

Predicting Parking Availability from Mobile Payment Transactions with Positive Unlabeled Learning

Jonas Sonntag,^{1,2*} Michael Engel², Lars Schmidt-Thieme¹

¹ Information Systems and Machine Learning Lab, University of Hildesheim

²Volkswagen Financial Services AG

jonas.sonntag@vwfs.io, michael.engel@vwfs.io, schmidt-thieme@ismll.uni-hildesheim.de

Abstract

Cruising for parking in city centers is a problem for many motorists and for communities that need to reduce emissions. A widespread provision of parking assistance to address this problem requires a scalable system to generate availability information. Existing approaches to estimate the availability of parking spaces use supervised learning and depend on ground-truth labeling processes involving sensors or manual data collection. This dependency constraints the widespread roll-out and operation of such systems as the ground-truth data collection for model training, monitoring and retraining is prohibitively expensive. We describe a parking availability prediction system for paid on-street parking zones that does not depend on costly ground-truth labeling. The new approach uses solely data from parking ticket bookings via a mobile phone app. Every parking transaction serves as an implicit signal for the availability of one parking spot shortly before the booking. The system leverages this weak supervision signal by applying algorithms and metrics for positive-unlabeled learning (PU-learning). This approach enables the deployment in diverse regions, as well as the scalable monitoring and retraining of models. We evaluate our framework on a public dataset from Seattle.

Introduction

Cruising for parking is a major problem in urban areas. Drivers waste time and energy when searching for parking. The additional traffic causes congestion and the emissions pollute the environment. To address these problems, cities and companies alike work on smart parking systems, that optimize the parking process. The critical input for smart parking systems is accurate availability information, hence prediction of parking availability is an active research field in the transportation and applied machine learning domains.

We focus on the problem of determining the availability of parking spaces in paid on-street parking zones which are sometimes also called paid curb side parking areas. These parking zones represent a set of parking spaces often covered by the same parking ticket machine. There is no standard for how city administrations divide streets into paid



Figure 1: Parking zones in Seattle. Each line corresponds to one parking zone for which we aim to predict the availability status.

parking zones and little published data on how large these parking zones are or how many parking spaces they contain. Most high-demand on-street parking spaces in city centers and areas close to other points of interest are paid parking zones and hence they represent a highly relevant scope for availability prediction. An example for this parking zones in Seattle is shown in Figure 1. In Seattle one parking zone usually covers one street segment from one intersection to the next.

Various supervised learning approaches have been proposed that use data from sensors and cameras or parking meter transactions as a central data source, respectively. The dependency of the supervised methods on costly labeled data is a prohibitive factor to the widespread and sustainable deployment. Parking patterns are likely to vary between cities and regions. Hence, any supervised parking availability prediction method would require an economical ground-truth sampling scheme to validate the prediction quality when a model is deployed in a different region than the one it was trained on. Second, urban parking is arguably a non-stationary environment and data shifts are very likely to occur over time. The sustainable operation of a parking availability prediction system requires a scalable way to monitor the quality of the parking-availability information and to retrain the models if the prediction quality decreases.

Our objective is to enable the widespread, economical,

*As part of the Volkswagen Financial Services AG researchers program
Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

and sustainable deployment of an availability prediction system for paid on-street parking using machine learning. Meeting this objective requires achieving a prediction quality sufficient for practical purposes with data that can be collected over time in various regions at a very low cost. Our success criterion is therefore not to exceed the prediction performance of supervised parking availability models, but to eliminate the dependency on sources of labeled data that prevent an economical and widespread deployment.

In this article, we describe a novel availability prediction system for paid on-street parking zones that is based solely on ticket booking data recorded by a mobile phone app. We propose a labeling rule based on the mild assumption that at least one parking spot was free for one minute immediately before a transaction started to construct positive labels for parking spot availability. This labeling rule enables us to create training and evaluation data from a single, inexpensive and readily available data source and thereby resolves the dependency on costly ground-truth data collection.

The labeling rule generates data with only positive and unlabeled examples of parking availability. We cannot construct negative examples in the same way since the absence of parking transactions may indicate that a parking zone is fully occupied but also that it is vacant due to a temporary lack of demand. This requires the application of methods for positive-unlabeled (PU) learning and evaluation that were proposed in recent contributions to this active field of weakly supervised machine learning research.

