

***PE7046***

***MSc Research Project***

**Student Name:** Bethany Welsh

**Supervisor Name:** Saqib Hussain

**Second Marker Name:** Maria Salama

**A deep learning approach to predicting domestic violence instances in the Northeast of England during football events.**

**2024/2025**

# Declaration

I declare the following:

(1) that the material contained in this dissertation is the end result of my own work and that due acknowledgement has been given in the bibliography and references to ALL sources be they printed, electronic or personal.

(2) The Word Count of this Dissertation is 10,959

(3) that unless this dissertation has been confirmed as confidential, I agree to an entire electronic copy or sections of the dissertation to being placed on the eLearning Portal (Blackboard), if deemed appropriate, to allow future students the opportunity to see examples of past dissertations. I understand that if displayed on the eLearning Portal it would be made available for no longer than five years and that students would be able to print off copies or download.

(4) I agree to my dissertation being submitted to a plagiarism detection service, where it will be stored in a database and compared against work submitted from this or any other Department or from other institutions using the service.

In the event of the service detecting a high degree of similarity between content within the service this will be reported back to my supervisor and second marker, who may decide to undertake further investigation that may ultimately lead to disciplinary actions, should instances of plagiarism be detected.

(5) I have read the Northumbria University Policy Statement on Ethics in Research and Consultancy, and I confirm that ethical issues have been considered, evaluated and appropriately addressed in this research.

**SIGNED: B L WELSH**

**DATE:** 07/04/2025

# Abstract

In recent years the use of predictive models in policing has increased, however, an accurate model to predict domestic violence has not been produced. Using data provided by Northumbria Police, the project aimed to develop a predictive model to forecast domestic violence instances that are associated with football matches in the Northeast of England. Research that has been previously conducted into this subject suggested that there was a link between football matches and increased levels of domestic violence, however, predictive models remained underdeveloped in this area, with limited success. Three deep learning models were utilized in the research - Multi-Layered Perceptron(MLP), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) - to predict the likelihood of domestic violence occurring during football matches in the North East Demonstrating the highest level of accuracy using the data collected, was the RNN model, with a Mean Absolute Error (MAE) of 0.2235, performing better than the MLP and CNN models which had MAEs of 51.68 and 0.9856 respectively. Although the RNN model provided accurate results, further research is needed into this topic, using alternative algorithms such as XGBoost or Long Short-Term Memory (LSTM) to aid in a more reliable prediction. As well as this, more data is required from minor leagues to improve upon accuracy as it would allow more teams in the Northeast to be checked for a correlation. Data could be collected via team instead of league.

**Table of Contents**

[**Declaration 1**](#_gvmeb0i1kpzv)

[**Abstract 2**](#_18yx8thzpvla)

[**1.0 Introduction 5**](#_vbuw47dma1vp)

[1.2 Definitions of Domestic Abuse 6](#_qj3nvxte8orw)

[1.3 Background 7](#_mljxpeub2mn2)

[**2.0 Literature Review 9**](#_bl35h97l4ar8)

[2.1 Predictive Policing 10](#_lsbwn1m2qx5s)

[2.1.1 Predicting Policing, Domestic Violence, and Sport 11](#_vuu417q1o056)

[2.2 Algorithms Used in Predictive Policing 13](#_eeqtjm8w2c7w)

[2.2.1 Multi-Layer Perceptron 13](#_hj71dhcspi3s)

[2.2.2 Convolutional Neural Networks 15](#_am9exkfuiegl)

[2.2.3 Recurrent Neural Network 16](#_yk9tabzv28z)

[2.3 Bias In Predictive Policing 18](#_djjnyalyxdli)

[2.4 Challenges in Implementing Deep Learning for Domestic Violence Prediction 20](#_6c8g62qzremh)

[**3.0 Project Rationale 22**](#_3eigm52ha84z)

[**4.0 Research Approach 24**](#_bsp1j19lptb5)

[**5.0 Research Aim and Objectives 25**](#_da9trcm64g9t)

[5.1 Research Aim 25](#_js8y7yikxxqv)

[5.2 Research Objectives 25](#_6pne4vfb6wyx)

[**6.0 Methodology 27**](#_esjez1oz69ei)

[6.1 Ethical Implications 28](#_1ol9heto8tyl)

[6.1.1 Data Storage 28](#_rkcfj2exfb9v)

[6.1.2 False Positives, False Negatives, and Model Accuracy 29](#_5wpjgn6ks12)

[**7.0 Results 30**](#_mpwkix3hy7os)

[7.1 Data Visualisation 30](#_55cbqlwdg7u9)

[7.1.1 Domestic Violence Instances 30](#_zg4apklzq85k)

[7.1.2 Newcastle Domestic Violence and Football Matches 31](#_mn3eyv8wrdox)

[7.1.3 Sunderland Domestic Violence and Football Matches 33](#_p9k8iurovfbs)

[7.2 Models 35](#_izofwi7qlj6h)

[7.2.1 Multi-Layer Perceptron 35](#_16gljqvlafhk)

[7.2.2 Convolutional Neural Network 37](#_v5arfr3ksz29)

[7.2.3 Recurrent Neural Network 38](#_ygom9aurrdo1)

[**8.0 Application 40**](#_pkrul7h8n654)

[**9.0 Analysis and Discussion 42**](#_jplpkypbm9gd)

[9.1 Data Visualisation and Pearson Correlation Coefficient 42](#_olxs0pz6hx86)

[9.2 Models and Accuracy 44](#_f3bmvwvciemx)

[**10.0 Conclusion 47**](#_fiop9gx5vw7h)

[**References 50**](#_j15cwguoyjbh)

[**Appendix A. Research Proposal 60**](#_6gbtcvo8aqv9)

[**Appendix B. Gantt Chart 76**](#_essqy2ujupr9)

[**Appendix C. Code 77**](#_p9gy8aieslns)

# 

# 1.0 Introduction

Multiple attempts have been made to create an accurate model that predicts crime using artificial intelligence (Gerber, 2013; Towers et al., 2018; Karystiantis et al., 2021; Thornton, 2017). The growth of big data in recent years has allowed predictive models to be used in policing. Approaches of predictive analysis surrounding crime often integrate social media posts (Gerber, 2013) or police records (Towers et al., 2018) into their dataset, in order to forecast criminal activity or hot spots. Accuracy of these models ranges from 70-81%. These models tend to focus on crime in general, however, it is possible to apply these machine learning techniques to predict domestic violence instances, similarly to how the police have used predictive policing to intercept illicit drugs and weapons (Karystiantis et al., 2021). Recent efforts to predict domestic violence have had limited success, with some models producing weak predictive accuracy (Karystiantis et al., 2021). In one case, 89% of fatal cases were incorrectly labeled as not a high risk (Thornton, 2017).

Violence and sport go hand-in-hand, with player violence or crowd violence being widely acknowledged and focused upon (Neville, 2015). However, there have been links made to sporting culture and domestic violence since the 1980’s (Sabo and Runfola, 1980). Evidence points to an increased rate of domestic violence perpetrated by male athletes, as well as some evidence confirming that domestic violence increases throughout wider society during sporting events (Neville, 2015). Neville (2015) notes that the ‘holy trinity’ of sport, alcohol, and hegemonic masculinity are often used to explain the increase in domestic violence where sporting events are concerned. Kirby, Francis, and O’Flaherty (2013) used Poisson and negative binomial regression models to investigate if empirical evidence exists to support the view that football (soccer) tournaments are associated with a rise in reported domestic abuse incidents. To do this, Kirby, Francis and O’Flaherty (2013) focused upon the FIFA World Cup Tournaments between 2002 and 2010, and domestic abuse incidents reported to a police force in the Northwest of England. The study found a 26% rise in instances of domestic violence when the England national team won or drew on match day, however, it increased to 38% when England lost (Kirby, Francis and O'Flaherty, 2013). Kirby, Francis and O’Flaherty (2013) focused upon the FIFA World Cup Tournaments between 2002 and 2010, and domestic abuse incidents reported to a police force in the Northwest of England. However, this research encompasses only three tournaments - due to the FIFA World Cup being held every four years. More research into this area would be beneficial to confirm a correlation, especially with other leagues and cups in the sport. This project aims to delve into popular football events in the North-East, focusing on Northumbria Police statistics, thus, confirming if there is a correlation between football and domestic violence in the region.

A predictive model of domestic violence instances by location in the Northeast, in relation to football matches through historical police records marked as domestic violence instances and timelines of football events, will be created. Datasets will be combined to create a model for domestic violence prediction that encompasses sporting events. This research will compare three different deep learning and artificial intelligence methods: MLPs, CNNs, and RNNs to uncover which algorithm is best suited for the task.

## 1.2 Definitions of Domestic Abuse

Domestic abuse, also known as intimate partner abuse, is defined as the following in UK law:

“*2. Behaviour of a person (“A”) towards another person (“B”) is “domestic abuse” if—*

1. *A and B are each aged 16 or over and are personally connected to each other, and*
2. *the behaviour is abusive.*

*3. Behaviour is “abusive” if it consists of any of the following—*

1. *physical or sexual abuse;*
2. *violent or threatening behaviour;*
3. *controlling or coercive behaviour;*
4. *economic abuse (see subsection (4));*
5. *psychological, emotional or other abuse;*

*and it does not matter whether the behaviour consists of a single incident or a course of conduct.*”

(Domestic Abuse Act, 2021)

## 1.3 Background

Intimate partner violence (or domestic violence) is a “global epidemic” (Forsdike, O’Sullivan, and Hooker, 2022) with the most recent Office for National Statistics figures (2023) showing that one in four women and one in seven men will be a victim of domestic abuse in their lifetime. The World Health Organisation (WHO) defines domestic violence as “behaviour by an intimate partner or ex-partner that causes physical, sexual or psychological harm, including physical aggression, sexual coercion, psychological abuse and controlling behaviours” (WHO, 2021).

Despite the likelihood of domestic violence occurring within society, tools that can predict future instances are lacking (Karystianis et al., 2021). Through the growth and availability of big data as well as the creation of predictive policing throughout recent years, it is possible to apply machine learning techniques to predict domestic violence instances, similarly to how the police use predictive policing to intercept illicit drugs and weapons (Karystiantis et al., 2021).

Approaches of predictive analysis surrounding crime often apply machine learning to Tweets (Gerber, 2013) or police records (Towers et al., 2018) in order to forecast criminal activity or hot spots, with the accuracy of these models ranging from 70-81%. However, recent efforts regarding the prediction of domestic violence have had limited success, with some models producing weak predictive accuracy (Karystiantis et al., 2021). In one case, 89% of fatal cases in the Thames Valley Police area were incorrectly labelled as not a high risk (Thornton, 2017). Ringland (2018) attempted to predict if there would be repeat victimisation of survivors of domestic abuse by asking binary questions. If a survivor answered ‘yes’ to 12 or more questions, this was an indicator that repeat victimisation would occur, however, the predictive results were poor as this method was a modified risk assessment (Ringland, 2018). Several papers reviewed revealed that current predictive methods for identifying domestic violence risk before it occurs does not display a usable accuracy for policing methods - with Thornton’s (2017) model from Thames Valley Police producing a false positive rate of up to 99%. There are several reasons as to why these methods lacked accuracy and according to Swallow (2017), this is due to those who experience intimate partner violence may have difficulty in defining what constitutes domestic violence, therefore being unable to answer a risk assessment in a factual way. In predictive policing, there is the application of predictive models though it has not yet been utilised for the prediction of domestic violence such as the models produced by Gerber (2013), Towers et al. (2018), and Karystiantis et al. (2021).

At present, there is a potential gap in research that has been identified - the creation of a predictive model to aid police in pursuing perpetrators of domestic violence. This project will focus on the prediction of domestic violence occurring during notable football matches due to the significant research that instances of reporting increase during these events. The project will be limited to the Northeast of England, with a focus on Northumbria Police data.

# 2.0 Literature Review

Algorithms are used for prediction in everyday life. One example is the use of predictive algorithms in healthcare to aid in diagnosis. Machine learning techniques can aid doctors and nurses to interpret medical images, such as CT scans, mammography, and MRI images, accurately and efficiently (Rana and Bhushan, 2023). Lopez-Ubeda et al. (2021) used Natural Language Processing (NLP) and Electronic Health Records (EHR) to train an SVM model that aids doctors in diagnosis of cancers. Overall, the model had a 92.2% accuracy in classifying CT scans and an 86.9% accuracy in classifying the MRI scans. Machine learning is one of the more promising computer-based approaches for predicting mental health issues (Chung and Teo, 2023). Chung and Teo (2023) used varying machine learning techniques to predict mental health issues and found that Neural Networks showed the highest accuracy at 78%.

Other areas in which predictive algorithms are used are the weather, forecasting of stock markets and shares, and in policing. Using machine learning to predict the weather has been essential to early warning systems that are built to protect the lives and livelihood of people from extreme weather, but it has been less useful in providing a more accurate forecasting model (AbdulRaheem et al., 2022). In their research, AbdulRaheem et al. (2022) used decision trees and K-nearest neighbours to predict the weather forecast, with the former achieving 100% accuracy and the latter having its highest accuracy at 78%. An example of machine learning being used to predict stocks and shares would be Usmani et al. (2016) who used Support Vector Machines, Radial Basis Function and Layer Perceptrons, both single and multi-layer, to predict the market performance of the Karachi Stock Exchange (KSE) at the end of each day. Usmani et al. (2016) found that it was possible to predict the performance of the KSE-100 using machine learning with the Multi-Layer Perceptron having the highest accuracy of all algorithms tested at 77%.

Policing has also seen a rise in the use of predictive methods and algorithms to aid law enforcement with preventing crime. Although accuracy is significantly lower within predictive policing on average, research in this area is still emerging and has potential for improvement.

## 2.1 Predictive Policing

Meijer and Wessels (2019) note that over the past few years, an increasing number of police forces globally have adopted software that guides either decision making via statistical data. This has come to be known as ’predictive policing’ (Meijer and Wessels, 2019).

Predictive policing can be defined as the following:

“the use of historical data to create a forecast of areas of criminality or crime hotspots, or high-risk offender characteristic profiles that will be one component of police resource allocation decisions. The resources will be allocated with the expectation that, with targeted deployment, criminal activity can be prevented, reduced, or disrupted.”

(Ratcliffe, 2019: 349)

Hamilton (2021) notes that there are three main types of predictive policing: forecasting hotspots; predicting who is likely to become a victim; and predicting who is likely to offend. The most common is forecasting hotspots - this is due to the latter two types being less researched due to the practicality of investigation (Hamilton, 2021) as well as issues with bias within predictive policing, which will be discussed in section 2.3.

By finding hotspots of crime, police departments can analyse historical statistical data to predict in which geographic areas instances of criminal activity may occur (Meijer and Wessels, 2019). In Richmond, Virginia, predictive policing was used to forecast where gun crime would occur on New Year’s Eve 2006 (Pearsall, 2010). Surveillance routes were adapted in line with these predictions, and it was considered a success, with gun crimes decreasing 49% and 246% more weapons seized (Pearsall, 2010). Only one-third as many police officers were deployed compared to previous years, however, effectiveness and efficiency increased regardless of the decrease in police presence due to predictive policing, saving the police around $15,000 in overtime pay (Harris, 2010).

Analysing data and forecasting crime can be achieved in a multitude of modalities - one of these is via machine learning. Machine learning (ML) is the term used for algorithms that produce outcomes that are based on patterns of data (Alikhademi et al., 2021) - usually from intelligence gained from ‘big data’ (Hardyns and Rummens, 2018). Eulluri, Madalapu and Roy (2019) used ML - both traditional algorithms and deep learning - to predict the effect of weather on the likelihood of crimes committed in New York City. Both traditional algorithms and deep learning algorithms had a high correlation in relation to weather and crime (Eulluri, Madalapu, and Roy, 2019). Khan et al. (2019) also used ML to predict crimes, however, their research was focused in metropolitan cities using the Karachi region in Pakistan as their dataset. Khan et al.’s (2019) research used R and WEKA to predict the likelihood of street crime such as phone snatching and theft, pulling together large amounts of data and producing a predictive model with an accuracy of around 70%. This model used clustering and classification machine learning algorithms, such as K-means and Naive Bayseian, to predict the crime rate for the following week (Khan et al., 2019). In their model, Khan et al. (2019) used 80% of their collected data for training, saving 20% for testing purposes, which they found to be best for accurate results. Another example of ML being utilised for predictive policing is Williams (2018) research into the effectiveness of predictive policing in the case of re-offending in the Dyfed-Powys Police force. The model created by Williams (2018) used Random Forest and XGBoost algorithms, as well as Feedforward Neural networks to predict the likelihood of an offender re-offending chance. Williams (2018) found that there was a clear case in favour of predictive policing being used to prevent crime and recidivism rate.

## 2.1.1 Predicting Policing, Domestic Violence, and Sport

The majority of predictive methods where domestic violence is concerned, are focused upon using risk assessments to predict the likelihood of instances. There have been recent efforts regarding the prediction of domestic violence for the police, however, these have had limited success, with some models producing weak predictive accuracy (Karystiantis et al., 2021). Currently there is one tool used in the United Kingdom to predict domestic violence (Hamilton, 2021). This is called the Domestic Abuse, Stalking, Harassment and Honour-Based Violence (DASH) instrument and it ranks risk based on predictive items regarding current situations, dependents, domestic violence history, and certain characteristics of the abusers (Hamilton, 2021). DASH is a risk assessment based tool that was created by Laura Richards in 2009 and is widely used in police forces across the United Kingdom. The domestic violence section of the instrument is 27 questions and places the victim in a standard, medium or high-risk category depending upon their answers (Richards, 2009). However, Thornton (2017) investigated the use of this method in Thames Valley and found that 89% of fatal cases were incorrectly labelled as medium risk when they were, in fact, high risk. If these had been correctly labelled as high risk, there is a possibility that these cases would not have led to a fatality (Thornton, 2017). Ringland (2018) also aimed to predict if there would be repeat victimisation of survivors of domestic abuse by asking yes or no questions, which provided poor predictive results. Several papers reviewed show that the current predictive methods for identifying domestic violence do not display a usable accuracy for policing methods - with some models producing a false positive rate of up to 99% (Thornton, 2017).

