

Open access

4,026 Views | 19 CrossRef citations to date | 2 Altmetric

Listen

Articles

# Predicting the spatiotemporal legality of on-street parking using open data and machine learning

Song Gao , Mingxiao Li, Yunlei Liang, Joseph Marks, Yuhao Kang & Moying Li

Pages 299-312 | Received 17 Sep 2019, Accepted 04 Oct 2019, Published online: 18 Oct 2019

Cite this article <https://doi.org/10.1080/19475683.2019.1679882>

Check for updates

Full Article

Figures & data

References

Citations

Metrics

Licensing

Reprints & Permissions

View PDF

View EPUB

STRACT

Formulae display:  MathJax

## We Care About Your Privacy

We and our 848 partners store and/or access information on a device, such as unique IDs in cookies to process personal data. You may accept or manage your choices by clicking below, including your right to object where legitimate interest is used, or at any time in the privacy policy page. These choices will be signaled to our partners and will not affect browsing data. [Privacy Policy](#)

We and our partners process data to provide:

Use precise geolocation data. Actively scan device characteristics for identification. Store and/or access information on a device. Personalised advertising and content, advertising and content measurement, audience research and services development.

List of Partners (vendors)

I Accept

Essential Only

Show Purpose



retail stores, health-care services, accommodation and food services are positively associated with the number of parking violation tickets.

1. Introduction

KEYWORDS:

Open data

data fusion

machine learning

urban computing

2. Related work

3. Data

4. Methods

# 1. Introduction

5. Results

6. Prototype

Parking is an important element in the transportation system and plays an important

7. Conclusion

role in people's travel decisions. Parking availability information and pricing can

influence people's departure and arrival time, travel mode choices, and activity

Acknowledgements

duration. Almost all U.S. cities have minimum parking requirements for each type of

Disclosure statement

land use, which determines the minimum number of parking spaces that should be

provided by land developers. Most on-street parking is free or underpriced compared to garages and parking lots and therefore it is often over-demanded. This makes searching

Additional information

for a parking spot in metropolitan areas a great challenge comparable to the Hunger

Footnotes

Games, especially in highly populated areas such as downtown districts, job centres,

References

etc. First, the supply and demand of parking spaces is unbalanced with the increasing

number of vehicles but limited parking facilities in urban areas. Second, the

urbanization process is accelerating in most metropolitan areas and attracting more job

opportunities, human flows, business and social activities. These popular destinations

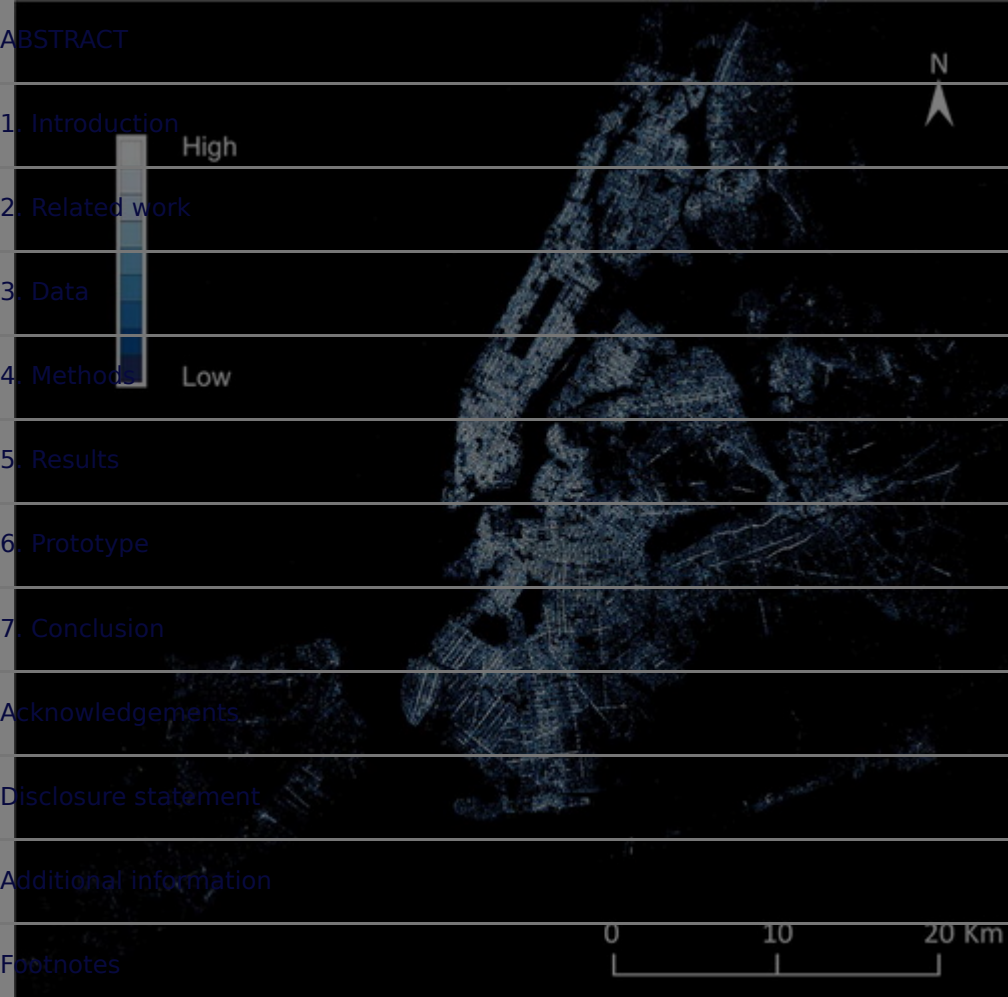
together with underpriced parking generates more travel demand and parking needs.

Third, parking availability and legality are highly variable spatially and temporally.

Figure 1. X ation tickets

in the Ne





Display full size

On-street parking is often a cost-effective choice compared to parking facilities such as garages and parking lots. However, limited space and complex parking regulation rules make the search process very difficult. Moreover, there is almost no real-time information about available on-street parking spots and it is legal or illegal to park. Even if the driver is a local resident, compound parking rules can still surprise the driver

and gen  
proximit  
periods.  
cities in  
generate  
show  
two  
overlayi  
The Man  
parking  
spatiote

x

les,  
tain time  
ket cost  
kets  
3. Figure 1  
ts over the  
ets without  
ets in NYC.  
n-street  
nderstand the  
l reliable

machine learning models to predict the spatiotemporal legality of on-street parking using NYC open data<sup>2</sup>.

1. Introduction

Another important aspect that needs attention when analysing the point-based parking violation data is the scale effect and the modified area unit problem (MAUP). The scale

effect may cause variations in statistical results among different spatial aggregation

levels and MAUP affects results when point-based measures are spatially aggregated

into different zoning configurations (Openshaw [1984](#); Fotheringham and Wong [1991](#)).

How much detail a machine learning model needs while still producing an acceptable error rate is worth of exploration. To this end, we propose a data-driven framework for

understanding and predicting the spatiotemporal legality of on-street parking by

training machine learning models using the NYC parking tickets open data. And four

types of spatial analysis units (i.e. point, street, census tract, and grid) are used to

examine the impact of spatial aggregation scale in machine learning predictive models.

Additional information

The remainder of the paper is organized as follows. First, in [Section 2](#), we present the

literature review on parking availability prediction studies in transportation and

computer sciences. Then, in [Section 3](#), we introduce the datasets and preprocessing

steps in order to feed the data into machine learning models. In [Section 4](#), we

formulate the problem into regression and classification approaches, and briefly

introduce a set of machine learning models used in this study. In [Section 5](#), we test our

framework at different spatial scales and compare the model performance. In [Section 6](#),

we show the designed Web prototype for exploring the parking legality information in

NYC. Finally, we conclude this work with some considerations on the potential of this

work for [Section 7](#).

