A LARGE SCALE CLASSIFICATION OF PUBLIC SPACES USING A STUDY OF PERCEPTION AND PUBLIC DATA FROM INEGI OF THE CITY OF PUEBLA, SAN PEDRO CHOLULA AND SAN ANDRES CHOLULA

A. Oliart Ros[1], T. González Cacho[1], D. Sol Martínez[1], D. Clavijo Plourde[1,]

[1] ITESM, 72453 Puebla, México - (aoliart, tgonzalezc, dsol, dclavijo)@tec.mx

KEY WORDS: Perception, Machine Learning, Classification, City Planning, Public Places, Multi Layer Perceptron, Crowdsourcing

ABSTRACT:

This work aims to create a methodology to automatize the classification of public spaces using a perception test and data obtained from city census information, in our case, from the Mexican National Institute of Statistics and Geography (INEGI). Nowadays there is no well defined process in decision making when planning the creation or development of public spaces. For this reason, a study to measure the human perception was made in order to gather data about what people perceived about five variables: architectural beauty, pollution, fun, wealth and safety. The information obtained was used to create a Machine learning model that could find a relation between the perception obtained and the census dataset. This first attempt aims to find key insights needed to develop a more complex methodology to classify, at a greater scale, public places in terms of their safety or architectural value and which socio-demographic data defines this perception.

1. INTRODUCTION

The design of public spaces to improve the quality of life has become one of the main research topics for architects, urbanists and city planners. On the last decade, urban studies, linked to smart cities and information technologies, have introduced new ways to understand the public space to respond to user requirements. Human perception of the space is becoming a new variable for urban design. This variable allows government and society to work together in defining the best plan for every space according to the users.

Society has an important role in the improvement of public spaces. Nevertheless, if the efforts are not part of a methodology, it becomes hard to give continuity and replicate good practices. In order to aid the community, it is very useful for the government to have a structured manner to ask people for feedback regarding public space design. Different people can have different perceptions of the same place, moreover, one person can have different perceptions of the same place depending on different variables, like time, the activities on it or other variables, like safety or beauty. Our approach includes perception as an important variable to describe public spaces and perception will be described in our work as urban qualitative data.

On the other hand, the methodology includes quantitative data which can describe facts that do not depend on perception. Big amounts of data are collected by governments related to the number of inhabitants, number of shops and businesses, the volume of traffic and other variables. These variables describe the urban quantitative data. Our approach looks for the relationships between the qualitative and quantitative data. In other words, we want to find the relationship between what has been measured and what has been perceived. The relationship between both kinds of data will allow us to describe what happens in a urban area. This is the first step to understand a public space. The work aims to improve the understanding of public space perception and the variables that define it. The goal is to determine the impact level and relationship between the

variables that affect the urban space perception. Other levels of understanding will help us make predictions and develop a sensitivity analysis. The last one will help us anticipate the behavior of a community in relation to a public space, and improve its urban design.

Our approach allows the generation of qualitative and quantitative data by using a tested methodology based on crowdsourcing ((Naik et al., 2014), (Barwell et al., 2012)). The strategy is to gather information about what people think of their surroundings in five specific variables: safety, fun, pollution, wealth and architectural beauty; secondly, train a multilayer perceptron network using the result of the perception test and urban descriptors; and finally, measure the performance of the models obtained. This implementation could lead to create a large scale classification of public spaces which certainly would impact the process of city planning.

2. STATE OF THE ART

The sensorial user experience of the public space research is relatively new for architectural and engineering fields, being more common in other disciplines, such as industrial design, psychology and sociology, among others. The human perception studies of the city begin more seriously during the 70s. Almost 20 years ago, the expectations were that the continuous improvements in technology would lay the foundations to create, at this time, convincing virtual environments. However, scientists were just becoming aware that this advances were an almost optimal configuration for the human perception, cognition and behavioral studies (van Veen et al., 1998).

The first mentions on the topic of city perception date from the 1940s. At this period the perception was mainly focused on visual criteria. "The process of visual perception has been the focus of attention and research for the experimental psychologist", but there has been little attention received by other scholars of human behavior (Schachtel, 1941). A few years later,

Guy Debord, a socialist theorist and member of the situationist group, proposed a novel way of touring the city and sharing human spaces. Debord seeks and shapes conventional activity patterns, through the technique of "derives", which are small walks through different urban environments based on psychological reviews. He calls this way of navigating through the psychic state of the behavioral psychogeography environments, which is what differentiates them from a conventional walk (Debord, 2014).