In summary our main contributions are the following:

- We propose a data labeling rule for parking availability based on paid parking transactions which eliminates the dependency on costly manual or sensor based ground-truth data collection.
- We (re)frame the prediction of parking availability in paid parking zones as positive unlabeled (PU) learning task as opposed to existing supervised learning approaches.
- We describe a deployment-ready system including scalable model learning, verification, and monitoring components.
- We show that the proposed system performs well using data from Seattle.

Related Work

Parking Availability

Parking availability prediction is an active research field in applied machine learning. Several approaches were proposed that can be grouped by the ground-truth data collection variant they use.

One approach is to directly observe availability based on fixed sensor installations like cameras or in-ground sensors (Shinde et al. 2016; Bura et al. 2018), which constantly provide accurate parking data in real time. The high costs of installing and maintaining such systems prevent a widespread use beyond some pilot cities (Nandugudi et al. 2014).

The second central data source are parking meter transactions, which exist in abundance, as they are recorded in most urban regions during the payment process. It has been

shown that meter transactions are a good predictor of occupancy (Yang and Qian 2017; Sonntag and Schmidt-Thieme 2020). The existing approaches argue that they do not need ground-truth data during deployment, however they apply supervised learning and hence depend on ground-truth labels collected by sensors to train models and to monitor their service. Hence, the deployment of these models is limited to areas where at least historic ground-truth labeled data is available, which significantly narrows the scope.

One approach to limit the dependency on ground-truth data was recently introduced by Zhang et al. (2020). The authors propose a semi-supervised learning algorithm relying only on few sensor data and regard all other locations as unlabeled. A similar idea was developed by Ionita et al. (2018) where the authors transfer parking information from sensed areas to non-sensed areas by calculating similarity values between neighbourhoods based on background data. However both methodologies can not be trained in the complete absence of sensor data and equally important there is no way to monitor the accuracy of such a solution during deployment.

An other variant is the use of the sensor data created by mobile phones and vehicles to detect available parking spots (Carnelli et al. 2017; Krieg et al. 2016). A practical deployment, however, has to deal with the highly irregular spatio-temporal coverage of this data. It also depends on a curated representation of parking restrictions to not show available spaces where parking is prohibited (Arora et al. 2019).

Positive-Unlabeled Learning

Since one major aspect of our contribution is the reframing of parking availability prediction as positive-unlabeled learning problem we want to introduce some important concepts without going into details. For a detailed overview about PU-learning we recommend Bekker and Davis (2020). In general PU-learning tries to solve a classification problem in a setting where only some positive labels are available while the unlabeled data contains both positive and negative observations.

One important concept to enable learning in that setting is to make assumptions about the labeling process. Most methods like (Elkan and Noto 2008; Lee and Liu 2003) require the so called SCAR-assumption, which states that the labeled positives are Selected Completely At Random among all positives or more formally $p(x|y = 1) = p(x|y = 1, l = 1)$ where $l = 1$ is the property of the observation being labeled. More recent research tries to relax this assumption and allows a bias in the labeling process (Kato, Teshima, and Honda 2019).

An other important concept in PU-settings is the class prior $\pi = p(y = 1)$ which is required as an input for many model training approaches (Kiryo et al. 2017; Kato, Teshima, and Honda 2019). The problem of class prior estimation is therefore also well studied in the literature (Christoffel, Niu, and Sugiyama 2016; Ramaswamy, Scott, and Tewari 2016).

PU-learning arises naturally in many applications like identifying disease genes (Yang et al. 2012), text classification (Li and Liu 2003) and targeted marketing (Fei et al. 2013). To the best of our knowledge we are the first to propose PU-

learning in the transportation domain.

Our Methodology

We propose a prediction system for the availability of paid parking spaces using solely the transaction data that is automatically generated by the mobile app PayByPhone¹. The main idea of our approach is to leverage the transactions as a signal for parking space availability and use positive-unlabeled machine learning approaches to infer the current parking situation.

Our framework covers all components of a machine learning workflow as introduced by Ashmore, Calinescu, and Paterson (2019) including data management, model learning, model verification and monitoring. This section gives a detailed description of all the components. We summarize our system in Figure 3.

The system as described here is currently implemented to enable parking availability services in the PayByPhone App.

