Estimating trip attractors for non-commute travel using interpretable Machine Learning techniques on Open Data

A case study of public transport in Greater London

The-Huan Hoang
Supervised by Dr. Elsa Arcaute

A dissertation presented in partial fulfillment
of the requirements for the degree of
Master of Science in Urban Spatial Science

Bartlett Centre for Advanced Spatial Analytics (CASA)
University College London
United Kingdom
23 August 2024

Abstract

Abstract goes here...

Declaration

I, The-Huan Hoang, declare that the work presented in this dissertation is my own. Where claims, information or findings have been derived from other authors and sources, I confirm that this has herein been indicated in writing, attributing due credit. In the spirit of fostering reproducible research, all scripts used for analysis and their outputs are made available on GitHub at https://github.com/shaunhoang/casa-tfl-dissertation.

Acknowledgements

I am grateful for the guidance and insightful feedback from my supervisor, Dr. Elsa Arcaute, who has shaped the direction of my work from its inception. I would also like to acknowledge the invaluable support from the Public Transport Service Planning team at Transport for London. Their expert advice and real-world insights into the data science behind transport planning have greatly enriched my knowledge, as well as this research project.

Contents

1 Introduction **8**

2 Literature Review **9**
2.1 Urban amenities as attractors . . . 10
2.2 Transport accessibility as attractors . . . 11
2.3 Public transport and the walkable city . . . 11
2.4 Explainable Artificial Intelligence for spatial systems . . . 12
2.5 Open data and reproducibility . . . 13
2.6 Summary . . . 13

3 Methodology **14**
3.1 Defining the spatial unit of analysis . . . 15
3.2 Data Preprocessing . . . 16
3.2.1 Aggregating target variables . . . 16
3.2.2 Feature engineering . . . 18
3.2.3 Incorporating spatial lags as a feature . . . 23
3.3 Model selection and interpretation . . . 24

4 Results **27**
4.1 RQ1: Predicting attractiveness from spatial features . . . 27
4.2 RQ2: Identifying local attractors and hotspots . . . 32

5 Discussion **36**

6 Conclusion **39**

A Appendix **40**
A.1 Open data sources . . . 40
A.2 True to Data and True to Model in SHAP technique . . . 42
A.3 Supplemental Figures . . . 43

List of Figures

3.1 Overview of the feature engineering pipeline 14
3.2 Overview of model training and interpretation 15
3.3 Methodology to generate 16,890 isochrones as the spatial unit of analysis 16
3.4 Two example isochrones as spatial units of analysis located in Central London . . 16
3.5 Public transport arrivals on an average Saturday, spatially aggregated to the spatial units. *Top* Total, *Bottom* By time band. 17
3.6 Distribution of the Total arrivals as raw count (left) and log-transformed (right). Similar transformations are carried out for arrivals by each of the four time bands 17
3.7 Population, Amenity and Transport POI density by type (Source: UK Census 2021, OSM) 18
3.8 Normalised diversity indices for Amenity POIs, Transport POIs by mode and by route options. *Top*: Mapping top 10% values, *Bottom*: Distribution of diversity indices 20
3.9 Top 10% values of 3 diversity indices Borough: Hammersmith and Fulham 21
3.10 Network graphs of TfL Rail (left) and Bus (right) networks Constructed based on TfL API data 21
3.11 Mapping three centrality measures of spatial units on the Rail (left) and Bus (right) networks 22
3.12 LISA cluster map of Total arrivals showing local sptial autocorrelation Global Moran's I: 0.742 (High spatial autocorrelation) 24
3.13 Spatial vs. non-spatial nested cross-validation using 5-folds for hyperparameter tuning and performance evaluation. This technique helps ensure the training and test sets in both the inner and the outer folds are spatially clustered to avoid data leakage Source: Schratz et al. (2018) 25

4.1 Correlation matrix of the input features and target variable Total arrivals (log) . . 28
4.2 Observations with significant overprediction and underprediction (N=576 of 16,890) Fine-tuned XGBoost model, Total arrivals (log) 29
4.3 Summary of feature importance based on SHAP values. Red-Blue scale represents the feature value, and the x-axis represents the SHAP value, i.e., impact on final model prediction of Total arrivals (log) 30
4.4 Feature importance in predicting Total arrivals (log) Features clustered by correlation and ordered by mean absolute SHAP values 31
4.5 Local SHAP values for 2 spatial units Kew Gardens and Covent Garden 32
4.6 Mapping local SHAP values for 2 features: Outdoor leisure POIs and Entertainment leisure POIs. Target: Total arrivals (log) 33
4.7 Mapping local SHAP values for 2 features: Outdoor leisure POIs and Entertainment leisure POIs, weighted by number of arrivals 34

4.8 Raw vs. weighted SHAP for 2 features Outdoor leisure POIs and Entertainment leisure POIs Total arrivals (log) 35

5.1 Hot spots for destination groups: Amenity vs. Residential vs. Transport Midday vs Late 36
5.2 Identifying high activity areas in Wood Green 37

A.1 Map of the spatial k-fold cross-validation splits for hyperparameter tuning and performance assessment (k=10) 43
A.2 Summary of feature importance based on SHAP values of a sample. Red-Blue scale represents the feature value, and the x-axis represents the SHAP value, i.e., impact on final model prediction of arrivals in each time band (log) 43

List of Tables

3.1 Summary of the 25 main features engineered for the model 23
3.2 Mapping spatial lag features to predict corresponding target variables 24

4.1 Comparing model R^2 results of 3 models predicting Total arrivals (raw) and Total arrivals (log) using spatial k-fold cross-validation for tuning and evaluation 27
4.2 Finetuned XGBoost model evaluation results for All target variables (log) using nested spatial k-fold cross-validation for tuning and evaluation 28

A.1 Summary of data used in the analysis and their sources 40
A.2 Summary of Amenity and Transport POI types used in the analysis 42

1. Introduction

[to be completed]

2. Literature Review

The task of estimating trip attraction constitutes an integral component of the demand forecast framework, termed the Four-Step Model (FSM), which has been in ubiquitous use since the 1950s. More specifically, the FSM begins with the Trip Generation step, where the number of trips generated by each traffic analysis zone is estimated, and Trip Distribution allocates these trips to their destination zones before proceeding to mode choice determination and route assignment. Planners accomplish the task of Trip Distribution using a variety of spatial interaction models, most prominent among which is the classic *Gravity Model* and its derivatives inspired by Newton's law of universal gravitation (Erlander and Stewart, 1990). Originally formulated to estimate flows among cities in a region, it has been adapted to intraurban mobility and suggests that the number of trips between two locations is proportional to the difference in *mass* of each location and inversely proportional to the distance between them.

However, it is worth noting that the definition of *mass* in any gravity-based spatial interaction model has seen many interpretations depending on the types of travel demand of research interest. For example, a spatial interaction model for commuting trip forecasts may consider the emissivity of residential areas as origins and the attractiveness of employment centres as destinations. In fact, the estimation of trip attraction in transport planning literature has predominantly revolved around commutes (i.e., primary travel demand) from residential zones to zones where employment centres or education institutions are located. On the one hand, this choice is a practical one first and foremost, as commuting trips represent a large share of regular intracity mobility flows, whose high volume within short peak periods has important implications on road capacity management or public transit operations, among others. Due to its importance, the calibration of the spatial interaction model for commuting flows is further aided by nationwide flow datasets, such as those present in the UK Census [1], which provides detailed and explicit Origin-Destination data to serve as ground truth. In the absence of this declared data, researchers have also turned to spatial features to calibrate commute trip distribution, as explored by Yang et al., 2014.

On the other hand, a plethora of potential trips generated for other purposes is often underresearched, partially due to the complexity of non-commute trip purpose identification, classification and quantification: Non-commute trips can be defined as trips that are not carried out to fulfil time-bound and regular work or education obligations and can include trips for shopping, leisure, and social activities. Therefore, as opposed to commuting trips, Non-commute trips are highly irregular both spatially and temporally, i.e., they can vary in terms of distance, duration, and mode of transport from one to another, even for the same destination between different days. Nevertheless, Non-commute trips are an important component of urban travel behaviour and can have a significant impact on the overall system. Advances in the field have brought about im-

[1] An interactive visualisation of the UK Census 2021 origin-destination data can be found at https://www.ons.gov.uk/visualisations/censusorigindestination/

provements to methodologies such as travel diary surveys[2], which collect data on personal and household travel behaviour, including Non-commute trips, to allow for comprehensive planning. By design, however, they are limited in scale.

Overall, this reveals a gap in our understanding of the comprehensive intracity travel behaviour on an aggregate scale. With the number of trips for shopping, leisure, and social activities increasing in recent years and the number of work-related trips decreasing as flexible work arrangements become more prevalent (Wöhner, 2022), with recognised health benefits to workers (MacLeod, Cole, and Musselwhite, 2022), it is important to understand the factors that influence non-commute trip attraction and to incorporate them into urban and transport planning models. Within the scope of this research, we hypothesise that a given area's non-commute trip attractiveness *mass* is associated with the built environments, which include their amenity and public transport connectivity profiles.

2.1 Urban amenities as attractors

Cervero and Kockelman (1997) was one of the foundational works to provide an econometric estimation of how three groups of factors pertaining to the built environments—namely Density, Diversity and Design—are associated with travel demand to a traffic analysis zone for different purposes (work/non-work) and with different mode choices (private vehicles/public transport). This study is novel for demonstrating that changes in the built environment—such as density, diversity, and walkability have a tangible impact on travel behaviour, especially non-work one, which supports the idea that designing more compact, mixed-use, pedestrian-friendly neighbourhoods can significantly influence whereto and how people choose to travel. The components of density and diversity refer to the number and the variety of destinations available in a given area, respectively, while the design component refers to the physical layout of the area, including the presence of sidewalks, bike lanes, and other pedestrian-friendly infrastructure.

There have been many recent attempts to link the presence of urban amenities to trip attraction and calibrate how we define urban attractors. Beyond the pure correlation observed between the density of points of interest (POIs) in a region and intensity of trips made to a given area (Melikov et al., 2021), the typology of the POIs differs between urban clusters that see distinct patterns of inflow trips made. For example, after Aaqib Javed et al. (2020) have clustered urban areas in one city into Global (e.g., airports), Downtown (d.g., central business district) and Residential attractors based on the volume, spatial dispersion and distance of the trips made using origin-destination information extracted from call data, they found that the differences in the types of POIs in each of these clusters were also statistically significant. The typology of amenity POIs explored in this study included not only Retail, as is common when thinking about the main purpose for non-commute trips, but also hospitals, public services, restaurants, religious institutions, and other categories. This suggests that the type of urban amenities available in an area can influence the volume of trips made to that area and that different types of amenities may attract different types of trip purposes.

Lastly, we must also consider the amenity diversity factor, apart from the typology of the amenities themselves, in the amenity profile of areas as a potential attractor. More specifically, for non-commute trips where locations are not necessarily prescribed or obligated, and the destination choice can be influenced by personal preferences and circumstantial convenience, people may be

[2]An example of these is the UK National Travel Survey, accessible at https://www.gov.uk/government/collections/national-travel-survey-statistics

more likely to visit more vibrant areas that offer a variety of services and activities where multiple tasks could be carried out, as opposed to areas where one type of destinations is overrepresented. This subjective perception of the quality of urban areas is consistent with the concept of 15-minute cities, which purports that mixed-use areas with a diversity of amenities within a 15-minute walking distance are more attractive areas to live in—and by extension to visit—than single-use ones. (Khavarian-Garmsir et al., 2023)

2.2 Transport accessibility as attractors

If the density and diversity of amenities were the only explanatory factors for non-commute trip demand, it would be tempting to conclude that having the amenities alone is the sole determinant of why people go to one place and not another for their diverse needs. In reality, neighbourhoods do not exist in a vacuum but are intricately connected to one another by the transportation network. In simple terms, the consideration of distance as friction in the classic Gravity Model attests to the fact that the more accessible a destination is, the more likely it is to attract travel demand, especially trips with higher elasticity of choice such as non-commute. Public transport accessibility has been shown to contribute to the growth of intercity importance of certain areas as opposed to others and preserve travel demand variability over time (Zhong et al., 2015), which means that it reinforces the prominence of well-connected areas as urban attractors, presumably with similar density and diversity of destination points of interest.

Jayasinghe, Sano, and Rattanaporn (2017) went one step further to purport that the network centrality property of an urban area is enough to predict travel demand to a high degree of accuracy without the need to consider other attractors such as employment centres and amenities, thanks to its high correlation. More specifically, the study looked at four centrality measures: Connectivity, Choice, Global Integration and Local Integration, which are analogous to the concept of degree centrality, betweenness centrality, closeness centrality, and closeness centrality in a confined radius, respectively. Among them, Global Integration (closeness centrality) was found to be the most significant predictor, suggesting that the role of the transport network in shaping urban travel demand at the aggregate scale is paramount, perhaps just as important as socio-economic factors (Convery and Williams, 2019). However, this study using car trip data is not representative for all mobility patterns especially when public transport and walking are considered. Therefore, it is logical to investigate the role of network centrality in conjunction with other variables.

2.3 Public transport and the walkable city

Walkability contributes to the quality of life in cities, with the "15-minute city" being one of the most prominent urban planning concepts that garnered attention in recent years. A resident who lives in a neighbourhood that has all the amenities they need within a 15-minute walk or cycle can ideally reduce their reliance on cars and public transport to a minimum. However, this does not mean that the role of public transport is diminished within the 15-minute city vision. On the contrary, public transport is an essential component of a sustainable system, which provides an efficient and affordable way for people to travel longer distances in the city, effectively agglomerating a larger urban system made up of many 15-minute cities. In other words, public transport is the backbone of the 15-minute city, connecting neighbourhoods and enabling people to travel further than they can on foot or by bike. This, in turn, has multiple implications for the local communities:

- From the lens of urban planning and traffic management, the higher the volume of public transport trips made to an area, the higher the amount of pedestrian traffic it receives on top of the local population who patronises their local amenities. Understanding the trip attraction factors can help prioritise investments in public transport capacity planning, pedestrian or cycling infrastructures and services for these individuals, which can lead to a virtuous cycle of increased active travel and reduced car dependency.
- From the lens of fomenting local economies, increased pedestrian traffic injected by public transport can also be seen as an additional source of footfall, invaluable for businesses with storefronts. Findings on the economic benefits of walking and cycling done by Transport for London[3] showed that retail rents increased by 7%, and office space rents increased by 4% in areas that encouraged non-motorised access, suggesting that the street improvements translated into much more desirable spaces.

It is also worth mentioning that footfall estimation itself is a prolific field of research within retail analytics, which counts on a variety of technologies such as WiFi tracking, video analytics, and mobile phone data to estimate the number of people who pass by a given location. The insights derived from footfall data can be used to inform decisions on store location, opening hours, staffing levels, and marketing strategies, among others. Uncovering patterns of public transport trip attractiveness, which in turn is an indicator of non-local pedestrian inflows, could be used to complement traditional footfall data and provide a more comprehensive retail footfall estimate.

2.4 Explainable Artificial Intelligence for spatial systems

Up to this point, our visited literature has shown that both urban amenities and transport accessibility are important factors in determining the attractiveness of an area to non-commute trips, albeit in isolation. In reality, the interaction between these two factors has not been well studied in conjunction. In the interest of addressing the research questions, we are presented with a dilemma of interpretability versus prediction accuracy in the choice of modelling techniques.

Spatial statistical methods, commonly used in policy analysis, can help us quantify and interpret the relationship between the dependent variables and independent variables in the form of explicit coefficients. However, even spatially aware methods such as Spatial Lag Model (SLM) or Geographically Weighted Regression (GWR) are limited in their assumptions of linearity and issues with multicollinearity (Wheeler and Tiefelsdorf, 2005), in order to provide robust interpretations.