During the 60s Kevin Lynch presented his book, The image of the city, which sets a precedent to link architecture with human perception. Lynch believes that the environment is not foreign to us, but complementary. "We are not mere observers of this show, we are part of it, on stage with the other participants. [...] almost any sense is in operation and the image is the combination of all of them" (Lynch, 2014).

Lynchs perspective, that every image is composed of three elements: identity, structure and meaning, is reinforced by the ecological perspective of Gibson (1979) according to which all the information that a person needs to perceive of the environment is already contained in the impact produced by an environmental optical pattern (Gibson, 2014). However, Gibson, returns to the visual, leaving aside the multisensory influence.

Certainly, most of the perception studies before 1970 were primarily focused on the visual criteria, however Cowan and Stewart, in their book The city and the senses: urban culture since 1500, make a historical review based on the sensory dimension of life in the city according to the five senses. Since then they recognize that although psychologists mentioned that the bulk of perception is conceived by sight, the other four senses (touch, taste, sound and smell) contribute to the particular construction of the sensory (Cooper, n.d.).

Accepting the premise that perception is multisensory, perception is discerned as conscious and not conscious. "It is based on the rejection of the hypothesis that the phenomenal experience is identical or is a direct reflection of the representations produced by perceptual processes" (Marcel, 1983). Based on these premises, studies of the first urban models that consider conscious and non-conscious perceptual processes are initiated.

Added to the perceptive processes is the search for the representation of them. Since 1960, Lynch proposed the creation of cognitive maps of the city. These maps focused mainly on the conscious perception of each person. The main components and the cartography chosen to represent these findings were mainly mental maps or cognitive maps. These maps reflect the subjective perception of people in small urban segments (Downs, Stea, 1975). The environmental perception includes the set of attitudes, motivations and values that influence the different social groups; when defining the perceived environment. Which affects not only the knowledge of the environment; but also its behavior within it (Rapoport, 1978). In 2010, the artist and designer Kate Mclean, creator of smell maps of cities around the world, began to "focus on human experience of urban smellscapes rendering it as "eye-visible" through mapping" as she explains in its webpage. Mclean develops a methodology to represent the smell of the city taking as background the Theory of the derive of Deborb. McLeans maps settled a pathway to communicate graphically the results of the perception research.

In 1988 Jack L. Nasar started to conduct an investigation to determine how citizens evaluate the cities where they live. "To improve community appearance, planners need to know how the public evaluates the cityscape: their evaluative image of the city." (Nasar, 1990). The evaluation of the perceived image of the city, denotes the importance to study and understand the sensory response to improve the quality of life in a city. The research problem lies in finding the factors that define what people like and dislike. In addition to improve the quality of life of the citizens, understanding the sensorial response allows the city planners to take decisions that enhance tourism and diversify or reinforce the local economy.

According to Nasar, there are two essential aspects for the planning of a city: how one imagines the city and how it is evaluated. The study was conducted in two cities in the United States of North America: Knoxville and Chattanooga of the state of Tennessee, where 220 residents and 180 visitors were interviewed and 160 evaluation maps were generated. Some of the results were expected to support the original hypothesis: people like vegetation and do not like advertisements. Nevertheless, just by plotting these results graphically and geographically it aims to pass from the common believe to make it a reality in the city.

The evolution of these studies improved the development of more structured processes to infer the sensorial response of public space. New proposals, such as the Sensorial Urban Lab (LUS), provide a methodology for obtaining reliable data to understand and interpret the sensorial response of a public space and patterns to design it according to the results to improve life quality. A study carried out by the town planner Talía González Cacho, with the methodology of the Sensorial Urban Lab (González Cacho et al., 2017), shows the results of the users of the linear park and bikeway in Puebla. Users were interviewed using a set of pre-selected images, to inquire about their likes and dislikes. Once this test is done, the interview continues using the model of creating mental maps of the space in question, followed by direct questions about how the interviewee is affected with what surrounds him. After the interview, the researchers process the information gathered in sets called cross observations so that finally, an expert can read them and denote insights or descriptors of the place. Although this methodology concluded relevant information, one of the disadvantages of the model is the large amount of time that must be spent for each interview, so, carrying this methodology on a large scale makes it very complicated.

Recent studies, like The Collaborative Image of The City: Mapping the Inequality of Urban Perception (Nasar, 1990) and UrbanGems (Quercia, 2012) use crowdsourcing in order to obtain what does the citizen feels about their surroundings. Moreover, César A. Hidalgo, the author of a second study called Deep learning the city: Quantifying Urban Perception at a Global Scale (Dubey et al., 2016) merged a perception test with a Siamese-Like convolutional neural architecture (also known as ConvNet) aiming to relate what people perceived from a collection of images, with the visual information contained in those images, like geometric forms, in order to automatize, at a great scale, the classification of public places trying to recreate people's perception about safety and other variables.

