Use of Machine Learning in Exploring Spatial (In)Justices[1]

Debzani Deb
Department of Computer Science
Winston-Salem State University
Winston-Salem, USA
debd@wssu.edu

Russell M. Smith
Department of History, Politics and Social Justice
Winston-Salem State University
Winston-Salem, USA
smithrm@wssu.edu

***Abstract*— In light of recent local, national and global events, spatial justice provides a potentially powerful lens by which to explore a multitude of spatial inequalities. For more than two decades, scholars have been espousing the power of spatial justice to help develop more equitable and just communities. However, defining spatial justice and developing a methodology for quantitatively analyzing spatial justice is complicated and no agreed upon metric for examining spatial justice has been developed. Instead, individual measures of spatial injustices have been studied. One such individual measure of spatial justice is economic mobility. Recent research on economic mobility has revealed the importance of local geography on upward mobility and may serve as an important keystone in developing a metric for multiple place-based issues of spatial injustice. As a result, this paper seeks to explore place-based variables within individual census tracts in an effort to understand their impact on economic mobility and potentially spatial justice. The methodology relies on data science and machine learning techniques and the results show that the deep leaning model is able to predict economic mobility of a census tract based on its spatial variables with 89% accuracy. In the end, this research will allow for comparative analysis between differing geographies and also identify leading variables in the overall quest for spatial justice.**

Keywords— spatial justice, economic mobility, data science, geographic information systems

I. Introduction

Spatial justice as a theoretical concept holds much promise for exploring, understanding and solving issues of spatial inequality in a wide variety of landscapes [1]. According to Rocco, "Spatial Justice refers to general access to public goods, basic services, cultural goods, economic opportunity and healthy environments" [2]. Numerous scholars have used the concept to call for a more equitable future for millions of people across the globe, as a theory by which planners should create more spatially just cities and as a political agenda to drive social change [3] [4]. Achieving spatial justice would be a means by which to address the inequitable distribution of public and private goods, services and resources.

However, the 'real world' application of spatial justice leaves much to be desired. From questions about the definition of spatial justice, to issues of tackling past, current and/or future spatial injustices, to making the larger public aware of the concept and its potential; spatial justice as a working concept is still in its infancy. With this in mind, the goal of this paper is to advance our collective understanding of spatial justice as it relates to measuring spatial justice. It is critical to develop quantitative techniques through which spatial justice can be explored across a variety of geographic landscapes. This paper begins this process by utilizing a combination of data science, machine learning, and geographic information systems techniques to explore place-based variables that influence spatial justice in order to make objective data-driven decisions on issues of spatial justice/injustice.

The uneven distribution of public and private goods and services across the urban landscape has created numerous issues for the larger society. From poor performing schools, to issues of gentrification, to food insecurity, spatial inequalities have become a byproduct of the capitalistic, market-driven, private property economic system that has become common across the globe. Numerous scholars have explored these singular issues of spatial inequalities including studies on education, health, housing, food, transportation and parks [5, 6, 7, 8, 9, 10]. However, little scholarship has been focused on exploring how these manifestations of spatial inequality are connected.

The one exception has been recent research focused on exploring issues of economic mobility in the United States [11] While not an outright measure of spatial (in)justice, the examination of intergenerational economic mobility is rooted in many factors that are at their core spatial. Reference [11] study on how a person's probability of moving from the bottom 20% of the income ladder to the top in a generation revealed the importance of five factors: residential segregation, income inequality, primary school education, social capital and family stability. Additionally, the study found that *where one is born and raised* has a causal effect on long-term economic outcomes. Being born on the 'right' side of the tracks/river/highway can

[1] NSF Award 1600864, FY21-23 UNC ROI Award

978-1-7281-8470-8/20/$31.00 ©2020 IEEE
DOI 10.1109/ICMLA51294.2020.00191

determine a person's ability to reach higher income levels than his/her parents. Indeed, this work places local geography at the center of the debate regarding economic mobility [11].

The fact that local geography affects economic outcomes is likely unsurprising to the millions of Americans who face an array of spatial injustices on a daily basis, which include unfair siting of environmental hazards, targeted school district assignments, and exclusionary zoning practices [1, 12, 13, 14]. Hence, understanding the particular features of local geographies that either promote or hinder upward mobility is critical and has been the focus of recent studies. As a result, this paper presents a data-driven approach that seeks to explore place-based variables within individual census tracts in an effort to understand their impact on economic mobility and potentially on spatial justice. In the end, this research will allow for comparative analysis between differing geographies and also identify leading variables in the overall quest for spatial justice.

