

多変量LSTMによる短期交通量予測： 鎌倉市の観光交通を対象として

小川 晃平¹・福田 大輔²

¹学生会員 東京工業大学修士課程 環境・社会理工学院 (〒152-8552 東京都目黒区大岡山2-12-1)

E-mail: k.ogawa@plan.cv.titech.ac.jp

²正会員 東京工業大学准教授 環境・社会理工学院 (〒152-8552 東京都目黒区大岡山2-12-1)

E-mail: fukuda@plan.cv.titech.ac.jp

近年のセンシング技術の発展によって多様な交通情報をリアルタイムで収集することが可能となり、蓄積されたビッグデータを使用した交通状態の短期予測に関する研究が進展している。しかし、観光地の交通量予測への適用事例は少なく、オーバーツーリズムの評価等への活用にも至っていない。観光地の交通量を高い精度で予測するためには、曜日や季節、天気、イベントの有無などの多様な特性を包括的に考慮することが必要であり、本研究ではそれらを考慮可能な予測手法として多変量型LSTMを用いたモデル分析を行った。具体的には、神奈川県鎌倉市中心部を対象に、市内の交通量データ、降水量や社会イベント情報などの複数の入力に対して学習を行い、市内中心部における交通量の60分先予測値を出力するモデルを構築した。時間帯別、季節別のモデル出力誤差率などを検証して、本モデルの予測性能を確認したところ、混雑状況下における交通量の予測を誤差率15.33%で行うことができた。

Key Words : *deep learning, traffic volume prediction, long short-term memory*

1. はじめに

近年、センサーや観測機器に代表されるセンシング技術の発展により膨大な交通情報をリアルタイムで収集することが可能となった。蓄積されたビッグデータを使用して将来予測を実施することはAIの典型的な活用事例であり、交通の分野でも日々研究が進められている。しかし観光交通が卓越した地域における短期交通量予測の適用事例は未だ少なく、オーバーツーリズムの評価等への活用にも至っていない。通学や通勤による交通混雑は、発生する時間帯がある程度特定されているのに対し、観光交通による交通混雑はこの限りではない。特に、曜日や季節、天気、イベントの有無によって観光交通の変動は大きく異なる。そのため、このような特徴を持つ交通変動をいち早く検出する短期交通量予測システムの確立が求められている。

本研究では、数ある観光地の中でも神奈川県鎌倉市を対象として、交通混雑緩和のための政策実施を想定した短期交通量予測モデルの構築に取り組む。交通量を動的に予測する代表的手法である交通シミュレーションなどでは、対象地の道路ネットワークデータやOD交通量など詳細な情報を準備する必要がある。道路ネットワーク

やOD表を構築すること自体は決して難しいものではないが、これらは時間的に不変のものではない。道路ネットワーク自体の変化は目まぐるしいものではないが、OD表等の交通インプット情報は曜日や季節の移り変わりと共に大きく変化する。本研究の分析対象である鎌倉市を例に挙げると、市内に多く存在するあじさいの名所には開花時期の6月から7月上旬には例年多くの観光客が訪れる。鎌倉市に限らず一般に観光交通需要には大きな季節変動が存在するため、厳密に交通量を予測するためには曜日や季節ごとのOD表を多数構築する必要があるが、容易な作業ではない。

そこで本研究では、分析手法として深層学習に着目する。本研究では、鎌倉市中心部における車両の時間的な集中度合いを短期で予測するための分析枠組みを構築することを目的としており、本モデルは経路・リンク交通量やエリア内の車両総旅行時間を短期予測するものではない。むしろ、観光交通需要の季節性、イベントの有無、降水量の影響等を考慮することで、短期の交通量予測の精度がどこまで向上しうのかを明らかにしたいと考えている。深層学習は、入力されたデータを学習することでその特徴を包括的に考慮し、極めて詳細な予測を実施する。そのため上述の分析目的に適した手法であること

に加えて、交通需要マネジメント政策をリアルタイムで実施できることが期待される。以上から、本研究では深層学習を用いて鎌倉市中心部の交通量を短期予測するモデルを構築する。交通量などのデータは時間方向に強い依存性を持っていることから時系列データと呼ばれる。深層学習において時系列データの予測に適しているモデルにはLong short-term memory (LSTM)¹⁾がある。後述するように、LSTMを活用した既往研究はこれまでも見られるが、その多くは通勤交通等が支配的な都市部道路等における短期の交通量や旅行速度を主眼としたものであり、主に一変量型のLSTMに基づく短期予測が行われている。これに対し本研究では、鎌倉市中心部に位置する下馬の交通量、下馬に繋がる周辺部複数地点の交通量、降水量、市内で開催されるイベントの有無などを入力データとして下馬の60分先の交通量予測値を出力する枠組みを、多変量型LSTMに基づいて構築する。