3. METHODOLOGY

The study has two main components: the perception test and the information obtained from the mexican National Institute of Statistics and Geography, INEGI by its name in Spanish. The research guidelines for making this investigation were:

- Selection of city hotspots.

- Perception test based on 5 variables: fun, health, pollution, architectural beauty and safety.

- Living place commodities.

- Socio-demographic descriptors.

- Machine learning model implementation.

3.1 Selection of City Hotspots

This approach contains 124 semi-randomly selected points of study or urban hotspots (figure 1). The selection conditions where:

- Each hotspot had to be within the boundaries of the city of Puebla, San Andrés Cholula or San Pedro Cholula.

- Be a crossroad.

- Be at least 2 square blocks away from other hotspots.

Figure 1. Hostpots location.

3.2 Perception Test Implementation

The perception test consisted on a website (figure 2) where a pair of images from different hotspots were shown and one of five different questions related to the five evaluated variables was asked (eg. Which image looks wealthier?). Once the user clicked on the image which best answered the question, a new pair of images was shown and the question changed. Also, the site stored the winner image, the loser image and the variable in question. One of the test goals was to be as simple as possible in order to obtain the greater number of answers or clicks. On the other hand, a user could answer as many questions as she or he would like to. In addition, a 2 seconds counter limited the response time in order to decrease the number of fake answers.

Each hotspot had 4 different images (in order to get a 360 kind of view), one for each street side. The images were obtained using the Street View Static API (Google Maps. Street View Static API) which allows to get a street view image of a pair of coordinates (example given in figure 3).

The website was conformed by four pages; home page, where the description of the project was written; test page, where the perception test took place; the collection of images, where the users could see all collection images; and the maps page, where hotspots where showed using a Google Maps widget.

The test was projected to have at least 10,000 clicks. Consequently, the ranking from which the best and the worst image
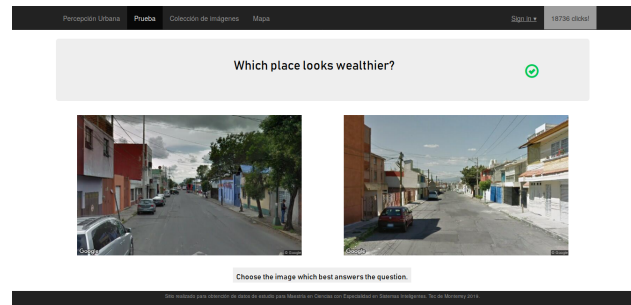

Figure 2. Perception Test Website.

Figure 3. Four different images from one hotspot. Heading and fov are two parameters used to configure the resulted image from de Street View Static API.

were label was made by using the Katz Centrality theorem (juyong Park, Newman, 2005) (equation 1).

$$C_{katz} = \sum_{k=1}^{\infty} \sum_{j=1}^{n} a^k A_{ij}^k \tag{1}$$

3.3 Living place commodities and socio-demographic descriptors

Living place commodities and socio-demographic descriptors were obtained from the public database of INEGI (https://www.inegi.org.mx/) but in two different sections; the first can be found in the National Inventory of Living (INV) and the latter in the National Directory of Statistics of Economic Units (DENUE). Both were obtained from a census done on 2010 and a study from 2015.

3.4 Machine Learning Model Implementation

Scikit Learn (Buitinck et al., 2013) is an open software repository written in Python programming language that comes with different machine learning techniques ready to use. It is commonly used for prototyping, so it comes very handy when needed to find the most suitable technique for data processing. For this research, it was used the Multi Layer Perceptron Classifier (or MLPC) due its capacity to learn non-linear models. This

model has a high quality training algorithm that uses backpropagation and can be configured with up to 22 parameters.

For this study, it was expected to create 5 different models; one for each perception variable. In order to get the most accurate model configuration, there were made several experiments changing 4 main parameters: the activation function, the solution method, maximum number of iterations and the number of hidden layers. On the other hand, 75% of the dataset will be used to train the network while the other 25% will be used to measure the performance of the created model.

The presented methodology and following results were obtained from a master thesis called: Clasificación de espacios urbanos a gran escala a partir de un estudio de percepción y datos del INEGI de la ciudad de Puebla, San Pedro Cholula y San Andrés Cholula (Clavijo Plourde, 2019).

