DOI: 10.1111/mice.12972

INDUSTRIAL APPLICATION

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING

WILEY

Interpretable machine learning learns complex interactions of urban features to understand socio-economic inequality

Chao Fan[1] | **Jeffrey Xu**[2] | **Barane Yoga Natarajan**[2] | **Ali Mostafavi**[1]

[1]Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, Texas, USA

[2]Department of Computer Science and Engineering, Texas A&M University, College Station, Texas, USA

Correspondence
Ali Mostafavi, Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, 77840, USA.
Email: amostafavi@civil.tamu.edu

Funding information
National Science Foundation under Grant, Grant/Award Number: CMMI-1846069; Microsoft Azure AI for Public Health Grant; Texas A&M University X-Grant, Grant/Award Number: 699

Abstract

Inequality in cities is a phenomenon arising from the complex interactions among urban systems and population activities. Conventional statistics and mathematical models like multiple regression models require assumptions of feature interactions with specified mathematical forms that may fail to fully capture complex interactions of heterogeneous urban components, creating challenges in systematically assessing socio-economic inequality in cities. To overcome the limitations of these conventional mathematical models, in this work, we propose an interpretable machine learning model to capture the complex interactions of urban variables and the main interaction effects on socio-economic statuses. We extract urban features from high-resolution anonymized mobile phone data with billions of activity records related to people and facilities in 47 US metropolitan areas and predict the attributes of urban areas from six income and race groups. We show that socio-economic inequality in cities can be effectively measured by the predictability of trained machine learning models in controlled experiments. We also examine the tradeoff between spatial resolution, sample size, and model accuracy; test the presence of influential features; and measure the transferability of the trained models to identify the optimal values for controlled factors. The results show that metropolitan areas share similar patterns of inequality, which could be moderated by improved polycentric facility distribution and road density. The generality of associated factors and transferability of machine learning models can help bridge data gaps between cities and inform about inequality alleviation strategies. Despite similarities, 50% to 90% of variations among cities are still present, which shows the need for localized policies for inequality alleviation and mitigation. Our study shows that machine learning models could be an effective approach to examine inequality, which opens avenues for more data-centric and complexity-informed planning, design, policymaking, and engineering toward equitable cities.

1 | INTRODUCTION

Inequality in metropolitan cities has become one of the cornerstone social and economic issues of our age, prompting a debate about the measurement and solutions and fueling public discontent with the built environment and society (Woetzel et al., 2017). Despite great effort (Acemoglu & Robinson, 2009; Balland et al., 2020) having been applied to research and practice for measuring and mitigating inequality, systematic divergence from the optimal equality of facility services and life opportunities in cities still exists (Mirza et al., 2021), a situation that is not well understood. We hypothesize that a contribution to such divergence arises from neglecting to examine equality as an outcome of the complex interactions between the population and the built environment in urban areas (Fan et al., 2021a). Capturing this mechanism by a computational metric can help measure and explain the presence of inequality, pinpoint potential solutions for mitigating inequalities, and inform policy and design promoting equitable cities (Xue et al., 2022).

Understanding and improving socio-economic equality in metropolitan cities is a long-lasting challenge (Acemoglu & Robinson, 2009). A growing and diverse number of studies (Gazzotti et al., 2021) have been investigating this phenomenon over the past two decades. Conventional research (Marger, 1999) mainly focuses on a theoretical understanding of social and economic inequality. The problems of inequality arise with the stratification of socio-economic classes and relations, characterized by income concentration (Thomas & Emmanuel, 2014). The most controversial topics related to income inequality previously focused on the distribution of wealth. As research progressed and cities developed, studies on this front started addressing the inequalities present in people's lives, such as satisfactory public services (Anand & Ravallion, 1993), accessibility to life needs, availability of social capital (Dahl & Malmberg-Heimonen, 2010), and opportunities of higher education (Triventi, 2013). The economic inequality intertwining social needs increases the complexity of the inequality assessment problem. Literature (Cingano, 2014) has been attempting to establish connections between urban features and socio-economic status of people. Theoretical studies, however, are not fully considering socio-economic inequality as a multidimensional phenomenon.

Urban inequality represents the level of disparity in diverse socio-economic contexts across different areas of a city, which has been unveiled in a variety of aspects including infrastructure services and population activities (Casali et al., 2021). The infrastructure statuses and human activities are heterogeneous and dynamic, leading to high variations in socio-economic patterns (Niu et al., 2020). Recall that inequality is defined from such a variation that exists in the relationship between urban components and socio-economic patterns. Quantifying the variation in socio-economic patterns is one of the key steps to evaluating inequality in cities (Li et al., 2019). In addition, the interactions between the built environment and population activities are nuanced and non-linear as a result of the different paces of the dynamic urban components including socio-economic activities of populations and the evolution of the infrastructure and the environment (J. Wang et al., 2019). To understand the inequality arising from intertwined urban features, it is critical to capture the variation and the non-linearity in the interactions between heterogeneous urban components.

Machine learning, a method that captures information from a portion of samples and predicts the labels of the rest, provides an effective way to assess the variation present in the data samples (Adeli & Hung, 1994; Rafiei & Adeli, 2016). Inequality, in the socio-economic context, could be well considered as the variations in the relationship between input features and output labels of data samples. Hence, machine learning could be very helpful to address inequality in cities (Zhou & Liu, 2019). On the other hand, machine learning models are created in the way the complex and non-linear interactions of the features are modeled in an automated manner, without theoretical assumptions for formulating the equations (Rafiei & Adeli, 2017). Such an automated learning process is promising to connect the interactions of urban features with the non-linear model structures of machine learning (Ahmadlou & Adeli, 2010). Considering these capabilities, we could claim the fundamental connection between the inequality of cities and the predictability of machine learning models to inspire the adoption of machine learning to assess inequality.

Examining socio-economic inequality as a phenomenon based on population activity and built environment features cannot be fully implemented without the support of sufficient fine-scale data. Prior to the age of smart devices and technologies, it was notoriously difficult to collect and analyze fine-grained data about urban components, such as facilities and population activities and their interactions (Esmalian et al., 2022). The digital footprints that accumulate and aggregate on smartphones provide an efficient and effective proxy for investigating issues of inequality, as the mobile phone data reveal patterns of human movements and activities at greater temporal and spatial granularity while ensuring anonymity and user privacy (Moro et al., 2021). In addition, the availability of place data that describe the location, category, and brand of a place enables specifying the distribution of urban facilitie

and the development of the built environment, as well as population life activities. To harness the potential of these emerging location-based datasets, an increasing number of studies (Aleta et al., 2020) have employed these data in multiple research domains and have validated the scale and accuracy of these data. In particular, existing literature (F. Wang et al., 2019) has demonstrated that the location-based data could be highly demographically representative. Hence, the use of fine-scale location-based data can transform conventional measurement and understanding of inequality at a scale and in ways never attempted before (Milanovic, 2016).

More recently, benefiting from the explosion of urban data, data-driven inequality research (Fan et al., 2021b) has been growing significantly, and a transition from theoretical to data-driven inequality research has emerged (Mirza et al., 2021). One stream of work adopts and analyzes location-based data, such as mobile phone data and geotagged social media data. Researchers in this stream quantify the connection inequality of neighborhoods (Q. Wang et al., 2018), income inequality for resilience to natural disasters (Yabe & Ukkusuri, 2020), the racial inequality of probabilities of becoming infected in pandemics (Millett et al., 2020), and economic inequality of innovation activities and products (Balland et al., 2020) in cities. Another stream of research relies on public utility and empirical data, such as facility locations and survey data. These studies capture the inequality of facility distributions (Xu et al., 2020) and income inequality of hazard exposure (Rasch, 2017). These studies are largely based on datasets that document only single aspects of urban systems, such as social and physical connections (Dong et al., 2019), access to services (Johar et al., 2018), and interactions with the environment (Rao et al., 2017).

Cities, however, are complex systems involving a variety of interconnecting components, such as facilities, infrastructure, and populations (Pan et al., 2013). Devoting efforts to understanding and seeking equality based on individual components of cities is not nearly enough. An optimal socio-economic equality knowledge and solution require an integrative consideration of all urban components and their non-linear interactions. The question arises as to whether it is possible to predict the socio-demographic status of areas based on features related to population activities and the built environment and their interaction. This question is far from being answered by extant research due to the absence of consensus on ways of measuring inequality by concurrently incorporating features of the built environment and population activities, as well as the non-linear interactions among the features. Traditional linear mathematical models are insufficient to encode the non-linearity in urban systems in examining inequality.

Conventional mathematical models like multiple regression models have been widely adopted to examine the effect of independent variables on the dependent variable in the context of social science and urban development. In these complex study areas, independent variables also commonly interact with each other. That means, the relationship between an independent variable and the dependent variable changes when the independent variable interacts with another independent variable and the value of the third variable changes. This type of effect makes the underlying mechanism of variable relationships more complex. But this is, in fact, how the real world behaves, and it is critical to incorporate it into the model. Conventional mathematical models call it interaction effect.

