

Determining Urban Emotion using a Supervised Learning Approach: A Case Study around Majitar, East

Supriya Choudhury
Sikkim Manipal
University
Dept. of CSE, SMIT
Majitar, Sikkim

Mohan P. Pradhan
Sikkim University
Dept. of CA
Gangtok, Sikkim

Pratikshya Sharma
Sikkim Manipal
University
Dept. of CSE, SMIT
Majitar, Sikkim

S. K. Kar
Sikkim Manipal
University
Dept. of CSE, SMIT
Majitar, Sikkim

ABSTRACT

Perception and expectation of citizens is an important factor in urban settlement, planning and management. Hence, there is a need of a participatory citizen centric planning of urban settlement based on spatial data. These perception and expectation may be represented in terms of emotions. Determining Urban Emotions is an approach which can be used to map different types of emotions associated with urbanization. In the recent years, some new methods have been presented for the area of urban and spatial planning, which resulted in a fundamental change of the issues and understanding of urban planning. Geographical information system acts as a key factor for analyzing urban emotions from various types of data. This paper presents the supervised learning approach for determining urban emotions using K-Nearest Neighbor algorithm.

Keywords

Urban Planning, Spatial planning, Smart city, Urban Emotions, K-Nearest Neighbor algorithm.

1. INTRODUCTION

An Urban Emotion is one of the emerging approaches that combine the concepts of spatial planning, geographic information systems, computer linguistics, sensor technology methods and real world data, where spatial planning considers all social and spatial structures within the city and helps in collecting various forms of data in context to the city. It involves both the spatial and temporal patterns that help in research activities in identifying processes and to characterize special social-cultural movements and developments. Geographic information system consists of two distinct disciplines geography and information system. It is an information system designed to work with data that are referenced by spatial or geographical coordinates. Computer linguistics is an interdisciplinary field which is concerned with the statistical and rule based modeling of natural language from a computational perspective. Real world data is an umbrella term used for different types of data that are collected in conventional randomized controlled trials. It can be technical sensor data, crowd sourced data, human- sensor data or social data, etc. It can be used for decision making.

The main idea behind this approach is the involvement of people of a particular location into various planning processes. Urban Emotion deals with different expectations of people regarding a particular location and what additional features can be added to the locality. It explains the potentiality of integrating objectively quantifiable emotions in context of citizen participation. Determination of Urban Emotion figures out the use of real world data.

Citizen's perception and urban space when linked together

triggers an emotional reaction and creates its own atmosphere in the observer. Urban emotion aims to understand how people's feelings get affected by features of the current environment, green spaces, air pollution, water pollution, noise pollution, affects of industrialization, land degradation, road condition, and other geographical factors.

Better urban planning approaches are needed to build a city into a smart city. Smart cities are the cities that are able to operate in a sustainable, efficient and intelligent manner and require smart infrastructure with advanced sensing capabilities that extend beyond mere technical subtitles, thereby possibly benefitting architects and citizens of the cities. It means smart citizens can make intelligent cities. It relies on the idea that only citizens can make a city really intelligent. It needs to be tackled both from technological view point and human centric view point that a city requires smart citizens to be intelligent themselves.

2. PROBLEM DEFINITION

People's feelings and emotions generally changes with the geographical location. How people gain perception with context to the city is always been an issue in urban planning and management. The problems of considering subjective measurements and views provided by the citizens by involving them into planning processes represent great challenges for efficient urban planning. The wide range of the problems of extracting human emotion in context to the city may make good understanding of different expectation of the people. Assessing human emotions with relation to various geographical data is an important issue in urban planning. Urban Emotion adds a new information layer which will help in urban planning.

3. PROPOSED APPROACH

To solve this problem, different types of data of various facilities are used which are considered to be essential for urban planning and settlements. This case study is conducted for better understanding of the developments that took place so far in context to each facility, what are the different expectation of the people regarding each facility, how much people are emotionally attached to that location, how much knowledge do they have regarding Majitar location, what additional features and developments they want in that locality.

Urban emotions can be categorized by studying the demands of the individuals availing the following six facilities:

- (i) Educational Facility
- (ii) Entertainment Facility
- (iii) Health Facility

- (iv) Industrialization Facility
- (v) Shopping Mart Facility
- (vi) Transportation Facility

Urban emotion can be determined using the following approach:

Step1. Find and determine the types of emotions to be analyzed.

Step2. Building up questionnaires for emotions.

Step3. Determination and expression of emotions by the Sample Data Set.

Step4. Determining relationship between emotions by analysis of the expressions using a computational technique.

4. METHODOLOGY USED

The k-nearest neighbor (KNN) classifier is a type of supervised machine learning approach and a non-parametric lazy learning algorithm where, non-parametric means it does not make any assumptions on the data distribution and lazy means it does not use training data objects to perform generalization. The lack of generalization means it keeps all the training data which are needed during the testing phase. Its training phase is pretty fast. It makes decisions based on the entire training data set (in the best case a subset of them). It is versatile and very simple to understand but it works incredibly well in practice.