2. 既往研究と本研究の位置づけ

(1) 既往研究

過去数年に渡り、深層学習を利用した交通状態の短期予測に関する研究が数多く行われてきた。例えばFu et al.²⁾は、カリフォルニアの都市高速道路に設置された固定感知器から5分ごとの通過交通量を取得し、自己回帰和分移動平均モデルと深層学習モデルLSTM、GRUそれぞれを用いて予測を実施し、その精度を比較した。予測精度比較の指標には平均絶対誤差 (Mean Absolute Error / MAE) と平均二乗誤差 (Mean Squared Error / MSE) が用いられ、LSTMは既存の回帰モデルよりも約5.8%低いMAEを示したことが確認された。またZhao et al.³⁾では、同じくLSTMによって北京内の主要道路3地点の交通量が予測された。1時間あたりの交通量が約2,000台に達する地点で平均絶対誤差率 (Mean Absolute Percentage Error / MAPE) が約6.41%という結果が算出された。

先に挙げた二つの例は、過去の交通量データのみを入力データとして予測を行う一変量型モデルであったが、降水量や曜日、時間帯などの要因が交通量の変化に影響を与えることは容易に想像できる。複数の入力から交通現状値を出力した例としてはKe et al.⁴⁾やXu et al.⁵⁾などがある。前者はFusion convolutional long short-term memory network (FCN-Net) という畳み込みLSTM層、LSTM層、Convolutional Neural Network (CNN) 層を組み合わせた複数層モデルを提案し、交通需要、平均旅行時間、時間帯・曜日、天気の入力から交通需要を出力した。後者は交通需要、天気、大気環境データ、土地利用データを入力として、交通需要を出力するステーションフリー自転車シェアのための交通需要予測モデルを提案した。推定の結

果より入力データが交通需要単独の予測よりも複数データの入力による予測の方がMAPEが低下することが明らかとなった。

(2) 本研究の位置づけ

既往研究のレビューより、交通量の短期予測にはLSTMに代表される時系列学習モデルが適していることが示唆された。しかし先に述べたように、交通量の変動が大きい観光地を対象とした予測には季節性や時間帯、天気など交通情報以外の要因を考慮することが重要となる。そこで本研究では、Ke et al.⁴⁾およびXu et al.⁵⁾を拡張し、LSTMのモデル構造のもとで多変量型の交通量短期予測モデルを提案する。なお、本研究の先行研究であるOgawa and Fukuda⁶⁾では、1年間のデータを使用して分析を実施していたが、学習に用いるデータと予測性能の評価に使用されるデータに同じ季節が含まれておらず、季節性に対する検証を正確に実施できていないと考えられる。そのため本研究では新たに、1年間以上のデータを基に予測モデルの構築を行った。

3. 短期交通量予測の詳細

(1) 分析対象

本研究では鎌倉市中心部に位置する複数地点の通過交通量を短期予測する深層学習モデルの構築に取り組む。鎌倉市は年間の入込観光客数が2,000万人を超える人気観光都市である。四方を海と山に囲まれた閉鎖的な地形と単純な道路ネットワーク構造により車両の進入口が限定されてしまい、交通混雑が発生しやすい背景を持つ。交通混雑の緩和には計画的な道路規制や混雑課金等の政策実施が必要となるが、本研究ではその前段階の道路状況の分析として、市内中心部の交通量の60分先の予測値を出力する。予測対象は「下馬 (横浜鎌倉線下り)」、 「下馬 (横浜鎌倉線上り)」、 「下馬 (横浜鎌倉線上下合計)」の3種類の交通量である。なお、予測先の時間を60分と設定したのは、都心から鎌倉市中心部までの車での所要時間が約70分であることから、60分先の交通量を出力することでドライバーのトリップ選択行動を補助できると判断したためである。また、予測精度の評価には式(3a)、(3b)に示す二乗平均平方根誤差 (Root Mean Squared Error / RMSE) と平均絶対誤差率 (MAPE) を用いることとした。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{obs,i} - y_{pred,i})^2} \quad (3a)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_{obs,i} - y_{pred,i}|}{y_{obs,i}} \quad (3b)$$

(2) データの概要

a) 交通量データ

日本道路情報センター (JARTIC)⁷⁾から図-1に示す7地点の交通量データを取得した。このエリアでは現在、路上カメラによる道路状況の撮影が行われており、将来的には映像解析による交通データが入手可能とされているが、今回の分析では固定感知器による交通量データを使用する。データ期間は2017年1月1日から2018年12月31日である。交通量は5分単位で記録されていたが、本分析では15分単位で集計しデータセットを作成した。分析を実施した当初は24時間分の交通量データを使用する予定であったが、1日の中で比較的交通量が少い深夜や早朝の時間帯を入力データとする必要がないと考え、午前8時から午後8時までの交通量を使用した12時間周期のデータを学習に使用した。