Based on the objective of characterizing spatial justice, and empirically evaluating the impact of spatial factors on economic mobility and in turn spatial justices, the contributions of this paper are as follows. First, spatial justice is effectively characterized by a set of attainable public goods, basic services, cultural goods, economic opportunity and healthy environments factors within a specific geographic area. Second, the feasibility of using these spatial factors in distinguishing upward economic mobility is carefully assessed by performing a correlation analysis. Third, a machine learning based approach is presented to predict upward economic mobility for children who grow up in a particular geographic area based on place-based variables extracted in step 1. A set of classification algorithms such as k-Nearest Neighbor (kNN), Support Vector Machine (SVM), Random Forest (RF), and Deep Neural Network (DNN) are considered by comparing their testing accuracies in economic mobility classification.

The rest of this paper is organized as follows. Section 2 discusses related works, then Section 3 elaborates the detailed methodology. Section 4 discusses the experimental setup and evaluation results. Section 5 concludes the paper.

II. Background and Related Works

While not explicitly called 'spatial justice', the theoretical concept of prioritizing the geography of justice has been around for several decades and finds its modern-day roots in the work of Lefebvre, Harvey and Pirie [15, 16, 17]. Lefebvre developed a concept called the 'right to the city', in which he calls upon society to reclaim the city for 'all' in the face of increasing levels of commercialization, privatization and public-private partnerships [15]. Harvey builds upon Lefebvre's 'right to the city' and believes that geography cannot remain disengaged, impartial and objective, when many ills confront cities across the planet. As a result, he calls on geographers and others to bring together spatial and social analysis to improve urban spaces [16]. Pirie discusses the idea of 'territorial social justice' and is perhaps the first person to use the term spatial justice in an academic paper. He states that, "Surely it would be another string in their bow if geographers could answer questions such as these: is a person's living at place x just? Is the spatial distribution of grocery stores just? Is the siting of some new airport just? Is the re-siting of the hospital just? Is the removal and rehousing of squatters just? These questions range over the justness of absolute and relative location as well as over the justness of processes of siting and relocation" [17].

However, a concrete definition is still being developed. In his book, Seeking Spatial Justice, Soja does a masterful job discussing the importance of spatial justice, applications of spatial justice and the need for planners to engage in proactive spatial justice efforts, but his pivotal work leaves much to be desired as it relates to providing a concrete definition of spatial justice [1]. The closest Soja comes is an affirmation of what spatial justice should be… justice has a geography and that the equitable distribution of resources, services, and access is a basic human right. Meanwhile, Rocco states that "Spatial Justice refers to general access to public goods, basic services, cultural goods, economic opportunity and healthy environments" [2]. Similar to Soja's idea but with a little bit more detail.

Absent a fully formed and agreed upon definition, most scholars have opted to provide the characteristics that would help make a place more spatially just. These characteristics tend to focus on three fundamental components: access, equity and opportunity. Soja was interested in how differing geographies have access, opportunities and equity as it relates to resources and services. Rocco goes a step further and adds public goods, basic services, cultural goods, economic opportunity and healthy environments to the list of features that the population should have equal access to, opportunities for and equitable distribution of. Fainstein offers her own opinion on how planners can contribute to what she calls 'The Just City" by focusing on three factors: democracy, diversity and equity [18]. In the end, the concept of spatial justice and related ideas provides an interesting lens for exploring issues of geographic inequality.

In addition to defining the term spatial justice, an important consideration is how to measure it. Specifically, which locational based variables are worthy of study. As stated above, individual studies of spatial injustices are quite common. These studies exploring spatial inequalities across geographies focus on environmental injustices, education, healthcare, transportation, and parks to name a few [12, 5, 6, 8, 9, 10]. However, they do not provide a complete picture of the spatial injustices that may be occurring at local geographies and this paper seeks to begin developing a more robust and holistic exploration of spatial injustices in the belief that communities that suffer from one spatial injustice often have additional underlying concerns of injustice.