It classifies data objects (or cases) based on their similarity to other data objects. In machine learning, it is used to recognize patterns of data without requiring an exact match to any stored patterns, or cases. Similar near each other and dissimilar objects are distant from each other. Thus, the distance between two objects (or cases) is a measure of their dissimilarity. The data objects that are near each other are said to be “neighbors”. A new data objects are termed as holdout and are used for testing purpose. Whenever a new data object (case) is presented, its distance from each of the objects (cases) in the model is computed. The classifications of the most similar cases (the nearest neighbors) are tallied and the new object case is placed into the class (or category) that contains the greatest number of nearest neighbors.

For supervised learning, the dataset of size 100 is split into two parts: 50% training and 50% testing (or holdout). There are some assumptions that are followed before using this algorithm. It is assumed that the data is in a feature space; they have a notion of distance which can be a commonly used Euclidean distance. The basic KNN computes the similarity measure based on this distance measure. This metric makes KNN very sensitive to noisy features. Each of the training data consists of a set of object (or features) and class label associated with each object. KNN can work equally well with arbitrary number of classes. It consists of a single number (k) which decides how many neighbors (that are defined based on the distance metric) influence the classification. If k=1, it is simply called as the Nearest Neighbor algorithm.

The main objective is to find the ‘k’ training samples in order to the k-nearest neighbors based on a distance measure, that is, Euclidean Distance. For that purpose, we sort the distances of all training samples to the new instance using Quick Sort algorithm and determine the kth minimum distance. We find the majority of the class (or category) of k-nearest neighbors for data classification [11]. Then, ranking and categorization of new data objects (or points) is performed.

K-nearest neighbor Algorithm:

1. Determine the ‘k’ number of neighbors.
2. Calculate the distances between the new input data and all training data using Euclidean distance [11].
3. Sort the distances and determine ‘k’ nearest neighbors based on the kth minimum distance [11].
4. Gather the categories (or classes) of those neighbors.
5. Determine the category (or class) based on majority vote.

5. CASE STUDY AREA

The methodology is implemented in the Majitar. It is a small village in East District of Sikkim, located at 27.1894°N 88.4978°E. The nearest towns to Majitar are Rangpo (4 km away) and Singtam (7 km away) as shown in figure 1 and figure 2. It is about 200 metres (660 ft) above sea level which gives it a sub-tropical climate [12].

Due to its geographic location and low population density, it has been considered to have less urban settlements and development. Majitar is largely populated by Nepalese people, Bhutia people, Marwari people, and Bengali people. Nepali language is the predominant language but some people speak Hindi too.



Fig 1: Map of Majitar (Source: Google Maps)



Fig 2: Map of Majitar (Source: Google Earth)

6. RESULTS

For the implementation, a set of questionnaires for each facility was prepared. The number of questions varies from 10-12 numbers for each facility. The responses to each questions need to be provided in the range of 1-10 for specifying the quantities. Some questions require the responses in the form of nominal values such as worst, bad, average, good and better for better understanding of present condition and scenarios of each facility. And few questions

required to be answered as yes or no. Then, the data was collected from 100 people of age group 15-60 residing in Majitar locality by conducting a survey. A dataset of range 1 to 100 was prepared. Out of 100 people, 66 responses are provided by male participants and 34 are provided by female participants. For determination and expression of emotions by the sample set we tried to analyze the responses provided by the people of that particular location. The implementation has been done with the help of IBM Statistics SPSS software. Initially, the number of neighbors was determined, that is, k=3 for each facilities respectively.

During implementation, whenever a new data object (case) is presented, Euclidean Distance is used to compute to the distance of new object (cases) from each of the objects (cases) in the model and place it into the class (or category) that contains the greatest number of nearest neighbors. It is the distance between point's p and q. It is actually the length of the line segment connecting them (\overline{pq}). If $C_1 = (p_1, p_2, p_3, \dots, p_n)$ and $C_2 = (q_1, q_2, q_3, \dots, q_n)$ are two data objects. Then, the distance between the data objects can be given as:

$$d(p, q) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \dots + (b_n - a_n)^2}$$

$$d(p, q) = \sqrt{\sum_{i=0}^n (q_i - p_i)^2}$$

The results obtained by applying k nearest neighbor algorithm on various facility data using IBM SPSS tool can be given in the form of tables and charts for each facilities:

- i. **Case Processing Summary Table:** It summarizes the number of cases (data objects) included and excluded in the analysis in total and by training and testing (holdout) samples.
- ii. **Predictor Space Chart:** It is an interactive graph of the feature space (or subspace, if there are more than 3 features). Each axis represents a feature in the model and location of points in the chart shows the values of these features for cases in the training and testing (or holdout) partitions. It classifies the data object based on their nearest neighbours in accordance with three predictor variables. It is a 3D interactive chart that identifies the three nearest neighbour to the data object. It is also termed as Feature Space Chart, where features are attributes of a data object that can be represented in a multidimensional space.
- iii. **Predictor Importance Chart:** It is a chart that helps us to focus our modelling efforts on the variables (or featured attributes) that matter most and consider dropping or ignoring those that matter least by indicating the relative importance of each variable (or featured attribute) in estimating the classifier model. Since, the values are relative, the sum of values of all variables (or featured attributes) on the display is 1.0. It does not relate to accuracy of the classifier model. It is also termed as Feature Importance (for a predictor). It is calculated by the ratio of the error rate or sum of squares of the classifier model with the predictor removed from the model to the error rate or sum of squares error for the full model.