2. Related work

Parking  
park



choice (

the U.S.

creates

travel m

cruising

In this article



balance the parking demand and supply (Litman [2018](#)). With better parking information provided, drivers can enhance their parking decision-making and prevents parking overflow from one place to another (Caicedo, Blazquez, and Miranda [2012](#)).

## 1. Introduction

## 2. Related work

There have been many research efforts towards improving the search efficiency for an available parking space given its impact on driving time, traffic, and even air pollution

## 3. Data

(Todorović and Panta [2006](#)). One study found that 30% of the average traffic cruising

in investigated areas was actually caused by searching for parking, with an average

## 5. Results

search time of 8.1 min (Shoup [2006](#)). Google AI research team developed a logistic

regression model to predict parking difficulty (e.g. limited parking or easy) by utilizing

anonymous aggregated trajectory information from mobile users who opt to share their

location data (Cook, Li, and Kumar [2017](#)). They used grids as a spatial unit for training

## Acknowledgements

the model. Using this, Google launched a new feature for the Google Maps App across

25 US cities that offers predictions about parking difficulty close to users' destination.

## Disclosure statement

There also have been studies focused on detecting the availability of parking spaces

either using instrument parking infrastructure with special sensors (Chatman and

## Footnotes

Manville [2014](#)) as well as using crowd-sensing solutions (Chen, Santos-Neto, and

Ripeanu [2012](#); Zheng, Rajasegarar, and Leckie [2015](#); Pflügler et al. [2016](#); Bock,

Attanasio, and Di Martino [2017](#)), but both types of studies rely on the existence of

predefined parking spaces or the development of mobile applications such as

PhonePark, iPark, and UPDetector (Xu et al. [2013](#); Yang, Fantini, and Jensen [2013](#); Ma,

Wolfson, and Xu. [2014](#)). This study focuses on determining the legality of a parking

space at a given time and day and the location at different spatial scales by utilizing

machine learning algorithms and using publicly available parking violation ticket

information.

Some of the machine learning algorithms used in this study are based on training.

regression models using a vector-based representation of the input data (e.g. [2017](#)).

With these models, we can predict the availability of parking spaces while driving.

These models are trained on historical data and can be used to predict parking availability.

Some of the machine learning algorithms used in this study are based on training.

regression models using a vector-based representation of the input data (e.g. [2017](#)).

With these models, we can predict the availability of parking spaces while driving.

These models are trained on historical data and can be used to predict parking availability.

Some of the machine learning algorithms used in this study are based on training.

regression models using a vector-based representation of the input data (e.g. [2017](#)).

With these models, we can predict the availability of parking spaces while driving.

These models are trained on historical data and can be used to predict parking availability.



modelling, and groundwater studies (Naghibi, Pourghasemi, and Dixon [2016](#)). In this study, we examine the importance of spatial resolution when incorporating spatial data with machine learning. At the time of writing, machine learning models examining the spatial heterogeneity at different spatial resolutions are still an emerging study area (Lu et al. [2018](#); Yang et al. [2019](#)). Such research incorporates Tobler's First Law of Geography (Tobler [1970](#)) and the scale effect, which is to say that more detailed spatial resolution should provide more related features and likely produce less error than less detailed spatial resolutions. The idea that a machine learning model is able to provide these predictions at all is Tobler's First Law of Geography in practice, and potentially transferable according to the Third Law of Geography which focuses on the similarity of geographic configuration of locations in spatial predictions (Zhu et al. [2018](#)). By examining different spatial resolutions or at aggregating into different spatial analysis units, we provide insights into how much detail a machine learning model needs while still producing an acceptable error rate.

### 3. Data

#### 3.1. Parking violation tickets open data

As mentioned above, we downloaded over 10.8 million parking violation tickets generated in NYC in the fiscal year 2017 and 11.7 million in fiscal year 2018 from the NYC Open Data platform. Each ticket contains information including a summons number, violation code, street address, ticketing time, vehicle plate, etc.

#### 3.2. Po

In order to be consistent with more parking tail stores, health-c er 137,000 point e<sup>3</sup>. The POs are fir em (NAICS) two-digi based on the NAICS tw ng Oil and Gas Extr Wholesale Trade 42



Scientific Tech 54, Administrative Support and Waste 56, Educational Services 61, Health Care and Social Assistance 62, Arts & Entertainment & Recreation 71, Accommodation & Food Services 72, Other Services 81, Public Administration 92. In addition, the category of Parking Lots and Garages (NAICS Code: 812,930) is treated as a separate POI category since it is directly related to parking activity. This gives a total of 24 categories of POIs at the root level of categorization as part of model features (See Figure 5 in detail).

### 3.3. Human mobility patterns

In addition to the static spatial distribution of POIs information, we also retrieved the fine-resolution visit patterns of all POIs from the aforementioned SafeGraph database which covers dynamic human mobility patterns of millions of anonymous smartphone users. For each POI, the records of aggregated visitor patterns illustrate the number of unique visitors and the number of total visits to each venue during the specified time window, which could reflect the attractiveness of each venue. The mean hourly visits over a week were recorded as a 168-dimensional vector to show the dynamic stream of visit patterns. If a visitor stays for multiple hours, a visit will be shown in each hour during which the visitor stayed.

Figure 2. The spatial distribution of parking violation tickets at four spatial scales and the temporal variation curves of tickets.



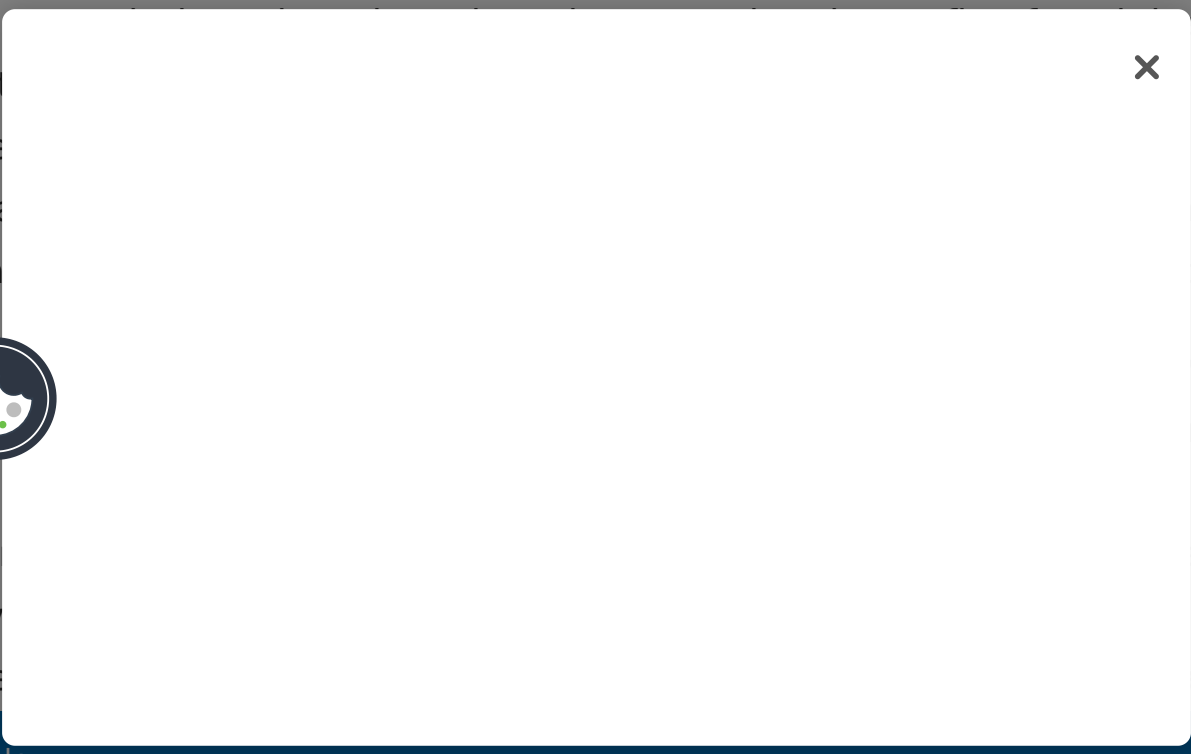
### 3.4. Data preprocessing

In this study, we discretized the study area into four spatial scales (point level, street level, census tract level, and 1 km grid level) and the time into 168 hourly slots (7 days of a week \* 24 h of a day) to capture the snapshots of the legality of street parking. To

