Article

Quantifying Liveability Using Survey Analysis and Machine Learning Model

Vijayaraghavan Sujatha [1],* , Ganesan Lavanya [1] and Ramaiah Prakash [2]

1   Department of Civil Engineering, Anna University—University College of Engineering,
    Ramanathapuram 623513, India
2   Department of Civil Engineering, Alagappa Chettiar Government College of Engineering and Technology,
    Karaikudi 630003, India
*   Correspondence: suja.3051@gmail.com

Abstract: Liveability is an abstract concept with multiple definitions and interpretations. This study builds a tangible metric for liveability using responses from a user survey and uses Machine Learning (ML) to understand the importance of different factors of the metric. The study defines the liveability metric as an individual's willingness to live in their current location for the foreseeable future. Stratified random samples of the results from an online survey conducted were used for the analysis. The different factors that the residents identified as impacting their willingness to continue living in their neighborhood were defined as the "perception features" and their decision itself was defined as the "liveability feature". The survey data were then used in an ML classification model, which predicted any user's liveability feature, given their perception features. 'Shapley Scores' were then used to quantify the marginal contribution of the perception features on the liveability metric. From this study, the most important actionable features impacting the liveability of a neighborhood were identified as Safety and Access to the Internet/Organic farm products/healthcare/Public transportation. The main motivation of the study is to offer useful insights and a data-driven framework to the local administration and non-governmental organizations for building more liveable communities.

Keywords: urban planning; liveability; supervised machine learning; online user survey

1. Introduction

Liveability is an abstract concept with multiple definitions and interpretations. Liveability is the degree to which a place fulfills the expectations of its residents for their well-being and quality of life. Myers [1] mentioned that liveability could be expressed as sustainability, quality of life, the "character" of place, the health of communities, etc. Liveability is an "ensemble concept". Balsas [2] measured the 'city-center liveability' through a set of Key Performance Indicators (KPIs). The online survey for this study was designed on the perception of liveability factors, such as preserving green spaces, reduction of traffic congestion, restoring community, promoting neighboring community's collaboration, and enhancing competitiveness at an economic level. Emphasis is given to safety, affordability of housing, and transportation options as important aspects of liveability. Tyce and Rebecca [3], in their research initiative, reviewed 237 sources related to liveability and found that the three most common categories used to define liveability were Transportation, Development, and Community features.

Philips, a global leader in energy technology and electronics, analyses regional trends in liveability, health, and well-being based on five factors: employment, community, physical health, emotional health, and family/friends [4]. This survey reveals regional liveability patterns. This survey's flaw is that it cannot rank cities. In 2008, Organization for Economic Cooperation and Development (OECD) developed a worldwide project for measuring

society's progress [5]. It proposed rethinking measuring methods and initiated an international conversation on economic, environmental, and social objectives and whether they are represented in national and international metrics. This led to a better life index and United Nations (UN) commissioned happiness studies. The OECD has identified 11 themes vital to material welfare (housing, money, employment) and human welfare (life satisfaction, safety, community, education, governance, health, environment, and work-life balance). This is a major worldwide index used to assess nations' welfare and quality of life, although the metrics utilized are not disaggregated. The UN-Habitat city prosperity index measures how well cities support their populations' well-being [6]. It expands economic prosperity to include infrastructure, quality of life, equity, and environmental sustainability. The Index's data provides useful information about population status (health, education, income, etc.) and resource provision, which are important indicators of well-being and liveability. Nestor et al. [7] have explained the impact of amenities and public goods in the neighborhood on happiness and satisfaction with several life domains like health, family, social life, work, and economic situation. The first step of the life satisfaction methodology is to ask the residents how satisfied their lives are. The place diagram developed by the Project for Public Spaces (PPS) [8] provides a measurement framework for the residents to evaluate any place through refined questionnaires. Different authors use specific indicators to evaluate well-being [9,10] which is more inclined to their studies. Kim et al. [11] aimed to contribute to sustainable development in Seoul for urban planning by analyzing a questionnaire collected from an electronic platform in which citizens in the city participated.

Most of the literature reviewed defined factors of liveability and weighed them based on their subject matter expertise or provided equal weight to derive liveability. However, these methods are neither easily transferable nor replicable due to the varying relative importance of factors with different geo-cultural contexts. Hence this study provides a framework to recalibrate the weights of the factors using the wisdom of the crowd (responses of the residents collected from a survey) through an ML prediction model. The prediction model was then transformed into an explainable model using Shapley Additive exPlanations (SHAP) Values [12]. Random Forest (RF) method is used in this study which reduces the generalization error of a forest of tree classifiers [13]. According to Lundberg and Lee [14], the SHAP values help us to get clear insights into the model by understanding the reasons behind predictions. The factors associated with liveability are then used only as falsifiable and testable hypotheses.

1.1. Purpose Statement

The intent of the study is:

- To design a tangible metric (the individual's preference to live in any place is defined as a tangible metric in this study) for liveability at an individual level and make it scalable across any administrative unit (the administrative unit here defines a postal code, a village, a city, or a state);
- To understand and quantify the marginal contribution of the different factors of liveability towards the designed metric.

1.2. Research Design

This study proposes the use of a survey design with close-ended questions randomly distributed to residents of different neighborhoods in the study area. The results of the survey were used to build Machine Learning (ML) and explainable models. The process explained in Figure 1 can be broken down into the following steps:

- Hypotheses generation;
- Design & distribution of the questionnaire;
- Collection & analysis of survey data (Feature Engineering & data munging);
- Fitting ML Classification model to predict liveability;
- Using the SHAP model to quantify the marginal contribution of the different factors.
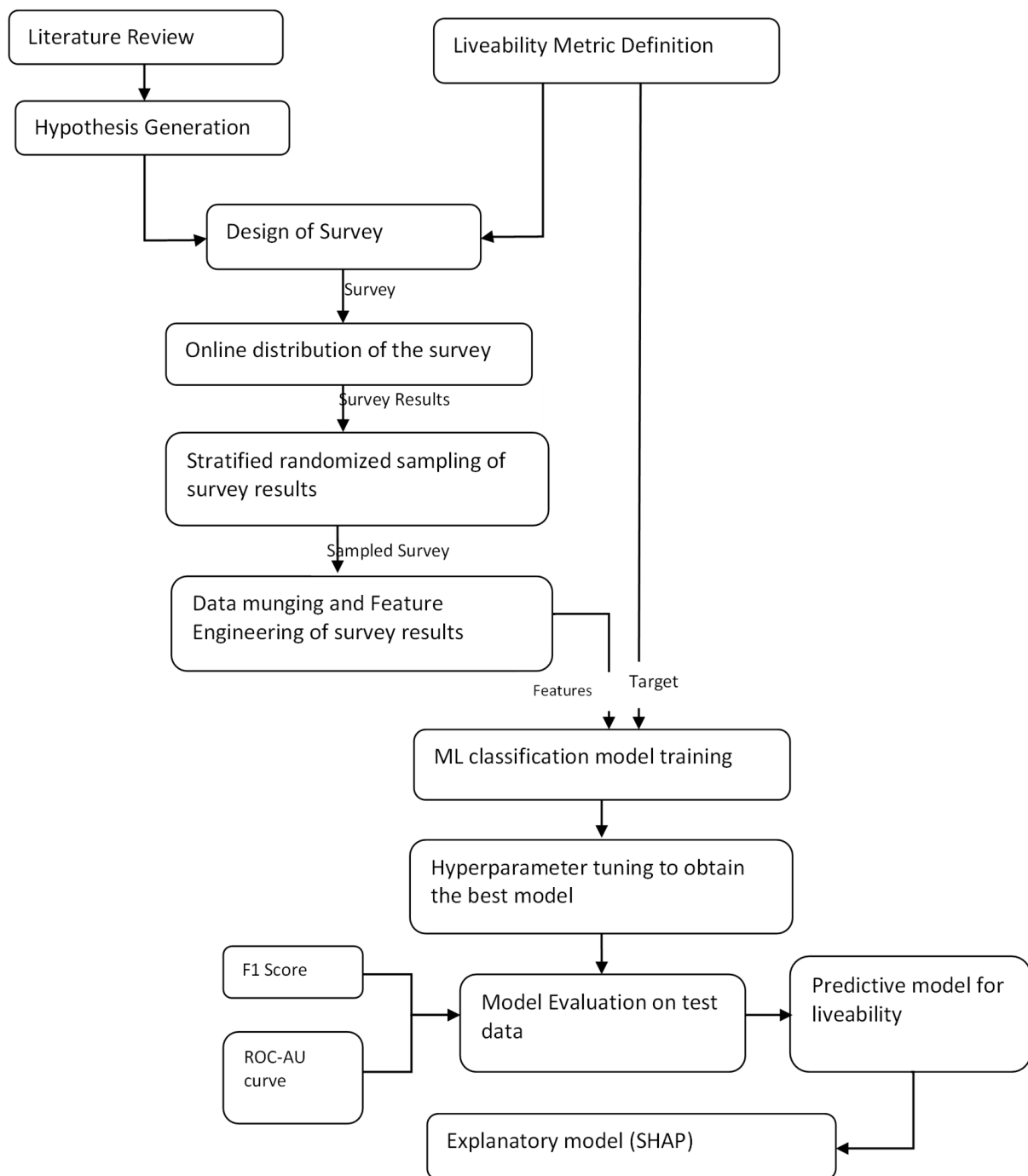
Figure 1. Research design flowchart.

2. Methodology

2.1. Hypotheses Generation

Demographics, access to transportation facilities and amenities, neighborhood characteristics, civic and social engagement, employment, and educational opportunities were hypothesized as the factors of liveability based on the literature review conducted [15], as shown in Table 1. Liveability feature was defined as the respondent's willingness to continue living in their location in the foreseeable future, to validate these hypotheses. Since the liveability feature was a binary response (yes/no), we were able to define liveability as a simple, measurable, and scalable signal, attributable to the factors of liveability.

Table 1. Liveability features derived from the responses to the questionnaire (Questionnaire Survey Link: https://forms.gle/etEuXGKQ97Ei69Re8).

Assumed Factors from Literature Reviews	Questionnaire Framework	Derived Features
Demographic identifier	• Pincode	Spatial Feature
Transportation Facilities	• Use of Public Transportation • Waiting time in Traffic	Perception Feature
Accessibility	• To Health care • To Market/grocery store • To Organic farm products	Perception Feature
Neighborhood	• No. of parks • No. of Libraries • Safety • Neighbors' dependability	Perception Feature
Ecology and Environment	• Water Quality • Air Quality • Open space/Green space	Perception Feature
Civic and Social Engagement	• Socio-economic Equality • Access to Internet • Cultural/sports/entertainment institutions • Tourist Attractions • Member of a Social Organization • Small-Scale household farming	Perception Feature
Opportunity	• Educational Opportunity • Employment Opportunity	Perception Feature
Likelihood to stay in their neighborhood	• Happy with the Neighborhood • Given an opportunity, "Will you shift?"	Liveability Metric

2.2. Design and Exploratory Analysis of the Questionnaire

The Indian state of Kerala was chosen as the site for the case study. In this study, to measure the perceptions of the residents regarding the quality of life in their city, an online survey was conducted. People who reside in the 14 districts of Kerala and were aged above 18 were randomly selected and invited to participate. A total of 3280 responses were received. Some of the residents took the survey through an online link sent to them. The rest of the respondents (about 2800 people) took the survey when they had visited the e-service centers called Akshaya centers in Kerala. These centers provide a range of government services, such as applying for a passport or paying utility bills. The survey was given to a diverse group of people, including those from different income levels and with varying levels of internet access. However, there may be biases in the survey results due to self-selection (only certain types of people may have voluntarily chosen to take the survey) and the fact that some people who may have needed to use the government services offered at Akshaya centers were not able to do so in person (for example, due to disability).

The residents were briefed the purpose of the survey and were required to provide the Pin code of their residence along with other questions which were of 2 formats. One was a Likert Scale question, on a scale of 0–5 (0 is the least likely and 5 is the most likely). The other format was a Yes or No type. A total of 3280 responses were received on which the preliminary Exploratory Data Analysis (EDA) was performed, to understand the nuances of the data better [16].

2.3. Building an ML Model

A supervised machine learning classification model uses sample data generated by a process and its known outcomes (labels) to create a model, which can predict the outcome for a future/unseen set of data [17]. Among many ML models that were used, Random Forest (RF) model was selected because most of the perception features were categorical Likert scale values, and they did not have a simple linear relationship with the target metric. Additionally, the performance values were high for this model from Table 2.

Table 2. Comparison of different ML model's Performance values.

Machine Learning Model	AUC	ACCURACY	SENSITIVITY	SPECIFICITY	F1-Measure
Random Forest	70	80.5	92.6	71.4	89.41
SVM	68.5	78.6	80	65	80.52
Decision Tree	65.6	77.6	88.3	70.5	84.7
Naïve Bayesian	69.5	70.52	77.84	75	79.5

An RF is nothing but an "ensemble" of Decision Trees [18]. With the aggregation of the results (via. majority voting of the classes) of all the unique decision tree models, the problem of overfitting the training data is overcome. In this study, the entire model was built using Python programming. There are a total of 20 features (including the target variable). Before applying the supervised learning algorithm, the following feature engineering techniques were applied to the raw data.