The interaction effect in conventional mathematical models is examined in a couple of ways, such as incorporating the multiplication of two variables in the regression model to consider both the main effect and the interaction effect of the variables at the same time. This conventional method works well to consider the interaction effects, indicating that the relationship between an independent variable and the dependent variable depends on the value of another independent variable. The conventional methods, however, have two assumptions. First, the methods assume that the interaction effects of the variables follow the multiplication relationship. Second, the value of the dependent variable is a linear combination of the main effects of individual independent variables and the interaction effects of multiple independent variables. That is, conventional mathematical methods require that the relationship between the dependent variable and the independent variables and the interactions of independent variables need to be specified before testing the models on real-world data.

The interactions of urban environment features are particularly complex. Without fully understanding the mechanism of how these features are interacted and influence the dependent variable, it is challenging and problematic to specify the relationship in the mathematical model, especially in a case of a great number of independent variables. To overcome the limitations of these conventional mathematical models, here, we propose an interpretable machine learning model to automate the process of capturing the complex interactions of independent urban variables and the main and interaction effects on the dependent variable (socio-economic attributes). The proposed machine learning method can encode both the built environment and population activity features. The method advances our understanding of variable interactions, which releases the constraints of specifying the interaction terms and the linear combination of multiple effects in existing mathematical models, which

will provide fundamental insights into interpreting the effects of urban development, human activities, and landscape change on socio-economic inequality in cities. With that, we claim that the proposed interpretable machine learning model outperforms conventional mathematical models.

The core idea of this study is that inequality can be identified and measured in cities using machine learning. Machine learning enables capturing various heterogeneous urban systems and population features and their interactions; if the socio-economic status of different areas could be predicted accurately by machine learning models using population activity and built-environment features and their non-linear interactions, then inequality exists. In other words, if equality is present, features of population activities and the built environment would not vary drastically across high-income versus low-income and minority versus non-minority areas. Hence, the prediction performance metrics of machine learning models could be used to measure the extent of inequality. The high predictability of models indicates greater socio-economic inequality in cities. Also, it could be evidence that inequality is a phenomenon that may not be attributed to individual features but rather to the complex interactions among various features in cities if individual features alone cannot explain the predictability of machine learning models.

We first created grid maps for 47 US Metropolitan Statistical Areas (MSAs), assigned socio-economic labels of census block groups (CBGs) to grid cells within block groups, and computed features for each grid cell. The considered features of urban components draw upon multiple sources of data, including 1 million points of interest (POIs) data, billions of anonymized mobile phone data, and more than 10,000 social-economic records for CBGs. The mobile phone data covers population activities during the first week of April 2019, which is considered a stable period, portraying regular human life activities. Two advanced machine learning (ML) models, XGBoost and neural network models, were trained and tested. We considered the predictability of the machine learning models, quantified by F1 scores, as a metric for evaluating models' prediction performance and, accordingly, as a measure of inequality in a city. To demonstrate the effectiveness and reliability of the metric, we investigated the tradeoff between grid size and accuracy and tested the influence of individual features on the predictability of the models. Furthermore, we demonstrated the cross-MSA generality of inequality patterns by training a model in one MSA and then applying it directly to other MSAs. The transferability of machine learning models can imply sharable inequality patterns and quantify variations across MSAs. We further examined the relationship between inequality metrics and urban characteristics, including road density and facility distribution in MSAs to explore potential solutions for alleviating inequalities. Finally, a conventional mathematical model, ridge regression model, is used to demonstrate the performance and capabilities of machine learning models in capturing the complex interactions of urban features. The study serves as an effort toward data-driven and ML-based scientific discovery to address urban policy challenges such as infrastructure planning to combat urban inequality.

2 | METHODS

2.1 | Data collection and processing

This study focuses on MSAs in the United States. We selected the MSAs based on three criteria. First, the population size of the MSA should be sufficiently large to serve as an object of study. Hence, the MSAs selected in this study are ranked in the top 50 in terms of the sizes of residential populations. Second, the selected MSAs should cover different regions of the United States, to consider the regional effects in concluding the general patterns of socio-economic inequality in cities. Finally, both public and private datasets should be available for the selected MSAs. Considering these criteria, we end up with 47 MSAs for analyses in this study. A complete list can be found in the Supplementary Information.

2.1.1 | Grid and label creation

To understand the fine-scale socio-economic disparities in cities, we divided the area of an MSA into grid cells of relatively equal size (see Figure 1). We considered one side of a grid cell as spanning a certain range of latitude or longitude. We started with 0.01 degree as the length of the side of grid cells and tested different values from 0.01 to 0.05 degrees with a step size of 0.01 degree. As the grid cells get larger, more facility and human activity information will be covered by a grid and integrated to represent the features of the grid cells. We used grid cells with a side of 0.01 for all analyses in this study and also showed that this is a proper selection for the size of grid cells.

To compare the features of different urban areas, we collected socio-economic public data including per capita income and race–ethnicity data from the US Census 2014–2018 (5 years) American Community Survey (ACS) at census tract level of spatial aggregation (United States Census Bureau, 2019). We focused on the three largest race–ethnicity groups as determined by self-identification in the Census: White, Black or African American, and Hispanic (Q. Wang et al., 2018). These three population

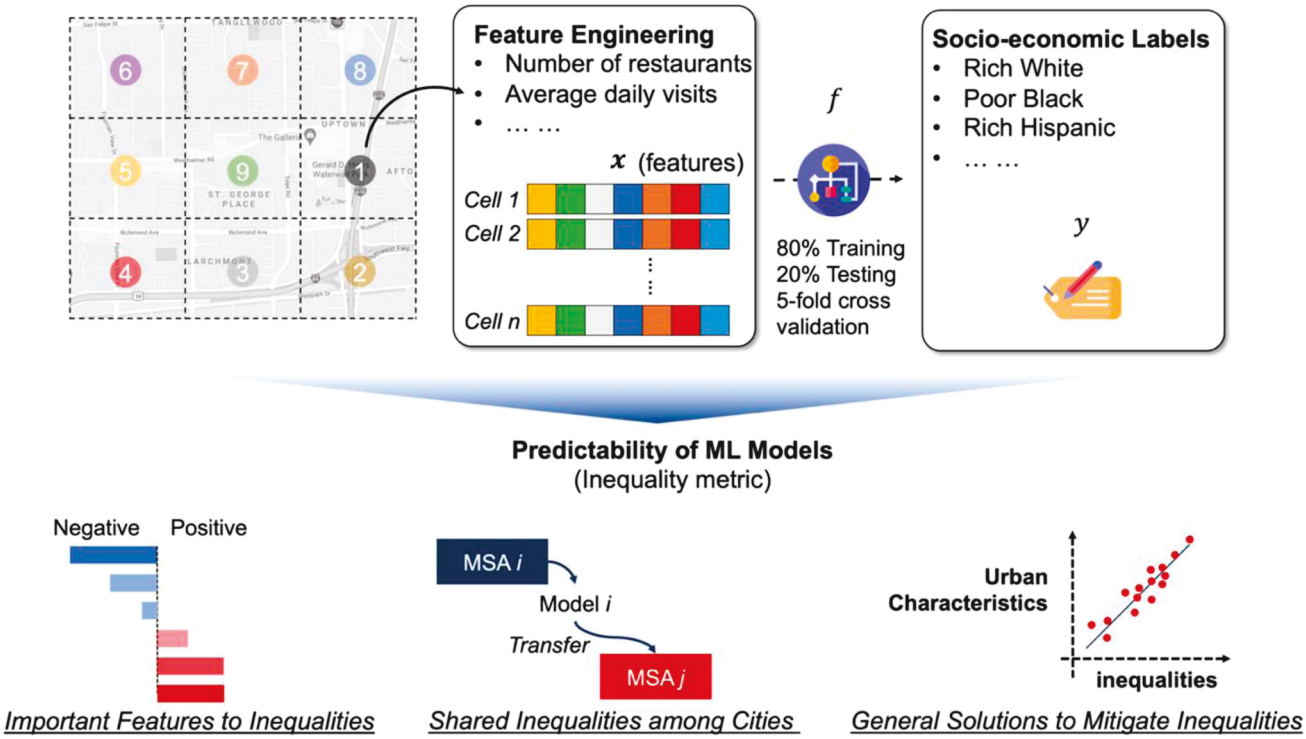

FIGURE 1 Illustration of the methodological framework. The upper panel shows a schematic of feature engineering, training, validation, and testing processes. We divide a metropolitan statistical area (MSA) into grid cells of equal size, extract the features related to facilities and human mobility, and convert the features into a vector for each grid cell. Each grid cell is labeled by one of the six labels related to income level and race. The lower panel of the figure shows three analyses using the F1 scores of the machine learning models as a metric of inequality. We interpret the importance of the features on the inequality of an MSA, evaluate the similarity of inequalities among MSAs, and identify general solutions for alleviating inequalities

subgroups are mutually exclusive: "Hispanic" including people of all races except White and Black, "Black" referring only to non-Hispanic Black people, and "White" including only non-Hispanic White people. The race that accounts for greater than 50% of people in a census tract reported in the Census data is considered the race label of this census tract. We similarly classified the census tracts as low-income or high-income based on whether the per capita income of the census tract is higher than the median of the MSA or not. We assign the label of a grid cell to the label of a census tract if the centroid of the grid cell falls into the polygon of the census tract. As such, the grid cells belonging to specific census tracts in an MSA are labeled by one of six socio-economic labels.