Data Management

We operate a system that receives all parking transactions that are paid via mobile phone in real-time. Since mobile phone transactions are already digitalised for parking enforcement purposes this is a straightforward and mild assumption. Each parking transaction comes with a location identifier, a start- and an expiration timestamp. Our goal is to predict the current availability status for all locations (i.e. whether there is at least one more free parking space) given only the transactional history.

The number of ongoing transactions does not directly translate into occupancy since people do not stick to their paid duration or do not pay at all (Yang and Qian 2017). Hence we lack ground-truth labels that are required for a standard supervised learning setting.

The key idea of our methodology is to define an implicit labeling approach that is based on the idea that a parking space was available right before someone parked and paid for a ticket. Namely we define the following labeling scheme:

Labeling Rule: *For each transaction T assign a positive label to the corresponding location at the previous timestamp.*

Figure 2 illustrates how this labeling approach works. Following this idea we can generate a positive availability label for every transaction that we receive and therefore constantly build and enlarge a training dataset. Since we consider parking locations with many parking spaces we can not generate negative labels (fully occupied) with the same logic which is why we are required to use positive-unlabeled (PU) learning methodologies to learn from that data. Furthermore we emphasize that our labeling is very sparse, i.e. we still don't have labels for the majority of location-timestamp combinations. The fact that we can observe only digital payments, hence a fraction of all parking transactions increases this challenge.

Parking transactions do not only provide labels but are also used to generate features for the model. Obviously one

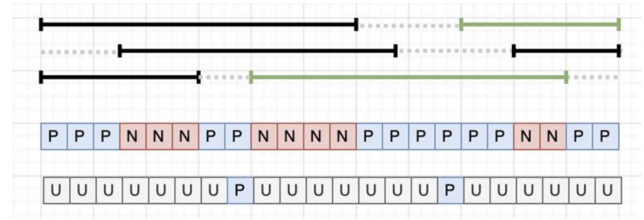


Figure 2: An example of our labeling approach. We show a parking location with three spaces over time. A black line indicates the parking session of an unobserved car while a green line is an observed parking session. The upper row shows the ground-truth label per timestamp while the lower row shows the labeled datapoints according to our labeling rule.

can consider the number of cars with a valid ongoing session as a signal for availability but one can also consider more complicated representations of the transactional history. In earlier research we showed that features generated from parking transactions can be exploited to learn about parking availability when there is a ground-truth dataset available for training (Sonntag and Schmidt-Thieme 2020). We describe our detailed feature engineering approach in the experiments section.

Model Learning and Verification

Since our labeling approach can only generate positive labels of availability, we are dealing with the well known positive-unlabeled (PU)-learning setting. In the last years a variety of models were proposed to deal with this special setting of semi-supervised learning. Our system in principal allows any PU-model to be trained on the dataset created by our labeling rule and to be deployed afterwards. The more tricky part is how to decide in the absence of ground-truth data

- which model to choose among all candidates,
- whether a certain model is good enough for deployment.

Hence we need an evaluation schema based only on positive and unlabeled data. This is a general problem in PU-settings and hence has already received some attention from a theoretical perspective. In our framework we use the results from Jain et al. (2016). They propose to calculate the false positive rate η and true positive rate γ based on the classification between labeled and unlabeled data (i.e. we regard all unlabeled data as negative). Then the precision p_{PU} and recall r_{PU} of the PU-model can be estimated based on the pre-estimated class prior π . For the recall w.r.t. PU-setting we have simply $r_{PU} = \gamma$ while for the precision we have $p_{PU} = \frac{\pi\gamma}{\eta}$. Hence the PU-adopted F1 score can be defined as

$$F1_{PU} = 2 \frac{p_{PU} \cdot r_{PU}}{p_{PU} + r_{PU}}. \quad (1)$$

We therefore can evaluate all models on a positive-unlabeled testset generated from transaction data and choose the best performing model as release candidate. Furthermore the PU-