2.3.1. Feature Engineering & Data Munging for ML Training

The Categorical (non-numerical) values must be converted to numerical values for a machine-learning model to work. 'Yes' or 'No' type of responses for questions such as, "Are you a member of any social organization?" were coded as 1/0, respectively. The missing values (NaN) were imputed with the median of the feature column. One of the final survey questions asks the residents whether they would be ready to move to a different place (supposedly with a better liveability) if given the opportunity. The response to this question enables the inference of their willingness to stay (rephrased as "Will u Stay?") by swapping the 0s and 1s (0 = No, 1 = Yes). This was chosen as the target metric ("liveability metric") of this study.

2.3.2. Fitting ML Classification Model to predict Liveability
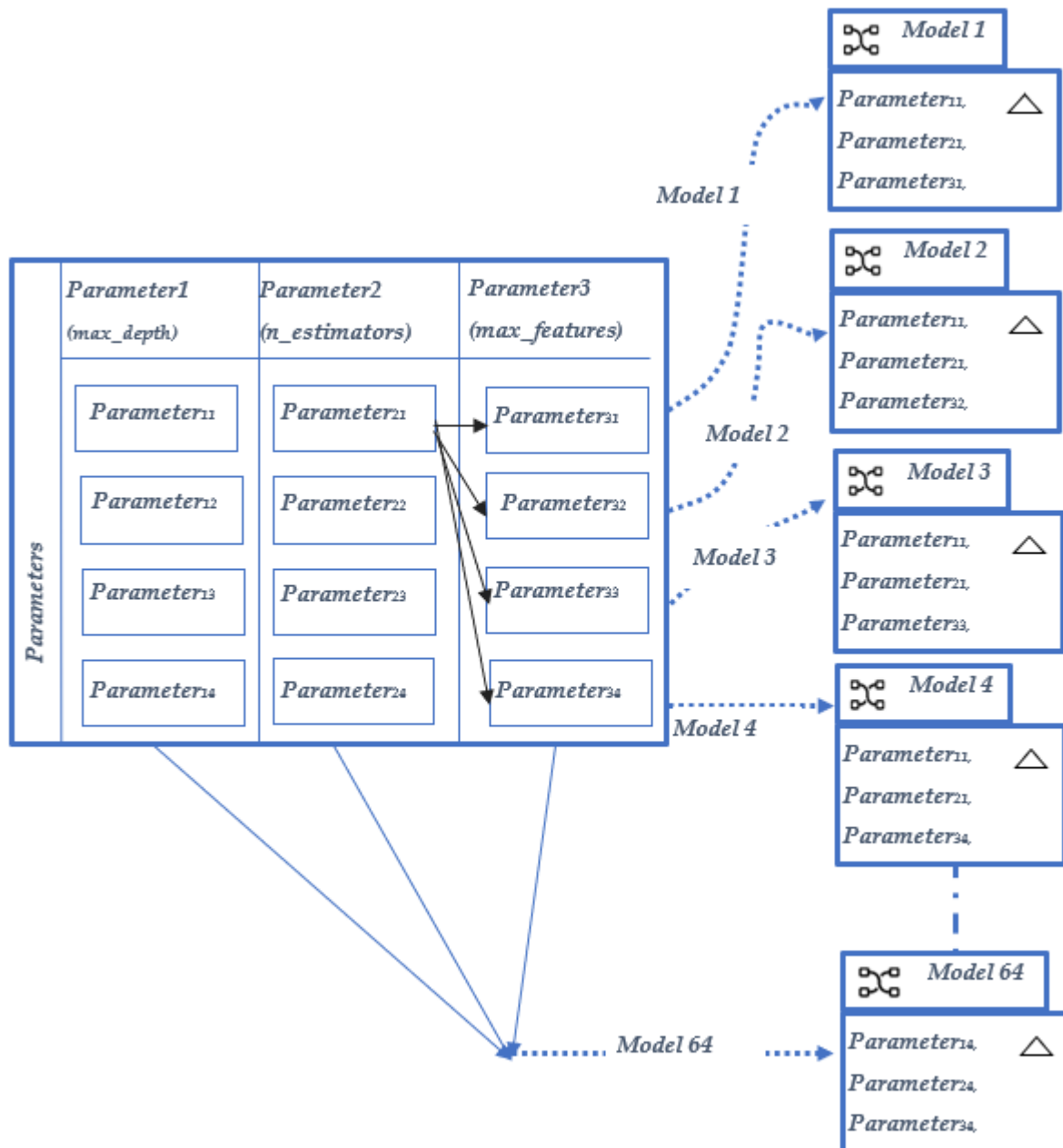Splitting the Train-Test Data

Initially, the dataset used for analysis was split into two subsets:

1. Training set: a subset from the dataset to train the model;
2. Test set: a holdout subset from the dataset to evaluate the model.

Hyperparameter Tuning

Hyperparameter tuning is the process of finding the ideal set of parameters for the model by evaluating it against a fixed criterion (such as F1 score or ROC_AUC) over a large set of parameter combinations. During the Hyperparameter tuning, the training data was equally split into k subsets [19]. The model was trained k times on k-1 unique subsets and cross-validated on the left-out subset each time. Each cross-validated model was trained on

a specific set of parameters chosen from all possible combinations of parameters as shown in Figure 2. Each hyperparameter combination yielded a model which was evaluated k times on the holdout dataset and the average of the evaluation metric (F1 & ROC-AUC) was used to choose the best parameters.

Figure 2. Finding the best parameters for the model by hyperparameter tuning.

2.3.3. Predicting the Test Data

The trained model was used to predict the liveability metric (target metric) based on the test data perception features. The predicted class and the actual class were compared to find the efficiency of the model. The interpretation of the model errors is as follows:

- False Positive (FP): (Type 1 Error) The model wrongly classified that the residents would stay when the resident indicated that they would not;
- False Negative (FN): (Type 2 Error) The model wrongly classified that the residents would not stay when the resident indicated that they would;
- The True Positives and True Negatives for the model should be inferred as follows:

- True Positive (TP): The model correctly classified that the residents would stay when the resident indicated that they would;
- True Negative (TN): The model correctly classified that the residents would not stay when the resident indicated that they would not.