図-1 交通量データ取得地点

表-1 鎌倉市で開催される主な社会イベント

イベント名	開催日 (2017年)	開催日 (2018年)
かまくらまつりパレード	4月9日～16日	4月8日～15日
大船祭り	5月21日	5月20日
鎌倉花火大会	7月19日	7月24日
あじさい祭り	6月中旬～ 7月上旬	6月中旬～ 7月上旬
大船 to 大船渡	9月24日	9月30日

b) 降水量データ

気象庁より1時間単位の降水量データを取得した。データ期間は2017年1月1日から2018年12月31日である。データ取得可能地点に鎌倉市が含まれていなかったことから、隣接する三浦市の降水量データを代替とした。

c) 鎌倉市イベントデータ

鎌倉市観光課に問い合わせたところ、表-1に示すイベントが来訪者が特に多いという回答を得たので、交通量に影響を及ぼすと判断し分析に取り入れることとした。

(3) 多変量LSTM

交通量などの時系列データを分析することに適した深層学習モデルの代表例⁸⁾にはLSTMがある。

LSTMはRecurrent Neural Network (RNN) が抱える勾配消失問題を解消した高次モデルであり、モデルの全体構造は図-2のようになる。隠れ層にはループ構造が備わっており、過去の情報を適切なタイミングまで保持して次の層へ伝播する仕組みである。ゆえにある時点 t の隠れ層 $h(t)$ は同時刻の入力層 $x(t)$ と一時刻前の隠れ層 $h(t-1)$ から信号を受け取ることとなる。なお保持する過去の情報量は自由に設定することができる。隠れ層内には図-3に示すLSTMブロックが設置されており、このブロックはメモリーセル、入力ゲート、出力ゲート、忘却ゲートで構成されている。式(3c)はメモリーセルの値、式(3d)は入力ゲートの値、式(3e)は出力ゲートの値、式(3f)は忘却ゲートの値、式(3g)は活性化された値、式(3h)は中間層の値、式(3i)は出力層の値、 U は入力層と隠れ層間の重み、 W は隠れ層間の重み、 V は隠れ層と出力層間の重み、 f, g は活性化関数、 b はバイアスペクトル、 \odot はアダマール積をそれぞれ表している。これらの装置には過去の信号をネットワーク内に保持する、保持した信号を必要なタイミングで次の層へ伝播させる、学習が完了し必要なくなった信号を消去するなどの役割が備わっている。これにより長期・短期の時系列データに対して効率よく学習を行うことが可能となった。

$$c(t) = i(t) \odot a(t) + f(t) \odot c(t-1) \quad (3c)$$

$$i(t) = \sigma(W_i x(t) + U_i h(t-1) + V_i c(t-1) + b_i) \quad (3d)$$

$$o(t) = \sigma(W_o x(t) + U_o h(t-1) + V_o c(t) + b_o) \quad (3e)$$

$$f(t) = \sigma(W_f x(t) + U_f h(t-1) + V_f c(t-1) + b_f) \quad (3f)$$

$$a(t) = f(W_a x(t) + U_a h(t-1) + b_a) \quad (3g)$$

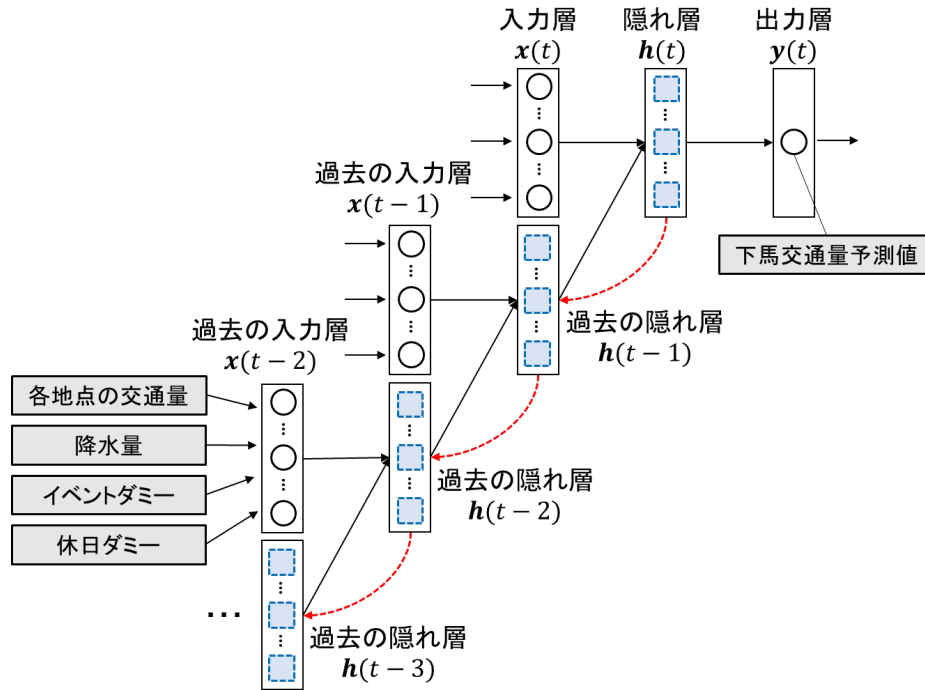


図-2 LSTMモデルの概略

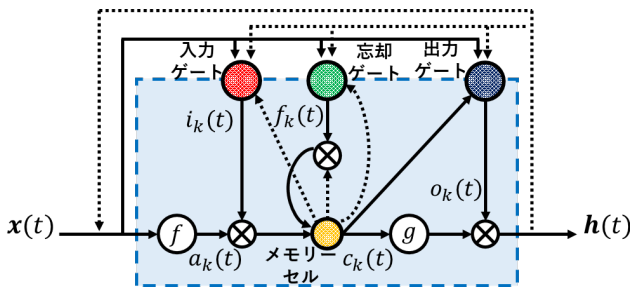


図-3 LSTMブロック

$$h(t) = o(t) \odot g(c(t)) \quad (3h)$$

$$y(t) = g(Vh(t) + c) \quad (3i)$$

(4) 本研究で扱うモデルの設定

a) 入力データ

先に述べたデータを入力項目として下馬の交通量を入力するモデルを構築する。図-2から分かるように、入力層には出力対象である下馬を含む各地点の交通量、降水量が15分刻みで挿入され、イベントの有無は該当する場合には1、そうでない場合は0が入る仕様とした。また休日ダミーも同様に、土曜日、日曜日、祝日の際には1、平日には0が記入されており、本モデルは入力層に複数のデータが挿入される多変量構造を有している。