In parallel, machine learning methods have gained popularity in the field of urban mobility research due to their ability to capture complex, non-linear relationships between spatial features with high prediction accuracy. However, the lack of interpretability of machine learning models has been a major drawback, especially in the context of urban planning, where the ability to understand the underlying factors that drive the model's predictions is crucial for decision-making. Among many opportunities to develop Machine Learning explanation techniques for causal inference in urban mobility is the need to evaluate causal effects of input features and derive actionable insights (Xin et al., 2022).

To address the interpretability challenge, Machine Learning interpretation techniques like **SHAP**. SHapley Additive exPlanations, is a unified approach to explain the output of any machine learning model—introduced by the seminal work of Lundberg and Lee (2017).based on

[3]https://tfl.gov.uk/corporate/publications-and-reports/economic-benefits-of-walking-and-cycling

cooperative game theory of Lloyd Shapley—, and aim to provide detailed attributions of input features to individual predictions made. Its introduction kick-started a growing body of literature on interpretable machine learning with spatial systems for regression with XGBoost (Li, 2022) or classification with Light GBM (Louhichi et al., 2023). One of the most notable recent examples of SHAP for urban mobility in action was the development of the Deep Gravity Model introduced by Simini et al., 2021, which used SHAP to not only explain how the predictions of the deep learning model for spatial interaction were influenced by the input spatial features such as land-use mix, points of interest in the origin and destination zones.

As we approach this research with an eye on modelling the real-world phenomenon and explaining the contributing factors, instead of GWR or SLM, we will use a machine learning model to predict non-commute trip attraction in urban areas to a high degree of accuracy, and then leverage SHAP to explain the predictions in an attempt to quantify the marginal importance and contribution of amenities and transport connectivity as attractors to an area's trip attractiveness.

2.5 Open data and reproducibility

Many recent advances in the field of urban mobility research have been made possible by the availability of proprietary datasets owned by private companies, such as mobile phone data, GPS data, and social media data. Although the high data granularity has opened doors to new research opportunities and the development of new advanced techniques, its use also raises concerns about data privacy, data ownership, and data sharing. Moreover, it heightens the barrier to entry for independent researchers who would like to replicate the studies for their cities or regions. In contrast, open data not only democratises independent and participatory research in the public interest but also breaks down many reproducibility barriers when applying established methodologies to other localities (Yadav et al., 2017).

On one side, we have readily available open datasets maintained by local or federal government agencies in many countries. On the other, we have global platforms like **OpenStreetMap** and **Overture Maps Foundation** that provide a wealth of spatial data for cities around the world. The use of open data to encourage reproducibility will be a key focus of this research to ensure that the findings can be replicated, unburdened by commercial data availability and access barriers. It also enables flexible preprocessing of the amenity or transportation features to model the intracity mobility behaviour that is pertinent to the locality's needs while adhering to the same model training and analysis methodology. Adhering to the spirit of openness and future reproducibility across different urban contexts, all of the datasets used in this research will be entirely sourced from open data hubs and platforms. The associated scripts used for data processing, model training and prediction analytics are available on GitHub.

2.6 Summary

To address the research questions posed and informed by past literature, we set out to fine-tune and select a performant model to predict non-commute trip attractiveness, where we hypothesise that the built environment factors including urban amenities and transport connectivity, are important trip attractors. The use of a machine learning model in conjunction with explainable machine learning techniques such as SHAP, allows us to capture complex, non-linear relationships, as well as analyse contributions of input features to the predictions made, in order to uncover underlying mobility patterns.

3. Methodology

In this section, we will first describe the methodology used to preprocess data from various sources to engineer the features for the predictive model as informed by past literature. Second, we will outline the training model selection and evaluation process. Finally, we will discuss the methodology used to interpret the model predictions using SHAP techniques to answer the research questions. For ease of reference, the data preprocessing and feature engineering pipeline is summarised in Figure 3.1, and an overview of the model training and prediction interpretation methodology can be found in Figure 3.2. More detailed information on the data sources and the feature engineering process can be found in the Data Sources section of the Appendix.

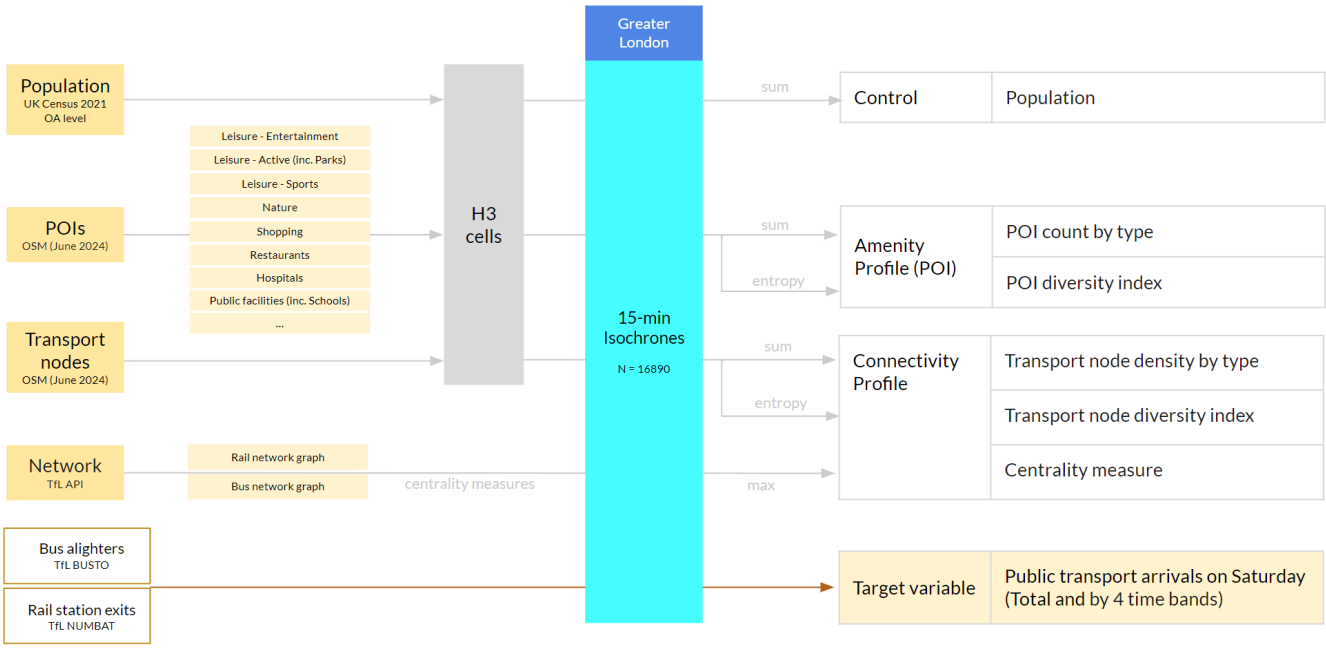

Figure 3.1: Overview of the feature engineering pipeline

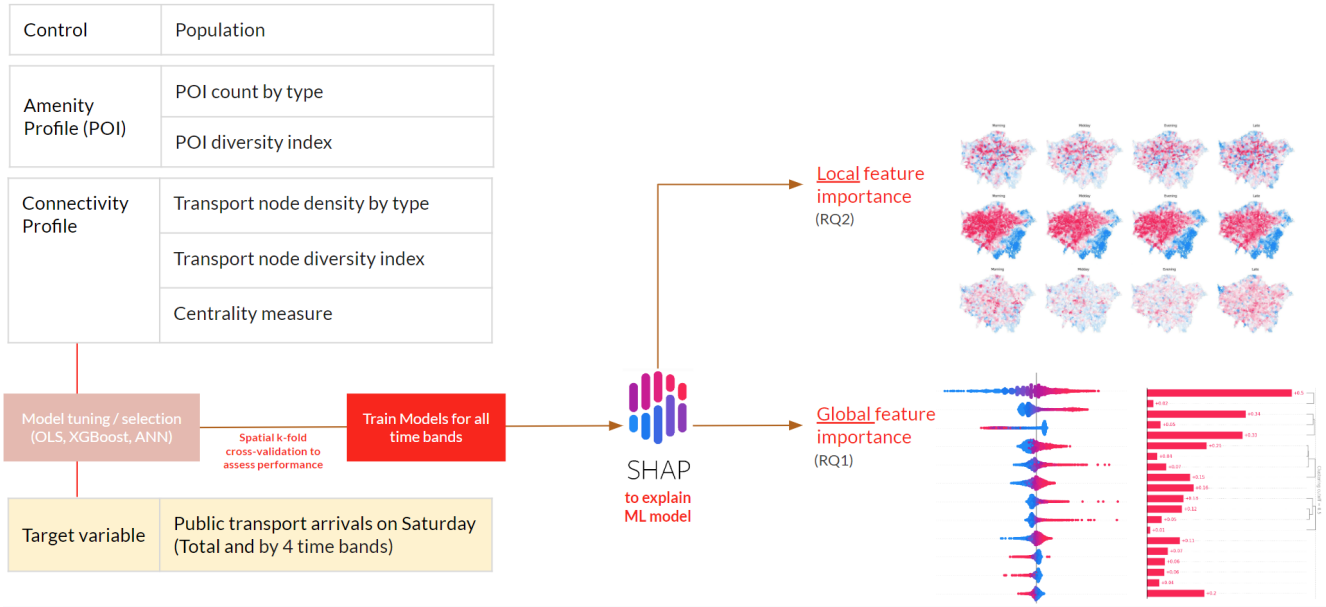

Figure 3.2: Overview of model training and interpretation

3.1 Defining the spatial unit of analysis

As features will be extracted and engineered from different data sources, the spatial unit of analysis—at which features are aggregated and the model is trained—is an important consideration. We will use 15-minute walking isochrones—i.e., the area that can be reached within a 15-minute walk from a point. They were generated based on initial seeds of 16,890 points distributed evenly across the Greater London region, using OpenStreetMap Pedestrian Network API (osmnx). The choice of using isochrones as the spatial unit of analysis for our model is motivated by the following reasons:

- First, the spatial unit should reflect how individuals interact with transport nodes and amenity POIs that are pedestrian-accessible, whereas Euclidean distance proximity overlooks natural barriers, such as rivers or private estates.
- Second, human mobility in cities is not bound by administrative boundaries such as Output Areas or boroughs, and the use of overlapping isochrones allows us to capture the spatial heterogeneity of the input features and the target variables with high granularity.
- Third, the 15-minute specification of the isochrones aims to be aligned with the 15-minute city concept that characterises desirable intra-neighbourhood mobility threshold.
- Lastly, the density of the spatial units is a compromise between the granularity and the computational resources required to process the data. The initial seeds of 16,890 points are centroids of H3 cells at resolution 9 needed to cover the Greater London area.

The detailed methodology to generate the isochrones is illustrated in Figure 3.3, and two example isochrones in Central London can be seen in Figure 3.4, showing how these spatial units may vary in shapes and sizes. In the next section, we will describe the data transformation and feature engineering process in more detail. Note that the terms spatial units of analysis or isochrones may be used interchangeably from this point on in our discussion.

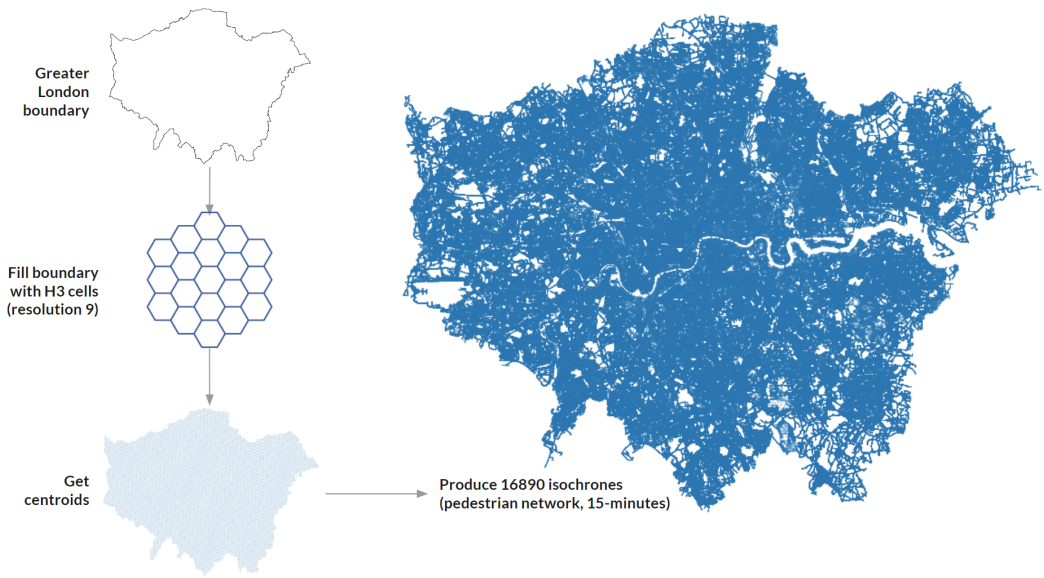

Figure 3.3: Methodology to generate 16,890 isochrones as the spatial unit of analysis

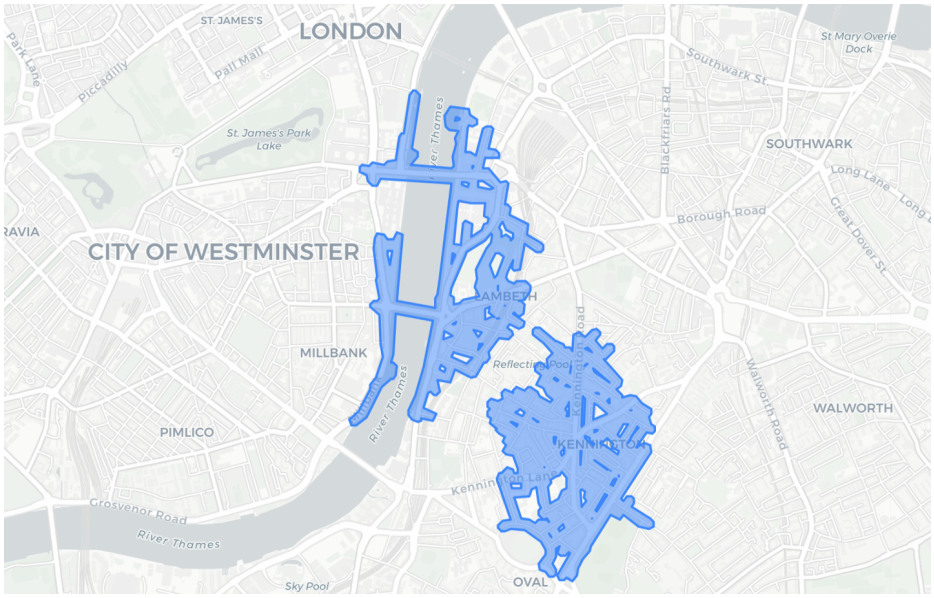

Figure 3.4: Two example isochrones as spatial units of analysis located in Central London

3.2 Data Preprocessing

3.2.1 Aggregating target variables

Total arrivals, i.e., the average numbers of passengers who exit a station or alight from a bus within a spatial unit on a Saturday, are the target variables for our model as a proxy for trip attractiveness. The aggregation results are illustrated in Figure 3.5, which shows the total figures as well as the figures aggregated into the four time bands. The total figures show a clear concentration of passenger arrivals in the Central London area, with the highest number of passengers arriving at transport nodes, while the time band figures show a similar pattern, with the highest number of passengers arriving during the Midday and Evening time band (10:00-19:00 and 19:00-22:00, respectively). Meanwhile, the Morning and Late time bands show lower numbers of

passengers using the system due to early hours in the day on weekends and sparser night services, respectively. Lastly, since the target variable has a highly right-skewed distribution, we opt for a log transformation (shown in Figure 3.6) to the target variable to normalise the distribution and improve the model's performance, which will be corroborated in the Results section.

