



Urban Human Mobility Modelling and Prediction: Impact of Comfort and Well-Being Indicators

Luís Rosa¹(✉), Fábio Silva^{1,2}, and Cesar Analide¹

¹ Centro ALGORITMI, Department of Informatics, University of Minho, Braga, Portugal

id8123@alunos.uminho.pt, analide@di.uminho.pt

² CIICESI, ESTG, Politécnico do Porto, Felgueiras, Portugal
fas@estg.ipp.pt

Abstract. There are increasingly more discussions on and guidelines about different levels of indicators surrounding smart cities (e.g., comfort, well-being and weather conditions). They are an important opportunity to illustrate how smart urban development strategies and digital tools can be stretched or reinvented to address localised social issues. Thus, multi-source heterogeneous data provides a new driving force for exploring urban human mobility patterns. In this work, we forecast human mobility using indoor or outdoor environment datasets, respectively, Metropolitan Transportation Authority (MTA) Wi-Fi and LinkNYC kiosks, collected in New York City to study how comfort and well-being indicators influence people's movements. By comparing the forecasting performance of statistical and Deep Learning (DL) methods on the aggregated mobile data we show that each class of methods has its advantages and disadvantages depending on the forecasting scenario. However, for our time-series forecasting problem, DL methods are preferable when it comes to simplicity and immediacy of use, since they do not require a time-consuming model selection for each different cell. DL approaches are also appropriate when aiming to reduce the maximum forecasting error. Statistical methods instead have shown their superiority in providing more precise forecasting results, but they require data domain knowledge and computationally expensive techniques in order to select the best parameters.

Keywords: Human mobility patterns · Device network datasets · Deep learning methods · Statistical methods

1 Introduction

Recent studies about human mobility, comfort and well-being and social interactions evaluated impacts on the perceptions of citizens [11, 12, 20, 21]. The match between human activities in city and urban infrastructures may be the main

contributor for these works. Additionally, human mobility is associated with a large personal and societal cost, with problems being attributed to a combination of individual factors (physical, cognitive and psychological) combined with environmental conditions [6]. As an example, the relationship between human mobility behavior and climate—namely, weather and environmental conditions when travel planning decisions are made. Meteorological effects could influence travel demand and route choices in various ways, including diversion to other trip modes or paths, or deferring and cancelling trips [23].

On the other hand, smartphones and embedded sensor systems have given researchers unprecedented access to new and rich datasets, recording detailed information about how people live and move through urban areas. We can select a number of examples that highlight how datasets generated from these devices are lending insight into individuals lives and urban analysis. For example, in [13], embedded sensors were used to measure the spatio-temporal patterns of an entire city’s usage of a shared-bicycle scheme. Other approaches used Bluetooth sensors to measure social interactions [10] or GPS sensors to show urban planning and design [5]. Lastly, [8] uses the dataset from public transport automated fare collection systems which was previously used to investigate travellers’ perceptions.

In this paper, we study mobility patterns evaluating and comparing the performance of classical and modern Machine Learning (ML) methods based on two approaches, univariate and multivariate, using two sets of time series forecasting datasets, which provide the temporal variations of census over time. The traditional methods considered are Autoregressive Integrated Moving Average (ARIMA) model, Autoregressive Integrated Moving Average model with exogenous variables (ARIMAX), Seasonal Autoregressive Integrated Moving Average (SARIMA) model and Seasonal Autoregressive Integrated Moving Average model with exogenous variables (SARIMAX). These are types of traditional time series models most commonly used in time series forecasting. DL models such as Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), hybrid CNN-LSTM and Bi-Directional LSTM were the ML techniques explored. These ML methods are capable of identifying structure and pattern of data such as non-linearity and complexity in time series forecasting. Each model must be studied and understood in a parameterized way so that the integration with any data set does not cause any problem. In fact, experimental results confirm the importance of performing a parametric grid search when using any forecasting method, as the output of this process directly determines the effectiveness of each model.

All this is possible thanks to the availability of data describing long-term human behaviour on mobile phones. The available data is based on a few years of network traffic generated by LinkNYC Kiosk devices, MTA Wi-Fi Locations, based on the city of New York, and context reactions of citizens via their smartphones. Basically, in a modern society where smartphones are widely used, understanding the impact of environmental factors, comfort and well-being

indicators has both theoretical and practical implications in understanding and modelling human behaviour.

