Tackling homelessness through AI powered social innovations: A novel and ground-breaking assessment of criminal victimization of homeless populations in los angeles employing predictive analytics and machine learning models such as ARIMA and LSTM

Ovais Rehman Shah, *Minot State University, Ovais.Shah@minotstateu.edu*
Lori Willoughby, *Minot State University, Lori.Willoughby@minotstateu.edu*
Nicholas Bowersox, *Minot State University, Nicholas.Bowersox@minotstateu.edu*

Abstract

Homelessness is a worldwide problem, and the number of homeless individuals in Los Angeles (LA) has increased dramatically in recent years. Although multiple studies have been undertaken to investigate various aspects of homelessness and its intersections with criminal victimization, none have used machine learning techniques to analyze the relationship between homelessness and victimization of homeless populations. To better understand the impact of victimization of homeless populations, we created a unified dataset by integrating three datasets obtained from Federal, State and Municipal Government Agencies to come up with significant findings. Feature Engineering was used to elicit the relationship between different dimensions of homelessness. Based on the Features extracted, Machine Learning Techniques were used to model the victimization of homeless people. Our findings revealed that victimization of homeless populations in LA is strongly linked to race, gender, age, and neighborhood divisions. Given that the main goal of this study is to help social service agencies implement social innovations, we applied two sophisticated machine learning approaches to anticipate the future of homeless people at risk: Auto-Regressive Integrated Moving Average (ARIMA) and Long Short Memory Networks (LSTM). Both models were trained from different angles to predict crime hotspots and the gender, race, and age group of more vulnerable people in the following two years. Finally, a few recommendations for social improvements that could improve services and programs for homeless victims of crime are presented to various departments and government bodies.

Keywords: Law, Machine Learning, Homelessness, Predictive Analytics, LSTM/ARIMA, Social Innovations

Introduction

Homelessness has become a critical global issue just like global warming, public health crises and climate crisis, etc. According to the United Nations Human Settlements Program, 1.1 billion people around the world are homeless (Wilde et al., 2020). A homeless individual is defined as someone living in unstable and un-permanent accommodations such as shelters, cars, parks, and abandoned buildings (United States Housing and Urban Development (HUD)). Although, homelessness in and of itself is a critical issue because it significantly burdens the community (Larimer et al., 2009) discovered an even more concerning issue is increased susceptibility of homeless populations to criminal victimization. According to the Los Angeles Police Department (LAPD), the crime rate against homeless populations in Los Angeles increased by 80

percent or more between 2019 and 2020. It is crucial to note that according to the Bureau of Justice Statistics (BJS), only 44% of all cases of crimes against homeless populations are reported. (Novac et al., 2006; Wachholz, 2005) posit that underreporting of crimes perpetuated against homeless populations is primarily driven by distrust of the police held by the homeless populations because they fear that they might not be believed or implicated of a crime through guilt by association. Moreover, the unwritten '*code of the street*' discourages police informants amongst homeless populations which makes it hard for law enforcement agencies in helping homeless victims of crimes. As a result of these complexities, determining the true figures is difficult (Leomporra & Hustings, 2018).

Given the seriousness of this social and global issue, it is essential that advanced intelligent algorithms be applied to current dataset to get important insights and create predictive forecasts that can aid policymakers, law enforcement authorities, and social service departments in taking appropriate action. Previously, common quantitative approaches such as interviews, statistical regression, and so on were used to get certain crucial insights on homelessness, but owing to the complexity of these approaches and the introduction of noise due to human involvement, the end findings were not very convincing or practical for social innovations. Applications of Artificial Intelligence (AI) to resolve social problems is crucial for guiding policy decisions around social services and law enforcement funding (Weinberger, 2018). In our assessment of multiple sources, no research could be found to date that used AI to explore the relationship between homelessness and criminal victimization of homeless populations.

This research makes a significant contribution to the existing corpus on social innovation by being the first of its kind to use machine learning techniques and predictive analytics to study the relationship between homelessness and victimization of homeless individuals. For the purposes of this study, a unified dataset was constructed by integrating three datasets collected from Housing and Urban Development Department (HUD), Los Angeles Homeless Services Authority (LAHSA), and Los Angeles Police Department (LAPD). By using feature engineering techniques, key features were extracted from the integrated dataset. Furthermore, two well-known machine learning forecasting algorithms Auto-Regressive Integrated Moving Average (ARIMA) and Long Short Memory Networks (LSTM) are employed to predict the types of personas that are susceptible to become victims of crime in homeless shelters. The goal of this study is to aid in service delivery and unlock legal, social services and policy-making innovations by predicting the number of homeless people in the coming years susceptible to crime. Furthermore, the study also helps identify the personas of people susceptible to crime to provide better social services to the most vulnerable.

Literature Review

Homelessness became a societal concern after the Great Depression of the 1930s, when over 1.2 million people lost their homes due to unemployment. During World War II, many homeless individuals became soldiers, but after the war, due to racial segregation in the United States, black veterans became homeless once more (Sugrue, 2014). Between the 1950s and 1970s, impoverished communities in major cities became homeless as large areas with skid rows were demolished to make space for the rich urban elite. According to the US Government, homelessness is classified as chronic (long-term) or episodic (short-term) in studies. According to Aron et al. (2001), 83% of chronic homeless people have had drug or alcohol problems at some point in their lives, and 53% have a mental health problem. According to studies by Aron et al. (2001) and Howard (2013), homeless populations with mental problems are more likely to become victims of crime.

