

Neural Network

Mid-term Presentation

工科所 R04525102 張哲瑜
資管所 R05725003 葉文傑

Agenda

1. Methodology
2. Tool Kit
3. Performance
4. Discussion
5. Future Work

Methodology

- 預測 Oct. 18 到 Oct. 24 之 08:00 - 10:00 及 17:00 - 19:00 的數據。
- 比賽中有提供Aggregate的Code,
利用它可以生成Training Data每個Time Window的平均數據。
- 目前採用最直觀的想法, 就將題目視為時間序列的資料, 並透過時間序列的模型分析。
 - 固定路線後, 時間序列中把 Window設定成一星期。
 - 舉例來說找出 Oct. 18 之前每週的08:00的時間當成歷史資料, Oct. 18 08:00當成預測的資料。

Date	Travel_time
2016/7/19	176.96
....	
2016/10/4	206.53

Methodology

- ARIMA (自動回歸整合移動平均模型)

一種用於穩健時間序列的熱門預測模型，這個模型一旦被識別後就可以從時間序列的過去值及現在值來預測未來值，整合了AR(自動回歸) MA(移動平均)兩種預測時續資料的模型。

- AR 為利用過去時間的變數來預測未來的結果

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

- MA 為利用過去時間的加權誤差平均來預測未來的時間點結果

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}$$

Methodology

- ARIMA(p,d,q)的三種參數：

p 表示自動回歸 (AR) 項數

d 表示做幾次差分，差分越多，時間序列也越趨於穩健

q 表示移動平均 (MA) 項數

Tool Kit

- Language: Both Python 3
- Library: Pandas, matplotlib, Pyflux
- Machine: Mac Pro Mid 2014 (Program can within 1 minutes)

Performance

- Task 1: Travel Time Prediction
 - 0.3072 (Rank:378) \rightarrow 0.2239 (Rank:319)
- Task 2: Volume Prediction
 - 0.2832 (Rank:279) \rightarrow 0.2182 (Rank:252)
- Parameter
 - Bic matrix v.s manually setted

Discussion

- Daily Evaluation
 - As mentioned, Parameter p, q need to be tested.
 - Enthusiasm
- Data Sparse
 - Task 1: 7 points
 - Task 2: 4 points

Future Work

- Reduce Noisy Data
- Try Other Time series Model (e.g Dynamic autoregression model)
- Take other factors into account (e.g Weather data, Vehicle Type and so on)
- Testing Data may also provide good information.