Recent academic research into economic mobility may provide an opportunity to bridge the gap that exists in the literature and begin the process of understanding the complex relationship numerous factors have on creating spatial injustices for certain communities. Reference [11] highlights the importance of place on economic mobility for the poorest populations in the United States [11]. Specifically, Chetty et al. explored a wide variety of place-based variables and the influence they have on economic opportunity for the poorest populations. In the end, their research found that place matters. Reference [19] determined that five factors are strongly correlated to these results: residential segregation, income inequality, school quality, social capital and family structure. Several of these factors have a strong place-based component including: residential segregation, income inequality, school quality and social capital from which this study will build upon. Additionally, [19] stated the need for research on the relationship between economic mobility and location at "narrower geographies" in an effort to understand the microgeographic attributes that may influence economic mobility and in turn spatial injustices.

Researchers have identified many location-based factors relevant to economic mobility and as a result - potentially spatial justice, such as educational opportunities, de facto and de jure racism, quality of family networks, and specific geographic characteristics [20, 21, 22, 23, 24, 25]. For example, [20] found a relationship between urban form and economic mobility. This work determined that the more compact a geographical area is, the higher upward economic mobility tends to be for residents in that area. In other words, sprawling built environments inhibit upward mobility and may play a role in exacerbating spatial inequalities. In the end, as a result of the diversity of variables that influence economic mobility it provides a starting point for building a more robust model and understanding of spatial justice.

Inspired by the sporadic connections made by the previous researches between location-based factors and economic mobility, the research presented in this paper aims to explore these relevancy in detail and utilizes a data-science and machine learning based approach to empirically evaluate the impact of place-based variables mentioned in Rocco's framework on economic mobility and potentially spatial justice.

III. Methodology

The methodology is proposed according to the following three steps: first, a list of features based on Rocco's elaboration on spatial justice are constructed (Section III.A); second, the feasibility of using these spatial variables in distinguishing different economic mobilities is evaluated by a correlation analysis (Section III.B); and third, economic mobility of an individual census tract is classified by applying a set of predictive models (Section III.C) utilizing spatial factors as features. The results are presented in Section 4. The study focuses on data acquired from 2157 census tracts within the USA state of North Carolina (NC).

A. Data Set and Preprocessing

The spatial feature set utilized in this study is constructed based on Rocco's reference of spatial justice as general access to public goods, basic services, cultural goods, economic opportunity, and healthy environments. Table I provides an overview of Rocco's characterization factors and corresponding location-based feature variables that were used in this study to measure them. The feature dataset is obtained from two primary sources. First, a majority of the feature values were obtained from NC OneMap, a repository of geographic based data located in the State of North Carolina [26], and were collected during 2018-2020 period. The NC OneMap website is the authoritative data collection for spatially based data in North Carolina and includes spatially referenced data on a multitude of categories including: boundaries, community safety, education, environment, recreation and transportation. Some of the economic opportunity feature values such as *Mean Travel Time to Work*, *Total Jobs* and *Job Density* were collected during 2015 and were obtained from reference [11]. *Upward Economic Mobility*, the outcome variable of interest for this study, is also calculated with reference [11]'s columns measuring income ranks for parents and children by census tract. More specifically, the target variable is calculated with income ranks for parents and children by census tract, and then is labeled as 0 when the difference between children and parents' income is equal or less than zero (meaning no or downward economic mobility), and 1 for when the difference in income rank is positive (meaning upward economic mobility). Reference [11] provides these and other mobility statistics, collected during 2015, in the hopes that researchers will use them to further shed light on intergenerational income mobility at the local level. This paper has adopted them as a surrogate for spatial justice at the census tract level.

TABLE I. Spatial Feature Varibles

Rocco's Factors	Spatial Feature Variables
Public Goods	Number of Schools, Water and Sewer Service, Fire Stations, Hospitals, Medical Facilities, Correctional Facilities
Basic Services	Gas Stations, Food Desert, Limited Broadband
Cultural Goods	Libraries, Colleges, Non-Public Schools
Economic Opportunity	Mean Travel Time to Work, Total Jobs, Jobs Density, Area covered.
Healthy Environment	Underground storage tanks, Brownfields, NPDES Sites, Hazardous Waste Facilities, Landfills

Once constructed, the dataset contains 2157 rows corresponding to NC census tracks with 22 numerical columns describing the feature variable as depicted in Table I and one binary target variable depicting *Upward Mobility*. During the preprocessing, the feature dataset is further explored for null values and outliers, and distribution of each feature column is explored. The data distribution reveals a large disparity in terms of scale among feature columns. To address this issue, the feature dataset is further normalized and standardized by utilizing python scikit-learns' StandardScaler function to get all features to have the same standard scale. In the original dataset, only 11% census tracts indicated *Upward Mobility,* while the rest demonstrated otherwise. Therefore, this research also enforced balanced classes by storing only as many samples from the negative class (*Upward Mobility* = 0, signifying no or downward economic mobility) as there are samples of the positive class (*Upward Mobility* = 1, signifying upward economic mobility).