b) 保持する過去の情報量

交通量は日単位あるいは週単位で規則的に変化することが予想されるため、保持する情報量をそれらの周期に合わせることで交通量の変動を効率的に学習できると考えられる。そのため本分析では保持する過去の情報量を1日とし、日単位の周期性を捕捉する学習設定を施した。

c) ハイパーパラメータ

ハイパーパラメータとはモデル自体で最適化することのできない学習設定であり、隠れ層の数やブロック数などが該当する。本分析では隠れ層を1層から6層の6種類、隠れ層内に設置するブロックを4個、8個、16個、32個、64個、128個の6種類を用意し、予測精度が最良となる組み合わせを調査する。

d) 損失関数

予測モデルの本来の目的は真値に近い値を出力することであるが、今回の分析のように観光地の混雑緩和を想定した予測を実施する場合には、観測値より小さな予測値の出力を避けることが望ましいという考え方が存在する。これは予測した交通量よりも多くの交通量が通過する状態が継続すると、混雑の発生を検知するタイミングが遅れてしまう恐れがあるためである。この問題に対処すべく本分析では損失関数に着目する。回帰問題において一般的に扱われる損失関数は式(3j)で表される平均二乗誤差 (Mean Squared Error / MSE) である。nはサンプル数、

$y_{obs,i}$ は観測値、 $y_{pred,i}$ は予測値を表している。これは線形回帰モデルやニューラルネットワークなど様々なモデルで用いられている。これに対して式(3k)で表される平均二乗対数誤差 (Mean Square Logarithmic Error / MSLE) は予測値が実測値を超過しやすいという特徴を持っている。ゆえに予測値の過小問題を回避することが期待されるため、本分析では学習用の損失関数としてMSLEを使用することとした。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{obs,i} - y_{pred,i})^2 \quad (3j)$$

$$MSLE = \frac{1}{n} \sum_{i=1}^n \{ \log(1 + y_{obs,i}) - \log(1 + y_{pred,i}) \}^2 \quad (3k)$$

e) 訓練データとテストデータ

深層学習の目的は、データを学習することでその特徴を把握し識別や予測を実施することであるが、このとき既知のデータのみでは未知のデータ（入力）に対して識別や予測を行うことが難しい。したがって、データから適切な予測がなされているかを判断するためには与えられたデータ以外の未知のデータに対する予測性能を確認する必要がある。しかし、真に未知のデータを入手することは困難であるため、実際には与えられた全データを訓練データとテストデータに分割することで疑似的に未知のデータを作成することが一般的な手法とされている。

ゆえに観光交通予測モデルとしての有用性を正確に評価するためにはテストデータの期間を各季節や特に混雑が予想されるあじさいの開花時期に設定し、予測性能を確認することが必要だと考えられる。

4. 予測性能の確認と考察

(1) 分析の手順

本研究で構築した学習モデルは、鎌倉市中心部に位置する下馬の下り交通量、上り交通量、両者を合計した交通量の60分先の予測値を出力する。

まず前章で述べたハイパーパラメータに関して、モデルの予測性能が最適となる設定を明らかにすべく、基礎的な分析を実施した。そこで決定されたハイパーパラメ

ータを以降の分析でも使用することとした。

次に、観光交通を短期予測するうえで重要だと考えられる項目について評価を行う。観光交通需要が有する季節性の捕捉度を評価するために春（4月～6月）、夏（7月～9月）、秋（10月～12月）、冬（1月～3月）それぞれの季節をテストデータとした推定を実施した。

また、降水量や鎌倉市内で開催されるイベントが通過交通量に与える影響をモデルが捕捉しているか評価するため、降水量と社会イベントを入力データから除去したモデルにて推定を実施し、出力された予測交通量を確認する。

最後に、実際に交通混雑が発生していると考えられるあじさいの開花時期である6月中旬から下旬にかけての期間において、モデルの予測性能を確認する。

(2) 基礎推定結果

基礎推定は2017年1月1日から2018年9月30日までを訓練データ、2018年10月1日から12月31日までをテストデータとして、予測性能が最良となる隠れ層数とLSTMブロック数を調査した。推定結果を表-2に示す。「下馬下り」、「下馬上り」では隠れ層数を3層とした際、「下馬上下合計」では隠れ層数を4層とした際にMAPEが最良値を示した。また隠れ層内に設置されるLSTMブロックは全ての推定において64個が最適値であった。ゆえに以降実施する推定ではこの学習設定を用いることとする。

