Urban Planning
2025 • Volume 10 • Article 8551
https://doi.org/10.17645/up.8551

ARTICLE

Open Access Journal

What Is My Plaza for? Implementing a Machine Learning Strategy for Public Events Prediction in the Urban Square

Jumana Hamdani [1,2], Pablo Antuña Molina [3], Lucía Leva Fuentes [3], Hesham Shawqy [3], Gabriella Rossi [4], and David Andrés León [3]

[1] School of Information Systems and Technology Management, University of New South Wales, Australia
[2] Department of Spatial Planning, Blekinge Institute of Technology, Sweden
[3] IAAC—Institute for Advanced Architecture of Catalonia, Spain
[4] CITA, Royal Danish Academy, Denmark

Correspondence: Jumana Hamdani (j.hamdani@unsw.edu.au)

Submitted: 29 April 2024 **Accepted:** 22 July 2024 **Published:** 21 January 2025

Issue: This article is part of the issue "AI for and in Urban Planning" edited by Tong Wang (TU Delft) and Neil Yorke-Smith (TU Delft), fully open access at https://doi.org/10.17645/up.i388

Abstract
Plazas are an essential pillar of public life in our cities. Historically, they have been seen as public fora, hosting public events that fostered trade, interaction, and debate. However, with the rise of modern urbanism, city planners considered them as part of a larger strategic development scheme overlooking their social importance. As a result, plazas have lost their function and value. In recent years, awareness has risen of the need to re-activate these public spaces to strive for social inclusion and urban resilience. Geometric and urban features of plazas and their surroundings often suggest what kinds of usage the public can make of them. In this project, we explore the application of machine learning to predict the suitability of events in public spaces, aiming to enhance urban plaza design. Learning from traditional urbanism indicators, we consider factors associated with the features of the public space, such as the number of people and the high degree of comfort, which are evolved from three subcategories: external factors, geometric shape, and design factors. We acknowledge that the predictive capability of our model is constrained by a relatively small dataset, comprising 15 real plazas in Madrid augmented digitally to 2025 fictional scenarios through self-organising maps. The article details the methods to quantify and enumerate quantitative urban features. With a categorical target variable, a classification model is trained to predict the type of event in the urban space. The model is then evaluated locally in Grasshopper by visualising a parametric verified geometry and deploying the model on other existing plazas worldwide regarding geographical proximity to Madrid, where to share or not the same cultural and environmental conditions. Despite these limitations, our findings offer valuable insights into the potential of machine learning in urban planning, suggesting pathways for future research to expand upon this foundational study.

Keywords
data classification; event prediction; machine learning; Madrid; plaza; public squares; self-organising maps; urban planning

1. Introduction

The need for event prediction models in an urban environment is essential for all kinds of applications, from natural disaster preparation to urban management, planning, and development of smart cities (Mukhina et al., 2019). This rapid change in the relationship between events and public space highlights the features of the eventful city (Richards & Colombo, 2017) and how their form defines them, duration, content, and effects, determined to a certain extent by urban space and process (Richards & Palmer, 2010). The emphasis on people's behaviour in the public space and how it is affected by the built environment (Lynch, 1964) and its underlying elements that navigate the mode of human group life (Wirth, 1938) underscores the importance of designing urban spaces that foster social interaction and community engagement.

Traditional methods focus on trending major sporting and cultural events on a large scale (Smith, 2012, 2017), ignoring the continuity across scales to predict, prepare, and manage smaller-scale events and their outcomes. In addition, small-scale events remained unprecedented (Page & Connell, 2023) despite their essential effect on specific public spaces' publicness (Brighenti, 2010). In contrast to the traditional approach of other practices within the field of event prediction, our research tests the use of machine learning (ML) to overcome those limitations of scale, where those boundaries are diminished on the shared ground of the dataset.

Our main research question is: To what extent can ML models predict public events in urban plazas based on quantifiable urban features? This study aims to achieve two main objectives: (1) to explore the affordance of urban plazas and their contemporary design, use, and role; and (2) to test the application of ML for predicting public events in these spaces.

This study fills a critical gap in the literature by applying ML to predict the suitability of urban plazas for various public events, offering a novel methodological approach that integrates spatial and temporal data. Building on Donald Appleyard's theoretical framework (Appleyard, 1981) on street and public space activation, this research highlights the importance of designing public spaces that encourage various forms of social interaction and activities.

2. State of the Art

Modern studies have shifted from event detection, which focuses on identifying events after they occur, to event prediction, which aims to anticipate events before they happen. There are many existing works about event forecasting, including disease outbreaks (Achrekar et al., 2011; Zhao et al., 2015), crimes (Rumi et al., 2018; Wang et al., 2012), and other types of events (Dencik et al., 2018; Huang et al., 2017; Jin et al., 2014). For example, Smith (2012) and Wang et al. (2012) demonstrate how predictive models can forecast events such as protests, and cultural gatherings based on social media data and urban features. Existing methods for event prediction often rely on social media data like Twitter or Instagram (Kursuncu et al., 2018) but typically overlook spatial design features integral to public spaces (Ramakrishnan et al., 2014; Zhao et al., 2016).

Public spaces in historical districts, particularly in Mediterranean countries, play a vital role in urban life due to their historical and cultural significance. These spaces are often central to social interactions, cultural events, and community activities, making them ideal case studies for examining the impact of urban design on public space utilisation (Gehl, 2011; Lynch, 1964).

3. Methods

Our methodology hinges on a comprehensive approach to model development, encompassing the encoding of quantifiable plaza metrics in different cities, data augmentation, ML model training, and rigorous evaluation. The aim is to design a methodology that could be tailored to different scenarios and urban contexts. By choosing a single city to compose the dataset, we restrict the number of variables and features that appear throughout the process. Since our goal is the evaluation of urban plazas and outdoor activities, southern European cities are undoubtedly relevant. Due to its historical configuration and tourist interest, Madrid is taken as a case study, as its urban fabric comprises enough samples to establish a balanced initial dataset regarding size, shape, and urban situations (Figure 1).

Given the potential biases in our data, we limit our study to urban plazas in similar geographical and cultural contexts to Madrid. The selection of the indicators was guided by their relevance to public space utilisation. We categorised these indicators into external factors, geometric shape, and design features, and evaluated their influence on the suitability of plazas for different events. The indicators used in our classification include 1) External factors such as pedestrian traffic and proximity to amenities, 2) Geometric shape including size and layout, and 3) Design features like seating availability and shading. These indicators were chosen for their demonstrated impact on public space usage as identified in urban studies literature. The rationale for each indicator is based on its ability to influence the suitability of a plaza for various types of events.

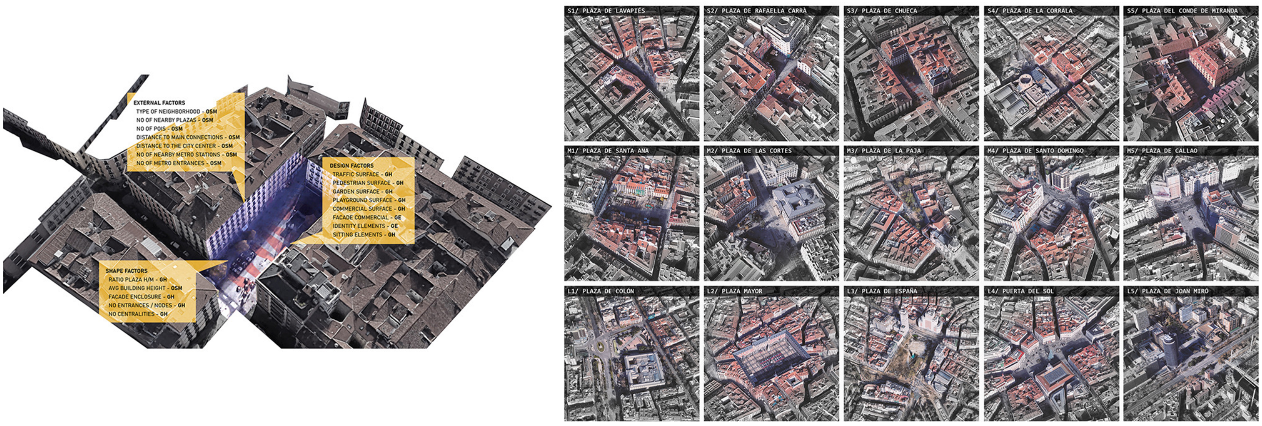

Figure 1. Initial dataset. On the left, plaza metrics are classified as External factors, Shape factors, or Design factors. Abbreviations: GH (Grasshopper plugin in Rhino software), OSM (Open Street Maps). On the right, satellite imagery of the 15 initial plazas grouped by size (first row large, second row medium, third row small). Source: Authors' creation from Google Earth.