4. RESULTS

4.1 National Inventory of living

After processing the information found in this database, we found 3 main living descriptors: household livelihood, household services and household age.

The percentage of the 3 living descriptors versus the studied hotspots resulted as follows:

- **Percentage of livelihood**:

$$h = \frac{\text{Habited households}}{\text{Total number of households}} \quad (2)$$

- **Percentage of basic services**:

$$s = \frac{\text{Sum of households with some service}}{5 * (\text{Total number of households})} \quad (3)$$

- **percentage of x age of population**:

$$\text{Population(age)} = \frac{households(age)}{\text{Total number of households}} \quad (4)$$

This information was gathered by selecting the 4 nearest Basic Units of living, AGEB by its name in Spanish, for the 124 studied hotspots. The process to obtain this information is by placing a 20 m radius circle in the INV interactive map as shown in figure 4.

Figure 4. Interactive map from INV website.

4.2 National Directory of Statistics of Economic Units

Because the DENUE database could be massively downloaded and had the latitude and longitude coordinates, it was easier to get the total number of establishments around the point of study with all their characteristics. By using a python script there were selected the establishment that were at a maximum distance of 50 m from each hotspot. The resulted information obtained the:

- Total number of establishments per hotspot (figure 5).

- Economic activity for each establishment (there were found 20 different economic activities groups in the 124 points of study).

- Establishment size divided in:

    - Micro A (between 0 and 5 employees)
    - Micro B (between 6 and 10 employees)
    - Small A (between 11 and 30 employees)
    - Small B (between 31 and 50 employees)
    - Medium A between 51 and 100 employees)
    - Medium B between 101 and 250 employees)
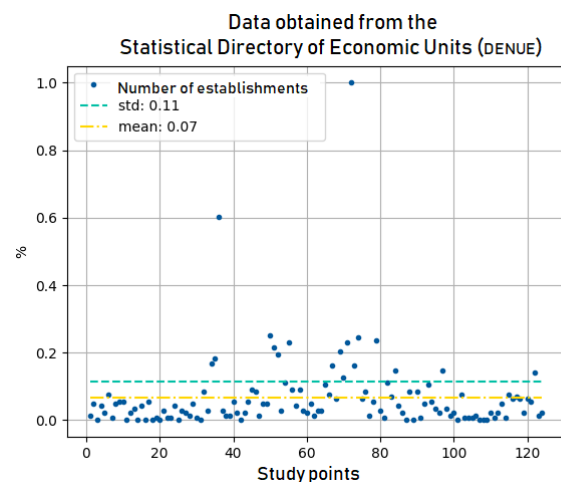   - Large (more than 251 employees)

Figure 5. Number of establishments per hotspot.

4.3 Perception test

The perception test was taken 9,758 times in a lapse of 4 weeks. This 9,758 clicks were divided as follows:

- Architectural beauty: 2,345 clicks.

- Pollution: 317 clicks (footnote: there was an error in the website that make to jump this question).

- Fun: 2,370 clicks.

- Wealth: 2,392 clicks.

- Safety: 2,334 clicks.

The results are shown in the figures: 6, for architectural beauty; 7, for fun; 8, for pollution; 10, for safety and 9 for wealth. The perception rank for each hotspot was obtained by computing the average value of its images and the then normalized on a scale from 0 to 1, where 1 represents the highest level of perception and 0 the lowest.
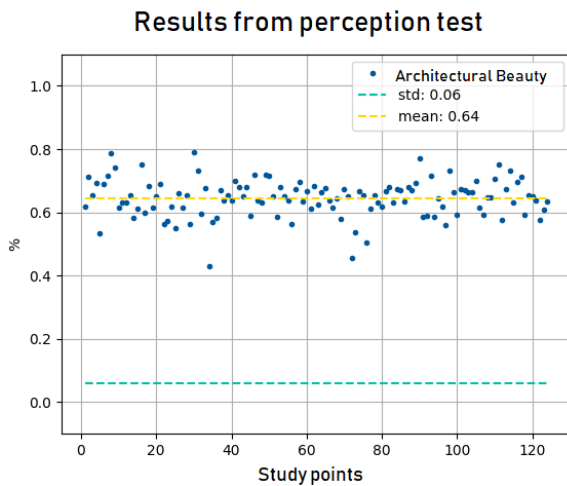
Figure 6. Architectural beauty perceived from the points of study.
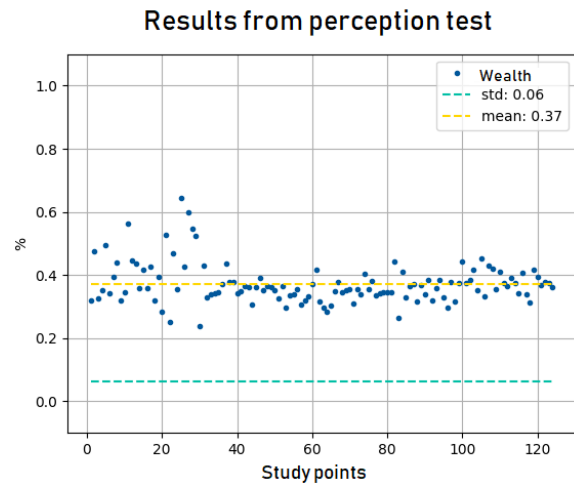