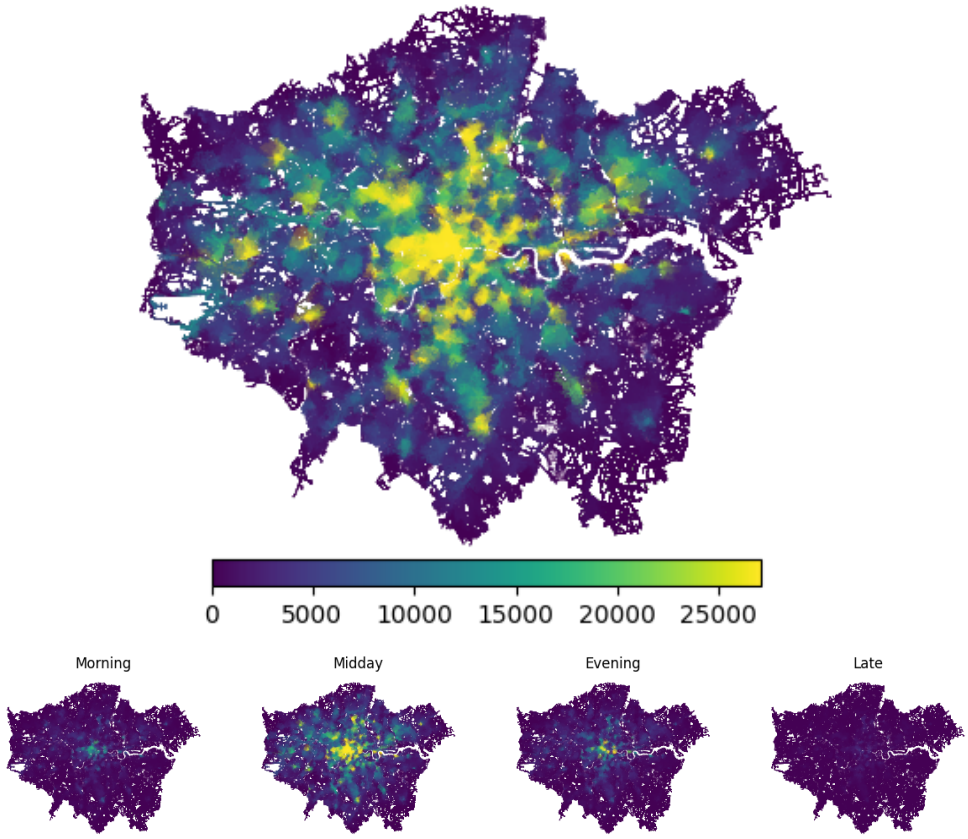

Figure 3.5: Public transport arrivals on an average Saturday, spatially aggregated to the spatial units. *Top* Total, *Bottom* By time band.

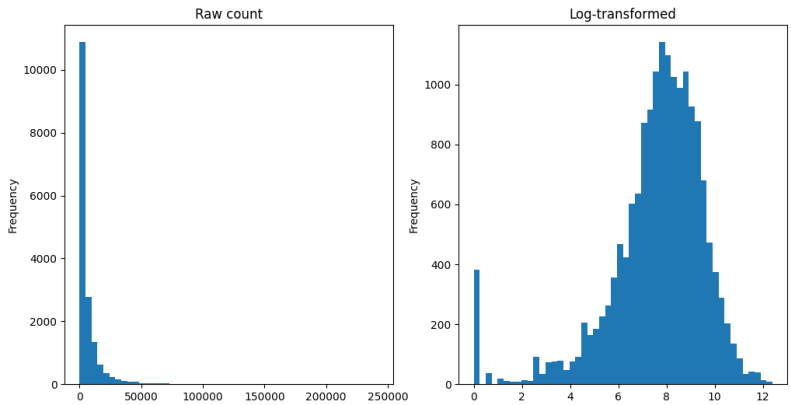

Figure 3.6: Distribution of the Total arrivals as raw count (left) and log-transformed (right). Similar transformations are carried out for arrivals by each of the four time bands

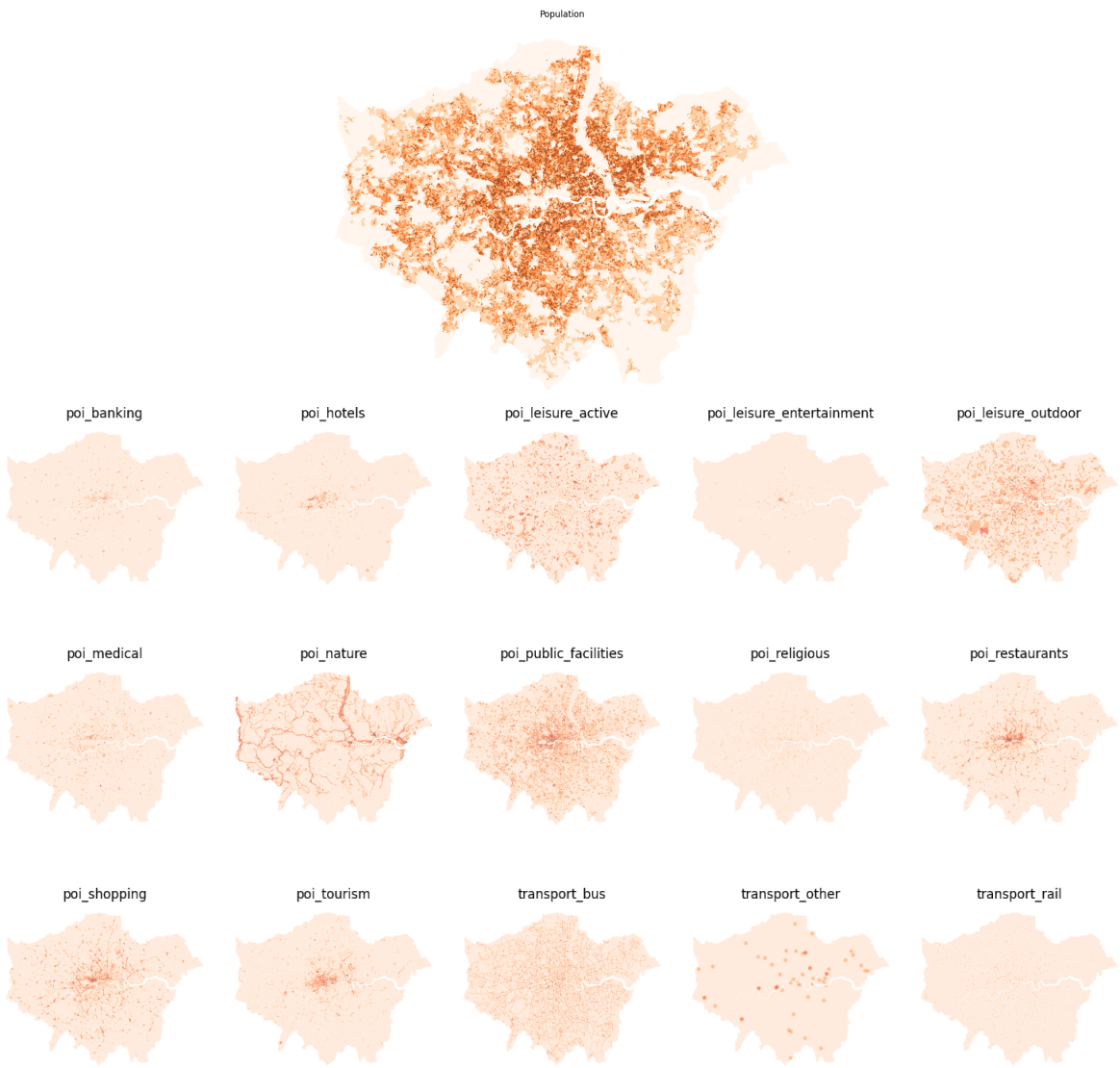

Figure 3.7: Population, Amenity and Transport POI density by type
(Source: UK Census 2021, OSM)

3.2.2 Feature engineering

Density-based features

The density of POIs is one of the first feature types of interest for our model, which represents the intensity of amenities or transport nodes of each type in an area that we hypothesise would attract visitors as destinations. Based on Cervero and Kockelman (1997)'s work, we expect a positive correlation between the density of POIs of all types to the total arrivals since an area with a high concentration of retail is expected to be more attractive than a similar-sized area with a low concentration of retail if the non-commute trip is for shopping purposes. Similarly, an area with a high concentration of transport facilities, such as stations and bus stops, is expected to accommodate a higher inflow volume of visitors. All density-related features are formulated by aggregating their counts for each spatial unit by type, resulting in a set of **16 features** (See Figure 3.7)

Diversity-based features

Diversity of Amenity POIs according to Cervero and Kockelman (1997) is also an important factor that may influence the trip attractiveness of an area for certain types of non-commute travel demand. The implication is that between two neighbourhoods, all else being equal, the area that is more diverse in the types of amenities it provides will be more attractive as a destination on average because of their vibrancy and the possibility it can afford visitors to accomplish different tasks in one trip.

What was left under-explored by literature is the diversity of public transport facilities present in a spatial unit, which may be an important trip attractor as it is a better proxy for transit accessibility than transport node density alone. Moreover, transport diversity can also capture the modal interchange behaviour of passengers whereby they alight in a certain area to switch between different means in order to reach their final destination elsewhere. As a hypothetical example, between two areas with a similar amenity density and diversity profile, the area that can be reached with multiple transit modes and routes should be more attractive to visitors, either because it makes that area more easily accessible or because it provides opportunities for passengers to switch between different modes of transport. The latter is of critical importance. Since there is no origin-destination matrix available for the study area, its inclusion potentially helps the model disentangle the impact of modal interchange behaviour on the total arrivals, as opposed to the trip attraction factor of the area itself as a destination.

Cervero and Kockelman (1997) conceptualised diversity in the form of *entropy*. A concept originated in information theory, entropy is the average amount of information conveyed by an event when considering all possible outcomes. Batty et al. (2014) further adapted it to convey complexity in a spatial system. Interestingly, they also considered the increase of spatial entropy to be a trade-off between the density of amenity types and the unique types of amenity present, which stands in competition with other density-based features created. Therefore, the inclusion of entropy as a measure of POI diversity could help us visualise this trade-off as the model is trained. We intend to replicate the approach used by both of the cited works above and adopt the formulation of entropy introduced by Shannon (1948) to calculate the diversity index of amenity POIs in each spatial unit as follows:

$$H = -\sum_{i=1}^{n} p_i \ln p_i \tag{3.1}$$

where p_i is the proportion of the i-th POI class in the total number of POIs in the spatial unit, and n is the total number of POI classes being considered. The entropy index ranges from 0 to $\ln n$, where 0 indicates no diversity and $\ln n$ indicates maximum diversity, which is then normalised to the 0-1 range. Based on this definition, we will engineer **3 features** related to the diversity of amenity POIs and transport POIs, further explained as follows:

- *Amenity POI*: 12 amenity POI classes are considered for the calculation of the diversity index for each spatial unit (see Table A.2). Looking at the 10% highest values of the diversity index in Figure 3.8, we can see that the most diverse areas are located throughout the Greater London area, which urban studies literature links to activity centres or neighbourhoods. As a whole, the distribution of the diversity index is right-skewed, indicating a pluricentric urban structure in Greater London.
- *Transport POI diversity by mode*: 3 transport POI classes (bus stops, rail stations, other hubs) are considered for the calculation of the diversity index for each spatial unit (see Table A.2). The top 10% values of the diversity index are mapped in Figure 3.8, showing that

the most diverse areas are located around important transport intermodal hubs, such as national rail stations, airports, and ferry piers.

- *Transport POI diversity by route option*: A more granular transport diversity index is calculated for the transport POIs, considering the different bus routes and train services the transport POIs in each spatial unit serve. The top 10% values of the diversity index are mapped in Figure 3.8, showing that the most transport-option diverse areas are not only intermodal hubs but also areas around arteries where many bus routes or rail services converge.

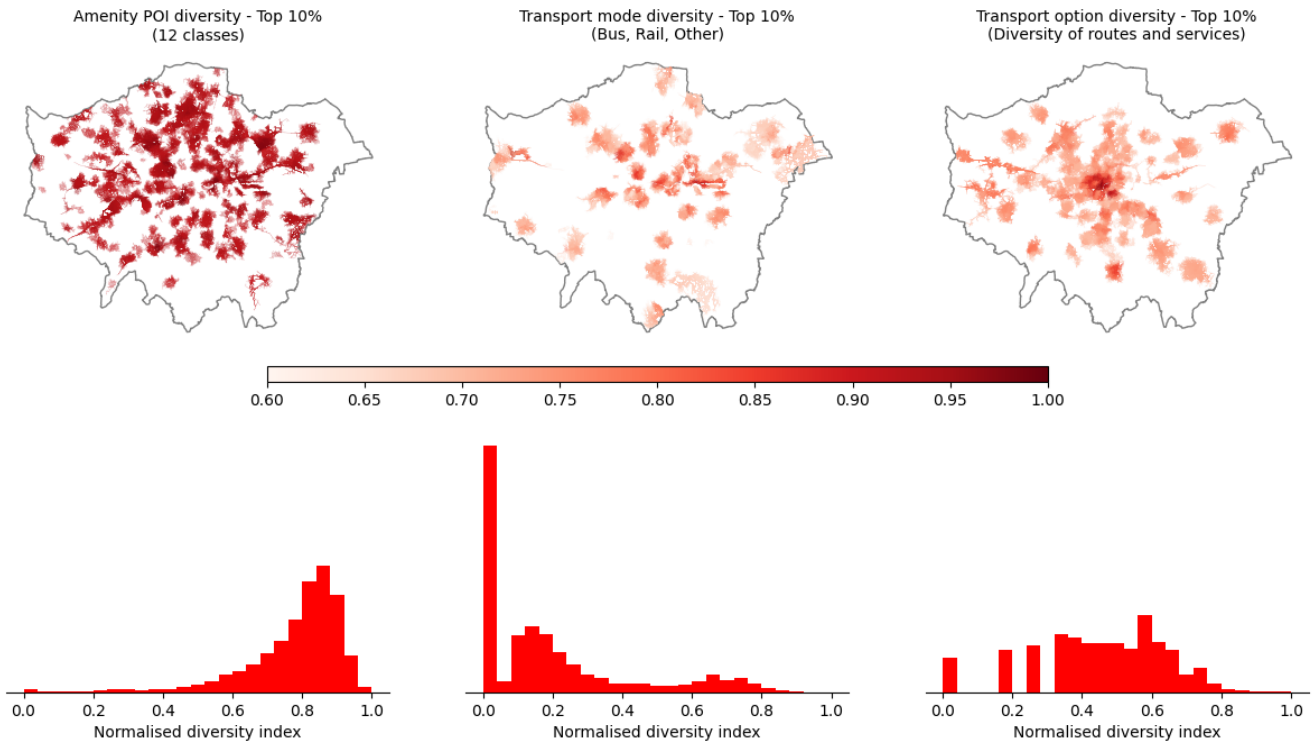

Figure 3.8: Normalised diversity indices for Amenity POIs, Transport POIs by mode and by route options. *Top*: Mapping top 10% values, *Bottom*: Distribution of diversity indices

The inclusion of both transport diversity indices aims to provide complementary information to the model, as they capture different factors to explain why an area receives more or less arriving passengers. High transport mode diversity indicates a high possibility that the area facilitates modal interchange for onward travel (such as long-distance intercity or cross-city trips), whereas high transport route diversity, in addition to that, indicates a high possibility that the area is locally well-connected with good frequency. Zooming into a specific area, such as the Hammersmith and Fulham borough, we can see that there is a slight divergence between the two transport diversity indices (Figure 3.9). The Shepherd's Bus area has a higher transport mode diversity index thanks to the existence of national rail services (Southern Railways) beyond Tube and buses, while the area around Earl's Court has a higher transport option diversity index due to the presence of multiple local bus routes converging in the area.

