SMU Data Science Review

Volume 8
Number 2 *Summer 2024*

Article 3

2024

Geospatial Temporal Crime Prediction Using Convolution and LSTM Neural Networks: Enhancing the Las Vegas Cardiff Model

Corey D. Holmes
Southern Methodist University, coreydaneholmes@gmail.com

Christian Orji
Southern Methodist University, christianuorji@gmail.com

Chris Papesh
University of Nevada, Las Vegas, chris.papesh@unlv.edu

Geospatial Temporal Crime Prediction Using Convolution and LSTM Neural Networks: Enhancing the Las Vegas Cardiff Model

Corey D. Holmes[1], Christian Orji[2], Chris Papesh[3]
[1] Master of Science in Data Science, Southern Methodist University,
Dallas, TX 75275 USA
[2] University of Nevada, Las Vegas
Las Vegas, NV 89154 USA
daneh@mail.smu.edu
corji@mail.smu.edu
chris.papesh@unlv.edu

Abstract. According to the Department of Justice, more than half of violent crimes go unreported to law enforcement in the United States (Kollar et al., 2018). This data gap reduces the opportunity to implement proven solutions in the areas with the greatest need. In 1996, Dr. Shepherd developed the Cardiff Model with the aim of bringing together hospitals, law enforcement, and community leaders through the sharing of data. We partnered with ongoing efforts to implement the Cardiff Model in Las Vegas, Nevada. Our goal was to provide a geospatial temporal model that can predict the next 30 days of crime. By utilizing the Metropolitan Police Department's (LVMPD) violent crime database, we were able to use a combination of long short-term memory (LSTM) and convolutional neural network (CNN) models to predict where and when violent crimes are likely to occur. Our total crime LSTM model produced an RMSE of 8.621 over a 30-day horizon. When incorporating the spatial component, our CNN and LSTM model produces an MSE of 0.0009 over the same horizon. These findings show that with sufficient latitude and longitude tracked violent crime data, we're able to accurately produce predictive heat maps. This establishes a framework to expand on the current process which develops heat maps aggregated over historical time periods. By adding to the existing drug overdose heat maps built by Grard et al. (2023), we hope to provide local leadership with the necessary tools to achieve similar reductions in violent crimes seen in Cardiff Projects across the globe.

1 Introduction

According to the Department of Justice, more than half of violent crimes go unreported to law enforcement in the United States (Kollar et al., 2018). This, inherently, leaves a significant gap in our understanding of where and the volume to which crimes are occurring. In the 2019 FBI Crime Report, it's estimated that nationwide violent crimes reached 1.2 million (Violent Crime, 2019). This translates to a 366.7 per 100,000 people occurrence rate. Given the Department of Justice's estimated reporting rate, those numbers could be severely underestimated.

In 1996, Dr. Shepherd developed and implemented a model for pooling the data resources to provide a better picture of violent crimes in Cardiff, Wales. A core tenant of the, now named, Cardiff Model is the ability to create violence maps with the combined data of law enforcement and hospitals. These maps provide not only the police, but community leaders with a better and more informed picture of violence in their neighborhoods. In the time since its creation, the Cardiff Model has been implemented in cities around the world. In the CDC's study of 14 similar cities to Cardiff, Wales, they found a "32% reduction in police recorded injuries" and a "42% reduction in hospital admissions for violence-related injuries" (Kollar et al. 2018).

To date, sixteen different cities across the United States have Cardiff Model projects underway (Sixteen US Cities in National Cardiff Violence Prevention Network, 2023). The model was officially adopted into US policy in 2018. Looking at the same 2019 FBI crime reports for Las Vegas, we see a violent crime rate of 525.7 per 100,000 people (Violent Crime, 2019). This is 1.43 times the national average. It's for this reason that Chris Papesh of the University of Nevada, Las Vegas (UNLV), started the process of bringing the Cardiff Project to Las Vegas. Over the last several years, they have begun the hard work of forging connections with critical stakeholders at the Las Vegas Metropolitan Police Department (LVMPD), Las Vegas hospital systems, and community leaders. One of the fruits of those efforts has come in the creation of an overdose heat map tool generated from hospital data (Girard et al., 2023).

Las Vegas offers a unique cocktail of social and political circumstances that appear to reinforce cycles of violence and crime. Long known as “Sin City”, the Las Vegas Convention and Visitors Authority has advertised the well-known slogan, “what happens here stays here”. Its tolerance towards high-risk activities has made Las Vegas a destination for many who seek out these activities. While Las Vegas offers significant low-risk activities like restaurants and shows, the primary attraction is gambling. Nevada gambling goes back nearly to its statehood. Only five years in, Nevada officially legalized gambling in 1869 (Task Force & Gemignani, n.d.). By the late 1940s, Nevada began to see increased growth due to California beginning to shut down illegal casino operations (Task Force & Gemignani, n.d.). According to the UNLV Center for Gaming Research, between 1984 and 2023, Las Vegas strip casino revenue went from $2.16 billion to $23.75 billion (University Libraries, University of Nevada, Las Vegas, n.d.). The US Bureau of Labor Statistics estimates the 1984 dollar to be worth $0.34 compared to 2023 (CPI Inflation Calculator, n.d.). Adjusting for inflation, this brings the revenue growth to 270% over the 19-year period of the study. Next on the list of high-risk activities are sex, drugs, and alcohol. Nevada is the only state with legalized prostitution. While it is no longer legal within Las Vegas, legal prostitution is available only 60 miles west of Las Vegas in Nye County. For anyone who has spent time in Las Vegas casinos, it is well known that alcohol is complementary when playing. In 2012 the Las Vegas-Paradise MSA showed statistically significant high rates of binge alcohol use (25.6%) in persons over the age of 12 (SAMHSA, 2012). This was compared to a national average of 23.2%. Furthermore, according to CDC data, over 109,000 nationwide fatalities in 2022 were due to drug overdoses, with synthetic opioids like fentanyl contributing to nearly 70% of these deaths (Tanz et al., 2024b). According to the latest 2021

CDC death rates, Nevada's drug overdose death rate per 100,000 is 29.2 (Drug Overdose Mortality by State, n.d.). This compares to the national average of 20.1 per 100,000. Legalized gambling, legalized sex, and higher than average alcohol and drug use gives Nevada a particularly unique cross section of overlapping high risk activities. We explore their connection to cycles of crime in the upcoming literature review section.