perform  
of each t  
since we  
more tha  
tickets a

At th  
of day  
used as  
spatiote  
street ty  
characte

coordinates  
API. Then,  
that were  
e number of  
h time slot.  
(j), the time  
n(j,h,d) was  
length stlen,  
s follows: [





features were more complex. At these levels, we not only considered their spatiotemporal characteristics  $p_j$ ,  $th$ , and  $td$ , their street attributes with the summation of street length  $sum\_st\_len$  and street area  $sum\_st\_area$ , but also considered the dynamic human mobility patterns and the POI distributions in the corresponding spatial unit. The POI data were aggregated to each spatial unit and represented as 24 features. In addition, the number of visits observed in the specified unit  $sum\_visit_j$ , the number of unique visitors  $sum\_visitor_j$ , the number of visits of corresponding to the time of day  $visit(j,h)$  and the number of visits of corresponding to the day of week  $visit(j,d)$  were added to model the human mobility patterns. Thus, the training features can be represented as follows:  $[p_j, th, td, sum\_st\_len, sum\_st\_area, POI1, \dots, POI24, sum\_visit_j, sum\_visitor_j, visit(j,h), visit(j,d)]$ . [Table 1](#) shows the number of samples at each spatial scale and the ratio of positive and negative cases for parking legality classification (See [Section 5.3](#) for more details).

Additional information

**Table 1. The number of samples at different spatial scales and the ratio of positive and negative cases for parking legality classification.**



Download CSV

Display Table

Figure 2 shows the spatial distributions of georeferenced parking violation tickets at four scales and their temporal variation curve of the summation of tickets. The figure of point level data was derived using kernel density estimation, and the other three figures were derived using the Haralick texture features. In order to show the spatial distribution of tickets in each scale, we used the Haralick texture features: safe, low probability of getting a ticket, 'dangerous' interval of class set at 0.001, 0.001, 0.001, 0.001.



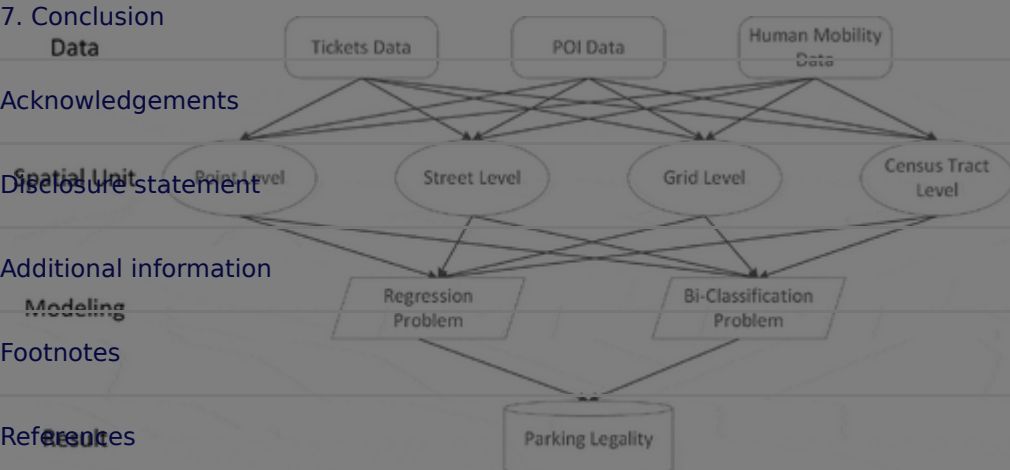
4. Met

In this article



In this section, we provide the details of our proposed framework to understand and predict the parking legality. In this framework, we approach the parking legality prediction as a regression problem to predict the number of parking violation tickets given a location and time, and as a binary classification problem to interpret parking legality. To this end, a series of machine learning models are trained in four spatial scales with processed datasets that are mentioned in the previous section. The architecture of our proposed framework is shown in Figure 3.

Figure 3. The proposed parking legality predictive framework using multi-source data and machine learning.



Display full size

### 4.1. Prediction of parking violation counts

To investigate the impact of the spatial scale on different regression models, we selected the following six machine learning models for this study:

MLR: Multiple Linear Regression

SVM: Support Vector Machine

Decision Tree



**Random Forest:** The random forest (RF) constructs a multitude of decision trees and outputs the results by computing the mean of the predictions of each individual tree (Breiman [2001](#)). RF is trained on different parts of the same training set, with the goal of reducing the variance.

**GBRT:** The gradient boosted regression trees (GBRT) use a gradient boosting method to construct a set of decision trees as base learners and outputs the result by computing the sum of the base learners (Friedman [2001](#)). XGboost uses a more regularized model formalization to control over-fitting issue and thus is chosen in this study (Chen and Guestrin [2016](#)).

**DNN:** Deep neural network (DNN) is a multi-hidden-layer artificial neural network whose artificial neurons can respond to a surrounding unit within a portion of the coverage (Goodfellow, Bengio, and Courville [2016](#)). As shown in Figure 4, we constructed a DNN architecture consisting of four fully connected dense layers with reclinear activation functions and two dropout layers with a 0.5 rate to regularize the DNN and improve the generalization error. The output layer uses a linear activation function for regression and a sigmoid activation function to produce a probability between 0 and 1 for binary classification using a threshold of 0.5. The mean square error (MSE) is used as the loss function for regression training while the cross-entropy is used for classifier training.

Figure 4. The multilayer architecture of deep neural network used in this study.



Evaluating the performance of the models and comparing them with existing models. The results of the training the models are shown in the following table.

MSE = 1.1

Where  $y_i$  is the number of observed parking violation tickets at the location  $i$ ,  $\hat{y}_i$  is the corresponding predicted value, and  $N$  is the number of fed data into each model.

1. Introduction

## 4.2. Prediction of parking legality

2. Related work

In addition to interpreting parking legality as a regression problem by estimating the number of violation tickets, it can also be interpreted as a binary classification problem of whether the corresponding time and place can be legally parked by analysing the historical parking violation ticket information. Therefore, a specified location and time with at least one ticket is marked as 'positive case' (i.e. the machine learning binary classification when the number of tickets  $\geq 1$ ) to represent 'Risky Parking' and the others will be marked as 'negative case' (i.e. the number = 0 and no ticket issued) to represent 'Legal Parking'.

Acknowledgements

In addition to the aforementioned SVM, random forest, and DNN machine learning approaches that can also be used for the classification problem, four additional classification models are used in this study:

Additional information

**KNN:** k-nearest neighbours algorithm (KNN) classifies an object by a vote of its neighbours, with the object being assigned to the most common class among its k nearest neighbours in feature space (Cover and Hart et al. [1967](#)).

Footnotes

References

**Logistic regression:** It uses a logistic function to model the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables (Berkson [1944](#)).