3.1. Model Establishment and Indicators

To assess which aspects could be more relevant for the event suitability prediction we used Gehl's classification of outdoor activities (Gehl, 2011), which differentiates between necessary, optional, and social activities. Necessary activities occur in almost any condition and the physical environment only slightly influences their occurrence. Optional activities are those that people choose to do if the conditions are favourable. Finally, social activities depend on the presence of people and how they interact. With this definition in mind, optional activities are the ones greatly influenced by the physical nature of public space, so they will be the main focus of this research.

To underpin this classification, Gehl refers to a series of indicators to assess the quality and effectiveness of urban spaces:

- Degree of Comfort: Measures how comfortable people feel in a given urban space;
- Level of Activity: This indicator looks at the types of activities occurring in an urban space;
- Social Interaction: Measures frequency and quality of social interactions taking place in the space;
- Safety and Security: Evaluates how safe and secure people feel;
- Accessibility and Connectivity: Assesses how easily people can reach and move through the space.

For the specific case of Madrid, the Degree of Comfort and Level of Activity were chosen as the key indicators. Their evaluation does not include a fixed set of features; they are context-dependent and local urban knowledge is required to find the most appropriate ones. For the case of Madrid, we selected a series of features commonly used in urban studies and also specific to Madrid (Higueras García et al., 2017). They are categorised into external, geometric shape, and design factors (Figure 2).

- External factors:
  - Nearby plazas (*nearby_plazas*): Number of plazas located in the surroundings.
  - Points of interest (*poi*): Surrounding locations labelled of interest by Open Street Map.
  - Distance to the city centre (*dist_citycenter*).
  - Public transport lines (*transport_lines*): Number of transportation modes that reach the site.
  - Public transport accesses (*tr_stop/entrances*): Public transport stops in the surroundings.
  - Distance to public transport stops (*dist_connections*): Average distance to public transport stops.
- Geometric shape factors:
  - Buildings' height/plaza's width (*h_div_side*): Estimates how open the space feels in elevation.
  - Percentage of perimeter length surrounded by buildings (*per_building*): Gauges how open the space feels in the plan.
  - Number of entrances to the public space (*no_entrances*).
  - Number of centres (*centralities*): Related to the plaza's plan geometry, broken-down polycentric plan geometries are not ideal for some activities.
- Design factors:
  - Traffic surface (*traffic_srf*): Percentage of surface area dedicated to vehicular traffic.
  - Pedestrian surface (*pedestrian_srf*): Percentage of surface area dedicated to pedestrian traffic.
  - Garden surface (*garden_srf*): Percentage of surface area for green areas.
  - Playground surface (*playgr_srf*): Percentage of surface area dedicated to children-safe areas.

- Commercial surface (*commercial_srf*): Percentage of surface area dedicated to commercial activities.
- Sitting furniture (*sitting_furn*): Percentage of surface area with sitting furniture.
- Commercial facades (*fac_commercial*): Percentage of surrounding facades dedicated to retail space.
- Number of identity elements (*identity_elem*): Number of built identity elements within the plaza.

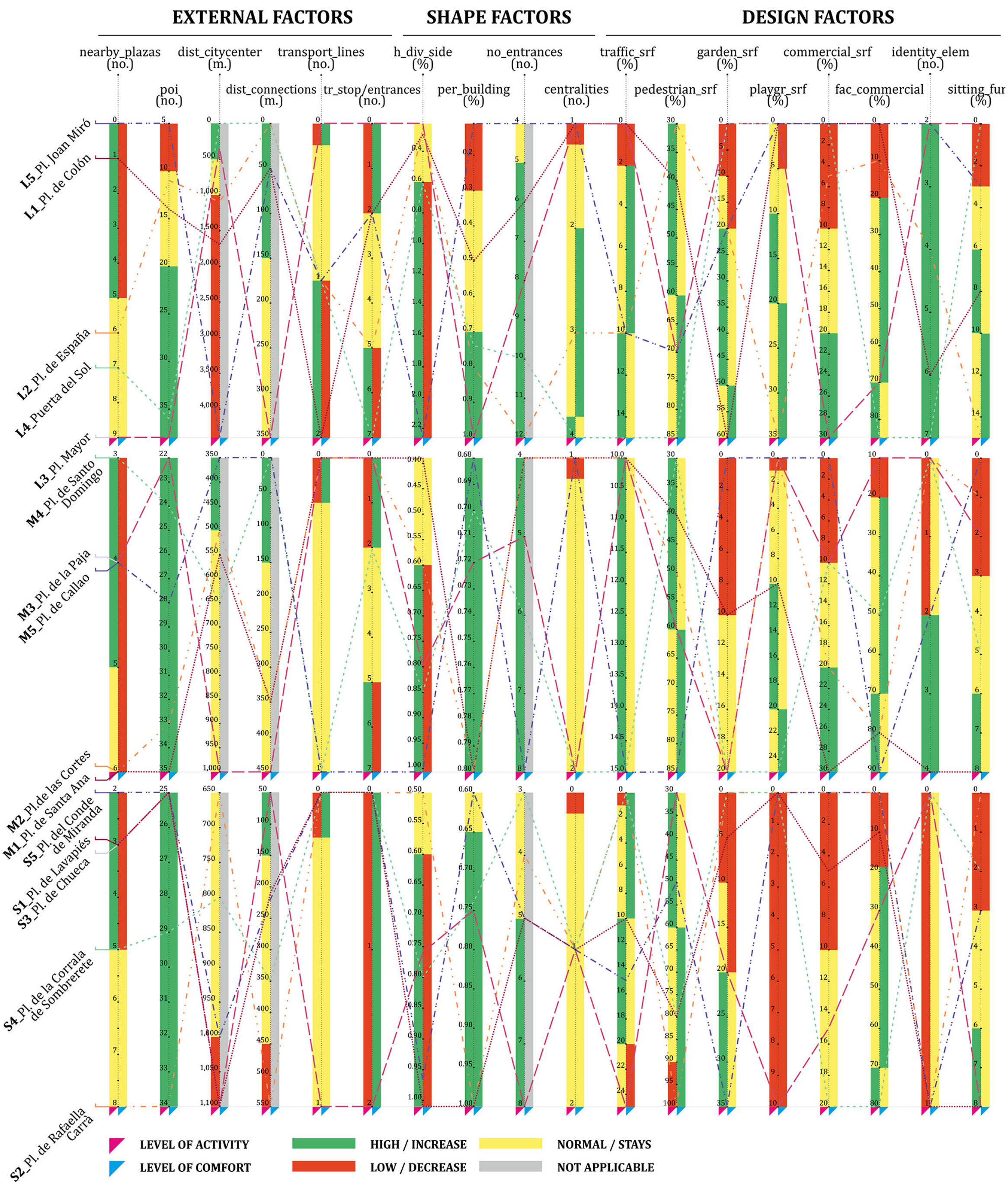

Figure 2. Multi-bar diagram showing the evaluation process (green high, yellow normal, red low, grey n/a) of each plaza (coloured lines), returning two values per input (columns); Activity Level (left bar/pink triangle) and Degree of Comfort (right bar/blue triangle). The increase/decrease of levels is driven by specific urban features such as pedestrian traffic, seating availability, and shading. The bars represent these levels concerning the type of events likely to occur.