There have been links made to sporting culture and domestic violence since the 1980’s (Sabo and Runfola, 1980). Evidence points to an increased rate of domestic violence perpetrated by male athletes, as well as some evidence confirming that domestic violence increases throughout wider society during sporting events (Neville, 2015).

Kirby, Francis, and O’Flaherty (2013) conducted quantitative analysis, using Poisson and negative binomial regression models, to investigate if evidence exists to support the view that football tournaments are linked with a rise in domestic abuse incidents reported to the police. Kirby, Francis and O’Flaherty (2013) focused upon FIFA World Cup Tournaments between 2002 and 2010 (2002, 2006, 2010), comparing the report rate of incidents to those reported when the tournaments were not occurring. The study found that domestic violence rose by 26% when the England national team won or drew on match day, however, it increased to 38% when England lost (Kirby, Francis and O'Flaherty, 2013). Another trend was apparent that with each new tournament, reports of domestic abuse incidents increased in frequency (Kirby, Francis, and O’Flaherty, 2013). However, this research encompasses only three tournaments - due to the FIFA World Cup being held every four years. More research into this area would be beneficial to confirm a correlation with more contemporary data. This project aims to investigate this further, not only with more recent data, but also with additional sports included to determine if the link to domestic violence is pertinent throughout all types of sport.

## 2.2 Algorithms Used in Predictive Policing

Research into the algorithms used in predictive policing indicate that multi-layered perceptrons, recurrent neural networks and convolutional neural networks are the most commonly used algorithms in predictive policing (Walczak, 2021).

A multi-layer perceptron (MLP) is a supervised learning feed forward network most commonly trained using back propagation (Walczak, 2021). An MLP is a commonly used type of artificial neural network (Popescu et al., 2009).

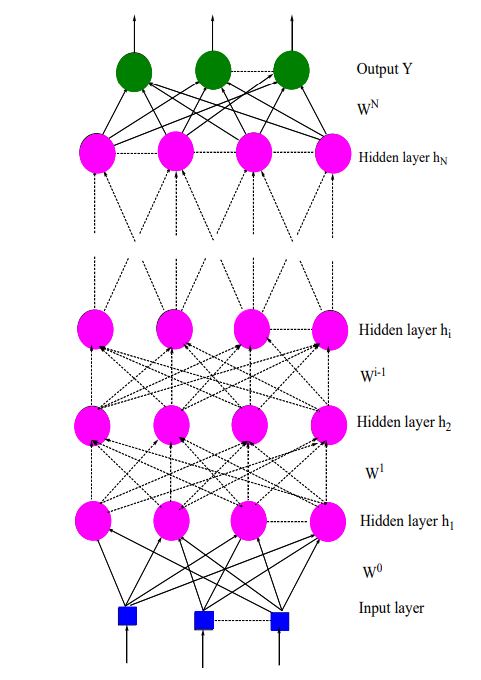
Convolutional Neural Networks (CNNs) are useful in a multitude of different applications including image classification, semantic segmentation and object detection (Wu, 2017). As well as this, a common use for CNNs is in natural language processing (Li et al., 2022). CNNs are a hybrid of unsupervised and supervised learning (Walczak, 2021).

Recurrent Neural Networks (RNN) are “an MLP-like architecture with backwards links to previous layers that enable better time series modeling” (Walczak, 2021). This algorithm is capable of running features and long-term dependencies from sequential and time series data (Salehinejad et al., 2018) and has been an important focus of research since the early 1990’s (Unadkat et al., 2001).

### 2.2.1 Multi-Layer Perceptron

A multi-layer perceptron (MLP) contains multiple layers of neurons. “MLPs are significant in machine learning because they can learn nonlinear relationships and data” which makes MLPs powerful models for classification, regression, and pattern recognition - all of which are important in prediction (Jaiswal, 2024). Delashmit and Manry (2006) note that each MLP must comprise a minimum of three layers, consisting of an input layer, one or more hidden layer(s), and an output layer. The number of neurons in the input layer equal to that of the number of neurons in the output layer, as demonstrated below in *Figure 1* (Ramchoun et al., 2016). The input layer determines the distribution of the inputs into each layer with each hidden node and output having thresholds associated with them (Delashmit and Manry, 2006).

Hirithik et al. (2022) found that MLPs can outperform the existing algorithms for police prediction and the algorithm also reduces the amount of error values. In their research, Hirithik et al. (2022) used MLP back propagation neural networks to predict crimes and to train classification. By reviewing the results, it can be observed that the MLP algorithm is the most accurate classifier for crime occurrence prediction as it contains the least error rate in comparison to other prediction algorithms tested (Hirithik et al., 2022).

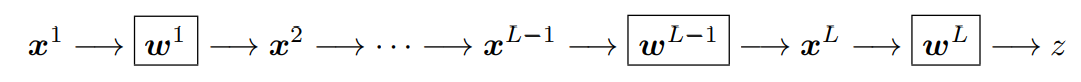


*Figure 1: A diagram of a feed forward neural network structure and multi-layer perceptron. (Ramchoun et al., 2016)*

Whilst MLPs are accurate for crime classification in research conducted by Hirithik et al. (2022), Banerjee (2024) notes that there are a few limitations of using this method. One of these weaknesses is overfitting, where a model is trained on a dataset to the extreme point in which it can only analyse that specific set of data and is prone to error if not using the training data, especially when trained on smaller datasets (Banerjee, 2024). Another problem that arises with MLPs is the issue of gradient vanishing where, when training, as the sequence length increases, the partial derivative of the loss function exponentially decreases, and - in some cases - may completely halt the neural network from further learning (Banerjee, 2024).

### 2.2.2 Convolutional Neural Networks

In a Convolutional Neural Network (CNN), the input is sequentially processed through a layer - for example, a convolutional layer (Wu, 2017).



*Figure 2: Equation that illustrates how a CNN runs layer by layer (Wu, 2017)*

In *Figure 2*, the equation shows the way in which a CNN runs layer by layer, with representing the input. In Wu’s (2017) example, the is an image that is being input and processed through the first layer and provides the output which then acts as the input for the next layer of processing. This proceeds until all the layers have been finished which then outputs (Wu, 2017). Wu (2017) notes that additional layers are added for backward error propagation.

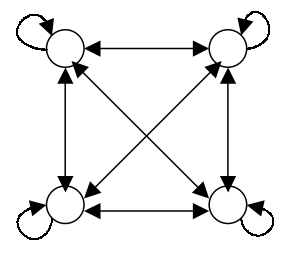
In recent years, there have been various improvements made to CNNs. For example, the convolutional layer has seen improvements by the creations of a Network in Network - or NIN (Gu et al., 2018). NIN was first proposed by Lin et al. (2014) as a general network structure that replaces the linear filter of the convolutional layer by a micro network.

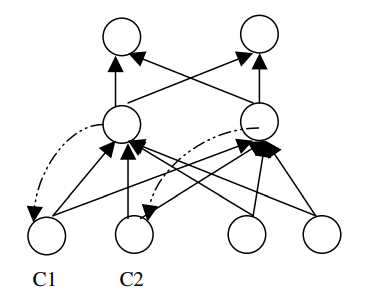
While CNNs were developed to classify images, the algorithm can be used for spatial data, since a map is considered as an image (Berk, 2021). Kounadi et al. (2020) found in their review of crime forecasting that spatial crime forecasting research has had significant growth. Loeffler and Flaxman (2018) note that the emphasis in predicting crime occurrence in specific locations has increased in recent years. Duan et al. (2017) developed a Spatio Temporal Crime Network (STCN) in an attempt to apply CNNs to automatically predict crime. The model can forecast crime risk for each region in an urban area, in this example New York, for the next day based upon the volume of high dimensional data that was received the previous day. The results showed that the STCN achieved between 88% and 92% accuracy when compared with felony and crime datasets from New York City in 2010 to 2015. The results were then used to visually explain to people the likelihood of crime in their area (Duan et al., 2017).

Similar to MLPs, an issue that can arise with CNNs is overfitting depending upon the size of the dataset that a model is being trained with (Shah et al., 2022). Using a single approach to CNNs can cause these issues, Shah et al. (2022) note that blended Long-Short Term Memory (LSTM) and CNNs will negate the overfitting potential and can positively impact the accuracy of the model.

### 2.2.3 Recurrent Neural Network

A Recurrent Neural Network (RNN) is a neural network with closed loop feedback connections (Fausett, 1994). The architectures of this algorithm range from fully connected nets (*Figure 3*) to partially connected nets (*Figure 4*), as well as multilayer feedforward networks similar to that of MLPs (Unadkat et al., 2001). RNNs are “capable of learning features and long-term dependencies from sequential and time-series data” (Salehinejad et al., 2017) and are, therefore, an excellent tool for prediction. Over the past 10 years, there has been a plethora of developments in RNNs, including ReNet, which is an RNN-based alternative to CNNs (Visin, 2015), and gated orthogonal recurrent units (Jing, 2017).

*Figure 3: A diagram of a fully connected RNN (Unadkat et al., 2001).*



*Figure 4: A diagram of a partially connected RNN (Unadkat et al., 2001).*

Meskela et al. (2020) designed a time-series crime prediction model utilizing Long Short-Term Memory (LSTM) RNNs. In their research, the data focused upon violent crime in Addis Ababa, Ethiopia and their model predicted the likelihood of crime on a monthly, daily and hourly basis (Meskela et al., 2020). By using LSTM RNN algorithms, Meskela et al.’s (2020) hourly prediction model had a 97% accuracy, providing a “high quality of precision.” Keerthika et al. (2024) conducted similar research in India with an aim to decrease crime. In their research, Keerthika et al. (2024) optimised confidence rates and compared these against unoptimized confidence rates. The optimised confidence rate was worked out as:

Where in:

(Keerthika et al., 2024)

This was then quantified as an average and compared against the CF.

By comparing the two confidence rates, Keerthika et al. (2024) were able to predict the likelihood of a certain crime if the two confidence rates were similar. By investigating this monthly, Keerthike et al. (2024) were able to identify which crimes were committed the most in each month, with Harassment being more common in June than any other month whereas Petit Larceny was more likely to be committed in December. Unfortunately, the research conducted by Keerthike et al. (2024) does not produce suggestions or infer as to why this is the case. It would be beneficial for police if there was a known catalyst for the increase and decrease in crimes each month.

Similarly to MLPs, RNNs can experience a vanishing gradient, which can make it difficult to capture long-term dependencies (Amidi and Amidi, 2018). Amidi and Amidi (2018) note that using RNNs can make computation slow. Sharkawy (2020) argues that although an RNN is simple and powerful in theory, in practice it can be hard to train correctly due to vanishing gradients.

## 2.3 Bias in Predictive Policing

Although evidence has been found that predictive policing works, ML can reflect and entrench biases of humans, which has led to discussions and research regarding the fairness of predictive policing and the threat of bias towards certain groups in society (Alikhademi et al., 2021). Lum and Isaac (2016) note that communities that have a higher police presence will naturally have a higher arrest rate, therefore the dataset will reflect this Lum and Isaac (2016) argue that it appears to reflect a higher crime rate, however, what it truly reflects is greater police attention Studies have shown that this greater police attention - generally in tends to lean towards ethnic minorities and neighbourhoods with lower socio-economic status (Lum and Isaac, 2016). This was found to be common in corrupt police forces that have entrenched within them institutional racism and classism, where predictive policing being utilised exacerbates the corruption (Richardson et al., 2019). Perry et al. (2018) note that by using historical datasets, historical bias can also occur, which ignores the fact that society changes - for example, the motivations, means and types of perpetrators can change over time. This must be taken into consideration, otherwise the historical biases could create a feedback loop (Perry et al., 2018).

In order to combat this issue, a few solutions have been suggested to produce algorithmic fairness. Mehrabi et al. (2019) defines algorithmic fairness as “absence of any prejudice or favouritism toward an individual or a group based on their inherent or acquired characteristics.” Alikhademi et al. (2021) note that the methods suggested to improve algorithms within the criminal justice sphere can be split into two main groups: algorithm designs being thought about qualitatively in advance of bias being detected or programmatic interventions being placed in the algorithm to detect and/or correct the bias. A suggestion from the first camp of thought is Altman et al. (2018) proposing that algorithm designers would have to find areas in the algorithm design that could be subject to bias, for example, asking questions such as ‘what would the results look like if race was excluded from the training data?’ Those who fall into the second category have suggested modifications of the input provided to predictive policing algorithms. Ensign et al. (2017) investigated this using PredPol - which was used by the LAPD up until relatively recently. PredPol is designed to learn from new crime data that is gathered by patrolling officers so that the police can adapt (Alikhademi et al., 2021). They found that the algorithm was sending police officers back to the site they just patrolled, due to the new data input, leading to a vicious cycle. Ensign et al. (2017) proposed a new solution that alters the way in which data is added so that the more likely the police are sent to a given district, the less likely it is that it is incorporated into the data. Without the addition proposed by Ensign et al. (2017), PredPol’s predictions did not accurately predict the true crime rate - something that is concerning as the LAPD were actively using PredPol at the time. Other police forces have developed different predictive policing methods to attempt to eliminate bias. An example of this is Patternizr, a predictive policing application created by data scientists at the New York Police Department (NYPD), that uses the *modus operandi* (M.O.) of the crime, rather than the location or people, allowing the data to be less bias (Griffard, 2019). The M.O. can be defined as a “set of habits that the offender follows and is a type of motif used to characterise the pattern” (Wang et al., 2013). Although the developers of Patternizr note that the application is fair and an accurate predictive policing tool, there have been concerns raised as to whether this is true (Griffard, 2019). However, unlike PredPol, this is due to the application not being independently assessed by outside researchers as Patternizr has currently only been assessed by the NYPD (Griffard, 2019).

From the research into bias conducted above, this project will aim to limit bias by focusing upon the M.O. and crime of the perpetrator, rather than the type of person or location as the mitigating factor. Location will still be required to be in the dataset so that the places where the offences occur can be predicted. An anti-bias algorithm could also be used to negate bias from the public policing data collected for this project, however, it is yet to be decided if this would benefit the project.

## 2.4 Challenges in Implementing Deep Learning for Domestic Violence Prediction

Hui et al. (2023) note that there is a lack of data sources to conduct machine learning in intimate partner violence research. Berk et al. (2016) report that, in some cases, records are still in paper format and that there is a lack of electronic records which are essential to machine learning. Paper records would require scanning documents and typing handwriting - both are time consuming and can be inaccurate (Hui et al., 2023). This can impede the use of data-mining techniques to retrieve data from the sources (Hui et al., 2023).

There are further issues with the data, with people from different backgrounds being underrepresented or overrepresented in the dataset. Rahmen et al. (2023) note that education level was a factor in if women reported abuse to the police, with those with less of an education not knowing rights and relying on their partner. While this means that these women are less likely to report abuse to the police, it does increase their risk of being victims of domestic abuse (Rahmen et al., 2023).

One of the major issues with predicting domestic abuse and intimate partner violence is that it often occurs away from the public places and, most often, the only witnesses to the abuse are those involved or those who live in the same household (Dar, 2013). It is common for those who have experienced domestic abuse to have difficulty in identifying the behaviours as abusive, more likely if the abuse is financial or psychological, and some survivors of domestic abuse are reluctant to report (Swallow, 2017). ONS (2023) notes that only 18.9% of women who had experienced intimate partner abuse in 2022/23 reported the abuse to the police. Although the police attend a domestic abuse related call every 30-seconds, it is estimated that less than 24% of domestic abuse crimes are reported (Refuge, 2024). SafeLives (2024) notes that on average, victims of domestic abuse will experience 50 incidents, ranging from physical to psychological, before reporting. This can lead to the data being skewed and not being necessarily accurate within the context of wider society. Although the dataset would be incomplete in regard to wider society, it is necessary to use this dataset as any attempt to account for this, such as oversampling, could lead to further bias as with PredPol (Ensign et al., 2017).

# 3.0 Project Rationale

Most research into this subject is done from a social science perspective, focussing upon qualitative methods to collect and display data. Research that has been conducted from a computer science standpoint about predicting domestic violence has not been accurate, with models obtaining high rates of false positives - up to 99% (Thornton, 2017). This project aims to change this by producing a model that will provide accurate predictive results regarding the likelihood of domestic violence instances for a wide range of sporting events by comparing MLPs, CNNs, and RNNs.

Using major sporting events as a marker for the likelihood of domestic violence instances has been researched in a quantitative manner, however, using these major sporting events to predict these crimes is an area that is lacking. Kirby, Francis and O’Flaherty (2013) found trends that with each FIFA tournament, the reports of domestic abuse incidents increased. However, as noted previously, the research only encompasses three tournaments (2002, 2006, and 2010).

Further investigation into this would be beneficial with more recent crime data - which is something that this project aims to do with the use of historical, as well as the most recent, public crime data (2019-2024).

Previous research into this topic has been either qualitative experiences recorded from a social science point of view, or quantitative research based upon risk assessments and small samples of data. This study will use a wider range of data, both historical and contemporary, as well as using machine learning to train a model to predict the likelihood of violence based on location.

By successfully creating this project and producing accurate results, this project would prove that predicting instances of domestic violence, by location and during major sporting events, is possible and plausible. As well as this, it would also prove a direct link between the two variables of domestic violence and sport, with successful predictions implying that the variables are dependent. Current research suggests that this is not possible with Thornton (2017) noting that being able to predict “domestic violence based on intelligence from prior police contacts does not appear possible at present.” Although in recent years, predictive policing using artificial intelligence has increased in reliability (Walczak, 2021).

This research project will aim to create a model that predicts the likelihood of domestic violence occurrences during major sporting events using deep learning. Creating this project successfully would be a meaningful contribution to the ways in which crime can be forecast and police can allocate resources.

