outline

- Background
- Objective
- Short-term and Long-term Prediction
- Restricted Boltzmann Machine
- Methodology
- Study Flow
- Experiment result
- Conclusion
- Living & Traveling in Taiwan

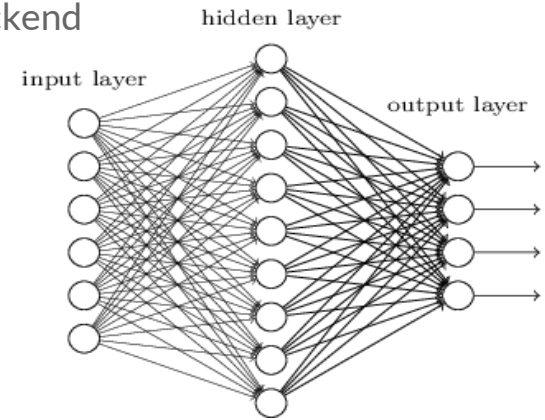
Background

Location-Based Service Technology

- Travel Planning Support
- Logistic Support

Deep Learning

- Finding Hidden features from data
- Training Predicting model
- Keras Library and Theano backend



DM Lab

Research Objective

construct a prediction model that collects the transportation data and predicts the traffic states (including car speed, and traffic congestion) during the period from 6.00 am to 9.00 am.

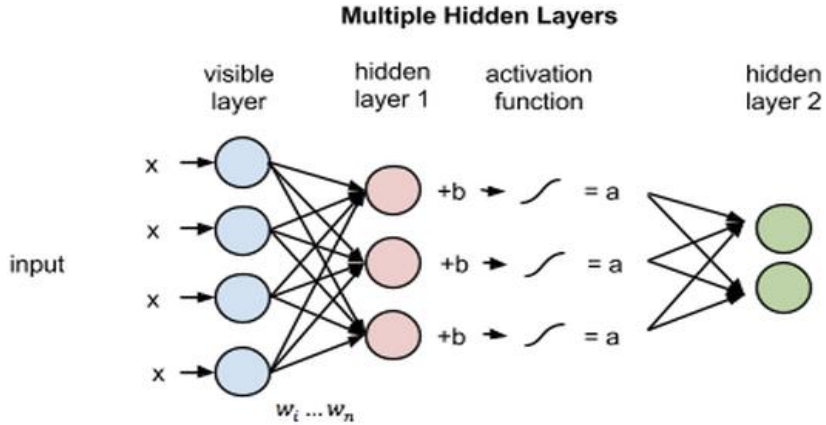
Short-term and Long-term Prediction

- 5-Minute Prediction
- Any Minute Prediction

“Information ahead of time improves planning travel routes better”

Restricted Boltzmann Machine

- Visible layer and Hidden Layer
- Restricted = no connection in the same layer
- Stochastic Gradient Descent as the model's optimizer
- Sigmoid activation function



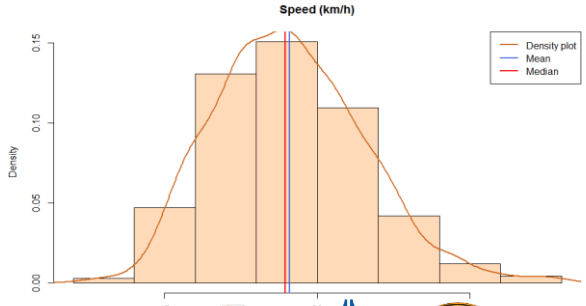
Methodology

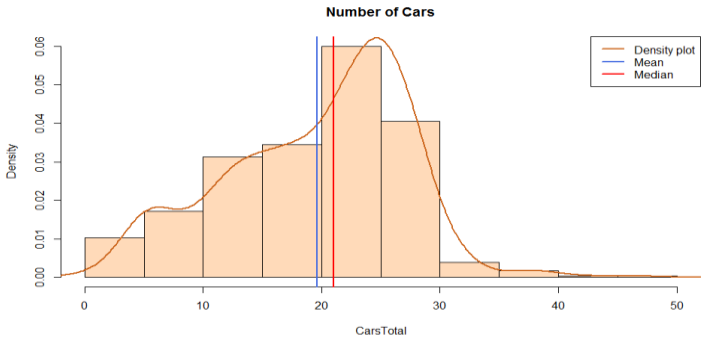
Dataset

- Sensor data were collected from Taiwan's highway.
- Various features include Time, Speed, Number of cars, Number of cars for three car types(small, medium, and large).
- Only data records in range of 6:00 a.m. to 9:00 am are considered.
- We use 75% of data as training data set and 25% of data as testing data set.