- iv. **Classification Table:** It displays the cross classification of observed versus predicted values of the categorical target attribute by partition.
- v. **Error Summary Table:** It displays the error associated with the model; sum of squares for a continuous target and the error rate for a categorical target.

In this paper, the result obtained for three facilities are shown as below:

6.1 Educational Facility Data

It consists of 11 numbers of questions. Q1 to Q4 are about how many number of primary schools, secondary schools, higher-secondary schools, engineering and medical institutes, respectively are available at Majitar. Q5 to Q8 are about how many more of them should be there at Majitar. Q9 is about availability of government funding regarding this facility. Q10 is about whether there is any development so far in this facility. Q11 is about present condition of this facility.

The number of cases overall classified in training data sample is 54 and in testing is 40. The number of valid cases is 94 and invalid (or excluded) cases are 6 as shown in table 1.

Table 1. Case Processing Summary for Educational Facility Data

Case Processing Summary			
		N	Percent
Sample	Training	54	57.4%
	Holdout	40	42.6%
Valid		94	100.0%
Excluded		6	
Total		100	

In training sample data set, the percentage of correctly classified data objects is 44.4% and incorrectly classified data objects are 55.6%. In testing sample data set, the percentage of correctly classified data objects is 45.0% and incorrectly classified data objects are 55.0%. The accuracy (or percentage of correctness) of training data sample is 44.4% and of testing data sample is 45.0%. A class having highest measure of precision (or producer accuracy) is 'Good' in case of both training and testing samples. A class having highest measure of recall (or user accuracy) is 'Average' and 'Good' in case of both training and testing samples as shown in table 2 and table 3.

Table 2. Classification Table for Educational Facility Data

Classification Table							
Partition	Observed	Predicted					Percent Correct
		Average	Bad	Better	Good	Worst	
Training	Average	8	2	4	3	0	47.1%

	Bad	7	2	0	0	0	22.2%
	Better	5	0	0	3	0	0.0%
	Good	2	0	1	14	0	82.4%
	Worst	1	1	1	0	0	0.0%
	Overall Percent	42.6%	9.3%	11.1%	37.0%	0.0%	44.4%
Testing (Holdout)	Average	7	2	1	1	0	63.6%
	Bad	5	0	2	0	0	0.0%
	Better	0	0	1	1	0	50.0%
	Good	3	0	1	10	0	71.4%
	Worst	3	1	0	2	0	0.0%
	Missing	0	0	0	0	0	-
	Overall Percent	45.0%	7.5%	12.5%	35.0%	0.0%	45.0%

Table 3. Error Summary for Educational Facility Data

Error Summary	
Partition	Percent of Records Incorrectly Classified
Training	55.6%
Holdout	55.0%

The predictor space chart of educational facility data is a lower-dimensional projection of the predictor space, which contains a total of 10 predictors as shown in figure 3. Since, the number of neighbors, $k=3$. Therefore, the three features attributes (or predictors) that are used to build the k-nearest neighbor classifier model (in lower dimension) are Q3 (No. of higher secondary schools available, Q5 (No. of primary schools more should be there) and Q9 (Availability of government funding).

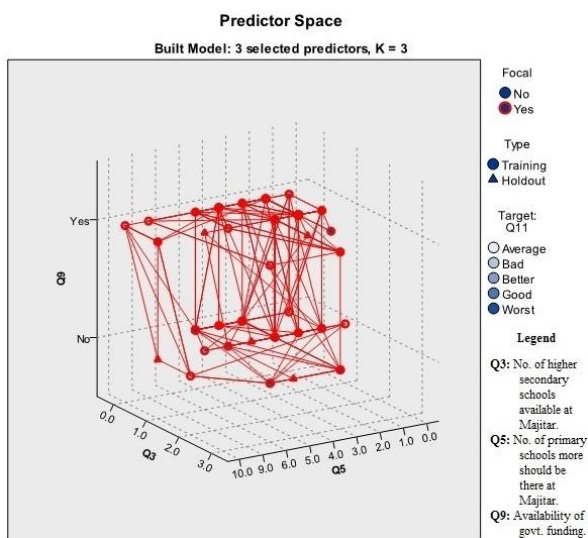


Fig 3: Predictor Space Chart for Educational Facility Data

The 10 important feature attributes which are considered by the classifier to have importance in urban development of this facility are: Q9 (Availability of government funding), Q3 (No.

of higher secondary schools available at Majitar), Q8 (No. of engineering and medical institute more should be there at Majitar), Q7 (No. of higher secondary schools more should be there at Majitar), Q5 (No. of primary schools more should be there at Majitar), Q2 (No. of secondary schools available at Majitar), Q6 (No. of secondary school more should be there at Majitar), Q4 (No. of engineering and medical institutes available at Majitar) and Q10 (Development so far). The most important amongst them is Q9 (Availability of government funding) and least important amongst them is Q1 (No. of primary schools available) as shown in figure 6.