**Naive Bayes:** Naive Bayes is a probabilistic classifier that makes classifications using the Maximum A Posteriori decision-making Decision-making (Marsen [1961](#)).

**SGD:** Stochastic Gradient Descent (SGD) is a vector optimization algorithm that updates the model parameters at a fixed interval of time and space (Zhang [2017](#)).

**Evaluation:** The performance of the proposed model is evaluated using the True Positive Rate (TPR) and True Negative Rate (TNR) as the overall legality classification accuracy.



$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (4)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

## 1. Introduction

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

## 2. Related work

## 3. Data

# 5. Results

## 4. Methods

Figure 5. The result of selected variables vs. R-squared at the census tracts level and at the 1 km grid level.

## 6. Prototype

## 7. Conclusion

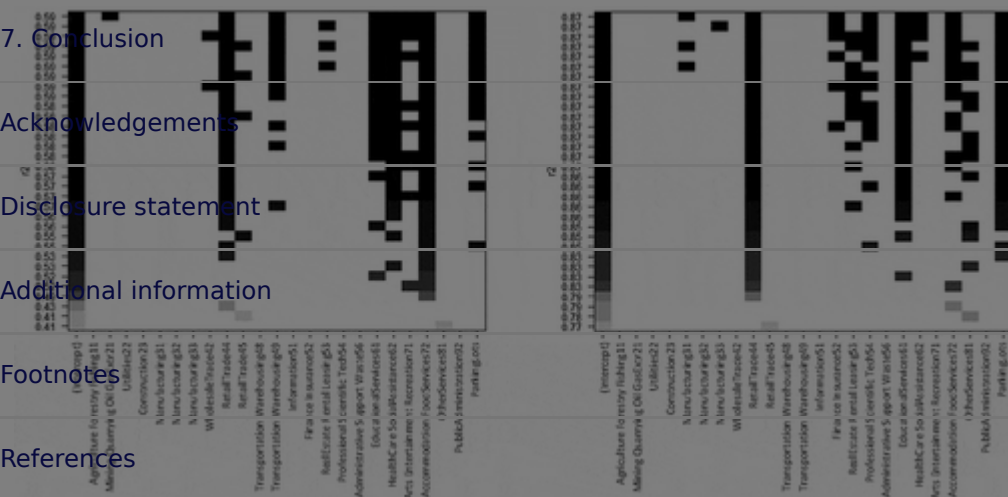
## Acknowledgements

## Disclosure statement

## Additional information

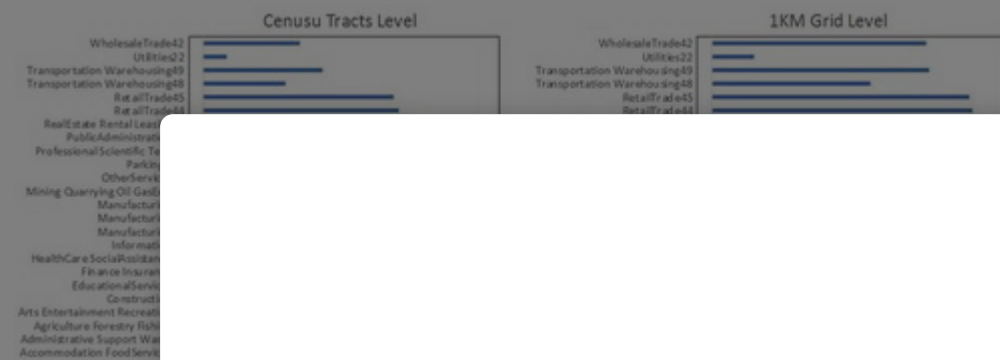
## Footnotes

## References



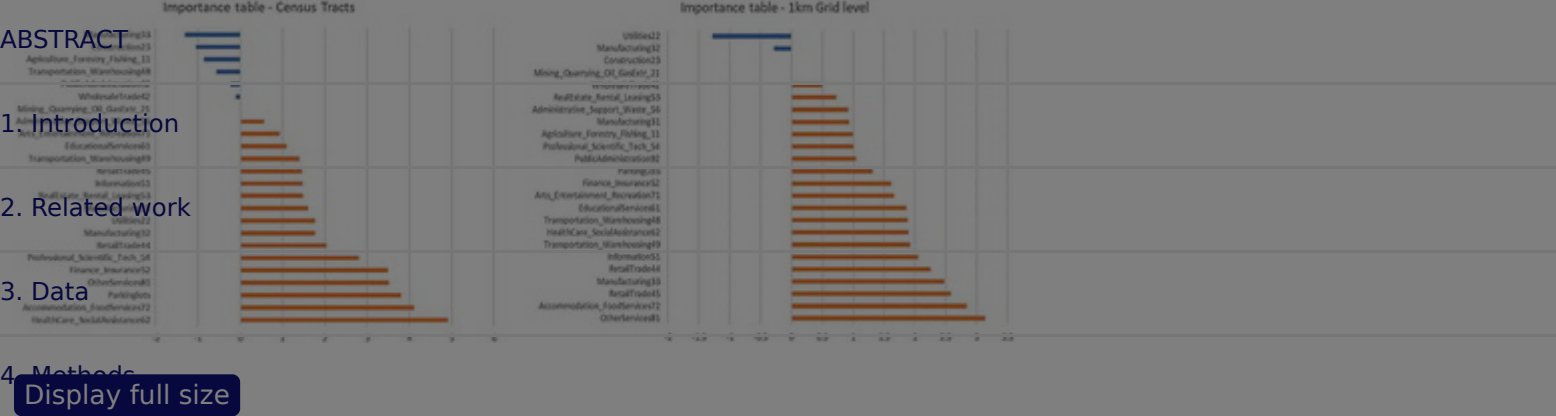
Display full size

Figure 6. The correlation analysis at the census tract level and the 1 km grid level.



Display full

Figure 7. The result of selected variables vs. R-squared at the census tracts level and at the 1 km grid level.



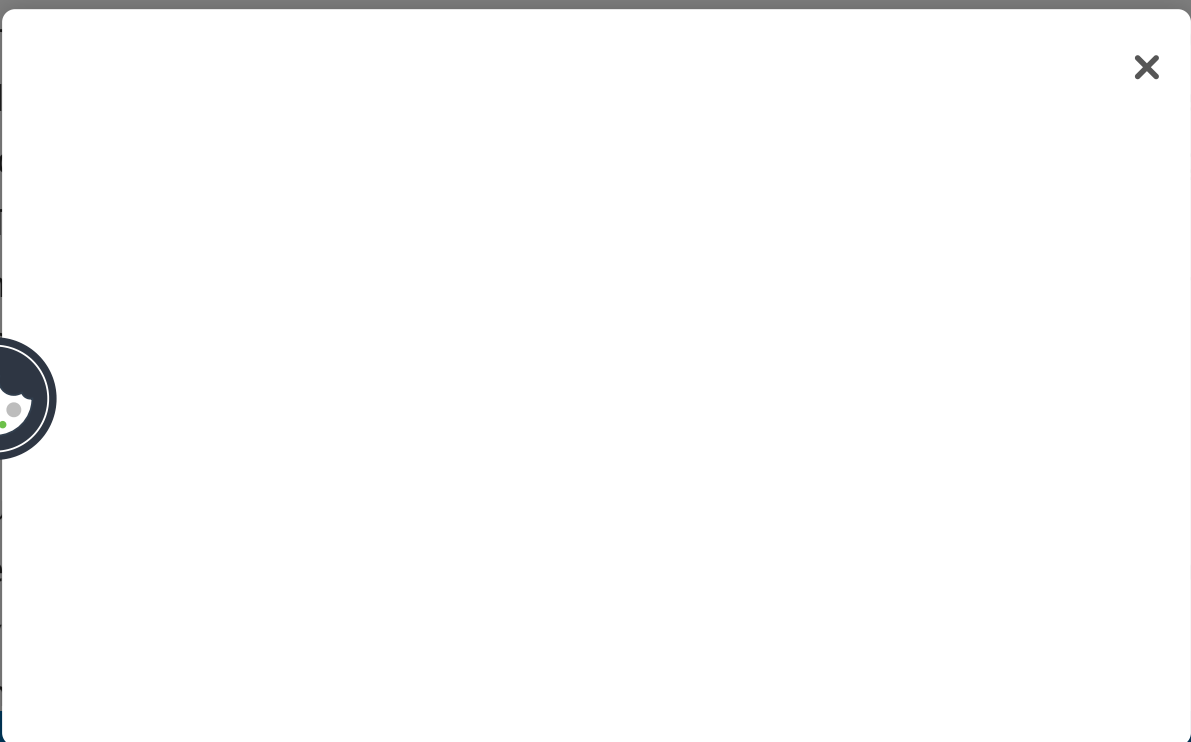
## 5.1. Relationship between parking violation and POIs and visit patterns