The variables configuring our External factors were extracted from open data (OpenStreetMap) using Python's APIs. Nevertheless, due to the lack of urban data, most of the features encoded as Shape and Design factors have been geometrically rebuilt and subsequently compiled using Rhino/Grasshopper software (Version 7; Robert McNeel & Associates, 2022).

3.2. Model Fitting and Evaluation

The model was trained using a classification approach, where the selected target classes represent five different suitable events classified as optional activities following Gehl's classification. The proposed event types were chosen to reflect the programmatic diversity plazas can accommodate, focusing on those that require minimal fixed infrastructure and can attract varied demographic groups. The selected classes were: 1_Concert, 2_Market, 3_Urban performance, 4_Protest, and 5_Nothing.These categories were chosen based on common types of events identified in previous studies on public space utilisation (Smith, 2017; Wang et al., 2012). This selection allows us to examine the adaptability of public spaces to different events while excluding others like sporting activities, which typically require specialised facilities. These scenarios reflect a diverse range of events that plazas can accommodate, focusing on those that require minimal fixed infrastructure and can attract varied demographic groups.

Performance was assessed via accuracy, precision, and recall metrics, critical for evaluating the model's effectiveness in predicting event types based on plaza characteristics. Variability in event type predictions was analysed against key metrics of plaza design—Level of Activity and Degree of Comfort—to understand how different factors influence event suitability, as detailed in Figure 3.

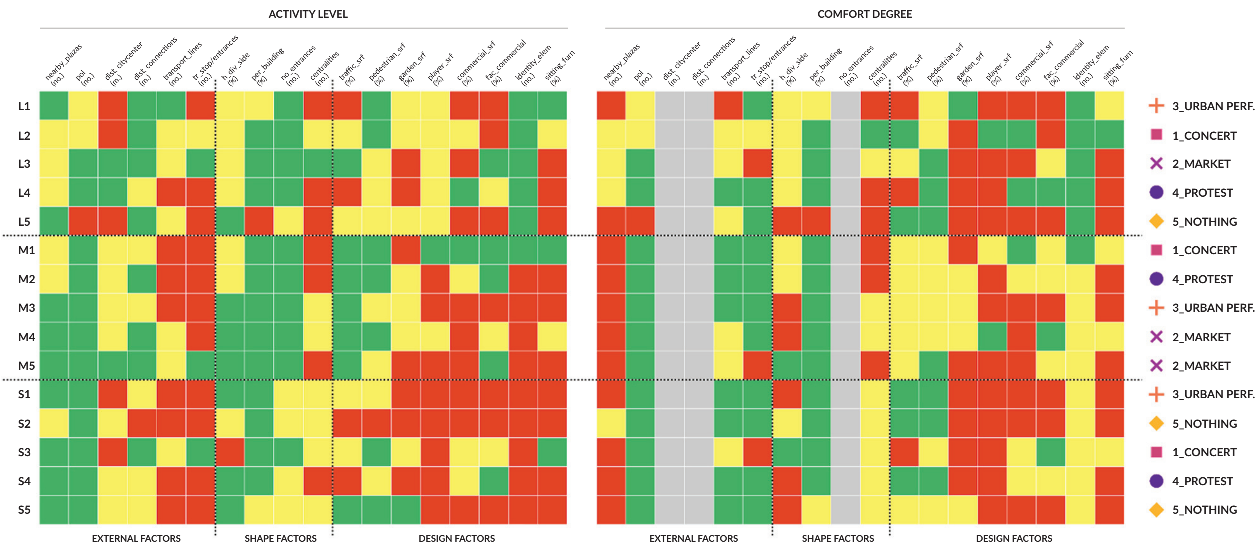

Figure 3. The figure shows the most likely event per plaza based on our model's predictions, though multiple events can occur in reality. It shows the result (green high, yellow normal, red low, grey n/a) of the evaluation process (Figure 2) of each plaza (row) per input (columns) according to its Activity Level (left matrix) and Degree of Comfort (right matrix), from which the output is assigned (right column).

3.3. Data Augmentation and Synthesis

To add variability to the initial sample in terms of shape and urban situations, the selection of the initial dataset includes 15 plazas differentiated by size: small (< 2000 m^2), medium (2000–7000 m^2), and large (> 7000 m^2). Each of them was analysed and evaluated according to the parameters established in the previous sections. Finally, the most suitable output out of the five proposed classes is assigned according to the result obtained regarding its two main indicators, Level of Activity and Degree of Comfort, which were identified based on their empirical relevance in urban studies.

There are two major drawbacks to curating a valid dataset for a single city. Firstly, encoding existing plazas, collecting, and formatting relevant metadata into a dataset is tedious. Secondly, the actual number of squares present in a single city is too few to become a valid dataset on its own. To address the challenge of a small initial dataset, we employed a self-organising map (SOM) for data augmentation, expanding our dataset from 15 initial plazas (seeds) to 2025. This approach allowed us to synthetically generate a diverse array of plaza scenarios, enhancing the model's learning potential. Of the numerous virtual sample generation methods that have proven effective for model training (DeVries & Taylor, 2017), the SOMs (Kohonen, 1995) relative mapping algorithm was used as the augmentation model in this research thanks to its Grasshopper implementation (Food4Rhino, 2021). After testing different combinations to determine which map size best allows the seeds to occupy the map boundaries, we established a square ratio of 45x45 as it produces the most distributed results.

The categorical data had to be coded as discrete quantitative variables so as not to introduce errors into the initial dataset. In addition, interpolating the discrete variables from the initial plazas yielded continuous variables in the data augmentation process, which had to be refined *a posteriori* to adopt the discrete value of their closest neighbour as a 2-dimensional interpolation method.

SOMs are a relative dimensionality reduction technique where initialisation plays a defining role in the final convergence of the map. To ensure that none of the initial seeds are underrepresented in the new synthetic dataset, there needs to be an iterative refining process where parameters could be tweaked and seeds removed until the SOM algorithm converges in a dataset that properly represents the problem.

3.4. Dataset Analysis

The produced synthetic dataset is subjected to statistical and graphical analysis to ensure it is coherent. This stage was essential for uncovering useful information and coming up with our conclusions at the end. Through this process, features were organised, transformed, cleaned up, added, or subtracted to understand the model predictability behaviour.

We analysed the correlation between our dataset's response and input variables using pair plots and correlation heatmaps. The goal was to detect features that might be redundant and therefore be able to reduce the dimensionality of our dataset through feature engineering. We also graphically analysed the distribution of the different categories' data points to understand possible challenges that could hinder model training.

3.5. ML Models

We trained different ML models on our synthetic dataset to test its quality, both shallow and deep. The three shallow learning models tested were Scikit-learn's Logistic Regression, Random Forest Classifier (Scikit-learn, 2022), and XGBoost Classifier. In addition, we chose a simple Artificial Neural Network (ANN; Tensorflow, 2022) as a deep learning model with cross entropy as a loss function. Finally, we tested two versions of our dataset with a train/test split of 80/20% (1620 train/405 test):

Raw dataset: Original dataset with all the selected 20 features.

Reduced dataset: Following the latest principal component analysis (PCA), the features were reduced to the 14 most relevant ones.

4. Results and Discussion

Following the previously presented method, we tested the approach with the 15 initial plazas.

4.1. Results From the Data Augmentation (SOM)

As mentioned, before reaching the final synthetic dataset we used for model training, we refined the initial parameters fed to the SOM algorithm multiple times. In our first attempts to produce our dataset, most seeds clustered and merged in the same synthetic data point while only two of the large ones (Plaza de Colón and Plaza de España) were distributed around the map. These two seeds and their associated programs were overrepresented compared to the rest that clustered. Despite fine-tuning the initial features, these two outlying plazas conditioned and unbalanced the dataset (Figure 4). To prevent this, they were excluded from the dataset.