2.1.2 | Mobility data for activity features

Urban systems are spatially diverse in terms of population activities and facility distributions. Here, we characterize each grid cell based on these two dimensions. To understand the inequality of population activities, we employed mobile phone data from Cuebiq, a data intelligence company that collects location data from mobile phone users who opt in to share their data anonymously through a General Data Protection Regulation- and California Consumer Privacy Act-compliant frameworks. The current daily active user count collected by Cuebiq is roughly 15 million in the United States. The data sample has a wide set of attributes, including anonymized device identifier (ID), latitude, longitude, visited place ID (if the user visited a specific POI), UTC (coordinated universal time) time of observation, and the duration of each visit/stop (e.g., dwelling time). The data were shared under a strict contract with Cuebiq through their academic collaborative program in which they provide access to de-identified and privacy-enhanced mobility data for academic research. Cuebiq's responsible data-sharing framework enables us to query anonymized, aggregated, and privacy-enhanced data, by providing access to an auditable, on-premises sandbox environment (Moro et al., 2021). All researchers processed, aggregated, and analyzed the data under a non-disclosure agreement and were obligated not to share data further and not to attempt to re-identify data.

It is important to capture population activities in regular conditions when no external extreme events perturb

human activities. Considering that we extracted the Cuebiq mobility data from April 1, 2019, to April 7, 2019 (7 days) for selected MSAs, there are no particular events for MSAs in this time period, to the best of our knowledge. Also, we took the data for 7 days in order to account for the variation of population activities on weekdays and weekends. Using these data, we first assigned each visit or stop point to a defined grid cell. Then, we calculated a vast number of features related to population activities, such as the mean daily number of visits to a grid cell, the average duration of each visit in a grid cell, and the maximum daily number of stops in a grid cell. In addition, Cuebiq provides an estimation of the residential areas of mobile devices, which allows us to estimate the number of residents in each grid cell. The complete list of population activity features is provided in the Supplementary Information. The representativeness of the Cuebiq mobility data has been demonstrated by multiple prior studies (Aleta et al., 2020; F. Wang et al., 2019). They found that Cuebiq data are valid to describe human activities as one of the urban components (Deng et al., 2021). Hence, the features generated using these datasets should be representative and valid for our analyses.

2.1.3 | POI data for facility-relevant features and metrics

To capture the distribution of facilities in urban areas, we adopted the 6.5 million active POI data in the United States from Cuebiq. The dataset includes basic information about the POIs, such as POI IDs, location names, geographical coordinates, address, brand, and North American Industry Classification System (NAICS) code to categorize the POIs. The NAICS code is the standard used by Federal statistical agencies in classifying business establishments, such as retail trade, health care facilities, education, and entertainment places (United States Census Bureau, 2017). In this study, we selected 10 important types of POIs that are closely relevant to human lives: restaurants, schools, grocery stores, churches, gas stations, pharmacies and drug stores, banks, hospitals, parks, and shopping malls. We counted the number of POIs in each grid cell as their facility features.

By knowing the grid cell location of each POI, we further adopted a metric, urban centrality index (UCI), to characterize the distribution of the facilities in an MSA. UCI is the product of the local coefficient and the proximity index (Pereira et al., 2013). The local coefficient is computed based on the number of POIs within each grid cell, and the proximity index is computed based on the number of POIs within each grid cell along with a distance matrix that considers the distance between grid cells. The indices are formulated as follows:

$$\begin{aligned} LC &= \frac{1}{2}\sum_{i=1}^{N}\left(k_i - \frac{1}{N}\right) \\ PI &= 1 - \frac{V}{V_{\max}} \\ V &= \mathcal{K}' \times D \times \mathcal{K} \end{aligned} \tag{1}$$

where N is the total number of grid cells in an MSA; $\mathcal{K}$ is a vector of the number of POIs in each grid cell, and k_i is a component of the vector $\mathcal{K}$; D is the distance matrix between grid cells; $V_{\max}$ is calculated by assuming that the total POIs are uniformly settling on the boundary of the MSA. LC is the local coefficient, which measures the unequal distribution; PI is the proximity index, which solves the normalization issue; and V is the Venables index (Pereira et al., 2013). The value of UCI ranges from 0 to 1. The values close to 0 indicate polycentric distributions, while the values close to 1 indicate monocentric distributions.

2.1.4 | Other datasets and metric calculations

To calculate other metrics, we employed datasets from multiple commonly adopted platforms. In particular, we extracted data from Open Street Map (Open Street Map, 2021) to calculate the density of road segments in urban grid cells. We estimated complete road networks from the raw data by assembling road segments. Since the lengths of road segments created by the source are close to each other, we approached the road density by dividing the number of road segments by the areas of an MSA. To estimate the status of the economic development of the MSA, we adopted the 2018 data of gross domestic product (GDP) for each MSA (Bureau of Economic Analysis, 2018). The data are provided by the Bureau of Economic Analysis in the US Department of Commerce.

The socio-demographic data obtained from US Census 2014–2018 (5 years) ACS is also used to calculate the ethnicity entropy for an MSA. We first generated the distribution of population sizes for all race–ethnicity subgroups. Then, the Shannon entropy function is applied to calculate the ethnicity entropy $H(R)$:

$$H(R) = -\sum_{i=1}^{n} P\left(r_j\right) \log P\left(r_j\right) \tag{2}$$

where r_j is the race–ethnicity category, which occurs with probability $P(r_j)$ calculated by the proportion of people in the population of an MSA.

2.2 | Inequality characterization

The analyses employing the features and labels for urban grid cells consist of two components: (1) measuring inequality of each MSA using a quantitative metric, and (2) examining inequality within and across MSAs to explore potential inequality-alleviating solutions. This section provides an overview of the methods adopted to conduct experiments in these two components of analyses.

2.2.1 | Machine learning models

Machine learning models take as inputs the features of urban grid cells and learn the non-linear relationships among the features and the labels (Ramchandani et al., 2020). If the machine learning model in controlled experiments can reveal the socio-economic disparities of grid cells based on the input features and their non-linear relationships, it is an indicator of inequality in a city. In other words, in the presence of equality, the model should not be able to predict the socio-economic status of grid cells based on the input features. Accordingly, the predictability of socio-economic status based on the input features in the machine learning models is an indication of the existence of inequality, and thus the prediction performance measure could be a metric for measuring the inequality of the cities with regard to the complex interactions of the features. Hence, we consider the F1 score, which is a metric for the predictability of machine learning models, as the metric of inequality of the cities (see Figure 1).

In this study, the F1 scores in each socio-economic class are calculated individually first in a one-vs-rest manner. In each class, the positive label is the class label, and the negative label includes the rest socio-economic labels. Then, true positives are the ones where the model correctly predicts their real positive socio-economic label, and true negatives are the ones where the model correctly predicts a real negative label. False positives are the ones where the model incorrectly predicts the positive label, and false negatives are the ones where the model incorrectly predicts the negative label. Both false positives and false negatives indicate that the machine learning model cannot distinguish the socio-economic label well. True positives indicate a good performance of the model. Hence, both precision (considering true positives and false positives) and recall (considering true positives and false negatives) are equally important to the model. F1 score, the harmonic mean of precision and recall, conveys the balance between the precision and the recall of the machine learning models. In addition, data samples for different socio-economic labels are highly imbalanced. F1 score has been designed to work well on imbalanced data, compared to the accuracy of a machine learning model. The greater the F1 score in a model of a city, the greater the inequality.

To obtain valid and reliable results, this study adopts two widely used machine learning models: XGBoost and neural networks. The XGBoost model, a scalable tree boosting system, is an efficient and easy-to-use algorithm that delivers high performance and accuracy (Chen & Guestrin, 2016). We tend to have hundreds of thousands of samples (i.e., urban grid cells) in each MSA, leading to time-intensive model training processes. The XGBoost model could quickly execute and perform well in prediction tasks. Hence, this study mainly uses the results of XGBoost to characterize and understand the inequality in MSAs. Neural networks, composed of an input layer, a hidden layer, and an output layer, can efficiently identify important information from inputs leaving out redundant information. Through an embodied activation function, the neural networks are capable of capturing the non-linear relationship between the input features and output labels. Recognizing the benefits of the neural network model, we employed this model for validating the results generated from XGBoost, further enhancing the reliability of the findings and implications obtained from this study. The ridge regression model is a conventional mathematical model that is good at avoiding overfitting by regularizing the coefficient estimates (Hoerl & Kennard, 1970). The results of the ridge model help to demonstrate the performance and capabilities of the machine learning models in capturing complex urban feature interactions.

We implemented these machine learning models using an open-source Python package, scikit-learn (Pedregosa et al., 2011). We first randomly split the data into two sets, train and test; 80% of the samples are in the training set, and 20% of the samples are in the testing set. We further adopt the cross-fold validation to train the machine learning model and tune its hyperparameters. We divide the training set into five subsets of equal size. Four out of five subsets are used for training, and the remaining one is used for validation. With this process, the model would be further applied to the testing set and compute the F1 score for each city. In addition, the results of the machine learning model, especially the F1 scores for different MSAs, are validated through the training and testing of different machine learning models, neural networks and XGBoost.

