講演番号：9

都営バスのリアルタイム運行データを用いた渋滞検知の試作

Prototype of Traffic Jam Detection Using Real-time Operation Data of Toei Buses

大箸 大生† 田中 悠† 笠松 大佑†
Daisei OHASHI† Haruka TANAKA† Daisuke KASAMATSU†
† 創価大学理工学部情報システム工学科
† Faculty of Science and Engineering, SOKA University

1. はじめに

本研究の目的は，渋滞の発生頻度が高い都営バスのリアルタイム運行データを対象として渋滞検知の試作に取り組むことである．

2. 課題

先行研究[1]では，西東京バスのリアルタイム運行データを用いた渋滞検知の試作が提案されている．西東京バスを対象にリアルタイム到着時刻と正解データとなる渋滞情報を約 1 カ月間収集してデータセットを作成した．リアルタイム到着時刻はバスナビと呼ばれる Web サイトから取得し，正解データとなる渋滞情報は，NAVITIME と呼ばれる Web サイトから取得した．実験結果では，全区間に対する F1 スコアが高かったが，一部の停留所区間においては精度が低下した．原因として，渋滞のデータ数が少ないといったデータの偏りによる検知精度の低下が考えられる．よって，郊外エリアのバス運行の実データでは，渋滞の頻度が少ないため渋滞検知の精度を計測する実験が十分に行うことができない．本研究の課題は，より渋滞の発生が多いと考えられる都内を対象エリアとして研究を行い渋滞検知の精度を向上することである．

3. 提案手法

提案手法の学習パイプラインを図 1 に示す．複数のデータ提供元からデータ収集を行い，取得したデータの前処理、整理、結合を行い，そのデータを用いて機械学習を行う．

学習データとして，収集したデータの整理、統合を行い，バス停名，日付，曜日，バスの到着予定時刻，到着時刻と時刻表の誤差，バス停周辺の混雑数，降水量，の 7 つのデータを特徴量として使用する．正解データとして，バスの渋滞情報を用いる．特徴量選択として，フィルタ法を用いて特徴量の最適化を行う．

学習方法として，Python 用の機械学習ライブラリである scikit-learn を使用する．機械学習の一種であり，教師あり学習として分類・回帰が行えるランダムフォレスト，SVM，決定木やロジスティクス回帰，アンサンブル学習であり弱学習器を組み合わせ学習を行う AdaBoost，同じくアンサンブル学習でひとつ前の学習器の誤りを次の学習器で修正するのを繰り返し行う XGBoost を使用する．

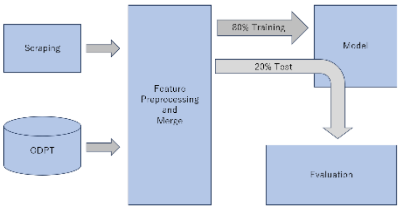

図 1 学習パイプライン

表 1 バス停の全区間に対する結果

アルゴリズム	acc	pre	rec	F1
SVM	0.690	0.681	0.967	0.799
ランダムフォレスト	0.940	0.964	0.940	0.952
AdaBoost	0.940	0.958	0.954	0.956
XGBoost	0.946	0.960	0.953	0.957
ロジスティクス回帰	0.900	0.939	0.906	0.922
決定木	0.950	0.960	0.956	0.958

4. 実験

北 47 系統バス停留所の「北千住駅前」から「公園前」までの全区間に対して分類器を使った精度，各区間ごとでの精度の二種類を求めた．学習データとテストデータの分け方として，前半 80%を学習データ，後半 20%をテストデータとした．

バス停留所の全区間に対して，複数の学習モデルを適用した実験結果を表 1 に示す．決定木が 0.958 の F1 スコアとなり最も高い精度を得た．先行研究と同じ特徴量数を用いた最高 F1 スコアと比較すると，約 0.041 向上した．

5. まとめ

本研究では，渋滞の発生頻度がより高い都営バスを対象にしてデータ収集を行い渋滞検知の試作を行った．渋滞のデータ数が少ない先行研究と比較し，F1 スコアの向上を確認することができた．

参考文献

[1] 野田悠人，笠松大佑 ”西東京バスのリアルタイム運行データを用いた渋滞検知の試作，” 電子情報通信学会東京支部学生会研究発表会，19，2023.

Copyright © 2024 IEICE