When repeating the analysis with the same parameter distribution, we observed how the seeds mixed and occupied more significant portions of the board. However, they did not span toward two of the board edges, resulting in most events being underrepresented in the initial dataset and having a significant part of the synthetic data points distant from any labelled seed, meaning that most of the data generated was not within the parameter bounds of the 13 plazas used (15 initial ones minus the 2 outliers) and did not configure a valid dataset (Figure 4).

In the third attempt, we tried to find ways to occupy the board edges. We introduced two plazas from the large category, the same as the ones we removed in the previous stage. These additions added more variability to the initial plaza seeds and located themselves in the board edges without compromising the entire dataset. Finally, some of the metrics initially considered were modified or refined to avoid highly correlated values in the dataset (Figure 4).

Throughout the dataset curation process, shallow learning models were trained in an iterative process. We used a logistic regression (Scikit-learn, 2022) and an XGBoost Classifier (Chen & Guestrin, 2016), with different data representations to understand how relevant the selected features were and which ones were more meaningful for the classification.

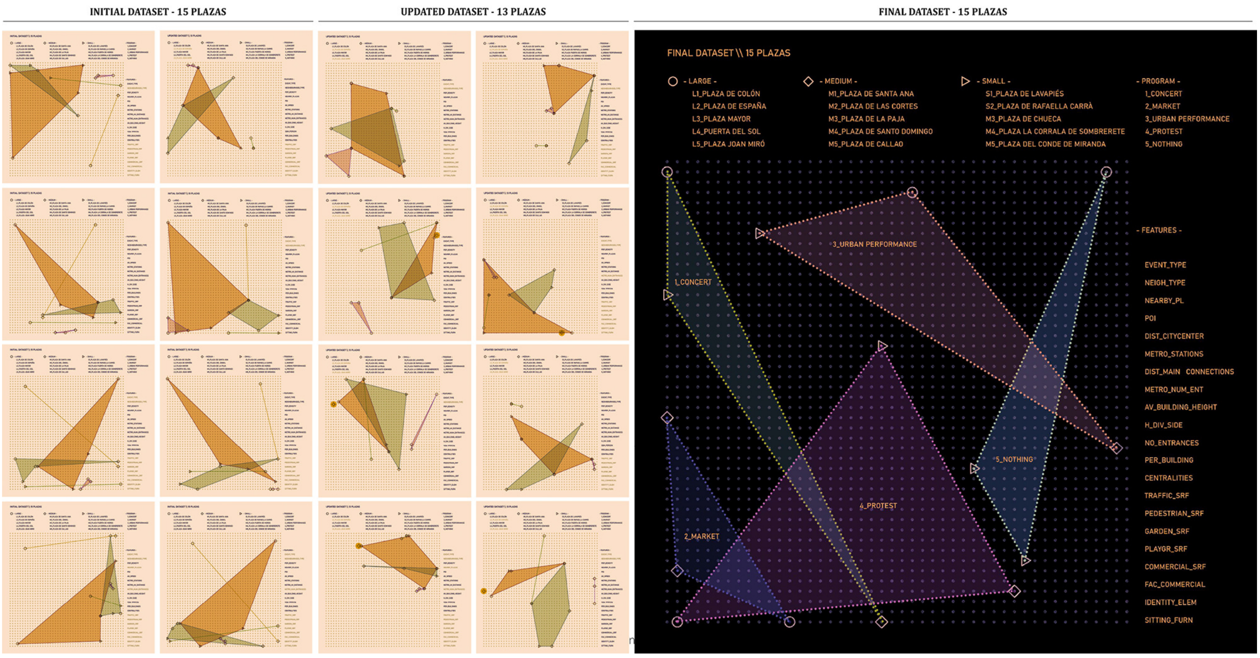

Figure 4. Results of the tests performed using the Kohonen Maps plugin. This figure shows the positions of the initial seeds relative to each program. On the left are different feature configurations with 15 and 13 plazas. On the right, the SOM displays the final dataset that was used after the data augmentation process. The x-axis represents the various features of the plazas, while the y-axis indicates the different event types predicted by the model. From the results, we observe that there is no significant correlation between the size of the plaza and the suitability of an event happening. Instead, the suitability for events appears to be influenced more by the combination of external factors, geometric shapes, and design features. This analysis highlights the complexity of predicting event suitability, suggesting that a multifaceted approach considering various urban features is necessary for accurate predictions. Source: Authors' creation from Grasshopper.

4.2. Results From the Dataset Analysis

The Correlation Matrix Heatmap showed the covariance between the different metrics and, in the case of observing a high proportionality (direct or indirect), we analysed whether this covariance implied causality or was due to the effect of a third variable, such as the type of event. For instance, we observed a high proportionality between the number of nearby plazas and the number of POIs (*nearby_pl* & *poi*). An exclusive heatmap for each type of event revealed that both variables present covariance for each output case, so we simplified our model by reducing one of these parameters. In the individual analysis by event type, some relationship was observed between other pairs of features. Still, its relevance for the model was discarded since this covariance did not appear in the overall matrix.

In our initial PCA of the input variables (Jolliffe, 2004), we observed that while most indicators contribute similarly to the overall prediction of event suitability when considering all event types collectively, there are distinct patterns when examining each event type individually. In particular, certain events are more strongly associated with specific parameters. For example, concerts are often related to larger plaza sizes and proximity to transportation hubs, while markets are more influenced by pedestrian traffic and the availability of amenities such as shops and cafes. Figure 5 illustrates these relationships by showing the loadings of each indicator on the principal components, highlighting which parameters are most influential for different event types.

In the relationship matrix between the features and the PCA analysis we observed that, for the first three iterations, none of our parameters stood out as highly determinant for the prediction of our target, which reveals that the definition of our initial metrics included enough variability to the model (Figure 5).

When graphically analysing the distribution of different features in our dataset, the output (Figure 6) reveals two challenges for the model training that the SOM induced in the dataset:

- Firstly, most classes overlap significantly throughout the different features. There is no unique feature that helps the classification.
- This is aggravated by the fact that within the same category there are gaps between the data points.

One of the main conclusions we drew from the synthetic data augmentation process, beyond the importance of defining the initial features and how to measure them, was the crucial role of urban data analysts in interpreting whether the dataset is valid and adequately represents the problem to be addressed. The researchers in charge of this process must be knowledgeable about the research topic and understand what they can do to curate a satisfactory dataset. This suggests that for most urban studies that could be investigated using this methodology, the best way to proceed is to create a specific dataset for each. Attempting to create a generalised one may not be satisfactory for all cases.

During the data augmentation process, the SOM algorithm could spawn unrealistic values that hinder learning. These data points tend to be located in the perimeters of the SOM far away from the initial seeds. In these areas, the vector interpolation of the algorithm could generate values outside of the boundaries set by the initial seeds. After achieving a satisfactory SOM result, the dataset needs to be carefully reviewed and the unrealistic data points excluded.