The FBI's national crime report indicates that the overall number of violent crimes in 2019 has leveled off, reaching the same levels as in 2015 (Crime in the United States, 2019). Although we don't have current crime rates, it was recorded that there were 366.7 incidents per 100,000 individuals in 2019. In comparison, Las Vegas recorded 525.7 incidents per 100,000 individuals in the year. This represents a 43% increase over the national average. Now, more than ever, a strong and capable Cardiff Model implementation is needed in the city of Las Vegas. It is our aim to continue the work of SMU data scientists before us and increase the tools available to community leaders through predictive modeling. We will be using available data sets to provide a better understanding of violent crimes in Las Vegas communities. We will be utilizing data from the LVMPD violent crime database to build predictive models using a combination of neural networks that capture both the geospatial and temporal components found in the data. By building our model with a foundational LSTM neural network, we enable future development that incorporates temperature and event data. Towers et al. (2018) showed strong correlation between crime levels and events like holidays and festivals. They also identified that the inclusion of temperature forecasts can increase the short-term prediction performance of violent crime models. They also showed that precipitation forecasts may provide additional short term prediction benefits for assault and batteries. We chose to exclude them from our current model due to the additional complexity it introduces to the long-term deployment of the model. Since Las Vegas is primarily a tourist city, the various events it hosts are likely to provide relevant data for future iterations of the model. Furthermore, the underlying data structure required for building the recurrent neural network (LSTM) portion of our model facilitates the inclusion of future hospital data with minimal rework.

Through the creation of these tools, we aim to accomplish three things. First, we want to provide a future looking toolset for the established hotspot policing (HSP) strategy meetings outlined in Corsaro et al. (2023). Second, we hope to show the capability and benefits available to any parties hesitant to get involved. Most importantly, we hope to reduce the number of violent crimes and overdoses in Las Vegas, Nevada. Previous studies have established an overlap between victims and victimizers (victim-offender overlap) (Averdijk et al., 2016). Research is uncertain on if this is a causal relationship or due to some underlying trait or environmental correlation (Turanovic & Pratt, 2013). None the less, it is our hope that by reducing the number of victims, we'll be able to break the cycle of victims becoming victimizers.

2 Literature Review

2.1 Violent Crime and Overdoses

In our introduction, we outlined Las Vegas's unique offering of readily available high-risk activities like gambling, sex, alcohol, and drugs. When combining these with violent crime rates 43% higher than the national average, it creates an environment susceptible to self-reinforcing tendencies. At the forefront of this discussion is the overlap of these high-risk activities and the underlying trait of self-control. Self-control is the ability for an individual to change their response to align to some outside standard or goal (Findley et. al., 2018). In this same study, Findley et. al. (2018) ranks each state by the two underlying factors initiation self-control and inhibition self-control. They showed Nevada to rank last in both. On this standardized score, Nevada had a (-2.82) initiation and (-2.77) for inhibition. For comparison, the closest states on initiation self-control were West Virginia with (-2.24) and on inhibition self-control was Delaware with (-2.13). Rhode Island ranked number one with 1.66 for initiation self-control and Texas with 1.72 for inhibition self-control. This makes Las Vegas have both one of the highest cross sections of readily available high-risk activities, 43% higher than national average violent crime rates, and last in trait self-control.

Turanovic and Pratt (2013) outline the overlap between victims and offenders. Their findings show that victims with low self-control showed an increased likelihood to use drugs and alcohol as a form of self-medication. This is relevant because the study showed that this increase was significant even when accounting for the trait of low self-control. They go on to outline that low self-control, victimization, and drug use independently increase the likelihood of future violent offenses. Furthermore, victims who use drugs and alcohol post-victimization, are shown to engage in violent activities at a much higher rate (Turanovic & Pratt, 2013). Alternatively, those who exhibit higher trait self-control are more likely to seek help in places that take longer to see a return and are less destructive. An example of this would be seeking therapy or attending group sessions with those who have experienced similar trauma. This is reinforced by Gottfredson and Hirschi (1990) who argue that those with low self-control engage in activities that provide immediate gratification rather than long term gains. Sex, drugs, and alcohol are primary outlets for immediate and short-term relief. While much blame can be pointed at a lack of self-control, Averdijk et al. (2016) showed that even when accounting for external factors like self-control, anxiety, depression and many other pre-existing conditions, victims still have a statistically higher probability to become victimizers. They posit that victims undergo a shift in their cost benefit analysis that skews their perception to the benefits of performing violent crime and away from the costs. The perceived benefits of engaging in violence begin to outweigh the perceived costs. This shift in perception may arise from feelings of injustice, the need for retribution, or a desire to reassert control over one's life. In this altered state of judgement, violent acts may be viewed as a necessary, even rational, response to one's circumstances. This transformation from victim to perpetrator, although not necessarily casual and possibly linked to an underlying disposition such as low self-control, has been supported by various research, highlighting a concerning trend that victims are at higher propensity to victimize others (Averdijk et al., (2016).

Within the normal victim offender overlap, we see some that are injected into the cycle due to drug use. It's estimated that between 3% and 19% of all people who take prescription pain medication will develop a subsequent addiction (Opioid Use Disorder, n.d.). Pierce et. al. (2017) outlines the relationship between heroin and opiate addicts to crime. They showed that those who are dependent on heroin or opiates have an increased relationship to criminal activity than those who are not dependent. These activities are typically related to gaining additional finances to obtain more drugs. The inverse was also seen. Those who are involved in criminal activities were more likely to use drugs. While states such as West Virginia, Louisiana, Kentucky, and Tennessee have some of the highest death rates per 100,000 people in the country, Nevada still finds itself above the national average. At nearly 43% higher overdoses per 100,000 than the national average, Nevada is faced with an uphill battle. This increased proximity to drug use and the violence that comes with it further complicates the challenges facing community leaders. Adding to the complexity, the cycle of violence theory suggests that witnessing or experiencing violence can set off a chain of violent acts. Such behavior is not solely the result of personal vulnerabilities but is also shaped by environment and social context. High-crime areas with limited access to mental health care, substance abuse treatment, and support networks fail to provide the necessary tools for individuals to cope with and recover from trauma, possibly leading to an increased risk of substance abuse and violent behaviors. This situation is exacerbated in socially disorganized neighborhoods where systemic inadequacies fail to arrest the cycle of violence, allowing it to persist and even escalate.

In a study assessing the relationship between alcohol consumption and violence in Norway from 1880 to 2003, while accounting for variables such as unemployment and divorce, Elin K. Bye finds a clear correlation (Bye, 2007). Bye, employing ARIMA models on time-series data demonstrates that a one-liter increase in alcohol consumption per person per year is linked to an 8% increase in violence rates. This association persists even when controlling for potential confounders. Only divorce showed a significant association with violence rates among the seven considered confounders. The findings suggest a potentially causal relationship between alcohol consumption and violence and highlight the role of alcohol as a critical factor in explaining variations in violence over time. This is further supported by Norström et. al. (2010) who demonstrate that increased heavy episodic drinking is significantly associated with heightened violent behavior, especially among individuals prone to suppressing their anger.