The performance of a machine learning model may be influenced by many factors including the structure of the model, size of data, and so forth. The proposed method considers these uncertainties and controls them in generating the metric. We used the same model for learning the patterns of cities, the same size of grid cells in dividing urban spaces, and the same data sources for generating

features. Each MSA has more than 1000 grid cells so that the model can have sufficient data for training and validation. Hence, we could expect that the method proposed in this study is capable of capturing the actual inequality phenomenon in cities.

2.2.2 | Understanding of inequalities

As explained earlier, in this study, the F1 score of machine learning models quantifies the degree of inequality of each MSA. The next step is to identify potential solutions to alleviate inequalities in urban areas, which requires a thorough understanding of the underlying mechanisms of inequality within and across MSAs. Here, we propose three experiments to understand inequality from three different aspects.

In an MSA, inequality is shaped by both static features of facilities and dynamic features related to human activities. Examining the contributions of each feature to the inequality of the MSA is necessary for identifying alleviation solutions. To this end, we conducted experiments to measure the importance of features to the F1 score of the machine learning models. In these experiments, based on the trained models with all parameters and hyperparameters fixed, we set the values of one input feature to be zero for all samples and measure the predictability of the model (Lundberg & Lee, 2017). The decrease in F1 scores, to some degree, can indicate the importance of the features to the inequality of the MSA. Transforming the distribution of the important feature in areas of MSA would contribute to reducing the inequalities.

In addition to MSA-specific strategies, policies that are effective in more than one MSA would be beneficial for reducing policy-making efforts and enhancing the execution of policies at scale. Capturing the similarities of MSAs based on their inequality characteristics allows us to understand the effectiveness of cross-MSA policies. To this end, we employed the method of transferring machine learning models to different MSA and quantifying the similarities of inequalities across MSAs by the metric of model transferability. Specifically, we train the machine learning model by feeding in the samples from an MSA. Once the training process is done and all parameters are fixed, we feed in the sample from other MSA and measure the predictability of the model. The obtained F1 score could indicate the extent to which the patterns of the MSA on which the model is trained share similarities with the patterns of the MSA that the model is predicting. This quantitative metric offers us a generic metric to capture similarities of features shaping inequality, which could inform us about policy generalization and execution.

Finally, inequalities are not uniform among MSAs. The variations of urban characteristics across MSAs may tell us general approaches to mitigate urban inequalities. As such, we extend our analysis to capture the relationships between urban characteristics and F1 scores across MSAs. Here, we primarily look into: (1) the status of economic development quantified by GDP; (2) the scale of urban development quantified by the number of POIs in the MSA; (3) the connectedness of urban areas quantified by road density; (4) the diversity of residents quantified by ethnicity entropy; and (5) the geometric distribution of facilities quantified by the UCI. The calculation of these metrics is as aforementioned in previous sections. With all these characteristics of MSAs, to capture the relationships between inequality and urban characteristics, we employ an ordinary least squares (OLS) regression model to incorporate the interactions among multiple independent variables:

$$y_i \sim \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_3 x_{i,3} + \beta_4 x_{i,4} + \beta_5 x_{i,5} + \varepsilon_i \quad (3)$$

where y_i is the F1 score of MSA i; $x_{i,1}$ to $x_{i,5}$ are the variables of urban characteristics; β are coefficients; ϵ_i is the error term. In the regression, since the values of GDP, road density, and number of POIs have a much larger scale than other variables, we use logarithmic transformation of values.

3 | RESULTS

3.1 | Empirical statistics of features

The variety of datasets we gathered allowed us to capture different features of the cities. We first show examples of features mapped into the metropolitan area of Atlanta to gain a basic and empirical understanding of the distribution of facilities and human activities in an MSA. Figure 2 illustrates the extent to which densities of features vary across the areas of the Atlanta MSA. As we observed, the number of active residents varies across different regions of the MSA (Figure 2a). POIs are concentrated in the center of the MSA and expand like a tree from the center to the periphery of the MSA (Figure 2c). The main incentive for human movements is the visits to POIs, such as working and shopping, leading to agglomerated activities in the center of the MSAs with a high density of POIs (Figure 2b). Beyond activities in POIs, the footprints of people also include visits to friends and work commutes. Hence, the scale of population activities is broader than the locations of POIs. Finally, in Figure 2d, we show the residential areas labeled by socio-demographic groups. We find that White people account for the majority of the residential

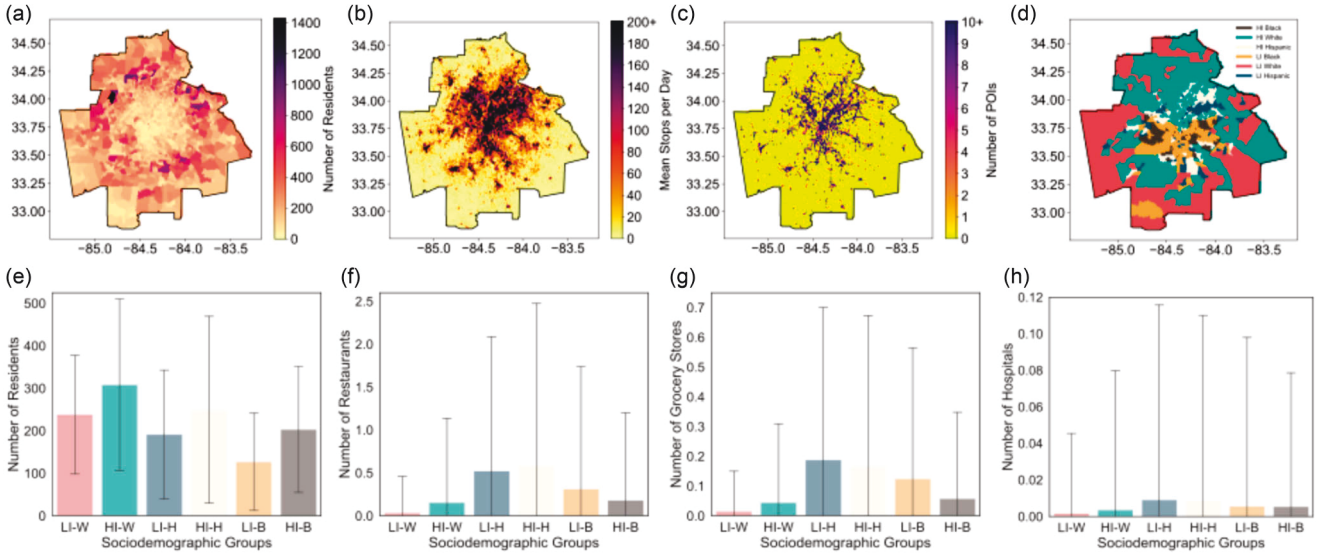

FIGURE 2 Spatial distribution of features and socio-economic characteristics of population groups in Atlanta MSA. (a) The distribution for the number of residents (mobile phone devices as a proxy) based on mobile phone data. The numbers of residents are aggregated at the census tract level. Because the areas of census tracts vary, the figure shows the only number of people in different regions of the MSA rather than the population density. (b) The distribution of the average number of stops per day in a grid cell. (c) The distribution of the number of points of interest in a grid cell. (d) The distribution of different income and race groups: HI represents high-income groups; LI represents low-income groups. (e–f) The distribution of some example features (i.e., number of residents (e), number of restaurants (f), number of grocery stores (g), and number of hospitals (h)) in each sociodemographic group: W represents White; H represents Hispanic, and B represents Black. The error bar represents the variance of samples.

area of the MSA. High-income White people are living in the North and close to the center of the MSA, while low-income White people tend to live on the periphery and the South of the MSA. Compared to the wide distribution of White people, the residential areas of Black people are more condensed, and high-income and low-income Black subgroups are intertwined in the center of the MSA. Hispanic subgroups occupy only a very small proportion of the area and are dispersed across the MSA. These observations inform us about the segregation and inequality of feature distributions and the complex association between urban features and socio-demographic groups.

In the next step, we first look at the features of facilities shared by different population groups. Figure 2e–h shows the differences in the number of example facilities in the grid cells occupied by different socio-demographic groups. We observe that the differences in facilities in residential areas of different socio-demographic people are not significant. Specifically, comparing the mean and variance in the average number of facilities in a grid cell, the differences may be present in the mean values. For example, grid cells of Hispanic people have more restaurants and grocery stores (Figure 2f,g). White people, high-income or low-income, have the minimum number of facilities in their residential grid cells. The variance across grid cells in a population group, however, is extremely large, making the differences in the number of facilities inconspicuous. This pattern is observed in all selected MSAs (more details can be found in the Supplementary Information). Such observation implies that inequality is not apparent and cannot be simply quantified through basic statistics and based on only one urban feature due to the complex interactions of urban features. Hidden and nonlinear mechanisms resulting in inequalities at the nexus of urban features and socio-demographic attributes exist and are underexplored without advanced methods capable of specifying the complex interactions of features.