# 4.0 Research Approach

Research will begin with a rigorous exploration of historical and contemporary research surrounding domestic violence prediction, as well as domestic violence surrounding major sporting events. A comprehensive literature review will follow this research.

Data will then be collected through a Freedom of Information Act via Northumbria police that will provide data regarding instances of domestic violence. A .csv file will be created that will outline football matches throughout England over the past five years, in major leagues and events - Premier League, Championship, World Cup, FA Cup, and Euros.

Following this, an experimental model with the goal of yielding predictions of domestic violence instances during football matches will be created. The output of this model intends to indicate the likelihood of these instances.

Deep learning will be used to accurately predict instances of domestic violence surrounding football matches. Three deep learning algorithms will be compared and used in this project: Multi-Layer Perceptrons (MLPs), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN).

# 5.0 Research Aim and Objectives

## 5.1 Research Aim

The project aims to provide a predictive model of domestic violence instances in the United Kingdom in relation to football matches through historical police records marked as domestic violence instances and timelines of football matches by comparing three deep learning algorithms that are commonly used in predictive policing. Datasets will need to be combined to create a model for domestic violence prediction that encompasses sporting events and for the aim to be completed. This research will compare three different deep learning and artificial intelligence methods: MLPs, CNNs, and RNNs.

## 5.2 Research Objectives

**Historical Police Records Dataset**

**Description:** Locating, downloading, importing, and classifying data. Preparing for training and testing. Data separated into a training and testing set. Full copy created to use in real time integration.

**Acceptance Criteria:** Data imported and denoised, set up and ready for training the model.

**Time Estimate**: 3 Weeks

**Historical Football Events Dataset**

**Description**: Locating, downloading, importing, and classifying data. Preparing for training and testing. Data separated into a training and testing set. Full copy created to use in real time integration.

**Acceptance Criteria:** Data imported and denoised, set up and ready for training the model.

**Time Estimate**: 3 Weeks

**Combining the Datasets**

**Description**: Joining the datasets together to see where and when major sporting events occurred.

**Acceptance Criteria:** Datasets successfully combined.

**Time Estimate**: 1 Week

**Model Training Using the Dataset**

**Description:** Using the model to train the dataset.

**Acceptance Criteria:** Model trained.

**Time Estimate:** 3 Weeks

**Real Time Major Sporting Event Integration**

**Description**: Importing data from Sporting Event API.

**Acceptance Criteria**: Data stored, and types match the current data.

**Time Estimate**: 3 Weeks

**Accuracy Testing**

**Description**: Testing accuracy at different stages of the project. These will be compared.

**Acceptance Criteria**: Accuracy values included in dissertation analysis.

**Time Estimate**: 1 Week

**Dissertation Writing**

**Description**: Writing sections of dissertation and completing throughout the research project. Will be completed simultaneously with other aspects of the project.

**Acceptance Criteria**: Final version completed and checked - handed into eLP.

**Time Estimate**: 12 Weeks

# 6.0 Methodology

Deep learning will be used to accurately predict instances of domestic violence surrounding football in the Northeast of England. Three deep learning algorithms will be used: Multi-Layer Perceptron (MLP), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Multi-layered perceptrons, recurrent neural networks and convolutional neural networks are the most commonly used algorithms in predictive policing (Walczak, 2021).

A multi-layer perceptron (MLP) is a supervised learning feed forward network most commonly trained using back propagation (Walczak, 2021). Back propagation is a machine learning technique that adjusts the weights in a neural network to improve predictions (Walczak, 2021). MLPs have developed consistently since their inception and have become a powerful technique for solving a variety of issues (Delashmit and Manry, 2006). An issue that can occur using MLPs is gradient vanishing where, when training, as the sequence length increases, the partial derivative of the loss function exponentially decreases, and - in some cases - may completely halt the neural network from further learning (Banerjee, 2024). By adding more nodes and slowing learning rate, this should allow for the gradient vanishing to be minimised.

CNNs can be used for a multitude of tasks - from image recognition to natural language processing and can process large amounts of data (Gu et al., 2018). This project will require a significant amount of data meaning that CNNs are an appropriate algorithm to use.

Recurrent Neural Networks (RNN) are “used for problems that are based on time, such as time-series forecasting” (Walczak, 2021). This algorithm is capable of running features and long-term dependencies from sequential and time series data (Salehinejad et al., 2018) and has been an important focus of research since the early 1990’s (Unadkat et al., 2001).

Each of the algorithms will be implemented and then evaluated to investigate the accuracy of each method for the purpose of predicting domestic violence and its relation to football matches. These algorithms were chosen as they are some of the most robust when researching deep learning algorithms for prediction. MLPs are a general-purpose model and are able to approximate any function, given that there are enough hidden layers and neurons in the model built (Cybenko, 1989). Banerjee (2024) notes that MLPs can be prone to overfitting, learning every aspect including the noise of the data, especially on smaller datasets. CNNs, on the other hand, are able to predict time series forecasting - for example, stock market prediction (Borovykh, Bohte, and Oosterlee, 2017). However, there are some disadvantages of using CNN models, for example, if the model has multiple neural layers, then the training may take a lot of time depending on GPU ability (Bhuiya, 2020). RNNs are able to handle sequential data, allowing them to capture any temporal patterns which are ideal for forecasting via time-series (Petneházi, 2019). A disadvantage of RNNs is that the model can develop a vanishing gradient (Bengio et al., 1994) which is when a neural network’s weights become extremely small during backpropagation.

Studies into domestic violence prediction have been based on risk assessments and statistical analysis. By utilising machine learning, as in predictive policing, and the use of CNNs, the research project aims to predict the geography of domestic violence in relation to football accurately.

The data that will be used in this project has been collected from publicly available sources and compiled into a .csv file. The police data has been obtained via the Freedom of Information Act from Northumbria Police, whereas the football data has been collected from a football statistics website (worldfootball.net). These will be converted to Pandas dataframes in Jupyter Notebooks and visualised, before being transformed and used to train and test prediction models.

## 6.1 Ethical Implications

### 6.1.1 Data Storage

The data used in this project will be from publicly available police records and a publicly available website that records match history of football, manually converted into a .csv file. This data will not contain sensitive personal information, such as names or addresses, that could identify a person. All data used in this project will be anonymous due to sensitivity and the project being presented in the public space. Northumbria Police anonymise their data before releasing it to the public and the police records were obtained via the Freedom of Information Act.

All data will be stored securely on a personal PC which will not be moved or leave the house and is password protected. Files will be backed up to the university OneDrive which ensures that the data is stored securely, and password protected. Data will be destroyed once the research project has concluded, and marks awarded. All data not being present in the research will be removed from all storage devices above and thoroughly digitally disposed of. No data will be collected on paper.

### 6.1.2 False Positives, False Negatives, and Model Accuracy

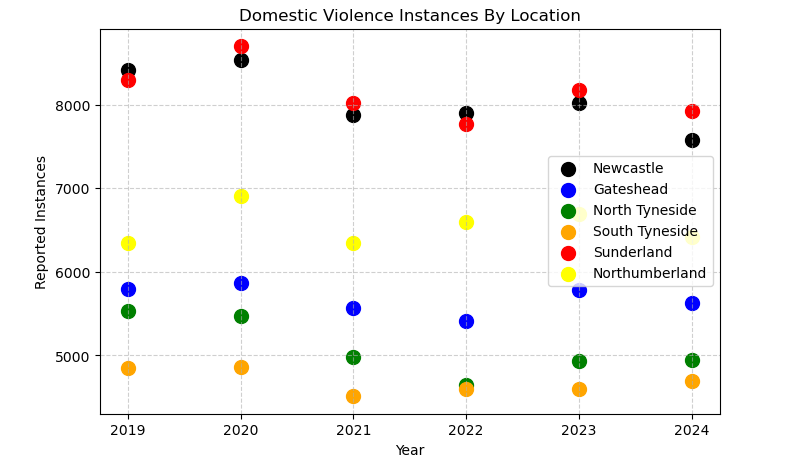
False positives and negatives would compromise the model’s accuracy. As stated previously, Thornton’s (2017) model from Thames Valley Police produced a false positive rate of up to 99%, and false negative rate of 89%, leading to fatal cases being incorrectly labelled as medium risk. It is important to ensure that there are limited false positives and negatives in the model as, if it were to be used by Northumbria Police, it may result in issues surrounding false accusations (positives) or not placing resources where needed (negatives). If police were to place resources exclusively in locations where the model stated there were going to be domestic violence instances and there were false negatives, this would mean that areas of the North-East where the model falsely stated would have less domestic violence would have to wait longer for responses than those that were manned. In some instances, this could mean life or death for victims of domestic violence, as proven in Thorton’s (2017) research.

# 7.0 Results

This section of the report outlines the results of both the data visualisation performed and the results of the predictive model created.

## 7.1 Data Visualisation

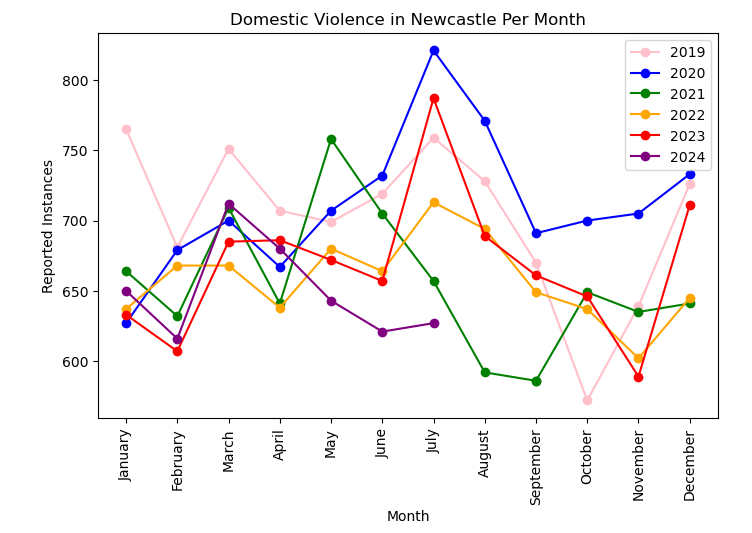
### 7.1.1 Domestic Violence Instances

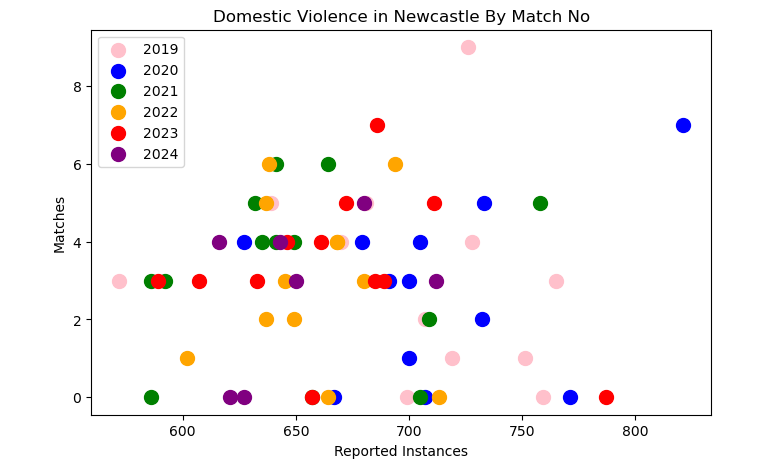


*Figure 5: Graph showing domestic violence instances in the Northumbria Police area per year, separated by district.*

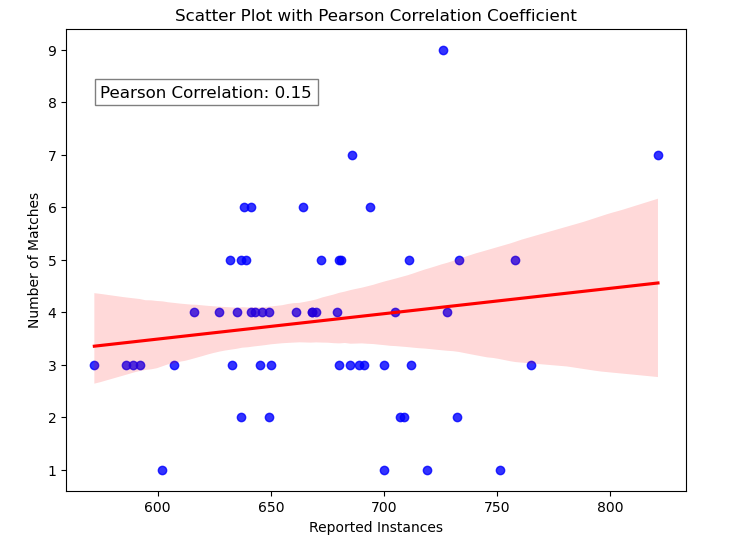
The above scatter graph (Figure 5) shows the domestic violence instances per year, separated by location. Newcastle and Sunderland show the highest rates each year, in comparison to the other areas in the Northeast. The area with the lowest amount of domestic violence instances per year is South Tyneside.

### 7.1.2 Newcastle Domestic Violence and Football Matches



*Figure 8: Line graph showing trends of domestic violence in Newcastle between January 2019 and July 2024.*

*Figure 9: Scatter graph outlining domestic violence incidents compared to matches played by Newcastle.*

*Figure 10: Scatter graph with Pearson Correlation Coefficient outlining correlation between Newcastle football matches and domestic violence instances*

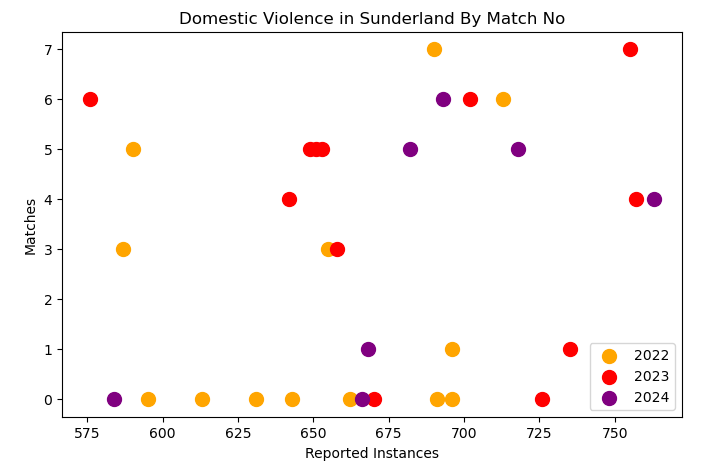
In Figure 8, the line graph shows the trends of domestic violence in Newcastle by month, by year. July 2020 shows a spike in domestic violence instances in Newcastle. The summer months tend to show an increase in reported instances.

Figure 9 shows the amount of football matches in the month compared to the number of reported instances of domestic violence in Newcastle. The graph shows that once football matches are between a frequency of 3-4 a month, the domestic violence instances also increase.

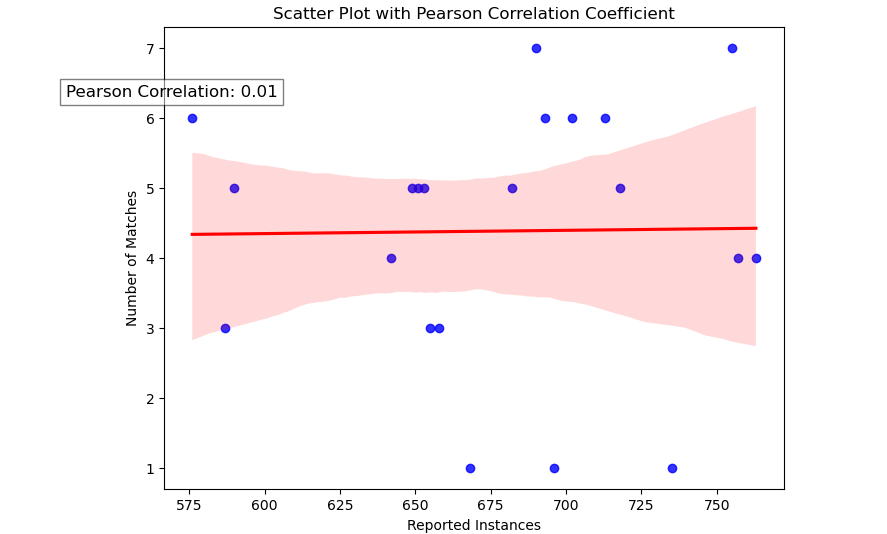
Figure 10 is a scatter graph that plots the Pearson Correlation Coefficient of matches and domestic violence instances in Newcastle. The Pearson Correlation Coefficient is 0.15, indicating a weak positive correlation, suggesting that there is a small positive relationship between the two variables.

### 7.1.3 Sunderland Domestic Violence and Football Matches

*Figure 11: Line graph showing trends of domestic violence in Sunderland between January 2022 and July 2024.*



*Figure 12: Scatter graph outlining domestic violence incidents compared to matches played by Sunderland*



*Figure 13: Scatter graph with Pearson Correlation Coefficient outlining correlation between Newcastle football matches and domestic violence instances*

In Figure 11, the line graph shows the trends of domestic violence in Sunderland throughout the months, by year. January 2024, August 2023 and December 2023 show a spike in domestic violence instances in Sunderland. 2022 tends to follow a similar pattern to the Newcastle graph where the amount of instances increases during the summer, however, this is different in 2023 and 2024, where instances are highest in December and January respectively.

Figure 12 shows the amount of football matches in the month compared to the amount of reported instances of domestic violence in Sunderland. The graph shows that there is a slight uptake in domestic violence as there are more matches, however, there is also a significant number of instances reported when there are not matches that month.

Figure 13 is a scatter graph that plots the Pearson Correlation Coefficient of matches and domestic violence instances in Sunderland. The Pearson Correlation is 0.01, which is an extremely weak positive correlation, suggesting that there is little to no link between the two in Sunderland.