B. Correlation Analysis

It is important that the constructed spatial features are evaluated in terms of their effectiveness and adequacy in characterizing upward economic mobility. In order to provide more insight in how effectively economic mobility varies with respect to individual spatial features, a correlation analysis is performed between the target *Upward Mobility* and the individual features and the outcome is shown in Table II. The most correlated or informative feature is *Jobs_total*, followed by the feature *Gas_Station_count* (strong negative correlation) and so on. The least correlated features are *Hospitals_count* and *NonPublic_Schools_count.* While this ranking does not mean that the few topmost ranked features constitute the most informative collection of features in characterizing *Upward Mobility,* as one can gain more information by combining features that complement each other, it is reassuring to observe the outcome variable being highly correlated with input features and to perceive the fact that there is further potential in utilizing these spatial features for the classification of economic mobility.

TABLE II. CORRELATION BETWEEN THE TARGET AND THE INDIVIDUAL FEATURES

Individual Feature	Correlation
Gas_Station_count	-0.406970
Fire_Stations_Count	-0.345119
NPDES_sites_count	-0.273248
Area	-0.270121
Medical_Facilities_count	-0.261866
Food_Desert	-0.247073
Limited_broadband_count	-0.229297
Public_Schools_count	-0.228285
Sewer_Plant_Count	-0.208590
Public_Water_Supply_Count	-0.204383
Landfills_count	-0.150027
Public_Library_Count	-0.149961
Colleges_count	-0.128597
Brownfield_count	-0.100617
Correctional_institutions_count	-0.088649
Hazardous_waste_count	-0.086062
Underground_Tank_Count	-0.072705
Hospitals_count	-0.023890
NonPublic_Schools_count	0.029359
Mean_commute_time	0.030846
Job_density	0.192169
Jobs_total	0.442078
Upward Mobility	1.000000

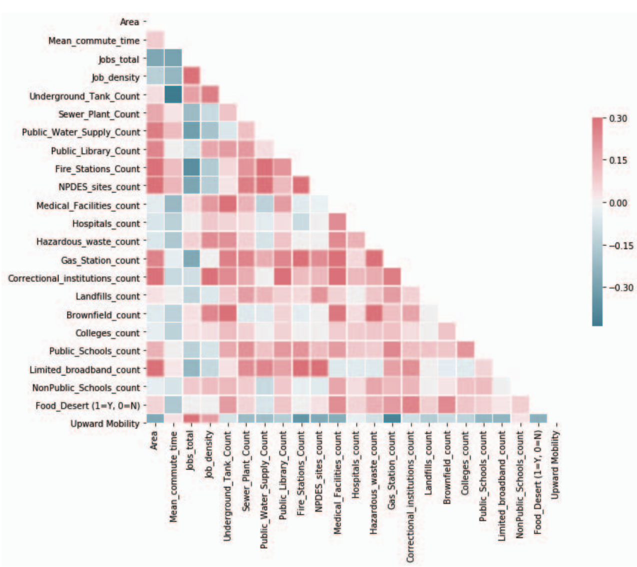

Fig. 1. Correlation matrix of all features

In order to better understand the adequacy of 22 spatial features identified in Section III.A, the correlation coefficients of all pairs of features is depicted in a color-coded plot as in Figure 1. White encodes that the feature pair is not correlated, red indicates a positive correlation and green indicates a negative correlation. The darker a color is, the larger is the absolute correlation coefficient. With this plot, one can spot a few highly correlated features that may represent similar and redundant place-based information. However, this research chooses not to use correlation as a guideline for feature selection as two correlated features can still improve classification when they are in the same collection of features [27].