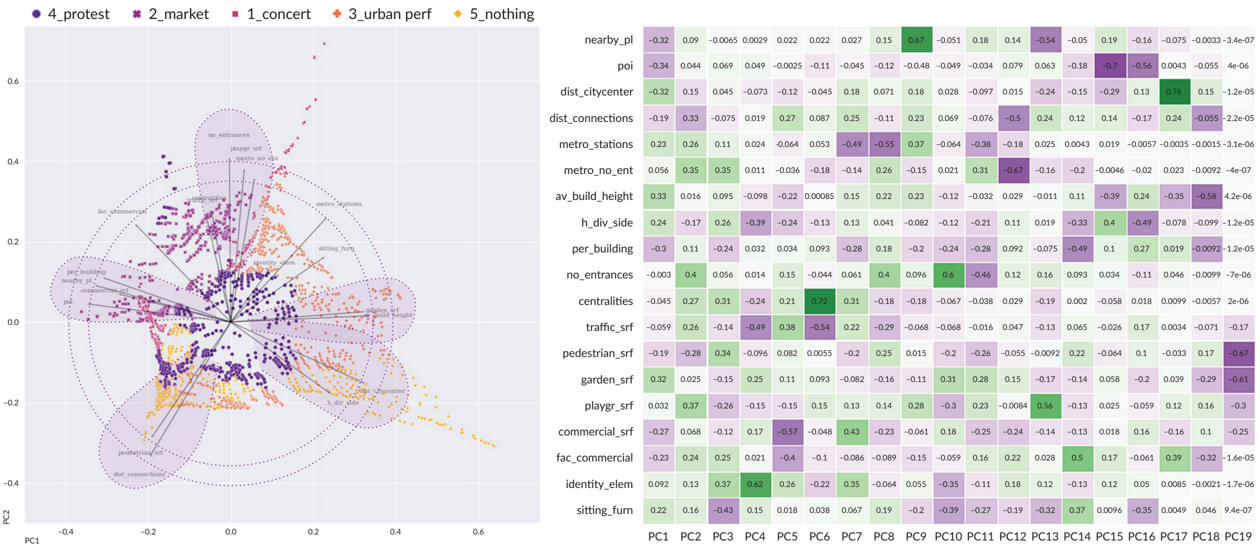

Figure 5. Results of the statistical analysis. On the left is the two-dimensional graph of the first two principal components, mapping the relationship between the output (coloured symbols) and the inputs (text). On the right, the relationship matrix illustrates the proportionality among each input (row) and each PCA analysis (columns) on a scale from 1.00 (green) to −1.00 (purple). The visuals indicate that there is no significant correlation between the size of a plaza and the suitability of an event occurring. Source: Authors' creation from Matplotlib and Seaborn Library.

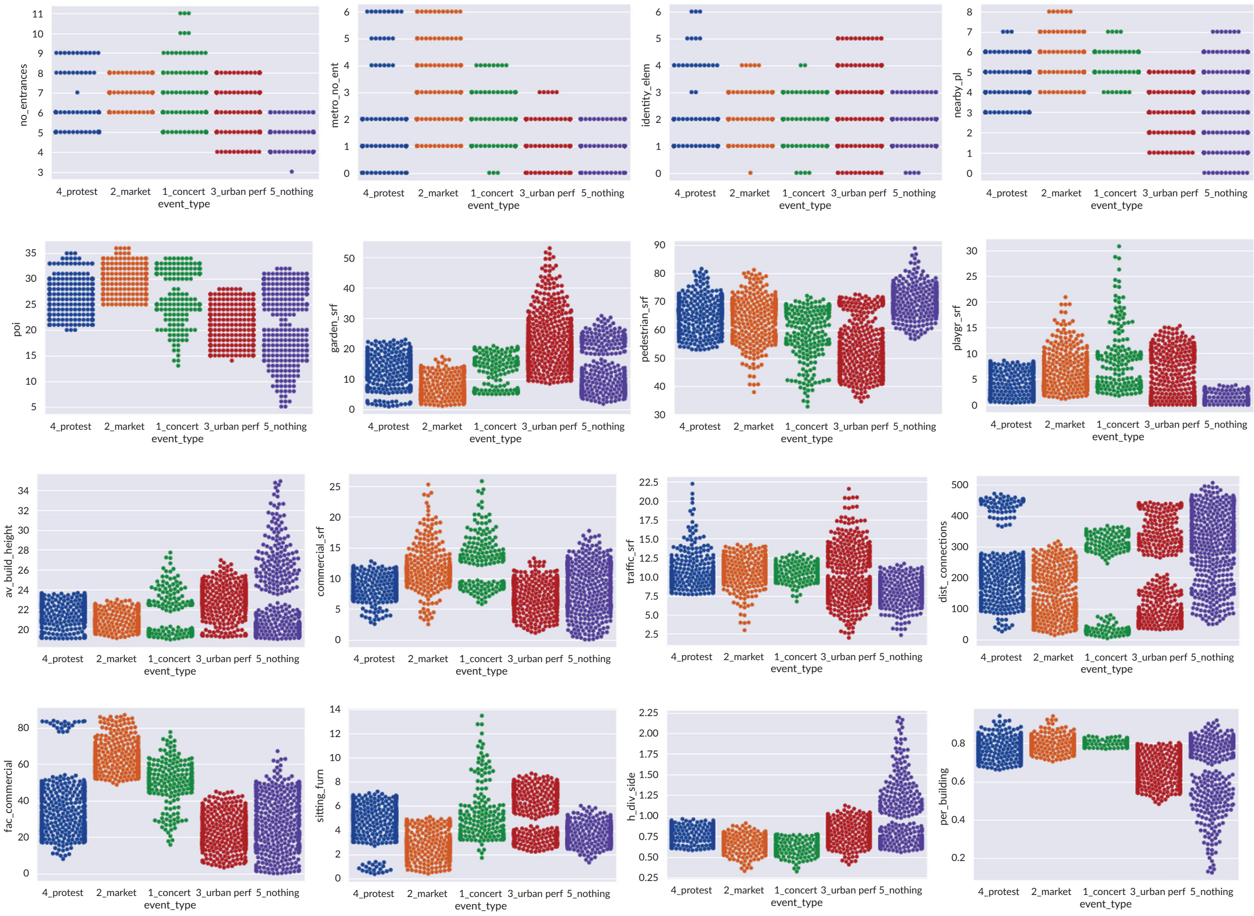

Figure 6. The exploratory data analysis of 15 out of the 20 features shows a great overlap between classes throughout features and the existing gaps between data points of the same class in different features. Source: Authors' creation from Matplotlib and Seaborn Library.

4.3. Results From the Learning Process

As mentioned in Section 3.5, during the learning process we tested multiple shallow and deep learning ML models and two versions of our dataset: a raw dataset with the initial 20 features considered and a reduced one with the 14 most relevant ones following the PCA.

With the raw dataset, the three shallow learning models had a good performance, over 0.93 (Table 1). XGBoost Classifier (score: 0.978) performed better than the linear regression (score: 0.938) and the Random Forest Classifier (score: 0.968) (Table 1). Precision and recall values are high throughout the different classes, especially with XGBoost, and, in general, the class 2_market stands out over the rest (Figure 7).

Table 1. The table compares train and test accuracies between the different models.

Model	Train accuracy	Test accuracy
Logistic regression	0.93703	0.93827
Random Forest Classifier	1.0	0.96296
XGBoost Classifier	1.0	0.97778

The ANN performed better with shallower architectures and the use of Dropout layers to try to reduce overfitting. The performances, in general, were below the three shallow learning models explored and, in any case, outperformed by the two best-performing ones (Random Forest Classifier and XG Boost). The batch size hyperparameter was changed to verify that the learning process was not stuck in local minima. Still, learning did not improve. As seen in Figure 7, all architectures tend to overfit after a few epochs, meaning that the complexity of the ANN appears excessive with the current features.

Good accuracy and low loss were considered in choosing the best-performing deep learning model. It is important not only to label the plazas adequately but also, if failing on the labelling, to not have them labelled with a program that did not suit them. The selected architecture is composed of a Dense layer of 64 neurons (activation function: relu), a Dropout layer, a Dense layer of 32 neurons (activation function: relu), and lastly, a Dense layer of five neurons (activation function: softmax), one per class.

When using the PCA Reduced dataset, the three shallow learning models and the best-performing ANN performed worse than the original dataset, meaning that the considered features were all relevant to the problem definition.