¹<https://www.paybyphone.com/>

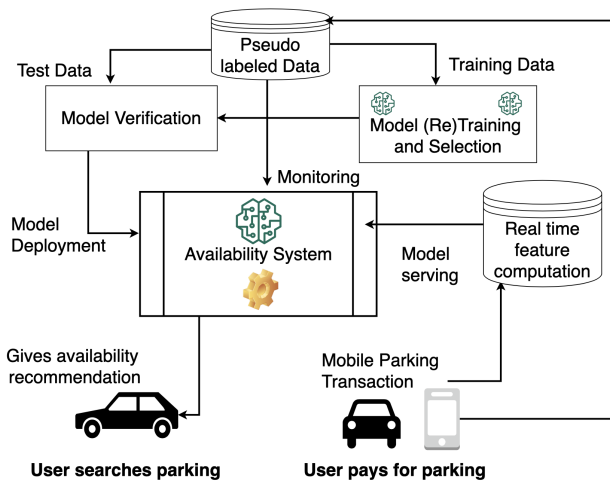


Figure 3: Our proposed parking system architecture

F1-score enables us to set a threshold of a minimum performance before deploying a new model to production.

Deployment and Monitoring

With the approach described so far we can generate PU-models for every city that allows mobile phone payment for parking. Furthermore we can decide whether we want to include a given city based on the model performance on the PU-testset. Since mobile parking transactions are already digital we can easily get access by deploying a simple data forwarding pipeline and hence calculate features in realtime. With access to transaction data all other components can be build using standard machine learning architectures and best practices.

Serving a parking availability system in production requires the capability of constantly monitoring the quality and staleness of the model (Breck et al. 2017). For parking availability systems based on supervised learning this used to be only possible with significant time delay and high costs, since new hand-labeled data needed to be acquired for validation. However given our evaluation schema based on equation (1) we can easily monitor the performance over time and intervene if quality deteriorates.

Experiments

Based on the real-word ground-truth data from Seattle we empirically prove that our system achieves good performance and our evaluation schema provides meaningful ranking of different models. Furthermore we show that our model verification process identifies areas with poor performance during training and hence prevents a bad user experience.

Data

We use the public available Seattle Annual-Parking-Study-Data² as ground-truth for evaluation purposes. The data

²<https://data.seattle.gov/Transportation/Annual-Parking-Study-Data/7jzm-ucez>

is manually collected on the street by the Seattle Department of Transportation starting in 2014. Most data is from spring each year while some datasets include summer studies. Some insights and further information on this study are publicly available³.

We were provided corresponding mobile parking transaction data from 2015 to 2019 by the company PayByPhone. The rate of mobile transactions among all paid parking transactions varies between 20 and 50 percent depending on the area.

The data is sampled from 25 different areas in Seattle where each area has it’s own parking policy in terms of pricing, opening hours and maximum duration of stay. One parking zone in Seattle usually corresponds to one street segment from one intersection to the next (as shown in Figure 1).

In the study the city reports the current number of cars at a given street block. Since parking spaces are not necessarily marked we estimate the capacity of a block with the maximum number of cars that that were observed in parallel.

We create the labeled dataset by creating a labeled datapoint for every transaction one minute before the transaction starts and add the same amount of unlabeled data by randomly selecting location-time pairs where no transactions were observed.

Based on the transactions and our methodology we provide a positive label to roughly 2.5 percent of the observations.

Although the transaction data is not completely public the city publishes recent transaction data⁴ with the same format than the data we used.

Feature Engineering

Considering a location l at a given timestamp, we have access to all previous transactions T . We consider all parking transactions in T that are valid for location l , have already started and are not expired for more than one hour. For those transactions we calculate

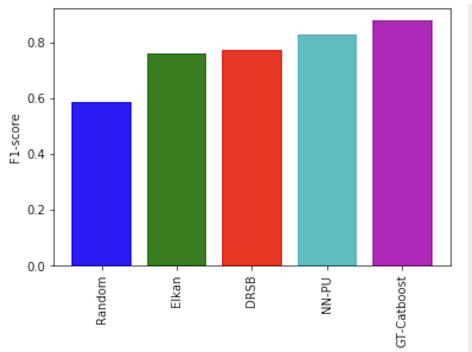
- Remaining valid time in minutes
- Time passed since transaction started in minutes

Both features are computed per transaction and we use padding to account for different numbers of transaction per timestamp. These features represent the current occupancy level based on transactions where we account for the fact that the probability of a car still being parked at time t changes with the length and the remaining time of a parking session (Yang and Qian 2017). We use the following additional features per location and timestamp