(3) 季節性に対する予測性能

各季節におけるモデルの予測性能を把握するために、2018年の1月から3月、4月から6月、7月から9月、10月から12月それぞれの期間をテストデータとした推定を実施した。この際、データ数を同条件として各季節に対する予測性能を比較するため、訓練データの量を統一した。2018年の1月から3月をテストデータとした推定の訓練データは2017年の1月から12月であり、2018年の4月から6月をテストデータとした推定の訓練データは2017年4月から2018年の3月となっている。ゆえにどの推定も12か月の訓練データと3か月のテストデータというデータ分割で統一されている。

推定の結果を表-3に示す。下馬下り交通量に対する予測性能は四季を通じて安定していることが確認された。

表-2 各モデルにおける基礎推定結果

	隠れ層数	LSTM ブロック数	平均交通量 [台/15分]	RMSE [台/15分]	MAPE
下馬下り	3層	64個	95	11.38	9.84%
下馬上り	3層	64個	99	15.48	13.07%
下馬上下合計	4層	64個	194	19.82	8.09%

表-3 季節性に対する予測性能

	テストデータ期間	平均交通量 [台/15分]	RMSE [台/15分]	MAPE
下馬下り	冬 (1月~3月)	94	12.16	12.08%
	春 (4月~6月)	96	11.62	9.95%
	夏 (7月~9月)	98	12.33	10.37%
	秋 (10月~11月)	95	11.71	10.31%
下馬上り	冬 (1月~3月)	97	17.41	20.12%
	春 (4月~6月)	101	15.00	12.24%
	夏 (7月~9月)	104	14.64	10.92%
	秋 (10月~11月)	99	15.25	13.83%
下馬上下合計	冬 (1月~3月)	191	26.04	14.10%
	春 (4月~6月)	197	19.52	8.08%
	夏 (7月~9月)	202	19.34	7.56%
	秋 (10月~11月)	194	19.51	7.88%

ゆえに本モデルは、観光交通需要の季節変動を十分に捕捉していると考えられる。一方で、下馬上り交通量と下馬上下を合計した交通量に対する予測性能は、冬期(1月~3月)における性能が他の季節と比較して悪いことが読み取れる。しかしこの結果はモデルによる不備ではなく、使用データに起因する問題だと考察する。前述したように、本分析に使用している交通量データはJARTICより取得した断面交通量データであるが、2018年1月の下馬上り交通量データがデータセットから欠落しており入手することができなかった。そのため本分析では、2017年1月の下馬上り交通量データを2018年1月の交通量として代用したが、これにより2年分の交通量データ内に本来存在する変動の特徴が乱れてしまい、冬期の予測性能が低下したと考えられる。2018年2月、3月におけるMAPEを算出したところ、下馬上りは13.71%、下馬合計は9.40%であり、他の季節に対する予測性能と遜色ない結果が確認された。ゆえに代用された2017年1月の交通量データがモデル学習に悪影響を及ぼしたと考察する。

(4) 降雨・社会イベントが観光交通に及ぼす影響

これまでの分析では、降雨や社会イベントの開催が観光交通量の変動に影響を与えるという仮定の下、これら

表-4 下馬合計交通量予測値 (降水量対象実験)

	交通量予測値
降水量入力モデル	9,275 [台/日]
降水量除去モデル	9,318 [台/日]

表-5 下馬合計交通量予測値 (社会イベント対象実験)

	交通量予測値
イベントダミー入力モデル	8,582 [台/日]
イベントダミー除去モデル	8,246 [台/日]

を入力変数として設定した。そのため、実際に降水量データや社会イベントダミーが予測性能に与える影響を把握すべく、降水量データ、社会イベントダミーを入力したモデルと、降水量データを入力から除外したモデル、社会イベントダミーを入力から除外したモデルそれぞれの交通量予測値の比較を行った。

まず降雨が連続した2018年9月2日から4日までの下馬上下合計の交通量予測値を比較した結果を表-4に示す。降水量を考慮することにより、1日あたりの交通量予測値が約40台減少していることが読み取れる。一般に、雨天時は晴天時よりも道路容量が低下し、また観光を目的とした外出機会が減少すると考えられるため、通過交通量は減少すると考えられる。降水量の対象実験の結果で確認された減少割合は微々たるものであったが、ここで把握すべきは降水量が観光交通に与える影響であり、出力された交通量予測値は想定される変化を示した。

次に、かまくらまつりパレードが開催された4月8日から15日までの8日間の交通量予測値を比較した結果を表-5に示す。降水量とは対照的に、社会イベントの開催を考慮することで、1日あたりの交通量予測値が約340台増加することが確認された。社会イベントが開催される日は平時よりも多くの観光客が来訪すると予想されるため、通過交通量は増加すると考えられる。

すなわち、いずれの場合においても、構築した学習モデルは直感に合致した交通変化を示しており、本モデルが両者の影響を適切に捉えている可能性が示唆される。

(5) 混雑時期の交通量に対する予測性能

a) あじさい開花時期における1週間

これまでの分析から本モデルが観光交通が持つ季節性を捕捉し、降雨や社会イベントの影響を予測に反映していることが確認された。しかし観光交通予測モデルとし

表-6 あじさい開花時期の交通量に対する予測性能

	平均交通量[台/15分]	RMSE[台/15分]	MAPE
下馬下り	97	10.46	9.26%
下馬上り	99	13.37	10.57%
下馬上下合計	199	18.50	7.31%