There is a tradeoff between these errors. Multiple useful metrics like Precision, Recall, F1 Score, etc., can be calculated based on this tradeoff. In this model, Precision is the proportion of people who were correctly classified as they will be staying, to those that the model predicted to stay. Recall or True Positive Rate (TPR) is the proportion of people who were correctly classified as will be staying, to those that indicated to stay. To accommodate the tradeoff between Precision and Recall, a comprehensive metric such the F1 Score can be used, which is nothing but the harmonic mean of the model's Precision and Recall.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{1}$$

False Positive Rate (FPR) is the proportion of people who were wrongly classified as saying they would stay to those that indicated they would not. Yet another comprehensive metric to evaluate the performance of the model is the Receiver Operating Characteristics— Area Under the Curve (ROC-AUC). ROC handles the tradeoff between FPR and TPR [20]. The higher the AUC, the better the model is at predicting 'Will stay' classes as 'Will stay' and 'Will not stay' classes as 'Will not stay'. For this study F1 score and ROC Curve (AUC) were used as evaluation measures.

2.4. SHAP Model

While the ML Classification model is good at predicting the target metric for a given set of perception features, the model does not have a good way of explaining the impact and directionality of perception features on the liveability metric at an individual level [20]. Shapely (SHAP) values, a game-theoretical approach towards the ML model's explainability, solve this problem [21,22]. SHAP helps us to understand what decisions the model is making. SHAP trains *2^Features* models and finds out the marginal contribution of each feature toward that observation's prediction. In this study, SHAP represents the marginal contribution of each perception feature to the liveability metric (target variable).

3. Results and Discussions

3.1. Descriptive Statistics

From the summary of responses of the residents, the following observations have been tabulated in Table 3.

A high correlation was found between accessibility features, which was evident from the fact that accessibility connected to the transportation network of a place.

3.2. Random Forest Classifier Model Predictions

After the feature engineering steps, the dataset was split into Train and Test datasets. Hypertuning of parameters was performed to get the optimal model.

3.2.1. F1 Score and ROC Curve

The optimal model chosen had an F1 score of 81.41% and a ROC-AUC of 70%. The classification threshold of 0.491 (shown in Figure 3) was found to be the optimal one based on the max F1 score. This threshold yielded a better classifier which meant that only those who had a prediction probability greater than 49.1% were predicted to live in their location. The area under the curve in the ROC-AUC (Figure 4) was found to be around 70% which meant 70% of the dataset would be correctly classified in relation to their respective classes without an error.

Table 3. Descriptive statistics from the summary of responses of the residents.

Features	Description	
	Positive Responses (Approx. %) (Only Ratings 5 and 4)	**Other Responses**
Residents who use public transportation	28%	rest all used private modes of transportation
Residents who waited for a long time in traffic	14%	52% of the residents never waited in traffic.
Residents who could find an ample open space/green space for exercise/walking/jogging/cycling in their neighborhood	44%	
Residents who could feel that there is the availability of good quality drinking water and air	73%	
Residents who had access to a grocery store or a market	82%	
Residents who had access to health care services	73%	
Residents who had access to the internet	88%	
Residents who said that they have experienced perfect socioeconomic equality	26%	26% claim to experience high socioeconomic inequality
Residents who rated the availability of cultural, arts, sports, or entertainment institutions in their city	90%	
Residents who rated that they had a good or great educational opportunity in their location	60%	only 5% were lacking good educational opportunities
Residents who rated the employment opportunities in their location		80% of them felt it was not very high
Residents who rated their access to farm products	30%	
Residents who rated the safety of their neighborhood	54%	Only 7% of the respondents felt that their neighborhood is not safe
Residents who were happy with their neighborhood	94%	
Residents who wanted to continue living in their location (Liveability metric)	65%	

3.2.2. Liveability Confusion Matrix

The optimal model chosen was evaluated against the test data and the confusion matrix was plotted for the same, which provided a grid of correctly and wrongly classified data against their true labels. From Figure 5, it can be observed that among the 656 residents in the test dataset, 405 were correctly predicted to stay (TP) and 150 were correctly classified as will not stay (TN). The number of residents who were wrongly predicted to not stay (FN) was 38 and 63 residents were wrongly classified as saying they would stay (FP).
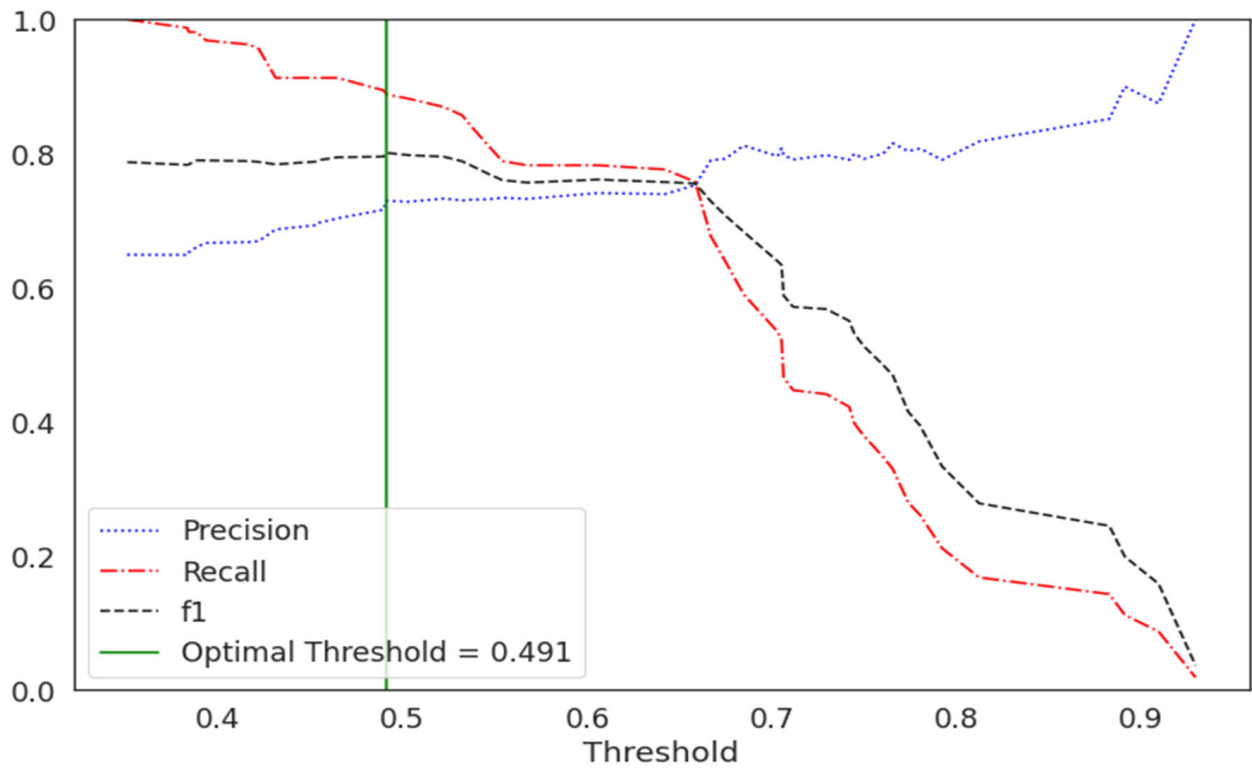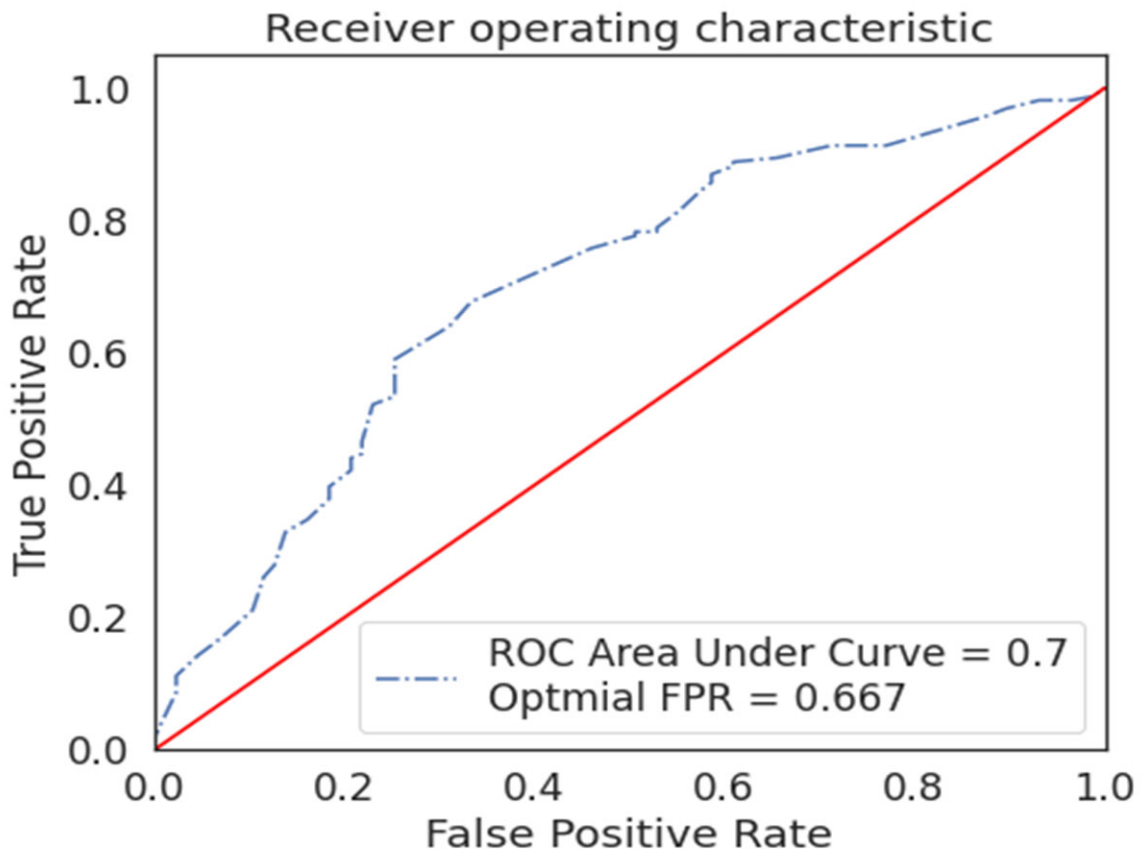
Figure 3. Precision-recall Curve.