Urban Emotions determined for this facility with respect to the expressions (or responses) provided by the people of Majitar are: More numbers of engineering and medical institutes, higher secondary, primary, and secondary schools should be add to the locality of Majitar respectively and there should be more government funding provided for that purpose.

6.2 Entertainment Facility Data

It consists of 11 numbers of questions. Q1 is about availability of amusement park at Majitar, Q2 gives the number of amusement park available, and Q3 queries whether more amusement park are necessary at Majitar. Q4 is about the availability of Children Park at Majitar, Q5 gives the number of Children Park available, and Q6 queries whether more children park are necessary at Majitar. Q7 is about the availability of Multiplex at Majitar, Q8 gives the number of Multiplexes available, and Q9 queries whether more multiplexes are necessary at Majitar. Q10 is about availability of sufficient area for additional features and Q11 is about present condition of this facility.

The number of cases overall classified in training data sample is 54 and in testing is 44. The number of valid cases is 98 and invalid (or excluded) cases are 2 as shown in table 4.

Table 4. Case Processing Summary for Entertainment Facility Data

Case Processing Summary

		N	Percent
Sample	Training	54	55.1%
	Holdout	44	44.9%
Valid		98	100.0%
Excluded		2	
Total		100	

In training sample data set, the percentage of correctly classified data objects is 48.1% and incorrectly classified data objects are 51.9%. In testing sample data set, the percentage of correctly classified data objects is 38.6% and incorrectly classified data objects are 61.4%. The accuracy (or percentage of correctness) of training data sample is 48.1% and of testing data sample is 38.6%. A class having highest measure of precision (or producer accuracy) is 'Bad' in cases of both training and testing samples. A class having highest measure of recall (or user accuracy) is 'Bad' and 'Worst' in both cases of training and testing samples as shown in table 5 and table 6.

Table 5. Error Summary for Entertainment Facility Data

Error Summary	
Partition	Percent of Records Incorrectly Classified
Training	51.9%
Holdout	61.4%

Table 6. Classification Table for Entertainment Facility Data

Classification Table							
Partition	Observed	Predicted					Percent Correct
		Average	Bad	Better	Good	Worst	
Training	Average	0	8	0	0	0	0.0%
	Bad	0	25	0	0	1	96.2%
	Better	0	3	0	0	0	0.0%
	Good	0	1	0	0	0	0.0%
	Worst	2	13	0	0	1	6.2%
	Overall Percent	3.7%	92.6%	0.0%	0.0%	3.7%	48.1%
Testing (Holdout)	Average	0	7	0	0	0	0.0%
	Bad	0	15	0	0	1	93.8%
	Better	0	0	0	0	0	0.0%
	Good	0	5	0	0	2	0.0%
	Worst	0	12	0	0	2	14.3%
	Missing	0	0	0	0	0	-
Overall Percent	0.0%	88.6%	0.0%	0.0%	11.4%	38.6%	

The predictor space chart of entertainment facility data is a lower-dimensional projection of the predictor space, which contains a total of 10 predictors as shown in figure 4. Since, the number of neighbors, $k=3$. Therefore, the three features attributes (or predictors) that are used to build the k-nearest neighbor classifier model (in lower dimension) are Q2 (No. of amusement parks available), Q3 (Amusement park is necessary) and Q6 (Children park is necessary).

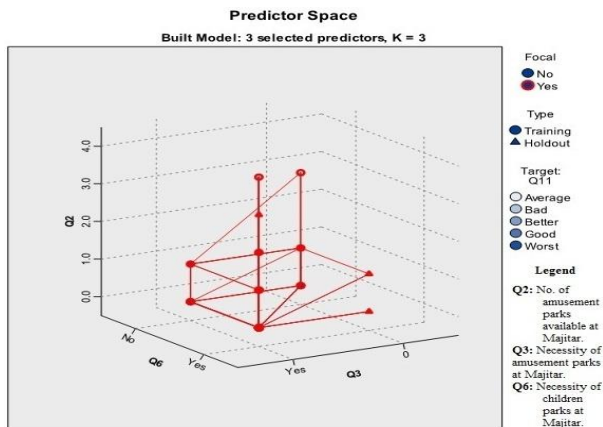


Fig 4: Predictor Space Chart for Entertainment Facility Data

The 10 important feature attributes which are considered by the classifier to have importance in urban development of this facility are: Q6 (Necessity of children parks at Majitar), Q9 (Necessity of multiplexes at Majitar), Q8 (No. of multiplexes available at Majitar), Q7 (Availability of multiplexes at Majitar), Q5 (No. of children parks available at Majitar), Q3 (Necessity of amusement parks at Majitar), Q2 (No. of amusement parks available at Majitar), Q10 (Availability of sufficient area), Q1 (Availability of amusement parks at Majitar) and Q4 (Availability of children parks at Majitar). The most important amongst them is Q6 (Children park is necessary) and least important amongst them is Q4 (Availability of children park) as shown in figure 7.

Urban Emotions determined for this facility with respect to the expressions (or responses) provided by the people of Majitar are: More children parks, multiplexes and amusement parks should be add to the locality of Majitar respectively and sufficient area should be provided for the same to the locality of Majitar.