Data Pre-processing

- Smoothing speed with LOWESS
- Selecting speed which has Pearson correlation more than 0.8





DM Lab

Methodology

Pearson Correlation

- Calculate linear correlation between two variables.
- Value between -1 and +1

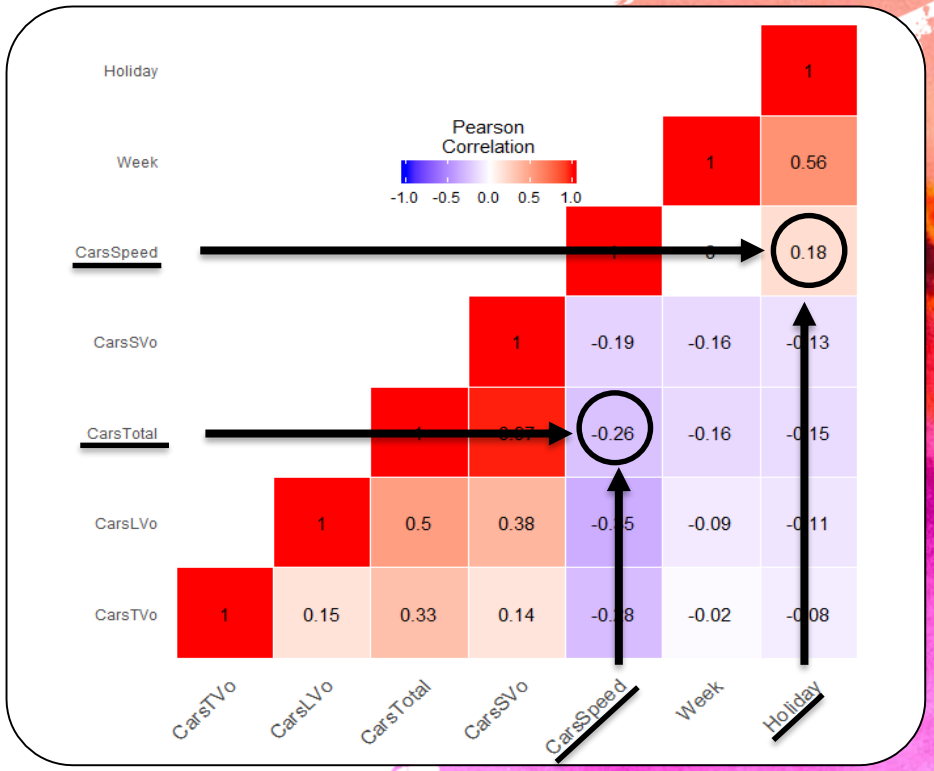
$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where:

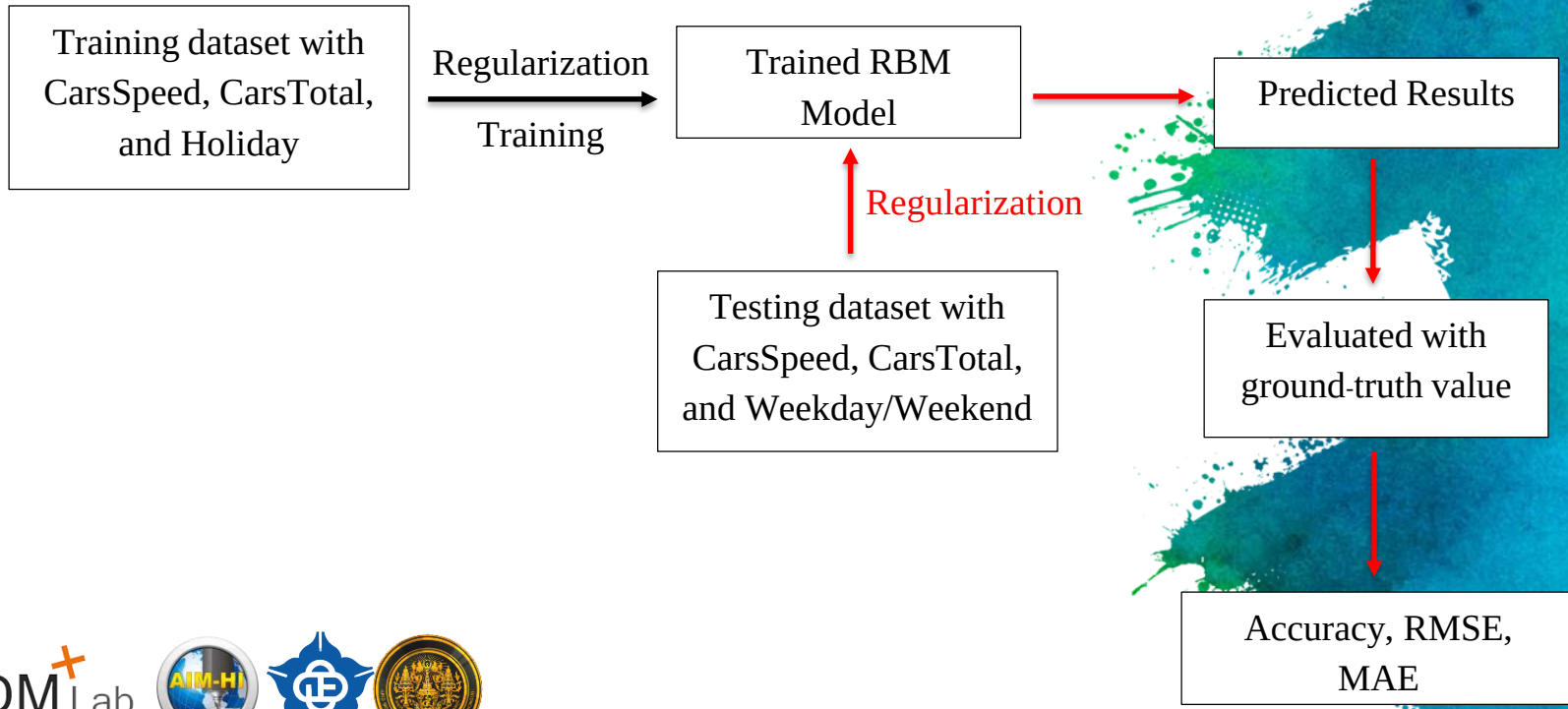
- n is the number of samples
- x_i, y_i are the single samples indexed with i
- $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (the sample mean); and analogously for $\bar{y}$