## 7.2 Models

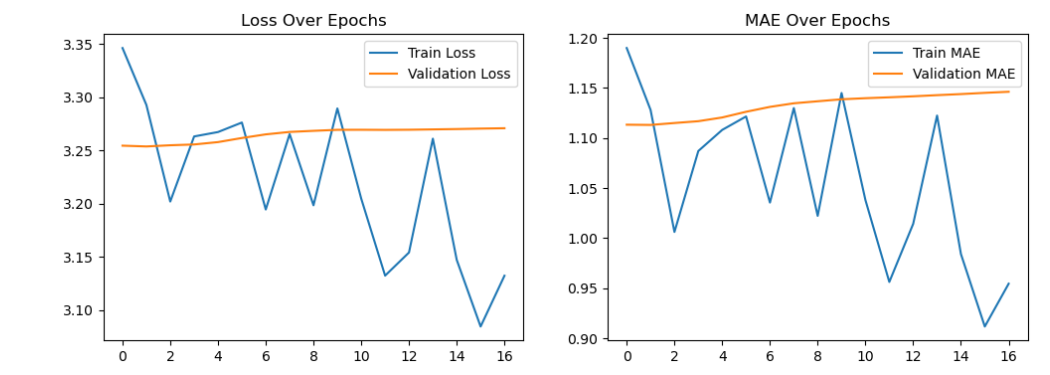
The models used were Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN).

|  |  |  |
| --- | --- | --- |
| Algorithm / Model | Mean Absolute Error (MAE) | Mean Absolute Percentage Error (MAPE) - rounded 1 d. p |
| Multi-Layer Perceptron | 51.68 | 20.8% |
| Convolutional Neural Network | 0.9856 | 5.9% |
| Recurrent Neural Network | 0.2235 | 0.1% |

*Figure 14: Table outlining results of each model*

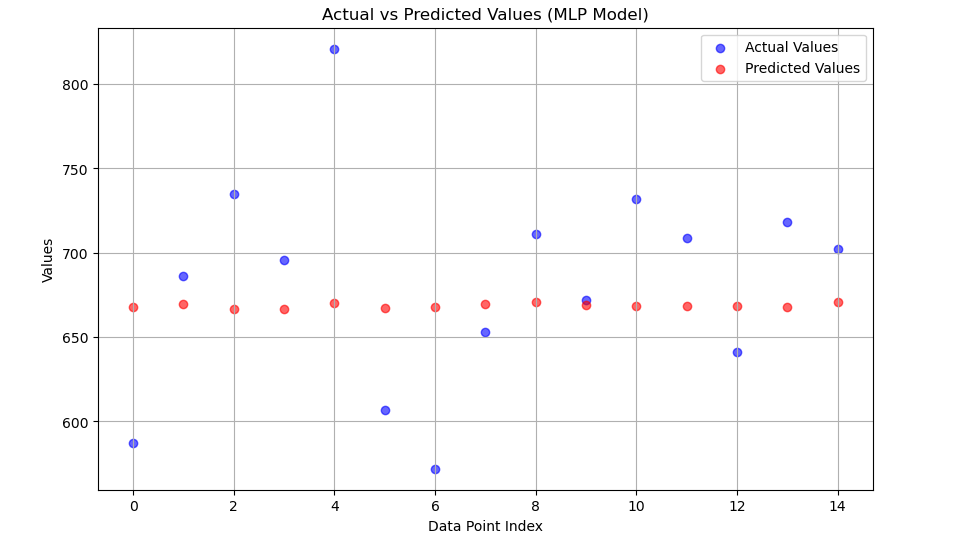
The MAE of MLPs is significantly higher than that of CNNs and RNNs, sitting at 51.68 MAE with an MAPE of 20.8%. A MAE of 51.68 shows that the model is not as accurate as others produced. The CNN model has an MAE of 0.9856 and an MAPE of 5.9%, compared to the RNN MAE of 0.2235, with an MAPE of 0.1%. From the results above, it displays that a RNN has a greater chance of producing accurate predictions surrounding domestic violence and football.

### 7.2.1 Multi-Layer Perceptron

****

*Figure15: Line graphs depicting the loss and mean absolute error over the course of MLP model training*

In Figure 15, the first graph depicts the loss over epochs, and the blue line indicates that there has been unstable training due to an inconsistent decrease. The orange line suggests that the model has not been improving. The second graph depicts the MAE over the epochs and suggests similar issues within the blue line as in the first graph. The validation MAE line (orange) increases gradually, meaning the model is generalising unseen data poorly.

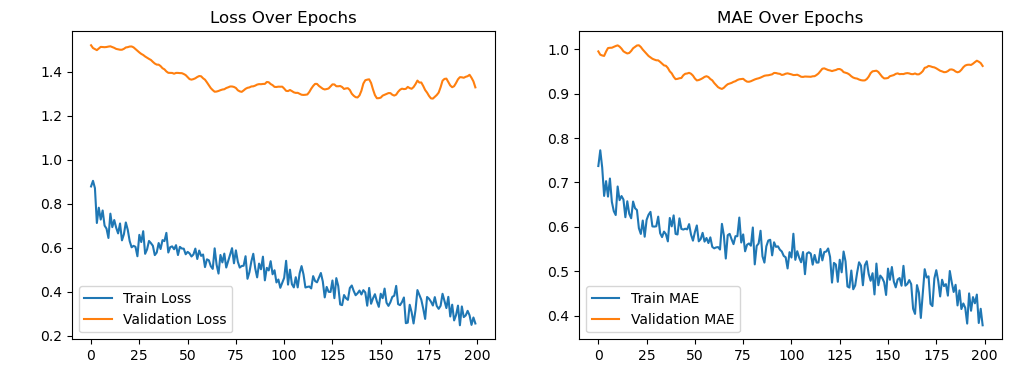


*Figure16: Scatter graph depicting the actual values vs the predicted values of the MLP Model*

Figure 16 shows the actual values against the predicted values and shows that the actual values have more variance than the predicted values. This implies that the model is too smooth and is failing to capture any fluctuations.

### 

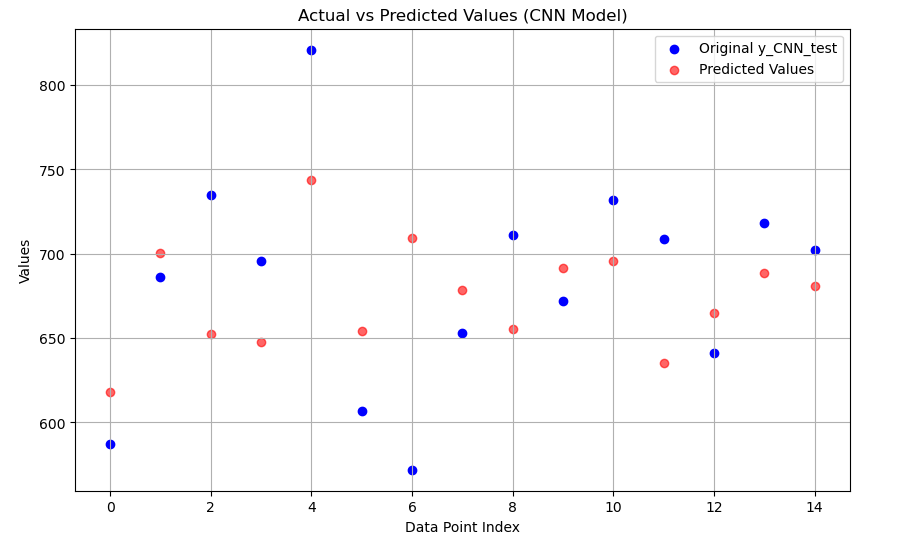
### 7.2.2 Convolutional Neural Network



*Figure 17: Line graphs depicting the loss and mean absolute error over the course of CNN model training*

Figure 17 shows both the loss over epochs and MAE over epochs. Train loss (blue line) in the first graph indicates that the model is learning due to it consistently decreasing over time. Validation loss does not decrease in the same way, remaining consistent with a few fluctuations.

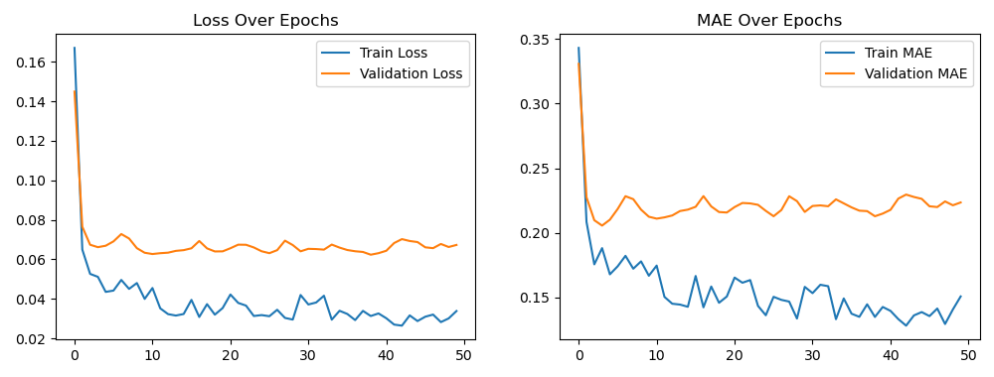
In the second graph, the training MAE implies that the model is well trained, with the MAE decreasing steadily. Validation MAE is similar to that of validation loss.



*Figure 18: Scatter graph depicting the actual values vs the predicted values of the CNN Model*

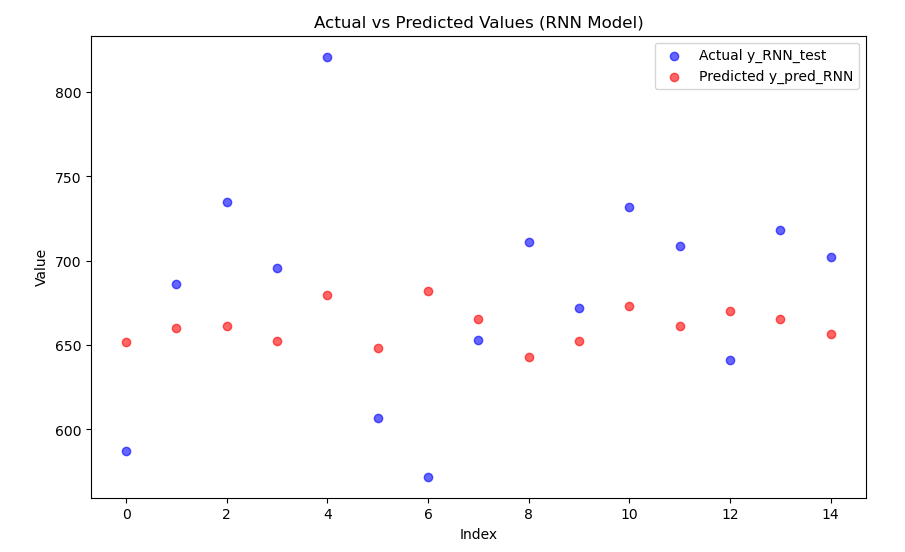
It is possible to see in Figure 18 that the model is capturing some of the trends better than the MLP model. The predicted values are more spread out and align closely with actual values. However, the predictions still appear somewhat biased towards a specific range and do not capture the full variance seen in the actual values.

### 7.2.3 Recurrent Neural Network

****

*Figure 19: Line graphs depicting the loss and mean absolute error over the course of RNN model training*

Figure 19 shows the loss and MAE over Epochs of the RNN model. On the left, the graph shows that the model is learning from the training data (train loss), but the validation loss flattens instead of following the training line, indicating that the model stops improving on unseen data. However, the graph on the right shows that the model is improving, with the train MAE line decreasing over time. Validation MAE decreases initially but then fluctuates which suggests that it is not improving after a certain point.

****

*Figure 20: Scatter graph depicting the actual values vs the predicted values of the RNN Model*

Figure 20 shows the predicted values against the actual values. The predicted values are close to the mean, suggesting that the model is too smooth, failing to capture large fluctuations.

# 8.0 Application

Whilst creating the MLP model, further research was conducted, and it was noted that using XGBoost may have been a better option when compared to the MLP model. This is due to the data being tabular and a relatively small dataset. A weakness of using MLPs for prediction models is the risk of overfitting, especially when trained on smaller datasets (Banerjee, 2024). XGBoost, also known as Extreme Gradient Boosting, is a machine learning library that uses gradient boosting and can be used for both classification and regression tasks (Chen and He, 2024). Gradient boosting is a supervised learning algorithm that aims to find the best function that minimises the expected loss between input data and predictions produced by the model (Chen and Guestrin, 2016). XGBoost is known for its scalability and allows research to be continued with further data (Chen and Guestrin, 2016). In further research, it would be suggested that this be used in place of the MLP model. It was decided that the XGBoost algorithm would not be used due to prior research conducted into predictive policing models that used MLP, CNN, and RNN, as stated in section 2.1, and the research conducted in this dissertation was to ascertain which of these models was best suited for predicting domestic violence occurrences surrounding football matches.

The FA Cup and World Cup datasets were not large enough to accurately develop a model and predict domestic violence surrounding the events. The World Cup is held every four years, most recently in 2022. Data collected for the World Cup was from the years 2018 and 2022, however, the Northumbria police data went back to 2019, meaning there was only one year available to train in a model. The results from this would not have been accurate due to the limited amounts of data. The FA Cup is held annually; however, not all teams will get through the qualifying rounds. As the data was manually collated, it was not possible, due to time restrictions, to manually input the over 800 matches a year, with around 700 teams in the cup overall. It was decided that the best course of action was to start from the First Round Proper, which occurs after the qualifying stage, as there was too much data to manually input, this takes the amount of matches to around 160 matches. In some cases, North-Eastern teams such as Newcastle United or Sunderland AFC are exempt for the First and Second Round Proper, meaning that the dataset was not robust. In order to further this research, it would be beneficial to review the data produced and collated for the project. Non-league teams would allow other areas of the North-East to be analysed and predicted. In the Northumbria police data, there were six areas in the dataset, two being Newcastle and Sunderland who have football teams in the Premier League and Championship respectively. Other areas in the North-East tend to have teams that are in lower divisions and leagues such as National League and National League North, which data was not collected for. By collecting data for lower divisions and leagues, the data would be more robust for the Northeast. Further research could also include expansion into other constituencies to analyse the Premier League and Championship with additional data. Vabalas et al. (2019) note that a larger dataset has “greater statistical power for pattern recognition.” Larger datasets are becoming more common due to easier and more widely available methods of collection, such as the Internet of Things (Vabalas et al., 2019). A risk of producing a prediction model with a smaller sample size is that there is the likelihood of overfitting and issues with accuracy (Vabalas et al., 2019). CNNs especially are usually trained using larger datasets (Barbedo, 2018). By having a larger dataset, the CNN may have a higher accuracy.

# 9.0 Analysis and Discussion

The following section of this report will discuss and analyse the results outlined above.

## 9.1 Data Visualisation and Pearson Correlation Coefficient

Figure 10 shows the correlation between domestic violence and football matches in Newcastle between 2019 and 2024 seasons, by month and by year. The Pearson Correlation Coefficient is 0.15, suggesting a weak positive correlation between the two variables of matches played and reported domestic violence instances. There is a slight upward trend indicating that there is an increase in instances of domestic violence when football matches are played, however, the trend is weak and not a strong predictor that the variables are related. Analysing the graph, it appears that there is a lot of dispersion of the data, meaning that other factors could be affecting the reported instances.

To improve upon this, checking for outliers may help seeing the correlation clearer as they can weaken an observed correlation. It is possible that the outliers are due to other variables other than matches played and domestic violence instances. One example could be if Newcastle won, drew, or lost the match. This would need to be investigated further and could be an area for further research, adding to the research Kirby, Francis, and O’Flaherty’s (2013) conducted into the England national team winning, losing, or drawing matches, but on a more localised level.

Figure 13 shows the correlation between domestic violence and football matches in Sunderland between 2022-2024 seasons, by month and by year. The Pearson Correlation Coefficient of this dataset is 0.01, suggesting that there is no correlation between the two variables. The observed correlation is more likely due to random chance than an actual trend or correlation. On Figure 13, the shaded red area (which represents the confidence interval of the correlation) shows uncertainty around the trendline.

To possibly improve upon the trend calculation, it would be beneficial to find more data regarding Sunderland matches. As Sunderland were promoted to the Championship in 2022, the team would have played in another league previously, meaning that their games would be part of a different dataset. It may also be beneficial to collect data from each of the football teams in the North-East, rather than using the league tables and statistics. Having a larger dataset may help find a more accurate correlation and allows for the identification of more patterns, and trends - which, in turn, leads to a deeper understanding of the data being visualised. Unwin (2020) notes that larger data sets can be analysed, visualised and graphed in a way that allows them to play a valuable role in “diagnosing the strengths and weaknesses” of the data.

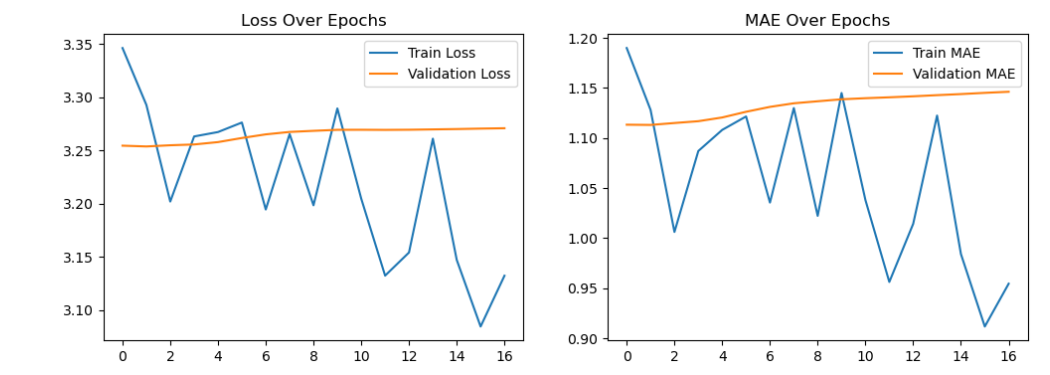
The Pearson Correlation is only one type of correlation and others could be considered. An example of this would be the Spearman correlation which may have been more appropriate for this visualisation. Spearman correlation is dissimilar to Pearson as it is not a measure of linear relationships (Hauke and Kossowski, 2011). It does not require an assumption of a relationship between the two variables to measure the correlation (Hauke and Kossowski, 2011). In order for the Pearson Correlation to accurately correlate between the two variables, the relationship must be linear (i.e. a straight line on the scatter plot) whereas the correlation can be non-linear in Spearman (Hauke and Kossowski, 2011). This may have been a better correlation coefficient to use due to the correlation possibly being non-linear. In future work, the Spearman correlation would be best practice for the visualisation of the data.