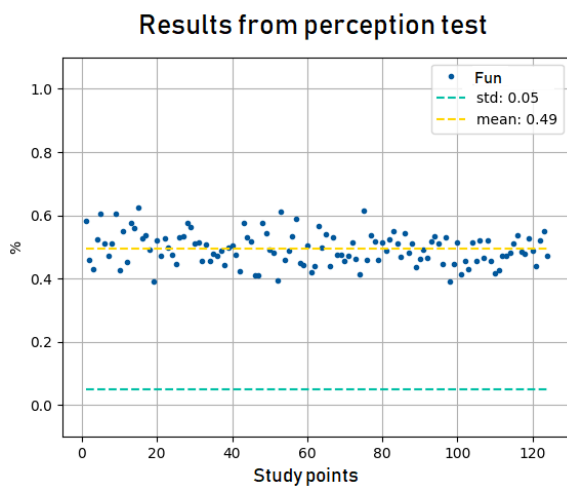Figure 9. Wealth perceived from the points of study.

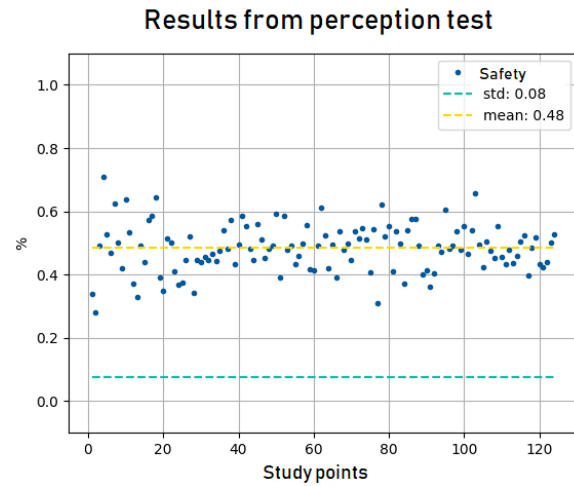Figure 7. Fun perceived from the points of study.

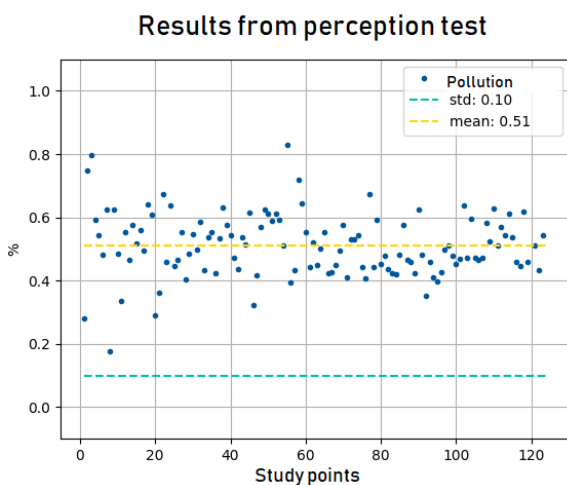Figure 10. Safety perceived from the points of study.

Figure 8. Pollution perceived from the points of study.

4.4 Matching perception test results with INV and DE-NUE values

One interesting collateral result was the 170 correlations obtained from the resulted ranks from the perception test and the values from the INV and DENUE databases. Figures 11, 12 and 12 shows 3 of this correlations.