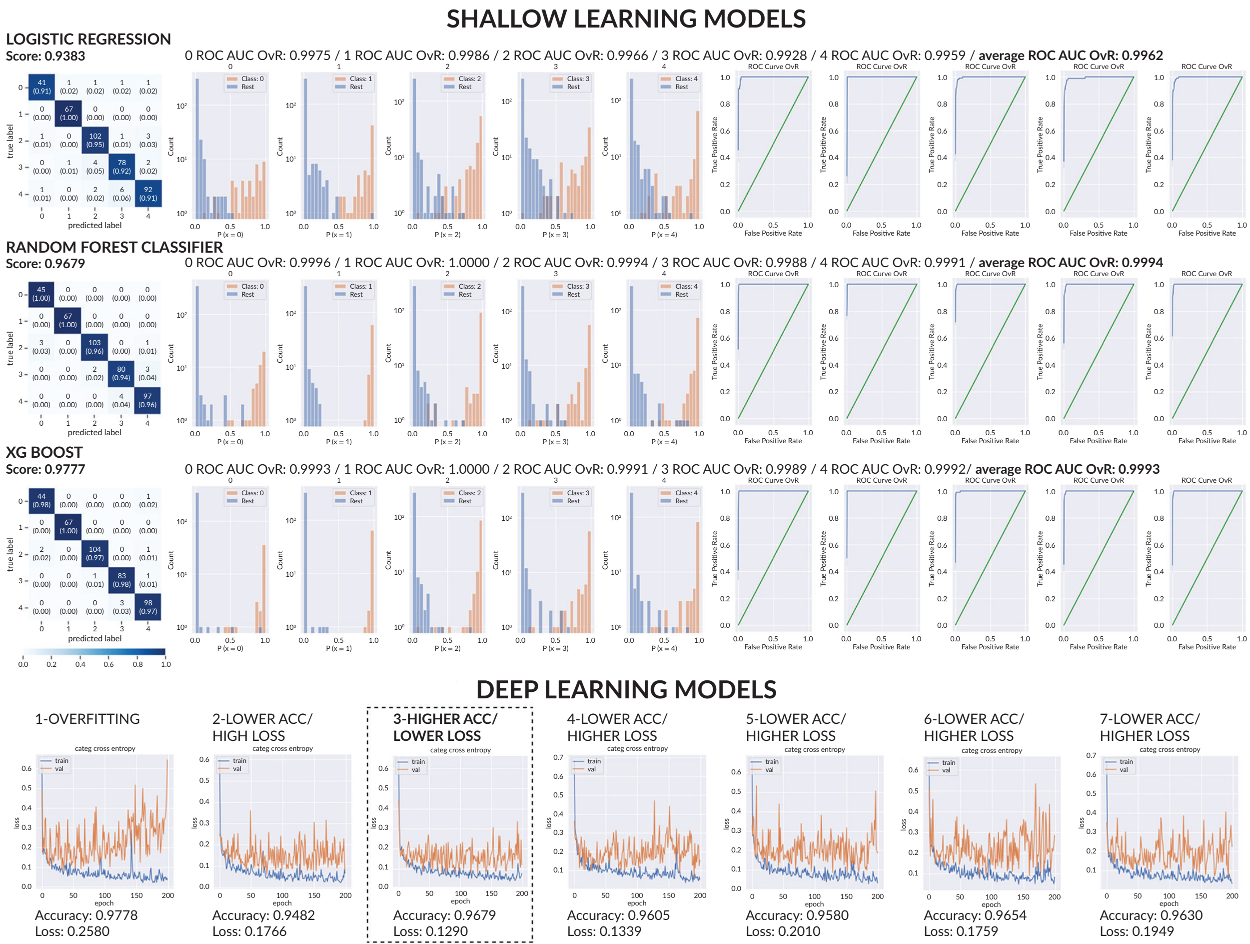

Figure 7. This figure compares both types of ML models used—shallow and deep. As mentioned, the three shallow learning models used—the upper part of the image—perform better than the different ANN architectures used. For the shallow learning models, the first two classes (1_concert and 2_market) are more clearly depicted than the rest. The deep learning models tend to overfit.

The three shallow learning models seemed to have greater potential and were further explored to understand their confidence when labelling the different data points. The Area Under the Receiver Operating Characteristics (AUC-ROC) was used following a One vs. Rest approach. As seen in Figure 7, the first two classes—1_concert and 2_market—are more appropriately depicted with the selected attributes, as there is a clear distinction between positives and negatives. In the remaining classes—3_urban performance, 4_protest, and 5_nothing—there is a greater overlapping threshold, and in the last two the model's confidence drops. It is also worth mentioning how the best-performing model, XGBoost, has a lower AUC-ROC average than the Random Forest Classifier model.

Our experiments showed that shallow learning models tend to overperform ANNs when training on our synthetic dataset. This is because they are easier to tune and can pick up the inherent patterns of our data, given our limited number of data points and features. On the other hand, deep learning models tend to overfit the sample space due to their high complexity and large number of parameters. Unlike shallow learning models, which are simpler and more constrained, deep learning models can learn intricate patterns and noise in the training data. The model that performed better was the XG Boost.

To finalise the process of data gathering, dataset curation, and ML model training, we deployed the XG Boost model in a designer-friendly interface like Rhino (Robert McNeel & Associates, 2022) via Grasshopper using the Hops component. This tool outputs the predicted class given the necessary information, plaza coordinates, and the features previously explained (Figure 8).

This presented a double use. First, the research team can easily input existing plazas to test the model's accuracy. It has the potential to allow other designers with knowledge from certain public squares to do the

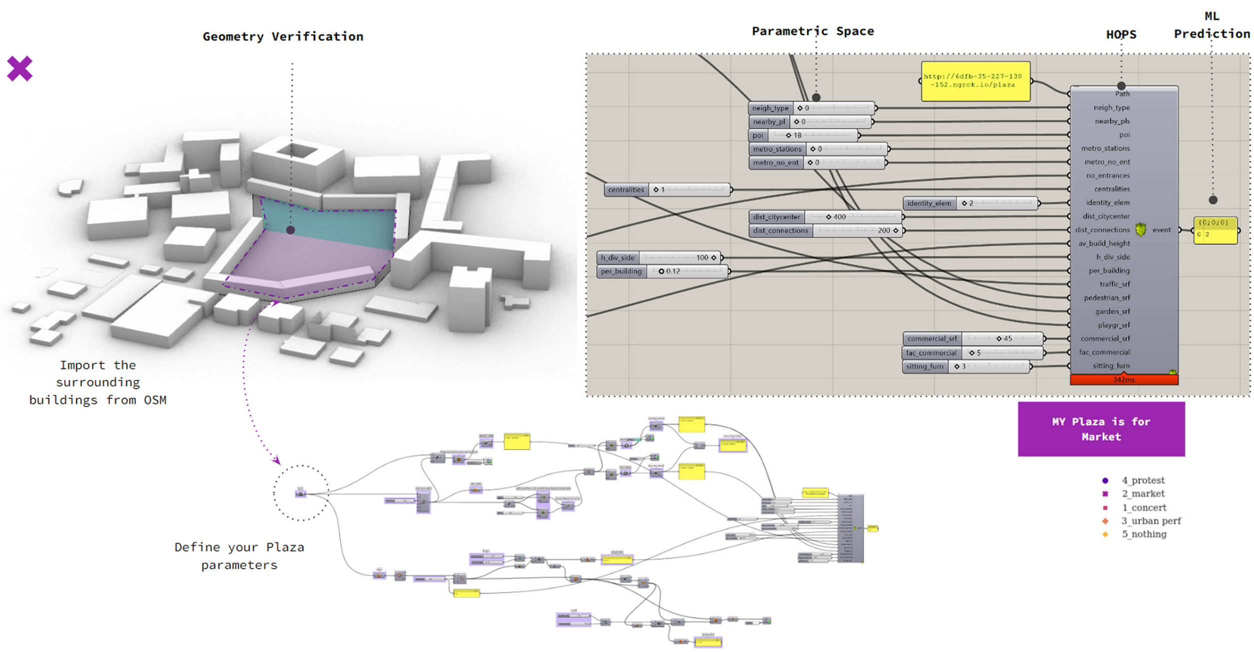

Figure 8. The Hops component from Grasshopper allows the user to predict the use of a plaza just by inputting the necessary features that can also be extracted from Rhino geometry. Source: Authors' creation from Grasshopper.

same and help improve the process. Second, it is the first step to making a tool for preliminary urban studies during the initial design stages of a project.

The model was tested in nine real plazas which do not stem from our synthetic dataset. Having created the dataset exclusively from plazas in Madrid, a poorer performance was expected the further away we moved from that city, especially from Mediterranean European plazas, as they share cultural, social, and historical backgrounds as well as climatic and physical similarities.