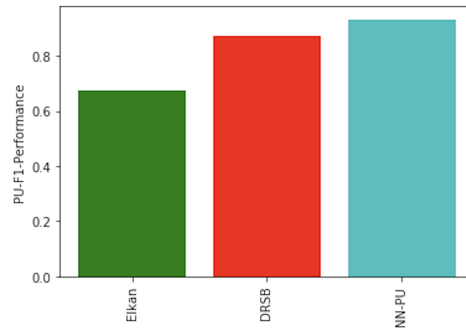
- number of ongoing transactions
- location identifier (categorical)
- hour of the day
- weekday
- year

³www.seattle.gov/Documents/Departments/SDOT/-ParkingProgram/PaidParking/SDOT_AnnualReport2018.pdf

⁴<https://data.seattle.gov/Transportation/Paid-Parking-Transaction-Data/gg89-k5p6>



(a) F1 score evaluated on ground-truth data in Seattle



(b) PU-F1-score evaluated in our PU-Framework based solely on transactions in Seattle Belltown

Figure 4: PU models show good performance when tested on ground-truth data (4a) and model selection in the absence of ground-truth data based on PU-F1 score preserves ordering (4b).

The identifier enables us to learn location specific availability patterns while the other features can capture temporal trends and seasonality in the data.

We want to emphasize the fact that although we deal with time dependant observations we consider our setting not as a time series prediction problem. In classical time series prediction one is able to observe a process x_t and aims to predict a future state x_{t+h} while in our setting we are not able to observe the process of availability regularly for $t < t_0$. We leave the question whether PU-based time series classification (Nguyen, Li, and Ng 2011) can increase the performance for future research.

Model Candidates

There exists various methods in the literature for learning on positive and unlabeled data. For this work we considered three different methodologies with different assumptions and strengths.

1. With **Elkan** we denote the methodology proposed by Elkan and Noto (2008). It's considered as one of the pioneers work in PU-learning and is supposed to work well under the standard SCAR assumption. As base models to classify labeled vs unlabeled data we used linear regression, random forest and SVM during hyperparameter training and select the one that performs best on $F1_{PU}$.
2. **DRSB** is the density ratio estimation proposed by Kato, Teshima, and Honda (2019) which is able to handle data with a selection bias in the labeling process. Since this methodology requires a pre-estimated class prior we follow their work and use the km2 estimator proposed by Ramaswamy, Scott, and Tewari (2016).
3. **NN-PU** refers to the non-negative risk estimator implementation in combination with neural network architectures for PU-learning by Kiryo et al. (2017). We considered network architectures with 3 and 6 layers with 100 nodes each and ReLU activation function during hyperparameter training. As for the DRSB method we used the class prior estimation from (Ramaswamy, Scott, and Tewari 2016).

Model	Whole city	Uptown (best area)	Denny Triangle (worst area)
Random	0.58/0.64	0.61/0.65	0.51/0.63
Elkan	0.75/0.67	0.79/0.66	0.62/0.72
DRSB	0.77/0.87	0.82/0.80	0.68/0.88
NN-PU	0.82/0.93	0.89/0.91	0.71/0.90
GT-Catboost	0.88/	0.91/	0.83/

Table 1: Model performance in Seattle

We furthermore consider non-PU methodologies as references

1. **GT-Catboost** is the Catboost model by Prokhorenkova et al. (2018), which is a state-of-the-art gradient boosting algorithm, trained on fully labeled ground-truth data. Since all labels are made available during training we can not hope to outperform this approach, but we consider it as an indicator of the predictability of the dataset and as an upper bound of what one can achieve with massive investments in ground-truth data.
2. With **Random** we denote a random model which guesses the availability state of a parking location with equal probability on each class. We consider this as a lower bound to prove that our labeling approach provides useful information to the model.

Results

In order to provide a first indication that our approach works, we regard Seattle as the pilot city to roll-out our availability system and train one model for the whole city.