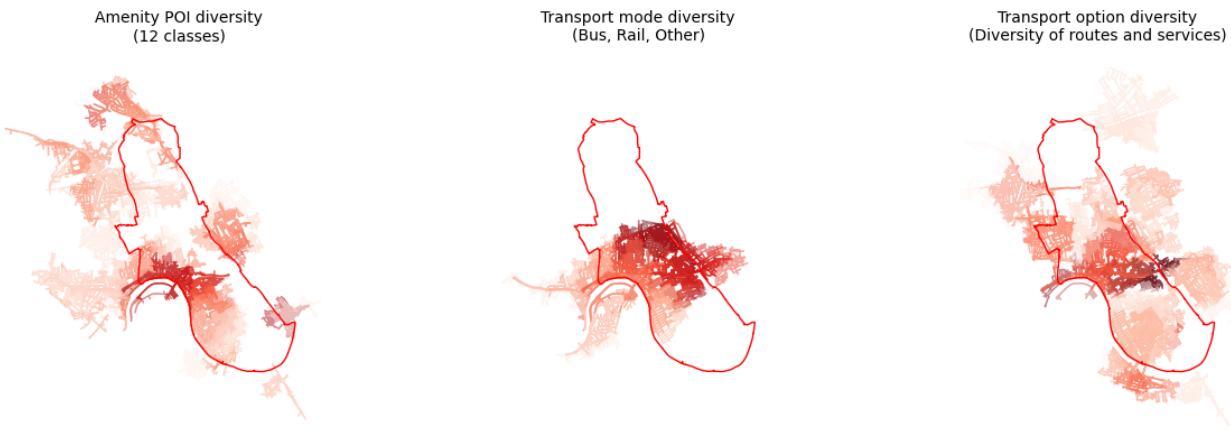

Figure 3.9: Top 10% values of 3 diversity indices
Borough: Hammersmith and Fulham

Centrality-based features

Whereas the public transport diversity index, previously discussed, captures the variety of transport options available in an area, it obscures one aspect of accessibility in relation to other areas in London using the public transport network. Centrality-based features add yet another layer of information to the model, capturing the importance of a bus stop or a station in relation to the whole network. For example, between two hypothetical areas with similar diversity of transport options (e.g., both are served by 5 bus routes), the area with a higher centrality measure may be more attractive to visitors because it is more 'accessible' from other areas on average.

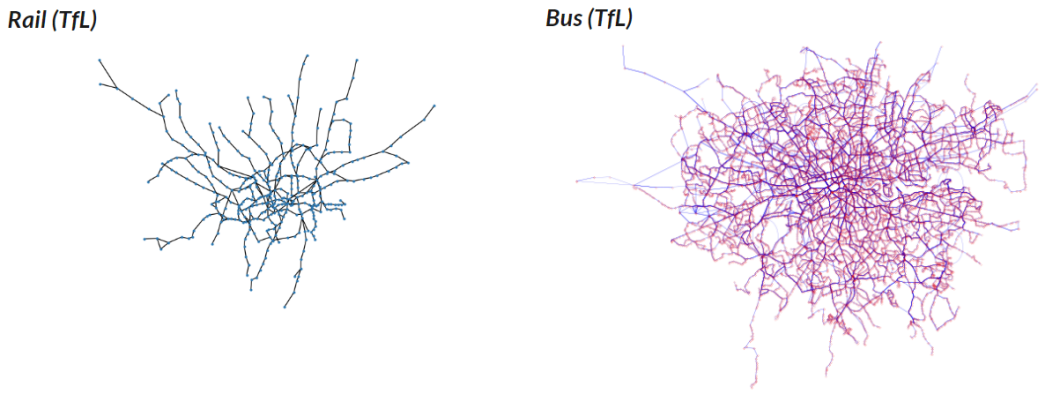

Figure 3.10: Network graphs of TfL Rail (left) and Bus (right) networks
Constructed based on TfL API data

To achieve that, first, we construct the network graph of the TfL Bus and the TfL Tube and Rail networks from open data (See Figure 3.10), with nodes representing the stops or stations, and the edges representing routes or lines. Due to the lack of data related to travel time and scheduling, we abstain from combining the two networks into a multi-layered network, which would have allowed us to consider the whole public transport network as one connected component. Instead, we will calculate the following centrality measures for all the nodes in the Rail and Bus networks separately, resulting in a set of 6 node attributes:

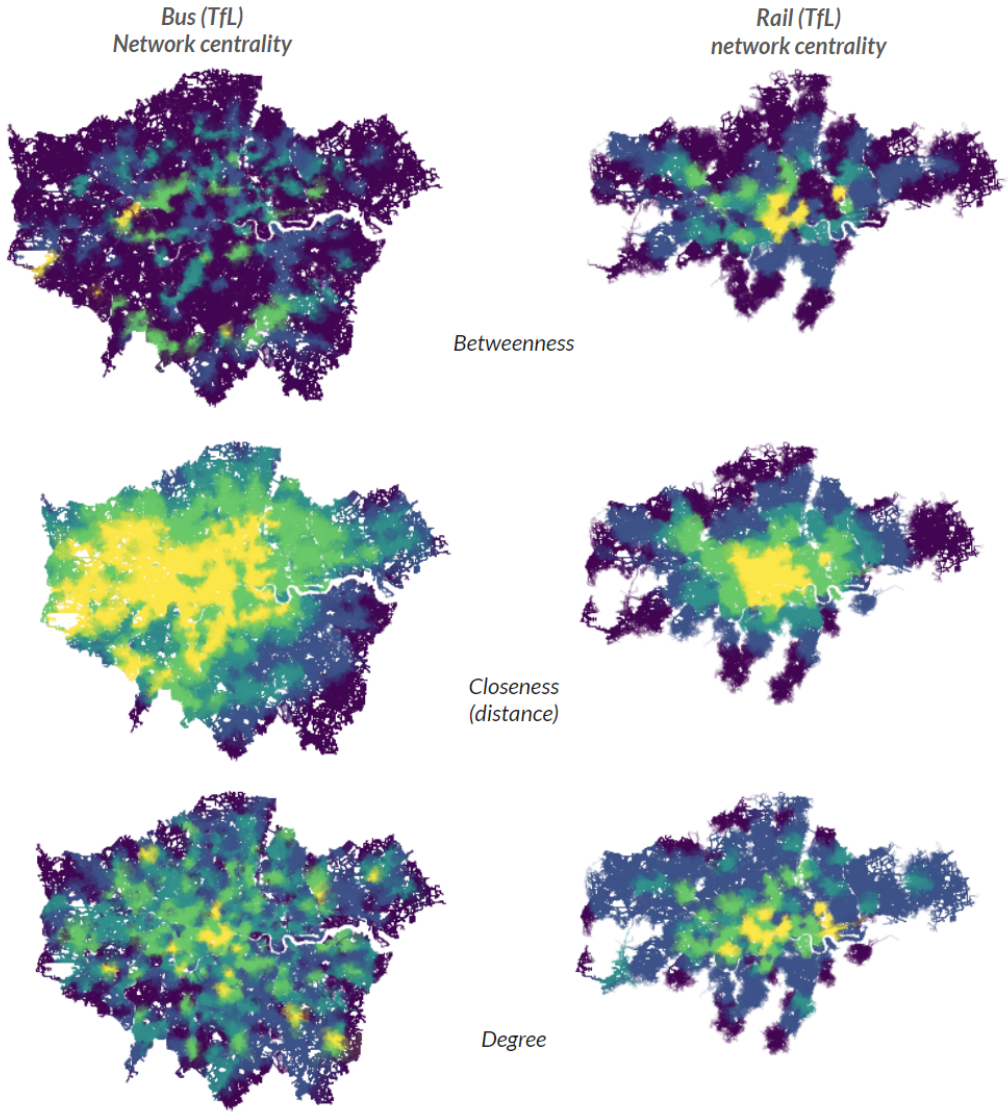

Figure 3.11: Mapping three centrality measures of spatial units on the Rail (left) and Bus (right) networks

- *Degree centrality*: The number of direct connections a node has in the network. A high degree of centrality indicates that the node is well-connected to other nodes in the network (i.e., many routes available).
- *Closeness centrality*: The average shortest path length from a node to all other nodes in the network. A high closeness centrality indicates that the node is close to other nodes in the network (i.e., centrally-located).
- *Betweenness centrality*: The number of shortest paths that pass through a node in the network. A high betweenness centrality indicates that the node is an important bridge between other nodes in the network (i.e., transport hubs).

Finally, in order to incorporate calculated attributes of the nodes into **6 features** of centrality, we will consider all the nodes that are in each of the spatial units and extract the maximum attribute values. In other words, the centrality feature of a spatial unit has the value of the rail or bus transport node with the highest centrality measure of each of the six attribute. The mapping of the centrality measures to the spatial units is illustrated in Figure 3.11.

3.2.3 Incorporating spatial lags as a feature

In summary, the feature engineering process thus far has resulted in a total of **25 features** that will be used to train the model to predict the (log-transformed) public transport arrivals in a given spatial unit. The features are summarised in Table 3.1.

No.	Feature
1	Population density
2-13	Amenity POI density by type (12)
14-16	Transport POI density by type (3)
17	Amenity POI diversity index
18	Transport POI diversity index by mode
19	Transport POI diversity index by route option
20-25	Bus and Rail centrality measures (Degree, Closeness, Betweenness) (6)

Table 3.1: Summary of the 25 main features engineered for the model

When working with machine learning models with data that exist in a closed spatial system such as the Greater London area, it is important to consider the spatial autocorrelation in the target variable. In the context of our analysis, this means that the number of passengers arriving in an area is likely to be correlated with the number of passengers arriving at nearby areas. Although deep learning models' parameters and hyperparameters can be made more complex and tuned to reach a high degree of prediction accuracy without considering the spatial aspect of the dataset, accounting for potential spatial autocorrelation could reduce the probability of biased outcomes even with a relatively low-complexity and high-interpretability architecture, e.g. tree-based models (Meyer et al., 2019).

The indication of spatial autocorrelation in the target variables is confirmed by the Local Indicators of Spatial Association (LISA) analysis, which shows the presence of spatial clusters in the target variables (Figure 3.12), using Local Moran's I indicators. If spatial information among these spatial units is provided to the model during tuning and evaluation, we could improve accuracy further, especially in these localities that make up more than 50% of the study area.

To address this issue, Liu, Kounadi, and Zurita-Milla (2022) has proposed incorporating spatial lag into the model as a feature. More specifically, the authors have concluded that when spatial features were induced in the random forest model, the global spatial autocorrelation among the predictions' residuals was significantly reduced, also leading to higher validation accuracy. For our analysis, the spatial lag feature of each spatial unit is calculated as the average number of passengers arriving at neighbouring isochrones[1]. The creation of a **Spatial Lag features** is replicated for the total arrival counts, as well as for each time band, resulting in a set of additional features to be included to train the model in predicting the corresponding target variable.

[1]The choice of the distance band of 750m was calibrated with the objective to ensure all spatial units have neighbours, i.e., there are no islands.

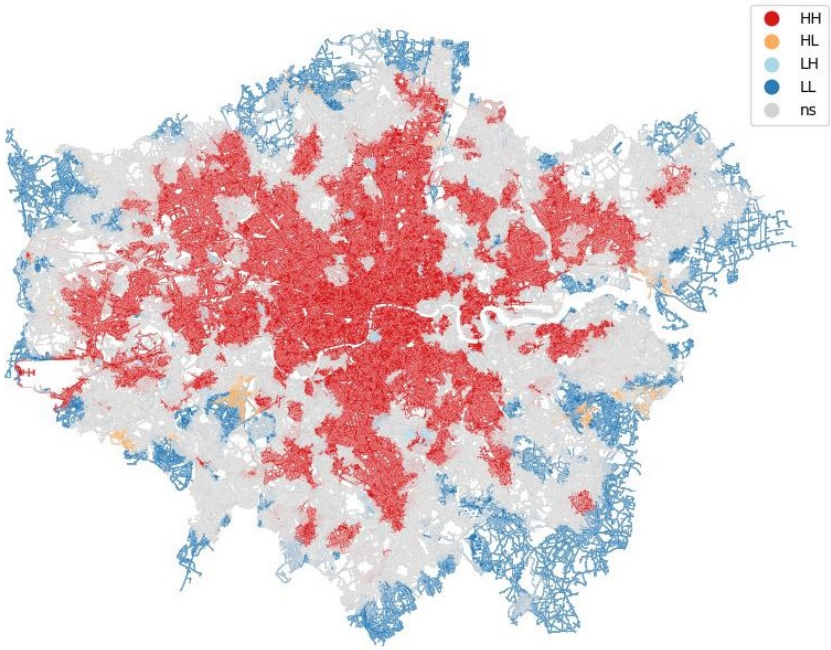

Figure 3.12: LISA cluster map of Total arrivals showing local sptial autocorrelation
Global Moran's I: 0.742 (High spatial autocorrelation)

No.	Target variable	Main Features		Spatial Lag Feature
1	Total arrivals	Population, amenity and accessibility profile (25 features)	+	Spatial Lag - Total arrivals
2	Morning arrivals			Spatial Lag - Morning arrivals
3	Midday arrivals			Spatial Lag - Midday arrivals
4	Evening arrivals			Spatial Lag - Evening arrivals
5	Late arrivals			Spatial Lag - Late arrivals

Table 3.2: Mapping spatial lag features to predict corresponding target variables

3.3 Model selection and interpretation

As previously explored through relevant literature, we are faced with an important decision when attempting to model the non-commute travel demand in the Greater London area based on variables related to the built environments, such as the amenity profile and the connectivity profile. More specifically, the choice is between using a machine learning model—which prioritises prediction accuracy and can handle heterogeneous data, and a spatial-statistical one—which prioritises interpretability and requires more stringent requirements of the explanatory variables. Although spatial statistical models such as the Spatial Lag Model or Geographically Weighted Regression are standard for such analysis, Li (2022) has shown that machine learning models can capture spatial effects if spatial information is provided and the appropriate Machine Learning explanation technique is used. The aim of this work is to demonstrate the use of SHAP explanation on a fine-tuned prediction model to provide insights into the geographically varying effect of the features on the predictions. Figure 3.2 provides an overview of the model training and interpretation process.

Model evaluation

Firstly, we will compare the performance of three machine learning models: *Linear Regression, Gradient Boosting, and Neural Network* in predicting Total arrivals. As we are working within a spatial system with a high degree of spatial autocorrelation demonstrated earlier, special attention will be paid to spatial cross-validation in hyperparameter tuning and evaluation because random cross-validation and train-test splits are not able to prevent train-test data leakage, which inflates the accuracy score.

More specifically, the model tuning and evaluation process will be carried out using a nested 10-fold cross-validation approach introduced by Schratz et al. (2018), with R^2 as the performance metric. The nested cross-validation approach is illustrated in Figure 3.13, where the inner loop is used to tune the hyperparameters of the model, and the outer loop is used to evaluate the model's performance. The spatial 10-fold cross-validation splits for our datasets (see Figure A.1 in Appendix) ensure that the training and test sets in both the inner and the outer folds are spatially distinct. The model and its hyperparameters with the highest R^2 score will be selected as the best model to fit the entire dataset for further analysis[2].