Policies aimed at breaking this cycle must address both the individual and environmental factors at play. They should not only focus on providing comprehensive support services but also on improving community resources and social infrastructure. Such holistic approaches may offer the dual benefit of aiding individual recovery while fostering community resilience against the perpetuation of violence. Considering these insights, there is a pressing need for a concerted effort that spans across multiple sectors-healthcare, law enforcement, social services, and community organizations- to develop interventions that are sensitive to the nuanced relationship between victimization and subsequent violence. Strategies that emphasize rehabilitation over punishment, and that recognize the social determinants contributing to the cycle of violence, will be pivotal in transforming the narrative for victims and preventing the transition to victimizer.

2.2 Policy and Policing

Bohnert et al. (2021) explores the relationship between policing practices and overdose mortality in urban neighborhoods. Data from New York City's police precincts from 1990 to 1999 were analyzed to determine if there was a correlation between misdemeanor arrest rates-a measure of police activity- and drug overdose deaths. The study found that higher levels of police activity were associated with increased overdose mortality. This suggests that intense policing, possibly creating a fear of arrest among drug users, could deter people from seeking medical help for overdoses. The implication of these findings is significant for public health and law enforcement policies, particularly in urban areas where drug is prevalent. The researchers suggest that while aggressive policing can lower crime rates, it might simultaneously increase the risk of drug overdose fatalities, indicating a complex balance between law enforcement and community health outcomes. These findings are consistent with the Cardiff Model which seeks to work with community leaders to drive change. Increased policing at the cost of increased overdoses is still a net loss for the community of Las Vegas. As discussed previously, drug use is another vector into the cycle of crime. Effective policy and process must simultaneously reduce crime and treat the causes to have a lasting effect.

One of the primary tools in the belt of law enforcement to combat crime is through hotspot policing (HSP). Braga and Wisburd (2022) showed a statistically significant reduction in crime in areas that received HSP. Additionally, they showed that the adjacent areas did not show a statistically significant increase in crime. It is reasonable to assume an overall reduction in crime rather than a shift in its spatial attribute. More specifically to Las Vegas, Corsaro et al. (2023) showed statistically significant reductions in calls for service of violent incidents and overall calls for service in areas that received HSP. Particularly important to the findings was that even areas with higher-than-normal policing also saw statistically significant decreases. This indicates no evidence of an observed cap to the effectiveness of HSP. The paper by Joseph G. Bock, titled "The efficacy of violence mitigation: A second look using time-series analysis," published in Political Geography, reassesses previous findings on violence mitigation efforts in the Horn of Africa. While earlier research by Meier, Bond, and Bond (2007) found a positive correlation between violence mitigation and organized raids, suggesting that mitigation efforts might inadvertently contribute to violence, Bock's analysis introduces a different statistical approach. By employing a "de-trending" method commonly used in economics and finance, Bock's study finds an opposite and statistically significant result: violence associated with organized raids is negatively correlated with mitigation efforts when data are de-trended for time and seasonality. This implies that, contrary to previous findings, mitigation efforts are negatively associated with violence, challenging the notion that such efforts do more harm than good. Bock's analysis underscores the importance of considering temporal dynamics in peace research and suggests that the timing and nature of mitigation efforts are crucial for their success in preventing violence.

2.3 Cardiff Model

To combat the rising crime rate in his community, Dr. Shepherd of Cardiff University in Wales developed a system that brought together data sources from hospitals and law enforcement to build more complete crime maps (Kollar et al., 2018). A core tenant of this model was the idea of data sharing. By bringing together the desperate data sources, all invested parties would gain a better understanding of where crime was occurring in their community. This information could then be used by law enforcement and community leaders to seek answers as to the why of hotspots and treat them accordingly. This method proved to be so successful that it has gone on to be implemented from the "Netherlands to Australia and South Africa" (Sixteen US Cities in National Cardiff Violence Prevention Network, 2023). It is also noted that the United States adopted the Cardiff Model as official policy in 2018. This resulted in the CDC creation of the toolkit referenced elsewhere in this study. To date, sixteen US cities have ongoing Cardiff Projects. In the UK, the recent 2022 study showed that across 14 similar cities to Cardiff, Wales, information sharing also lead to cost savings in addition to the reductions in crime (The Cardiff Model for Violence Prevention - Cardiff University, 2022). In an older but independent study, Boyle et al. (2013) were able to find a reduction in crime in a similar city to Cardiff. However, they were not able to attribute causal effects to the implementation.

Previous works done by Grard et al. (2023) have aimed to curb the drug overdose metrics outlined above through the creation of drug overdose heat maps in Las Vegas. Grard et al. (2023) utilized the chief complaint field of medical records to create heat maps but due to the inconsistent nature of the records, they were unable to build more predictive models from the source material. One avenue to improve the data quality coming from hospitals was assessed by Nguyen et al. (2022). They showed that by creating a short screening for nurses to fill out, they were able to gather data beneficial for implementing the Cardiff Model. Beyond that, they found that nursing staff found the additional screening to be in alignment with their overall mission and didn't interfere with their workflow. These findings are important because the primary gap in implementing the Cardiff Model is effectively utilizing the hospital data. Nguyen et al. (2022) reaffirmed the sentiment shared by Grard et al. (2023) that hospital records were inconsistent and difficult to generate predictive analytics from.

2.4 Spatial/Temporal Modeling

Traditional crime modeling techniques, as detailed by Dakalbab et al. (2022) in their analysis of 128 studies, predominantly rely on tools like ArcGIS to identify temporal and geographic crime hotspots. This approach, known as crime density prediction, involves calculating the number of crime incidents within specific areas, such as neighborhoods or sections of a map, relative to the population. While this method helps in pinpointing regions for targeted policing strategies like hotspot policing, its major limitation is the lack of temporal analysis. It can differentiate crime rates by days of the week but fails to forecast future trends based on historical data. Prathap (2023) supported these findings and highlighted the integration of Kernel density estimation (KDE) to enhance pattern recognition and hotspot detection, allowing for adjustable metrics in the analysis. The study by Dakalbab et al. (2022) also noted a preference for supervised machine learning (ML) algorithms among researchers, emphasizing the importance of creating interpretable models, especially at a time when law enforcement's public credibility is under scrutiny.

When digging further into the methodology of crime science, we see most analysis being done at the week level (Curiel, 2021). To drive down to the daily or hourly level, research must account for the higher prevalence of zero values. While crime is prevalent on the week and month scale, it is much rarer in these smaller windows of time. Curiel (2021) outlines the trade-off between these windows of time by noting that the meaningfulness might be lost as the window is expanded to increase the number of occurrences. For example, is it relevant to know how many crimes happen between 12:00 AM and 10:00 AM? Does this allow authorities to create reactive action items? These questions must be considered when setting the window size. One method for handling this zero-occurrence phenomenon is to map the zero values as negative values (Liang et al., 2022). This maintains the relative importance of each measure while not causing as many issues with ML models. This technic is referred to as the Priori Knowledge-based Data Enhancement (PKDE) strategy. Liang et al. (2022) also used a Neural Attentive framework to generate their hourly crime predictions.