In addition to that we consider each area separately to provide some statistics about model quality and verification. We report ground-truth results in terms of F1-score as well as PU-based performance in terms of estimated PU-F1 score with equation (1) in Table 1 for the whole city as well as for two selected areas, namely Uptown where we achieve the best results and Denny Triangle where we achieve the lowest performance after removing outliers in the model verifi-

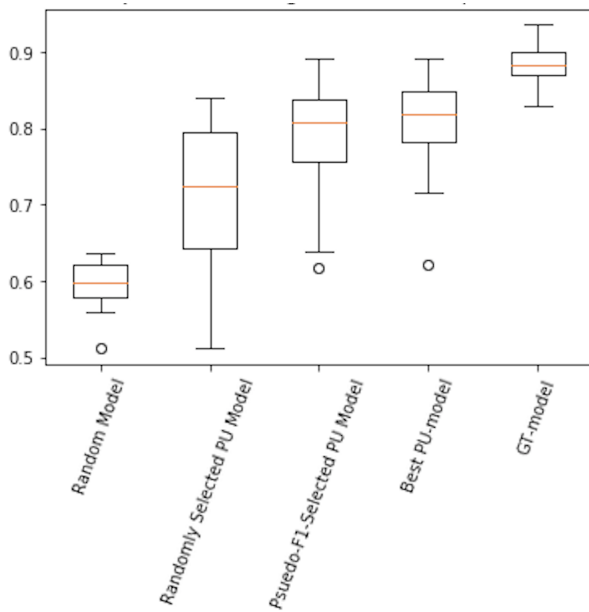


Figure 5: Distribution of F1 performance in 19 areas of Seattle

cation step.

We first take a closer look at the complete city of Seattle. As expected the model trained on ground-truth performs best and hence should be favored when smart sensor data allows the collection of such data.

However we find that PU-learning trained on the data created by our labeling rule provides strong performance given the huge amount of unlabeled data and the high scalability of this approach. Comparing the three candidate models we find that every considered PU-model is significantly better than no model (random guessing) while the more sophisticated models outperform the relatively simple model by Elkan and Noto (2008).

For the purpose of evaluation and model selection it is important to note that we find the same order of the models when considering the PU-F1 score calculated on the pseudo-labeled testset (see Figure 4).

We now take a closer look at different areas and consider each one separately, i.e. for each area we train all models and select the one that performs best based on the pseudo-labeled dataset with PU-F1-score. After removing areas with less than 150 ground-truth observations we conduct the experiments in the remaining 19 areas.

The distribution of the performance of the PU-models among different areas as shown in Figure 5 proves that the PU approach works well in the vast majority of the areas. The average improvement of the PU-model chosen based on PU-metrics over a random model is above 30 percent.

However we also notice outliers which indicates that the training can fail for certain areas or cities. We therefore need to investigate whether we can identify such cities already in our verification step to prevent a deployment of our system in that specific area. That is why we investigate the relation-

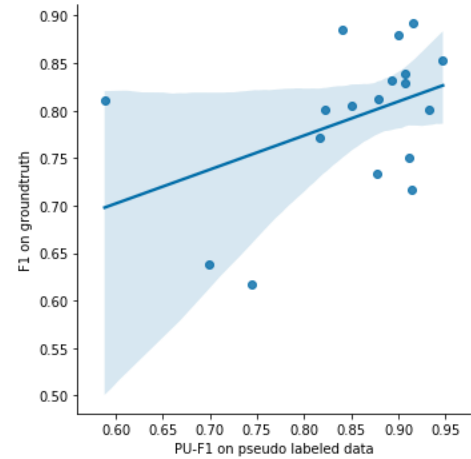


Figure 6: Correlation between PU metric and on-street performance

ship between PU-metrics and on-street performance in the next section more systematically.

Model evaluation with PU-metrics

Model evaluation based on PU-metrics is used in our framework for model selection as well as model verification. We first investigate the issue of model selection during training. Figure 5 shows the distribution of the achieved on-street F1-scores among all areas for PU-models selected by our evaluation schema, randomly selected PU-models among all three candidates and PU models chosen based on their (unknown) ground-truth performance, hence always choosing the best model.

We find that PU-F1 score gives good indications about ground-truth performance with an accuracy of correctly chosen models of 72 percent and an average drop of F1-performance of only two percent compared to the best performing PU-models. On the other hand choosing models with PU-metrics gives an eleven percent boost compared to randomly chosen PU-models. Hence we find that PU-F1-score provides a reliable and inexpensive way to select models for production.

However one has to be careful when interpreting the PU-F1-score since we found that this metric is systematically overestimating the actual performance on the street (see Figure 6). Also the correlation between the two metrics is not as clear as one would hope. A possible reason for this behaviour is the area-dependant class prior estimation-error since the class prior is critical for PU-metric calculation. We leave the question how to optimally estimate the class prior and it's influence in the parking availability case for future research.