C. *Classification Algorithms*

This study utilized four different classifiers such as k-Nearest Neighbor (kNN), Support Vector Machine (SVM), Random Forest (RF), and Deep Neural Network (DNN) in classifying upward economic mobility based on the spatial features. The choice for these classifiers was driven by various reasons as discussed below.

kNN is robust to work with and provides a fast classification. The kNN classifier takes every single spatial vector (composed of 22 place-based variables within a particular census tract) and locates it in feature space with respect to all training observations. The classifier identifies the k training observations that are closest (based on Euclidian distance) to the new observation. Then, it selects the label (upward mobility) that the majority of the k closest training observations have. This procedure requires no explicit training phase and the classifier merely stores all training observations and their labels in order to make predictions. For large datasets, the limitation of this method can be that not all data can be stored. This is not a problem for the presented research as both the feature space and data set are comparably low-dimensional. However, in future when the study will be extended to cover all geographical census tracts within USA or other countries, the kNN approach needs to be investigated more to assess its feasibility in making timely predictions.

Support vector machines are popular and powerful binary classifiers. SVMs divide the feature space by a hyperplane such that the margin between the two classes is maximized, i.e., SVMs squeeze a maximally thick layer between the boundary observations of both classes, known as support vectors. In contrast to kNN, SVM generalizes from the observed data, i.e., it does not store the individual observations once the training is performed and only saves the decision hyperplane. For more robustness against outliers, a small number of boundary observations are tolerated within the margin. A parameter C controls the trade-off between maximizing the margin and minimizing the number of such exceptions. For classes that are not linearly separable in feature space, the standard scalar products involved in the computation of the hyperplane can be replaced with 'kernels'. Kernels implicitly relocate the problem in another high-dimensional space where the classes are separable. In the same step, the kernel maps the found hyperplane back to feature space. The presented research used a Gaussian radial-basis function (*rbf*) as the kernel, parameterized by the width parameter *gamma*. The expectation is that, with the right values for parameters *C* and *gamma*, SVM improves accuracy for borderline data whose feature space location is between mobility-classes.

Random Forest is an ensemble tree-based learning algorithm. The RF Classifier consists a set of decision trees, each of them built over a random extraction of the observations from the dataset and a random extraction of the features. Not every decision tree in the set utilizes all the features or all the observations in the training dataset, and this guarantees that the trees are less correlated and more independent, and therefore less prone to over-fitting. Each tree uses a sequence of yes-no questions based on a single or combination of features in order to divide the training observations. At each node, the tree divides the dataset into 2 buckets, each of them hosting observations that are more similar among themselves and different from the ones in the other bucket. Therefore, the importance of each feature is derived from how "pure" each of the buckets is. The most widely used impurity measure is the *Gini impurity*, which is also utilized in this study. The classifier aggregates the votes from different decision trees to decide the final class of the test object. Random forests are one of the most popular machine learning algorithms because of the good predictive performance and their resistance to outliers. The insensitivity to outliers is a desirable characteristic for the current search, because some census tracts may have extremely high values for some spatial variables and the expectation is that the Random Forest will perform better in such cases.

Finally, deep learning, one of the most popular branches of machine learning, is also explored in order to investigate the potential of predicting upward economic mobility based on place-based features. The current study utilized a deep neural network (DNN) model containing four dense layers with *LeakyReLU* activation, as specified in Table III, with a single output signifying upward mobility or not. The model is optimized using *binary cross entropy* and is fitted with *Adam* version of stochastic gradient descent with *learning rate* being 0.0002 and *momentum* being 0.5. In total, the proposed DNN can handle 679,937 parameters and can train them all. Regularization technique *Dropout* is utilized to avoid overfitting and to get a more robust network that generalizes better.

Deep learning typically shines in learning complex patterns and relationship, when there are large number of labeled data is available. This is not specifically true for the current research as it relies on data extracted from 2157 census tracts only, however, the extended research will analyze all USA census tracts, which certainly will be a sizable problem for deep learning. This is enough justification for the current research to explore deep learning and evaluate its potential in discovering hidden patterns in the constructed dataset composed of many apparently independent variables.