For POI-based and mobility-based analyses, the POIs mentioned in [Section 3](#) are aggregated to the census tracts level and the 1 km grid level, respectively, to conduct analysis. After the aggregation, for each analysis unit (a census tract or a grid cell) there is a sum of parking tickets and 24 categories of POIs in each unit.

The multi-linear regression is conducted first to obtain an overall relationship between the parking violation and its surrounding environments. Then, the correlation analysis and the importance ranking using the random forest regression method are also implemented to identify critical factors for the parking violation.

### 5.1.1. Census tracts level

In the census tracts level, when using all 24 categories of POIs to fit the number of parking tickets in the linear regression, an adjust R2 around 0.59 is obtained. To select the most important variables, an exhaustive search for the best subsets of variables for predicting



increase of parking violation, which may not be intuitive. This might indicate that the current parking lots are still not enough for the parking needs of citizens in New York.

1. Introduction

## 5.1.2. 1 km grid level

2. Related work

3. Data At the 1 km grid level, the result of multiple linear regression using all 24 variables is

4. Methods around 0.87, which indicates a good estimation. Figure 5 shows the best subsets of variables for prediction. The frequently selected variables are a little different from the

5. Results result of the census tracts level, which are Retail Trade, Educational Services,

6. Prototype Accommodation and Food Services, and Parking Lots. However, by examining the

7. Conclusion multicollinearity of different combinations of variables, there always exists high

colinearity with the Variance Inflation Factor of over four among variables.

Acknowledgements

in this case, the results from correlation analysis (Figure 6) and the importance ranking

Disclosure statement

from random forest regression (Figure 7) may be more reliable. The results of the

Additional information

correlation analysis at the 1km grid level are very similar to that of the census tracts

level (except for the ParkingLots) where Accommodation and Food Services, Retail

Footnotes

Trade and Other Services are the top three most correlated variables to parking ticket

References

numbers. These three factors are also the top three important factors in the importance

ranking.

In general, most categories of the POIs will lead to an increase in parking violation, especially when there are accommodation, food, and retail stores. Also, the existence of parking lots cannot prevent the parking violation, which shows that the need for parking lots is still not satisfied at the current stage.

## 5.2. Re

As show

different

the decr

km g

mode

training

predicti

that the

km grid



level, which has a small gap. The results also show that the random forest model works the best with the minimum RMSE and outperforms other competitors by a large margin across all the spatial scales. Note that for the support-vector regression (SVR), we tried different kernel types (e.g. linear, second-degree polynomial, Gaussian RBF) and reported the best model result. The decision tree model built with recursive greedy algorithms ended up with a very complex structure with 36,797 and 13,459 leaf nodes at the census tract level and at the grid level, respectively. XGBoost and DNN are just as good as or better than simple MLR by a small margin at different spatial scales.

## 6. Prototype

## 7. Conclusion

## Acknowledgements

## Disclosure statement

[Download CSV](#)

[Display Table](#)

## Additional information

## Footnotes

## References

Table 3. Prediction of parking violation ticket counts using all features with different machine learning models.

[Download CSV](#)

[Display Table](#)

In addition, as expected, these machine learning models using all available features (Table 3) achieve a better performance compared with the same model but fed with only the local features. This is because the RMSE of the model using all features is significantly lower than the RMSE of the model using only local features. The RMSE of MLR reduced by 25% and 31% at the census tract level and at the grid level, respectively. This confirms our hypothesis that the machine learning models using all available features outperform the models using only local features. The results also show that the random forest model works the best with the minimum RMSE and outperforms other competitors by a large margin across all the spatial scales. Note that for the support-vector regression (SVR), we tried different kernel types (e.g. linear, second-degree polynomial, Gaussian RBF) and reported the best model result. The decision tree model built with recursive greedy algorithms ended up with a very complex structure with 36,797 and 13,459 leaf nodes at the census tract level and at the grid level, respectively. XGBoost and DNN are just as good as or better than simple MLR by a small margin at different spatial scales.



Figure 8. The learning curves for predicting the number of parking violation tickets with cross-validation at different spatial scales.

1. Introduction

2. Related work

3. Data

4. Methods

5. Results

6. Prototype

7. Conclusion

Acknowledgements

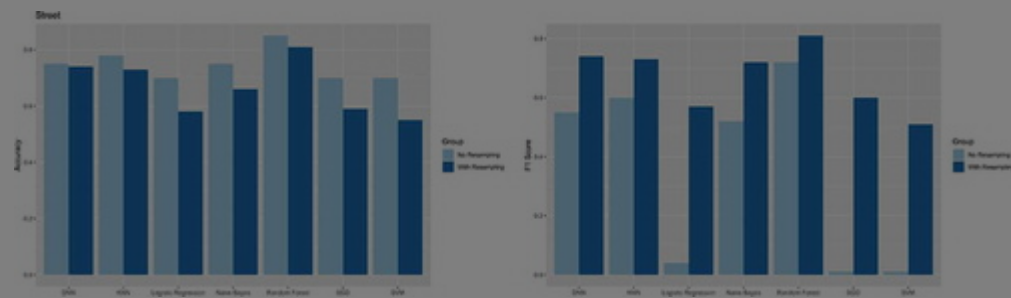
Disclosure statement

Display full size on

Footnotes

Figure 9. The accuracy and F1-score of different prediction models comparison using imbalanced and balanced samples with resampling process at the street scale.