In a recent study done by Jagait et al. (2021) on load forecasting for the electric grid, they found that by using an ARIMA and RNN ensemble, they were able to produce a model more accurate than the sum of its parts. It enabled them to model the underlying trends while still being able to include more current external events. The study by Hu et al. (2022) investigates spatial-temporal patterns of fatal drug overdose risk in British Columbia from 2015 to 2018. Key findings indicate that rural areas face a higher risk of fatal overdoses compared to urban centers, possibly due to less access to harm reduction services. The research utilized logistic regression and Generalized Additive Models (GAM) to analyze the data, with results presented as heatmaps to identify high-risk regions. The presence of harm reduction sites correlated with lower overdose risks. The study emphasizes the need for targeted harm reduction services in rural areas to mitigate the increasing trend of fatal overdoses province wide.

This study aims to give meaningful and actionable data to the key stakeholders of the City of Las Vegas. By providing a geospatial temporal predictive model, we hope to enable the LVMPD to more efficiently allocate their limited resources. Additionally, we hope to provide insights into spatial relationships that can assist community leaders in coming up with practical solutions to meet the needs of their community and curb violent crime rates.

3 Method

3.1 Data Handling

For this study, we utilized data from the Las Vegas Metropolitan Police Department, encompassing detailed crime reports from January 1, 2019, to February 29, 2024. The original data set contained over 1.3 million crime events categorized into 38 different types. After filtering to only include violent crimes (robbery, assault, homicide, and kidnapping) within the chosen location range, we were left with 82,184 individual incidents. Location ranges were established by picking four corners on the map that would encompass the city and fringe communities. The resulting latitudinal and longitudinal ranges were (35.96, 36.32) and (-115.37, -115.01), respectively. This excluded some data that was significantly outside the city. We then counted the number of individual grids when evenly spaced by .01 steps in both longitudinal and latitudinal dimensions. This resulted in a 36 by 36 grid with dimensions of approximately 0.9 by 1.1 km (0.6 by 0.7 miles) in height and width, respectively. We then aggregated the incidents by date and then by date and location. The corresponding arrays had dimensions [1886, 1] and [1886, 36, 36].

3.2 Model Design

Our ensemble model is composed of two distinct neural networks. The first is a specific kind of recurrent neural network (RNN) called Long Short-term Memory (LSTM). LSTMs are especially useful for long sequential data sets because they are robust against the vanishing gradient problem that traditional RNNs often encounter. This model was used to predict the total violent crime rate by day.

The second model was a combination of convolution networks (CNN) and LSTM layers. By passing each 36 by 36 layer through the CNN layers, the model was able to extract important structural features that account for spatial elements. These features were then passed to the LSTM to establish historical patterns within each of these layers. The result is a model that can take a 36 by 36 grid of crime values and output a new 36 by prediction of values. The input data was normalized between 0 and 1. Values below 0.01 were rounded to zero. This was done to account for fractional crime values so low that they significantly clouded the final results. The data was then normalized so that each layer of the [30, 36, 36] prediction array individually had a sum of 1. This, in effect, translated the spatial data into a probability distribution. For the final geospatial temporal prediction, we distributed the total crime prediction from our first LSTM across the second model's probability distribution to achieve total crime predictions by location and day.

3.3 Plotting Predictions

In order to map the data properly, each cell prediction was translated back into its coordinate dimensions. Each total was assigned the centered latitude and longitude of its respective cell and the entire matrix was translated back into dictionary with each row being unique to the location and date. This step was essential to translate the data back into a usable form for plotting tools.

4 Results

4.1 Exploratory Findings

The preliminary analysis revealed distinct seasonal patterns in crime rates, notably with yearly and weekly trends. Both ARIMA and LSTM models successfully demonstrated predictive capabilities, capturing both the trend and the cyclic nature of crime occurrences with significant accuracy. Figure 1 shows a clear peak around the summer months of each year.

Daily Violent Crime Totals

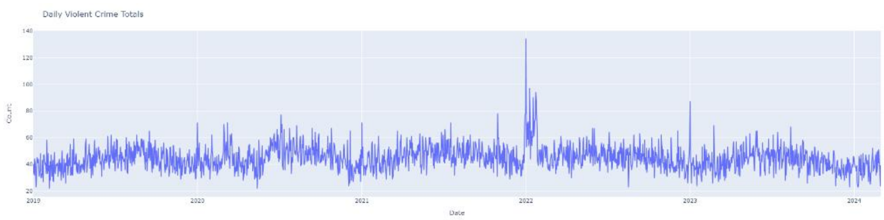

Figure 1 – shows the daily violent crime totals from Jan 2019 through the end of Feb 2024.

Seeing the underlying trends within the week is a bit more challenging. For this, we look to the Auto-Correlation (ACF) and Parzen Frequency Window plots. Figure 2 shows peaks at lag 7, 14 and 21 in the ACF and peaks in the frequency about every 0.14. This corresponds to 1/7 frequency.

Total Daily Violent Crimes - ACF and Parzen Frequency Window

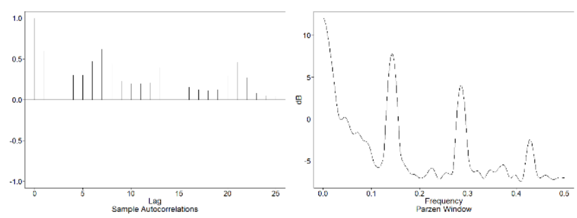

Figure 2 – shows the strong underlying weekly seasonality present in the data.

Furthermore, by looking at total violent crimes by day of the week and month of the year, in Figure 3 and 4 respectively, we can notice some underlying trends. For the weekly view, we can see slightly higher rates between Saturday and Monday with the middle of the week being lower. This same trend can be seen in the monthly data that shows a subtle hump between May and September. One notable outlier appears to be the month of January. We believe this is being skewed by the significant spike in violent crimes that we see in January of 2022. Our research has found no reason to exclude this data from the model, but it should be noted that this time period saw numbers significantly higher than any other time period before or after it.

Crime Counts by Day of the Week

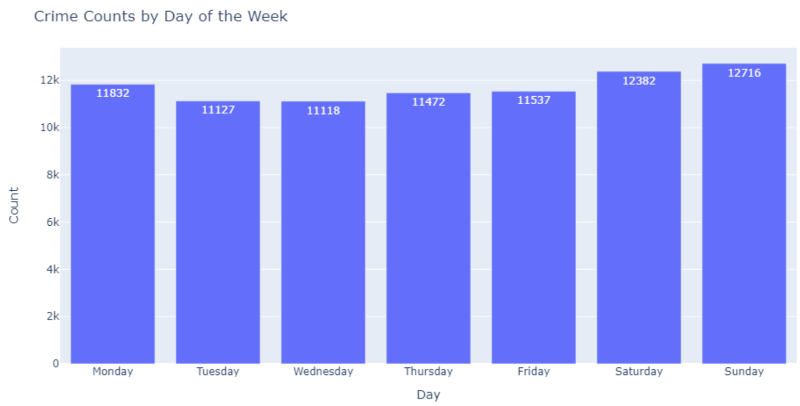

Figure 3 – shows the total violent crimes by day of the week.