6.3 Health Facility Data

It consists of 12 numbers of questions. Q1 queries to know whether Majitar is the birthplace of the participants in order to have a rough idea about how much experience do the participant holds about this place. Q2 and Q3 are about whether there is any hospital and health centers available there and how of them are in Majitar. Q4 and Q5 are about availability and number of pharmacies at Majitar. Q6 is about the availability of 24-hrs ambulance service at Majitar. Q7 queries how many more hospitals or health centers should be there at Majitar. Q8 is about the availability of I.C.U and advanced surgical unit at nearest hospital or health centers. Q9 is about more facilities which are to be added to health facility at Majitar. Q9 has four parts: Q91 for the need of more qualified doctors, Q92 for the need of advanced laboratories for testing and diagnosis, Q93 for the need of more advanced equipments to perform sophisticated surgeries and Q94 for the need of air ambulance for serious and critical patients. Q10 is about the rating of present condition of this facility (that includes cleanliness and surrounding of hospitals or health centers). Q11 is about availability of government funding regarding this facility. Q12 is about whether there is any development so far in this facility.

The number of cases overall classified in training data sample is 54 and in testing is 45. The number of valid cases is 99 and an invalid (or excluded) case is 1 as shown in table 7.

Table 7. Case Processing Summary for Health Facility Data

Case Processing Summary			
		N	Percent
Sample	Training	54	57.5%
	Holdout	45	45.5%
Valid		99	100.0%
Excluded		1	
Total		100	

In training sample data set, the percentage of correctly classified data objects is 27.8% and incorrectly classified data objects are 72.2%. In testing sample data set, the percentage of correctly classified data objects is 26.7% and incorrectly classified data objects are 73.3%. The accuracy (or percentage of correctness) of training data sample is 27.8% and of testing data sample is 26.7%. A class having highest measure of

precision (or producer accuracy) is ‘Average’ and ‘Better’ in case of training sample and is ‘Better’ in case of testing sample. A class having highest measure of recall (or user accuracy) is ‘Average’ and ‘Better’ in case of training sample and is ‘Better’ in case of testing sample as shown in table 8 and table 9.

Table 8. Classification Table for Health Facility Data

Classification Table							
Partition	Observed	Predicted					Percent Correct
		Average	Bad	Better	Good	Worst	
Training	Average	5	1	8	0	0	35.7%
	Bad	3	1	2	0	0	16.7%
	Better	6	1	9	1	1	50.0%
	Good	0	2	6	0	1	0.0%
	Worst	2	0	5	0	0	0.0%
	Overall Percent		29.6%	9.3%	55.6%	1.9%	3.7%
Testing (Holdout)	Average	2	0	6	2	0	20.0%
	Bad	0	0	4	0	0	0.0%
	Better	5	0	10	0	0	66.7%
	Good	2	1	6	0	0	0.0%
	Worst	2	1	4	0	0	0.0%
	Missing	0	0	0	0	0	-
	Overall Percent		24.4%	4.4%	66.7%	4.4%	0.0%

Table 9. Error Summary for Health Facility Data

Error Summary	
Partition	Percent of Records Incorrectly Classified
Training	72.2%
Holdout	73.3%

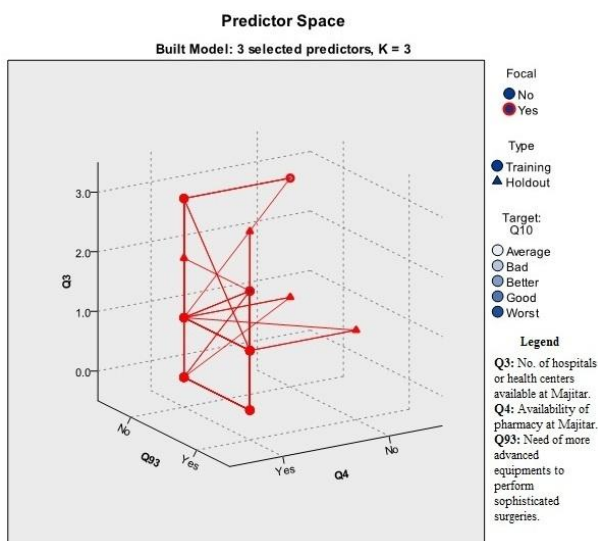


Fig 5: Predictor Space Chart for Health Facility Data

The predictor space chart of transportation facility data is a lower-dimensional projection of the predictor space, which contains a total of 14 predictors as shown in figure 5. Since, the number of neighbors, $k=3$. Therefore, the three features attributes (or predictors) that are used to build the k -nearest neighbor classifier model (in lower dimension) are Q3 (No. of hospitals or health centers available), Q4 (Availability of pharmacy) and Q93 (Need of more advanced equipments to perform sophisticated surgeries).