The rest of the paper is planned as follows: Sect. 2 focuses on a study about crowdsensed data from mobile devices and different human mobility forecasting methods such as DL and statistical models. In the next section, we execute an experimental case study that covers the benefits of Neural Network (NN) and statistical techniques in human mobility. In Sect. 4, we discuss the results of the case study. In Sect. 5, the conclusion summarizes the article’s arguments before extending the debate further by offering trajectories for future investigation on the prediction of human mobility.

2 State of the Art

For the realization of this project some concepts should be defined. In order to clarify their meaning and guarantee the quality of the project, the next sub-chapters introduce two crowdsensing infrastructures and indicators which are crucial to the understanding of the present work.

2.1 Crowdsensing Infrastructures

Contributing to this literature, this article investigates the human mobility that is captured with the development of the new public Wi-Fi infrastructure which is gradually making an appearance in cities across the world; such an infrastructure is growing steadily across New York City in recent times, and is called LinkNYC or Link [14]. This network infrastructure has been adapted and deployed to provide a free Wi-Fi service. It has transformed the way information is delivered in city streets, and supporting civic engagement has become a core part of our research. With thousands of screens encouraging New Yorkers to interact and offering helpful resources, it can provide strong participation of citizens in this work. As we see in Fig. 1, there are more than 1,800 LinkNYC kiosks around the city, including hundreds in Brooklyn.

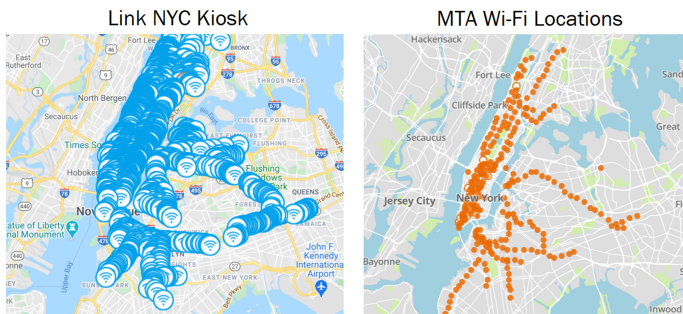


Fig. 1. Map of LinkNYC kiosk and MTA Wi-Fi locations.

In its turn, Transit Wireless' mission keeps millions of New York City subway riders connected, safe and informed via Wi-Fi network connectivity [7]. Figure 1 shows the 282 stations more than 100 ft below ground, and 109 stations above ground with endless miles of tunnels and bustle [3]. It only contains stations that are considered Wi-Fi-ready.

In both infrastructures, wireless network availability along with crowding in streets or public transport provides crowdsensing research opportunities based on people flow or passenger volume. Because rush hour (peak) is characterised by big spikes of demand concentrated in rather short time periods, leaving the transport network under-utilised before and after such spikes, it presents a real opportunity to understand human movements.

2.2 Well-Being and Comfort

The advances in mobile computing and Artificial Intelligence (AI) techniques enable people to probe the dynamics of human movements in a city. We can analyse the impact of well-being and comfort indicators in these dynamics using crowd sensing with the two datasets proposed in this paper.

Elena Alatarsteva and Galina Barysheva [1] argue that the modern man can be defined with regard to two levels of well-being: internal (subjective) and external (objective). In the external strand, well-being could be characterized by wage levels, residence conditions, educational opportunities, the environment, safety and civil rights. In its turn, the internal strand is conceptualized only as an internal state of an individual. However, other authors from different branches specify the definition of this concept. Their articles categorized it into different classes: Community Well-being [2], Economic Well-being [15], Emotional Well-being [22], Physical Well-being [18], Development and activity [17], Psychological Well-being [19] and Work Well-being [4]. Although these classes categorize well-being in multiple ways, they have common points.

On the other hand, regarding comfort, it is difficult to reach a consensus from literature on its definition. Some papers show factors that influence comfort. One of them shows that different activities can influence comfort, concluding that characteristics of the environment and the context can change how people feel [24]. Although it is often considered a synonym for well-being, it classifies the atmosphere that surrounds the human being. However, a mental health organization in the UK has argued that "it is important to realize that well-being is a much broader concept than moment-to-moment happiness" [9]. In other words, the comfort indicator is characterized by an extensive variety of factors, which associates it with a long-term context, e.g. a person may find himself comfortable but unhappy (and vice versa).