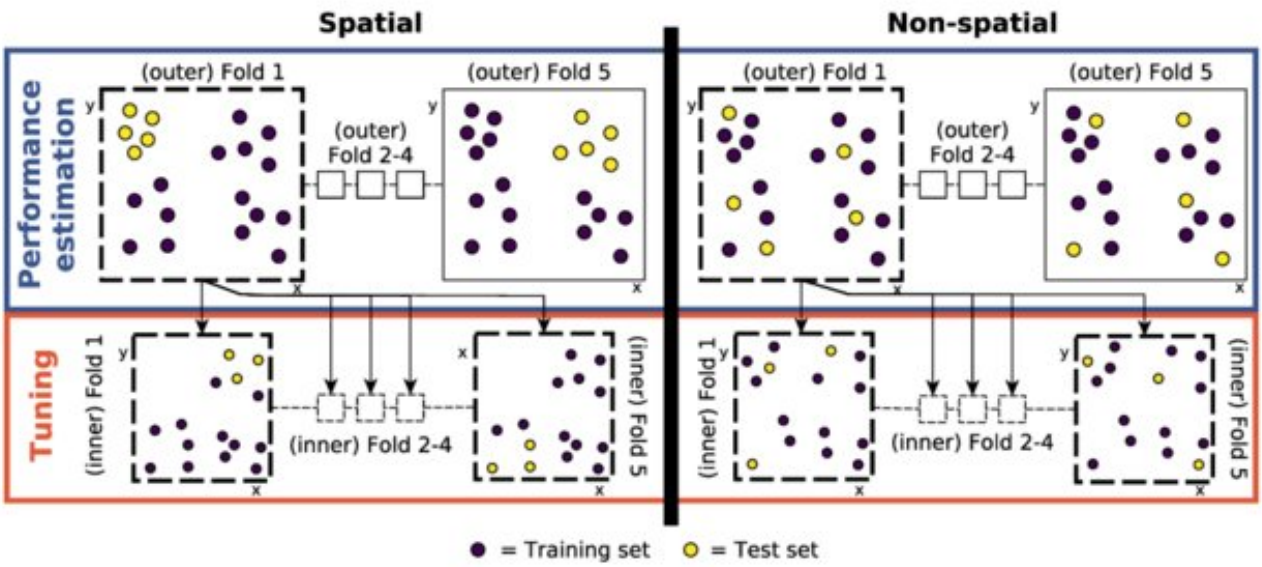

Figure 3.13: Spatial vs. non-spatial nested cross-validation using 5-folds for hyperparameter tuning and performance evaluation. This technique helps ensure the training and test sets in both the inner and the outer folds are spatially clustered to avoid data leakage
Source: Schratz et al. (2018)

Model interpretation

Rooted in the work on game theory, SHAP tackles the challenge of fair credit allocation within a cooperative game. In this reimagined context, features in our data are the players, the set of features used for training represents the coalition, and a model prediction is the game. For each observation, the SHAP values quantify the average marginal contribution of a feature to achieve

[2]Although the model tuning and selection process is only performed on Total Arrivals as the target variable for simplicity, the chosen model will be applied to predict arrivals by time bands.

the final prediction, providing a nuanced understanding of each feature's impact on the model's predictions at the local level.

Among other properties contributing to its popularity in Explainable AI research, it is model-agnostic and additive, meaning that the sum of the SHAP values for all features is equal to the difference between the model's prediction and the baseline prediction. While computation was initially a barrier, approaches like Kernel SHAP and Tree SHAP made available in the SHAP package have made estimation more efficient. SHAP's local interpretability and integration with popular machine learning libraries also make it a powerful tool for spatial data where feature attribution can be visualised geographically. These properties of SHAP allow us to derive our methodology to interpret the model's prediction in the context of the features engineered to answer the research questions:

- SHAP-based *global feature importance* will be used to identify the most important features (amenity and connectivity profiles) to the model's prediction of inflow travel demand, allowing us to understand the general patterns and relationships between the features and the target variables. Since it is important that the model in question attains a high level of accuracy before the SHAP values can be considered insightful, due care will be taken to ensure that the model is well-tuned and evaluated before proceeding with the SHAP analysis.
- SHAP-based *local feature importance* will be used to interpret the model's prediction for individual spatial units, allowing further investigation into identifying local spatial trip attraction patterns. With spatial clustering analysis such as HDBSCAN on the local SHAP, we can also identify clusters where our main features of interest have high (or low) importance compared to others, effectively identifying hotspots of high trip attraction factors for non-commute travel demand.

The next section will present the results of the model training and interpretation process, followed by a discussion of the findings and their implications for the research questions and beyond.

4. Results

4.1 RQ1: Predicting attractiveness from spatial features

Prediction model evaluation and selection

As previously explained, the choice of a suitable machine learning model is critical, as Machine Learning explanation techniques such as SHAP would be less insightful if the predictions produced high inaccuracies. The evaluation results in Table 4.1 show that the Gradient Boosting tree-based model (XGBoost) outperforms the other models in predicting the target variable *Total arrivals (log)* with an R^2 of 0.792. Meanwhile, the simple Linear Regression model performs more poorly in comparison due to underfitting (i.e., consistently low accuracy with training data folds), and the Multilayer Perceptron Artificial Neural Network model, also performs poorly but due to overfitting (i.e., high accuracy achieved with training data folds but low accuracy with validation data folds). XGBoost also outperforms itself when predicting the pre-transformed target variable with an R^2 of 0.633. This shows that while XGBoost has long been popular thanks to its fast convergence and high validation accuracy, transforming the highly right-skewed target variable can further improve prediction performance.

Model	R^2 **(raw)**	R^2 **(log)**
Linear Regression	0.576	0.697
XGBoost	0.633	**0.792**
ANN	0.526	0.410

Table 4.1: Comparing model R^2 results of 3 models predicting Total arrivals (raw) and Total arrivals (log) using spatial k-fold cross-validation for tuning and evaluation

On managing overfitting risks, XGBoost is a standout among other models also for its built-in L1 and L2 regularisation hyperparameters, which penalise the model for complexity and suppress highly correlated input features. This is particularly important when considering the complex patterns of correlations among our input features, as shown in Figure 4.1. Here we observe groups of highly correlated features such as Population-Public Facilities-Religious POIs, Entertainment-Restautants-Shopping POIs, or among the different public transport centrality measures. Therefore, with XGBoost, we had the opportunity to further fine-tune two hyperparameters *alpha* and *lambda*—which represent L1 and L2 regularisation thresholds, respectively—to achieve even better prediction results. As done previously, nested spatial k-fold cross-validation is used to fine-tune the model and validate the results. The fine-tuned XGBoost model improves R^2 up to **0.823** (an approx. 3pp increase) in predicting total arrivals from the pre-tuned model and performs consis-

Timeband	R^2	MSE	MAE
Total	**0.823**	0.687	0.570
Morning	**0.809**	0.614	0.571
Midday	**0.828**	0.629	0.549
Evening	**0.819**	0.687	0.593
Late	**0.824**	0.635	0.601

Table 4.2: Finetuned XGBoost model evaluation results for All target variables (log) using nested spatial k-fold cross-validation for tuning and evaluation

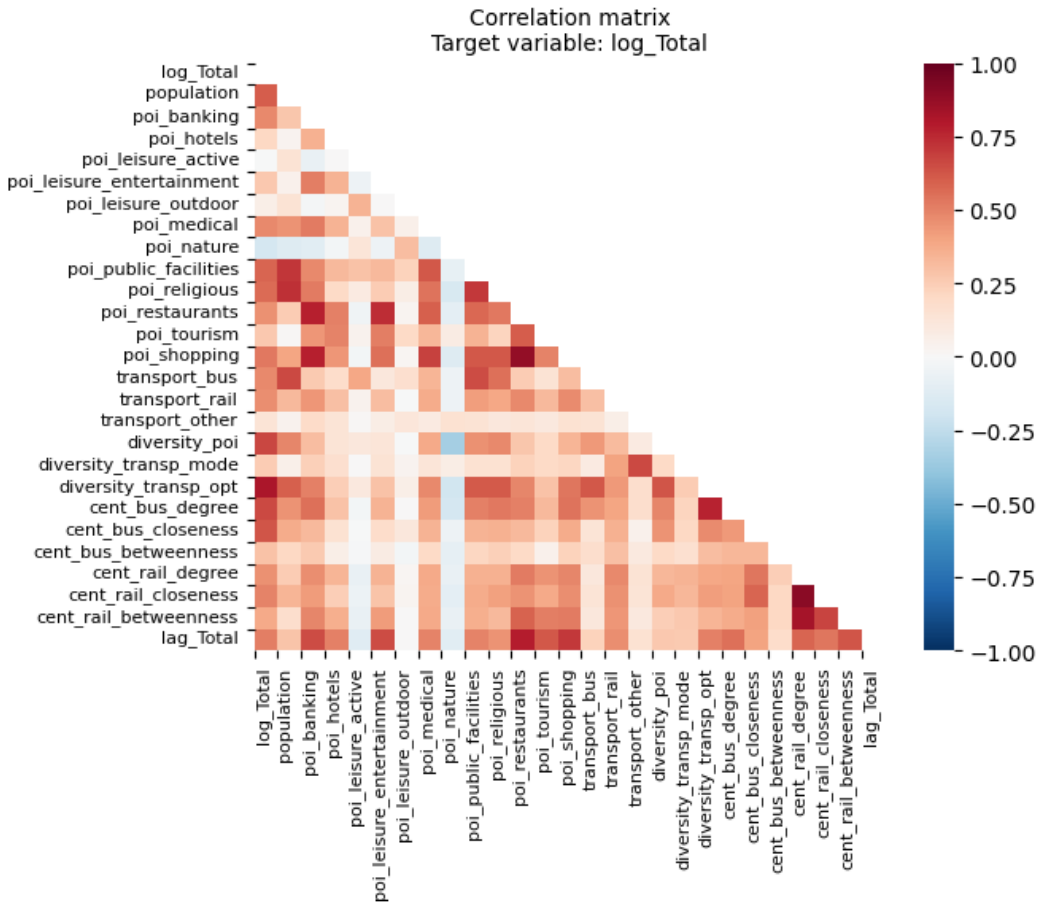

Figure 4.1: Correlation matrix of the input features and target variable Total arrivals (log)

tently well in predicting arrivals by time band, as shown in Table 4.2. The results suggest that the fine-tuned XGBoost model is fit for purpose and enables us to address the first research question: OPen data sourced built environment features are adequate predictors of public transport inflows (as a proxy for trip attractiveness) to a high degree of accuracy.

Before further analysis, it is important to recognise where the model falls short and what that may tell us about its potential lack of generalisability. Figure 4.2 shows the distribution of observations whose prediction residuals are Tukey's fence's 'outliers', i.e., the target variables are significantly overpredicted or underpredicted. These are mostly concentrated in Outer London where poor predictions go in both directions. While it is not easy to immediately unravel, possible

root causes include:

- *Low Spatial Autocorrelation:* There is consistency between areas with low spatial autocorrelation of the target variable (see Figure 3.12, grey regions), and where the model predicts more poorly. This may indicate incorporating spatial lag as a feature does improve the model's performance only where the spatial autocorrelation is present. In its absence, the model has a higher chance of poorer predictions based on the main features alone.
- *Omitted features:* It is possible that adding variables such as car ownership and socioeconomic factors can help the model predict two hypothetical areas with the same built environment profiles but different public transport usage. This is especially true in Outer London, where car ownership is higher, and public transport services are less frequent.
- *Incomplete target variable:* It is probable that a spatial unit with its built environment profile does have the predicted public transport trip attractiveness in reality. However, since only data from TfL-operated services are considered, omitting trip data from other rail operators—more prevalent in Outer London and especially south of the Thames—it may manifest as an apparent overprediction.

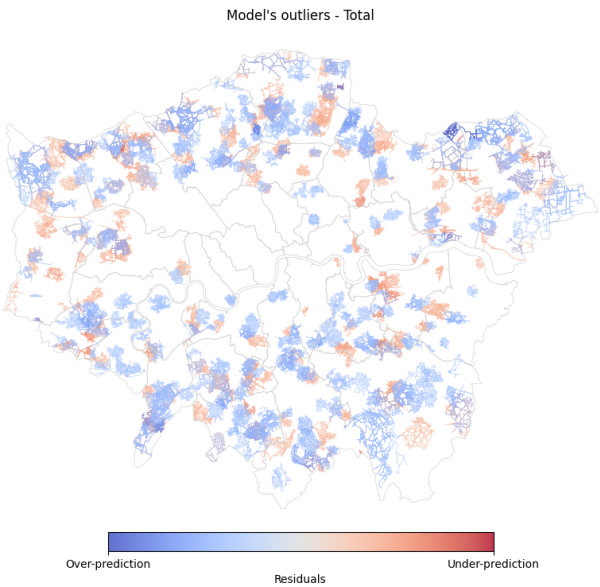

Figure 4.2: Observations with significant overprediction and underprediction (N=576 of 16,890) Fine-tuned XGBoost model, Total arrivals (log)

Nonetheless, these 'outliers' account for less than 3.5% of the total observations, and the model's performance is still satisfactory in explaining the large majority of the spatial units in the Greater London area.

Global feature importance

After fitting the fine-tuned XGBoost model to the entire dataset, we applied SHAP to obtain the SHAP values for each feature and observation. This allows us to evaluate the importance of each feature in predicting the target variable and understand the direction of the relationship between the feature and the target variable.

Figure 4.3 shows the summary of feature importance for the target variable *Total arrivals (log)* from most to least important. First, it is evident that the connectivity profile features have outsized importance with the *transport route option diversity index (diversity_transp_opt)* on top in predicting trip attractiveness. This is consistent with the reality that people tend to alight/exit at transport-rich areas where they have multiple bus routes or rail services for onward journeys, representing the most prominent interchange behaviour. The second most important feature is the spatial lag of the target variable (i.e., mean of target variables of neighbouring), which partially corroborates its influence to the model's performance hypothesised earlier when analysing outliers. The most important amenity POI feature is *shopping POIs*, which is also consistently important across all time bands (see Figure A.2 in Appendix), followed by *banking POIs*, *active leisure POIs*, *medical POIs*, *restaurant POIs*, *nature POIs*, and *outdoor leisure POIs*. This suggests that for the whole of Greater London, total arrivals as a proxy for trip attractiveness are driven mainly by, first and foremost, the need to interchange for onward travel and, secondly, by the availability of certain key types of amenities.