We deployed the model in nine plazas worldwide grouped by their geographical proximity to Madrid. These were:

Spanish:

- Plaza de la Corredera, Córdoba, Spain—37.8836° N, 4.7746° W.
  Approximate distance to Madrid: 300 km.
- Plaça de Catalunya, Barcelona, Spain—41°23′12″ N, 2°10′12″.
  Approximate distance to Madrid: 510 km.

European Mediterranean coast:

- La Place Masséna, Nice, France—43.6977° N, 7.2703° E.
  Approximate distance to Madrid: 980 km.
- Piazza del Duomo, Milan, Italy—45.4642° N, 9.1897° E.
  Approximate distance to Madrid: 1,190 km.
- Klafthmonos square, Athens, Greece—37.9794° N, 23.7311° E.
  Approximate distance to Madrid: 2,380 km.

European:

- Platz der Republik, Berlin, Germany—52.5186° N, 13.3732° E.
  Approximate distance to Madrid: 1,870 km.
- Sergels Torg, Stockholm, Sweden—59.3324° N, 18.0645° E.
  Approximate distance to Madrid: 2,600 km.

International:

- Union Square, San Francisco, USA—37.7879° N, 122.4075° W.
  Approximate distance to Madrid: 9,340 km.
- Hachiko Square, Tokyo, Japan—35° 39’ 32.7168’ ’ N, 139° 42’ 1.6920’ ’ E.
  Approximate distance to Madrid: 10,790 km.

The main challenge during this process was finding a unique program for each square. The selected uses were chosen according to the news available for each one of them. Even though this method is not reliable for validation, the accuracy shift is observed when applying the model to plazas outside Madrid, as seen in Figure 9. This can be attributed to differences in cultural, social, and environmental factors. For example, plazas in Northern Europe may have different usage patterns and climatic conditions compared to those in Southern Europe, affecting the model’s performance. Further tests should be conducted to support these results, and more data should be collected from diverse geographical locations to improve the model’s generalisability.

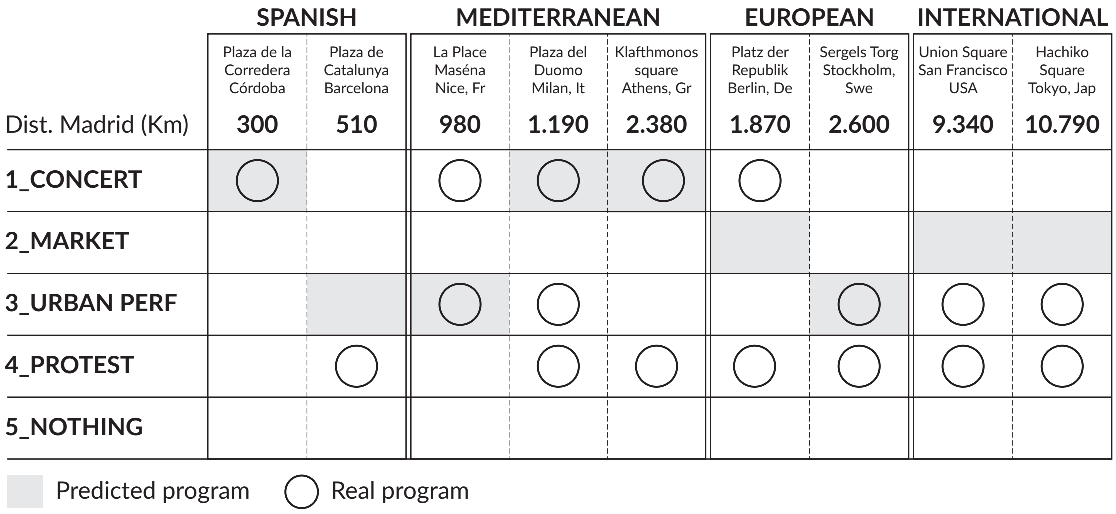

Figure 9. The figure shows the real plazas grouped by their geographical relationship to Madrid, the base of the dataset. The coloured rectangles are the predicted label by the XGBoost model. In contrast, the circles show the relevant programs on the squares.

In addition, the multitude of programs for which information was found highlights the difficulty of associating a single label with public space.

5. Conclusions

This study presents an initial exploration into the application of ML for predicting events in urban plazas, demonstrating the potential of data-driven approaches in urban planning. While our study primarily focuses on the application of ML for event prediction, it also touches upon the design and usability of urban plazas. However, due to space constraints, the latter aspect is briefly addressed and will be the focus of future research.

As an initial approach, the results have been satisfactory. However, the methods chosen for the dataset generation and the strategies for solving the issues during its curation are characterised by the team's common cultural background and similar urban experiences. This raises awareness of the possible bias that might have been induced while numerically encoding complex urban features.

The aforementioned bias could have been caused by using plazas solely from Madrid as seeds for the SOM, which already skewed the dataset and feature selection toward European Mediterranean cities. A possible solution to overcome this problem could be to use more seeds from different geographical areas for the SOM generation. However, as presented in Section 4.2, people's urban experience has a cultural component, and encoding these social nuances might overcomplicate the dataset curation and hinder the model training. In this sense, the research team believes that having local urban designer teams with multiple social backgrounds generating bespoke datasets for different geographic areas that share common cultural contexts could reduce complexity and allow the use of simpler ML models.

As mentioned in Section 3.3, after using the SOM relative mapping algorithm, there was a manual input process in which some data was refined and included *a posteriori* using the Euclidean distance from data points to seeds.

Depending on the feature selection used for creating the SOM, the resulting data distributions varied greatly, highlighting the complexity of including the right metrics to analyse public spaces without oversimplifying the parameters that define them. This questions the adequacy of this type of data augmentation algorithm for this research, and the feasibility of other methods needs to be further studied.

It was also mentioned that data augmentation processes might generate unrealistic values in the dataset that hinder learning and that need to be removed. Further studies should be carried out to look for more suitable data augmentation processes.

Concerning the ML models, two main conclusions are drawn. Firstly, from a theoretical perspective, the beauty of urban spaces is the diversity of uses that can take place on different dates or share the same time span. A multiclass classification model capable of outputting a single class per plaza is a rather simplistic way of understanding urban life, even though it sets a promising initial step. A multiclass multi-label classification model, capable of outputting various suitable scenarios for a given urban space, is a promising further step to improve the dataset curation and model selection.

Future steps include adding more features that better depict the different classes, especially the underperforming ones, 3_urban performance, 4_protest, and 5_nothing, as mentioned in Section 3.5. These features could also include more complex and intangible characteristics like sentiment analysis that could numerically encode a subjective yet shared opinion over a given space. This would mean the inclusion of analysis parameters that characterise the mental image of the public space based on an identity hierarchy established by the users.

Finally, the generation of a validation dataset made from real plazas poses two significant challenges. Firstly, assigning a unique class to an existing public square is complicated, and gathering sufficient reliable data would be extremely time-consuming and complex as the information is not readily available. Secondly, the class selection might be biased by the urban experiences of the individuals tagging the space. With these two drawbacks, the validity of the dataset could be easily compromised if not periodically verified. Therefore, we acknowledge the limitations of our work, particularly the small dataset size and the methodological challenges in data augmentation. Future research should aim to collect more comprehensive datasets, possibly through collaborative urban data initiatives, to enhance model accuracy.

To conclude, this article presents a methodology that opens up a new branch of research in the urban studies field: event suitability prediction for urban spaces. The different phases proposed, from data collection to model training and deployment, imply a series of challenges that will have to be solved for each specific application case: data scarcity, tedious dataset curation, model overfitting, etc. However, the greatest difficulty lies in encoding physical spaces that host unexpected and changing events with measurable values. In this regard, we believe this methodology requires local knowledge brought in by urban designers to reach its full potential. All stages of the process must be closely followed by designers knowledgeable about the research topic and capable of detecting possible failures or biases throughout. This combination of ML and local knowledge has the potential to strengthen the impact of urban studies.

Acknowledgments

We would like to thank the School of Information Systems and Technology Management, UNSW Sydney for their support.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The data in this article is available at the Institute for Advanced Architecture of Catalonia (IAAC).