Crime Counts by Month of the Year

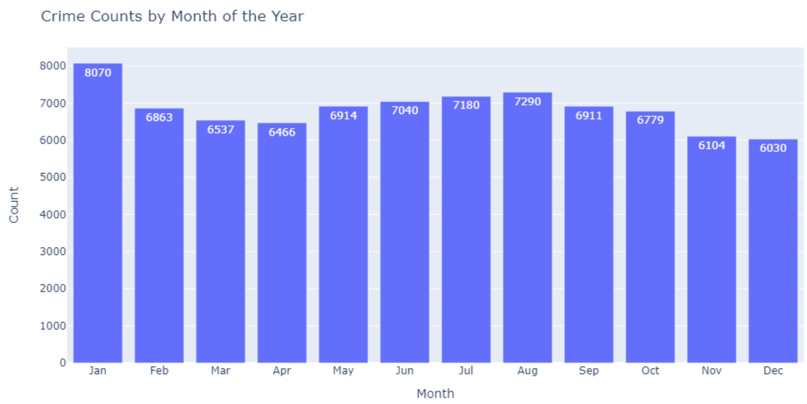

Figure 4 – shows the total violent crimes by month of the year.

While we saw clear trends in the aggregated violent crime rates, the question remained if this would translate to the spatial distribution as well. Figure 5 shows the same totals found in Figure 3 distributed across the 36 by 36 grid.

Violent Crime Heat Map by Day of the Week

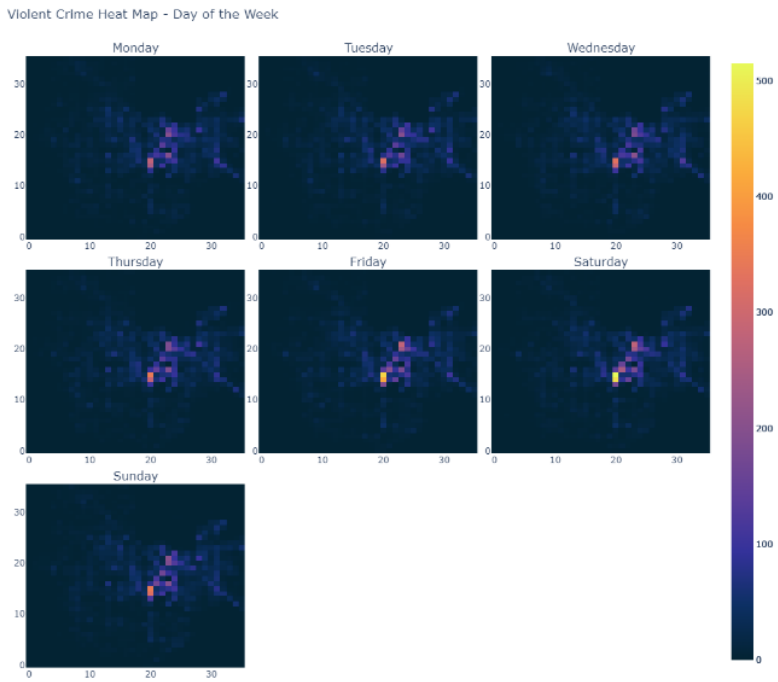

Figure 5 – shows the total violent crimes by day of the week distributed across the 36 by 36 latitude and longitude grid.

Now that we can see where crimes are occurring, Fridays and Saturdays appear to have more localized crime in two particular regions. When overlaying these grids onto the map, we can see that the lower and upper hot spots correlate to The Strip and Downtown, respectively. This was a pattern that we saw repeated when including spatial data. These two regions were so densely packed with crime that many of the other regions were washed out by comparison.

In Figure 6, we take the same approach, but for the month of the year. Again, we see that January produces significant spikes around The Strip and Downtown. This was followed by June and July showing the next highest peaks in the same regions. When looking at it from the opposite perspective, April appears to be the safest time to visit.

Violent Crime Heat Map by Month of the Year

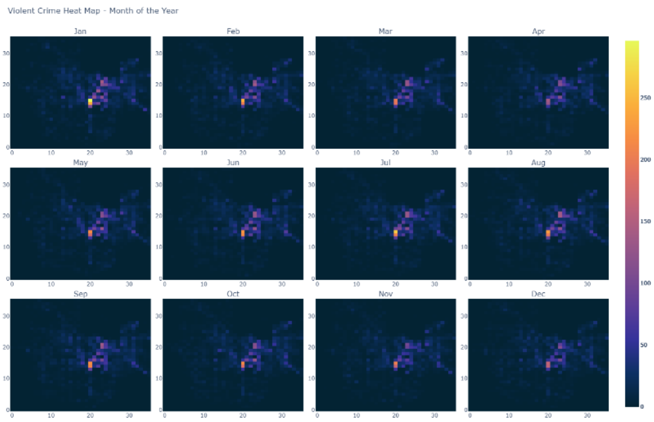

Figure 6 – shows the total violent crimes by month of the year distributed across the 36 by 36 latitude and longitude grid.

Translating the collective data into a density plot on the map shows (Figure 7) these same two regions as hot spots. However, this time, we can see a bit more detail now that the zero values are removed. Another, smaller hotspot, just east of The Strip is coming through. Additionally, we can see outlining neighborhoods with pockets of violent crime.

Las Vegas Heat Map (2019-2024) – Raw

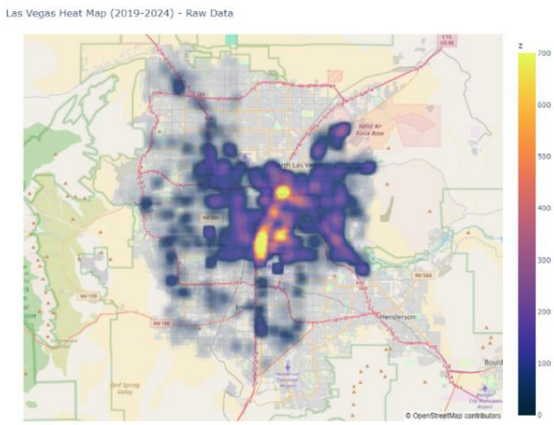

Figure 7 – shows the density plot over the map of Las Vegas for all violent crimes between 2019-2024.

4.2 Model Output and Results

Daily total Crime

Our initial LSTM model had an RMSE of 8.621 when predicting the last 30 days of known data in our dataset. When we look at the actual verse predicted values in Figure 8, we can see that the relative magnitude is similar, and it captures the overall trend within the 30 day period. We can see that the model is clearly not a perfect representation of the data. There are peaks and valleys that aren't properly represented. That being said, the model does appear to do a good job of following historical trends and even manages to predict nearly identical values in the early teens of February.