As we see, comfort and well-being are distinct terms, but we believe that from them our experimental case study can allow understanding the mechanism hidden in human mobility that affects New York City both at community and individual levels.

3 Experimental Case Study

This experimental case is particularly useful in investigating “how” and “why” questions concerning human mobility behaviours. As a qualitative research methodology, this case study focuses on understanding these phenomena in broader circumstances than those in which it is located. Our study aims to investigate the comfort, well-being and motivation through questionnaire-based online surveys, and further understand a complex social phenomenon in human mobility: how citizens react in indoor and outdoor environments, and why.

3.1 Data Collection

We designed and conducted this study involving LinkNYC kiosk data contributed by one hundred thousand of users, while MTA Wi-Fi Locations captured fifty thousand interactions with smartphones in subway locations. In order to enrich our dataset, well-being and comfort metrics were gathered via questionnaire-based mobile applications [20, 25]. In these individual forms, users were asked about their comfort and well-being voluntarily based on the environment they were in. In order to collect respondents’ attitudes and opinions, these works adapted a response scheme like Likert scale, commonly used in opinion polls.

Other information can be considered, like the weather. We used an API so that the information gathered was even wider. This includes, for example, the Meteostat API that enables the collection of a vast amount of data associated with weather conditions such as date, temp, heating degree, cooling degree, precipitation, snowfall and snow/ice depth. Archived data is provided for many legacy weather stations.

3.2 Data Pre-processing

This study involved the daily participation of citizens that connected to LinkNYC Kiosks and Wifi Metro Stations and used the application with the questions stated above during the period from 1 January 2017 to 31 December 2019. The collected dataset contains 1054 lines and a total of 23 features. But because data is taken from multiple sources which are in different formats, it is simply unrealistic to expect that the data will be perfect. Therefore, first of all, the following steps of data processing were done:

- Elimination of irrelevant variables: Some variables like the wifi status, tablet status, phone status, which is not relevant to the prediction, were deleted;
- Duplicate values: Some rows were duplicate data. We removed them to not give to data object an advantage or bias, when running machine learning algorithms;
- Handling of missing values: In the treatment of missing values, we replaced these values by the preceding value, due to the fact the data is captured sequentially. This method results in less introduction of variability in the dataset;

- Handling non-numerical data: Since DL models only accept numbers, we applied One Hot Encoder method to perform pre-processing in several features represented by strings;
- Target encoding: Since the target presents values in a certain way sorted from 1 to 5, a label encoding technique was used to normalize these values (thus transforming these values into classes 0 to 4);
- Splitting the dataset: We split the dataset into a 70:30 ratio. This means that you take 70% of the data (2 years) for training the model while leaving out the rest 30% (1 year);
- Cross-validation: 10-fold cross-validation to divide the model tests 10 times.

The data preprocessing transforms the data to bring it to such a state that the machine can easily parse it. In other words, the features of the data can be easily interpreted by ML algorithms. In this case, we wanted to study if the treated data was relevant to the prediction of physical well-being. Therefore, we used NN and dynamic regression models where the order of the treated data is quite relevant, although no shuffle has been done. In addition to pre-processing, other special precautions regarding the way data had to be processed were taken, which we will detail in the next subsection.

3.3 Building the Models

This step is the most important and most meticulous requirement of the entire research. With this, the aim of this work was to relate univariate and multivariate analysis in daily census in different environments (indoor and outdoor). In univariate time series dataset is generally provided as a single column of data, in this study, it's "census" column. On the other hand, a multivariate time series covers several variables such as census, temperature, heating degree, cooling degree, comfort, social interaction, physical, financial life, work, psychology, satisfaction that are recorded simultaneously over time.

In DL predictions and being a multiclass classification problem, the loss function is therefore categorical_crossentropy. Furthermore, in the final layer, a softmax activation function was used. Here we have to take into account the type of this activation function and the loss function, as the incorrect use of these can lead to false results. With the use of values in MinMaxScaler technique, the final step was validating and tuning the models. In these approaches, the objective was to experiment with some combinations in order to find a good fit. The number of layers, the number of neurons, the windows size, epochs, batch size, among other, in DL models, were tested together.

In the case of auto-regression, components are specified in the model as a parameter. The notations used by ARIMA and ARIMAX models are number of delayed observations, number of times that gross observations are differentiated, size of the moving media window and, besides these, the SARIMA and SARI-MAX models add the number of iterations for each seasonal period parameter.