Kershaw et al. (2003) presented the findings of an ONS (Office for National Statistics) survey of homeless individuals in Greater Glasgow, Scotland in which homeless communities were interviewed face to face on mental health, nutrition, social functions, drug usage, and other topics. Shinn et al. (1998) used interviews

to investigate predictors of entrance into shelters and subsequent housing stability. Machine learning is more efficient and capable of extracting crucial insights and forecasting actionable outcomes from very large datasets of social applications than traditional quantitative models (Hindman, 2015). Although machine learning has not been used to investigate the relationship between homelessness and victimization of homeless individuals, it has been used to investigate several other aspects of homelessness. Kuhn and Culhane (1998) used nearest centroid sorting cluster analysis technique to study the stay patterns of homeless populations. They grouped homeless people as transitionally homeless, episodically homeless, and chronically homeless to study their shelter occupancy, mental health, age group, etc. Furthermore, many researchers have used machine learning approaches to investigate various aspects of homelessness in recent years. Table 1 compiles a small number of significant works.

Table 1 Review of investigation of several aspects of homelessness by using machine learning techniques.

Author and year	ML/AI Approach	Homelessness aspect investigated
Slater (2011)	Socio-Economic Analysis (Qualitative)	The study used qualitative interviews, newspaper publications and synthesis are research inputs to argue that race-based gentrification led to eviction which in turn pre-disposed homeless individuals to crime and racial and political agendas.
Fowler et al. (2016)	Regression Analysis	Evaluated the prevalence and nature of problems faced by 265 homeless youth in a large mid-western metropolitan dataset. Model predicted housing instability and emotional and behavioral problems faced by homeless communities. Findings were used as inputs for improved foster-care planning for homeless youth.
Greer et al. (2016)	COX regression model	Tracked 11,105 families over 2 years who entered the homeless shelter in New York City. COX Regression Predictive Model built to predict the shelter entry rate. Findings from the model were leveraged as inputs into optimization of shelter services and planning.
Hong et al. (2018)	Logistic regression and K-Means	Created a unified comprehensive database of 6,000 homeless families. ML Model was applied to identify predictors around shelter re-entry, shelter term of stay, citizenship, age, medical conditions, and unemployment. Results of the Model were used to characterize shelter resident's health needs.
Chan et al. (2017)	Logistic regression and decision trees.	HUD dataset was used as an input into a ML model that maximized the number of homeless individuals assigned to a homeless program. A greedy algorithm was used to predict housing assignment settings for one million homeless youth cases. Outcome resulted in decision aids for housing prioritization amongst homeless youth.

VanBerlo et al. (2021)	HIFIS-RNN-MLP	Novel machine learning approach to predict a homelessness score from Canadian Homelessness Management System. 6,521 homeless individuals were used alongside shelter client attributes and dynamic features with a fine balance between recall and precision. First of its kind black-box machine learning model.
Wilde et al. (2020)	Forward-chaining temporal cross validation approach	Data Driven ML model was created in the UK with Homeless Link to build a human-center machine learning system to prioritize the alerts for volunteers that a new rough sleeper would enter the shelter system. Recommendation and classification of homeless rough sleepers was used to connect them to essential outreach services.
Yue et al. (2020)	Random forest	Information from Homeless Information Management Systems and Medicaid claims database was used to create a model around diseases, services and socio-economic scores around chronic illness for homeless population of 95,1717 that enrolled in California's Whole Person Care (WPC) Pilot Program.
Fisher et al. (2020)	Q-Learning (MDQL) & Modified Neural Fitted Q-Iteration (MNFQ) Algorithms	Simulation of transitions of individuals in a Canadian City between states of homelessness, staying in a shelter, staying in a hospital, or on the street and being housed over a course of 2 years. Outcome resulted in deep-learning model that can be applied to real-world data and test the impact of social policies pertaining to homelessness.

Auto-Regression Integrating Moving Average (ARIMA) is a univariate model that was proposed by Gwilym, Jenkins, and George Box in the 1970's. It is often used to forecast future trends and patterns in the economic, financial, and demographic research fields (Raza & Islam, 2018). It forecasts future values based on a linear mixture of its own historical values and previous errors (Jordan, 2018). Because it utilizes previous values of itself and previous mistakes, we hypothesized that the ARIMA model would be effective in predicting future trends against victimization of homeless individuals. The findings of ARIMA are cross verified through Long Short-Term Memory Units (LSTM). LSTM possesses the concept of long-term memory as well as short-term memory. LSTM performs better when long-term predictions are required while ARIMA is more suitable for short-term predictions. According to Alcocer, Leal & Torres (2019), LSTM models have achieved excellent time-series forecasting over ARIMA.

Data

Machine learning approaches rely heavily on the quality and diversity of the datasets used for training. To acquire optimal insights and make reliable forecasts, the dataset is constructed by integrating three datasets from three different sources.