However it is also clearly visible that exceptional bad ground-truth performance is reflected in a low PU-F1 metric. With regard to model verification we find that setting a threshold of PU-F1-score around 0.75 eliminates the two worst-performing areas while removing only one area with good performance.

Conclusion

We describe a deployment-ready system for the prediction of the availability of parking spaces at the parking-zone level. The system depends solely on the stream of digital parking fee payment transactions, in our example made via the mobile parking app PayByPhone. The system stands out in that it resolves the dependency on ground-truth labels which is the bottleneck of existing systems that apply supervised learning methods to predict parking availability. A key advantage besides saving the cost for ground-truth data collection is the systems scalability, as mobile parking payment services are available in many cities worldwide and the relevant data is already digital.

The core principle is to enhance the parking payment transaction data with the help of a labeling rule. The labeling rule is based on the mild assumption that each payment transaction indicates the availability of at least one parking space at the respective location shortly before, and assigns a positive availability label, respectively. We further described a complete model training and evaluation schema based on the only positive and unlabeled (PU) data that we created by applying the labeling rule. Experiments with PU-learning methodologies from the literature evaluated in different areas in Seattle show that the approach yields an accuracy suitable for practical use.

While there is some research about minimizing the dependency on ground-truth data, e.g. by using it only for training or exploiting semi-supervised learning approaches, we are the first to propose an end-to-end parking availability system that can even be trained and monitored in the complete absence of labeled ground-truth data. The scalability of our approach comes at the cost of some prediction accuracy loss compared to the supervised learning setting, however still achieves decent results.

We furthermore argue that this paper can be regarded as a proof-of-concept that an intuitive designed labeling rule based on transactions provide valuable information that can be leveraged by standard PU-learning methods. We believe that the performance can be further increased by building PU-models that are designed for our specific problem type, e.g. by regarding parking occupancy as a time series. Also the problem of class prior estimation in the case of parking availability should be further investigated since it's a major part of model training as well as evaluation. We were able to show that a given state-of-the art approach yields good results but also leads to an overestimated PU-F1-score during evaluation. To investigate the performance of different class prior estimations and the influence to the overall system is an other question for further research.

An other interesting approach for further research is to design additional labeling rules based on parking domain knowledge and use the data programming framework introduced by Ratner et al. (2016) which can deal with noisy heuristic labels.

References

Arora, N.; Cook, J.; Kumar, R.; Kuznetsov, I.; Li, Y.; Liang, H.-J.; Miller, A.; Tomkins, A.; Tsogsuren, I.; and Wang,

Y. 2019. Hard to Park? Estimating Parking Difficulty at Scale. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2296–2304.

Ashmore, R.; Calinescu, R.; and Paterson, C. 2019. Assuring the machine learning lifecycle: Desiderata, methods, and challenges. *arXiv preprint arXiv:1905.04223*.

Bekker, J.; and Davis, J. 2020. Learning from positive and unlabeled data: a survey. *Mach. Learn.* 109(4): 719–760.

Breck, E.; Cai, S.; Nielsen, E.; Salib, M.; and Sculley, D. 2017. The ML test score: A rubric for ML production readiness and technical debt reduction. In *2017 IEEE International Conference on Big Data, BigData 2017, Boston, MA, USA, December 11-14, 2017*, 1123–1132. IEEE Computer Society.

Bura, H.; Lin, N.; Kumar, N.; Malekar, S.; Nagaraj, S.; and Liu, K. 2018. An Edge Based Smart Parking Solution Using Camera Networks and Deep Learning. In *Proc. IEEE Int. Conf. Cognitive Computing (ICCC)*, 17–24. doi:10.1109/ICCC.2018.00010.

Carnelli, P. E.; Yeh, J.; Sooriyabandara, M.; and Khan, A. 2017. ParkUs: A Novel Vehicle Parking Detection System. In Singh, S. P.; and Markovitch, S., eds., *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*, 4650–4656. AAAI Press.

Christoffel, M.; Niu, G.; and Sugiyama, M. 2016. Class-prior estimation for learning from positive and unlabeled data. In *Asian Conference on Machine Learning*, 221–236.