References

Achrekar, H., Gandhe, A., Lazarus, R., Yu, S.-H., & Liu, B. (2011). Predicting flu trends using Twitter data. In *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 702–707). IEEE. https://doi.org/10.1109/INFCOMW.2011.5928903

Appleyard, D. (1981). *Liveable streets*. University of California Press.

Brighenti, A. (2010). *The publicness of public space: On the public domain*. Università di Trento. http://eprints.biblio.unitn.it/1844

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. arXiv. https://doi.org/10.48550/ARXIV.1603.02754

Dencik, L., Hintz, A., & Carey, Z. (2018). Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom. *New Media & Society*, *20*(4), 1433–1450. https://doi.org/10.1177/1461444817697722

DeVries, T., & Taylor, G. W. (2017). *Dataset augmentation in feature space*. arXiv. https://doi.org/10.48550/ARXIV.1702.05538

Food4Rhino. (2021). [Computer software].

Gehl, J. (2011). *Life between buildings: Using public space*. Island Press.

Higueras García, E., Llés Lazo, C., Rodríguez García, M. A., Biencinto, J., & Lage Gómez, P. (2017). *Paisaje urbano, necesidades y funcionalidad del sistema de plazas del entorno de Sol de Madrid*. UPM. https://oa.upm.es/81319

Huang, J., Peng, M., Wang, H., Cao, J., Gao, W., & Zhang, X. (2017). A probabilistic method for emerging topic tracking in Microblog stream. *World Wide Web*, *20*(2), 325–350. https://doi.org/10.1007/s11280-016-0390-4

Jin, F., Wang, W., Zhao, L., Dougherty, E., Cao, Y., Lu, C.-T., & Ramakrishnan, N. (2014). Misinformation propagation in the age of Twitter. *Computer*, *47*(12), 90–94. https://doi.org/10.1109/MC.2014.361

Jolliffe, I. T. (2004). *Principal component analysis* (2nd ed.). Springer.

Kohonen, T. (1995). *Self-organizing maps* (Vol. 30). Springer. https://doi.org/10.1007/978-3-642-97610-0

Kursuncu, U., Gaur, M., Lokala, U., Thirunarayan, K., Sheth, A., & Arpinar, I. B. (2018). *Predictive analysis on Twitter: Techniques and applications*. arXiv. https://doi.org/10.48550/ARXIV.1806.02377

Lynch, K. (1964). *The image of the city*. MIT Press. https://mitpress.mit.edu/9780262620017/the-image-of-the-city

Mukhina, K. D., Visheratin, A. A., & Nasonov, D. (2019). Urban events prediction via convolutional neural networks and Instagram data. *Procedia Computer Science*, *156*, 176–184. https://doi.org/10.1016/j.procs.2019.08.193

Page, S., & Connell, J. (Eds.). (2023). *The Routledge handbook of events* (2nd ed.). Routledge.

Ramakrishnan, N., Butler, P., Muthiah, S., Self, N., Khandpur, R., Saraf, P., Wang, W., Cadena, J., Vullikanti, A., Korkmaz, G., Kuhlman, C., Marathe, A., Zhao, L., Hua, T., Chen, F., Lu, C.-T., Huang, B., Srinivasan, A., Trinh, K., . . . Mares, D. (2014). *"Beating the news" with EMBERS: Forecasting civil unrest using open source indicators*. arXiv. http://arxiv.org/abs/1402.7035

Richards, G., & Colombo, A. (2017). Rethinking the eventful city: Introduction. *Event Management*, *21*(5), 527–531. https://doi.org/10.3727/152599517X15053272358997

Richards, G., & Palmer, R. (2010). *Eventful cities*. Routledge. https://doi.org/10.4324/9780080940960

Robert McNeel & Associates. (2022). *Rhinoceros* (Version 7) [Computer software]. https://www.rhino3d.com

Rumi, S. K., Deng, K., & Salim, F. D. (2018). Crime event prediction with dynamic features. *EPJ Data Science*, *7*(1), Article 43. https://doi.org/10.1140/epjds/s13688-018-0171-7

Smith, A. (2012). *Events and urban regeneration: The strategic use of events to revitalise cities*. Routledge.

Smith, A. (2017). *Events in the city: Using public spaces as event venues*. Routledge.

Wang, X., Gerber, M. S., & Brown, D. E. (2012). Automatic crime prediction using events extracted from Twitter posts. In S. J. Yang, A. M. Greenberg, & M. Endsley (Eds.), *Social computing, behavioral-cultural modeling and prediction* (Vol. 7227, pp. 231–238). Springer. https://doi.org/10.1007/978-3-642-29047-3_28

Wirth, L. (1938). Urbanism as a way of life. *American Journal of Sociology*, *44*(1), 1–24. https://doi.org/10.1086/217913

Zhao, L., Chen, F., Lu, C.-T., & Ramakrishnan, N. (2016). Multi-resolution spatial event forecasting in social media. In *2016 IEEE 16th International Conference on Data Mining (ICDM)* (pp. 689–698). IEEE. https://doi.org/10.1109/ICDM.2016.0080

Zhao, L., Chen, J., Chen, F., Wang, W., Lu, C.-T., & Ramakrishnan, N. (2015). SimNest: Social media nested epidemic simulation via online semi-supervised deep learning. In *2015 IEEE International Conference on Data Mining* (pp. 639–648). IEEE. https://doi.org/10.1109/ICDM.2015.39

About the Authors

Jumana Hamdani is an architect and a researcher at the Blekinge Institute of Technology, Spatial Planning Department. She is a coordinator at the SESA Project: Smart Energy Solutions for Africa and a PhD candidate at the ARC Centre Arch_Manu and the School of Information Systems, UNSW Sydney. Her design science research focuses on AI and digital sustainability in cities, the AEC sector, and related business models. Over the past two decades, she has designed several projects in Asia, Africa, and Europe. She is a member of the Swedish Association of Architects and the Chamber of Architects Baden-Württemberg in Germany.

Pablo Antuña Molina is a computational designer with experience in international architecture studios like UNStudio, BIG, and OMA. His multifaceted role encompasses various fields of computational design applied to architecture, from model analysis to data management and documentation. He is part of the Faculty Team in the Advanced Computation for Environmental and Structural Design module at the MaCAD programme of the Institute of Advanced Architecture of Catalonia.

Lucía Leva Fuentes is a computational designer and BIM manager at CCA, specialising in generative design and digital twins for efficient construction. She earned her Master's in Advanced Computation for Architecture & Design in 2022, focusing on graph theory for sustainable urban design. She is part of the Faculty Team in the Advanced Computation for Environmental and Structural Design module at IAAC's MaCAD programme.

Hesham Shawqy is a computational design specialist at Grimshaw, London, focusing on computational design, artificial intelligence, and web development. He also teaches the Complex Forming seminar at the MaCAD programme of the Institute of Advanced Architecture of Catalonia. He earned his Master's degree in Advanced Computation for Architecture & Design from IAAC in 2021, with research centred on digitising handcrafts using ML and robotic fabrication.

Gabriella Rossi is a doctoral fellow at CITA, Royal Danish Academy. Her research focuses on ML as an emerging modelling paradigm, for AEC modelling and fabrication, with a focus on architectural datasets. She teaches the AI and Architecture module at the MaCAD programme of the Institute of Advanced Architecture of Catalonia both as a thesis supervisor and as a lead for the Data Encoding & ML seminar. She is a member of the Danish Association of Architects.

David Andrés León is a private consultant for digital technologies for the AEC industry with a particular focus on web development and user interfaces for AI. His goal is to empower designers with digital tools to leverage the design process. He is also currently director of the Master of Advanced Computation for Architecture and Design (MaCAD) at the Institute for Advanced Architecture of Catalonia (IAAC), a member of the faculty of the Norman Foster Foundation and the Universitat Politècnica de Catalunya (UPC). He has been involved in various research collaborations and co-authored numerous publications in the field of architectural research.