Data Sources

Los Angeles Homeless Services Authority (LAHSA). This data set contains critical information about the homeless population, including socioeconomic status, mental health, age, race, gender, work status, and other details. This dataset makes it easy to track and record homeless cases in Los Angeles, as well as the homelessness rate. This dataset, however, is biased because it only contains the records of registered homeless individuals with the City of Los Angeles homeless services or homeless shelters .

Housing and Urban Development Department (HUD). This data set supplements the LAHSA data set. It provides yearly housing inventory counts, point-in-time counts, and yearly homeless assessments for each city and state for the past ten years.

Los Angeles Police Department (LAPD). This data collection covers nearly a decade of crimes in the city of Los Angeles. This data set contains a transcription of reports that were written on paper. As a result, bias might be reflected in the statistical analysis approach due to typographical mistakes.

Feature Extraction

All three datasets from the aforementioned sources spanning over the last 10 years (2010-2020) for the city of Los Angeles were merged to create a comprehensive description of the homeless population at risk of criminal victimization, The final dataset that was pre-processed for featured engineering entailed 2 million records which were ingested as a consolidated result set from the Los Angeles Police Department , Los Angeles Homeless Services Authority (LAHSA) and Housing and Urban Development (HUD) unified dataset. From these 2 million records, three types of features are extracted: demographic features, crime features and geographic features. The demographic information includes number of shelter check-ins, homeless count, etc. The crime related features cover victim sex, victim age, type of crime, type of weapon used, etc. The geography features cover number of housing units available in shelters, Number of CoC – Continuum of Care Locations, etc.

Data Preprocessing

Three techniques were employed to preprocess the data in order to make it ready for machine learning models: normalization, cleaning, and analyzing the collinearity of the data. The following equation is used to normalize the data:

$$final\ feature = \frac{Raw\ data}{Victims\ per\ neighborhood}\ X\ 100$$

To clean the data, null values and anomalies were addressed. Finally, collinearity was removed by plotting two features and using a correlation matrix.

Methodology

The ultimate purpose of this study is to assist police and social service agencies in developing interventions by identifying people in a homeless community who are at a higher risk of criminality. Moreover, the research aims to forecast the likelihood of crime against homeless individuals over time. Because the ARIMA model has widely been used in the literature to address social science and criminal justice concerns, we chose to use it to forecast the homeless community at a high risk of crime. Furthermore, while ARIMA is better at short-term forecasting, LSTM is utilized as a confirmer and long-term forecaster.

Auto-Regression Integrating Moving Average (ARIMA)

ARIMA is a statistical method that is used to analyze and forecast time series. It is a single-variate model that forecasts future values based on a linear combination of its own past values and past errors (Jordan, 2018). The type of ARIMA model is decided by three integer parameters: p (number of lag observations), d (degree of differencing), and q (size/order of moving window). The model uses a stationary time series (moving average) of the crime data set committed against people in homeless shelters to forecast future trends in victimization of homeless individuals based on past values and errors. The forecasting equation for the crime data set constructed with the ARIMA model is given below:

$$(1 - B)^d . b_t = S_t$$

The strategy for applying ARIMA model to the homeless victimization dataset is taken from (Jordan, 2018), which is presented in Figure 1.
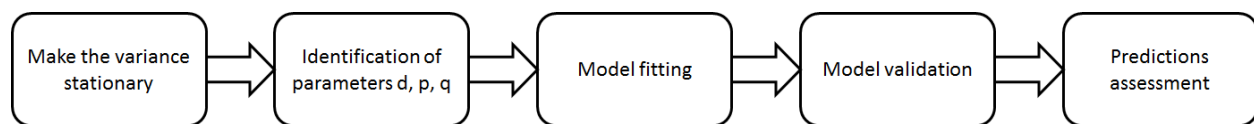

Figure 1. Steps to apply ARIMA model to homeless victimization dataset.

Long Short-Term Memory Units (LSTM)

LSTM is a kind of Recurrent Neural Network (RNN) proposed by (Hochreiter & Schmidhuber, 1997). RNNs have a memory system that helps to learn patterns in sequential data and predict the future using those patterns (Tan et al., 2020). However, LSTM is comprised of two types of states: cell states (long-term memory) and hidden states (short-term memory). One of the major benefits of LSTM is that it allows you to decide how much memory to use or forget (De, 2020). A typical LSTM network consists of a cell, an input gate, an output gate, and a forget gate. The architecture of LSTM is illustrated in Figure 2. The cell remembers values over time as information flows through it. The input gate checks information flowing into the cell. The forget gate controls how long the information can stay in each cell. The output gate checks the cell's values to compute the final output of the next cell.
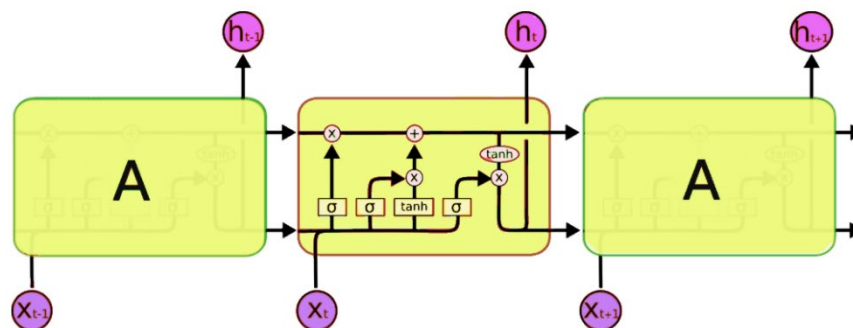

Figure 2. LSTM architecture.