## 

## 9.2 Models and Accuracy

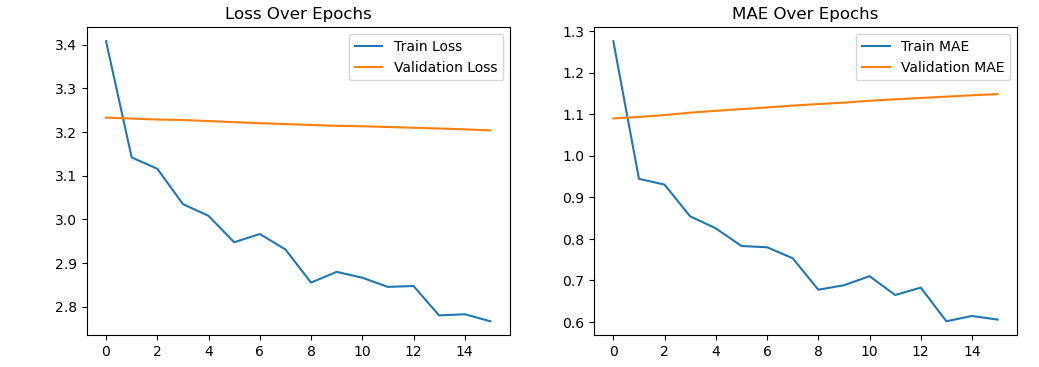
The Multi-Layer Perceptron (MLP) Model training was unstable, resulting in a low magnitude of variance. This failed to detect any spikes and anomalies in the dataset. It’s possible that this could be due to an issue with the data sample size being too small. The learning rate for this model had to be set low (0.00001) as the model was not training in an accurate way if the learning rate was any higher than this.



****

*Figure 21: Code above that outlines the compiling of the model with a learning rate of 0.00001, and a line graph showing the loss and MAE of the model with this learning rate.*





*Figure 22: Code above that outlines the compiling of the model with a learning rate of 0.0001, and a line graph showing the loss and MAE of the model with this learning rate.*

From Figures 21 and 22, it is possible to see the issues caused by a higher learning rate. In Figure 22, both the Train Loss and Train MAE are decreasing, which implies that the model is learning on the training data, however, the decrease is flat, unlike that of the model in Figure 21 with the spiking. This suggests that the model is overfitting, possibly due to outliers in the data, meaning that it is not learning patterns and is memorising the training data instead. By not learning the patterns in the data, this would mean that the model would not be accurate in predicting the likelihood of domestic violence instances. If the learning rate is too high, it can cause the model to train too quickly before optimising the weight of data, which explains why the validation loss and MAE do not improve in Figure 22. Brownlee (2020) notes that a learning rate is “is a configurable hyperparameter used in the training of neural networks that has a small positive value, often in the range between 0.0 and 1.0.” By lowering the learning rate, the weights of the neural network will be updated at a slower speed which can avoid overfitting (Mishra, 2023). This was the reason a lower learning rate was chosen as the version with the 0.0001 learning rate appeared to be overfitting based on its output.

The Convolutional Neural Network (CNN) performs better than the MLP as it recognises patterns in sequential data, as well as capturing variations in prediction. However, based on the graphs in 9.2.2, the model appears to be overfitting and memorising patterns, instead of generalising. CNNs are generally used for spatial relationships rather than long-term dependencies, which is something the model is known to struggle with. To improve upon this model, a lower learning rate could be used (Mishra, 2023) or it could be combined with another algorithm to create a hybrid model, for example, with Long Short-Term Memory (LSTM) (Brownlee, 2020).

The Recurrent Neural Network (RNN) performs the best of all three models at modelling the sequences and time dependencies. The MAE of the RNN was 0.2235, compared with the MLP and CNN models which had MAEs of 51.68 and 0.9856 respectively. This model trained in a more stable way than the MLP and CNN models. However, the predictions were still too smooth, meaning that it has not fully captured extreme variations - such as sudden spikes in reporting of domestic violence. RNNs tend to also have limitations in remembering long-term dependencies which is due to the model tending to suffer from vanishing or exploding gradients (Jain, 2024). This means that the data becomes either overly large or overly small as it works through the neural network (Jain, 2024). Even though the model failed to capture the magnitude of variance for some predictions, the accuracy of the model overall indicates that it is possible, potentially with a larger dataset, to accurately predict domestic violence instances using football matches as a feature.

From this research, it is possible to determine that deep learning performs better than traditional machine learning methodologies for predictive models regarding domestic violence and football events, with the RNN operating at a higher level of accuracy than the MLP and CNN.

Due to the size of the dataset being insufficient, this likely led to a pattern of overfitting, extending through all three models that were trained within this project (Vabalas et al., 2019). In order to rectify this issue, it is recommended to use either a larger time scale or a more granular approach to gathering football match data (Vabalas et al., 2019).

Further steps could be taken by using either Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) models to better handle the long-term dependencies (Brownlee, 2020). Another possibility is exploring hybrid models, for example combining CNNs and LSTMs. Lilhore et al. (2024) note that their hybrid LSTM and CNN model, which aims to identify vocal changes in those with Parkinson’s Disease, achieved better results (90.8%) when compared to non-hybrid models. LSTMs are a type of RNN that is designed to handle long-term dependencies in data, and it aims to solve the vanishing gradient problem that many RNNs face (Brownlee, 2020). This makes it effective at time-series forecasting, event predictions and pattern recognition. LSTMs are widely used in prediction research such as stock market forecasting, weather prediction and natural language processing (NLP) (Brownlee, 2020).

# 10.0 Conclusion

Previous research suggested that there was a link between domestic violence and football events - with domestic violence being a “global epidemic” (Forsdike, O’Sullivan, and Hooker, 2022). The Office of National Statistics (2023) reported that one in four women and one in seven men will be a victim of domestic abuse in their lifetime. Despite this, tools used to predict future instances of domestic violence within predictive policing are lacking and those that have been developed have been to varying degrees of success (Karystianis et al., 2021). It was found in the literature review that approaches of predictive analysis in crime generally applied machine learning to Tweets (Gerber, 2013) or police records (Towers et al., 2018) in order to forecast criminal activity and the accuracy tended to range from 70-81%. There has been recent research into the prediction of domestic violence, but it has been to a limited success, with some models producing weak predictive accuracy (Karystiantis et al., 2021).

The aims of this project were to create a predictive model that would be able to predict if a domestic violence incident would occur due to the football match in the Northeast of England. This is due to there being a gap in predictive models that accurately can predict domestic violence instances during sporting events. By producing this model, this project hoped to bridge that gap and, in doing so, aid the police in predicting instances of domestic violence.

Three deep learning algorithms were used in this project: Multi-Layered Perceptrons (MLP), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). These algorithms were chosen as they were the most commonly used algorithms in predictive policing (Walczak, 2021). Each of the algorithms were implemented and then evaluated to investigate the accuracy of each method for the purpose of predicting domestic violence and its relation to football matches.

The data visualisation of the matches and domestic violence instances correlation showed a weak positive correlation in Newcastle (0.15), but the data from Sunderland suggested that there was no link between domestic violence and the football matches (0.01). This could have been due to varying data sizes, as Sunderland’s dataset only began in 2022, due to the team not being in the football leagues that were used for data collection prior to that year. It would be beneficial in future research to obtain more data from the minor football leagues to ascertain if there is a correlation between Sunderland matches and domestic violence. Further to this, ensuring that outliers were checked may allow for a stronger correlation between both Sunderland and Newcastle’s football matches and domestic violence.

Out of the models instantiated, the RNN model predicted the most accurately, with a Mean Absolute Error (MAE) of 0.2235, in comparison with the MLP and CNN models which had MAEs of 51.68 and 0.9856 respectively. The MLP model had unstable training producing smooth predictions, likely due to the data set being too small. The CNN performed better, capturing variations in predictions, however, it does appear that the model is overfitting. The RNN model trained in a significantly more stable way that both the MLP and CNN models, however the predictions were still relatively smooth compared to the original data, implying that it was not capturing any sudden spikes. However, even though the model did not capture the magnitude of variance in predictions, this does not necessarily mean that it is a non-viable deep learning algorithm for this application. To realise the full accuracy of this model and others suggested in future research it is critical that the dataset be expanded to avoid potential overfitting.

As of completion of this project, the models produced are not ready to be used as a way to predict domestic violence in an accurate way. However, the increased accuracy relative to alternative methods of predicting domestic violence indicates the possibility of a viable model. Further research would ideally need to be taken, using other algorithms such as XGBoost, Long Short-Term Memory (LSTM), or Gated Recurrent Unit (GRU). XGBoost can be used for both classification and regression tasks, and it is known for its scalability (Chen and He, 2024). This would be beneficial to further research as it would allow the research to be continued with additional data in the future which could aid in producing a higher accuracy within the model. LSTMs and GRU can also be used for classification and regression tasks, and both algorithms tend to handle long-term dependencies better than MLP, CNN, and RNN (Brownlee, 2020). Another possibility when conducting this research further would be to combine LSTMs and CNNs in a hybrid model which would aid with the vanishing gradient known to occur in neural networks (Brownlee, 2020). Additionally, it would be beneficial to collect more data, especially data from leagues and events that are not as major as the Premier League, Championship, World Cup, FA Cup, and Euros. This would allow for additional training and testing data, which should improve the model’s accuracy, as not only will Sunderland have a higher amount of data, it would allow for other areas of the Northeast to be included in the dataset used for testing and training.

# References

AbdulRaheem, M., Bamidele Awotunde, J., Adeniyi, A. E., Oladipo, I. D., and, Adekola, S. O. (2022) ‘Weather prediction performance evaluation on selected machine learning algorithms’, *IAES International Journal of Artificial Intelligence (IJ-AI)* Vol. 11, No. 4, December 2022, pp. 1535~1544

Alikhademi, K. *et al.* (2021) ‘A review of predictive policing from the perspective of fairness’, *Artificial Intelligence and Law*, 30(1), pp. 1–17. doi:10.1007/s10506-021-09286-4.

Amidi, A., and Amidi, S (2018) Recurrent Neural Networks Cheat sheet. Available: <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks> (Accessed: 10 December 2024)

Banerjee, S., (2024) Exploring the Power and Limitations of Multi-Layer Perceptron (MLP) in Machine Learning. Available: <https://shekhar-banerjee96.medium.com/exploring-the-power-and-limitations-of-multi-layer-perceptron-mlp-in-machine-learning-d97a3f84f9f4> (Accessed: 10 December 2024)

Barbedo, J. G. A. (2018) Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture. Volume 153, October 2018, Pages 46-53. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0168169918304617> (Accessed 09 March 2025)

Bengio, Y., Simard, P., and Frasconi, P. (1994) "Learning long-term dependencies with gradient descent is difficult," in *IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 157-166, March 1994*, doi: 10.1109/72.279181.

Berk, R. A. (2021) Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement. *Annual Review of Criminology.* Vol 4: 209-237.

Bhuiya, S. (2020) Disadvantages of CNN models. *Medium*. Available: <https://sandeep-bhuiya01.medium.com/disadvantages-of-cnn-models-95395fe9ae40> (Accessed: 20 March 2025)

Borovykh, A., Bohte, S., & Oosterlee, C. W. (2017). *Conditional time series forecasting with convolutional neural networks*. arXiv preprint arXiv:1703.04691.

Brownlee, J. (2020) How to Develop LSTM Models for TIme Series Forecasting. Machine Learning Mastery. Available: <https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/> (Accessed: 18 March 2025)

Brownlee, J. (2020) Understand the Impact of Learning Rate on Neural Network Performance. Machine Learning Mastery. Available: <https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/> (Accessed: 18 March 2025)

Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. August 13-17, 2016, San Francisco, CA, USA. Available: <https://dl.acm.org/doi/pdf/10.1145/2939672.2939785> (Accessed: 08 March 2025)

Chen, T. and Hu, T. (2024) xgboost: eXtreme Gradient Boosting. Available: <https://cran.ms.unimelb.edu.au/web/packages/xgboost/vignettes/xgboost.pdf> (Accessed: 08 March 2025)

Chung, J. and Teo, J. (2023) ‘Single classifier vs. ensemble machine learning approaches for mental health prediction’, *Brain Informatics*, 10(1). doi:10.1186/s40708-022-00180-6.

Cybenko, G. (1989). *Approximation by superpositions of a sigmoidal function*. Mathematics of Control, Signals, and Systems.

Dar, N (2013) A study on domestic violence; unveiling the scars of the victim women. *International Journal of Physical and Social Sciences*, 3(12),342-358

Delasgmit, W. H., and Manry, M. T. (2006) *Recent Developments in Multilayer Perceptron Neural Networks.* Proceedings of the 7th Annual Memphis Area Engineering and Science Conference.

Dian, L., Tao H., Cheng, E., Zhu, J., Gao, C. (2017) Deep Convolutional Neural Networks for Spatiotemporal Crime Prediction. *Proceedings of the International Conference on Information and Knowledge Engineering (IKE); Athens*, (2017).

Fausett, L., (1994) *Fundamentals of Neural Networks*. Prentice Hall, Englewood Cliffs, NJ.

Forsdike, K., O’Sullivan, G. and Hooker, L. (2022) ‘Major sports events and domestic violence: A systematic review’, *Health & Social Care in the Community*, 30(6). doi:10.1111/hsc.14028.

Gerber, M.S. (2014) ‘Predicting crime using Twitter and kernel density estimation’, *Decision Support Systems*, 61, pp. 115–125. doi: 10.1016/j.dss.2014.02.003.

Griffard, M. (2019) A Bias-Free Predictive Policing Tool?: An Evaluation of the NYPD's Patternizr. Fordham Urban Law Journal. Vol 47:1. pp: 43-83. Available: [https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2779 context=ulj](https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2779&context=ulj) (Accessed: 13 December 2024)

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., and Chen, T. (2018) Recent Advances in Convolutional Neural Networks. *Pattern Recognition*. Vol 77: 354-377.

Hamilton, M. (2021) ‘Predictive policing through risk assessment’, *Predictive Policing and Artificial Intelligence*, pp. 58–78. doi:10.4324/9780429265365-4.

Hardyns, W., and Rummens, A. (2017) Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges. *Eur J Crim Policy Res* 24, 201–218.<https://doi.org/10.1007/s10610-017-9361-2> (Accessed: 13 December 2024)

Harris, C. (2010) Richmond, Virginia, Police Department Helps Lower Crime Rates with Crime Prediction Software. Available:<https://www.govtech.com/public-safety/richmond-virginia-police-department-helps-lower.html> (Accessed: 13 December 2024)

Hauke J., Kossowski T., Comparison of values of Pearson’s and Spearman’s correlation coefficient on the same sets of data. Quaestiones Geographicae 30(2), Bogucki Wydawnictwo Naukowe, Poznań 2011, pp. 87–93, 3 figs, 1 table. DOI 10.2478/v10117-011-0021-1, ISBN 978-83-62662-62-3, ISSN 0137-477X

Hirithik, B., Aakash, S. B., Hariharan, S., and Mathumathi, M. M. E. (2022) Crime Prediction and Analysis Using Multi-Layer Perceptron. *International Journal of Advances in Engineering and Management (IJAEM)*. Volume 4, Issue 4 Apr 2022, pp: 1061-1069 Available: <https://ijaem.net/issue_dcp/Crime%20Prediction%20and%20Analysis%20Using%20Multi%20LayerPerceptron.pdf> (Accessed: 03 January 2025)

Home Office (2012) *New definition of domestic violence*. Available: <https://www.gov.uk/government/news/new-definition-of-domestic-violence> (Accessed: 11 November 2024)

Hui, V., Constantino, R. E., and Lee, Y. J. (2023) Harnessing Machine Learning in Tackling Domestic Violence—An Integrative Review. *Int. J. Environ. Res. Public Health 2023*, 20(6), 4984; <https://doi.org/10.3390/ijerph20064984> (Accessed: 11 November 2024)

Jain, A. (2024) Problems with RNN. Medium. Available: <https://medium.com/@abhishekjainindore24/problems-with-rnn-9db2a0ba58f5> (Accessed: 18 March 2024)

Jaiswal, S. (2024) Multilayer Perceptrons in Machine Learning: A Comprehensive Guide. DataCamp. Available: <https://www.datacamp.com/tutorial/multilayer-perceptrons-in-machine-learning> (Accessed: 03 January 2025)

Jing, L., Gulcehre, C., Peurifoy, J., Shen, Y., Tegmark, M., Soljaˇci´c, M., and Bengio, Y. (2017) “Gated orthogonal recurrent units: On learning to forget,” arXiv preprint arXiv:1706.0276

Karistianis, G. *et al.* (2021) ‘Utilising text mining, data linkage and deep learning in police and health records to predict future offences in family and domestic violence’, *Frontiers in Digital Health*, 3. doi:10.3389/fdgth.2021.602683.

Khan, J.R., Siddiqui, F.A., Mahmood, N., Muhammad, S., and ul Arifeen, Q. (2019) ‘Predictive Policing: A Machine Learning Approach to Predict and Control Crimes in Metropolitan Cities.’ *University of Sindh Journal of Information and Communication Technology (USJ CT).* 3(1), 17-26. ISSN-E: 2523-1235, ISSN-P: 2521-5582

Kirby, S., Francis, B., & Flaherty, R. (2014). Can the FIFA World Cup football (soccer) tournament be associated with an increase in domestic abuse? *Journal of Research in Crime and Delinquency*, 51(3), 259–276. 10.1177/0022427813494843

Keerthika, V., Geetha, A., Deepak Raj, D. M. (2024) PREDICTIVE CRIME ANALYSIS: Statistical Approach to Forecast Crime Hotspots Using Recursive Neural Network in Deep Learning. *2024 Second International Conference on Advances in Information Technology (ICAIT-2024).* Available: <https://ieeexplore.ieee.org/abstract/document/10690551> (Accessed: 10 December 2024)

Kounadi, O., Ristea, A., Araujo, A., and Leitner, M. (2020). A systematic review on spatial crime forecasting. Crime Sci. 9, 1–22.

Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J. (2022) A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems* (Volume: 33, Issue: 12, December 2022)

Lilhore, U.K., Dalal, S., Faujdar, N., Margala, M., Chakrabarti, P., Chakrabarti, T., Simaiya, S., Kumar, P., Thangaraju, P., Velmurugan, H. (2024) Hybrid CNN-LSTM model with efficient hyperparameter tuning for prediction of Parkinson’s disease. *Sci Rep* 13, 14605 (2023). https://doi.org/10.1038/s41598-023-41314-y

Lin, M., Chen, Q., Yan, S. (2014) Network in network. *Proceedings of the International Conference on Learning Representations (ICLR).*

Loeffler, C., and Flaxman, S. (2018). Is gun violence contagious? A spatiotemporal test. *J. Quant. Criminol.* 34, 999–1017. doi: 10.1007/s10940-017-9363-8

López-Úbeda, P. *et al.* (2021) ‘Automatic Medical Protocol Classification Using Machine Learning Approaches’, *Computer Methods and Programs in Biomedicine*, 200, p. 105939. doi: 10.1016/j.cmpb.2021.105939.

Lum, K. and Isaac, W. (2016) Predictive policing reinforces police bias. Human Rights Data Analysis Group

Meijer, A. and Wessels, M. (2019) ‘Predictive policing: Review of benefits and drawbacks’, *International Journal of Public Administration*, 42(12), pp. 1031–1039. doi:10.1080/01900692.2019.1575664.

Meskela, T. E., Afework, Y. K., Ayele, N. A., Teferi, M.W., Mengist, T.B (2020) Designing Time Series Crime Prediction Model using Long Short-Term Memory Recurrent Neural Network. *International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878, Volume-9 Issue-4, November 2020

Mishra, m. (2023) The Learning Rate: A Hyperparameter That Matters. Medium. Available: <https://mohitmishra786687.medium.com/the-learning-rate-a-hyperparameter-that-matters-b2f3b68324ab> (Accessed: 18 Mar 2025)

Neville, F. (2015) ‘Preventing violence through changing social norms’, in P.D. Donnelly and C.L. Ward (eds.) *Oxford Textbook of Violence Prevention: Epidemiology, evidence, and policy*. Oxford, United Kingdom: Oxford University Press.

Office for National Statistics (2023) *Crime and justice*, *Crime and justice - Office for National Statistics*. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice> (Accessed: 07 November 2024).

Pearsall, B. (2010). Predictive policing: The future of law enforcement. *National Institute of Justice Journal*, 266(1), 16–19.

Petneházi, G. (2019). "Recurrent neural networks for time series forecasting." *arXiv preprint arXiv:1901.00069*.

Perry, W., McInnis., B, Price, C., Smith, S., and Hollywood, J. (2018) Predictive Policing: the role of crime forecasting in law enforcement operations. RAND Corporation, Tech. rep

Popescu, M., Perescu-Popescu, L., Balas, V. E., Mastorakis, Nk. (2009) Multilayer Perceptron and Neural Networks. *WSEAS TRANSACTIONS on CIRCUITS and SYSTEMS.*

Protasiewicz, J. (2025) Python AIL Why is Python So Good for Machine Learning? Available: <https://www.netguru.com/blog/python-machine-learning> (Accessed 20 February 2025)

Rana, M. and Bhushan, M. (2022) ‘Machine learning and deep learning approach for medical image analysis: Diagnosis to detection’, *Multimedia Tools and Applications*, 82(17), pp. 26731–26769. doi:10.1007/s11042-022-14305-w.

Rahman, R., Khan, M. N. A., Sara, S. S., Rahman, M. A., and Khan, Z. I. (2023) A comparative study of machine learning algorithms for predicting domestic violence vulnerability in Liberian women. *BMC Women's Health (2023)* 23:542 <https://doi.org/10.1186/s12905-023-02701-9> (Accessed 20 January 2025)

Ramchoun, H., Idrissu, M. A. J., Ghanou, Y., and Ettaouil, M. (2016) Multilayer Perceptron: Architecture Optimization and Training. *Modeling and Scientific Computing Laboratory, Faculty of Science and Technology*. University Sidi Mohammed Ben Abdellah, Fez, Morocco. Available: <https://reunir.unir.net/bitstream/handle/123456789/11569/ijimai20164_1_5_pdf_30533.pdf?sequence=1&isAllowed=y> (Accessed: 10 December 2024)

Ratcliffe, J. H. (2004). The hotspot matrix: A framework for the spatio‐temporal targeting of crime reduction. *Police Practice and Research*, [5](https://doi.org/10.1080/1561426042000191305)([1](https://doi.org/10.1080/1561426042000191305)), 5–23. doi:10.1080/1561426042000191305

Refuge (2024) *Facts and statistics*, *Refuge*. Available at: https://refuge.org.uk/what-is-domestic-abuse/the-facts/ (Accessed: 07 November 2024).

Richards, L. (2009) *Domestic Abuse, Stalking and Harassment and Honour Based Violence (DASH, 2009) Risk Identification and Assessment and Management Model.* Available:<https://reducingtherisk.org.uk/wp-content/uploads/2022/08/DASH-2009.pdf> (Accessed: 08 November 2024)

Ringland, C. (2017) *The Domestic Violence Safety Assessment Tool (DVSAT) and intimate partner repeat victimisation*. Sydney: NSW Bureau of Crime Statistics and Research.

Sabo, D., & Runfola, R. (1980). *Jock: Sports and male identity*. Prentice‐Hall.

Safelives (2024) *How long do people live with domestic abuse?*, *SafeLives*. Available at: <https://safelives.org.uk/about-domestic-abuse/what-is-domestic-abuse/facts-and-figures/length-of-abuse/> (Accessed: 07 November 2024).

Salehinejad, H., Snakar, S., Barfett, J., Colak, E., Valaee, S. (2018) Recent Advances in Recurrent Neural Networks. Available: <https://arxiv.org/abs/1801.01078> (Accessed: 11 December 2024)

Shah, J., Vaidya, D. and Shah, M. (2022) ‘A comprehensive review on multiple hybrid deep learning approaches for stock prediction’, Intelligent Systems with Applications, 16, p. 200111. doi: 10.1016/j.iswa.2022.200111.

Sharkawy, A. (2020) Principle of Neural Network and Its Main Types: Review. Journal of Advances in Applied & Computational Mathematics 7(1):8-19. August 2020. Available: <https://www.researchgate.net/publication/343837591_Principle_of_Neural_Network_and_Its_Main_Types_Review> (Accessed: 10 December 2024)

Sheremetov, D., and Bitkina, A. (2023) The Python Advantage: Why it’s the Top Choice for AI and ML. Available:<https://onix-systems.com/blog/python-is-best-for-ai-ml-and-deep-learning> (Accessed: 11 November 2024)

Swallow, J. *et al.* (2017) *An exploratory study of women’s experiences regarding the interplay between domestic violence and abuse and sports events*. thesis. University of Chester.

Thornton, S. (2017) ‘Police attempts to predict domestic murder and serious assaults: Is early warning possible yet?’, *Cambridge Journal of Evidence-Based Policing*, 1(2–3), pp. 64–80. doi:10.1007/s41887-017-0011-1.

Towers, S., Chen, S. and Malik, A. (2016) ‘Factors influencing temporal patterns in crime in a large American city; a predictive analytics perspective’, *SSRN Electronic Journal* [Preprint]. doi:10.2139/ssrn.2833583.

Unadkat, S. B., Ciocoiu, M, M., and Medsker, L. R. (2001) *Chapter 1: Introduction*. In Medsker, L. R. and Jain, L. (2001) *Recurrent Neural Networks: Design and Application*. CRC Press, Washington D.C.

United Kingdom Government (2022) Domestic Abuse Act 2021. Available: <https://www.legislation.gov.uk/ukpga/2021/17/part/1> (Accessed: 30 November 2024)

Unwin, A. (2020). Why Is Data Visualization Important? What Is Important in Data Visualization? Harvard Data Science Review, 2(1). <https://doi.org/10.1162/99608f92.8ae4d525> (Accessed: 15 March 2025)

Vabalas, A., Gowen, E., Poliakoff, E., and Casson, A.J. (2019) Machine learning algorithm validation with a limited sample size. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0224365> (Accessed: 10 March 2025)

Visin, F., Kastner, K., Cho, K., Matteucci, M., Courville, A., and Bengio, Y. (2015) “Renet: A recurrent neural network-based alternative to convolutional networks,” arXiv preprint arXiv:1505.00393.

Walczak, S. (2021) Predicting Crime and Other Uses of Neural Networks in Police Decision Making. *Forensic and Legal Psychology.* Vol 12. Available: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.587943/full> (Accessed: 10 December 2024)

Wang, T., Rudin, C., Wagner, D., and Sevieri, R. (2013) Learning to Detect Patterns of Crime. In: Blockeel, H., Kersting, K., Nijssen, S., Železný, F. (eds) *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2013. Lecture Notes in Computer Science*, vol 8190. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40994-3\_33

Williams, I. (2018) *Machine learning algorithms for crime prevention and predictive policing*. PhD Thesis, Cardiff University (Accessed: 10 November 2024)

World Health Organisation (2021) Violence against women, World Health Organization. Available at:<https://www.who.int/news-room/fact-sheets/detail/violence-against-women> (Accessed: 10 November 2024).

Wu, J. (2017) Introduction to Convolutional Neural Networks. LAMDA Group. National Key Lab for Novel Software Technology. Available: <https://cs.nju.edu.cn/wujx/paper/CNN.pdf> (Accessed: 02 December 2024)

# Appendix A. Research Proposal

**A deep learning approach to predicting domestic violence instances in the United Kingdom during major sports events.**

**1.1 Background**

Intimate partner violence (or domestic violence) is a “global epidemic” (Forsdike, O’Sullivan, and Hooker, 2022) with the most recent Office for National Statistics figures (2023) showing that one in four women and one in seven men will be a victim of domestic abuse in their lifetime. The World Health Organisation (WHO) defines domestic violence as “behaviour by an intimate partner or ex-partner that causes physical, sexual or psychological harm, including physical aggression, sexual coercion, psychological abuse and controlling behaviours” (WHO, 2021).

Despite the likelihood of domestic violence occurring within society, tools that can predict future instances are lacking (Karystianis et al, 2021). Through the growth and availability of big data as well as the creation of predictive policing throughout recent years, it is possible to apply machine learning techniques to predict domestic violence instances, similarly to how the police use predictive policing to intercept illicit drugs and weapons (Karystiantis et al, 2021). However, predictive policing in instances of illicit drugs and weapons has been to varying degrees of success (Karystiantis et al, 2021).

Approaches of predictive analysis surrounding crime often apply machine learning to Tweets (Gerber, 2013) or police records (Towers et al, 2018) in order to forecast criminal activity or hot spots. The accuracy of these models ranges from 70-81%. However, recent efforts regarding the prediction of domestic violence have had limited success, with some models producing weak predictive accuracy (Karystiantis et al, 2021). In one case, 89% of fatal cases were incorrectly labelled as not a high risk (Thornton, 2017). Ringland (2018) attempted to predict if there would be repeat victimisation of survivors of domestic abuse by asking binary questions. If a survivor answered ‘yes’ to 12 or more questions, this was an indicator that repeat victimisation would occur, however, the predictive results were poor (Ringland, 2018). Several papers reviewed revealed that current predictive methods for identifying domestic violence risk before it occurs does not display a usable accuracy for policing methods - with some models producing a false positive rate of up to 99% (Thornton, 2017). There are several reasons as to why these methods lacked accuracy. According to Swallow (2017), those who experience intimate partner violence may have difficulty in defining what constitutes domestic violence, therefore being unable to answer a risk assessment in a factual way.

Violence and sport go hand-in-hand, with player violence or crowd violence being widely acknowledged and focused upon (Neville, 2015). However, there have been links made to sporting culture and domestic violence since the 1980’s (Sabo and Runfola, 1980). Evidence points to an increased rate of domestic violence perpetrated by male athletes, as well as some evidence confirming that domestic violence increases throughout wider society during sporting events (Neville, 2015). Neville (2015) notes that the ‘holy trinity’ of sport, alcohol, and hegemonic masculinity are often used as a way to explain the increase in domestic violence where sporting events are concerned. Kirby, Francis, and O’Flaherty (2013) conducted quantitative analysis, using Poisson and negative binomial regression models to investigate if empirical evidence exists to support the view that football (soccer) tournaments are associated with a rise in reported domestic abuse incidents. To do this, Kirby, Francis and O’Flaherty (2013) focused upon the FIFA World Cup Tournaments between 2002 and 2010 and domestic abuse incidents reported to a police force in the North West of England. The study found a 26% rise in instances of domestic violence when the England national team won or drew on match day, however, it increased to 38% when England lost (Kirby, Francis and O'Flaherty, 2013). Another trend was apparent that with each new tournament, reports of domestic abuse incidents increased in frequency (Kirby, Francis, and O’Flaherty, 2013). However, this research encompasses only three tournaments - due to the FIFA World Cup being held every four years. More research into this area would be beneficial to confirm a correlation, especially with other types of sport. Previous papers have focused upon domestic violence and football, further investigation should add other sports into the equation in order to confirm a correlation between domestic violence and sports in general. This project aims to delve into a wide range of popular sporting events, including, but not limited to: football (soccer), rugby, ice hockey, American football, basketball, and baseball. Thus, confirming if there is a correlation between all sports and domestic violence, as well as investigating if the correlation is only present for specific sports.

At present, there is a potential gap in research that has been identified - the creation of a predictive model to aid police in pursuing perpetrators of domestic violence. This project will focus on the prediction of domestic violence occurring during sporting events due to the significant research that instances of reporting increase during these events.

**2.1 Motivation for the Project**

Although there is some research into this area, most research comes from a social science field, focussing upon qualitative methods to collect and display data. Research that has been conducted from a computer science standpoint about predicting domestic violence has not been accurate, with models obtaining poor predictive results or high false positives. This project aims to change this by producing a model that will provide accurate predictive results regarding the likelihood of domestic violence instances for a wide range of sporting events by utilising Convolutional Neural Networks (CNNs).

Using major sporting events as a marker for the likelihood of domestic violence instances has been researched in a quantitative manner, however, using these major sporting events to predict these crimes is an area that is lacking. Kirby, Francis and O’Flaherty (2013) found trends that with each FIFA tournament, the reports of domestic abuse incidents increased. However, as noted previously, the research only encompasses three tournaments (2002, 2006, and 2010) and further investigation into this would be beneficial with more recent data which is something that this project aims to do with the use of historical, as well as the most recent public crime data.

A vast amount of previous research into this topic has been either qualitative experiences recorded from a social science point of view, or quantitative research based upon risk assessments and small samples of data. This study will use a wider range of data, both historical and contemporary, as well as using machine learning to train a model to predict the likelihood of violence based on location.

By successfully creating this project and producing accurate results, this project would prove that predicting instances of domestic violence, by location and during major sporting events, is possible and plausible. As well as this, it would also prove a direct link between the two variables of domestic violence and sport, with successful predictions implying that the variables are dependent. Current research suggests that this is not possible with Thornton (2017) noting that being able to predict “domestic violence based on intelligence from prior police contacts does not appear possible at present.” Predictive policing, in general, has been in place for many years, with the police force using predictive tools to counter drug and weapon crimes to varying degrees of success (Karystiantis, 2021).

This research project will aim to create a model that predicts the likelihood of domestic violence occurrences during major sporting events using deep learning. Creating this project successfully would be a meaningful contribution to the ways in which crime can be forecast and police can allocate resources.

**3.1 Research Question**

The research question for this project is as follows:

*Is it possible to accurately predict instances of domestic violence during sporting events and see which areas of the country the police are most needed?*

**4.1 Research Aim**

The project aims to provide a predictive model of domestic violence instances by location in the United Kingdom in relation to major sporting events through historical police records marked as domestic violence instances and timelines of major sporting events. Datasets will need to be combined to create a model for domestic violence prediction that encompasses sporting events and for the aim to be completed.

**5.1 Research Objectives**

**5.1.1 Historical Police Records Dataset**

**Description:** Locating, downloading, importing, and classifying data. Preparing for training and testing. Data separated into a training and testing set. Full copy created to use in real time integration.

**Acceptance Criteria:** Data imported and denoised, set up and ready for training the model.

**Time Estimate**: 3 Weeks

**5.1.2 Historical Major Sporting Events Dataset**

**Description**: Locating, downloading, importing, and classifying data. Preparing for training and testing. Data separated into a training and testing set. Full copy created to use in real time integration. Find API for sporting events for real-time integration later on.

**Acceptance Criteria:** Data imported and denoised, set up and ready for training the model.

**Time Estimate**: 3 Weeks

**5.1.3 Combining the Datasets**

**Description**: Joining the datasets together to see where and when major sporting events occurred.

**Acceptance Criteria:** Datasets successfully combined.

**Time Estimate**: 1 Week

**5.1.4 Model Training Using the Dataset**

**Description:** Using the model to train the dataset.

**Acceptance Criteria:** Model trained.

**Time Estimate:** 2 Weeks

**5.1.5 Real Time Major Sporting Event Integration**

**Description**: Importing data from Sporting Event API.

**Acceptance Criteria**: Data stored and types match the current data.

**Time Estimate**: 4 Weeks

**5.1.5 Accuracy Testing**

**Description**: Testing accuracy at different stages of the project. These will be compared.

**Acceptance Criteria**: Accuracy values included in dissertation analysis.

**Time Estimate**: 2 Weeks

**5.1.6 Dissertation Writing**

**Description**: Writing sections of dissertation and completing throughout the research project.

**Acceptance Criteria**: Final version completed and checked - handed into eLP.

**Time Estimate**: 8 Weeks

**6.1 Literature Review**

Meijer and Wessels (2019) note that over the past few years, an increasing number of police forces globally have adopted software that guides either decision making via statistical data. This has come to be known as ’predictive policing’ (Meijer and Wessels, 2019).