3.2 | Measurement of inequality

To further decompose the inequality in cities, we trained three extensively adopted and technically mature models: two machine learning models, XGBoost and neural network models, and one conventional model, ridge regression model. The predictability of these models, given features in urban grid cells, is considered a metric of inequality in an MSA. The machine learning models are well-trained in the same way for different cities. All the metric values for evaluating the model performance are obtained when the models are optimized and convergent. We only compare the inequality metric of cities when all other influential factors, such as model types, grid size, and features, are controlled. Showing the influence of these

factors on model performance is to help select the proper model, grid size, and features for this study. Under this context, the poor performance of the model can indicate less inequality since all other influential factors are controlled well. The results of F1 scores are based on the testing data for each city. The inequality is pronounced if the machine learning model can obtain high predictability, indicated by a great F1 score. That is, the interactions among urban features can distinguish the residential areas with different socio-demographic population groups, reflecting the fact that inequalities of urban features in serving residents of subgroups present. Using the F1 score as the metric of inequality, we quantify the inequality of all selected MSAs by considering the nuanced relationships of urban features. However, as aforementioned in the Methods section, the ability to capture the complex relationships among urban features and the algorithmic advantages varies among machine learning models. Here, by training and testing the models, we found that the ridge and neural network models have similar performance across all MSAs; and the XGBoost model outperforms conventional ridge models by about 25% in the majority of the selected MSAs (Figure 3a). The XGBoost models achieve an average of 0.8 for F1 scores among selected MSAs, meaning that the model can explain 80% of the variations of labels based on input variables. In view of the outstanding performance of the XGBoost models, we used the results of XGBoost to analyze the inequality of MSAs in this study, and the results of the other two models to validate the outcomes of XGBoost models.

The predictability of the machine learning models may be influenced by factors such as the size of grid cells or specific features that undermine the importance of the complex interactions of urban features. To examine the robustness of the models and the results, we applied the models to samples generated from different sizes of grid cells. Figure 3b displays the relationship between F1 scores and the size of grid cells for three examples of MSAs. We observe that the performance of the XGBoost model decreases when the size of the grid cell increases. There is a jump in performance at around the grid size of 0.02 and 0.03. Decreases in model performance could be attributed to the lack of grid cells (samples) to train the model and also the aggregation of features that reduces the disparities among grid cells. Such a negative correlation between model performance and grid size provides us with the rationale for selecting a proper grid size for measuring the inequality of MSAs. Based on the results, 0.01 and 0.02 would be proper grid sizes. Thus, for all the analyses in this study, we used 0.01 as the size of the grid cell so that the results generated from the machine learning models could be comparable and informative.

In addition, individual features may also influence the performance of the model due to the strong correlation between individual features and labels. Here, we examined the contributions of individual features while fixing the parameters of the well-trained model. The trained model preserves both the complex interactions of the features and the contributions of individual features. Figure 3e shows the decrease in model performance by removing specific features. The elimination of features related to general human activities, such as mean stops, mean visits, average visit time, and the number of residents, could lead to decreases in F1 scores. But the decreases do not significantly influence the performance of the model, compared to the high predictability of the model. For example, for the results of the XGBoost model, the average influence of the number of residents on F1 scores is below 0.3. In Figure S1a,b, we also plot the influences of the features on the F1 scores for the ridge and neural network models. The average influences of the features are even much lower than 0.2. Compared to the average F1 score of XGBoost, which is 0.81, we consider that urban features do not have a significant influence on the model performance. In addition, the specific types of POIs and visits to these types of POIs do not make too much difference to the F1 scores. In general, individual features cannot explain the inequality of each MSA well. This result implies that inequality is a phenomenon arising from non-linear interaction among various urban features. Hence, inequality should be attributed to hidden complex interactions of the urban features rather than individual attributes.

3.3 | Transferability of inequality

We mapped the F1 scores of the MSAs obtained from the XGBoost model in Figure 3c. There are 22 MSAs from the South, 12 MSAs from the West, nine MSAs from the Midwest, and six MSAs from the Northeast of the United States. We observed significant regional patterns from the map: MSAs in the US West tend to have higher F1 scores than MSAs from other regions, and Northeast MSAs tend to have lower F1 scores. That means, socio-economic inequality is greater in the MSAs in the US West, and socio-economic inequality is lesser in the MSAs in the Northeast, compared to the MSAs in other regions. To further explore this observation, we plotted the relationships among F1 scores, regions, and the GDP in Figure 3d. In addition to the regional patterns, we also find that lower GDP is correlated with higher F1 scores, while higher GDP is correlated with lower F1 scores. This association is not very strong since we selected MSAs with the largest populations. The weak negative correlation can still demonstrate the

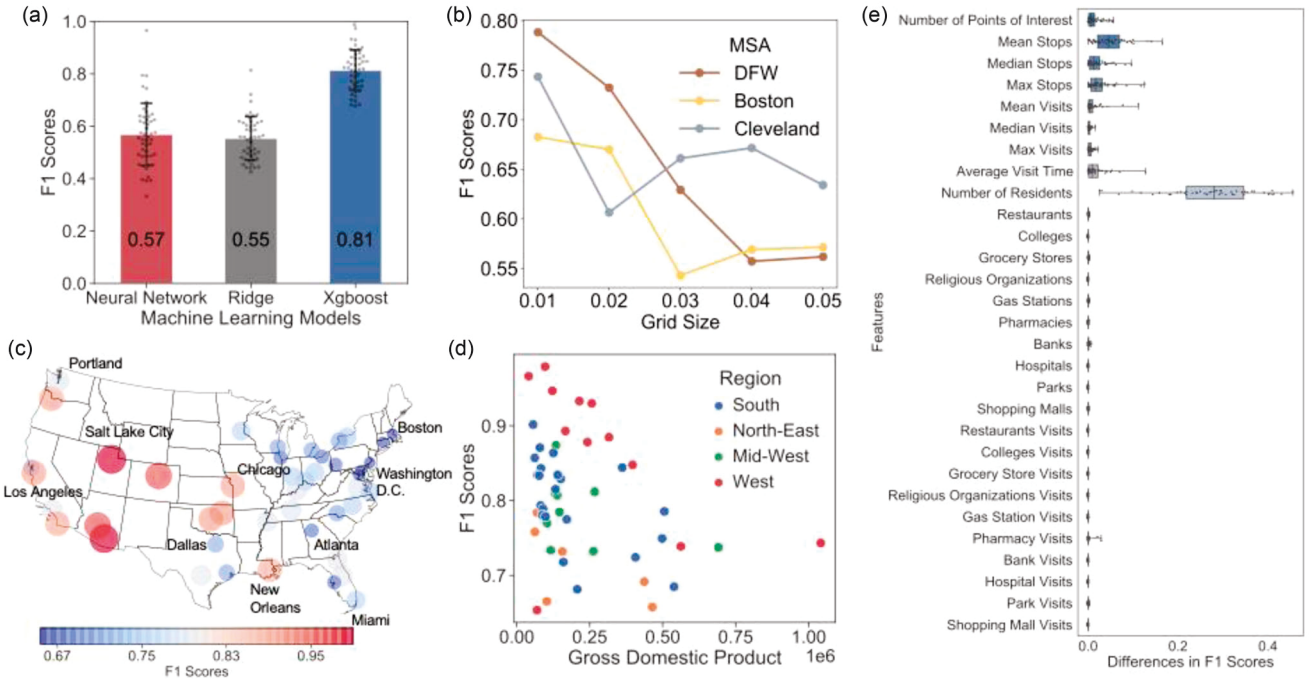

FIGURE 3 Results of model training and testing. (a) F1 scores of three models: neural networks, ridge classifier, and XGBoost for each MSA. The numbers on the bars are the mean values of the F1 scores for all selected MSAs. The dots on top of each bar represent the F1 scores of the MSAs. The error bars show the variance of the F1 scores. XGBoost achieves the best performance among the three models. (b) Results of testing the effect of grid size on the performance of the XGBoost model in three example MSAs: Dallas-Fort Worth MSA, Boston MSA, and Cleveland MSA. The sizes of the grid cells are measured by the differences in the longitude and latitude of the corner points on one side of a grid cell. Hence, the values on the x-axis represent the differences in degree in the geographical coordinate systems. (c) A geographical map shows the F1 scores for selected MSAs in the United States. (d) The relationships between F1 scores and gross domestic product of MSAs in four regions of the United States: South, Northeast, Midwest, and West. (e) Importance of features to the F1 score of the XGBoost models for each MSA. The x-axis is the difference between the original F1 scores and the F1 scores after dropping a specific feature from the input (decrease of predictability of the XGBoost model). The y-axis represents the features selected to be removed for understanding its contribution to the inequality of the MSAs.

association between the extent of socio-economic inequality and the GDP of the MSAs. These regional patterns of inequality motivate us to consider the common characteristics shared by MSAs.