We opted to use LSTM for victimization prediction and homeless prediction over time because of its applicability to various time series data in the literature (Malhotra et al., 2020). Furthermore, in this research, LSTM is employed as a confirmer of ARIMA findings.

Experimentation, Analysis, and Findings

In this section, the key insights are extracted with the help of data analytics and victimization of homeless individuals is forecasted using ARIMA and LSTM.

Applied Data analytics and findings

The aggregated data from three sources is analyzed from a variety of angles. However, due to the limited space available, just a few key relationships and findings will be discussed in this paper.

General findings.

With the help of a correlation matrix, it was discovered that there is a strong correlation between the victim's age, crime location, and time of day. Furthermore, the nature of crime and the type of crime weapon are also highly correlated. By doing a comparative analysis, we found that homeless males are 1.5 times more likely to be victimized than homeless women. Another interesting finding is that the rate of victimization of homeless individuals is growing exponentially over the time. For example, the number of cases reported in 2010 was fewer than 1000, whereas the number reported in 2019 was around 5,800. We also studied the impact of neighborhoods on crime rate. We discovered that the crime rate in Central division is four times higher than in any other division. However, a few other key crime hotspots are LA neighborhoods such as Rampart, Newton, and Pacific. Some of the most common crimes committed against homeless individuals are assault with a lethal weapon, simple battery assault, robbery, intimate partner violence, theft, rape, or weapon brandishing. Figure 3 uses a heat map to show the relative frequency of crimes, and assault with a deadly weapon is highly prevalent. We also examined the associations between different types of crimes and different times of day. We discovered that crime rates are lowest in the early morning and steadily rise as the day progresses. The peak time for crimes is between 5:00 PM and 11:00 PM. Additionally, there is also a peak in crime against homeless individuals observed in the afternoon.
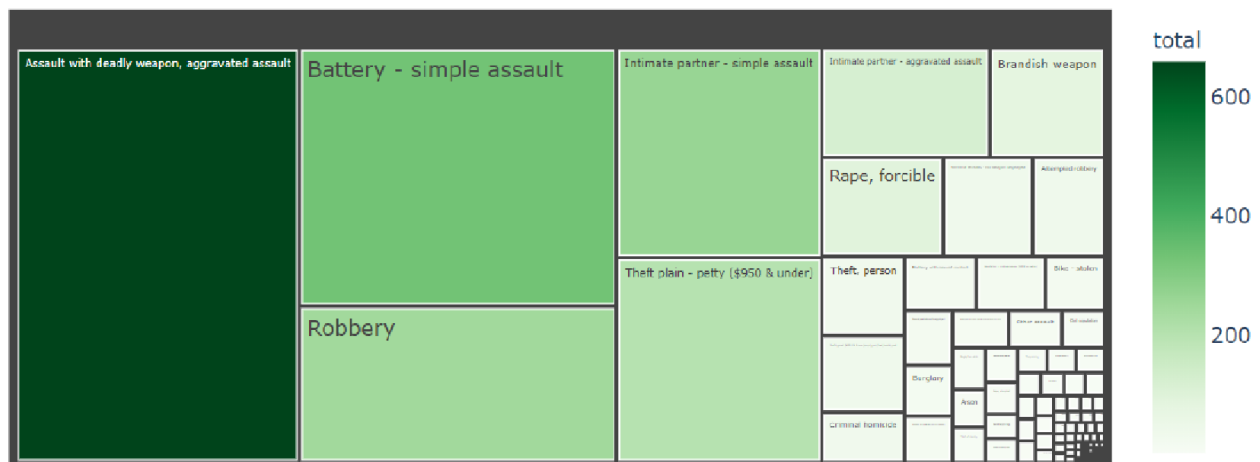

Figure 3. Heat Map demonstrating the crimes against homeless individuals. Darker colors represent the higher frequency crimes.

Homicides and homeless victimization

According to statistical analysis, the number of homicides among homeless individuals doubles every two years. In addition, when compared to West Valley, Devon Shire, Wilshire, and West Los Angeles neighborhoods of Los Angeles, our study reveals that Central, Hollenbeck, Newton, and Rampart neighborhoods have some of the highest homicide rates against homeless individuals. The analysis also reveals that guns and knives are some of the most predominant forms of weapons used to kill homeless individuals. Another important finding is that homeless people are more likely to be killed on the sidewalk, the street, a parking lot, an alleyway, or a park or playground versus inside a shelter o ra housing accommodation.

Gendered Dynamics of Homeless Victimization

We also analyzed the relationship of gender with victimization of homeless people. Our study reveals that men are more susceptible to victimization across all the top neighborhoods of Los Angeles except the financially affluent Harbor, Southwest, and 77th Street neighborhoods of Los Angeles. Moreover, men are the predominant victims of crimes such as homicide, aggravated and battery assault, robbery, and weapon brandishing while women are at higher risk of exposure to other crimes such as rape, aggravated sexual assault, and criminal threats.