3.4 Results

Since we want to classify the number of people some precautions have to be taken when we use DL (or NN models) and auto-regression models (or statistical models). Given that we are studying two datasets, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) errors were computed for each them. Essentially eight approaches are presented in Table 1.

Using the four NN models, besides building a predictive model that returns a minimization in error, we also adopt another data mining strategy based on the loss functions [16]. Basically, these two-fold approaches enable (i) presenting performance bounds of MAE, and (ii) demonstrating new properties of MAE that make it more appropriate than Mean Squared Error (MSE) as a loss function for Deep Neural Network (DNN).

Table 1. RMSE and MAE for Deep Learning and statistical models with univariate and multivariate time series.

		MTA Wi-Fi				LinkNYC Kiosks			
		Univariate		Multivariate		Univariate		Multivariate	
Model	Algorithm	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Deep learning	CNN	519.4	411.2	1145.7	970.4	659.3	532.1	1716.7	1407.4
	LSTM	969.5	742.3	1014.6	798.4	1375.4	752.9	1348.7	1143.3
	CNN-LSTM	1054.0	859.5	825.5	639.5	1173.0	968.4	1145.7	970.4
	Bi-Dir LSTM	32.3	24.6	131.4	101.9	1287.3	1104.8	256.9	203.7
Statistical	ARIMA	1111.6	899.4	-	-	1559.2	1201.6	-	-
	ARIMAX	-	-	1333.6	905.4	-	-	1617.0	1290.7
	SARIMA	1496.9	1186.1	-	-	1659.7	1388.9	-	-
	SARIMAX	-	-	1696.2	1211.0	-	-	1796.9	1441.0

Based on the above tables, we can draw two different perceptions concerning the experimental results. In this study, the whole experiment is carried out in two phases. The first phase of the experiment includes the eight models in the indoor environment dataset and then studying the outdoor environment dataset. Then, the models' performance is analyzed with metrics such as RMSE and MAE. Globally, these metrics show different performances between proposed types of models in study. In two datasets, the RMSE and MAE values are higher in autoregressive models than in DL models. However, we can find approximate values, for example, between the hybrid CNN-LSTM model and ARIMA models applied on MTA Wi-Fi dataset, using univariate time series. Or comparing accuracy between the LSTM and SARIMAX models applied on LinkNYC Kiosks dataset with all variables. But then we can find extreme values, in the case of MTA Wi-Fi dataset, between the Bi-Directional LSTM and SARIMA models using unique variable and, in LinkNYC Kiosks dataset, the Bi-Directional LSTM and ARIMAX models when applied in multivariate time

series. Although each metric has its own pros and cons, they are useful to address problems such as underfitting and overfitting which can lead to a poor performance on the final model despite the accuracy value. The quality assurance of results was only possible based on the loss functions.

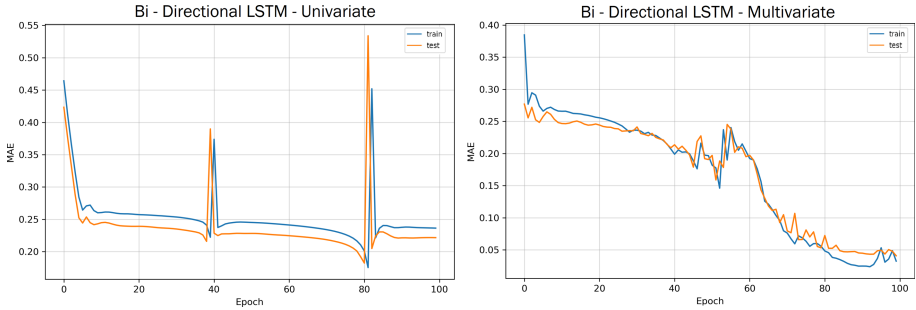


Fig. 2. Loss functions with lowest score based on MTA Wi-Fi dataset.

First of all, we choose the functions based on the number of variable (i.e., univariate and multivariate), and lowest score. In MTA Wi-Fi dataset, either with one or several variables, the Bi-Directional LSTM model presented the loss functions with lowest score. Figure 2 shows that, initially, the model has good performance, after 30 epochs it tends to converge, then it degrades. Taking Table 1 into account, CNN model for univariate model or hybrid CNN-LSTM model for multivariate also presents reasonable values and acceptable to be used for prediction and forecasting human mobility. They can be a good alternative for predictive modelling of human mobility.