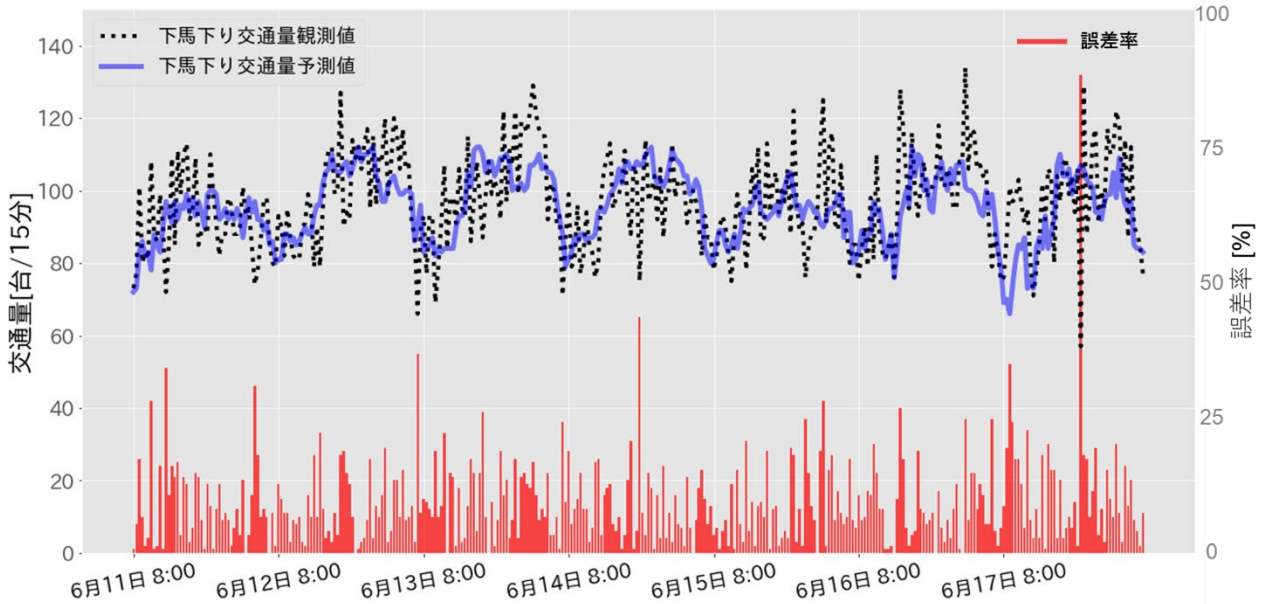


図-4 下馬下り車線交通量の予測値と誤差率（2018年6月11日～17日）

での適用可能性を評価するためには、混雑発生時におけるモデルの予測性能を確認する必要がある。そこで、あじさいの開花時期である6月中旬の1週間（6月11日～17日）をテストデータとして推定を実施した。結果を表-6に示す。これによると、基礎推定の結果よりも高い予測性能を示していることが分かる。Zhao et al. (2017)³⁾では、本分析と同様15分単位のデータから60分先の交通量予測値を出力しており、その際のMAPEは16~17%であった。予測対象とする地点やモデルの学習設定が同一でないため、単純に比較することは難しいが、本研究で構築したモデルが示すMAPEは7~10%であるため、既存モデルに劣らない性能を確保できたと考える。

また、推定された予測値に対して、時間帯別の適合度を確認するため、予測値のプロットと時間帯毎の観測値に対する予測値の誤差率を図-4に示す。左側の軸が下馬下り車線の断面交通量[台/15分]を右側の軸が誤差率[%]を表す。誤差率は式(4a)のように定義される。プロットの期間は2018年6月11日から6月17日までのあじさい開花時期の1週間である。黒破線は下馬下り車線の交通量の観測値、青実線はその予測値である。赤線の棒グラフで示される誤差率を見ると、8時や20時付近よりも日中の方が予測値の当てはまりが良いことが分かる。車両の流入量が多い時間帯に対して、精度の高い予測値が算出

されたことは、混雑緩和を目的とした交通政策の立案に対する本モデルの適用可能性を示唆する結果であると考えられる。

$$E^* = \frac{|y_{obs,i} - y_{pred,i}|}{y_{obs,i}} \times 100(\%) \quad (4a)$$

b) カメラ映像から確認された混雑発生日

前項では、あじさい開花時期の鎌倉市では観光混雑が発生しているという一般論に基づき推定を実施した。しかし実際の混雑状況を把握することは難しく、観光交通



図-5 鶴岡八幡宮東側の交通状況（2018年6月30日13時54分）

表-7 混雑発生時間帯における鶴岡八幡宮下り交通量に対する予測性能 (2018年6月30日)