IV. Experimental Evaluation

This section details an empirical evaluation of different machine learning algorithms on spatial data. The experiments were set up using Python and TensorFlow libraries. The curated dataset was divided into training (80%) and testing (20%) sets, and were utilized during model training and classification respectively. During the data split, stratified sampling was enforced to ensure that the right number of instances were sampled from each mobility subgroups to guarantee that the test

TABLE III. Structures and parameters used in DNN Model

Layer	Output_shape	No. of Parameters
Dense	(None, 1024)	23552
LeakyReLU	(None, 1024)	0
Dense	(None, 512)	524800
LeakyReLU	(None, 512)	0
Dropout	(None, 512)	0
Dense	(None, 256)	131328
LeakyReLU	(None, 256)	0
Dropout	(None, 256)	0
Dense	(None, 1)	257

set was representative of overall population. All classical classification algorithms (kNN, SVM, and RF) were fine-tuned by performing grid search over specified hyperparameter values. Grid search takes as input a large range of values for hyperparameters, uses cross-validation to train each model with all possible combinations of these hyperparameter values, and then identifies the model with best possible combination of hyperparameters. For kNN, the best value for number of neighbors was found to be 13. For SVM, the best values for *C* and *gamma* were identified as 10 and .001 respectively. For RF, the maximum depth of the tree, the number of features to consider when looking for the best split, and the number of trees hyperparameters were set to 30, 4, and 100 respectively by grid search.

The classifiers performances are evaluated using various standard evaluation metrics such as Precision, Recall, F1 and ROC score. In this study, Precision is the ratio of correctly predicted *Upward Mobility* observations to the total predicted *Upward Mobility* observations. Recall is the ratio of correctly predicted *Upward Mobility* observations to all actual observations with *Upward Mobility* labels. In other words, Precision and Recall are all interested in predicting the true answer of the positive label. F1 score takes both Recall and Precision into account, hence can be considered as a weighted average of them, and therefore it provides a useful accuracy indicator. The ROC curve is another common tool used with binary classifiers. The ROC curve plots the *true positive rate* (another name for Recall) against the *false positive rate* (FPR). The FPR is the ratio of negative instances that are incorrectly classified as positive. To visualize the performance of the classifier, Receiver Operating Characteristics (ROC) curve and Precision-Recall curves are introduced.

Table IV shows the evaluation metrics of the four classifiers utilized in this study. The best metrics in each column are shown in bold. In general, the DNN classifier performs better than the traditional methods considering all evaluation metrics. The deep learning model is fit for 100 epochs, however, *EarlyStopping Callback* is utilized to interrupt training when it measures no progress on the validation set for a number of epochs, and it rolls back to the best model in order to save time and avoid overfitting. For the presented DNN model, accuracy of training set reaches 94% after about 30 epochs, while accuracy of testing set ranges from 86% to 89%. These results show that the DNN network was not overfitted, and was able to achieve higher accuracy within a short period of time. Table IV also reveals that in case of kNN, SVM, and DNN classifiers, Recall is higher than Precision, which means that the real economically mobile census tracts are being classified as 'economically mobile' or 'spatially just' at a higher rate, whereas comparatively lower Precision means that the model is predicting many more census tracts as 'economically mobile' or 'spatially just' than actually there are. Figure 2 shows the Precision-Recall curves (PR-curves) for all classifiers. PR-curve is a very widely used evaluation method in machine learning. In general, the closer the curve is to the top-right corner, the more beneficial the tradeoff it gives between precision and recall. The PR-curve in Figure 2 shows the superiority of deep learning model in minimizing the number of false positives while ensuring high classification accuracy.

TABLE IV. CLASSIFICATION ACCURACY ON TEST DATA

Classification Algorithm	Precision	Recall	F1	ROC
k-nearest neighbor (kNN)	0.79	0.88	0.83	0.823
Support Vector Machine (SVM)	0.80	0.85	0.83	0.823
Random Forest (RF)	0.85	0.83	0.84	0.844
Deep Neural Network (DNN)	**0.88**	**0.89**	**0.88**	**0.885**