Total Crime Predictions – Last 30 Day Horizon

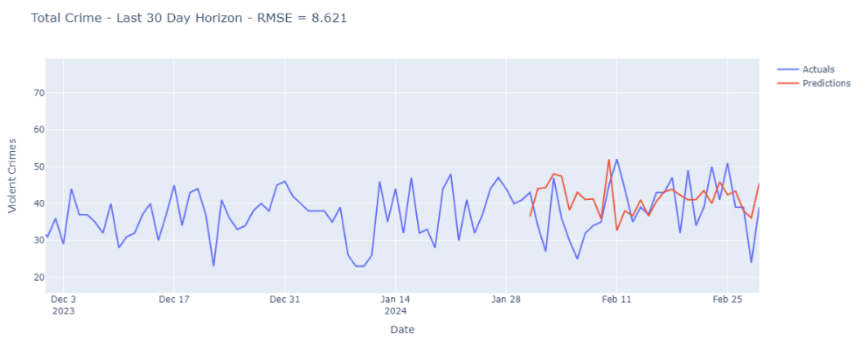

Figure 8 – shows the actual verse predicted values for the trailing 60 day plus 30 day horizon.

Using this same model to predict the next 30 days of violent crime values, the model produces values that appear consistent with long term trend seen in Figure 9. Furthermore, when zooming in (Figure 10), we see that the model predictions still appear to be visually similar to the trailing 60 days of values. With a respectable RMSE against the test set and visually sensible predictions, we saw fit to move forward with these values in our ensemble.

Total Crime – Future Prediction – 30 Day Horizon – Zoomed Out

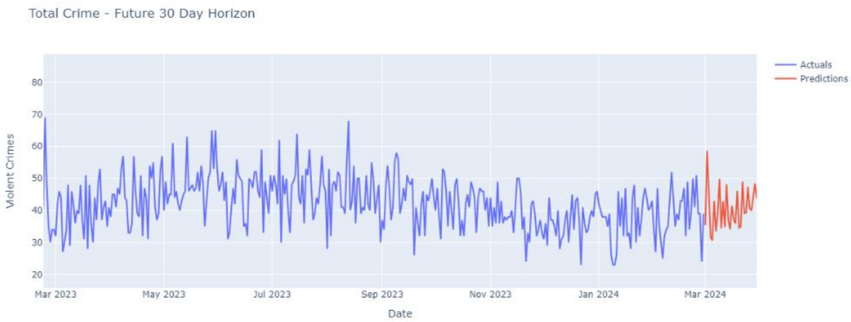

Figure 9 – shows the total violent crime predictions by day for the next 30 days.

Total Crime – Future Prediction – 30 Day Horizon – Zoomed In

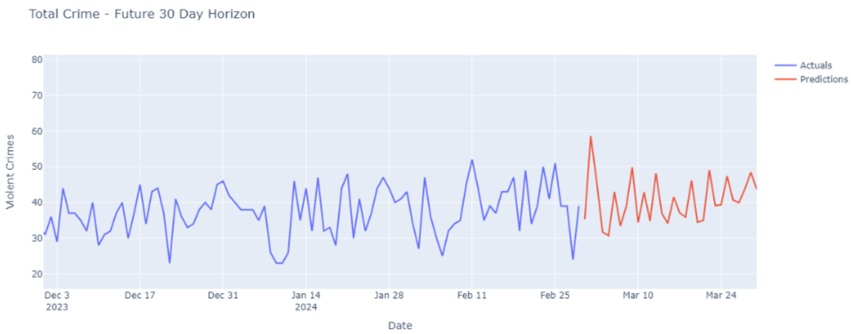

Figure 10 – shows the zoomed in view with trailing 60 plus 30 day horizon future predictions.

To further validate the quality of our model's predictive ability, we compared the underlying frequency of the original dataset to our predicted dataset. In order to have a dataset of sufficient size, we created a new 120 day horizon forecast. Our expectation is that both datasets will display similar underlying frequency plots. Figure 11 shows the Parzen Frequency Window for each dataset. In it, we can see strikingly similar peaks around 0.14 and 0.28. These represent a repeating underlying frequency of about 7 days (1/7 = 0.1429). Note that the predicted dataset fails to pick up the third peak. This represents a subtle reduction in the frequency quality the further our model gets from the last known data point.

Total Crime – Future Prediction – 30 Day Horizon – Zoomed In

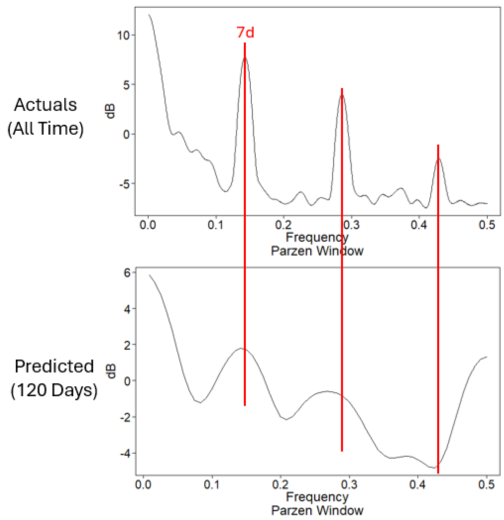

Figure 11 – shows the original Parzen Frequency window stacked on top of the 120 day prediction frequency window.

Next, we evaluated the same 120 day forecasted horizon dataset against the yearly trend we saw in the original data. In Figure 12, we overlayed the exact same parabolic shape next to each other. This indicates that our predictions are not only picking up the weekly frequency, but the yearly frequency found in the original dataset as well.

Total Crime – Future Prediction – 120 Day Horizon with Yearly Seasonal Trend Overlayed

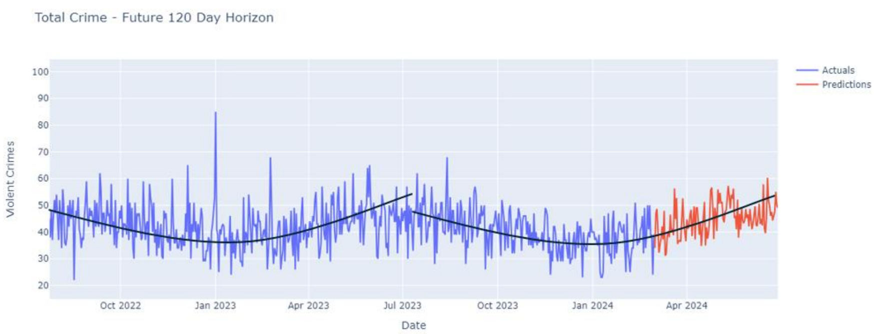

Figure 12 – shows the yearly seasonal trend in the original data as well as the predicted values following the same trend.