References

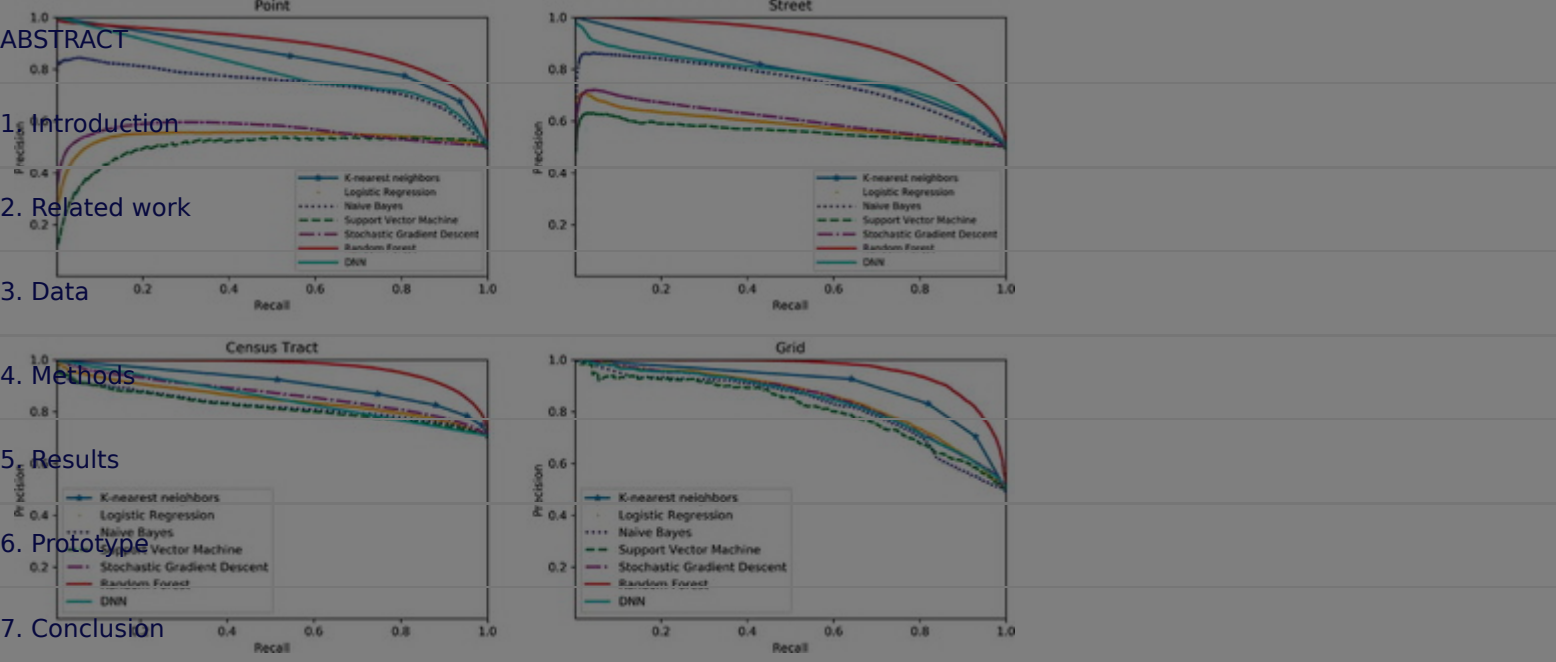


Display full

Figure 10  
different

results at





Display full size

Disclosure statement

### 5.3. Results for prediction of parking legality

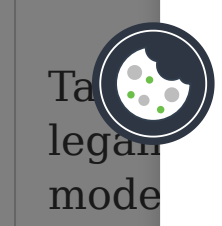
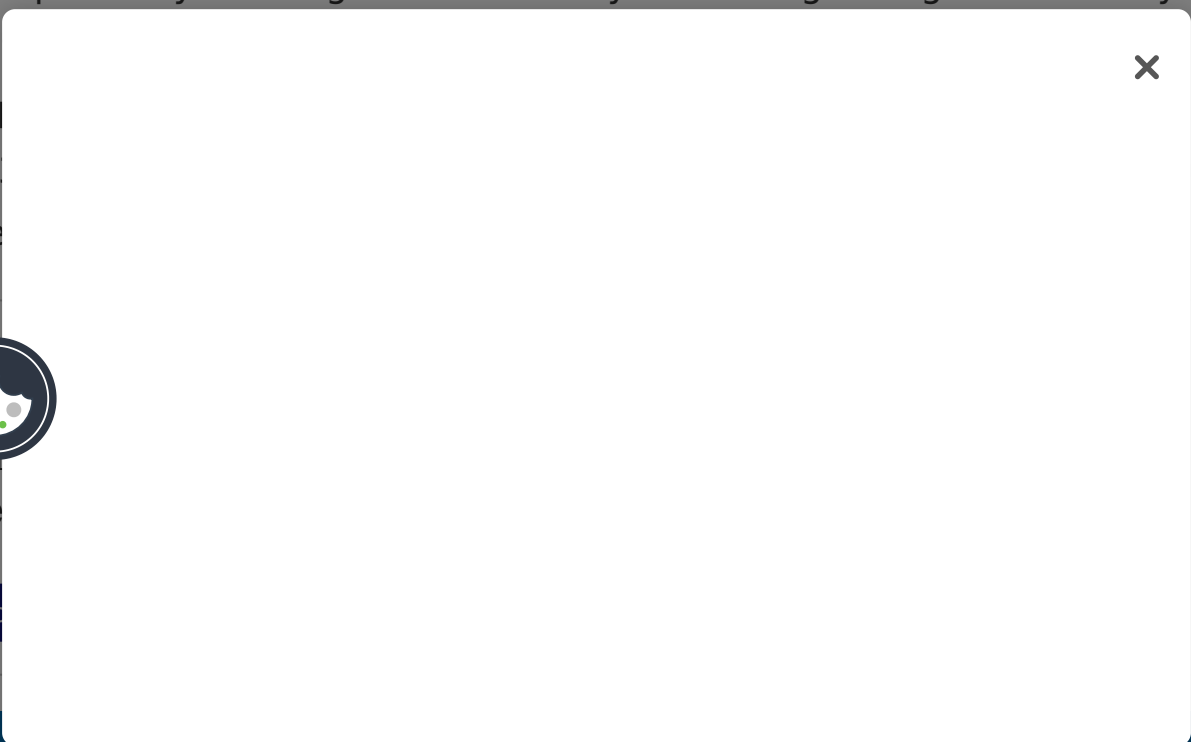
Additional information

Regarding the classification results, as shown in Table 4, random forest outperforms all other models and achieved both high accuracy scores (0.82, 0.85, 0.86, and 0.88) and

high F1-scores (0.82, 0.72, 0.90, and 0.88) across all four spatial scales. The KNN and the DNN also perform well and fall behind the random forest by a small margin. Note

that we chose  $k = 3$  as the number of nearest neighbours in feature space with regard to the temporal autocorrelation patterns of parking legality over time and its impact on model performance). The autocorrelation coefficient for parking legality with a temporal lag of 3 hours is 0.28, 0.40, 0.38, and 0.55 at the point, street, census tract, and grid levels, respectively. Although the Naive Bayes model gets a good accuracy and F1

scores and F1-scores (0.82, 0.72), it doesn't perform well at the census tract level (0.72, 0.65). We discuss these results in more detail in the