	平均交通量[台/15分]	RMSE [台/15分]	MAPE
全時間帯	83	17.13	16.74%
混雑発生時間帯	79	12.11	15.33%

予測モデルを検証する時期として適当であるか疑問の余地がある。そこで国土交通省により市内に設置された路上カメラの映像を確認し、混雑が発生している時間帯を把握した。図-5は鶴岡八幡宮東側に設置された路上カメラが捕らえた2018年6月30日13時54分頃の下り車線の交通状況である。画像奥に進行する車両が鶴岡八幡宮への流入交通である。撮影された時刻から1分近く車両が停止しており、動き出した後も車両の流入が継続していることが確認できた。待機列の直接的な原因は、鶴岡八幡宮交差点に設置された信号機だと思われるが、約30分に渡り車両の進行、停止が繰り返されていることが確認できた。そのため、この地点を下り交通の混雑位置であるとみなすこととする。カメラ映像から混雑が発生していると判定した時間帯は、8時30分～9時（第1混雑時間）、10時30分～11時（第2混雑時間）、13時30分～14時（第3混雑時間）、15時～15時30分（第4混雑時間）の4つの時間帯であった。

この時間帯におけるモデル予測性能を確認するために、2018年6月30日をテストデータとした推定を実施した。結果を表-7に示す。同じく下り車線の交通量である下馬下りの推定結果と比較するとMAPEが増加していることが分かる。鶴岡八幡宮下り交通量の平均値は下馬よりも小さく、混雑度合いが高い地点だと考えられる。そのため、交通量の変動が不規則でありモデルの予測性能が低下したと考察する。しかし混雑が発生している時間帯に限ったMAPEが全時間帯におけるMAPEよりも小さいこ

とが確認された。

また出力された予測値と観測値をプロットした結果を図-6に示す。黒破線は鶴岡八幡宮下り交通量の観測値を、緑実線はその予測値を、赤線の棒グラフは誤差率を、青枠が掛かる領域は混雑が発生している時間帯をそれぞれ表している。ここで、混雑が発生している時間帯における予測値の変動を確認する。誤差率の低い予測値を出力することは、観光交通予測モデルに要求される性能であるが、予測値の変動が観測値の増減に合致することも重要視されると考えられる。例えば、混雑が継続している場合は単位時間あたりの通過交通量は減少するが、出力される予測値が増加しているのは、仮に誤差率が小さい場合でも、交通状態を正確に予測しているとは言い難いためである。

まず第1混雑時間では、観測値が減少しているのに対して、予測値は僅かに増加した後に減少している。そのため徐々に誤差率が低下していることが分かる。第2混雑時間では観測値が減少している一方で、予測値はほぼ同じ値のままとなっている。そのため誤差率は時間が進むにつれて増加している。第3混雑時間では、観測値は減少した後、増加するV字型の変化を示しているが、予測値も同様の変化をしており、誤差率も低くなっている。第4混雑時間でも観測値はV字型の変化を示しており、予測値も類似した変動をしている。また、混雑開始時点の誤差率は高いが、徐々に低下していることが読み取れる。

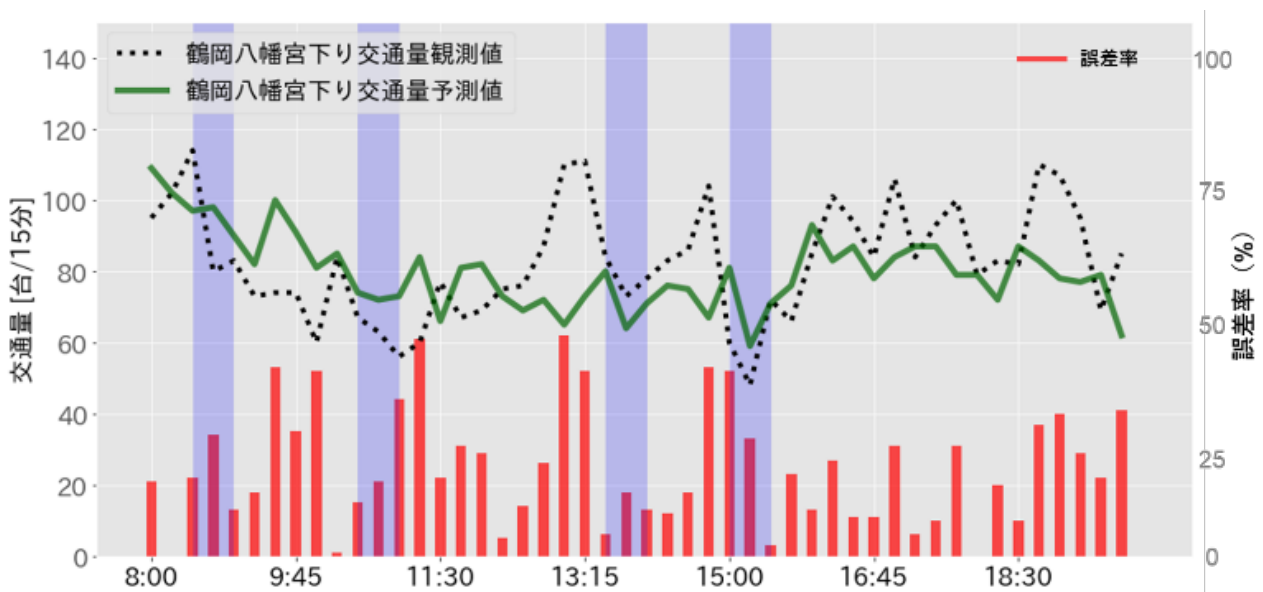


図-6 混雑発生時間帯における予測値の変動

以上から、混雑が発生している時間帯において、出力された予測値は概ね観測値と同様の変動を示していることが確認された。しかし誤差率は下馬下りよりも高くなっているため、今後さらなる改善が必要だと考えられる。