Elkan, C.; and Noto, K. 2008. Learning classifiers from only positive and unlabeled data. In Li, Y.; Liu, B.; and Sarawagi, S., eds., *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008*, 213–220. ACM.

Fei, H.; Kim, Y.; Sahu, S.; Naphade, M.; Mamidipalli, S. K.; and Hutchinson, J. 2013. Heat pump detection from coarse grained smart meter data with positive and unlabeled learning. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1330–1338.

Ionita, A.; Pomp, A.; Cochez, M.; Meisen, T.; and Decker, S. 2018. Where to Park?: Predicting Free Parking Spots in Unmonitored City Areas. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics, WIMS '18*, 22:1–22:12. New York, NY, USA: ACM.

Jain, S.; White, M.; Trosset, M. W.; and Radivojac, P. 2016. Nonparametric semi-supervised learning of class proportions. *arXiv preprint arXiv:1601.01944*.

Kato, M.; Teshima, T.; and Honda, J. 2019. Learning from Positive and Unlabeled Data with a Selection Bias. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

- Kiryu, R.; Niu, G.; du Plessis, M. C.; and Sugiyama, M. 2017. Positive-Unlabeled Learning with Non-Negative Risk Estimator. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*. URL <http://papers.nips.cc/paper/6765-positive-unlabeled-learning-with-non-negative-risk-estimator>.
- Krieg, J.; Jakllari, G.; Toma, H.; and Beylot, A. 2016. Unlocking the smartphone's senses for smart city parking. In *Proc. IEEE Int. Conf. Communications (ICC)*, 1–7. ISSN 1938-1883. doi:10.1109/ICC.2016.7511231.
- Lee, W. S.; and Liu, B. 2003. Learning with Positive and Unlabeled Examples Using Weighted Logistic Regression. In Fawcett, T.; and Mishra, N., eds., *Machine Learning, Proceedings of the Twentieth International Conference (ICML 2003), August 21-24, 2003, Washington, DC, USA*, 448–455. AAAI Press.
- Li, X.; and Liu, B. 2003. Learning to Classify Texts Using Positive and Unlabeled Data. In Gottlob, G.; and Walsh, T., eds., *IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico, August 9-15, 2003*, 587–594. Morgan Kaufmann.
- Nandugudi, A.; Ki, T.; Nuessle, C.; and Challen, G. 2014. Pocketparker: Pocketsourcing parking lot availability. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 963–973.
- Nguyen, M. N.; Li, X.-L.; and Ng, S.-K. 2011. Positive unlabeled learning for time series classification. In *Twenty-Second International Joint Conference on Artificial Intelligence*.
- Prokhorenkova, L. O.; Gusev, G.; Vorobev, A.; Dorogush, A. V.; and Gulin, A. 2018. CatBoost: unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada*.
- Ramaswamy, H. G.; Scott, C.; and Tewari, A. 2016. Mixture Proportion Estimation via Kernel Embeddings of Distributions. In *Proceedings of the 33rd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, volume 48 of *JMLR Workshop and Conference Proceedings*, 2052–2060.
- Ratner, A. J.; De Sa, C. M.; Wu, S.; Selsam, D.; and Ré, C. 2016. Data programming: Creating large training sets, quickly. In *Advances in neural information processing systems*, 3567–3575.
- Shinde, S.; Bhagwat, S.; Pharate, S.; and Paymode, V. 2016. Prediction of parking availability in car parks using sensors and IoT: SPS. *International Journal of Engineering Science and Computing* 6.
- Sonntag, J.; and Schmidt-Thieme, L. 2020. A machine learning approach to infer on-street parking occupancy based on parking meter transactions. In *23rd International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2020*.
- Yang, P.; Li, X.-L.; Mei, J.-P.; Kwok, C.-K.; and Ng, S.-K. 2012. Positive-unlabeled learning for disease gene identification. *Bioinformatics* 28(20): 2640–2647.
- Yang, S.; and Qian, Z. S. 2017. Turning meter transactions data into occupancy and payment behavioral information for on-street parking. *Transportation Research Part C: Emerging Technologies* 78: 165–182.
- Zhang, W.; Liu, H.; Liu, Y.; Zhou, J.; and Xiong, H. 2020. Semi-Supervised Hierarchical Recurrent Graph Neural Network for City-Wide Parking Availability Prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 1186–1193.