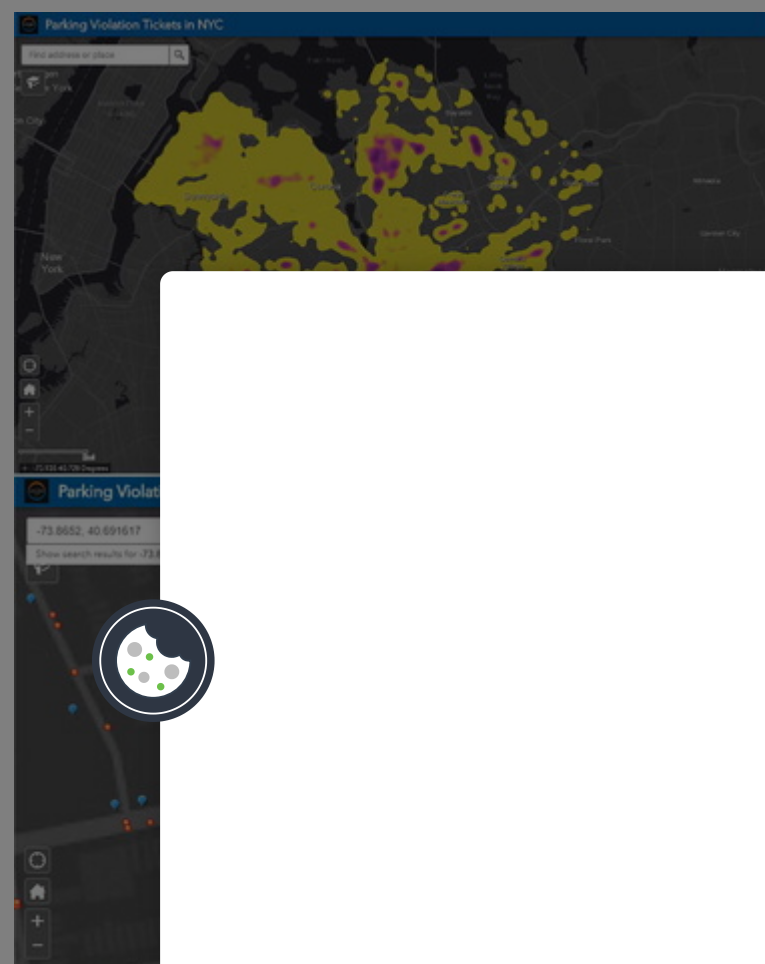


Download



Another important issue requiring attention is the presence of imbalanced training data in practice. In our case, there is a class imbalance between the positive class and the negative class for parking legality. The imbalanced training data can cause the accuracy paradox such that we get excellent accuracy but the accuracy is only reflecting the dominating class distribution. Taking the no-resampling street-level data as an example, as shown in [Figure 9](#), all the models got a high overall accuracy (over 0.7) but relatively low F1-scores, and some of the models (e.g. logistic regression, SVM and SGD) even got close to 0. That is mainly because of the imbalanced training samples between positive (29.7%) and negative (70.3%) classes. Therefore, we resampled the training data with a more balanced distribution for both positives (50%) and negatives (50%). The F1-score increased significantly since we get a better precision and recall performance. Alternatively, one may want to check the detailed precision-recall curves with different recall rates as shown in [Figure 10](#) to compare the model performance especially for highly skewed datasets (Davis and Goadrich [2006](#)). It shows that the random forest outperforms all other models with the highest precision value across different recall rates in parking legality prediction at four spatial scales.

[Figure 11](#). The Web interface of the multi-scale parking legality analysis platform based on Esri Web AppBuilder.



1. Introduction

2. Related work

## 6. Prototype

3. Data

4. Methods

With regard to the multi-level spatial variation of parking space availability and parking legality, we also worked to design and develop a parking legality Web GIS application by integrating the NYC parking violation open data with statistical analysis at different

5. Results

6. Prototype

7. Conclusion

Acknowledgements

Disclosure statement

Additional information

Footnotes

References

## 7. Conclusion

In this study, we propose a data-driven framework for understanding and predicting the spatiotemporal legality of on-street parking by training a set of machine learning models using the NYC parking violation ticket open data. The models are tested at four types of spatial analysis units (i.e. point, street, census tract, and grid) and the results confirmed the impact of spatial scale in machine learning predictive models. The more

detailed

error than

forest w

highest

Given a

than 2

which

that sev

accomm

violation

human r



frequency across space over time may have potential impact on generating the parking

violation tickets. With better on-street parking information provided in advance, drivers

1. Introduction

can enhance their parking decision-making. Our research may offer insights into

2. Related work

parking management policy such as parking regulation rules, pricing, and time

3. Data limitation to balance the parking demand and supply at different spatial scales using

3. Data

open data.

4. Methods

5. Results

## Acknowledgements

6. Prototype

7. Conclusion

The authors would like to thank NYC Open Data support and Safegraph Inc. for

providing the anonymous location data and POI visit patterns. Support for this research

Acknowledgements

was provided by the University of Wisconsin–Madison Office of the Vice Chancellor for

Disclosure statement

Research and Graduate Education with funding from the Wisconsin Alumni Research

Additional information

## Additional information

Footnotes

References

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Additional information

Funding

This work

Notes



1. [http://](#)

[neighbo](#)

2. [https:](#)

3. [https:](#)

In this article



# References

3. **Aksoy, S., and R. M. Haralick.** 2001. "Feature Normalization and Likelihood-based Similarity Measures for Image Retrieval." *Pattern Recognition Letters* **22** (5): 563-582. doi:10.1016/S0167-8655(00)00112-4.

4. [Web of Science](#) | [Google Scholar](#)

6. **Arnott, R., and E. Inci.** 2006. "An Integrated Model of Downtown Parking and Traffic Congestion." *Journal of Urban Economics* **60** (3): 418-442. doi:10.1016/j.jue.2006.04.004.

7. [Web of Science](#) | [Google Scholar](#)

9. **Berkson, J.** 1944. "Application of the Logistic Function to Bio-assay." *Journal of the American Statistical Association* **39** (227): 357-365.

10. [Google Scholar](#)

11. **Bock, F., Y. Attanasio, and S. Di Martino.** 2017. "Spatio-temporal Road Coverage of Probe Vehicles: A Case Study on Crowd-sensing of Parking Availability with Taxis." In *The Annual International Conference on Geographic Information Science*, 165-184. Wageningen, the Netherlands: Springer. doi:10.1177/1753193416662427.

12. [Google Scholar](#)

13. **Bock, F., Y. Attanasio, and S. Di Martino.** 2017. "Spatio-temporal Road Coverage of Probe Vehicles: A Case Study on Crowd-sensing of Parking Availability with Taxis." In *The Annual International Conference on Geographic Information Science*, 165-184. Wageningen, the Netherlands: Springer. doi:10.1177/1753193416662427.



14. **Breim, J.** 2017. "A Data-Driven Approach to GSPATIAL Data Analysis." *Journal of Data Science* **1** (1): 1-10. doi:10.1016/j.jds.2017.01.001.

7. Gaicedo, F., C. Blazquez, and P. Miranda. 2012. "Prediction of Parking Space Availability in Real Time." *Expert Systems with Applications* 39 (8): 7281-7290.

1. Introduction  
doi:10.1016/j.eswa.2012.01.091.

2. Related work | [Web of Science](#)® | [Google Scholar](#)

3. Data

8. Chatman, D. G., and M. Manville. 2014. "Theory versus Implementation in Congestion-priced Parking: An Evaluation of SFpark, 2011-2012." *Research in*

4. Methods  
5. Results  
*Transportation Economics* 44: 52-60. doi:10.1016/j.retrec.2014.04.005.

6. Prototype | [Web of Science](#)® | [Google Scholar](#)

7. Conclusion

9. Chen, T., and C. Guestrin. 2016. "Xgboost: A Scalable Tree Boosting System." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794. San Francisco, California: ACM.

Acknowledgements  
Disclosure Statement  
[Google Scholar](#)  
Additional information

Footnotes  
10. Chen, X., E. Santos-Neto, and M. Ripeanu. 2012. "Crowdsourcing for On-street Smart Parking." In *Proceedings of the second ACM international symposium on Design and analysis of intelligent vehicular networks and applications*, 1-8. Paphos, Cyprus: ACM.

[Google Scholar](#)

References  
1. Christiansen, P., Ø. Engebretsen, N. Fearnley, and J. U. Hanssen. 2017. "Parking Facilities and the Built Environment: Impacts on Travel Behaviour." *Transportation Research Part A: Policy and Practice* 95: 198-206.

2. Cook, ... parking  
Difficu ... predict.html

Goog

3. Co ... ning 20 (3):  
273-2

4. Cover ... n." IEEE



1. Introduction

5. Davis, J., and M. Goadrich. 2006. "The Relationship between Precision-Recall and ROC Curves." In Proceedings of the 23rd international conference on Machine learning, 233-240. Pittsburgh, Pennsylvania: ACM.

[Google Scholar](#)

4. Methods

6. Fotheringham, A. S., and D. W. S. Wong. 1991. "The Modifiable Areal Unit Problem in Multivariate Statistical Analysis." *Environment and Planning A* 23 (7): 1025-1044. doi:10.1068/a231025.