Geospatial Temporal Model

Figure 10 By passing the data through two CNN layers and one LSTM layer, we were able to achieve an average MSE score of 0.0009 across our 120 36 by 36 validation grids. Converting these predictions into a probability distribution produced the below heat map.

Probability Distribution Prediction – Horizon Day 1

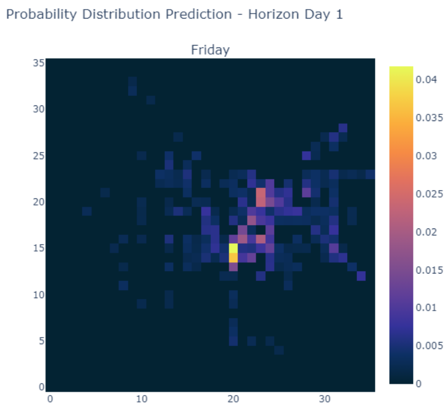

Figure 13 – shows the probability distribution of violent crimes for the first day of our 30 day horizon.

While this distribution does slightly change from day to day. It, correctly, shows that the likelihood of violent crimes is in The Strip and Downtown. In fact, when we compare this distribution to the overall heat maps, the resemblance is striking. To better capture the overall numbers, we multiply this probability distribution by the corresponding day value in our total crime LSTM model. That produced the 30 day aggregated view density plot for Las Vegas.

Violent Crime Predictions – 30 Day Aggregate Horizon

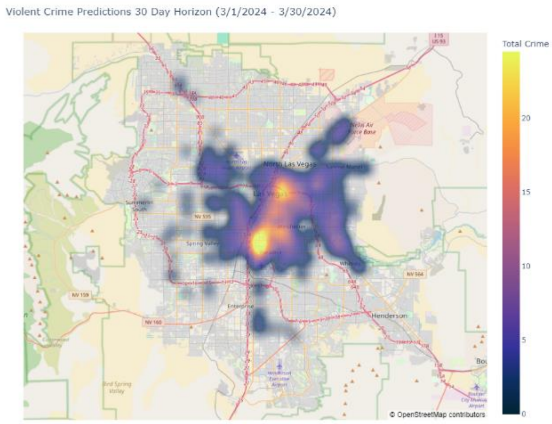

Figure 14 – shows the density plot for the next 30 days of aggregated violent crimes across the 36 by 36 grid.

For practical use, this data can also be viewed at a daily level. However, given the overwhelming influence of the hotspots shown in Figure 13 and Figure 7, the plots look very similar. It is, therefore, necessary to look at the scale to really understand minute differences in the predicted values.

5 Discussion

Both our overall crime prediction LSTM model and our geospatial temporal CNN/LSTM models successfully produce results with reasonable performance metrics and results in alignment with the findings of our exploratory data analysis. Now, each of these models should be circulated with police and community leaders for examination and further context. In the case of the LVMPD, these models should be evaluated for use within the current hot spot policing planning. For community leaders, we believe mapping points of interest like schools, churches, liquor stores, and clubs can help current and future Cardiff members to better build strategies around how to reduce violent crimes in their community.

Future research is ripe to incorporate external datapoints like temperature and events to see if they can further improve each model's predictive quality. Furthermore, these heat maps can build on the previous work of the previous Cardiff research group (Grard et al., 2023) to address the other half of this challenge (overdoses). With future iterations of this model, we believe it is possible to integrate both neural networks into a single geospatial temporal model without the need to ensemble the total crime values by day. We believe that this will require significant computing resources and time, but the data appears to show evidence that this is possible with much more tuning.

Given the nature of the data and the desired use, it is essential that we highlight some ethical concerns regarding the use of these models. We must reinforce to all parties who seek to utilize and build upon these model that they are not perfect. Any predictions must be reviewed by knowledgeable and informed consultants who use them as one of many sources of guidance. For both police and community leaders, it is to serve as a better understanding of where and when violent crimes are likely to occur. By making the methods and model design publicly available, all parties can engage in discussions that lead to greater transparency and collectively move the community towards a less violent future.

6 Conclusion

This research underscores the effective application of hybrid analytical models in understanding and predicting crime trends. The combination of CNN and LSTM models has proven effective in outlining crime patterns in both space and time. We believe this is an entirely new approach that has the potential to revolutionize the current approach of developing heat maps to support hot-spot policing. With the bare bones' requirements of the data set, it is our hope that its use case will have broad applicability and impact. Furthermore, we are interested to see if the predictive quality of the CNN/LSTM model will improve in cities with hot spots that are more distributed across the city and time.

Moving forward, it is our hope that future groups can expand on the findings of this model to incorporate hospital data to fill in the remaining half of unreported crimes (Kollar et al., 2018). Even within the current data set, there is a great opportunity to zoom the model into hourly increments. This can provide even further granularity and information for Cardiff leaders to support the city of Las Vegas. Even considering the limitations of our model, we are excited for the opportunities that it presents for the field of criminal science and current/future Cardiff models in the United States and beyond.

Acknowledgments

We sincerest gratitude to several individuals and organizations whose contributions have been invaluable to the success of this project. First and foremost, we thank our advisor and Cardiff leader, Chris Papesh. Without his vision and desire to bring the Cardiff Model to the city of Las Vegas, we would not have had such a great opportunity to apply our knowledge and efforts to such a worthy cause. Beyond that, countless potential future victims now have a better chance for their lives to be redirected towards a better outcome. Next, we want to thank the Las Vegas Metropolitan Police Department for providing access to detailed crime data, which formed the backbone of our analysis. Their choice to make this data publicly available helps fuel community engagement and research like ours.

Lastly, we are grateful to our families and friends for their patience and encouragement throughout the duration of this project. Their sacrifice and support have been an essential component to all that we were able to achieve during our research.