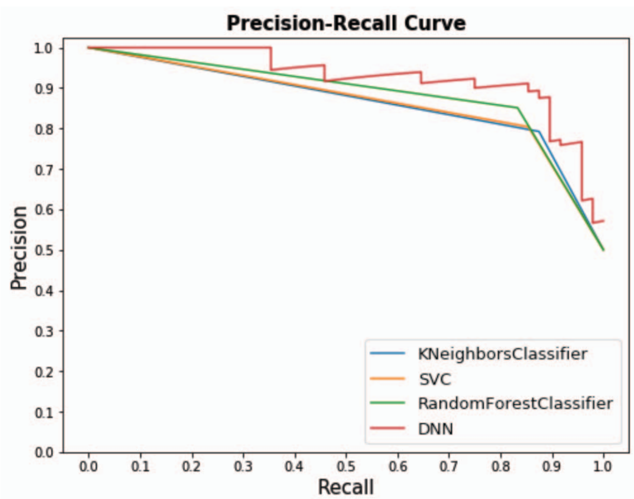

Fig. 2. Precision-Recall Curve

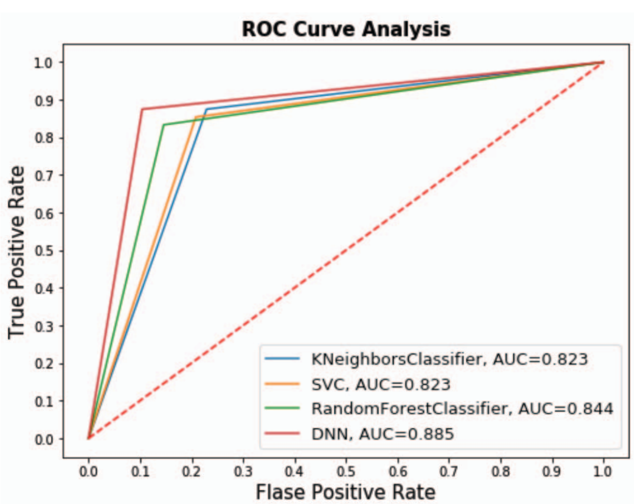

Fig. 3. ROC Curves and corresponding AUC values

In addition, ROC analysis is considered for all the classifiers. Figure 3 shows the ROC curves and the corresponding AUC values of all models. The ideal point in ROC space is the top-left corner. AUC is an important statistical parameter for evaluating classifier performance: the closer AUC is to 1, the better overall performance of established classifier. In the current work, as shown in Figure 3, the AUC value of DNN model is .885, which is higher than the other classical machine learning models with a significant margin (4% or more), indicating that the DNN model achieves better performance than the other classifiers. The deep parameterization process of DNN, where 679,937 parameters were trained and learned, cannot match that of RF, SVC or kNN which needs lesser number of parameters to be tuned. Based on this fact, DNN shows great potential for classifying upward economic mobility for children who grow up in a particular geographic area based on place-based variables.

V. CONCLUSION AND FUTURE WORKS

Through the application of data science and machine learning technologies, this study provides a way of characterizing spatial justice, and empirically evaluating the impact of spatial factors on economic mobility and in turn spatial justices. The specific contributions of this paper are: 1) effectively characterizing spatial justice by a set of attainable public goods, basic services, cultural goods, economic opportunity and healthy environments factors within a specific geographic area; 2) performing correlation analysis in assessing the feasibility of using these spatial factors in distinguishing upward economic mobility; and 3) presenting a machine learning based approach to automatically classify upward economic mobility for children who grow up in a particular geographic area based on place-based variables. The experimental results show a strong correlation between upward mobility and chosen array of spatial variables. The comparative performance analysis of four classification algorithms reveals the superiority of deep learning model in predicting upward mobility based on spatial variables with 89% accuracy. In future, the capability of these predictive models will be tested on a national-scale data set. The future works will also include fine tuning the DNN model and analyzing its ability to shed light into understanding the impact of each features on the predictive results.

The analysis conducted in this paper shows that it is possible to begin to develop a more robust and comprehensive spatial justice metric through the inclusion of the most critical locational variables in determining how 'just' a local landscape may be. The results of this study will be utilized to inform future research focused on the development of a spatial justice index. The spatial justice index is envisioned to be a metric that can be used to quantitatively compare diverse geographies based upon the level of spatial (in)justice at each location. Through the development of a spatial justice index, planners, public officials, policy makers, activists, concerned residents, etc. will be armed with a new tool to fight spatial injustices facing communities across the globe.

ACKNOWLEDGMENT

We would like to acknowledge the support provided by FY21-23 UNC Research Opportunities Initiative (ROI) award and NSF Award #1600864.

REFERENCES

[1] Soja, Edward W. Seeking spatial justice. Vol. 16. U of Minnesota Press, 2013.