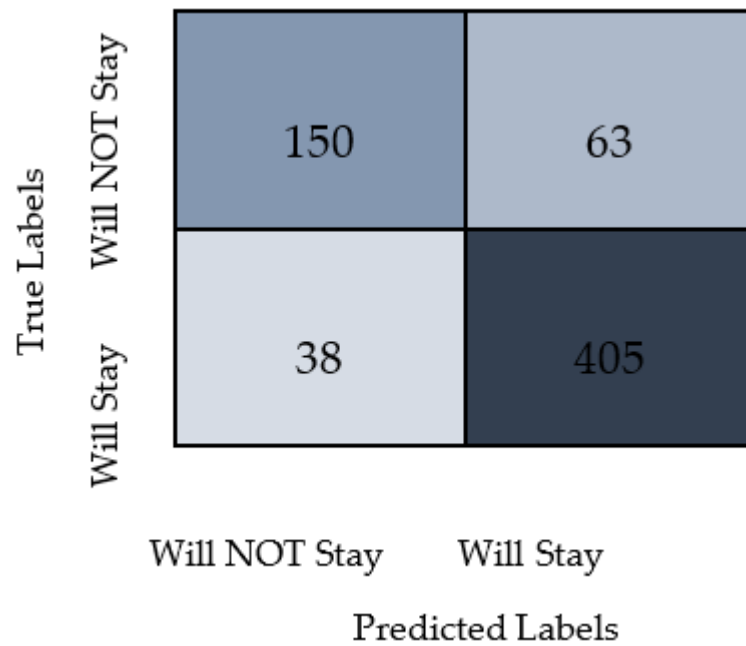**Figure 4.** ROC-area under curve.

Figure 5. Confusion matrix of the liveability.

3.2.3. Random Forest

In decision tree models, the Gini impurity is a measure of the impurity or disorder at a given node in the tree. It is used to evaluate the quality of a split, with higher values indicating a less pure or more mixed split and lower values indicating a more pure or less mixed split. The Gini impurity is calculated based on the probability of a randomly chosen element belonging to the wrong class if it was classified using the current node. If the probability is high, the Gini impurity is high, indicating a poor split. If the probability is low, the Gini impurity is low, indicating a good split. In the decision tree shown in Figure 6, the safety feature is the root node, which means it is the first feature used to split the data. This is because the Gini impurity for the safety feature is the lowest among the features considered, indicating that it is the most effective feature for separating the data into distinct classes.
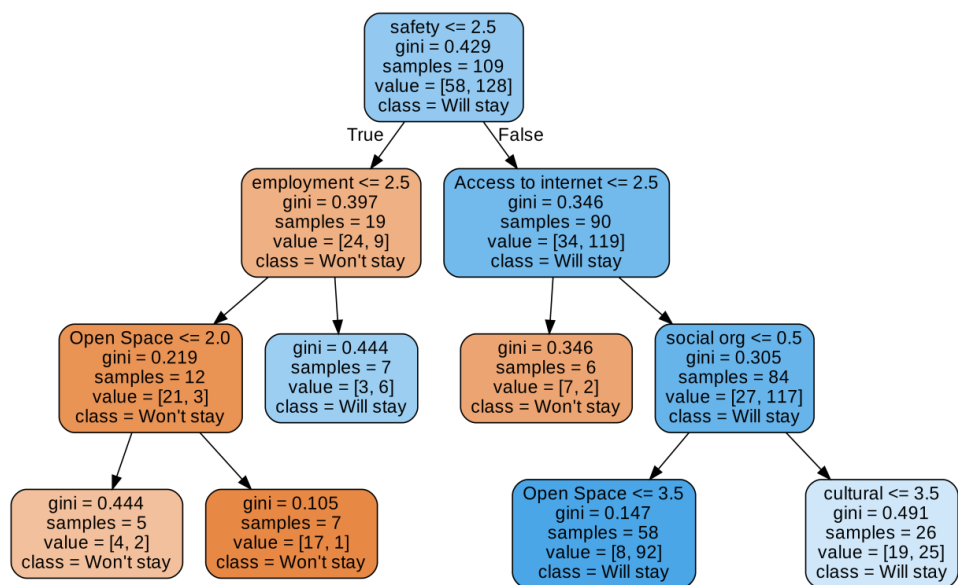

Figure 6. Decision tree.