To explore the similarities of inequality across a variety of MSAs, we conducted experiments on the transferability of the patterns. That is, to what extent the machine learning model trained with the samples of one MSA can predict the occupied population groups for grid cells in other MSAs. The transferability of the models helps us to understand the generalizability of the patterns across MSAs and regions. As most of the analyses and results are taken from the most populated MSAs, other MSAs can benefit from the identified and generalized patterns (Dong et al., 2019), if the shared inequality patterns can be captured. We trained the machine learning models using data samples from one MSA with both validation and testing processes. Then, we applied the fixed model to the data samples from another MSA. This process aims to address if the patterns from one MSA are transferrable to another MSA, which allows us to observe the variations of inequality in cities across the nation and motivates us to explore the factors related to variant inequalities. Hence, the results present in the paper are based on the performance of the models on the testing sets, either from the same MSA or a different MSA. Figure 4 summarizes the results obtained from cross-MSA experiments. As expected, all the models trained and tested on the same MSAs (diagonal) outperform models trained and tested in different MSAs. The performance of the models varies for different pairs of MSAs. The values on the upper left corner are closer to light blue, meaning that the F1 scores are close to 0.6 and the transferability is more evident, while most of the values on the right-hand side are dark red, meaning that the transferability is quite low (Figure 4a). These results imply that some MSAs share common characteristics shaping their inequalities, and thus the same inequality-alleviating measures could work across these MSAs. We also found that the transferability matrix is asymmetric. We show example MSAs that achieved the highest transferability and the

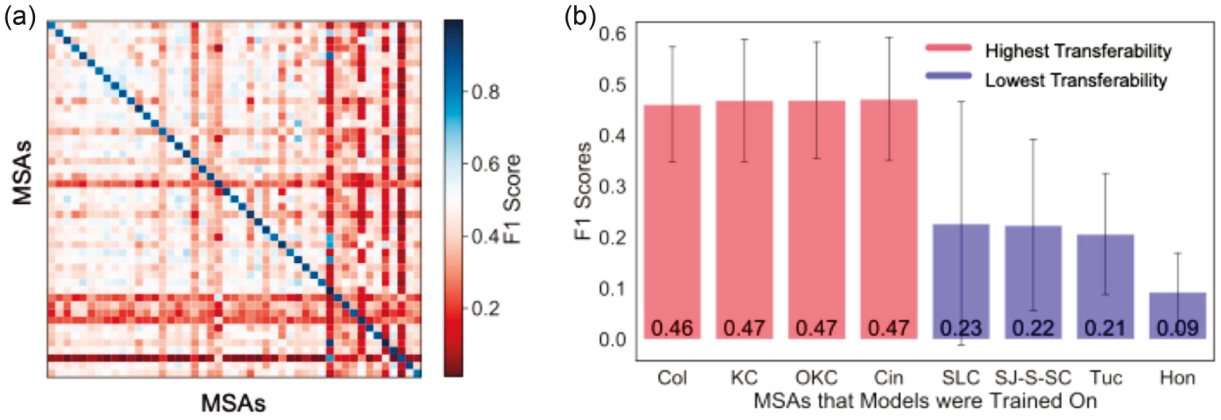

FIGURE 4 Shared inequality among selected MSAs measured by the transferability of machine learning models. (a) Pair-wise similarity of inequalities among MSAs. Each row represents the MSA where the model is trained, and each column represents the MSA where the trained model is adopted to make predictions. The color indicates the F1 scores. Here, the machine learning model is XGBoost. (b) Examples of transferability results for the top four and bottom four models for MSAs: Columbus (Col), Kansas City (KC), Oklahoma City (OKC), Cincinnati (Cin), Salt Lake City (SLC), SJ-Sunnyvale-SC (SJ-S-SC), Tucson (TC), and Urban Honolulu (Hon). The error bars show the variance of the F1 scores. The numbers attached at the bottom of the bars are the mean values of the F1 scores.

lowest transferability among the selected MSAs. Models trained on MSAs such as Columbus and Kansas City can learn the most common patterns of inequality, which could be applied to most of the other MSAs. However, models trained on MSAs such as Urban Honolulu are not able to capture the common inequality patterns of other MSAs since Urban Honolulu is in Hawaii, where the development and environment are different from cities in the US mainland.

3.4 | Relationship with urban characteristics

Considering the variety and transferability of models among MSAs, the next question is what inequality-alleviating strategies would be effective among MSAs consistent with their urban characteristics. To investigate this question, we computed the metrics of urban characteristics for MSAs, including the urban centrality index, road density, and ethnicity entropy, along with the number of POIs and GDP of the MSAs (more details can be found in the Methods section.) Results are summarized in Figure 5 and Table 1. The distributions of UCIs and the inequality extent measured by F1 scores are approximately normal, with histograms shown in Figure 5a. The Kendall rank correlation reaches 0.72, the Spearman rank correlation reaches 0.88, and the Pearson correlation coefficient approaches 0.89 for these 47 MSAs. All measures are statistically significant with $p < 0.01$, indicating a strong positive correlation between the UCI and the extent of inequality. UCI itself is not included in machine learning models. The strong correlation between UCI and the F1

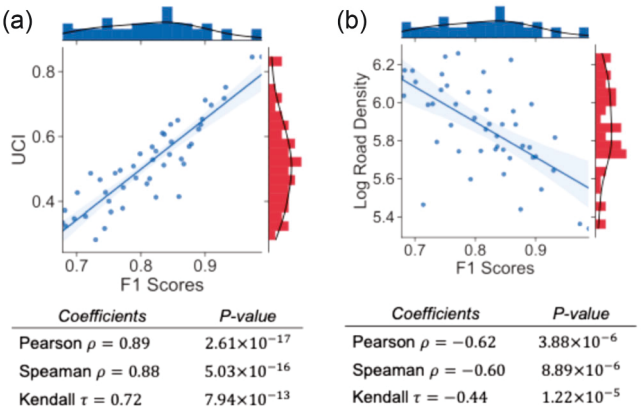

FIGURE 5 The relationship between urban characteristics and inequality (F1 scores). (a) The values of urban centrality index (UCI) as a function of F1 scores obtained from XGBoost models. (b) The logarithmic values of road density in grid cells are a negative function of F1 scores obtained from XGBoost models. The correlation analysis under the plots shows the exact statistics and *p*-values. Three statistical tests were conducted for each of the correlation analyses. All measures are statistically significant with $p < 0.01$. The UCI is strongly positively correlated with inequality, and road density is moderately negatively correlated with inequality for the selected 47 MSAs.

score serves as an important interpretation of the presence of inequality in cities. That is, a pronounced concentration of POIs greatly contributes to inequality in MSAs. Analyses on road density reveal another significant relationship. The distribution of road density is close to log-normal, while the distribution of F1 scores is normal (histograms in Figure 5b). The Kendall rank correlation reaches −0.44, the Spearman rank correlation reaches −0.60, and the

TABLE 1 Ordinary least squares regression between metropolitan statistical area characteristics and F1 scores of XGBoost models

	Dependent variable: F1 Scores of XGBoost models						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log GDP:				−0.004 (0.017)			0.033 (0.036)
Log POIs:			0.001 (0.020)			0.057 (0.044)	
Ethnicity Entropy:	0.031 (0.028)				0.020 (0.050)		
Log Road Density:		0.001 (0.031)			−0.209*** (0.040)	−0.249*** (0.051)	−0.237*** (0.052)
UCI:	0.521*** (0.039)	0.516*** (0.054)	0.515*** (0.041)	0.512*** (0.041)			
Constant:	0.509*** (0.040)	0.538** (0.204)	0.543*** (0.096)	0.568*** (0.097)	2.018*** (0.233)	2.034*** (0.229)	2.035*** (0.231)
Observations:	47	47	47	47	47	47	47
Adj. R^2:	0.796	0.79	0.79	0.791	0.355	0.377	0.364
F-stat:	90.77***	87.73***	87.73	87.87***	13.64***	14.9***	14.18***

Note: Standard errors are provided under coefficients in parentheses. All log values are log-based 10. GDP, gross domestic product; POIs, points of interest; UCI, coordinated universal time.
Significance Level
*$p < 0.1$
**$p < 0.05$
***$p < 0.01$.

Pearson correlation coefficient approaches −0.62, for 47 MSAs. These significant measures signify a moderate negative correlation between road density and inequality. That is, the increase in road density (as an indicator of urban development and connectivity) could contribute to alleviating socio-economic inequalities.

Coupled with other factors, we analyzed the extent to which urban characteristics can capture the inequality of MSAs. We examined the performance of multilinear models with different combinations of variables. Table 1 summarizes the results of the multilinear regression models using OLS. The first four models with the inclusion of UCI as a variable reach high-fitting performance with R^2 greater than 0.79, indicating that UCI can explain 79% of the inequality in MSAs. The coefficients for UCI are significant, showing a consistent result with the correlation analysis. Other variables, such as the number of POIs, GDP, and ethnicity entropy, are not significant, even though they may have positive and negative correlations with the inequality scores. The relationship between road density and inequality is not significant. This result implies that, although the correlation analysis finds the alleviating effect of road density on the inequality of MSAs, a polycentric distribution of POIs could moderate the effect of road density on inequality. The other three models exclude the UCI variable and examine the effects of road density coupling with other factors. In these models, the negative relationships between road density and inequality are significant, confirming our previous findings in the correlation analysis and making the road density weakly predictive of inequality. The R^2 of these models reaches more than 0.35, showing the moderate effect of expanding road density on the inequality of MSAs. Other factors, including GDP, and ethnicity entropy are still insignificant. To establish that the correlations between inequality and urban characteristics are sufficiently general, we tested these findings using the F1 scores obtained from neural networks and ridge models. The results are summarized in the Supplementary Information, Tables S1 and S2. These findings inform us about the potential of enhancing road density and POI distribution for inequality alleviation, which will be discussed in detail in the discussion section.