[2] Rocco de Campos Pereira, R. C. "Why discuss Spatial Justice in urbanism studies." Atlantis, 24 (4), 2014.

[3] Fainstein, Susan S. "The just city." International journal of urban Sciences 18, no. 1 (2014): 1-18.

[4] Williams, Justin. "Spatial justice as analytic framework." PhD diss., 2018.

[5] Wei, Yehua Dennis, Weiye Xiao, Christopher A. Simon, Baodong Liu, and Yongmei Ni. "Neighborhood, race and educational inequality." Cities 73 (2018): 1-13.

[6] Jones, A., H. M. Mamudu, and G. D. Squires. "Mortgage possessions, spatial inequality, and obesity in large US metropolitan areas." Public Health 181 (2020): 86-93.

[7] Rodríguez-Pose, Andrés, and Michael Storper. "Housing, urban growth and inequalities: The limits to deregulation and upzoning in reducing economic and spatial inequality." Urban Studies 57, no. 2 (2020): 223-248.

[8] Garcia, X., Garcia-Sierra, M., & Domene, E. (2020). Spatial inequality and its relationship with local food environments: The case of Barcelona. Applied Geography, 115, 102140.

[9] Liu, Chengliang, and Dezhong Duan. "Spatial inequality of bus transit dependence on urban streets and its relationships with socioeconomic intensities: A tale of two megacities in China." Journal of Transport Geography 86 (2020): 102768.

[10] Chang, Zheng, Jiayu Chen, Weifeng Li, and Xin Li. "Public transportation and the spatial inequality of urban park accessibility: New evidence from Hong Kong." Transportation Research Part D: Transport and Environment 76 (2019): 111-122.

[11] Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." American Economic Review 104, no. 5 (2014): 141-47.

[12] Bullard, Robert. "Dumping in dixie." Boulder, CO: Westview(1990).

[13] Weiher, Gregory. The fractured metropolis: Political fragmentation and metropolitan segregation. Suny Press, 1991.

[14] Orfield, M. "Metropolitics: A regional agenda for community and stability. Washington, DC: Brookings Institution." (1997).

[15] Lefebvre, Henri. Le droit à la ville. Vol. 3. Anthropos: Paris, 1968.

[16] Harvey, David. "Social justice and the city (revised edition)." Athens, GA: University of (1973).

[17] Pirie, Gordon H. "On spatial justice." Environment and Planning A 15, no. 4 (1983): 465-473.

[18] Susan, Fainstein. "The just city." Nueva York: Cornell University Press-Cornell Paperbacks (2010).

[19] Chetty, Raj, and Nathaniel Hendren. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." The Quarterly Journal of Economics 133, no. 3 (2018): 1107-1162.

[20] Ewing, Reid, Shima Hamidi, James B. Grace, and Yehua Dennis Wei. "Does urban sprawl hold down upward mobility?." Landscape and Urban Planning 148 (2016): 80-88.

[21] Altzinger, Wilfried, Jesus Crespo Cuaresma, Bernhard Rumplmaier, Petra Sauer, and Alyssa Schneebaum. Education and Social Mobility in Europe: Levelling the Playing Field for Europe's Children and Fuelling its Economy. No. 80. WWWforEurope Working Paper, 2015.

[22] Corak, Miles. "Income inequality, equality of opportunity, and intergenerational mobility." Journal of Economic Perspectives27, no. 3 (2013): 79-102.

[23] Black, Sandra E., and Paul J. Devereux. Recent developments in intergenerational mobility. No. w15889. National Bureau of Economic Research, 2010.

[24] Hardaway, Cecily R., and Vonnie C. McLoyd. "Escaping poverty and securing middle class status: How race and socioeconomic status shape mobility prospects for African Americans during the transition to adulthood." Journal of youth and adolescence 38, no. 2 (2009): 242.

[25] Delgado, Richard. "The myth of upward mobility." University of Pittsburgh Law Review 68, no. 4 (2007): 879.

[26] Portal, NC OneMap Geospatial. "NC OneMap Geospatial Portal." North Carolina Department of Information Technology, Data Division, Center for Geographic Information and Analysis (2020).

[27] Guyon, I. and Elisseff, A. "An introduction to variable and feature selection," Journal of Machine Learning Reserach., vol. 3, pp. 1157–1182, Mar. 2003.