Further split was done by a binary tree creation process and the branched nodes were further selected so that it has a lower Gini impurity than the parent node. The branching was continued till the node where it cannot be further split, and the Gini impurity cannot be smaller than the parent node. Those final nodes are called the leaves of the Decision tree. When test data is run through this decision tree model, each observation passes through the branch nodes and finally reaches a leaf node, which gives us the class of the observation (whether they will stay or not).

3.3. Feature Importance

Features such as safety, usage of public transport, traffic wait time, access to organic farm products, dependability of the neighbors, availability of cultural/arts/sports institutions, access to the internet, and access to health care came out to be the most important features per the random forest model's feature importance method as shown in Figure 7.
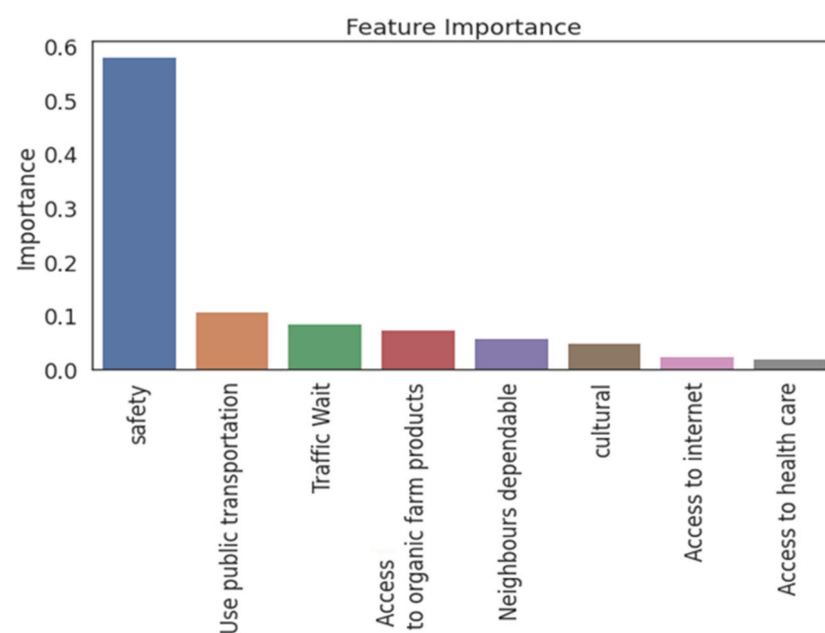
Figure 7. Feature importance.

3.4. Shapley Values and Their Analysis

The SHAP model provided the "feature importance" at the individual resident level. The 'feature importance' is nothing but the contribution of the feature values towards that resident's predicted liveability [23]. In Figure 8, the selected resident had a 66.1% predicted probability (base probability) of living in their location, in the absence of any other information. However, we know that the resident gave a rating of 5 for "safety" contributed 9% toward the final prediction probability of 71.1% from the base probability. Likewise, a rating of 2 for "traffic wait" and "access to farm products" contributed 2% and 1% toward the final prediction from the base probability. Whereas a 4 rating for "cultural" and "access to healthcare" contributed -3% and -1% each towards the final prediction from the base probability.

A partial dependence plot of a feature provides the partial dependence of the liveability metric on each distinct value of the feature. These plots provide an idea of each feature's marginal contribution for different levels of ratings. In the partial dependence plot for safety shown in Figure 9a, the x-axis shows the distinct values for the ratings provided for safety by all respondents, whereas the y-axis shows the SHAP values. The box plots in the figure show the distribution of SHAP values for each distinct rating, whereas the line plot shows the 95% confidence interval of the SHAP values for each rating, with the central point indicating the mean of the SHAP values.
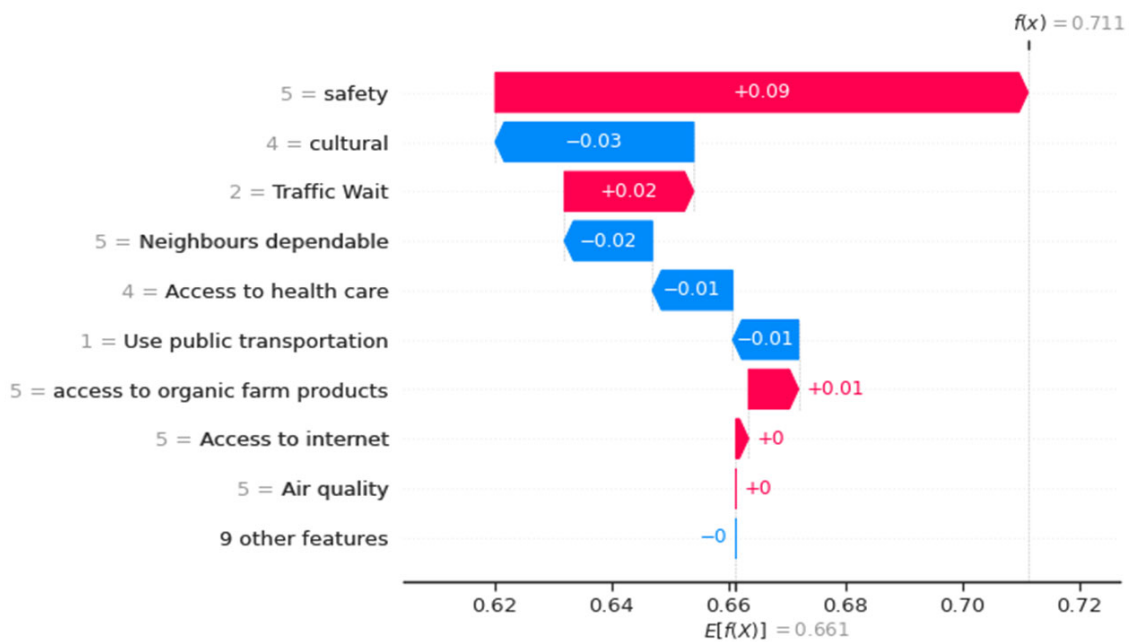
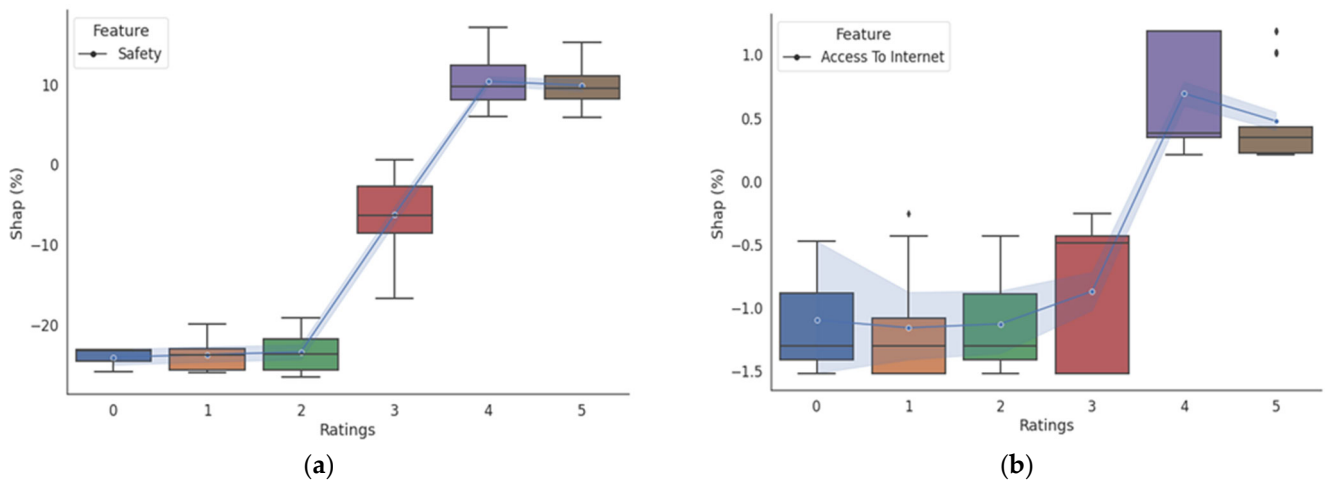**Figure 8.** Waterfall plot of the SHAP values generated for a resident.