The 10 important feature attributes which are considered by the classifier to have importance in urban development of this facility are: Q93 (Need of more advanced equipments to perform sophisticated surgeries), Q11 (Availability of government funding), Q4 (Availability of pharmacy at Majitar), Q3 (No. of hospitals or health centers available at Majitar), Q94 (Need of air ambulance for serious and critical patients), Q8 (Availability of I.C.U and advanced surgical unit at Majitar), Q12 (Development so far), Q92 (Need of advanced laboratories for testing and diagnosis), Q5 (No. of pharmacy available at Majitar) and Q7 (No. of hospitals or health centers more should be there at Majitar). The most important amongst them as considered by the classifier is Q93 (Need of more advanced equipments to perform sophisticated surgeries) and least important amongst them as considered by the classifier is Q7 (No. of hospitals or health center more should be there) as shown in figure 8.

Urban Emotions determined for this facility with respect to the expressions (or responses) provided by the people of Majitar are: More advanced equipments to perform sophisticated surgeries, air ambulance for serious (and critical patients), I.C.U and advanced surgical unit should be add to the hospital or health centers that are available at Majitar. From the development so far, we find that there are enough pharmacies but less number of hospital or health centers available at Majitar, so, more number of hospitals or health centers should be add to the locality of Majitar.

7. APPLICATIONS

There are various application areas of Urban Emotions. It helps in urban safety, settlements, planning, settlements and development. It is useful for traffic planning and people centric tourism. It may be used for assessing previous planning measures. It can be used to improve the quality of living of citizens.

8. CONCLUSION

Urban emotion is becoming one of the major areas of research that aims at quality planning and urban settlement prior to its implementation for the betterment of citizens and humanity at large. From the results of a supervised learning approach K-Nearest Neighbor Classifier algorithm, it was observed that the classifier finds out 10 important featured attribute in each facility which will effect in urban planning, development and urbanization at Majitar. More focus should be provided to featured attribute which is considered as most important by the classifier when the number of neighbors is $k=3$. By analyzing the results, it is found that featured attributes which has more importance will effect in urban development and those featured attributes are said to have the best response than the other attributes of each facilities. Development should be made to those featured attributes in basis of their importance in urban planning and development. From this study, it can be concluded that people of Majitar have demanded improvement as well as development in each facility. They have suggested many features which are need to be included at various facilities available at this locality. Thus,

Urban Emotions can act as new information layers within planning processes and determining it using a supervised learning approach will help in understanding urbanization

and in urban planning. In future, it will help in implementing the concept of 'Smart City' at Majitar, which will lead to better sustainability.