Feature Selection using Pearson Correlation

- CarsSpeed : Average Speed of Cars on the road
  - Correlation to CarsSpeed : 1
- CarsTotal : Number of Cars in sensor's range
  - Correlation to CarsSpeed : -0.26
  - Inverse of CarsSpeed
- Weekend : Data in weekday or weekend
  - Correlation to CarsSpeed : 0.18



System Flow


Training dataset with CarsSpeed, CarsTotal, and Holiday
Regularization
Training
Trained RBM Model
Predicted Results
Regularization
Testing dataset with CarsSpeed, CarsTotal, and Weekday/Weekend
Evaluated with ground-truth value
Accuracy, RMSE, MAE

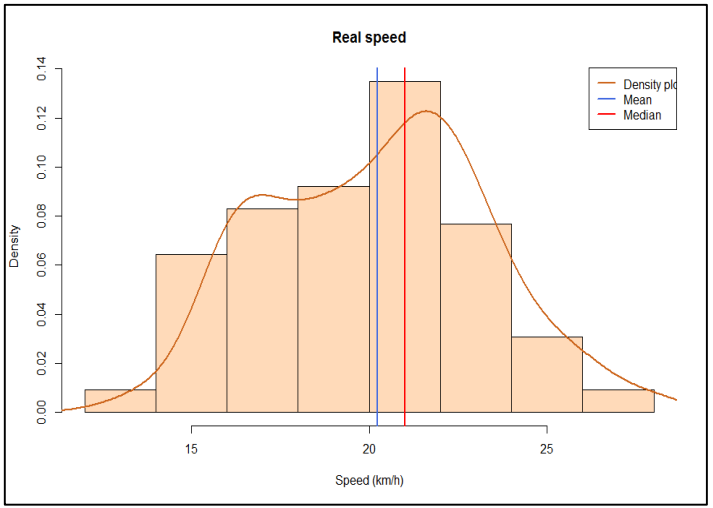
Experiment Results I

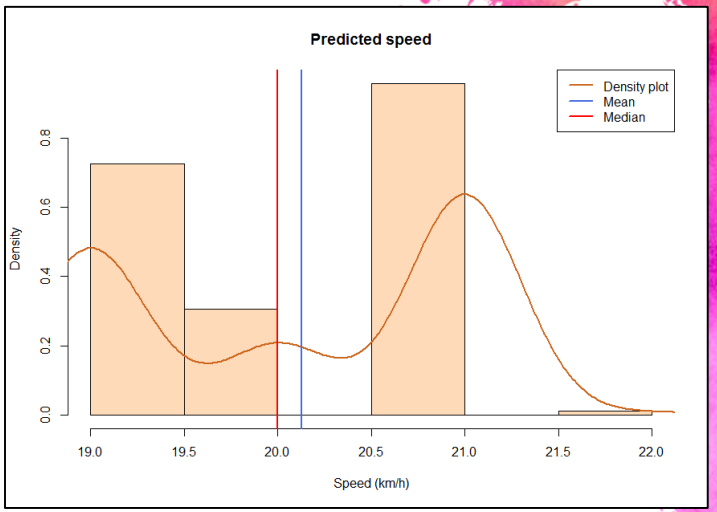
Example of speeds between real speed and predicted speed