Racial Dynamics of Homeless Victimization

We also tried to investigate the relationship among homeless victimization and racial dynamics. We found that Hispanics are at greater risk in Rampart, Newton, Hollenbeck, Foothill, Devonshire, Van Nuys, and West Valley areas of Los Angeles. Blacks are at greater risk in Central and 77 Street area. Whites are at higher risk in West Los Angeles, Pacific, Hollywood, and Devonshire. Furthermore, Black and Hispanic individuals are most susceptible to assault with deadly weapons and battery. Across the board, Blacks are more susceptible to counts of crime compared to Hispanics ranking second.

Age Dynamics of Homeless Victimization

Another factor that impacts victimization of homeless people is age. We found that adults and elders are at the highest risk of victimization across all the different Los Angeles neighborhoods. Teens are more susceptible to assault with a deadly weapon, battery or robbery, or intimate partner assault. Teens are more likely to be victimized on the sidewalk, street, or in a single-family dwelling, whereas the elderly and adults are more susceptible to violence in the parking lots, residences, and alleyways.

Forecasting using machine learning models

In this section, the crime against homeless people is forecasted for the next two years by using ARIMA model and LSTM model. The forecasts are based on gender, age group, racial group, and LA neighborhoods. Moreover, the outcomes of both models are compared against each other.

271

ARIMA model

As discussed in the preceding section, the ARIMA model is composed of three parameters called p, d, and q. The unified dataset contains around 2 million records. Because of the variance in the dataset, we opted to utilize Auto-ARIMA to choose parameters at random in an unsupervised manner rather than manually tuning them. The parameter settings for Auto-ARIMA are presented in Table 2.

LSTM model

LSTM in contrast to ARIMA is better for larger datasets and long-term predictions (Namini et al., 2018). The LSTM neural network model was trained on the LAPD crime victimization dataset to predict class labels such as age, race, descent, crime type, and weapon. These labels were utilized to get insight into the future of these parameters using predictive analytics. The LSTM model was trained using approximately half of the dataset. A normalization process was also carried out by deleting incorrect attributes. The LSTM unit was trained for 7.5 epochs on the training dataset to better predict crime rates where number of epochs were tuned with the help of loss graph for training vs. test datasets. The parameter settings for LSTM model are presented in Table 2.

Table 2: Parameter settings of ARIMA model and LSTM model.

Parameters of ARIMA Model		Parameters of LSTM Model	
Parameter	Value	Parameter	Value
Interval Start (p)	0	Epochs	10
Interval Max (p)	5	Batch size	3 months interval
Interval Start (q)	2	Optimizer	Adam
Interval Max (q)	5	Loss	MSE
Max(d)	2	LSTM	Units 50

Comparative results and discussion

With the help of data science analysis, we discovered in the preceding subsection that Central is the hotspot division for criminal acts against homeless people. The crime forecasting for the next two years in top neighborhoods of LA using ARIMA and LSTM are illustrated in Figure 4. Both models forecast that the number of cases in Central will rise in the next two years. However, when the growth rates of both curves are compared to their trends over the previous two years, LSTM appears to be more realistic. On the contrary, the cases in other divisions will follow a linear trend.
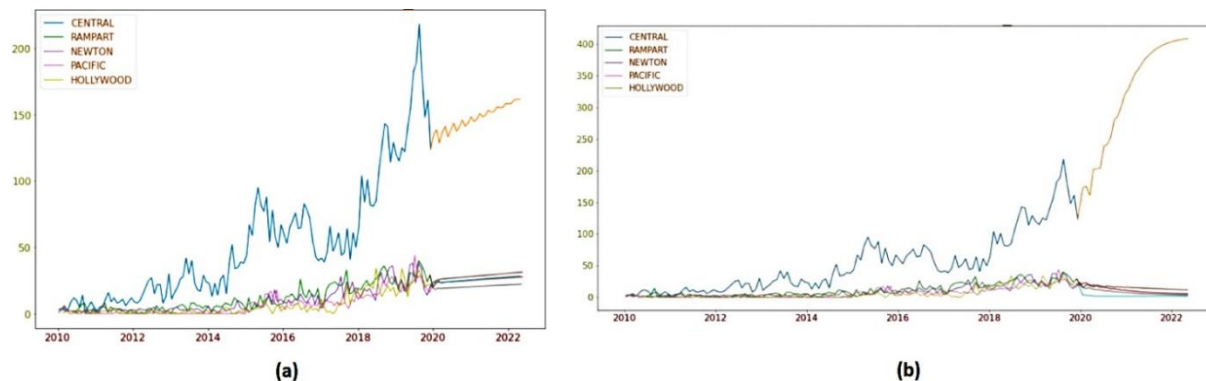

Figure 4. Two years of crime forecasting using ARIMA and LSTM in top five LA neighborhoods. (a) Prediction curves by ARIMA. (b) Predictions curves by LSTM.