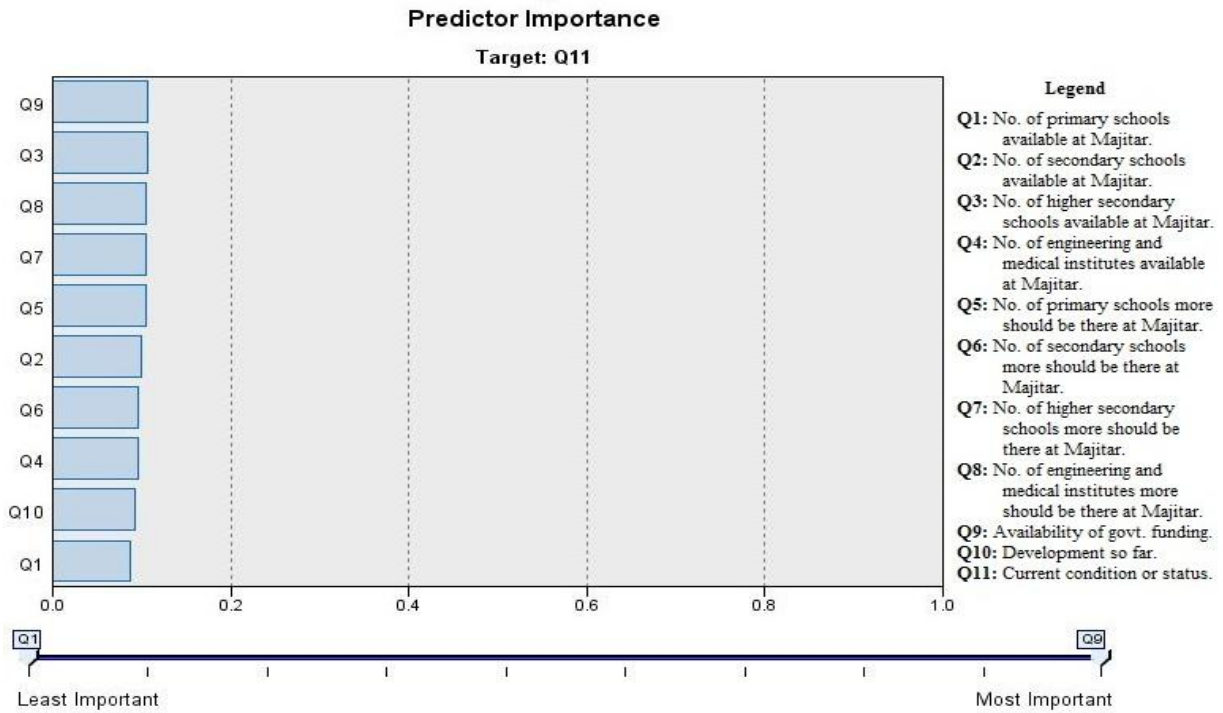


Fig 6: Predictor Importance for Educational Facility Data

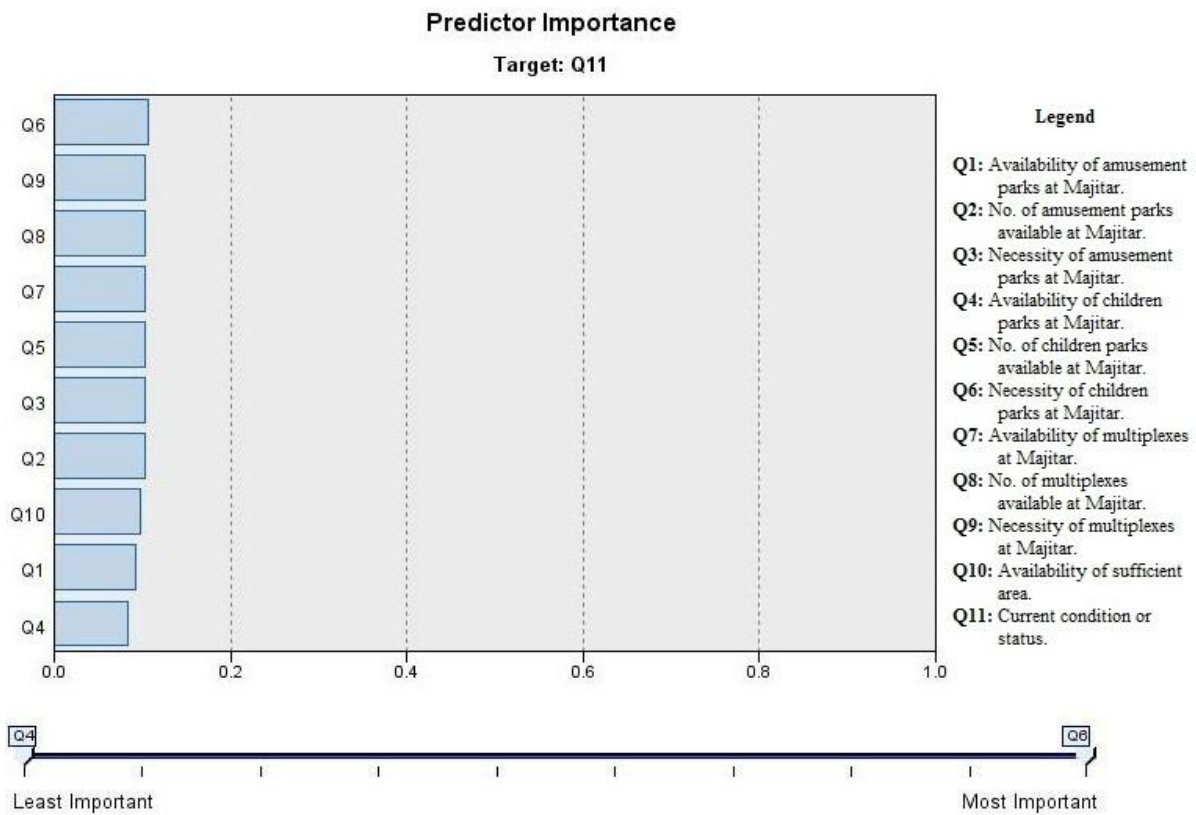


Fig 7: Predictor Importance for Entertainment Facility Data

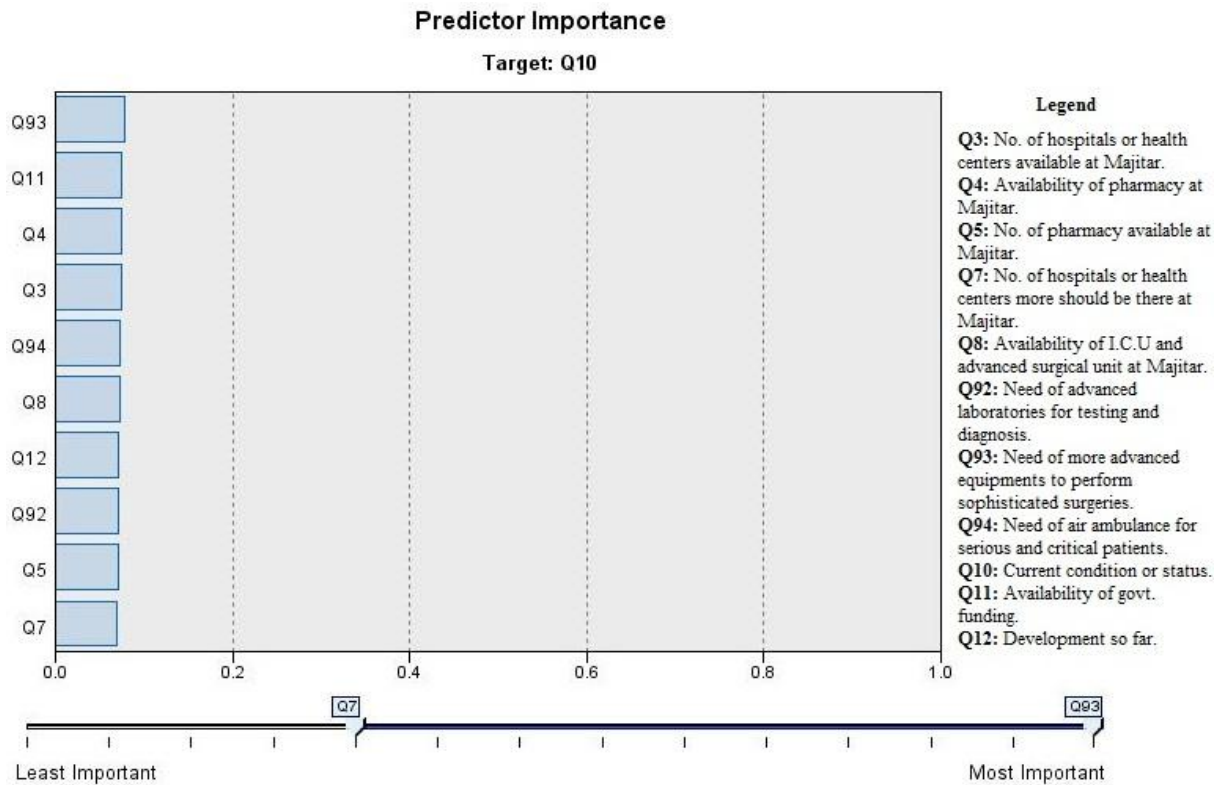


Fig 8: Predictor Importance for Health Facility Data

9. REFERENCES

- [1] Supriya Choudhury, Mohan P. Pradhan, S. K. Kar, "A Survey on Determining Urban Emotions using Geo-Data Classification: A Case Study around Majitar, East District, Sikkim", *International Journal of Computer Applications*, (0975 – 8887), Volume 135 – No.2, February, 2016, ISBN : 973-93-80891-05-5.
- [2] Peter Zeile, Bernd Resch, Linda Dorrzapf, Jan-Philipp Exner, Gunter Sagl, Anja Summa, Martin Sudmanns, *Urban Emotions–Tools of Integrating People’s Perception into Urban Planning*, Conference Proceedings REAL CORP 2015 Tagungsband, 5-7 May 2015, Ghent, Belgium. ISBN: 978-3-9503110-8-2 (CD-ROM); ISBN: 978-3-9503110-9-9 (Print).
- [3] Peter Zeile, Bernd Resch, Jan-Philipp Exner and Gunther Sagl, *Urban Emotions Benefits and Risks in Using Human Sensory Assessment for the Extraction of Contextual Emotion Information in Urban Planning*, Springer International Publishing, 2015.
- [4] Bernd Resch, Martin Sudmanns, Gunther Sagl, Anja Summa, Peter Zeile, and Jan-Philipp Exner, *Crowd-sourcing Physiological Conditions and Subjective Emotions by Coupling Technical and Human Mobile Sensors*, *GI Forum – Journal for Geographic Information Science*, 1-2015, Berlin, ISBN 978-3-87907-558-4, ISSN 2308-1708, doi:10.1553/giscience2015s514.