Figure 9. Partial dependence of liveability on: (**a**) safety (**b**) access to the internet.

For the safety feature, ratings of 0 to 2 had little or insignificant change in the SHAP values. These ratings bring down the base value of liveability by around 20% points. However, there was a huge jump in SHAP values—20% points—when the rating changed from 2 to 3 and 10% points when the rating changed from 3 to 4. As shown in Figure 9b, when the ratings for access to the internet were in the range of 0 to 3, the SHAP values were in the range of −12% to −9% points. In contrast, when the ratings changed from 3 to 4, there was a huge positive leap of 16% points in the SHAP values.

For access to organic farm products and access to health care, the SHAP values shot up when the ratings changed from 3 to 4 and from 4 to 5 correspondingly, as shown in Figure 10a,b. For the usage of the public transportation feature, a different trend was observed. When the ratings increased from 2 to 3, the SHAP values rose, and when they further increased from 3 to 4, the SHAP values fell, as shown in Figure 11.
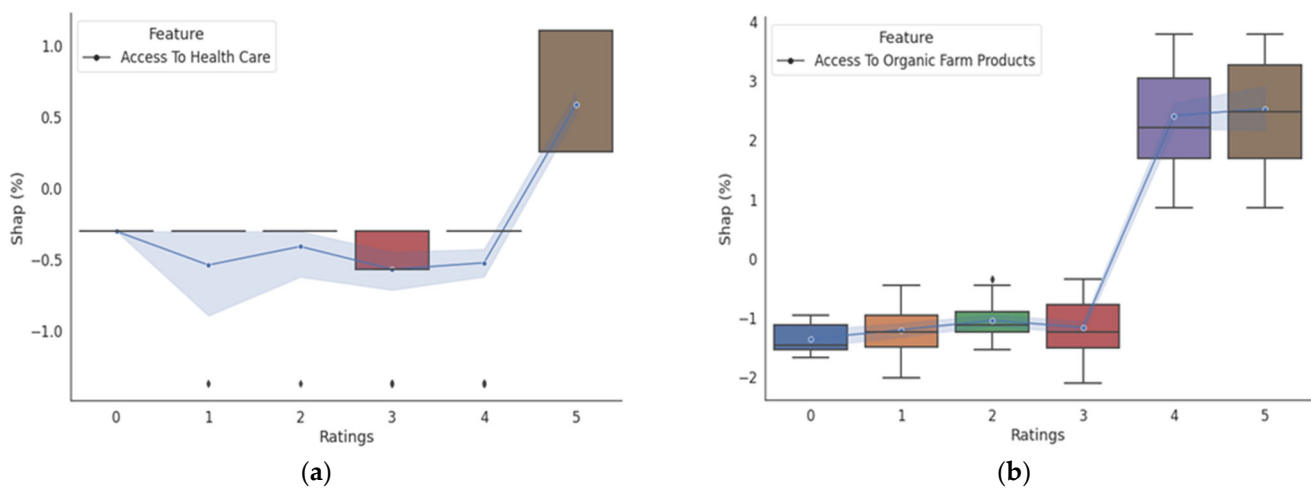
Figure 10. Partial Dependence of liveability on: (**a**) access to organic farm products, (**b**) access to healthcare.
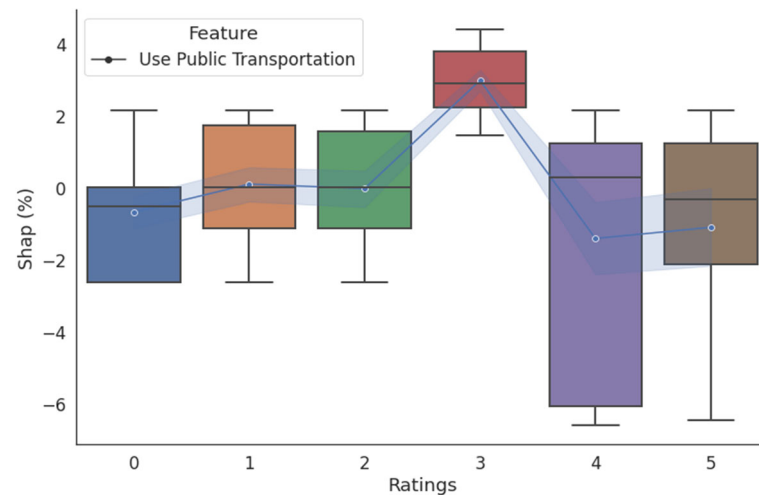
Figure 11. Partial dependence of liveability on the use of public transportation.