Top crimes are also forecasted using ARIMA and LSTM. The plotted and compared curves are shown in Figure 5. Both models predict that assaults with deadly weapon will become more frequent in the coming years. The results of both models for battery assault crime and theft pain, however, are inconsistent. Moreover, ARIMA predicts a considerable increase in battery assault and theft pain crimes in the next two years, but LSTM predicts a reduction in these crimes. Lastly, both models predict a steady growth in the remaining three crimes over the next two years.
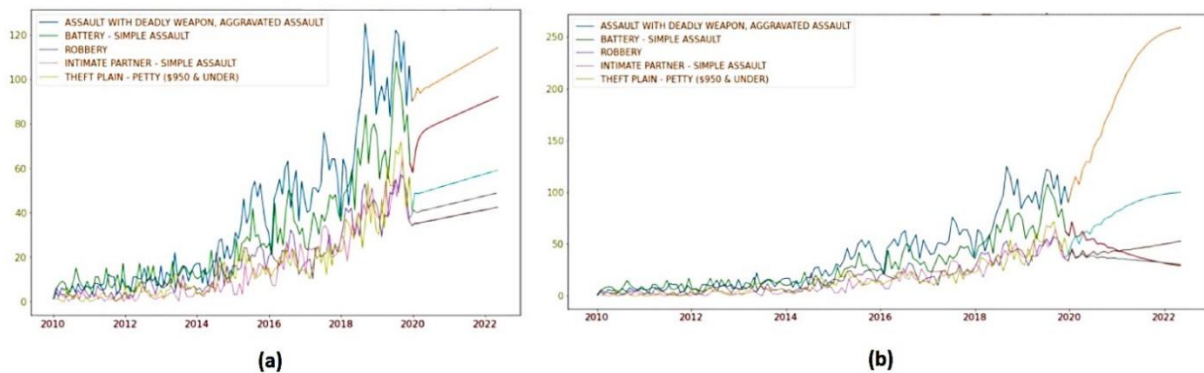

Figure 5. Two years of forecasting of top five crimes using ARIMA and LSTM. (a) Prediction curves by ARIMA. (b) Predictions curves by LSTM.

We also predicted how homeless persons will be victimized based on their gender. The comparative curves are plotted in Figure 6. According to ARIMA, the number of cases of both genders will steadily rise. LSTM, on the other hand, predicts an exponential growth in women cases over the following two years, rejecting the findings of ARIMA. As a result, we recommend LSTM over ARIMA.
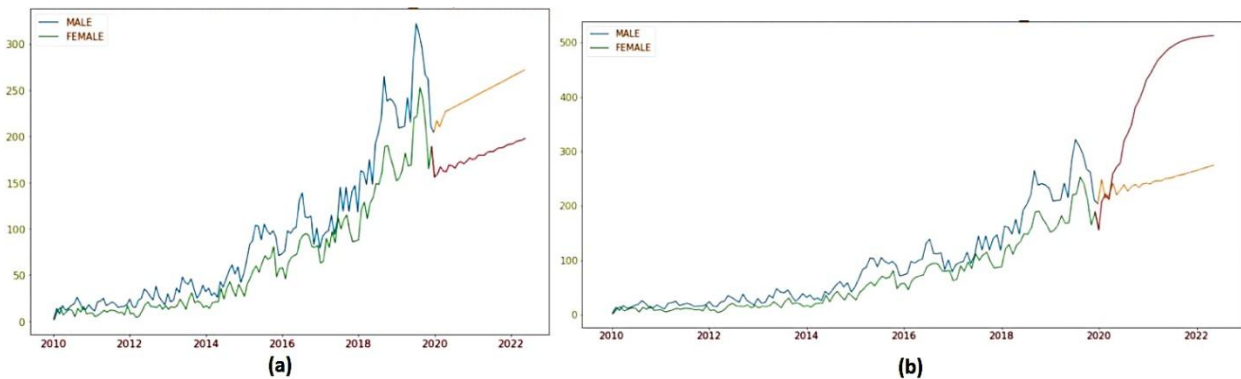

Figure 6. Two years of crime forecasting using ARIMA and LSTM for both genders. (a) Prediction curves by ARIMA. (b) Predictions curves by LSTM.

Victimization can sometimes be linked to the victim's race. As a result, we forecasted the number of cases against each race in the next two years. The results of both algorithms are presented in Figure 7. According to ARIMA, in the next two years, Blacks, Hispanics, and Whites will be more vulnerable to crime, while Asians and Others will be less susceptible. On the contrary, LSTM confirms ARIMA's findings, but forecasts that the crime rate against Blacks and Hispanics would rise exponentially in the next years.
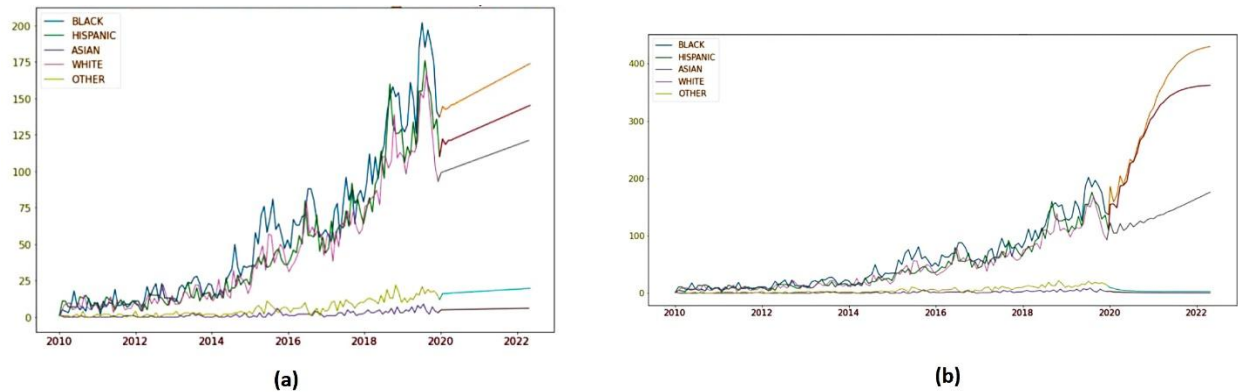

Figure 7. Two years of crime forecasting using ARIMA and LSTM for various racial groups. (a) Prediction curves by ARIMA. (b) Predictions curves by LSTM.