Figure 11. Relation between Architecture Beauty VS Number of establishments nearby. It can be seen that at lower Number of establishments, the higher perception of Architecture Beauty.
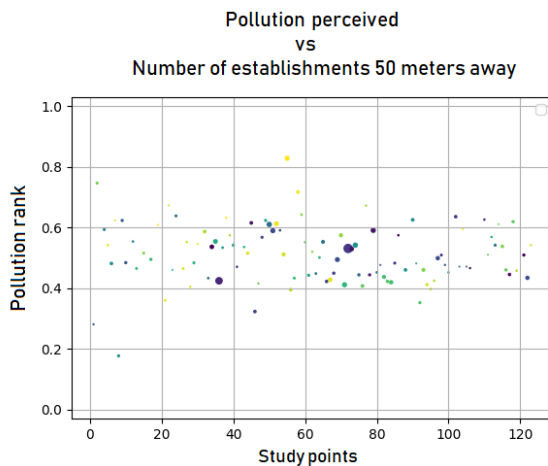
Figure 12. Relation between Pollution VS Number of establishments nearby. It can be seen that the higher perception of Pollution the greater Number of establishments
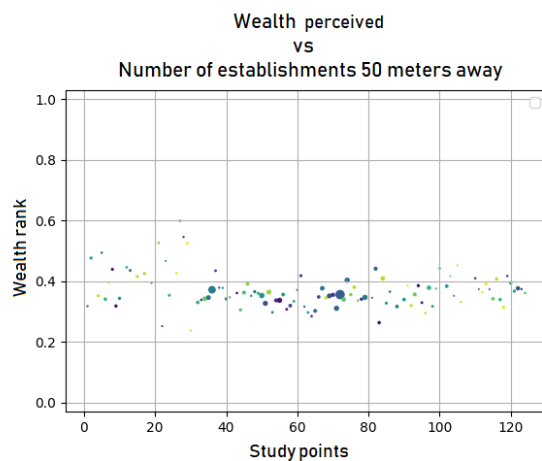

Figure 13. Relation between Wealth VS Number of establishments nearby. It can be seen that the Wealth perception has a inverse relation with the Number of establishments.

4.5 Multi Layer Perceptron Models

From the rank numerical value, each hotspot was categorically labeled from very low to top high values. This correlation is shown in table 1. This codification was used to train the classification model for each variable of study of perception. After testing different combinations of parameters, it was chosen for the activation function the sigmoid or logistic function, the gradient descent as the solution method, 1000 iterations and a single hidden layer of 34 perceptrons.

The prediction results from testing the 25% remaining data is shown in figures 14, 15, 16, 17 and 18 for each variable of perception.

5. FURTHER WORK

Every step during this study led to powerful insights that definitely could lead to new ways to improve data mining processes, optimize evaluation techniques and create other kinds of scalable urban projects. Each one of them leading to a better way

Range	Matrix	Meaning	Value
0 - .2	[1, 0, 0, 0, 0]	Very Low	0
.21 - .4	[0, 1, 0, 0, 0]	Low	1
.41 - .6	[0, 0, 1, 0, 0]	Medium	2
.61 - .8	[0, 0, 0, 1, 0]	High	3
.81 - 1	[0, 0, 0, 0, 1]	Very High	4

Table 1. Source: Own creation. Labeled ranks use to categorized the results from the perception tests.
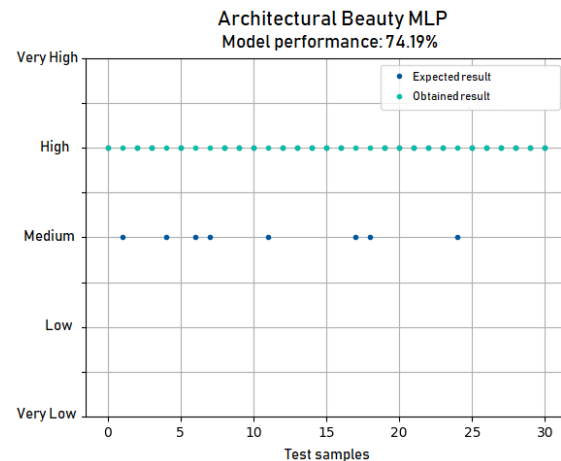

Figure 14. MLP for Architectural Beauty.
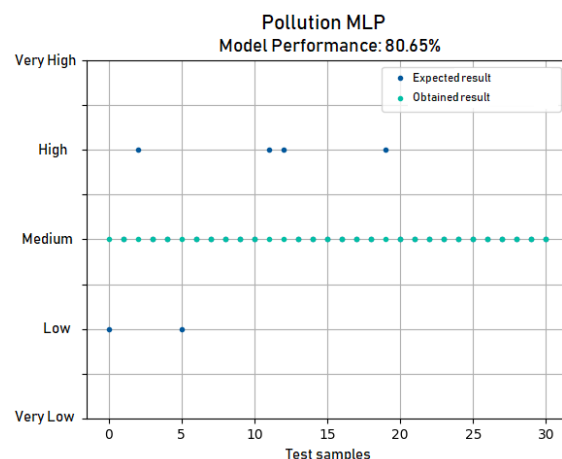Model Performance: 74.19%.