7. Conclusion

[Web of Science](#)® | [Google Scholar](#)

Acknowledgements

7. Friedman, J. H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *Annals of Statistics* 29: 1189-1232. doi:10.1214/aos/1013203451.

Additional information

[Web of Science](#)® | [Google Scholar](#)

Footnotes

8. Galton, F. 1886. "Regression Towards Mediocrity in Hereditary Stature." *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246-263. doi:10.2307/2841583.

[Google Scholar](#)

9. Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep Learning*. Cambridge, MA: MIT press.

[Google Scholar](#)

20. Hu, Y., et al. 2018. "Knowledge Discovery from Social Media: A Report the Knowledge Discovery (3): 16. doi:10.1007/978-1-4939-9739-7\_3



21. Huang, et al. 2018. "Knowledge Discovery from Social Media: A Report the Knowledge Discovery (3): 16. doi:10.1007/978-1-4939-9739-7\_3





2. Inci, E., and R. Lindsey. 2015. "Garage and Curbside Parking Competition with Search

Congestion." *Regional Science and Urban Economics* 54: 49–59.

1. Introduction  
doi:10.1016/j.regsciurbeco.2015.07.003.

2. Related work | [Web of Science](#)® | [Google Scholar](#)

3. Data

23. Jiang, Z., S. Shekhar, P. Mohan, J. Knight, and J. Corcoran. 2012. "Learning Spatial Decision Tree for Geographical Classification: A Summary of Results." In *Proceedings*

4. Methods  
5. Results  
6. Prototype  
of the 20th International Conference on Advances in Geographic Information Systems, 390–393. Redondo Beach, California: ACM.

[Google Scholar](#)

7. Conclusion

Acknowledgements

24. Litman, T. 2018. *Parking Management Best Practices*. New York, NY: Routledge.

Disclosure statement | [Google Scholar](#)

Additional information

25. Lu, F., K. Liu, Y. Duan, S. Cheng, and D. Fei. 2018. "Modeling the Heterogeneous Traffic Correlations in Urban Road Systems Using Traffic-enhanced Community

Footnotes  
References  
Detection Approach." *Physica A: Statistical Mechanics and Its Applications* 501: 227–237. doi:10.1016/j.physa.2018.02.062.

| [Web of Science](#)® | [Google Scholar](#)

26. Ma, S., O. Wolfson, and B. Xu. 2014. "UPDetector: Sensing Parking/unparking Activities Using Smartphones." In *Proceedings of the 7th ACM SIGSPATIAL international workshop on computational transportation science*, 76–85. Dallas-Fort Worth

27. Mao, H. 2016. *Report: The 1st ACM SIGSPATIAL International Workshop on Computing for Geographical Information Science (GIScience) 2016.*

28. Maron, J. 2016. *Proceedings of the ACM SIGSPATIAL International Workshop on Computing for Geographical Information Science (GIScience) 2016.* (JACM)

In this article



9. Mohri, M., A. Rostamizadeh, and A. Talwalkar. 2018. Foundations of Machine Learning.

Cambridge, MA: MIT press.

1. Introduction

[Google Scholar](#)

2. Related work

10. Naghibi, S. A., H. R. Pourghasemi, and B. Dixon. 2016. "GIS-based Groundwater Potential Mapping Using Boosted Regression Tree, Classification and Regression Tree, and Random Forest Machine Learning Models in Iran." Environmental Monitoring and Assessment 188 (1): 44. doi:10.1007/s10661-015-5049-6.

4. Methods

5. Results

6. Prototype

[PubMed](#)

[Web of Science ®](#)

[Google Scholar](#)

7. Conclusion

11. Openshaw, S. 1984. "The Modifiable Areal Unit Problem." Concepts and techniques in modern geography.

Acknowledgements

[Google Scholar](#)

Disclosure statement

Additional information

12. Pflügler, C., T. Köhn, M. Schreieck, M. Wiesche, and H. Krcmar. 2016. "Predicting the Availability of Parking Spaces with Publicly Available Data." Informatik 2016.

Footnotes

[Google Scholar](#)

References

13. Quinlan, J. R. 1986. "Induction of Decision Trees." Machine Learning 1 (1): 81-106. doi:10.1007/BF00116251.

[Google Scholar](#)

14. Shoup, D. 2017. The High Cost of Free Parking: Updated Edition. New York, NY:

Routledge



15. Shoup

doi:10

16.



16. Teodo

Opera

an Journal of

18.



7. Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46 (1): 234-240. doi:10.2307/143141.

1. Introduction

[Web of Science](#) <sup>®</sup> | [Google Scholar](#)

2. Related work

8. Vatsavai, R. R., A. Ganguly, V. Chandola, A. Stefanidis, S. Klasky, and S. Shekhar.

3. Data

2012. "Spatiotemporal Data Mining in the Era of Big Spatial Data: Algorithms and Applications." In *Proceedings of the 1st ACM SIGSPATIAL international workshop on*

4. Methods

analytics for big geospatial data, 1-10. Redondo Beach, California: ACM.

5. Results

[Google Scholar](#)

6. Prototype

7. Conclusion

9. Xu, B., O. Wolfson, J. Yang, L. Stenneth, S. Yu Philip, and P. C. Nelson. 2013. "Real-time

Acknowledgements

Street Parking Availability Estimation." In *2013 IEEE 14th International Conference on Mobile Data Management*, Vol. 1, 16-25. Milan, Italy: IEEE.

Disclosure statement

[Google Scholar](#)

Additional information

Footnotes

10. Yang, B., N. Fantini, and C. S. Jensen. 2013. "iPark: Identifying Parking Spaces from

References

Trajectories." In *Proceedings of the 16th International Conference on Extending Database Technology*, 705-708. Genoa, Italy: ACM.

[Google Scholar](#)

11. Yang, X., Z. Fang, Y. Xu, L. Yin, J. Li, and S. Lu. 2019. "Spatial Heterogeneity in Spatial Interaction of Human movements—Insights from Large-scale Mobile Positioning Data." *Journal of Transport Geography* 78: 29-40.

doi:10

12. Zhang

Gradie

conf

G

13. Zheng

Senso

Confe



1. Introduction

4. Zhu, A.-X., G. Lu, J. Liu, C.-Z. Qin, and C. Zhou. 2018. "Spatial Prediction Based on Third Law of Geography." *Annals of GIS* 24 (4): 225–240.

2. Related work

doi:10.1080/19475683.2018.1534890.

3. Data

Web of Science® | Google Scholar

4. Methods

5. Results

Download PDF

6. Prototype

7. Conclusion

Related research 

Acknowledgements

People also read

Recommended articles

Cited by  
19

Disclosure statement

Additional information

Footnotes

References



Large white modal window with a close button (X) in the top right corner.

[Information for](#)

[Open access](#)

[Authors](#)

[Overview](#)

[R&D professionals](#)

[Open journals](#)

[Editors](#)

[Open Select](#)

[Librarians](#)

[Dove Medical Press](#)

[Societies](#)

[F1000Research](#)

[Opportunities](#)

[Help and information](#)

[Reprints and e-prints](#)

[Help and contact](#)

[Advertising solutions](#)

[Newsroom](#)

[Accelerated publication](#)

[All journals](#)

[Corporate access solutions](#)

[Books](#)

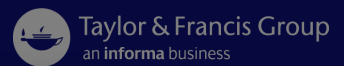
## Keep up to date

Register to receive personalised research and resources by email

 [Sign me up](#)



Copyright © 2024 Informa UK Limited [Privacy policy](#) [Cookies](#) [Terms & conditions](#)



Accessib

Registered  
5 Howick Pl



✕