3.5 | Model comparison

The machine learning models that this study focuses on are the neural network model and the XGBoost model. The ridge regression model is a conventional mathematical model because it is a type of linear regression technique used to solve some of the problems of OLS by imposing a penalty on regression. The form of the ridge model is clearly defined. Solving the ridge model is equivalent

to solving the coefficient for each independent variable. The results in Figure 3a show the lowest predictive performance of the ridge model, compared to the machine learning models like neural network and XGBoost models. In addition, comparing the results in Figures 3c,d and S2, we find that the ridge regression model cannot distinguish the degrees of inequality across US cities. Finally, based on the poor performance in Table S1, compared to the results in Tables 1 and S2, we find that the ridge regression model is limited in interpreting the factors influencing inequality in cities. Therefore, we prove that conventional regression models are not capable of capturing the complex interactions among the inputs. The proposed machine learning models outperform conventional mathematical models to measure and explain inequality in cities.

4 | DISCUSSION AND CONCLUDING REMARKS

Measuring and understanding the socio-economic inequality in cities is of great importance to policymaking, planning, and design toward equitable urban systems of facility services and life opportunities. When equality exists, people of different income levels and racial groups would have similar interactions with facilities and infrastructure to meet their life needs. In this study, we present a new computational method that leverages the interpretability of machine learning models to encode the high-dimensional and complex interactions of urban features to quantify and understand socio-economic inequality in 47 US metropolitan areas. Inequality is a multifaceted phenomenon that arises from the complex interactions among heterogeneous urban features. Different from existing works, the method proposed in this study allow us to integrate heterogeneous urban features and their complex interactions into a comprehensive and quantitative metric. The metric is capable of providing a holistic view of the inequality of intertwined urban components in a city and also allowing to transfer insights across cities.

We show that being able to predict the income and race label of an area based on population and the built environment features is an indicator of inequality. Accordingly, we demonstrate the effectiveness of using the predictability of machine learning models as a metric of inequality to integrate the non-linear relationships among urban components. We also examine the tradeoff between grid size and model accuracy and find regional patterns of inequality of MSAs. The results show that the predictability of machine learning models does not decline drastically if individual features are removed. This result provides evidence that inequality is a phenomenon influenced by the intertwined urban features rather than a consequence of individual features.

We conducted validation on different parts of the method to enhance the validity of the findings. First, the validation of the machine learning models has been conducted using five-fold cross-validation in training the models. Second, the results of the machine learning model, especially the F1 scores for different MSAs, are validated through the training and testing of different machine learning models such as neural networks and XGBoost and the comparison with the results of conventional mathematical models like the ridge regression model. Third, the strong correlations between F1 scores and facility distributions, and road density, which align with existing social science literature, could also support the validity of the method and findings in this study.

The objective of the proposed machine learning method for urban inequality is not to improve the prediction accuracy or other quantitative metrics of model performance. The proposed machine learning model overcomes the limitations of the conventional mathematical models that require specifying the form of feature interactions and compound effects on the dependent variable. In fact, it is improper to specify the forms of feature interactions and compound effects before being aware of the underlying mechanisms of these interactions. As such, existing mathematical models based on assumed formulae are not comparable with our model because the complex interactions of urban features are unknown.

The finding helps us rethink how inequality should be examined in cities. The transferability analyses of the models show that MSAs indeed share common patterns of inequality, implying that urban characteristics may influence the inequality of cities. Variations of inequality patterns, however, still exist because the models are not completely transferable. By examining the relationships between urban characteristics and the inequality metric, we develop a deeper understanding of inequality and identify general solutions for inequality mitigation. The results and findings of this study have notable implications that contribute to decision-making in various research and practical domains such as urban planning, infrastructure development, economic promotion, and government regulation.

With the growing availability of urban big data and the amplified complexity of urban systems, learning how urban components interact with and understanding the consequent impacts of complex interactions are particularly critical for optimizing the operations of urban systems and the decision-making of urban development. Our results suggest individual features cannot reveal the complexity of the urban systems and how inequalities emerge, and thus are not capable of quantifying the inequality

of cities properly. The inequality metric proposed in this study further understanding of the non-linear interaction of population activities and facility distributions and the effects on social-economic inequality of cities. The proposed metric provides a new perspective on evaluating the complex relationships of urban components and a novel approach to deriving knowledge of urban systems from large-scale multisource granular data. In particular, overcoming city-scale challenges such as inequality issues, a holistic perspective to think about the underlying mechanisms and solutions is required, as the interdependencies of the urban components are making a difference in the socio-economic outcomes of the whole city.

Another implication of our work is helping city planners and governments evaluate strategies for alleviating socio-economic inequalities in MSAs with the inferred relationship between urban characteristics and the inequality metric. Our study shows that better urban development and dispersed distribution of facilities could alleviate inequality of cities significantly. Changing the facility distribution from mono-centricity to poly-centricity could narrow the service gap between different areas of the cities and could intertwine with the regular life activities of the population. Increasing road density (as an indicator of urban development) could improve the accessibility of public services. On the other hand, the effects of facilities distribution may moderate the effects of road density on inequality. This finding raises a more practical way for alleviating and mitigating inequalities as dramatically changing the distribution of facilities in a city would lead to a worse impact on the economy than the benefits of mitigating inequality. Hence, given limited resources, policies that could increase road density and slightly change facility distribution at the same time may end up being cost-effective solutions, as these actions could reshape the mobility flows and visit patterns of the population. In addition, localized actions for each MSA are still needed since variations of inequality patterns are also observed in our study.

This study also has some limitations that need future research to overcome. First, human activities are not static features. Activities in different scenarios, such as gathering events, commuting peaks, and natural disasters could show a more comprehensive profile of population patterns and further make a difference in measuring socio-economic inequality. Future research could build upon our framework and extend the machine learning models to incorporate dynamic population activities. For example, the long short-term memory model could be adopted to encode time-series information on human activities (Alam et al., 2020). The understanding of inequality could be deepened by capturing more features about urban systems and populations. Second, this study considers each area of a city as independent. The physical adjacencies and social dependencies are not computed and included in our models, although these features are of importance to understanding the spillover effect of inequality. Future research could develop new computational models (Martins et al., 2020), such as graph neural networks to encode such relational information quantifying the inequality of cities.

ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Grant CMMI-1846069 (CAREER), the Texas A&M University X-Grant 699, and the Microsoft Azure AI for Public Health Grant. The authors also would like to acknowledge the data support from Cuebiq Inc. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Texas A&M University, Microsoft Azure, or Cuebiq Inc.

REFERENCES

Acemoglu, D., & Robinson, J. (2009). Foundations of societal inequality. *Science*, *326*(5953), 678–679. https://doi.org/10.1126/science.1181939

Adeli, H., & Hung, S. L. (1994). An adaptive conjugate gradient learning algorithm for efficient training of neural networks. *Applied Mathematics and Computation*, *62*(1), 81–102. https://doi.org/10.1016/0096-3003(94)90134-1

Ahmadlou, M., & Adeli, H. (2010). Enhanced probabilistic neural network with local decision circles: A robust classifier. *Integrated Computer-Aided Engineering*, *17*(3), 197–210.

Alam, K. M. R., Siddique, N., & Adeli, H. (2020). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*(12), 8675–8690. https://doi.org/10.1007/s00521-019-04359-7

Aleta, A., Martín-Corral, D., Pastore y Piontti, A., Ajelli, M., Litvinova, M., Chinazzi, M., Dean, N. E., Halloran, M. E., Longini, Jr., I. M., Merler, S., Pentland, A., Vespignani, A., Moro, E., & Moreno, Y. (2020). Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*, *4*(9), 964–971. https://doi.org/10.1038/s41562-020-0931-9

Anand, S., & Ravallion, M. (1993). Human development in poor countries: On the role of private incomes and public services. *Journal of Economic Perspectives*, *7*(1), 133–150. https://doi.org/10.1257/jep.7.1.133

Balland, P. A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, *4*, 248–254. https://doi.org/10.1038/s41562-019-0803-3

Bureau of Economic Analysis. (2018). *Gross domestic product (GDP) summary by metropolitan area*. U.S. Department of Commerce.