Figure 15. MLP for Pollution.
Model Performance: 80.65%.

of living, optimizing natural resources and improve citizen and community behavior.

From the previous research can be seen that urban development and city planning are very complex processes that will grow proportional to the size of its population and inversely proportional to its lack of natural resources. As a proposal for future work, it is intended to establish a methodology to obtain data, transform it to readable information and create knowledge. This implementation can be described in four different stages:

1. What happened? This information gives the actual state of the public place under study.

2. Why it happened? This question needs to correlate the gathered information, create relations between activities
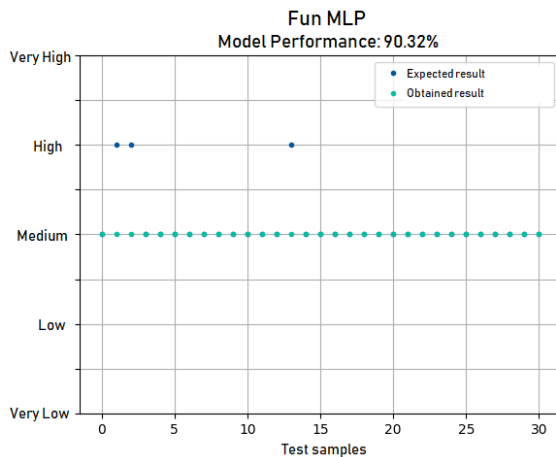
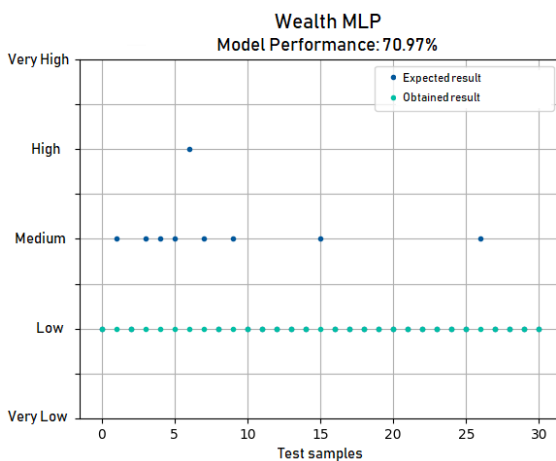Figure 16. MLP for Fun.
Model Performance: 90.32%.

Figure 17. MLP for Wealth.
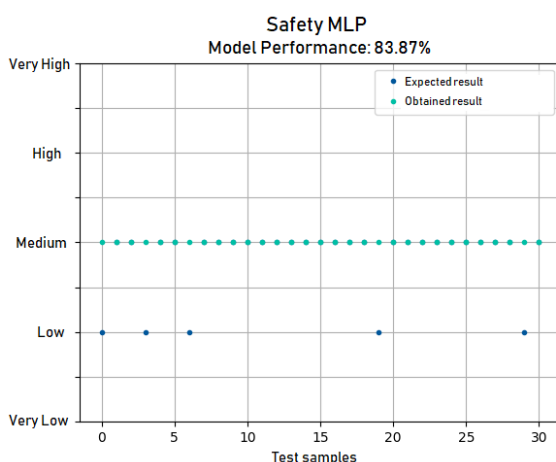Model Performance: 70.97%.

Figure 18. MLP for Safety.
Model Performance: 83.87%.

that took place in a period of time and determine which of them were the ones that lead to this present.

3. What will happen? If there is a known relation about pre-

vious activities and actual state, could be predicted the direction communities or specific places are heading?

4. How can it be changed? After making predictions, actions can be taken in order to change the future.

The present study presents a first approach to establish a process to resolve questions 1, 2 and 3; with more suitable machine learning techniques, deeper research and better raw data (more vast and recent) can manufacture a process in which cities growth and development could be planned years before problems come. One thing is certain, complex problems have to be resolved using robust processing techniques.

6. CONCLUSION

As many projects, this one had a profound problem gathering data. The first and most disreputable hassle was that it was not easy to find recent publicly available information that could describe the public space infrastructure. On the other hand, the National Institute of Statistics and Geography brings a suitable platform where a robust dataset can be found and explore. Yet, obtaining information from the INV section could only be made one by one because of the lack of coordinates in the AGEB description, new ways of gathering this data must be found in order to make easier the manipulation and preprocessing.

