
PREDICTING VIRGINIA BEACH RESIDENT SATISFACTION RATE WITH MACHINE LEARNING

A PREPRINT

Alan Zhai

Computer Science and Environmental Science
University of Virginia
asz9qm@virginia.edu

Shaun Kim

Department of Computer Science
University of Virginia
sk5ps@virginia.edu

Jaspreet Ranjit

Department of Computer Science
University of Virginia
jr4fs@virginia.edu

July 30, 2020

1 Abs

We are interested in city satisfaction score of the people of Virginia Beach varied by different aspects. As students who are interested in urban planning, it is important for us to know what people want in a city. More personally, in the future, when we are looking to settle down at certain locations, this model will provide us some guidance as to which city will provide a better quality of life. In macro-scale, although the scope of the data is limited to Virginia Beach, this model that we are going to build will have a greater impact and may be applicable to other urban areas across the country.

2 Dataset

This is a Virginia Beach Satisfaction survey, it contains 3028 rows and 69 columns. This dataset was published on March 3rd, 2018, while the survey itself was conducted in 2015. The purpose of the dataset is to provide timely and accurate city information to increase government transparency. The majority of the questions have categorical answers: such as satisfaction, ranging from very dissatisfied to very satisfied; agreement, ranging from strongly disagree to strongly agree; and yes-no questions.

<https://data.vbgov.com/Finance-Government/Citizen-Satisfaction-Survey/8utb-pcec>

3 Related Work

The use of machine learning in the field of education is quite common these days with Educational Data Mining becoming (EDM) an ever more robust field of Data Science. Here we list some of the many research papers and articles that have looked into our focus of interest:

1. ***Urban big data analytics and morphology (2019) by Martin Behnisch, Robert Hecht and Hendrik Herold***

This study essentially is cross-sectional and longitudinal analysis of urban land use patterns of 230 city regions in 34 European countries. By creating their own metrics that takes building stocks, human foot prints and activities into consideration and measuring how the cities influence human activities, the study concludes that lack of a basic public service could be associated with the spatial segregation of an area. This study in particular may not have used machine learning algorithms to draw such a conclusion, but it was most similar to our goal of predicting human satisfaction based on different aspects of the city.

2. ***The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images (2017) by Yu Ye, Wei Zheng***

The study on the visual quality of streets investigates the use of machine learning algorithms and artificial neural networks to train and evaluate a model that processes screenshots of street conditions in the town centre of Shanghai. The machine learning algorithm provided insight into the quality of the streets in a quantitative way, rather than relying on traditional subjective impressions and feelings. The model extracted relevant pixels that were representative of the visual quality of streets and used this information to come up with an evaluation of its condition. These predictions can be used to maintain street quality and expedite the urban planning process thus allowing city planners to evaluate and attend to low quality areas more efficiently. This study demonstrated the use of machine learning models for urban planning where the results have implications beyond improving the city planning process, but also indirectly contributing to higher satisfaction in residents.

3. ***Land-use-change modeling using unbalanced support-vector machines (2009) By Bo Huang, Chenglin Xie, Richard Tay, Bo Wu***

This study uses an unbalanced support-vector machine to predict land use change in order to provide a stronger model and address important issues in modeling land use change such as the imbalance of changed/unchanged cells. The main factors used to predict land use change are: economic factors, such as job growth and rent; social factors, such as affluence, and human attitudes; collective rule making factors, such as land zoning, and tenure; and other factors such as level of technology. Due to the ratio of unchanged cells, massively outweighing the number of changed cells, an unbalanced SVM was used instead of normal SVM. This allowed the model to sacrifice overall predictive accuracy in order to increase the minor predictive accuracy which is important for land use change modeling as changed cells only account for a small portion of the dataset. This demonstrates that a well optimized SVM out performs commonly used spatial logistic regression models in order to predict land use change.

Many of the urban planning academic journals and articles utilizes Machine Learning on geo-spatial data and how to maximize space and allocate resources. However, we were not able to find the direct correlation or an attempt to predict the satisfaction level based on different elements of the city. Thus, we hope our experiment to add value to the academic field.

4 Intended Experiments

Our main contribution from investigating the Citizen Satisfaction dataset, and developing a model using machine learning algorithms will be to predict citizen satisfaction on the basis of the provided features. Our primary experiment will be to analyze which aspects of Virginia Beach contribute most to resident satisfaction, develop a report of areas that require the most improvement, and use the model to predict people's satisfaction on the basis of the most relevant features. Not only this dataset provides lots of insight into the city of Virginia Beach, but also has a robust information of the demographic data as well. As a result, our second major contribution from our findings and analysis will be the investigation of the provided demographic data, and further analysis of which groups (based on gender, income etc.) are more satisfied. These experiments will provide a better understanding of what urban city planners need to direct their efforts towards to improve citizen satisfaction, and provide valuable insight as to what aspects contribute to a better standard for living, which goes beyond the scope of Virginia Beach, and can be applied to other cities.