References

1. AVERDIJK, M., VAN GELDER, J.-L., EISNER, M., & RIBEAUD, D. (2016). VIOLENCE BEGETS VIOLENCE ... BUT HOW? A DECISION-MAKING PERSPECTIVE ON THE VICTIM-OFFENDER OVERLAP. *Criminology (Beverly Hills)*, *54*(2), 282–306. https://doi.org/10.1111/1745-9125.12102
2. Bock, J. G. (2009). The efficacy of violence mitigation: A second look using time-series analysis. *Political Geography, 28*(3), 266-270. https://doi.org/10.1016/j.polgeo.2009.05.005
3. Boyle, A. A., Snelling, K., White, L., Ariel, B., & Ashelford, L. (2013). External validation of the Cardiff model of information sharing to reduce community violence: natural experiment. *Emergency Medicine Journal : EMJ*, *30*(12), 1020–1023. https://doi.org/10.1136/emermed-2012-201898
4. Bohnert, A. S. B., Nandi, A., Tracy, M., Cerda, M., Tardiff, K. J., Vlahov, D., & Galea, S. (2011). Policing and risk of overdose mortality in urban neighborhoods. Drug and Alcohol Dependence, 113(1), 62-68. https://doi.org/10.1016/j.drugalcdep.2010.07.008
5. Braga, A. A., & Weisburd, D. L. (2022). Does Hot Spots Policing Have Meaningful Impacts on Crime? Findings from An Alternative Approach to Estimating Effect Sizes from Place-Based Program Evaluations. *Journal of Quantitative Criminology*, *38*(1), 1–22. https://doi.org/10.1007/s10940-020-09481-7
6. Bye, E. K. (2007). Alcohol and violence: use of possible confounders in a time-series analysis. Addiction (Abingdon, England), 102(3), 369–376. https://doi.org/10.1111/j.1360-0443.2006.01701.x
7. Cardiff Model data collection procedures in a level I trauma adult emergency department. *BMJ Open*, *12*(1), e052344–e052344. https://doi.org/10.1136/bmjopen-2021-052344
8. Cesario, E., Catlett, C., & Talia, D. (2016). Forecasting Crimes Using Autoregressive Models. *The Institute of Electrical and Electronics Engineers, Inc. (IEEE) Conference Proceedings*, 795-.
9. Corsaro, N., Engel, R. S., Herold, T. D., Yildirim, M., & Motz, R. T. (2023). Hot spots policing in Las Vegas: results from a blocked randomized controlled trial in chronic violent crime locations. *Journal of Experimental Criminology*, *19*(1), 213–235. https://doi.org/10.1007/s11292-021-09485-8
10. Crime in the United States. (2019). ucr.fbi.gov. https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/tables/table-6

11. Dakalbab, F., Abu Talib, M., Abu Waraga, O., Bou Nassif, A., Abbas, S., & Nasir, Q. (2022). Artificial intelligence & crime prediction: A systematic literature review. *Social Sciences & Humanities Open*, *6*(1), 100342-. https://doi.org/10.1016/j.ssaho.2022.100342
12. Drug overdose mortality by state. (n.d.). https://www.cdc.gov/nchs/pressroom/sosmap/drug_poisoning_mortality/drug_poisoning.htm
13. Findley, M., & Brown, R. (2018). Fifty states of self-control: A U.S. statewide examination of the initiation and inhibition dimensions of self-regulation. The Journal of Social Psychology, 158(1), 23–36. https://doi.org/10.1080/00224545.2017.1297287
14. Gottfredson, M. R., & Hirschi, Travis. (1990). A general theory of crime. Stanford University Press.
15. Girard, J., Pandey, S., Bunn, Z., Papesh, C., Cheun, J., & Zhang, Y. (2023). *Identifying Locations of Drug Overdose in Las Vegas to Implement the Cardiff Violence Prevention Model*. SMU Scholar.
16. Hu, K., Klinkenberg, B., Gan, W. Q., & Slaunwhite, A. K. (2022). Spatial-temporal trends in the risk of illicit drug toxicity death in British Columbia. BMC Public Health, 22(2121). https://doi.org/10.1186/s12889-022-14586-8
17. Jagait, R. K., Fekri, M. N., Grolinger, K., & Mir, S. (2021). Load Forecasting Under Concept Drift: Online Ensemble Learning With Recurrent Neural Network and ARIMA. *IEEE Access*, *9*, 98992–99008. https://doi.org/10.1109/ACCESS.2021.3095420
18. Liang, W., Wang, Y., Tao, H., & Cao, J. (2022). Towards hour-level crime prediction: A neural attentive framework with spatial–temporal-categorical fusion. *Neurocomputing (Amsterdam)*, *486*, 286–297. https://doi.org/10.1016/j.neucom.2021.11.052
19. Mercer Kollar, L., PhD, A. Summer, S., MD, MSc, F. Jacoby, S., PhD, MPH, RN, & Ridgeway, G., PhD. (2018). Cardiff Model Toolkit. www.cdc.gov. https://www.cdc.gov/violenceprevention/pdf/cardiffmodel/cardiff-toolkit508.pdf
20. Norström, T., & Pape, H. (2010). Alcohol, suppressed anger and violence. Addiction (Abingdon, England), 105(9), 1580–1586. https://doi.org/10.1111/j.1360-0443.2010.02997.x
21. Opioid use disorder. (n.d.). https://www.psychiatry.org/patients-families/opioid-use-disorder#:~:text=Access%20to%20prescription%20opioids%20and,develop%20an%20addiction%20to%20them.
22. Pierce, M., Hayhurst, K., Bird, S. M., Hickman, M., Seddon, T., Dunn, G., & Millar, T. (2017). Insights into the link between drug use and criminality: Lifetime offending of criminally-active opiate users. Drug and Alcohol Dependence, 179, 309–316. https://doi.org/10.1016/j.drugalcdep.2017.07.024
23. Prathap, B. R. (2023). Geo-spatial crime density attribution using optimized machine learning algorithms. *International Journal of Information Technology (Singapore. Online)*, *15*(2), 1167–1178. https://doi.org/10.1007/s41870-023-01160-7

24. Prieto Curiel, R. (2023). Weekly Crime Concentration. *Journal of Quantitative Criminology*, *39*(1), 97–124. https://doi.org/10.1007/s10940-021-09533-6
25. SAMHSA. (2012). Substance Use and Mental Disorders in the Las Vegas-Paradise MSA. In Metro Brief [Report]. https://www.samhsa.gov/data/sites/default/files/NSDUHMetroBriefReports/NSDUHMetroBriefReports/NSDUH-Metro-Las-Vegas.pdf
26. Sixteen US cities in National Cardiff Violence Prevention Network - News - Cardiff University
27. Tanz, L. J., Gladden, R. M., Dinwiddie, A. T., Miller, K. D., Broz, D., Spector, E., & O'Donnell, J. (2024b). Routes of drug use among drug overdose deaths — United States, 2020–2022. Morbidity and Mortality Weekly Report, 73(6), 124–130. https://doi.org/10.15585/mmwr.mm7306a2
28. Task Force, C. C. B., & Gemignani, G. (n.d.). History of Nevada's gaming regulatory structure. https://tax.nv.gov/uploadedFiles/taxnvgov/Content/FAQs/NV-Gaming-History.pdf
29. *The Cardiff Model for Violence Prevention - Cardiff University*. (2022). Cardiff University. https://www.cardiff.ac.uk/documents/2665796-the-cardiff-model-for-violence-prevention
30. Towers, S., Chen, S., Malik, A., & Ebert, D. (2018). Factors influencing temporal patterns in crime in a large American city: A predictive analytics perspective. *PloS One*, *13*(10), e0205151–e0205151. https://doi.org/10.1371/journal.pone.0205151
31. Turanovic, J. J., & Pratt, T. C. (2013). The Consequences of Maladaptive Coping: Integrating General Strain and Self-Control Theories to Specify a Causal Pathway Between Victimization and Offending. *Journal of Quantitative Criminology*, *29*(3), 321–345. https://doi.org/10.1007/s10940-012-9180-z