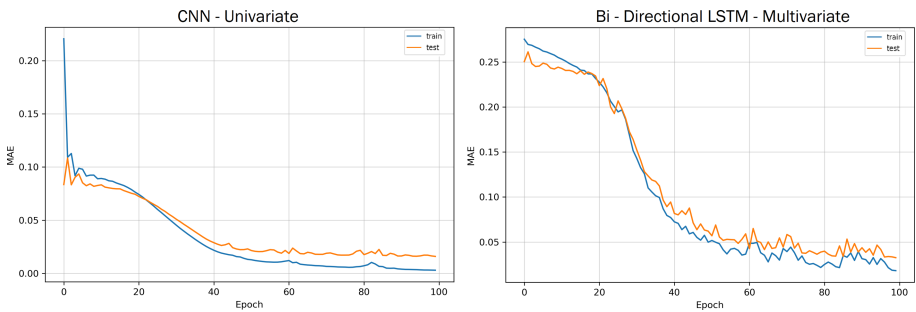


Fig. 3. Loss functions with lowest score based on LinkNYC Kiosks dataset.

As shown in Fig. 3, when LinkNYC Kiosks dataset only has a variable the lines of function in CNN model until 20 epoch seems to be converging, then it

tends to degrade. Whereas in the Bi-Directional LSTM model with multivariate (or multi variables) the lines of testing and training data never converge, the distance between them is decreasing over time. Additionally, we can see in Table 1 that the RMSE and MAE values in remaining models are worse than these models, making it hard to choose an alternative model.

On the other hand, we describe the forecasting performance of the statistical methods for a multi-step prediction task. The validation and consequently the final accuracy was obtained using the indoor and outdoor datasets. In particular, we consider 30-step-ahead forecasting, with a step equal to one day. We test the forecasting methods illustrated in Figs. 4 and 5 with each time series in our datasets.

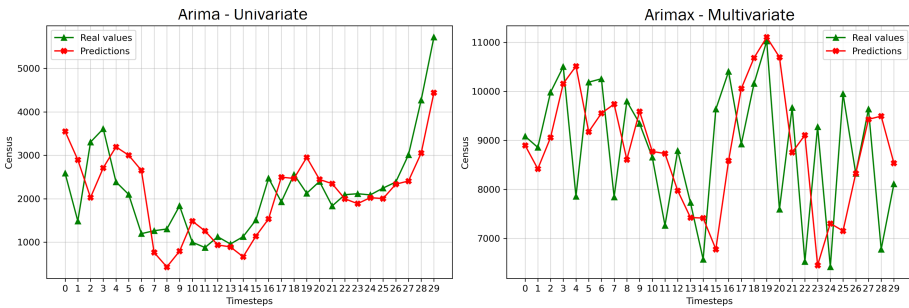


Fig. 4. Autoregression models for time series forecasting based on MTA Wi-Fi dataset.

As ours is a multi-step forecasting process, we also compute the forecasting error represented in Table 1. Based on them, when we applied Autoregression models on indoor dataset with a univariate, the lowest RMSE and MAE values obtained were 1111.6 and 899.4, but in multivariate values were 1333.6 and 905.4. This means that ARIMA and ARIMAX models presented the best results. In Fig. 4, ARIMA model (Univariate) predicted values closely match the actual values of Census. When the actual value changes direction, predicted value follows, which seems great at first sight. But in ARIMAX model (Multivariate), predicted values were worse. We can observe that predicted values didn't mimic the actual values.

In a bid to find a good model the same steps followed before were applied in the second approach presented in Fig. 5. Although, the ARIMA and ARIMAX models also present better results than SARIMA and SARIMAX models, if we compare with indoor environment dataset, globally, the RMSE and MAE values are worse. In other words, while ARIMA and ARIMAX have value pairs 1559.2 & 1201.6 and 1617.0 & 1290.7, respectively, the SARIMA is 1659.7 & 1388.9 and SARIMAX is 1796.9 & 1441.0, which means the first pair of statistical models presents a better performance.