We also looked for a relation between homeless people's age and their victimization. The prediction curves of both models are plotted in Figure 8. Both models suggest that the rate of crime against elders and adults will continue to rise, while the rate of crime against the other three age groups would remain stable.
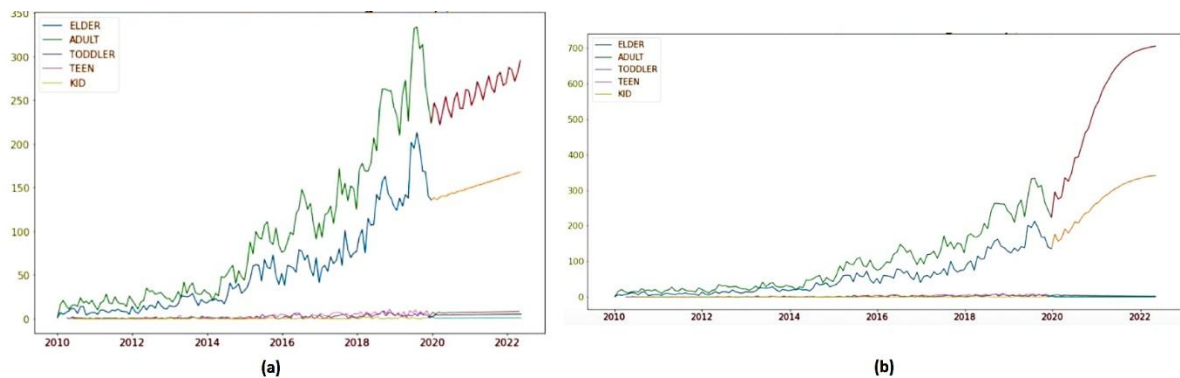

Figure 8. Two years of crime forecasting using ARIMA and LSTM for various age groups. (a) Prediction curves by ARIMA. (b) Predictions curves by LSTM.

Recommendations for Social Innovations

The objective of this study is threefold: (I) identification of correlated factors contributing to victimization of homeless people; (II) forecast the future of homeless people at risk using advanced machine learning techniques; and (III) make recommendations to social service departments and organizations on how to reduce homelessness and keep homeless people safe. In this section, we make a few recommendations for social innovations based on our findings.

The police department is recommended to be more active in neighborhoods highlighted by our models to reduce the crime rate. Moreover, by collaborating with social and health services, police and criminal justice

organizations may better plan protective services for homeless populations in high-risk age groups, racial groups, and genders.

According to a recent survey, the United States is among the top ten worst countries for racial equality that demonstrates how the country has prioritized solving racial imbalances (Drew, 2021). The predictions in this study about races at higher risk may aid concerned departments in taking timely action and pushing the United States in up the racial equality list of countries.

Government bodies such as HUD are recommended to use forecasts of these models to adjust the number of affordable units built each year and increase the inventory of existing houses for city inhabitants. This will improve the efficacy of existing social programs.

During our literature review, we discovered that one of the main causes of homelessness is unfair practices by larger investment firms, which result in higher rents and higher real estate prices. Hence, with the help of this study we recommend policymakers such as Real Estate Boards and the Los Angeles City Council, Homeless Prevention and Rapid Rehousing Program (HPRP), Homeless Emergency and Rapid Transition to Housing (HEARTH), and HUD to ally together to bring regulation against these unfair practices.

Conclusion

This study looks into the relationship between homelessness and victimization of homeless people. Data Science has revealed that there is a direct connection between victimization of homeless people and their age group, racial group, gender, and neighborhood division. We employed two powerful machine learning models named ARIMA and LSTM to forecast the homeless people at risk of victimization to aid social service departments. Machine learning algorithms predict that in the next two years, Blacks, Hispanics, and Whites will be increasingly vulnerable as victims of crime. Elders and adults, particularly women, will be at higher risk, according to our findings. Furthermore, the Central Division will be a crime hotspot in the future, with assault with a deadly weapon being the most common offense. Conclusively, in the next two years, elders or adult Black and Hispanic women in LA's Central division will be more vulnerable to assault with a deadly weapon. Our results should be used by police, criminal justice organizations, homeless care departments, and other concerned government bodies to help craft protective policing strategies, social policies and shelter services that enable homeless persons feel secure and safe and re-integrate back into society through combined agency of civil society social innovations.

References

Aron, L., Burt, M., Lee, E., & Valente, J. (2001). Helping America's homeless: Emergency shelter or affordable housing? The Urban Institute.