Predictive policing can be defined as the following::

“the use of historical data to create a forecast of areas of criminality or crime hotspots, or high-risk offender characteristic profiles that will be one component of police resource allocation decisions. The resources will be allocated with the expectation that, with targeted deployment, criminal activity can be prevented, reduced, or disrupted.”

(Ratcliffe, 2019: 349)

Hamilton (2021) notes that there are three main types of predictive policing: forecasting hotspots; predicting who is likely to become a victim; and predicting who is likely to offend. The most common form is forecasting hotspots, which is what this research shall investigate for domestic violence. Predicting who is likely to become a victim or an offender is not as researched as finding hotspots of crime is more practical for the police (Hamilton, 2021).

In terms of finding hotspots of crime, police departments can analyse historical statistical data to predict in which geographic areas instances of criminal activity may occur (Meijer and Wessels, 2019). It has been found that in some areas where predictive policing is used, crime decreases (Meijer and Wessels, 2019). An example of this would be Richmond, Virginia, USA in 2006 where predictive policing was used to forecast where gun crime would occur on New Year’s Eve (Pearsall, 2010). Surveillance routes were adapted in line with these predictions and it was considered a success, with gun crimes decreasing 49% and 246% more weapons seized (Pearsall, 2010). Only one-third as many police officers were deployed compared to previous years, however, effectiveness and efficiency increased regardless of the decrease in police presence due to predictive policing, saving the police around $15,000 in overtime pay (Harris, 2010)

There are multiple ways to analyse the data and forecast crime - one of these is via machine learning. Machine learning (ML) is the term used for algorithms that produce outcomes that are based on patterns of data (Alikhademi et al, 2021) - usually from intelligence gained from ‘big data’ (Hardyns and Rummens, 2018). Eulluri, Madalapu and Roy (2019) used ML - both traditional algorithms and deep learning - to predict the effect of weather on the likelihood of crimes committed in New York City. Both traditional algorithms and deep learning algorithms had a high correlation in relation to weather and crime (Eulluri, Madalapu, and Roy, 2019). Khan et al (2019) also used ML to predict crimes, however, their research was focused in metropolitan cities using the Karachi region in Pakistan as their dataset. Khan et al’s (2019) research used R and WEKA to predict the likelihood of street crime such as phone snatching and theft, pulling together large amounts of data and producing a predictive model with an accuracy of around 70%. This model used clustering and classification machine learning algorithms, such as K-means and Naive Bayseian, to predict the crime rate for the following week (Khan et al., 2019). In their model, Khan et al (2019) used 80% of their collected data for training, saving 20% for testing purposes, which they found to be best for accurate results. Another example of ML being utilised for predictive policing is Williams (2018) research into the effectiveness of predictive policing in the case of re-offending in the Dyfed-Powys Police force. The model created by Williams (2018),l used Random Forest and XGBoost algorithms, as well as Feedforward Neural networks to predict the likelihood of an offender re-offending chance. Williams (2018) found that there was a clear case in favour of predictive policing being used to prevent crime and re-offending.

Although evidence has been found that predictive policing works, ML can reflect and entrench biases of humans, which has led to discussions and research regarding the fairness of predictive policing and the threat of bias towards certain groups in society (Alikhademi et al, 2021). Lum and Isaac (2016) note that communities that have a higher police presence will naturally have a higher arrest rate, therefore the dataset will reflect this Lum and Isaac (2016) argue that it appears to reflect a higher crime rate, however, what it truly reflects is greater police attention Studies have shown that this greater police attention - generally in tends to lean towards ethnic minorities and neighbourhoods with lower socio-economic status (Lum and Isaac, 2016). This was found to be common in corrupt police forces that have entrenched within them institutional racism and classism, where predictive policing being utilised exacerbates the corruption (Richardson et al., 2019). Perry et al (2018) note that by using historical datasets, historical bias can also occur, which ignores the fact that society changes - for example, the motivations, means and types of perpetrators can change over time. This must be taken into consideration, otherwise the historical biases could create a feedback loop (Perry et al, 2018).

In order to combat this issue, a few solutions have been suggested to produce algorithmic fairness. Mehrabi et al. (2019) defines algorithmic fairness as “absence of any prejudice or favouritism toward an individual or a group based on their inherent or acquired characteristics.” Alikhademi et al (2021) note that the methods suggested to improve algorithms within the criminal justice sphere can be split into two main groups: algorithm designs being thought about qualitatively in advance of bias being detected or programmatic interventions being placed in the algorithm to detect and/or correct the bias. A suggestion from the first camp of thought is Altman et al (2018) proposing that algorithm designers would have to find areas in the algorithm design that could be subject to bias, for example, asking questions such as ‘what would the results look like if race was excluded from the training data?’ Those who fall into the second category have suggested modifications of the input provided to predictive policing algorithms. Ensign et al (2017) investigated this using PredPol - which was used by the LAPD up until relatively recently. PredPol is designed to learn from new crime data that is gathered by patrolling officers so that the police can adapt (Alikhademi et al, 2021). They found that the algorithm was sending police officers back to the site they just patrolled, due to the new data input, leading to a vicious cycle. Ensign et al (2017) proposed a new solution that alters the way in which data is added so that the more likely the police are sent to a given district, the less likely it is that it is incorporated into the data. Without the addition proposed by Ensign et al (2017), PredPol’s predictions did not accurately predict the true crime rate - something that is concerning as the LAPD were actively using PredPol at the time. Other police forces have developed different predictive policing methods to attempt to eliminate bias. An example of this is Patternizr, a predictive policing application created by data scientists at the New York Police Department (NYPD), that uses the *modus operandi* (M.O.) of the crime, rather than the location or people, allowing the data to be less bias (Griffard, 2019). The M.O. can be defined as a “set of habits that the offender follows, and is a type of motif used to characterise the pattern” (Wang et al, ). Although the developers of Patternizr note that the application is fair and an accurate predictive policing tool, there have been concerns raised as to whether this is true (Griffard, 2019). However, unlike PredPol, this is due to the application not being independently assessed by outside researchers as Patternizr has currently only been assessed by the NYPD (Griffard, 2019).

From the research into bias conducted above, this project will aim to limit bias by focusing upon the M.O. and crime of the perpetrator, rather than the type of person or location as the mitigating factor. Location will still be required to be in the dataset so that the places where the offences occur can be predicted. An anti-bias algorithm could also be used to negate bias from the public policing data collected for this project, however, it is yet to be decided if this would benefit the project.

The majority of predictive methods where domestic violence is concerned, are focused upon using risk assessments to predict the likelihood of instances. There have been recent efforts regarding the prediction of domestic violence for the police, however, these have had limited success, with some models producing weak predictive accuracy (Karystiantis et al, 2021). Currently there is one tool used in the United Kingdom to predict domestic violence (Hamilton, 2021). This is called the Domestic Abuse, Stalking, Harassment and Honour-Based Violence (DASH) instrument and it ranks risk based on predictive items regarding current situations, dependents, domestic violence history, and certain characteristics of the abusers (Hamilton, 2021). DASH is a risk assessment based tool that was created by Laura Richards in 2009 and is widely used in police forces across the United Kingdom. The domestic violence section of the instrument is 27 questions and places the victim in a standard, medium or high risk category depending upon their answers (Richards, 2009). However, Thornton (2017) investigated the use of this method in Thames Valley and found that 89% of fatal cases were incorrectly labelled as medium risk when they were, in fact, high risk. If these had been correctly labelled as high risk, there is a possibility that these cases would not have led to a fatality (Thornton, 2017). Ringland (2018) also aimed to predict if there would be repeat victimisation of survivors of domestic abuse by asking yes or no questions, which provided poor predictive results. Several papers reviewed show that the current predictive methods for identifying domestic violence do not display a usable accuracy for policing methods - with some models producing a false positive rate of up to 99% (Thornton, 2017).

One of the major issues with predicting domestic abuse and intimate partner violence is that it often occurs away from the public places and, most often, the only witnesses to the abuse are those involved or those who live in the same household (Dar, 2013). It is common for those who have experienced domestic abuse to have difficulty in identifying the behaviours as abusive, more likely if the abuse is financial or psychological, and some survivors of domestic abuse are reluctant to report (Swallow, 2017). ONS (2023) notes that only 18.9% of women who had experienced intimate partner abuse in 2022/23 reported the abuse to the police. Although the police attend a domestic abuse related call every 30-seconds, it is estimated that less than 24% of domestic abuse crimes are reported (Refuge, 2024). SafeLives (2024) notes that on average, victims of domestic abuse will experience 50 incidents, ranging from physical to psychological, before reporting. This can lead to the data being skewed and not being necessarily accurate within the context of wider society. Although the dataset would be incomplete in regards to wider society, it is necessary to use this dataset as any attempt to account for this, such as oversampling, could lead to further bias as with PredPol (Ensign et al, 2017).

There have been links made to sporting culture and domestic violence since the 1980’s (Sabo and Runfola, 1980). Evidence points to an increased rate of domestic violence perpetrated by male athletes, as well as some evidence confirming that domestic violence increases throughout wider society during sporting events (Neville, 2015).

As mentioned previously in this proposal, Kirby, Francis, and O’Flaherty (2013) conducted quantitative analysis, using Poisson and negative binomial regression models, to investigate if evidence exists to support the view that football (soccer) tournaments are linked with a rise in domestic abuse incidents reported to the police. Kirby, Francis and O’Flaherty (2013) focused upon FIFA World Cup Tournaments between 2002 and 2010 (2002, 2006, 2010), comparing the report rate of incidents to those reported when the tournaments were not occurring. The study found that domestic violence rose by 26% when the England national team won or drew on match day, however, it increased to 38% when England lost (Kirby, Francis and O'Flaherty, 2013). Another trend was apparent that with each new tournament, reports of domestic abuse incidents increased in frequency (Kirby, Francis, and O’Flaherty, 2013). However, this research encompasses only three tournaments - due to the FIFA World Cup being held every four years. More research into this area would be beneficial to confirm a correlation with more contemporary data. This project aims to investigate this further, not only with more recent data, but also with additional sports included to understand if the link to domestic violence is pertinent throughout all types of sport.

**7.1 Research Approach / Methodology**

Research will begin with a rigorous exploration of historical and contemporary research surrounding domestic violence prediction, as well as domestic violence surrounding major sporting events. Following this, an experimental model with the goal of yielding predictions of domestic violence instances during major sporting events will be created. The output of this model intends to indicate the location of these instances, as well as their existence. Data used to train this model will be from public police and arrest records and public information regarding the dates of major sporting events in the UK.

Deep learning will be used to accurately predict instances of domestic violence by location, surrounding major sporting events. The deep learning algorithm that will be used is Convolutional Neural Networks (CNN). CNNs can be used for a multitude of tasks - from image recognition to natural language processing and can process large amounts of data (Gu et al, 2018). This project will require a significant amount of data so, therefore, CNNs are the most appropriate algorithm to use.

Python will be used as the core programming language for this project due to its varying libraries such as MatPlotLib, for data visualisation, and TensorFlow, for deep learning algorithms. Sheremetov and Bitkina (2023) note that, when it comes to artificial intelligence (AI), machine learning (ML), and deep learning (DL) algorithms, Python has “unparalleled versatility.” Python is adaptable, robust, and has an extensive collection of libraries and frameworks - making it the ideal language to use for DL (Sheremetov and Bitkina, 2023).

Studies into domestic violence prediction have been based on risk assessments and statistical analysis. By utilising machine learning, as in predictive policing, and the use of CNNs, the research project aims to accurately predict the geography of domestic violence during sporting events.

**7.2 Ethical Implications**

The data used in this project will be from publicly available data sources such as police records and an API for major sporting events. This data will not contain sensitive personal information, such as names or addresses, that could identify a person. All data will be anonymous due to this being the way the data that will be used in this project is presented in the public space.

All data will be stored securely on a personal PC which will not be moved or leave the house and is password protected. Files will be backed up to an external hard drive which will be stored securely, and password protected. Data will be destroyed once the research project has concluded, and marks awarded. All data not being present in the research will be removed from all storage devices above and thoroughly digitally disposed of. No data will be collected on paper.

**8.1 Task List**

|  |  |  |
| --- | --- | --- |
| **Activity** | **Duration (Weeks)** | **Month** |
| Clarify Research Question and Project Themes | 2 | August |
| Research Proposal and Refining Research Question/Project | 4 | September |
| Literature Review | 4 | September / October |
| Research Methods | 2 | October |
| Data Collection | 2 | October |
| Refining Data and Data Modelling | 6 | October / November |
| Creating Model | 4 | November / December |
| Training Model | 4 | December / January |
| Testing Model | 4 | January / February |
| Dissertation Writing | 8 | January / February / March |
| Viva Preparation | 3 | March |
| Final Checks | 2 | March / April |
| Hand-In | 1 | April |

**9.1 References**

Alikhademi, K. *et al.* (2021) ‘A review of predictive policing from the perspective of fairness’, *Artificial Intelligence and Law*, 30(1), pp. 1–17. doi:10.1007/s10506-021-09286-4.

Dar, N (2013) A study on domestic violence; unveiling the scars of the victim women. *International Journal of Physical and Social Sciences*, 3(12),342-358

Forsdike, K., O’Sullivan, G. and Hooker, L. (2022) ‘Major sports events and domestic violence: A systematic review’, *Health & Social Care in the Community*, 30(6). doi:10.1111/hsc.14028.

Gerber, M.S. (2014) ‘Predicting crime using Twitter and kernel density estimation’, *Decision Support Systems*, 61, pp. 115–125. doi:10.1016/j.dss.2014.02.003.

Griffard, M. (2019) A Bias-Free Predictive Policing Tool?: An Evaluation of the NYPD's Patternizr. Fordham Urban Law Journal. Vol 47:1. pp: 43-83. Available: [https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2779 context=ulj](https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2779&context=ulj)

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., and Chen, T. (2018) Recent Advances in Convolutional Neural Networks. *Pattern Recognition*. Vol 77 : 354-377.

Hardyns, W., and Rummens, A. (2017) Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges. *Eur J Crim Policy Res* 24, 201–218.<https://doi.org/10.1007/s10610-017-9361-2>

Harris, C. (2010) Richmond, Virginia, Police Department Helps Lower Crime Rates with Crime Prediction Software. Available:<https://www.govtech.com/public-safety/richmond-virginia-police-department-helps-lower.html>

Hamilton, M. (2021) ‘Predictive policing through risk assessment’, *Predictive Policing and Artificial Intelligence*, pp. 58–78. doi:10.4324/9780429265365-4.

Karistianis, G. *et al.* (2021) ‘Utilising text mining, data linkage and deep learning in police and health records to predict future offences in family and domestic violence’, *Frontiers in Digital Health*, 3. doi:10.3389/fdgth.2021.602683.

Khan, J.R., Siddiqui, F.A., Mahmood, N., Muhammad, S., and ul Arifeen, Q. (2019) ‘Predictive Policing: A Machine Learning Approach to Predict and Control Crimes in Metropolitan Cities.’ *University of Sindh Journal of Information and Communication Technology (USJ CT).* 3(1), 17-26. ISSN-E: 2523-1235, ISSN-P: 2521-5582

Kirby, S., Francis, B., & Flaherty, R. (2014). Can the FIFA World Cup football (soccer) tournament be associated with an increase in domestic abuse? *Journal of Research in Crime and Delinquency*, 51(3), 259–276. 10.1177/0022427813494843

Lum, K. and Isaac, W. (2016) Predictive policing reinforces police bias. Human Rights Data Analysis Group

Meijer, A. and Wessels, M. (2019) ‘Predictive policing: Review of benefits and drawbacks’, *International Journal of Public Administration*, 42(12), pp. 1031–1039. doi:10.1080/01900692.2019.1575664.

Neville, F. (2015) ‘Preventing violence through changing social norms’, in P.D. Donnelly and C.L. Ward (eds.) *Oxford Textbook of Violence Prevention: Epidemiology, evidence, and policy*. Oxford, United Kingdom: Oxford University Press.

Office for National Statistics (2023) *Crime and justice*, *Crime and justice - Office for National Statistics*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice (Accessed: 07 September 2024).

Pearsall, B. (2010). Predictive policing: The future of law enforcement. *National Institute of Justice Journal*, 266(1), 16–19.

Perry, W., McInnis., B, Price, C., Smith, S., and Hollywood, J. (2018) Predictive Policing: the role of crime forecasting in law enforcement operations. RAND Corporation, Tech. rep

Ratcliffe, J. H. (2004). The hotspot matrix: A framework for the spatio‐temporal targeting of crime reduction. *Police Practice and Research*, [5](https://doi.org/10.1080/1561426042000191305)([1](https://doi.org/10.1080/1561426042000191305)), 5–23. doi:10.1080/1561426042000191305

Refuge (2024) *Facts and statistics*, *Refuge*. Available at: https://refuge.org.uk/what-is-domestic-abuse/the-facts/ (Accessed: 07 September 2024).

Richards, L. (2009) *Domestic Abuse, Stalking and Harassment and Honour Based Violence (DASH, 2009) Risk Identification and Assessment and Management Model.* Available:<https://reducingtherisk.org.uk/wp-content/uploads/2022/08/DASH-2009.pdf>

Ringland, C. (2017) *The Domestic Violence Safety Assessment Tool (DVSAT) and intimate partner repeat victimisation*. Sydney: NSW Bureau of Crime Statistics and Research.

Sabo, D. , & Runfola, R. (1980). *Jock: Sports and male identity*. Prentice‐Hall.