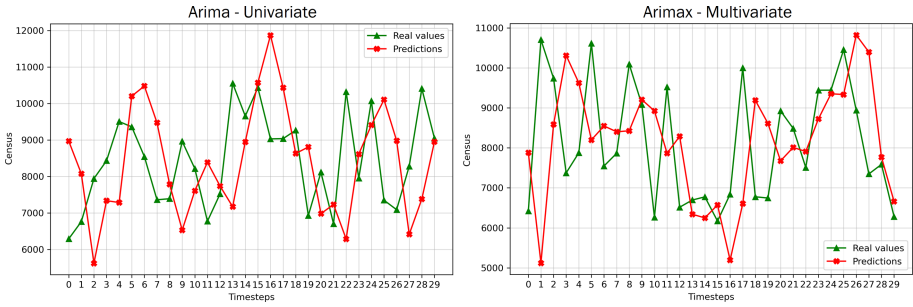


Fig. 5. Autoregression models for time series forecasting based on LinkNYC Kiosks dataset.

Figure 5 also compares predicted and actual census. We can observe that while the model outputs predicted values, they are not so close to actual values than occur in another dataset. But when it starts to generate values, the output almost resembles the sine wave. Later, in the last timestamp, values are similar.

4 Discussion

Something we can infer after the results is that the proposed DL techniques (especially Bi-Directional LSTM) may work better than statistical methods. In other words, experimental results of the proposed work show an improvement in the neural network over the statistical methods. Even changing the number of variables (i.e., change univariate to multivariate or vice-versa) and correct parameters, the performance of the neural network presents satisfactory results. Therefore, a neural network is fully modular. With Autoregressive Models, we forecast “only” on prior events, but these models are computationally intensive, more than NN models.

Models were trained to predict the Census of the next 30 days based on historical data. The census is a spatial-temporal popularity metric of human mobility. This metric captures the specifics of life within a human movements phenomena, and it is an empirical metric for people mobility of that particular area and time in the city. However, we can go further and based on the same datasets reach other interesting results. Adding other human mobility metrics such as displacement, perturbation and duration we can refine knowledge about people’s movements. As we mentioned, there are peaks of mobility, where there is variability in the density of people in an area of the city that may correspond to a smaller or greater collection of data in the interaction with the different infrastructures proposed in this paper. Thus, taking their effects into consideration in predicting human mobility will not only make it possible to improve the prediction accuracy but also many actions can be supported by the use of these metrics that may provide improvements to the planning in New York City.

5 Conclusions

In this article the study modelling and prediction were extended to several human mobility phenomena. It evaluates census using MTA Wi-Fi Locations and LinkNYC kiosks datasets. The experiments carried out have shown good results. Based on them, selected DL algorithms are more suitable, when compared to Autoregressive models. In addition, evaluating the RMSE and MAE results, enabled us to choose the best parameters. Consequently, they showed neural networks models provide better prediction accuracy than statical models.

In the future, unlike the data source presented in this work that requires a pre-connectivity to wi-fi, we hope to measure population using only device signals. They can give a better understanding of human mobility mainly based on census data and, consequently, stakeholders may be able to provide suitable responses to citizens (especially vulnerable ones), building and maintaining quality socially inclusive services and facilities. It means, planning and managing of pedestrian spaces should take into consideration the correct design of paths (also cycle paths), streets, common places, recognizing that the roads are both a social space and a space for mobility.

Acknowledgments. This work has been supported by FCT - Fundacao para a Ciencia e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020. It has also been supported by national funds through FCT – Fundação para a Ciência e Tecnologia through project UIDB/04728/2020.

References

1. Alartartseva, E., Barysheva, G.: Well-being: subjective and objective aspects. *Procedia - Soc. Behav. Sci.* **166**, 36–42 (2015). <https://doi.org/10.1016/j.sbspro.2014.12.479>, www.sciencedirect.com
2. Atkinson, S., et al.: Review team: What is Community Wellbeing? Technical report (2017)
3. Authority, M.T.: Transit Wireless Wifi: Product Reviews, Howtos & Buying Advice (2021). <https://transitwirelesswifi.com/>
4. Bartels, A.L., Peterson, S.J., Reina, C.S.: Understanding well-being at work: Development and validation of the Eudaimonic workplace well-being scale. *PLoS One* (2019). <https://doi.org/10.1371/journal.pone.0215957>
5. Blečić, I., Congiu, T., Fancello, G., Trunfio, G.A.: Planning and design support tools for walkability: a guide for Urban analysts (2020). <https://doi.org/10.3390/su12114405>
6. De Nadai, M., Cardoso, A., Lima, A., Lepri, B., Oliver, N.: Strategies and limitations in app usage and human mobility. *Sci. Rep.* **9**, 10935 (2019). <https://doi.org/10.1038/s41598-019-47493-x>
7. Department, M.R.E.: MTA Wi-Fi Locations (2021). <https://data.ny.gov/Transportation/MTA-Wi-Fi-Locations/pwa9-tmie>
8. Fadeev, A., Alhousseini, S., Belova, E.: Monitoring public transport demand using data from automated fare collection system (2018)
9. of Health, D.: What works well to improve wellbeing (2020). <http://whatworkswell.schoolfoodplan.com/>

