The Simpler The Better: A Unified Approach to Predicting Original Taxi Demands based on Large-Scale Online Platforms

Yongxin Tong\textsuperscript{1}, Yuqiang Chen\textsuperscript{2}, Zimu Zhou\textsuperscript{3}, Lei Chen\textsuperscript{4}, Jie Wang\textsuperscript{5}, Qiang Yang\textsuperscript{2,4}, Jieping Ye\textsuperscript{5}, Weifeng Lv\textsuperscript{1}

\textsuperscript{1} SKLSDE Lab, Beihang University, \textsuperscript{2} 4Paradigm Inc., \textsuperscript{3} ETH Zurich, \textsuperscript{4} Hong Kong University of Science and Technology, \textsuperscript{5} Didi Research

\{yxtong, lwf\}@buaa.edu.cn, \textsuperscript{2} chenyuqiang@4paradigm.com, \textsuperscript{3} zimu.zhou@tik.ee.ethz.ch, \textsuperscript{4} \{leichen,qyang\}@cse.ust.hk, \textsuperscript{5} \{wangjiejacob,yejieping\}@didichuxing.com

Target: UOTD Prediction

- OTD (Original Taxi Demand)
  - The number of taxi-calling orders \textit{submitted} to the online taxicab platform
  - Including orders that are cancelled finally

- UOTD (Unit Original Taxi Demand)
  - The number of taxi-calling orders \textit{submitted} to the online taxicab platform \textit{per unit time and per unit region}

Why we need to predict UOTD

- Expand Potential Market
- Assess Incentive Mechanisms
- Guide Taxi Dispatching

Key Methodology

- Two paradigms can be chosen
  - Complex (non-linear) models
  - Simple (linear) models

- In industrial practice, the latter one (Simple models + Massive features) is preferred
  - As it can Transform Model Redesign to Feature Redesign

Feature Engineering

- Basic Features
  - Temporal Features
  - Spatial Features
  - Meteorological Features
  - Event Features

- Combinational Features
  - Combine basic features based on business logics and data analysis
  - Express the joint influences of different basic features in a simple model

Model & Training

- Model
  - A linear regression model with high-dimensional features and a spatiotemporal regularizer

\[
\text{obj}_{\text{spatio-temporal}}(\mathbf{w}) = \sum_{X \subseteq D} \phi(\text{car}(\{w^T x| x \in X\}))
\]

- Distributed Training

Experiments