- [5] Gunther Sagl, Bernd Resch, and Thomas Blaschke, *Contextual Sensing: Integrating Contextual Information with Human and Technical Geo-Sensor Information for Smart Cities*, *Open Access Sensors*, 2015, 15, 17013-17035; doi: 10.3390/s150717013, ISSN 1424-8220.
- [6] Bernd Resch, Anja Summa, Gunther Sagl, Peter Zeile, Jan-Philipp Exner, *Urban Emotions–Geo-Semantic Emotion Extraction from Technical Sensors*, Human Sensors, Springer International Publishing, 2014.
- [7] Chrysaida-Aliki Papadopoulou and Maria Giaoutzi, *Crowd-sourcing as a Tool for Knowledge Acquisition in Spatial Planning*", *Future Internet* 2014, 6, 109-125; ISSN 1999-5903, doi:10.3390/fi6010109.
- [8] Benjamin S. Bergner, Jan-Philipp Exner, Martin Memmel, Rania Raslan, Dina Taha, Manar Talal, Peter Zeile, "Human Sensory Assessment Methods in Urban Planning – a Case Study in Alexandria", *Conference Proceedings REAL CORP 2013*, Tagungsband, 20-23 May 2013, Rome, Italy, ISBN: 978-3-9503110-4-4 (CD-ROM); ISBN: 978-3-9503110-5-1 (Print).
- [9] Bernd Resch, "People as Sensors and Collective Sensing-Contextual Observations Complementing Geo-Sensor Network Measurements", Springer International Publishing, 2013.
- [10] Peter Zeile, Martin Memmel, Jan-Philipp Exner, "A New Urban Sensing and Monitoring Approach: Tagging the City with the RADAR SENSING App", *Reviewed Paper of Conference Proceedings REAL CORP 2012*, Tagungsband, 14-16 May 2012, Schwechat, ISBN: 978-3-9503110-2-0 (CD-ROM); ISBN: 978-3-9503110-3-7 (Print).
- [11] About K-nearest neighbors algorithm, https://en.wikipedia.org/wiki/Knearest_neighbors_algorithm.
- [12] About Majitar, <https://en.wikipedia.org/wiki/Majitar>