5. おわりに

(1) まとめ

本研究では鎌倉市中心部の混雑を緩和するための交通政策実施を想定して、深層学習ベースの短期交通量予測の枠組みを構築した。時系列データ予測の代表的な深層学習モデルであるLSTMを用いたところ、本モデルは観光交通が持つ季節性を捕捉し、降雨や社会イベントの影響を予測に反映していることが確認された。また、あじさい開花時期における観測値と60分先予測値のMAPEが7.31%という結果が得られた。

また、予測値の適合度を時間帯ごとに確認すると、交通量の増減が大きい日中の誤差率が低いことが明らかとなった。さらに、混雑が発生している時間帯に対する予測性能を確認したところ、出力された予測値は概ね観測値と同様の変動を示し、日平均よりも低いMAPEを示した。これは混雑下における交通量の変化を、本研究で提案したモデルによって十分に捕捉できたためだと考えられる。

(2) 今後の課題

今回の分析では交通量データのみを使用しているため、混雑の発生をデータ内から判断することができず、詳細な交通状態を把握するまでには至っていない。ゆえに混雑発生を判断するために、同一のモデルを用いて、リンク旅行速度予測値の出力を実施した。旅行速度のデータはETC2.0プローブデータから今回の分析対象地と同じ地点のデータ取得した。しかしながら、取得したデータには欠損値や異常値が半数近く存在し、それらを補完した後に算出されたMAPEが30%以上であることから、十分な精度を得ることはできなかつたと判断される。そのためリンク旅行速度に対するモデルの予測性能を向上さ

せることが大きな課題であり、課題解消に向けて、より実用性を備えた観光交通予測モデルの構築を目指していきたい。

謝辞：本研究は、国土交通省・道路政策の質の向上に資する技術研究開発「学習型モニタリング・交通流動予測に基づく観光渋滞マネジメントについての研究開発」（代表：布施孝志）の一環として行われた。ここに記して感謝の意を表します。

参考文献

- 1) Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural Computation*, 9(8), pp1735-1780, 1997.
- 2) Fu, R., Zhang, Z. and Li, L.: Using LSTM and GRU neural network methods for traffic flow prediction, *The 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, Wuhan, pp. 324-328, 2016.
- 3) Zhao, Z., Chen, W., Wu, X., Chen, P.C.Y. and Liu, J.: LSTM network: A deep learning approach for short-term traffic forecast, *IET Intelligent Transport Systems*, Vol. 11, No. 2, pp. 68-75, 2017.
- 4) Ke, J., Zheng, J., Yang, H. and Chen, X.(M): Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach, *Transportation Research Part C*, Vol. 85 pp591-608,2017.
- 5) Xu, C., Ji, J. and Liu, P.: The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets, *Transportation Research Part C*, Vol. 95, pp. 47-60,2018.
- 6) Ogawa, K. and Fukuda, D.: Short-term traffic state prediction using the framework of gated recurrent unit: A case Study in Kamakura, Japan. *The 24th International Conference of the Hong Kong Society for Transportation Studies (HKSTS)*. Hong Kong, 2019.
- 7) 公益財団法人日本道路交通情報センター. “JARTIC インフォメーション”. 2019-12-19. http://www.jartic.or.jp/jartic_web/others/faqs.html, (参照 2020-01-17).
- 8) 巢籠悠輔：詳解ディープラーニング TensorFlow・Keras による時系列データ処理, マイナビ, 2017.

(2020年3月8日受付)

SHORT-TERM TRFFIC VOLUME PREDICTION USING MULTIVARIATE LSTM: A CASE STUDY ON THE TOURIST TRAFFIC IN KAMAKURA

Kohei OGAWA and Daisuke FUKUDA

Real-time spatio-temporal measurements of traffic information are available from in-ground detectors or GPS probes. One of the typical applications of artificial intelligence which learns and predicts using the accumulated big data and the accumulation of research has been advancing in the transportation field as

well. However, only few attempts have so far been made at the prediction of tourist traffic congestion, and the overtourism problem has not been resolved. Tourist traffic demand has a strong periodicity (for example, time of day, day of the week, season), and also some exogenous variables, such as weather conditions and social events, may have strong influences. Therefore, in order to predict tourist traffic with high accuracy, it is necessary to comprehensively analyze various information. In this study, we construct a multivariate and time-series deep learning model with real-time information in mind. This model outputs 60-minute forecasts of traffic volume in the city center by learning multiple inputs such as traffic volume data, precipitation and social event information in the city. Validation of the model output error rate for each time zone and season, and the prediction performance of the constructed model were confirmed. The results showed that the traffic volume prediction under congested conditions could be executed with an error rate of 15.33%.