Chan, H., Rice, E., Vayanos, P., Tambe, M., & Morton, M. (2017, October). Evidence From the Past: AI Decision Aids to Improve Housing Systems for Homeless Youth. In AAAI Fall Symposia (pp. 149-157).

De, P. (2020). Housing market crash prediction using machine learning and historical data [Master's Projects, San Jose State University]. https://doi.org/10.31979/etd.ujur-h4j5

Drew, K. (2021, April 13). Survey: U.S. Among 10 Worst Countries for Racial Equality. Usnews.Com. https://www.usnews.com/news/best-countries/articles/2021-04-13/us-is-one-of-the-10-worst-countries-for-racial-equality

Fisher, A., Mago, V., & Latimer, E. (2020). Simulating the evolution of homeless populations in canada using modified deep q-learning (mdql) and modified neural fitted q-iteration (mnfq) algorithms. IEEE Access, 8, 92954-92968.

Greer, A. L., Shinn, M., Kwon, J., & Zuiderveen, S. (2016). Targeting services to individuals most likely to enter shelter: Evaluating the efficiency of homelessness prevention. Social Service Review, 90(1), 130-155.

Hindman, M. (2015). Building better models: Prediction, replication, and machine learning in the social sciences. The ANNALS of the American Academy of Political and Social Science, 659(1), 48-62.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Howard, E. (2013). Homeless: Poverty and place in urban America. University of Pennsylvania Press.

Hong, B., Malik, A., Lundquist, J., Bellach, I., & Kontokosta, C. E. (2018). Applications of machine learning methods to predict readmission and length-of-stay for homeless families: The case of win shelters in new york city. Journal of Technology in Human Services, 36(1), 89-104.

Islam, K. & Raza, A. (2020). Forecasting crime using ARIMA model. arXiv preprint. https://arxiv.org/ftp/arxiv/papers/2003/2003.08006.pdf

Kershaw, A., Singleton, N., & Meltzer, H. (2003). Survey of the health and well-being of homeless people in Glasgow. International Review of Psychiatry, 15(1-2), 141-143.

Kuhn, R., & Culhane, D. P. (1998). Applying cluster analysis to test a typology of homelessness by pattern of shelter utilization: Results from the analysis of administrative data. American journal of community psychology, 26(2), 207-232.

Larimer, M. E., Malone, D. K., Garner, M. D., Atkins, D. C., Burlingham, B., Lonczak, H. S., ... & Marlatt, G. A. (2009). Health care and public service use and costs before and after provision of housing for chronically homeless persons with severe alcohol problems. Jama, 301(13), 1349-1357.

Leomporra, A. & Hustings, M. (2018). Vulnerable to hate: A survey of bias-motivated violence against people experiencing homelessness in 2016–2017 [Report]. National Coalition for the Homeless. https://nationalhomeless.org/wp-content/uploads/2019/01/hate-crimes-2016-17-final_for-web2.pdf

Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015, April). Long short term memory networks for anomaly detection in time series. In Proceedings (Vol. 89, pp. 89-94).

Shinn, M., Weitzman, B. C., Stojanovic, D., Knickman, J. R., Jimenez, L., Duchon, L., ... & Krantz, D. H. (1998). Predictors of homelessness among families in New York City: from shelter request to housing stability. American Journal of Public Health, 88(11), 1651-1657.

Siami-Namini, S., Tavakoli, N., & Siami-Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. 17th IEEE International Conference on Machine Learning and Application, 1394–1401. https://doi.org/10.1109/ICMLA.2018.00227

Slater, T. (2011). Gentrification of the city. The New Blackwell companion to the city, 1.

Sugrue, T. J. (2014). The origins of the urban crisis: Race and inequality in postwar Detroit - Updated edition. Princeton University Press.

Tan, Q., Ye, M., Ma, A. J., Yang, B., Yip, T. C. F., Wong, G. L. H., & Yuen, P. C. (2020). Explainable Uncertainty-Aware Convolutional Recurrent Neural Network for Irregular Medical Time Series. IEEE Transactions on Neural Networks and Learning Systems.

Torres, J. A. M., Leal, E. T., Galván, G. R., & Alcocer, U. M. R. (2019). UN ENFOQUE BASADO EN DATOS PARA PREDECIR EVENTOS DELICTIVOS EN CIUDADES INTELIGENTES (A DATA-DRIVEN APPROACH FOR PREDICTING CRIMINAL EVENTS IN SMART CITIES). Pistas Educativas, 41(134).

VanBerlo, B., Ross, M. A., Rivard, J., & Booker, R. (2021). Interpretable machine learning approaches to prediction of chronic homelessness. Engineering Applications of Artificial Intelligence, 102, 104243.

Weinberger, G. (2018). The Criminal Justice System and More Lenient Drug Policy (Doctoral dissertation, The Pardee RAND Graduate School).

Wilde, H., Chen, L. L., Nguyen, A., Kimpel, Z., Sidgwick, J., De Unanue, A., ... & Vollmer, S. (2020). A Recommendation and Risk Classification System for Connecting Rough Sleepers to Essential Outreach Services. arXiv preprint arXiv:2007.15326.

Yue, D., Pourat, N., Chen, X., O'Masta, B., Huynh, M., & Xin, K. (2020). How to Identify Homelessness Using Administrative Data. Health Services Research, 55, 139-140.