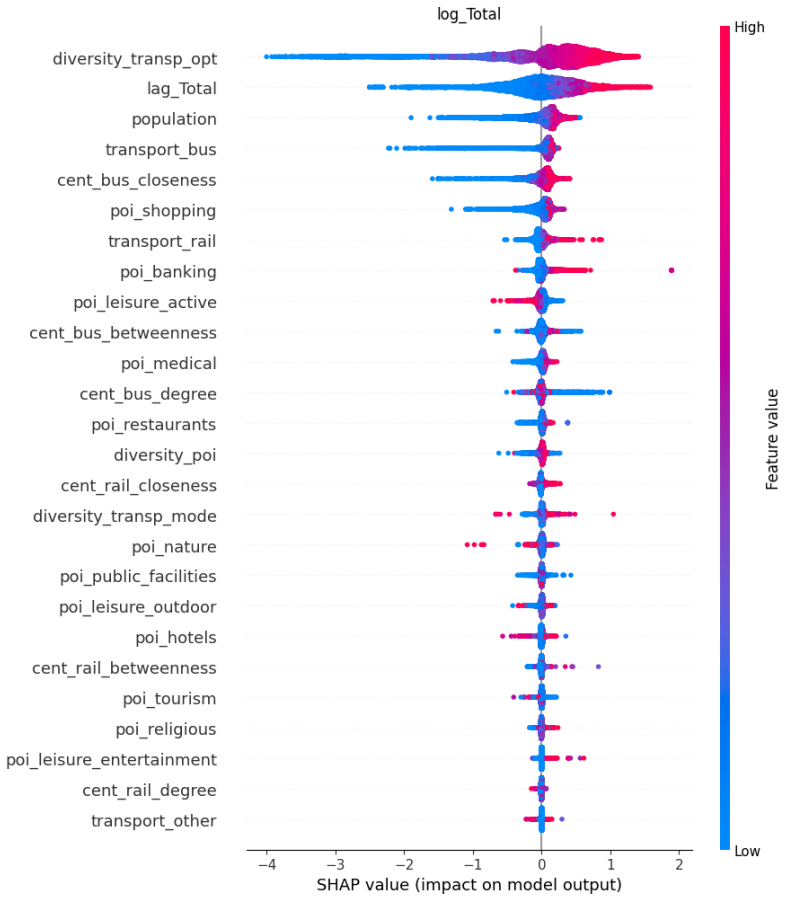

Figure 4.3: Summary of feature importance based on SHAP values. Red-Blue scale represents the feature value, and the x-axis represents the SHAP value, i.e., impact on final model prediction of Total arrivals (log)

At this point, it is important to acknowledge the effect of correlation between the input features as previously observed and how the XGBoost model learns on SHAP global explanations. Although XGBoost's regularisation hyperparameters are in place to avoid overfitting due to highly correlated features, they do not drop them from the model. Consequently, when SHAP is applied for a retroactive explanation, highly correlated features will show comparable importance. In such cases, it is, therefore, more insightful to consider the SHAP values of the correlated features in

clusters, as shown in Figure 4.4. We can see that there are differences from the previous graph, where the top most important features (or feature clusters) are as follows:

1. Transport route option diversity index
2. Spatial Lag
3. Population + Public Facilities (clustered, local residential amenity features)
4. Bus stop density + Bus stop closeness centrality (clustered, bus interchange features)
5. Density of Shopping, Banking and Restaurants POIs (clustered, commercial amenity features)

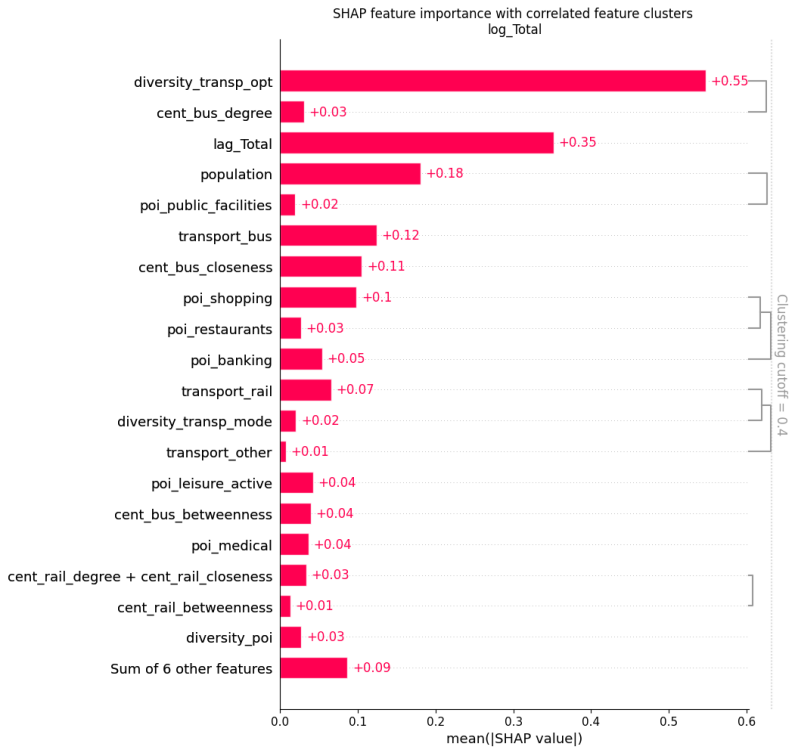

Figure 4.4: Feature importance in predicting Total arrivals (log)
Features clustered by correlation and ordered by mean absolute SHAP values

In summary, as it pertains to our first research question, the global feature importance analysis shows that built environment features are indeed important predictors of public transport inflow demand in London and that the XGBoost model is able to capture the complex relationships between these features and the target variable with high accuracy (¿80%). There are also opportunities to improve the prediction performance by addressing possibly omitted spatial and non-spatial features. The high importance of the spatial lag feature is both an advantage and a shortcoming: While it drastically improves the model's performance overall for Greater London, it penalises the model in areas where the spatial autocorrelation is low, reducing its generalisability to predicting the trip attractiveness of an isolated area without considering its neighbours. Finally, highly correlated features should be considered in clusters when analysing importance to avoid interpreting undue importance to a single feature.

4.2 RQ2: Identifying local attractors and hotspots

One of the highlights of SHAP is its ability to provide local explanations for individual predictions. When applied to a spatial system, these local explanations provide additional insights on spatial phenomena that are not well captured in a global feature importance analysis, namely spatial heterogeneity (i.e., varying impact of the target variable across different spatial units) and spatial non-stationarity (i.e., varying impact of an input feature on the target variable across different spatial units). SHAP's additive property allows us to decompose the prediction of a spatial unit into the sum of the SHAP values of each feature and exact the contribution of each feature to the prediction.

Figure 4.5 shows the local SHAP values for two spatial units, Kew Gardens and Covent Garden, for a closer look. They represent two distinct localities with supposedly different amenity and transport profiles. Kew Gardens is a quiet residential area in the public transport network, also known for its many renowned parks and gardens, while Covent Garden is a bustling commercial and entertainment district with high connectivity in the public transport network. The local SHAP values for Kew Gardens show a higher contribution of the population density and POI classes that are associated with neighbourhood amenities (stores, clinics, etc.) to the prediction of total arrivals, while the local SHAP values for Covent Garden show a higher contribution of POI classes such as restaurants, entertainment venues, etc. This is consistent with the expected profiles of the two areas and demonstrates the ability of SHAP to provide local explanations for individual predictions. In both cases, in line with global feature importance, the transport route option index's contribution is among the top, indicating that with the station exit and bus alighting dataset at hand, the interchange behaviour is still a significant pattern regardless of where in London the spatial unit is located.

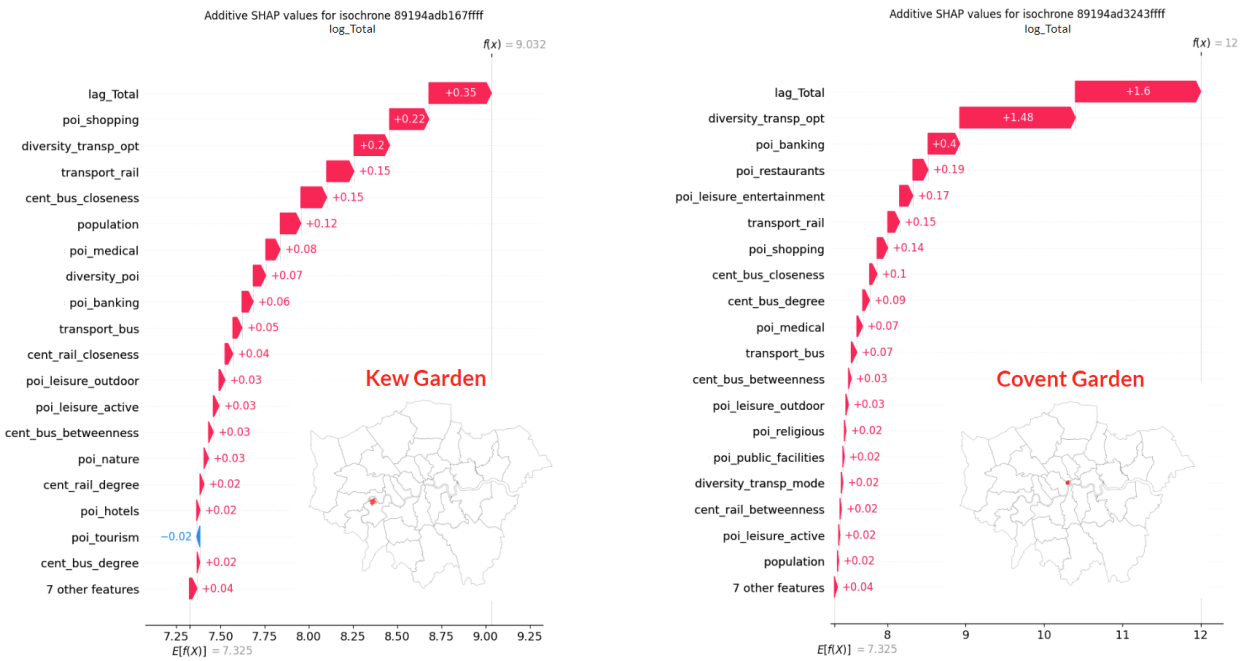

Figure 4.5: Local SHAP values for 2 spatial units
Kew Gardens and Covent Garden

By mapping the local SHAP values for all spatial units in the Greater London area, we can visualise the differing impact of the input features on the target variable across the region of interest. Figure 4.6 shows the local SHAP values for two features, Outdoor leisure POIs and

Entertainment leisure POIs, for the target variable of *Total arrivals (log)*. The map shows that the impact of these features on the target variable is not uniform across space, with some areas showing a positive impact (red) and others showing a negative impact (blue). Central London understandably commands the highest contributions of entertainment POIs to the prediction of trip attractiveness, but it also shows secondary areas of high impact in the outer boroughs, such as Brixton and Haringey. Outdoor leisure POIs, on the other hand, show a more dispersed impact citywide with a hotspot around Richmond.

On the temporal dimension across time bands, we can also observe the varying impact of the input features on the target variable across different time bands. Figure 4.7 depicts local SHAP values spatial distribution throughout the day normalised by the target variable (arrival by time band) to enable comparison. Entertainment POIs in Central London have the highest trip attractiveness in the Midday and especially Evening time bands. In comparison, Outdoor leisure POIs show a more consistent impact across time bands, with a notable increase in the intensity of certain hotspots and coldspots in the Late time band.

While it is not immediately clear why this surfaces, it reveals a shortcoming in the methodology when examining the temporal effect of the feature contribution to trip attractiveness among different features. According to global feature importance for the Late time band (see Figure A.2 in Appendix), *outdoor leisure POIs* seem to have an increased impact on the prediction contrary to the expectation that no passengers would be visiting outdoor leisure POIs after midnight. Since SHAP's goal is to ultimately approximate how a trained model makes decisions, it may overestimate the importance of existing features when omitted features would have been able to explain the variances better. This also means the choice of the model training methodology and the feature selection process so far may not be a one-size-fits-all for temporal datasets, and additional features might be necessary to avoid overinflating the contribution of certain features for individual time bands.

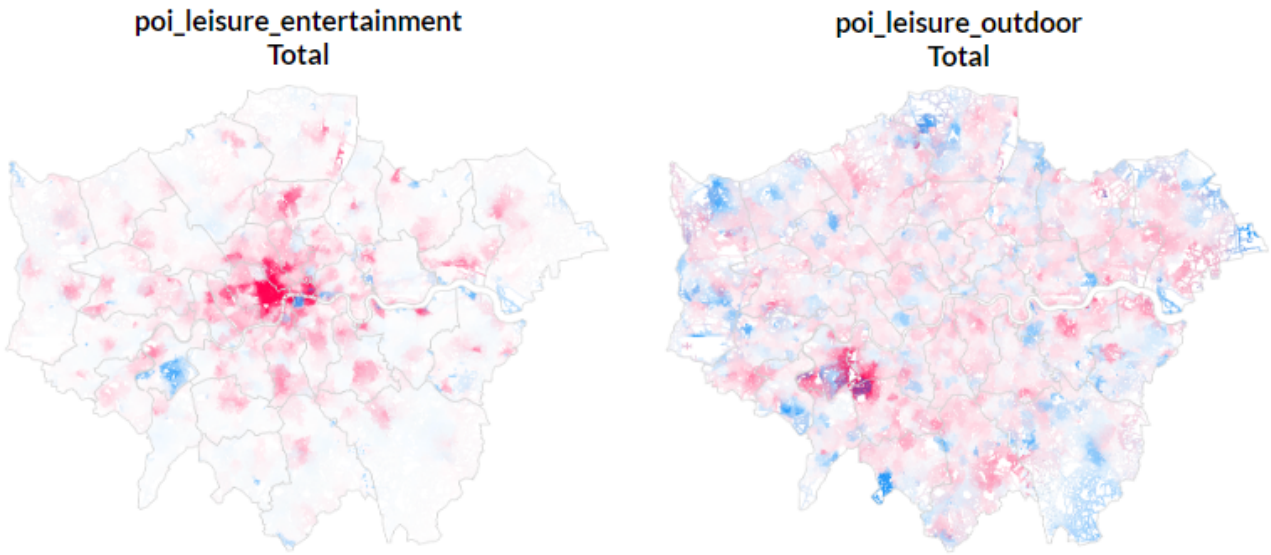

Figure 4.6: Mapping local SHAP values for 2 features: Outdoor leisure POIs and Entertainment leisure POIs. Target: Total arrivals (log)

Finally, the availability of raw SHAP values also enables other manipulations for specific purposes. For example, with different types of weightage, we can observe the impact of the features on the target variable in a different light. Figure 4.8 shows the comparison of raw and weighted SHAP values for two features, Outdoor leisure POIs and Entertainment leisure POIs, for the target variable *Total arrivals (log)*, weighted in two different ways:

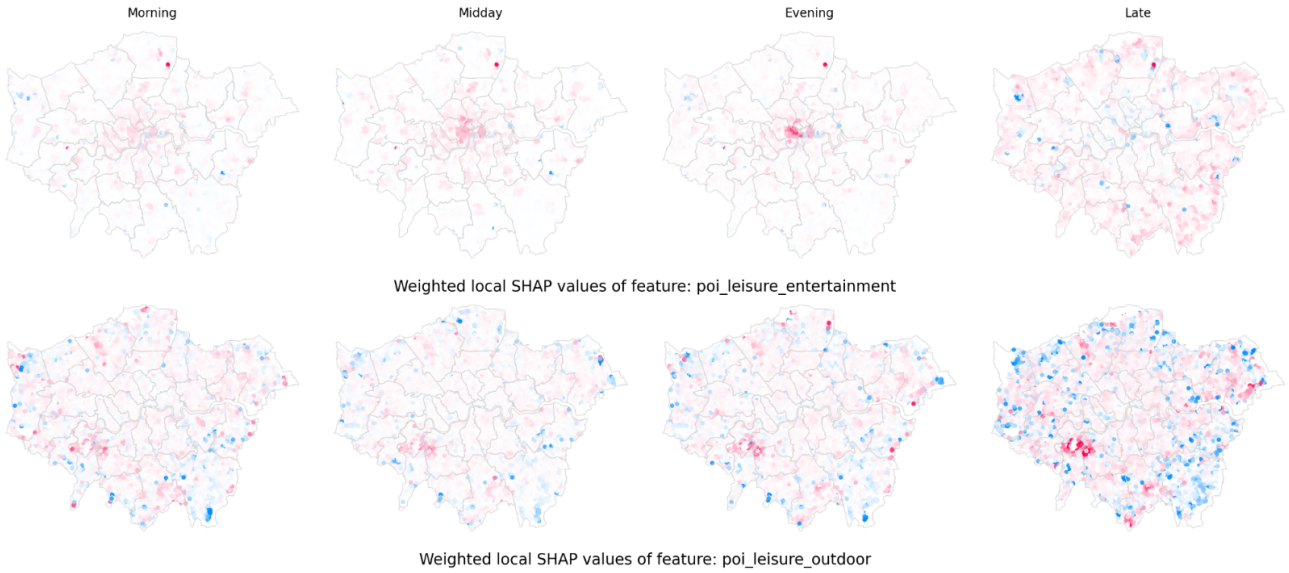

Figure 4.7: Mapping local SHAP values for 2 features: Outdoor leisure POIs and Entertainment leisure POIs, weighted by number of arrivals

1. Weighted by amenity diversity index: This can add more highlights to feature hotspots in areas with many other potential *contenders* for trip arrivals. Under this treatment, we can see that Entertainment POIs have an even more pronounced impact in Central London and clusters of small areas with high amenity diversity, indicating that these Entertainment POIs attract trips despite the fact that the area is saturated with other amenities.
2. Weighted by transport route option diversity index: This can add more highlights to feature hotspots in areas that are not well-served by many public transport routes. Under this treatment, we can see that Outdoor POIs on the outskirts of London now have a more pronounced trip attraction factor, indicating that despite the lack of public transport connectivity, these Outdoor POIs are still relatively popular.