10. Katevas, K., Hänsel, K., Clegg, R., Leontiadis, I., Haddadi, H., Tokarchuk, L.: Finding dory in the crowd: Detecting social interactions using multi-modal mobile sensing. In: SenSys-ML 2019 - Proceedings of the 1st Workshop on Machine Learning on Edge in Sensor Systems, Part of SenSys 2019 (2019)
11. Lawal, O., Nwegbu, C.: Movement and risk perception: evidence from spatial analysis of mobile phone-based mobility during the COVID-19 lockdown, Nigeria. *Geo-Journal*, 1–16 (2020). <https://doi.org/10.1007/s10708-020-10331-z>
12. Lee, K., Sener, I.N.: Emerging data for pedestrian and bicycle monitoring: sources and applications. *Transp. Res. Interdisciplinary Perspect.* **4**, 100095 (2020). <https://doi.org/10.1016/j.trip.2020.100095>
13. Loaiza-Monsalve, D., Riascos, A.P.: Human mobility in bike-sharing systems: structure of local and non-local dynamics. *PLoS One* **14**(3), e0213106 (2019)
14. NYC Department of Information Technology & Telecommunications: Find a Link (2021). <https://www.link.nyc/find-a-link.html>
15. Publishing, O.E.C.D.: OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth. OECD, June 2013
16. Parmar, R.: Common loss functions in machine learning (2018). <https://towardsdatascience.com/common-loss-functions-in-machine-learning-46af0fc4d23>
17. PBS: Public Broadcasting Service: Physical Well-Being and Motor Development (2019). <https://www.pbs.org/pre-school-u/pre-school-u-domains/physical-well-being-and-motor-development/>
18. Rath, T., Harter, J.: *The Economics of Wellbeing*. Gallup Press, New York (2010)
19. Ruggeri, K., Garcia-Garzon, E., Maguire, Á., Matz, S., Huppert, F.A.: Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries. *Health and Quality of Life Outcomes* (2020)
20. Sousa, D., Silva, F., Analide, C.: Learning user comfort and well-being through smart devices. In: Analide, C., Novais, P., Camacho, D., Yin, H. (eds.) *IDEAL 2020*. LNCS, vol. 12489, pp. 350–361. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62362-3_31
21. Thornton, F., et al.: Human mobility and environmental change: a survey of perceptions and policy direction. *Popul. Environ.* **40**(3), 239–256 (2018). <https://doi.org/10.1007/s11111-018-0309-3>
22. Trudel-Fitzgerald, C., Millstein, R.A., Von Hippel, C., Howe, C.J., Tomasso, L.P., Wagner, G.R., Vanderweele, T.J.: Psychological well-being as part of the public health debate? Insight into dimensions, interventions, and policy. *BMC Public Health* (2019). <https://doi.org/10.1186/s12889-019-8029-x>
23. Vanky, A.P., Verma, S.K., Courtney, T.K., Santi, P., Ratti, C.: Effect of weather on pedestrian trip count and duration: city-scale evaluations using mobile phone application data. *Preventive Medicine Reports* (2017)
24. Vink, P., Hallbeck, S.: Editorial: comfort and discomfort studies demonstrate the need for a new model (2012). <https://doi.org/10.1016/j.apergo.2011.06.001>
25. Woodward, K., Kanjo, E., Brown, D., McGinnity, T.M., Inkster, B., MacIntyre, D., Tsanas, T.: Beyond mobile apps: a survey of technologies for mental well-being. *IEEE Trans. Affect. Comput.* (2020). <https://doi.org/10.1109/TAFFC.2020.3015018>