- 10 examples

Real	20	21	22	23	24	23	22	20	19	20
Predict	21	21	19	19	19	19	19	19	20	20

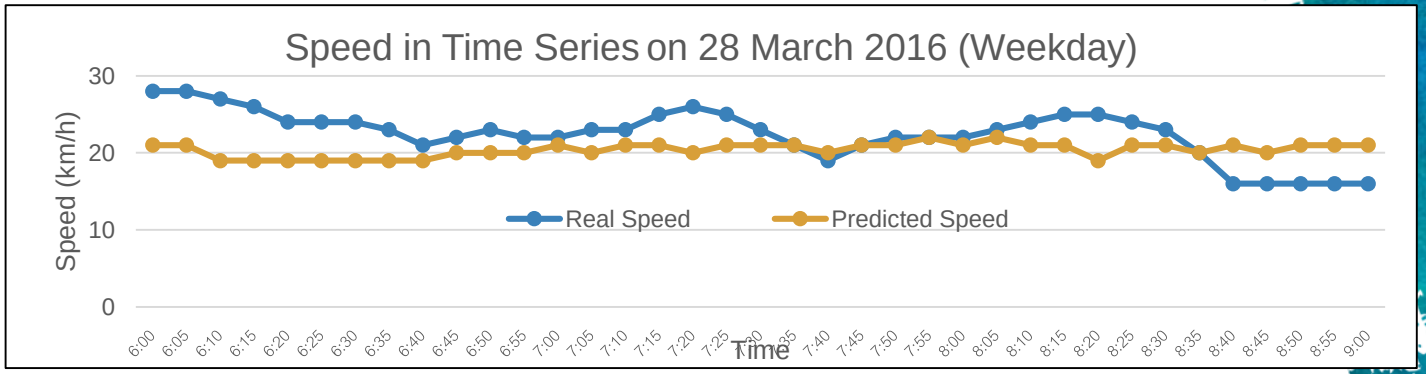
Histogram of speeds

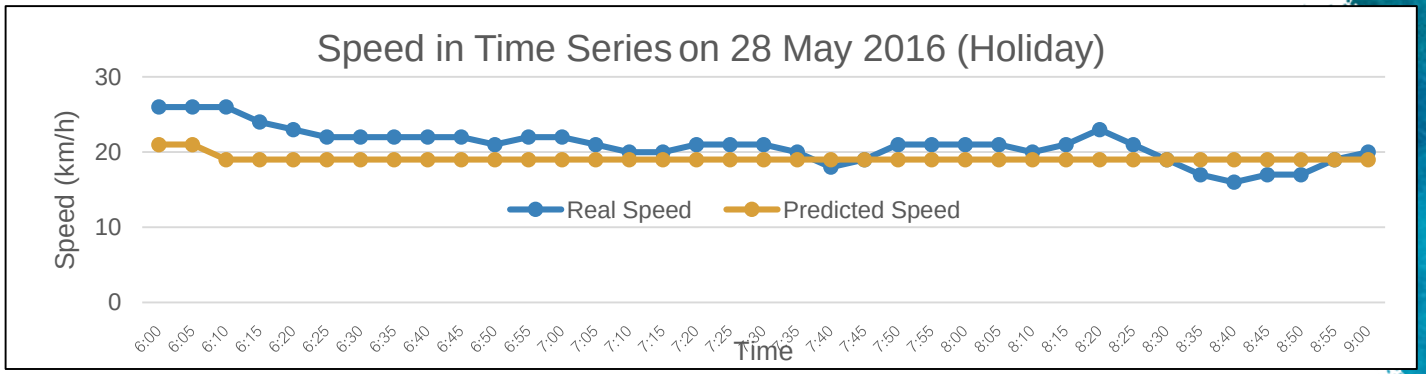




Experiment Results II

Compare real and predicted speed in weekday and weekend





Evaluation

Root Mean Square Error

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Mean Absolute Error

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|\,.$$

3.48%

Root Mean Squre Error (RMSE)

2.93%

Mean Absolute Error

Conclusion

- Prediction application can be built based on the RBM model to help plan and find the best route to reach the destination
- The model can be achieved up to 95%

Living and Traveling in Taiwan

Special thanks to..

× Data Management Laboratory

× Advanced Institute of Manufacturing with High- Tech Innovations

× National Chung Cheng University

× Dà chī restaurants, Family mart and 7-11

Travel Maps

Thanks!

Any questions?