4. Conclusions

By adopting an analytic approach to measuring liveability using a simple metric, this study was able to look at and understand the different hypotheses put forward as the factors of liveability. Furthermore, the study quantified the impact of each factor on the metric while disproving some of them as evidenced by the marginal contribution as well as the importance of the different factors towards the target metric (liveability). Based on the data collected, safety, usage of public transport, traffic wait time, access to organic farm products, dependability of the neighbors, availability of cultural/arts/sports institutions, access to the internet, and access to health care seemed to be important, whereas the other features were deemed less important/insignificant by the model. Among the important features selected by the model, only "actionable features" were considered to compute the marginal contributions. Actionable features are the ones, which can be impacted or influenced by an external agent such as a government authority or an NGO. For example, safety, access to the internet/organic farm products/healthcare/public transport, etc., were deemed "actionable", whereas features such as the dependability of neighbors and involvement in social organizations were deemed "non-actionable". The study provides some interesting insights into the marginal contribution it has on the liveability metric at different levels. For example, the marginal contribution of the access to the internet on liveability increased only if the rating was increased from 3 to 4, and not for any other increase in rating. Similarly,

safety could only improve liveability if it was rated above 3, and when it did, it improved liveability by almost 20% points.

The authors hope that the findings and insights from this study will help to inform the design of future neighborhoods and will drive policy changes in urban planning. Although the concept of liveability has been studied for many years, the dynamic nature of technology and people, calls for continuous research in this area. The concept of liveability or how suitable an area is for living, is influenced by a variety of factors. However, these factors and their importance may vary based on the cultural context of the area. This study proposes a generic framework for assessing liveability that can be adjusted based on survey responses from people living in a specific geographic and cultural context. In other words, the framework can be tailored to the specific cultural considerations of a particular area. Non-survey-based models of liveability can be constructed using features that do not require the collection of survey data, such as access to amenities and services, demographics, safety, and employment indices. These models can be useful because they can be developed without the need to gather information directly from individuals through a survey. Causal inference techniques are statistical methods that can be used to identify the causal relationships between different factors. For example, a causal inference technique might be used to determine whether access to amenities and services is a cause of liveability, or whether liveability is the result of other factors such as demographics or safety. These techniques can be useful for understanding the underlying causes of liveability in a given area and for identifying potential interventions that could improve liveability. These techniques can be further explored and researched.

References

1. Myers, D. Building Knowledge about Quality of Life for Urban Planning. *J. Am. Plan. Assoc.* **1988**, *54*, 347–358. [CrossRef]
2. Balsas, C.J. Measuring the livability of an urban centre: An exploratory study of key performance indicators. *Plan. Pr. Res.* **2004**, *19*, 101–110. [CrossRef]
3. Tyce, H.; Rebecca, L. What is livability? Research initiative 2015–2017: Framing livability. Sustainable Cities Initiative, University of Oregon. 2015. Available online: https://sci.uoregon.edu/sites/sci1.uoregon.edu/files/sub_1_-_what_is_livability_lit_review.pdf (accessed on 20 October 2018).
4. Philips Index for Health and Well-Being: A Global Perspective. Report by the Philips Center for Health and Well-Being. 2010. Available online: http://www.newscenter.philips.com/pwc_nc/main/standard/resources/corporate/press/2010/Global%20Index%20Results/20101111%20Global%20Index%20Report.pdf (accessed on 17 September 2018).
5. *Competitive Cities and Climate Change*; Organization for Economic Cooperation and Development (OECD): Milan, Italy, 2008; Available online: https://www.oecd.org/cfe/regionaldevelopment/50594939.pdf (accessed on 14 April 2019).
6. UN Habitat. City Prosperity Initiative. 2020. Available online: https://unhabitat.org/programme/city-prosperity-initiative (accessed on 17 September 2021).
7. Gandelman, N.; Piani, G.; Ferre, Z. Neighborhood Determinants of Quality of Life. *J. Happiness Stud.* **2011**, *13*, 547–563. [CrossRef]
8. PPS (Project for Public Space). *How To Turn a Place Around: A Handbook of Creating Successful Public Spaces*; PPS: New York, NY, USA, 2000.

9. Winkelmann, L.; Winkelmann, R. Why Are the Unemployed So Unhappy? Evidence from Panel Data. *Economica* **1998**, *65*, 1–15. [CrossRef]
10. Bruno, F.; Simon, L.; Alois, S. *Valuing Public Goods: The Life Satisfaction Approach*; Working Paper 1158; Center for Economic Studies and ifo Institute (CESifo): Munich, Germany, 2004.
11. Kim, B.; Yoo, M.; Park, K.C.; Lee, K.R.; Kim, J.H. A value of civic voices for smart city: A big data analysis of civic queries posed by Seoul citizens. *Cities* **2020**, *108*, 102941. [CrossRef]
12. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why Should I Trust You? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA USA, 13–17 August 2016; pp. 1135–1144. [CrossRef]
13. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 532. [CrossRef]
14. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4765–4774. [CrossRef]
15. Kaal, H. A conceptual history of livability. *City* **2011**, *15*, 532–547. [CrossRef]
16. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: New York, NY, USA, 2013; p. 426. ISBN 978-1-4614-7137-0. [CrossRef]
17. Dietterich, T.G. Ensemble Methods in Machine Learning. In Proceedings of the International Workshop on Multiple Classifier Systems, Cagliari, Italy, 9–11 June 2000; pp. 1–15. [CrossRef]
18. Lundberg, S.M.; Erion, G.G.; Lee, S.I. Consistent individualized feature attribution for tree ensembles. *arXiv* **2018**. [CrossRef]
19. Claesen, M.; Moor, B.D. Hyperparameter search in machine learning. *arXiv* **2015**. [CrossRef]
20. Fawcett, T. An Introduction to ROC analysis. *Pattern Recogn. Lett.* **2006**, *27*, 861–874. [CrossRef]
21. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2020**, *23*, 18. [CrossRef] [PubMed]
22. Excoffier, J.-B.; Salaün-Penquer, N.; Ortala, M.; Raphaël-Rousseau, M.; Chouaid, C.; Jung, C. Analysis of COVID-19 inpatients in France during first lockdown of 2020 using explainability methods. *Med. Biol. Eng. Comput.* **2022**, *60*, 1647–1658. [CrossRef] [PubMed]
23. Štrumbelj, E.; Kononenko, I. Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* **2014**, *41*, 647–665. [CrossRef]