Another limitation in the study was the strategy of semi-random selection of hotspots. For future work there must be selected precise point of study in order to have a higher contrast in the perception test. As a result of making a better selection of hotspots, the training set will carry richer information which will lead to machine learning models with a higher performance rate. In this case, almost all the points of study were ranked in the same range so the models created converge to a single number.

On the other hand, the study of perception resulted to be a very interesting, versatile, easy to use and a suitable way of crowdsourcing. The only thing to take into consideration is that in order to obtain a more confident result, the greater number of responses needed.

In conclusion, talking about city planning, every variable has to be profoundly study before taking an idea into action, so combining different machine learning techniques with urban theory could lead to find which socio-demographic variables have a deep relation with safety perception for example, and knowing which activities can improve citizen behavior should be a direct way to construct more robust and confident communities that can thrive from problems such as poverty, pollution and lack of services, or maybe increase the level of fun and happiness the citizens live day to day.

ACKNOWLEDGEMENTS

REFERENCES

Barwell, A., Quercia, D., Crowcroft, J., 2012. Urbangems: Crowdsourcing quiet, beauty and happiness. Accessed 16/03/2019.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., Varoquaux, G., 2013. API design for machine learning software: experiences from the scikit-learn project. Accessed 02/02/2019.

Clavijo Plourde, D., 2019. Clasificación de espacios urbanos a gran escala a partir de un estudio de percepción y datos del inegi de la ciudad de puebla, san pedro cholula y san andrés cholula. Master's thesis, Instituto Tecnológico y de Estudios Superiores de Monterrey, Vía Atlixcáyotl 5718, Reserva Territorial Atlixcáyotl, 72453 Puebla, Pue., Mexico. Master in Science with Specialization in Intelligent Systems.

Cooper, L. H., n.d. The City and the Senses: Urban Culture since 1500. *SIXTEENTH CENTURY JOURNAL*, 39(4), 1111 - 1112. http://0-search.ebscohost.com.millenium.itesm.mx/login.aspx?direct=true&db=edswah&AN=000263252200009&lang=es&site=eds-live.

Debord, G., 2014. *Theory of the dérive and definitions (1958).* Taylor and Francis.

Downs, R. M., Stea, D., 1975. Image and Environment: Cognitive Mapping and Spatial Behavior. *Geography*, 60(4), 327. http://0-search.ebscohost.com.millenium.itesm.mx/login.aspx?direct=true&db=edsjsr&AN=edsjsr.41415067&lang=es&site=eds-live.

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C., 2016. Deep Learning the City: Quantifying Urban Perception At A Global Scale. *ECCV*. http://streetscore.media.mit.edu/.

Gibson, J. J., 2014. *The theory of affordances (1979).* Taylor and Francis.

González Cacho, T., Sol Martínez, D., Sáenz de Miera Cuatloyol, J. A., Clavijo Plourde, D., García, H., 2017. URBAN MULTISENSORY LABORATORY, AN APPROACH TO MODEL URBAN SPACE HUMAN PERCEPTION. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W3, 29–35. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-4-W3/29/2017/.

juyong Park, Newman, M. E. J., 2005. A network-based ranking system for US college football. *J Stat Mech*, P10014.

Lynch, K. A., 2014. *The city image and its elements (1960).* Taylor and Francis.

Marcel, A. J., 1983. Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, 15(2), 238 - 300. http://www.sciencedirect.com/science/article/pii/0010028583900105.

Naik, N., Philipoom, J., Raskar, R., Hidalgo, C., 2014. Streetscore - Predicting the Perceived Safety of One Million Streetscapes. *CVPR Workshop on Web-scale Vision and Social Media*. http://streetscore.media.mit.edu/.

Nasar, J. L., 1990. The evaluative image of the city. *Journal of the American Planning Association*, 56(1), 41. http://0-search.ebscohost.com.millenium.itesm.mx/login.aspx?direct=true&db=bth&AN=9608140025&lang=es&site=eds-live.

Quercia, D., 2012. How to crowdource your happy space. Accessed 16/03/2019.

Rapoport, A., 1978. Nomadism as a Man-Environment System. *Environment and Behavior*, 10(2), 215-246. https://doi.org/10.1177/0013916578102005.

Schachtel, E., 1941. The Dynamic Perception and the Symbolism of Form. *Psychiatry*, 4(1), 79-96. https://doi.org/10.1080/00332747.1941.11022323.

van Veen, H. A., Distler, H. K., Braun, S. J., Bülthoff, H. H., 1998. Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Comp. Syst.*, 14, 231-242.