Overall, SHAP local interpretability with individual SHAP values for each spatial unit provides a more nuanced understanding of the impact of the input features on the target variable across space and, in our case, also time. This is particularly useful for identifying activity hotspots in the Greater London area, as it allows us to pinpoint areas where certain features have a high impact on the prediction of trip attractiveness in the form of *Total arrivals (log)*. Limitations exist, however, in the methodology as the comparative power of SHAP is diminished when models predicting many target variables are involved, as they are much more adept at enabling comparison of feature importance within the same model rather than across models, such as in the case of different time bands. In our discussion, we will also consider the implications and applications of these findings on urban analytics literature and transport planning.

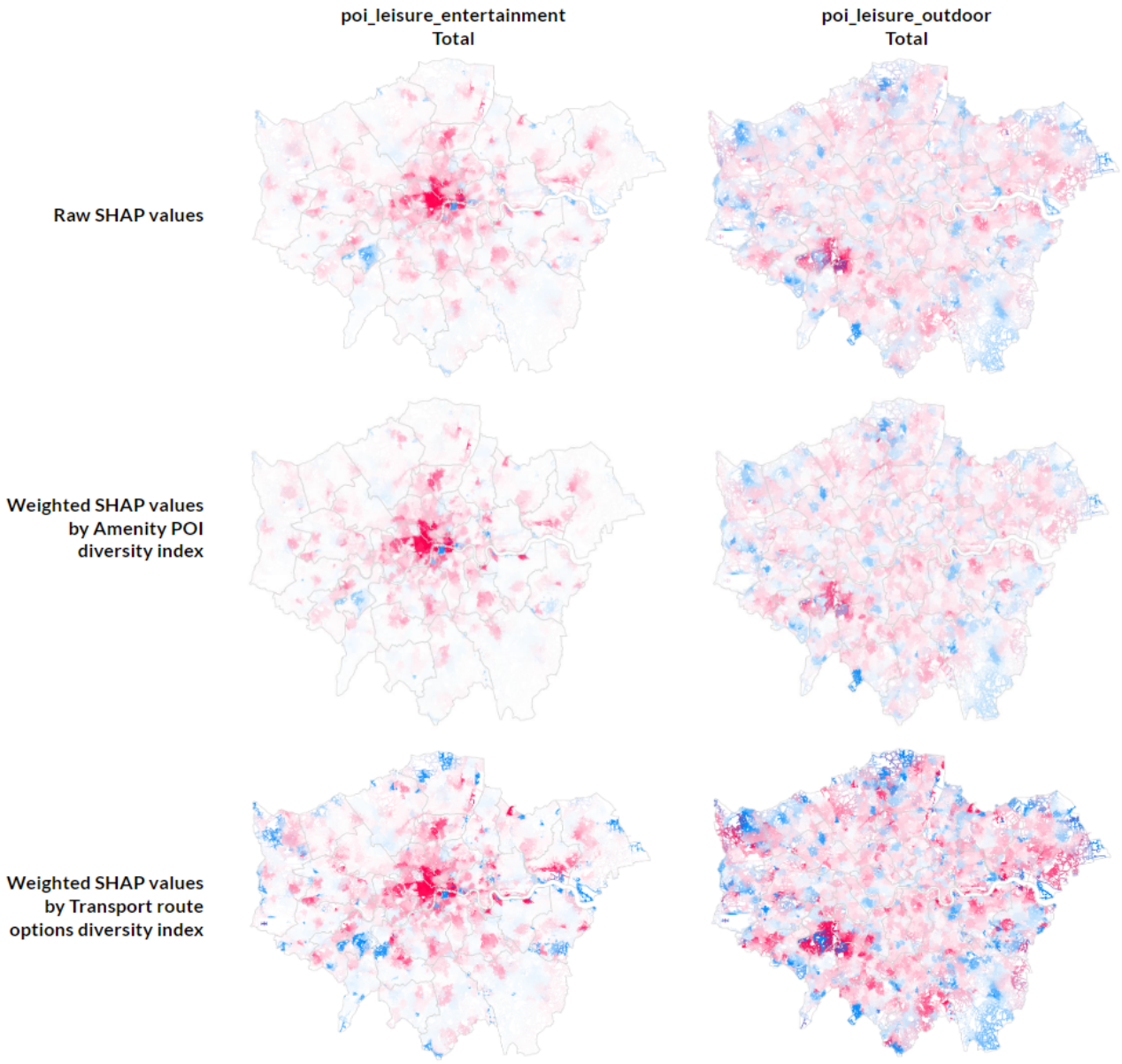

Figure 4.8: Raw vs. weighted SHAP for 2 features
Outdoor leisure POIs and Entertainment leisure POIs
Total arrivals (log)

5. Discussion

To recapitulate our findings thus far: First, we have established that provided with open data on population, amenity points-of-interest density and diversity, and public transportation accessibility indices, we can predict the total number of arrivals by public transport to a high degree of accuracy using an XGBoost model that is fine-tuned to overcome overfitting. Secondly, using SHAP for machine learning model explanation, we managed to extract more sophisticated feature importance indications than originally available with XGBoost. Furthermore, by examining local SHAP explanations, we can identify hotspots for different POI features overall and across different time bands, together with its shortcomings.

One of the benefits of using SHAP values as imperfect approximants to coefficients in a statistical model is their additive nature. This means that we can decompose the model's predictions into the sum of the SHAP values, with which we can control for unwanted effects. For example, without an explicit origin-destination matrix and dwell time data and using only station exits and bus alighting data, it is not immediately possible to say whether a person is exiting at a destination to attend an activity or to interchange to another mode of transport. Using SHAP to extract local importance with its additive property can help consider the impact of the amenity profile of the destination area on the trip attractiveness of the area while controlling for the connectivity profile of the area.

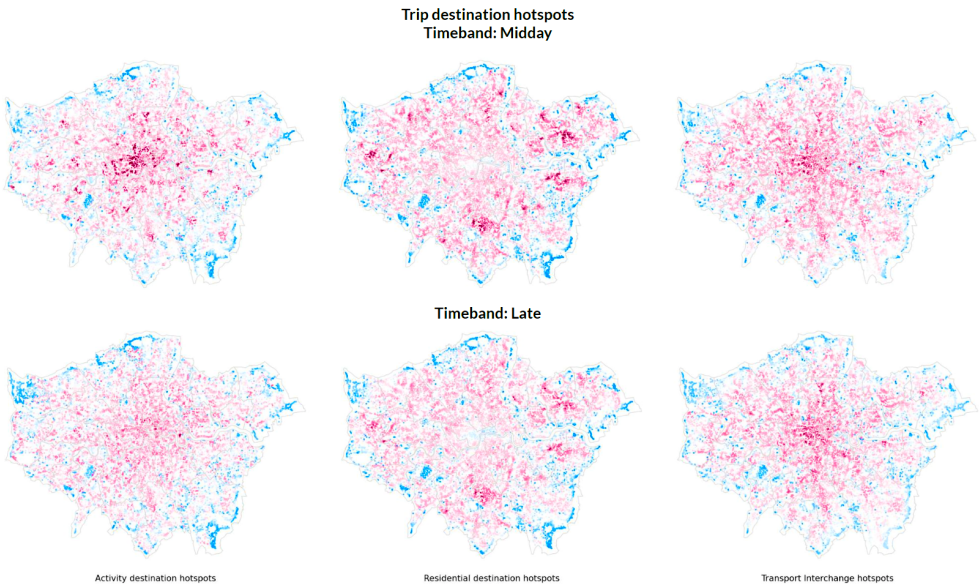

Figure 5.1: Hot spots for destination groups: Amenity vs. Residential vs. Transport Midday vs Late

One way to visualise it is by grouping local SHAP values by feature groups and visualising their spatial distribution. This can be seen in Figure 5.1 with features grouped into (1) amenity-related, (2) population, and (3) connectivity-related features. When comparing Midday and Late time bands, representing the peak and lowest volumes respectively, we can see that the trip attractiveness of areas because of their connectivity profile is unchanged, whereas the trip attractiveness of areas because of their amenity profile is significantly more pronounced in the Midday time band, despite the fact that connectivity-related features have high global importance. With the hypothesis that the only difference between these two time bands are the trip attraction generated by amenities as destination, this suggests that when grouped in this manner, the model as the whole seems to be able to isolate whether an area is attractive because of their amenities or their connectivity in the network.

This is useful for instances where we want to qualify certain popular areas as destination hotspots based on the availability of amenities alone while disregarding their connectivity-based attractors. Applying spatial clustering such as HDBSCAN of spatial units based on their amenity-related compound SHAP values can effectively surface activity hotspots in a given area in Greater London. Figure 5.2 demonstrates the activity hotspot in Wood Green with the highest mean amenity-related compound SHAP values, made up of 5 constituent spatial units (isochrones). This can be useful for local urban planners to identify destination areas that are popular for non-commute activities, such as parks, shopping, dining, or entertainment.

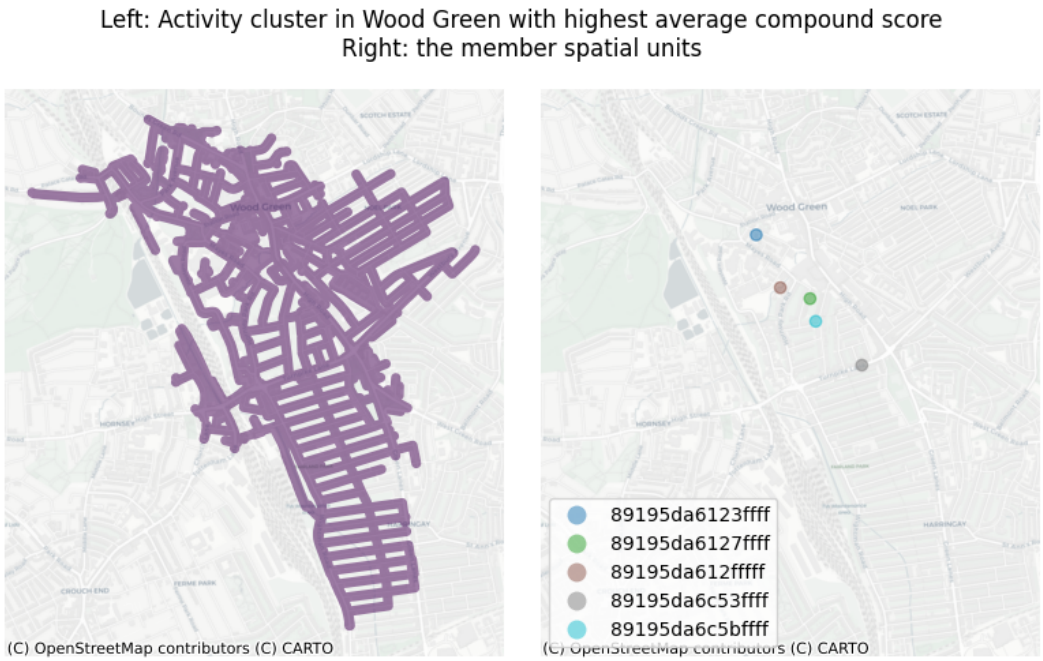

Figure 5.2: Identifying high activity areas in Wood Green

Finally, it is worth acknowledging the methodology's limitations, which affect its generalisability and applicability in certain cases.

- *Core assumptions on target variable selection*: In order to approximate non-commute trip attractiveness, we opted for data on a Saturday as a typical weekend day to remove the effect of employment centres from the analysis as much as possible. However, this does not account for travel demand from those who work on weekends.
- *Missing features*: The analysis of outliers suggests that there are other factors that may be influencing the number of arrivals in an area that are not accounted for in the model. These

could include socioeconomic factors, car ownership, or other urban morphological features that affect people's usage of public transport.

- *Generalisability*: The high feature importance of the spatial lag features across the board means that the model may not be able to accurately predict a new observation in isolation from its neighbours, nor can it be readily generalisable to other cities. Nevertheless, the model can be used to predict the trip attractiveness of a given area in Greater London as the city changes, giving transport planners a useful tool. For example, when a neighbourhood is undergoing redevelopment, we can predict the potential inflow of public transport trips based on the planned addition of amenities or transport facilities.
- *Causality*: However, the lack of consideration for causality means that we cannot infer that the presence of a certain amenity of a certain density will cause more arrivals to a certain area. Rather, we can only make predictions of potential trips attracted based on the presence of certain amenities.
- *POI classification*: The classification of amenity and transport POIs in this analysis adheres to Geofabrik's out-of-the-box OSM data classification. Future work could attempt to re-classify these POIs in a deliberate manner to capture the association between certain trip purposes of interest and the types of amenities that fulfil them as destinations.

Last but not least, the scope of the dissertation is limited to public transport trips and does not account for other types of mobility, such as cycling, walking, or driving, which may be more prevalent on weekends, which would provide a more comprehensive understanding of the non-commute trip attractiveness of an area regardless of mode of transport, and during time bands when public transport is less accessible such as late nights. Future work could incorporate mode-agnostic mobility data together with trip purpose inferences into a modified version of this methodology to provide a more comprehensive understanding of what draws people to a certain area in Greater London. Special care should be taken to ensure that the chosen features reflect the trip purposes and the transport modes of interest.

6. Conclusion

In conclusion, the methodology presented in this dissertation is a useful tool for transport planners to predict the trip attractiveness of a given area in Greater London based on its amenity and connectivity profile. By using SHAP values to extract local importance, we can identify destination hotspots based on the amenity profile of an area, which can be useful for local urban planners to identify areas that are popular for non-commute activities. However, the methodology is not without its limitations, and further work is needed to improve its generalisability and applicability in certain cases.

[to be completed]

A. Appendix

A.1 Open data sources

The analysis carried out will utilise the Greater London area as the working case study, serving as an example for further replications of the analysis for other localities. The Office for National Statistics (ONS) defines the Greater London metropolitan area to include the City of London and the 32 London boroughs, with a total area of 1,572 km^2 and a population of 9.4 million people, making it the largest in the United Kingdom. The choice of the Greater London metropolitan area is also motivated by the availability of open data sources from which necessary features could be derived and the complexity of the public transport network, coupled with rich usage data made available to the public. The data sources used in this analysis are summarised in Table A.1 and detailed below.

	Variables	**Sources**
Public transport arrivals (target)	Tube and Rail station exits Bus alightings	TfL network demand (NUMBAT) TfL network demand (BUSTO)
Population	Population count	UK Census 2021
Amenity Profile	POI count by type POI diversity index	OSM Amenity POIs
Public Transport (PT) Connectivity Profile	PT node count by type PT node diversity index PT network centrality	OSM Transport POIs and TfL

Table A.1: Summary of data used in the analysis and their sources

Transport for London (TfL) Open Data

We use the following GIS datasets from the TfL GIS Open Data Hub and TfL API to derive transport-related features for the study area including density of transport facilities, diversity of transport options, and accessibility in the form of centrality measures in the Rail and Bus networks separately. The networks of interest are defined as follows:

- *Rail network* includes the London Underground, London Overground, the Docklands Light Railway, and the Elizabeth Line, excluding stations and segments that are not managed by TfL or fall outside the Greater London area.