Casali, Y., Aydin, N. Y., & Comes, T. (2021). Zooming into socio-economic inequalities: Using urban analytics to track vulnerabilities––A case study of Helsinki. In R. Grace, K. Moore, & C. W. Zobel (Eds.), *ISCRAM 2021 conference proceedings–18th*

International conference on information systems for crisis response and management. (pp. 1028–1041). Virginia Tech.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA pp. 785–794. https://doi.org/10.1145/2939672.2939785

Cingano, F. (2014). *Trends in income inequality and its impact on economic growth*. (Working Paper No. 163), OECD Social, Employment and Migration, OECD Publishing. https://doi.org/10.1787/5jxrjncwxv6j-en

Dahl, E., & Malmberg-Heimonen, I. (2010). Social inequality and health: The role of social capital. *Sociology of Health & Illness*, *32*(7), 1102–1119. https://doi.org/10.1111/j.1467-9566.2010.01270.x

Deng, H., Aldrich, D. P., Danziger, M. M., Gao, J., Phillips, N. E., Cornelius, S. P., & Wang, Q. R. (2021). High-resolution human mobility data reveal race and wealth disparities in disaster evacuation patterns. *Humanities and Social Sciences Communications*, *8*(1), 144. https://doi.org/10.1057/s41599-021-00824-8

Dong, X., Morales, A. J., Jahani, E., Moro, E., Lepri, B., Bozkaya, B., Sarraute, C., Bar-Yam, Y., & Pentland, A. (2019). Segregated interactions in urban and online space. *EPJ Data Science*, *9*, 20. https://doi.org/10.1140/epjds/s13688-020-00238-7

Esmalian, A., Wang, W., & Mostafavi, A. (2022). Multi-agent modeling of hazard–household–infrastructure nexus for equitable resilience assessment. *Computer-Aided Civil and Infrastructure Engineering*, *37*(12), 1491–1520.

Fan, C., Lee, S., Yang, Y., Oztekin, B., Li, Q., & Mostafavi, A. (2021a). Effects of population co-location reduction on cross-county transmission risk of COVID-19 in the United States. *Applied Network Science*, *6*(1), 14. https://doi.org/10.1007/s41109-021-00361-y

Fan, C., Yang, Y., & Mostafavi, A. (2021b). Neural embeddings of urban big data reveal emergent structures in cities. ArXiv Preprint ArXiv:2110.12371, 1–17.

Gazzotti, P., Emmerling, J., Marangoni, G., Castelletti, A., van der Kaj-Ivar, W., Hof, A., & Tavoni, M. (2021). Persistent inequality in economically optimal climate policies. *Nature Communications*, *12*(1), 3421. https://doi.org/10.1038/s41467-021-23613-y

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Johar, M., Soewondo, P., Pujisubekti, R., Satrio, H. K., & Adji, A. (2018). Inequality in access to health care, health insurance and the role of supply factors. *Social Science & Medicine*, *213*, 134–145. https://doi.org/10.1016/j.socscimed.2018.07.044

Li, G., Cai, Z., Liu, X., Liu, J., & Su, S. (2019). A comparison of machine learning approaches for identifying high-poverty counties: Robust features of DMSP/OLS night-time light imagery. *International Journal of Remote Sensing*, *40*(15), 5716–5736. https://doi.org/10.1080/01431161.2019.1580820

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, CA pp. 4768–4777.

Marger, M. N. (1999). *Social inequality: Patterns and processes*. McGraw-Hill.

Martins, G. B., Papa, J. P., & Adeli, H. (2020). Deep learning techniques for recommender systems based on collaborative filtering. *Expert Systems*, *37*(6), e12647. https://doi.org/10.1111/exsy.12647

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization*. Harvard University Press. https://doi.org/10.4159/9780674969797

Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., Honermann, B., Lankiewicz, E., Mena, L., Crowley, J. S., Sherwood, J., & Sullivan, P. S. (2020). Assessing differential impacts of COVID-19 on black communities. *Annals of Epidemiology*, *47*, 37–44. https://doi.org/10.1016/j.annepidem.2020.05.003

Mirza, M. U., Xu, C., Bavel, B. V., van Nes, E. H., & Scheffer, M. (2021). Global inequality remotely sensed. *Proceedings of the National Academy of Sciences*, *118*(18), e1919913118. https://doi.org/10.1073/pnas.1919913118

Moro, E., Calacci, D., Dong, X., & Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large US cities. *Nature Communications*, *12*(1), 4633. https://doi.org/10.1038/s41467-021-24899-8

Niu, T., Chen, Y., & Yuan, Y. (2020). Measuring urban poverty using multi-source data and a random forest algorithm: A case study in Guangzhou. *Sustainable Cities and Society*, *54*, 102014. https://doi.org/10.1016/j.scs.2020.102014

Open Street Map. (2021). *OpenStreetMap*. https://www.openstreetmap.org/

Pan, W., Ghoshal, G., Krumme, C., Cebrian, M., & Pentland, A. (2013). Urban characteristics attributable to density-driven tie formation. *Nature Communications*, *4*(1), 1961. https://doi.org/10.1038/ncomms2961

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V., & others. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Pereira, R. H. M., Nadalin, V., Monasterio, L., & Albuquerque, P. H. M. (2013). Urban centrality: A simple index. *Geographical Analysis*, *45*(1), 77–89. https://doi.org/10.1111/gean.12002

Rafiei, M. H., & Adeli, H. (2016). A novel machine learning model for estimation of sale prices of real estate units. *Journal of Construction Engineering and Management*, *142*(2), 04015066. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001047

Rafiei, M. H., & Adeli, H. (2017). A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 3074–3083. https://doi.org/10.1109/TNNLS.2017.2682102

Ramchandani, A., Fan, C., & Mostafavi, A. (2020). DeepCOVIDNet: An interpretable deep learning model for predictive surveillance of COVID-19 using heterogeneous features and their interactions. *IEEE Access*, *8*, 159915–159930. https://doi.org/10.1109/ACCESS.2020.3019989

Rao, N. D., van Ruijven, B. J., Riahi, K., & Bosetti, V. (2017). Improving poverty and inequality modelling in climate research. *Nature Climate Change*, *7*(12), 857–862. https://doi.org/10.1038/s41558-017-0004-x

Rasch, R. (2017). Income inequality and urban vulnerability to flood hazard in Brazil*. *Social Science Quarterly*, *98*(1), 299–325. https://doi.org/10.1111/ssqu.12274

Thomas, P., & Emmanuel, S. (2014). Inequality in the long run. *Science*, *344*(6186), 838–843. https://doi.org/10.1126/science.1251936

Triventi, M. (2013). The role of higher education stratification in the reproduction of social inequality in the labor market. *Research*

in Social Stratification and Mobility, *32*, 45–63. https://doi.org/10.1016/j.rssm.2013.01.003

United States Census Bureau. (2017). *North american industry classification system (NAICS)*. United States Census Bureau. https://www.census.gov/naics/

United States Census Bureau. (2019). *American Community Survey 2014–2018 5-year estimates*. United States Census Bureau. https://www.census.gov/newsroom/press-releases/2019/acs-5-year.html

Wang, F., Wang, J., Cao, J., Chen, C., & Ban, X. (2019). Extracting trips from multi-sourced data for mobility pattern analysis: An app-based data example. *Transportation Research Part C: Emerging Technologies*, *105*, 183–202. https://doi.org/10.1016/j.trc.2019.05.028

Wang, J., Kuffer, M., Roy, D., & Pfeffer, K. (2019). Deprivation pockets through the lens of convolutional neural networks. *Remote Sensing of Environment*, *234*, 111448. https://doi.org/10.1016/j.rse.2019.111448

Wang, Q., Phillips, N. E., Small, M. L., & Sampson, R. J. (2018). Urban mobility and neighborhood isolation in America's 50 largest cities. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(30), 7735–7740. https://doi.org/10.1073/pnas.1802537115

Woetzel, J., Garemo, N., Mischke, J., Kamra, P., & Palter, R. (2017). *Bridging infrastructure gaps: Has the world made progress*. McKinsey & Company. https://www.mckinsey.com/capabilities/operations/our-insights/bridging-infrastructure-gaps-has-the-world-made-progress

Xu, Y., Olmos, L. E., Abbar, S., & González, M. C. (2020). Deconstructing laws of accessibility and facility distribution in cities. *Science Advances*, *6*(37), eabb4112.

Xue, J., Jiang, N., Liang, S., Pang, Q., Yabe, T., Ukkusuri, S. V., & Ma, J. (2022). Quantifying the spatial homogeneity of urban road networks via graph neural networks. *Nat Mach Intell*, *4*, 246–257. https://doi.org/10.1038/s42256-022-00462-y

Yabe, T., & Ukkusuri, S. V. (2020). Effects of income inequality on evacuation, reentry and segregation after disasters. *Transportation Research Part D: Transport and Environment*, *82*, 102260. https://doi.org/10.1016/j.trd.2020.102260

Zhou, Y., & Liu, Y. (2019). The geography of poverty: Review and research prospects. *Journal of Rural Studies*, *93*, 408–416. https://doi.org/10.1016/j.jrurstud.2019.01.008

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Fan, C., Xu, J., Natarajan, B. Y., & Mostafavi, A. (2023). Interpretable machine learning learns complex interactions of urban features to understand socio-economic inequality. *Computer-Aided Civil and Infrastructure Engineering*, 1–17. https://doi.org/10.1111/mice.12972