- *Bus network* includes stop locations and route data of all bus and tram services, excluding stops and route segments that are not managed by TfL or within the Greater London area.

TfL also provides annually-updated network demand datasets, which will be used to derive the target variables of our analysis. The datasets of special interest to our analysis are namely *Rail demand data (NUMBAT)* and the *Bus demand data (BUSTO)*, from which we will extract estimated counts of station exits and bus alightings at the station or stop level, respectively. The dataset's temporal granularity is also an important aspect to consider. To address non-commute trips for the analysis, we will specifically use **average Saturday demand**, representing the demand in a typical period without any major citywide disruptions or events[1]. To examine possible variations between different times of day, besides the total counts of exits and bus alightings in one day, we will aggregate the data into four larger time bands[2]:

- *Morning* (05:00 - 10:00): Limited mobility expected due to early hours
- *Midday* (10:00 - 19:00): High daytime mobility expected
- *Evening* (19:00 - 00:00): High evening mobility expected
- *Late* (00:00 - 05:00): Minimal mobility expected due to sparser night PT services

It is worth noting that despite the existence of the Oyster Card integrated fare system, TfL does not provide origin-destination data as open data. Therefore, station exits and bus alightings are not directly associated with the number of passengers traversing through the system but are treated as discrete events. For example, a person who takes a bus to reach a Tube station for onward travel and then exits to arrive at their destination would be counted as two separate data points, one alighting for the bus and one station exit for the Tube. Nevertheless, our model will take this into account when making poreditcion and allows us to investigate patterns and the contribution of modal interchange behaviour in our interpretation.

Lastly, we need to acknowledge the particularity of the TfL bus demand dataset (BUSTO). Although passengers do not 'tap off' when alighting from busses, TfL has adopted a methodology to estimate the number of passengers alighting at each bus stop based on individuals' next actions and other statistical methods. The methodology is described in the BUSTO User Guide and Data Dictionary

UK Census Data 2021

For the estimated population datasets, we rely on the most recent United Kingdom Census data administered in 2021. The data is ingested at the Output Area (OA) level, which is the smallest geographical unit for which census data is available. Further processing will aggregate the data to the desired spatial unit of analysis.

OpenStreetMap POIs

OpenStreetMap (OSM) is a collaborative mapping project that provides a free and open-source map of the world. The version of OSM data used in this analysis is packaged by an OpenStreetMap

[1] This Saturday travel demand may still include work trips for those working on Saturdays, albeit not concentrated into peak periods as with weekday travel demand. For the scope of this analysis, we will treat these potential work trips as non-commute.

[2] The choice of time band delineation is a compromise between different granularity levels afforded by the TfL rail and the bus demand datasets. There are opportunities to deepen the analysis by further segmenting the time bands into finer intervals if data permits.

community project Geofabrik and provides a snapshot of the OSM database as of June 2024. Geofabrik also standardises the categorisation of POIs' amenity attributes in the OSM database into larger classes[3]. Our analysis will use this default POI classification, extracting 12 amenity POI types. Lastly, apart from amenity POIs, we also make use of the OSM transport POIs point data to derive non-TfL transport-related features, such as national rail stations, ferry piers, long-distance bus stations and airports.

Combined with transport node data acquired from TfL open data sources, all extracted Amenity and Transport POI types of interest are summarised in Table A.2.

	POI Type	**Examples**
Amenity	*Public Facilities*	Post offices, schools, universities, libraries,...
	Medical	Hospitals, clinics, pharmacies,...
	Entertainment	Theatres, Night clubs, cinemas,...
	Outdoors	Parks, playgrounds,...
	Active	Sports centres, swimming pools, stadiums,...
	Restaurants	Restaurants, cafes, pubs, bars,...
	Hotels	Hotels, motels, hostels,...
	Shopping	Supermarkets, bakeries, florists, bookshops, malls,...
	Banking	Banks, ATMs,...
	Tourism	Tourist attractions, museums, monuments, viewpoints,...
	Religious	Churches, mosques, temples,...
	Nature	Riverbanks, beaches, hills,...
Transport	*Bus stops*	Bus stops and stations (TfL)
	Rail stations	TfL rail stations (TfL) and Non-TfL rail stations (OSM)
	Other transport hubs	Airports, ferry piers, long-distance bus stations (OSM)

Table A.2: Summary of Amenity and Transport POI types used in the analysis

A.2 True to Data and True to Model in SHAP technique

One ongoing debate regarding the use of SHAP for explainable is captured by Chen et al. (2020)'s work exploring 'True to Model' versus 'True to Data' explanations. The 'True to Data' approach produces *observational SHAP values* which reflect how a feature contributes to the model's prediction when considering the dependencies and correlations present in the training data, whereas the 'True to Model' approach produces *interventional SHAP values* which reveal how a feature influences the model's prediction if we manipulate it independently, regardless of how it might interact with other features in reality. In our case, as we are more interested in observing feature contribution to the model prediction in a controlled environment rather than observing the relationships between features, we will use the 'True to Model' approach to interpret the model's prediction made possible by the *shap* package. This means SHAP values are calculated by iteratively entering features from a sample dataset into the model and observing the change in prediction, effectively isolating the feature's contribution to the model's prediction rather than explaining the feature contribution using the tree paths alone.

[3]More information about Geofabrik's OSM data product can be found in Geofabrik OSM data dictionary

A.3 Supplemental Figures

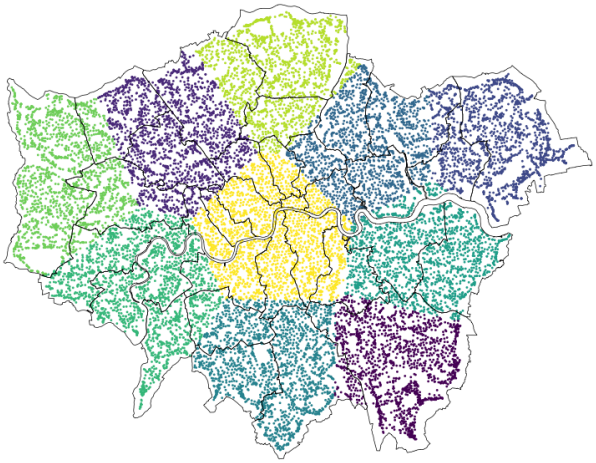

Figure A.1: Map of the spatial k-fold cross-validation splits for hyperparameter tuning and performance assessment (k=10)

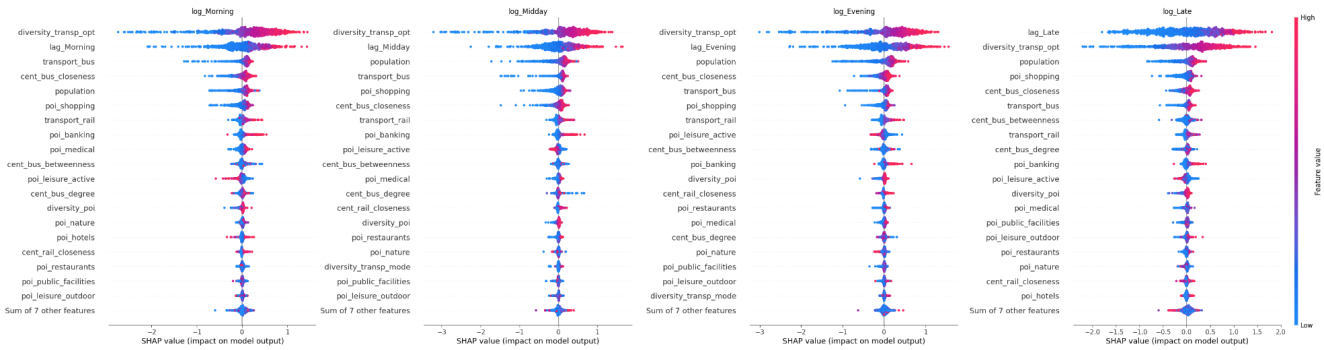

Figure A.2: Summary of feature importance based on SHAP values of a sample. Red-Blue scale represents the feature value, and the x-axis represents the SHAP value, i.e., impact on final model prediction of arrivals in each time band (log)

References

Aaqib Javed, Syed et al. (Apr. 6, 2020). "Estimation of Trip Attraction Rates and Models for Shopping Centers in Dhaka City". In: DOI: 10.5281/zenodo.3733088.

Batty, Michael et al. (Oct. 1, 2014). "Entropy, Complexity, and Spatial Information". In: *Journal of Geographical Systems* 16.4, pp. 363–385. ISSN: 1435-5949. DOI: 10.1007/s10109-014-0202-2. URL: https://doi.org/10.1007/s10109-014-0202-2 (visited on 08/01/2024).

Cervero, Robert and Kara Kockelman (Sept. 1, 1997). "Travel Demand and the 3Ds: Density, Diversity, and Design". In: *Transportation Research Part D: Transport and Environment* 2.3, pp. 199–219. ISSN: 1361-9209. DOI: 10.1016/S1361-9209(97)00009-6. URL: https://www.sciencedirect.com/science/article/pii/S1361920997000096 (visited on 06/14/2024).

Chen, Hugh et al. (June 29, 2020). *True to the Model or True to the Data?* DOI: 10.48550/arXiv.2006.16234. arXiv: 2006.16234 [cs, stat]. URL: http://arxiv.org/abs/2006.16234 (visited on 08/04/2024). Pre-published.

Convery, Sheila and Brendan Williams (Sept. 2019). "Determinants of Transport Mode Choice for Non-Commuting Trips: The Roles of Transport, Land Use and Socio-Demographic Characteristics". In: *Urban Science* 3.3 (3), p. 82. ISSN: 2413-8851. DOI: 10.3390/urbansci3030082. URL: https://www.mdpi.com/2413-8851/3/3/82 (visited on 06/18/2024).

Erlander, Sven and Neil F. Stewart (Dec. 1990). *The Gravity Model in Transportation Analysis: Theory and Extensions*. VSP. 252 pp. ISBN: 978-90-6764-089-3. Google Books: tId3PU1leR8C.

Jayasinghe, Amila, Kazushi Sano, and Kasemsri Rattanaporn (Oct. 1, 2017). "Application for Developing Countries: Estimating Trip Attraction in Urban Zones Based on Centrality". In: *Journal of Traffic and Transportation Engineering (English Edition)* 4.5, pp. 464–476. ISSN: 2095-7564. DOI: 10.1016/j.jtte.2017.05.011. URL: https://www.sciencedirect.com/science/article/pii/S2095756416302458 (visited on 05/17/2024).

Khavarian-Garmsir, Amir Reza et al. (Feb. 2023). "From Garden City to 15-Minute City: A Historical Perspective and Critical Assessment". In: *Land* 12.2 (2), p. 512. ISSN: 2073-445X. DOI: 10.3390/land12020512. URL: https://www.mdpi.com/2073-445X/12/2/512 (visited on 07/29/2024).

Li, Ziqi (Sept. 1, 2022). "Extracting Spatial Effects from Machine Learning Model Using Local Interpretation Method: An Example of SHAP and XGBoost". In: *Computers, Environment and Urban Systems* 96, p. 101845. ISSN: 0198-9715. DOI: 10.1016/j.compenvurbsys.2022.101845. URL: https://www.sciencedirect.com/science/article/pii/S0198971522000898 (visited on 07/07/2024).

Liu, Xiaojian, Ourania Kounadi, and Raul Zurita-Milla (Apr. 2022). "Incorporating Spatial Autocorrelation in Machine Learning Models Using Spatial Lag and Eigenvector Spatial Filtering Features". In: *ISPRS International Journal of Geo-Information* 11.4 (4), p. 242. ISSN: 2220-9964. DOI: 10.3390/ijgi11040242. URL: https://www.mdpi.com/2220-9964/11/4/242 (visited on 07/25/2024).

Louhichi, Mouad et al. (Jan. 1, 2023). "Shapley Values for Explaining the Black Box Nature of Machine Learning Model Clustering". In: *Procedia Computer Science*. The 14th International

Conference on Ambient Systems, Networks and Technologies Networks (ANT) and The 6th International Conference on Emerging Data and Industry 4.0 (EDI40) 220, pp. 806–811. ISSN: 1877-0509. DOI: 10.1016/j.procs.2023.03.107. URL: https://www.sciencedirect.com/science/article/pii/S1877050923006427 (visited on 07/24/2024).

Lundberg, Scott M and Su-In Lee (2017). "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html (visited on 07/24/2024).

MacLeod, Kara E., Brian L. Cole, and Charles Musselwhite (June 2022). "Commuting to Work Post-Pandemic: Opportunities for Health?" In: *Journal of Transport & Health* 25, p. 101381. ISSN: 2214-1405. DOI: 10.1016/j.jth.2022.101381. pmid: 35540370. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9074865/ (visited on 07/24/2024).

Melikov, Pierre et al. (Apr. 1, 2021). "Characterizing Urban Mobility Patterns: A Case Study of Mexico City". In: pp. 153–170. ISBN: 9789811589829. DOI: 10.1007/978-981-15-8983-6_11.

Meyer, Hanna et al. (Nov. 2019). "Importance of Spatial Predictor Variable Selection in Machine Learning Applications – Moving from Data Reproduction to Spatial Prediction". In: *Ecological Modelling* 411, p. 108815. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2019.108815. arXiv: 1908.07805 [cs, stat]. URL: http://arxiv.org/abs/1908.07805 (visited on 08/01/2024).

Schratz, Patrick et al. (Mar. 29, 2018). *Performance Evaluation and Hyperparameter Tuning of Statistical and Machine-Learning Models Using Spatial Data*.

Shannon, C. E. (July 1948). "A Mathematical Theory of Communication". In: *The Bell System Technical Journal* 27.3, pp. 379–423. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb01338.x. URL: https://ieeexplore.ieee.org/document/6773024 (visited on 08/01/2024).

Simini, Filippo et al. (Nov. 12, 2021). "A Deep Gravity Model for Mobility Flows Generation". In: *Nature Communications* 12.1, p. 6576. ISSN: 2041-1723. DOI: 10.1038/s41467-021-26752-4. URL: https://www.nature.com/articles/s41467-021-26752-4 (visited on 05/02/2024).

Wheeler, David and Michael Tiefelsdorf (June 1, 2005). "Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression". In: *Journal of Geographical Systems* 7.2, pp. 161–187. ISSN: 1435-5949. DOI: 10.1007/s10109-005-0155-6. URL: https://doi.org/10.1007/s10109-005-0155-6 (visited on 07/25/2024).

Wöhner, Fabienne (June 1, 2022). "Work Flexibly, Travel Less? The Impact of Telework and Flextime on Mobility Behavior in Switzerland". In: *Journal of Transport Geography* 102, p. 103390. ISSN: 0966-6923. DOI: 10.1016/j.jtrangeo.2022.103390. URL: https://www.sciencedirect.com/science/article/pii/S0966692322001132 (visited on 07/24/2024).

Xin, Yanan et al. (Oct. 18, 2022). *Vision Paper: Causal Inference for Interpretable and Robust Machine Learning in Mobility Analysis*. DOI: 10.48550/arXiv.2210.10010.

Yadav, Piyush et al. (June 1, 2017). *The Role of Open Data in Driving Sustainable Mobility in Nine Smart Cities*.

Yang, Yingxiang et al. (July 11, 2014). "Limits of Predictability in Commuting Flows in the Absence of Data for Calibration". In: *Scientific reports* 4, p. 5662. DOI: 10.1038/srep05662.

Zhong, Chen et al. (July 1, 2015). "Measuring Variability of Mobility Patterns from Multiday Smart-Card Data". In: *Journal of Computational Science*. Computational Science at the Gates of Nature 9, pp. 125–130. ISSN: 1877-7503. DOI: 10.1016/j.jocs.2015.04.021. URL: https://www.sciencedirect.com/science/article/pii/S1877750315000599 (visited on 06/18/2024).