Safelives (2024) *How long do people live with domestic abuse?*, *SafeLives*. Available at: https://safelives.org.uk/about-domestic-abuse/what-is-domestic-abuse/facts-and-figures/length-of-abuse/ (Accessed: 07 September 2024).

Sheremetov, D., and Bitkina, A. (2023) The Python Advantage: Why it’s the Top Choice for AI and ML. Available:<https://onix-systems.com/blog/python-is-best-for-ai-ml-and-deep-learning>

Swallow, J. *et al.* (2017) *An exploratory study of women’s experiences regarding the interplay between domestic violence and abuse and sports events*. thesis. University of Chester.

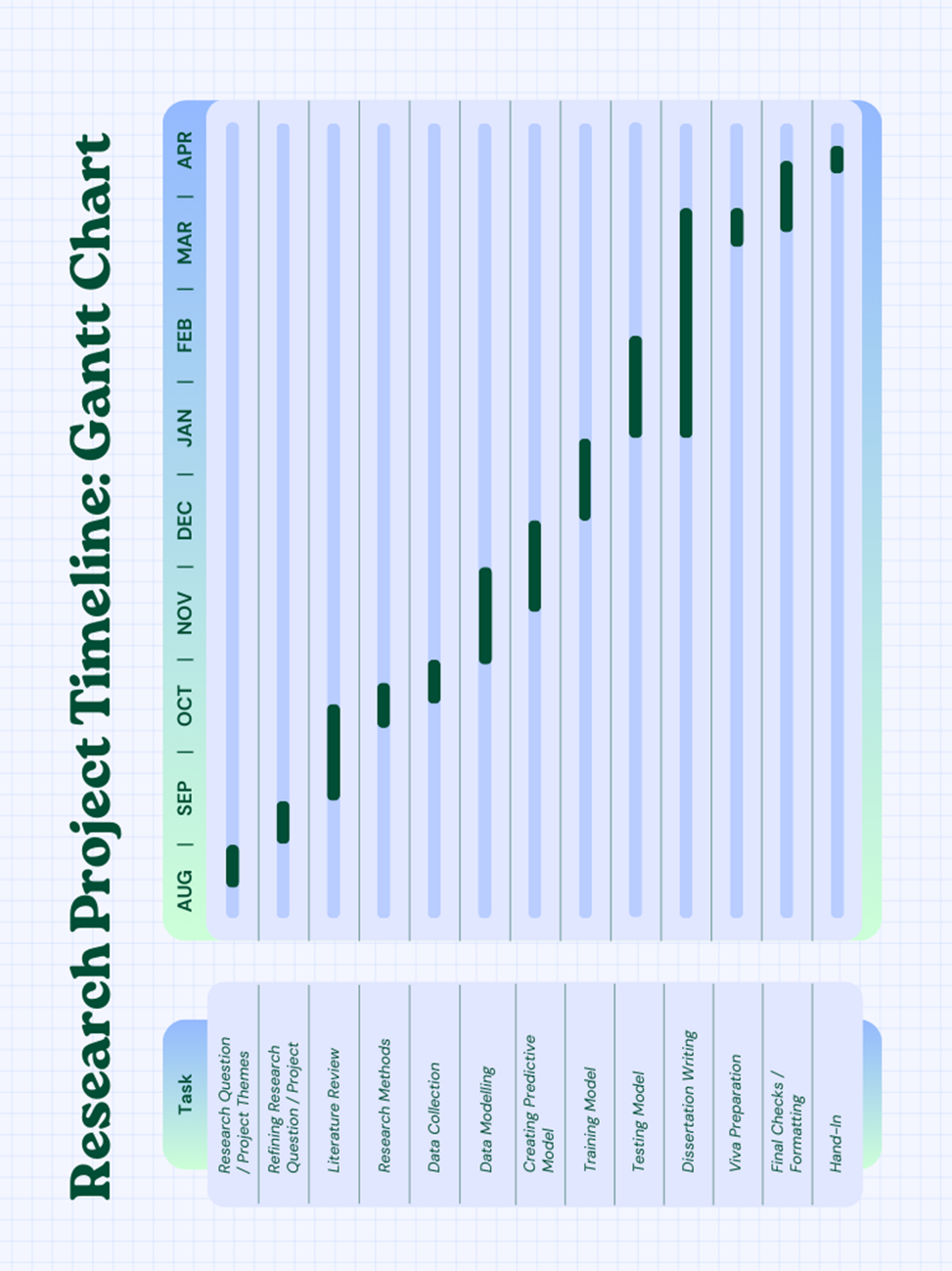
Thornton, S. (2017) ‘Police attempts to predict domestic murder and serious assaults: Is early warning possible yet?’, *Cambridge Journal of Evidence-Based Policing*, 1(2–3), pp. 64–80. doi:10.1007/s41887-017-0011-1.

Towers, S., Chen, S. and Malik, A. (2016) ‘Factors influencing temporal patterns in crime in a large American city; a predictive analytics perspective’, *SSRN Electronic Journal* [Preprint]. doi:10.2139/ssrn.2833583.

Williams, I. (2018) *Machine learning algorithms for crime prevention and predictive policing*. PhD Thesis, Cardiff University

World Health Organisation (2021) Violence against women, World Health Organization. Available at:<https://www.who.int/news-room/fact-sheets/detail/violence-against-women> (Accessed: 07 September 2024).

# Appendix B. Gantt Chart



# Appendix C. Code

*#importing libraries*

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

import seaborn as sns

from sklearn.model\_selection import train\_test\_split

from sklearn.preprocessing import StandardScaler

from sklearn.metrics import mean\_absolute\_error

import tensorflow as tf

from tensorflow import keras

from tensorflow.keras import layers

from sklearn.preprocessing import MinMaxScaler

from sklearn.model\_selection import train\_test\_split

from scipy.stats import pearsonr

*#loading data*

*#loading police data*

dv\_police\_df = pd.read\_csv('DV\_Data\_NP.csv')

*#loading football data*

fb\_world\_cup\_df = pd.read\_csv('FootballData\_WorldCup.csv')

fb\_premier\_df = pd.read\_csv('FootballData\_PremierLeague.csv')

fb\_championship\_df = pd.read\_csv('FootballData\_Championship.csv')

fb\_fa\_cup\_df = pd.read\_csv('FootballData\_FACup.csv')

fb\_euro\_df = pd.read\_csv('FootballData\_Euro.csv')

*#cleaning and dropping NAs from dataframe*

dv\_police\_df\_cleaned = dv\_police\_df.dropna()

fb\_world\_cup\_df\_cleaned = fb\_world\_cup\_df.dropna()

fb\_premier\_df\_cleaned = fb\_premier\_df.dropna()

fb\_championship\_df\_cleaned = fb\_championship\_df.dropna()

fb\_fa\_cup\_df\_cleaned = fb\_fa\_cup\_df.dropna()

fb\_euro\_df\_cleaned = fb\_euro\_df.dropna()

*#grouping police data by year and location, dropping month columns*

dv\_police\_df\_cleaned\_aggregated = dv\_police\_df\_cleaned.groupby(['Year', 'Location']).sum().reset\_index().drop(columns=['Month'])

*#aggregated police data, separated by location*

df\_police\_newcastle = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "Newcastle"]

df\_police\_gateshead = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "Gateshead"]

df\_police\_n\_tyneside = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "North Tyneside"]

df\_police\_s\_tyneside = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "South Tyneside"]

df\_police\_sunderland = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "Sunderland"]

df\_police\_northumberland = dv\_police\_df\_cleaned\_aggregated[dv\_police\_df\_cleaned\_aggregated["Location"] == "Northumberland"]

*#plotting domestic violence instances per year per location into a scatter graph*

plt.figure(figsize=(8, 5))

plt.scatter(df\_police\_newcastle['Year'], df\_police\_newcastle['ReportedInstances'], marker='o', color='black', s=100, label='Newcastle')

plt.scatter(df\_police\_gateshead['Year'], df\_police\_gateshead['ReportedInstances'], marker='o', color='blue', s=100, label='Gateshead')

plt.scatter(df\_police\_n\_tyneside['Year'], df\_police\_n\_tyneside['ReportedInstances'], marker='o', color='green', s=100, label='North Tyneside')

plt.scatter(df\_police\_s\_tyneside['Year'], df\_police\_s\_tyneside['ReportedInstances'], marker='o', color='orange', s=100, label='South Tyneside')

plt.scatter(df\_police\_sunderland['Year'], df\_police\_sunderland['ReportedInstances'], marker='o', color='red', s=100, label='Sunderland')

plt.scatter(df\_police\_northumberland['Year'], df\_police\_northumberland['ReportedInstances'], marker='o', color='yellow', s=100, label='Northumberland')

*#adding titles and labels*

plt.title('Domestic Violence Instances By Location')

plt.xlabel('Year')

plt.ylabel('Reported Instances')

plt.grid(True, linestyle='--', alpha=0.6)

plt.legend()

*#displaying the plot*

plt.show()

*#grouping football dataframe by season, winner and then sorting by the date of match for all football dataframes*

*#World Cup*

fb\_world\_cup\_group\_df = fb\_world\_cup\_df\_cleaned.groupby(["Season", "Winner"]).count().reset\_index()

fb\_world\_cup\_sorted\_df = fb\_world\_cup\_group\_df.sort\_values(by="DateofMatch", ascending=False)

fb\_world\_cup\_sorted\_df = fb\_world\_cup\_sorted\_df[fb\_world\_cup\_sorted\_df['Winner'] != "Draw"]

*#Premier Cup*

fb\_premier\_group\_df = fb\_premier\_df\_cleaned.groupby(["Season", "Winner"]).count().reset\_index()

fb\_premier\_sorted\_df = fb\_premier\_group\_df.sort\_values(by="DateofMatch", ascending=False)

fb\_premier\_sorted\_df = fb\_premier\_sorted\_df[fb\_premier\_sorted\_df['Winner'] != "Draw"]

*#Championship*

fb\_championship\_group\_df = fb\_championship\_df\_cleaned.groupby(["Season", "Winner"]).count().reset\_index()

fb\_championship\_sorted\_df = fb\_championship\_group\_df.sort\_values(by="DateofMatch", ascending=False)

fb\_championship\_sorted\_df = fb\_championship\_sorted\_df[fb\_championship\_sorted\_df['Winner'] != "Draw"]

*#FA Cup*

fb\_fa\_cup\_df\_group\_df = fb\_fa\_cup\_df\_cleaned.groupby(["Season", "Winner"]).count().reset\_index()

fb\_fa\_cup\_df\_sorted\_df = fb\_fa\_cup\_df\_group\_df.sort\_values(by="DateofMatch", ascending=False)

fb\_fa\_cup\_df\_top30\_df = fb\_fa\_cup\_df\_sorted\_df.nlargest(30, "DateofMatch")

fb\_fa\_cup\_df\_top30\_df = fb\_fa\_cup\_df\_top30\_df[fb\_fa\_cup\_df\_top30\_df['Winner'] != "Draw"]

*#Euros*

fb\_euro\_group\_df = fb\_euro\_df\_cleaned.groupby(["Season", "Winner"]).count().reset\_index()

fb\_euro\_sorted\_df = fb\_euro\_group\_df.sort\_values(by="DateofMatch", ascending=False)

fb\_euro\_sorted\_df = fb\_euro\_sorted\_df[fb\_euro\_sorted\_df['Winner'] != "Draw"]

*#bar chart showing Wins per team during the World Cup in the years 2018 and 2022*

plt.figure(figsize=(8, 5))

plt.bar(fb\_world\_cup\_sorted\_df['Winner'], fb\_world\_cup\_sorted\_df['DateofMatch'], color=['blue'])

*#adding titles and labels*

plt.title('Football Matches - World Cup')

plt.xlabel('Winner')

plt.ylabel('Amount of Wins')

plt.grid(True, linestyle='--', alpha=0.6)

plt.xticks(rotation=90)

*#displaying the plot*

plt.show()

*#bar chart showing Wins per team in the Premier League between 2018-2024*

plt.figure(figsize=(8, 5))

plt.bar(fb\_premier\_sorted\_df['Winner'], fb\_premier\_sorted\_df['DateofMatch'], color=['green'])

*#adding titles and labels*

plt.title('Football Matches - Premier League')

plt.xlabel('Winner')

plt.ylabel('Amount of Wins')

plt.grid(True, linestyle='--', alpha=0.6)

plt.xticks(rotation=90)

*#displaying the plot*

plt.show()

*#bar chart showing Wins per team in the Championship between 2018-2024*

plt.figure(figsize=(8, 5))

plt.bar(fb\_championship\_sorted\_df['Winner'], fb\_championship\_sorted\_df['DateofMatch'], color=['orange'])

*#adding titles and labels*

plt.title('Football Matches - Championship')

plt.xlabel('Winner')

plt.ylabel('Amount of Wins')

plt.grid(True, linestyle='--', alpha=0.6)

plt.xticks(rotation=90)

*#displaying the plot*

plt.show()

*#bar chart showing Wins per team in the FA Cup between 2018-2024*

plt.figure(figsize=(8, 5))

plt.bar(fb\_fa\_cup\_df\_top30\_df['Winner'], fb\_fa\_cup\_df\_top30\_df['DateofMatch'], color=['red'])

*#adding titles and labels*

plt.title('Football Matches - FA Cup')

plt.xlabel('Winner')

plt.ylabel('Amount of Wins')

plt.grid(True, linestyle='--', alpha=0.6)

plt.xticks(rotation=90)

*#displaying the plot*

plt.show()

*#bar chart showing Wins per team in the Euros between 2018-2024*

plt.figure(figsize=(8, 5))

plt.bar(fb\_euro\_sorted\_df['Winner'], fb\_euro\_sorted\_df['DateofMatch'], color=['purple'])

*#adding titles and labels*

plt.title('Football Matches - Euros')

plt.xlabel('Winner')

plt.ylabel('Amount of Wins')

plt.grid(True, linestyle='--', alpha=0.6)

plt.xticks(rotation=90)

*#displaying the plot*

plt.show()

*#Filtering premier league dataframe to Newcastle United games*

fb\_premier\_df\_filtered = fb\_premier\_df\_cleaned[(fb\_premier\_df\_cleaned['HomeTeam'] == "Newcastle United") | (fb\_premier\_df\_cleaned['AwayTeam'] == "Newcastle United")].reset\_index()

fb\_premier\_df\_filtered.head()

*#transforming Date of Match to correct format and grouping by Year, Month*

fb\_premier\_df\_filtered['DateofMatch'].astype(str)

fb\_premier\_df\_filtered['DateofMatch'] = pd.to\_datetime(fb\_premier\_df\_filtered['DateofMatch'], format='%d/%m/%Y')

fb\_premier\_df\_filtered['MonthNumber'] = fb\_premier\_df\_filtered['DateofMatch'].dt.month

fb\_premier\_df\_filtered['Year'] = fb\_premier\_df\_filtered['DateofMatch'].dt.year

fb\_premier\_df\_filtered['Month'] = fb\_premier\_df\_filtered['DateofMatch'].dt.month\_name()

fb\_premier\_df\_filtered = fb\_premier\_df\_filtered.groupby(['Year','Month']).size().reset\_index(name='NoOfMatches')

*#filling empty datapoints with 0s*

zero\_matches = [

{"Year": 2019, "Month": "May", "NoOfMatches": 0},

{"Year": 2019, "Month": "July", "NoOfMatches": 0},

{"Year": 2020, "Month": "April", "NoOfMatches": 0},

{"Year": 2020, "Month": "May", "NoOfMatches": 0},

{"Year": 2020, "Month": "August", "NoOfMatches": 0},

{"Year": 2021, "Month": "June", "NoOfMatches": 0},

{"Year": 2021, "Month": "July", "NoOfMatches": 0},

{"Year": 2021, "Month": "September", "NoOfMatches": 0},

{"Year": 2022, "Month": "June", "NoOfMatches": 0},

{"Year": 2022, "Month": "July", "NoOfMatches": 0},

{"Year": 2023, "Month": "June", "NoOfMatches": 0},

{"Year": 2023, "Month": "July", "NoOfMatches": 0},

{"Year": 2024, "Month": "June", "NoOfMatches": 0},

{"Year": 2024, "Month": "July", "NoOfMatches": 0},

]

zero\_matches\_df = pd.DataFrame(zero\_matches)

*#joining the two dataframes together*

fb\_premier\_df\_filtered = pd.concat([fb\_premier\_df\_filtered, zero\_matches\_df], ignore\_index=True)

fb\_premier\_df\_filtered.head(65)

*#filtering police dataframe to Newcastle*

dv\_police\_df\_filtered\_newcastle = dv\_police\_df\_cleaned[(dv\_police\_df\_cleaned["Location"] == "Newcastle")].reset\_index()

dv\_police\_df\_newcastle\_group\_by\_month = dv\_police\_df\_filtered\_newcastle.groupby(['Year', 'Month'])['ReportedInstances'].sum().reset\_index()

dv\_police\_df\_newcastle\_group\_by\_month.head(24)

*#combing dataset and then setting the month order, before sorting by year and month*

dv\_premier\_combined\_df = pd.merge(fb\_premier\_df\_filtered, dv\_police\_df\_newcastle\_group\_by\_month, on=['Year', 'Month'], how="inner")

*#setting month order*

month\_order = [

'January', 'February', 'March', 'April', 'May', 'June',

'July', 'August', 'September', 'October', 'November', 'December'

]

*#setting month column as categorical with ordered categories*

dv\_premier\_combined\_df['Month'] = pd.Categorical(dv\_premier\_combined\_df['Month'], categories=month\_order, ordered=True)

*#sorting dataframe by year and month*

dv\_premier\_combined\_sorted\_df = dv\_premier\_combined\_df.sort\_values(by=['Year','Month'])

dv\_premier\_combined\_sorted\_df.head(80)

*#splitting combined dataframe into years between 2019 to 2024*

dv\_premier\_combined\_df\_2019 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2019]

dv\_premier\_combined\_df\_2020 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2020]

dv\_premier\_combined\_df\_2021 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2021]

dv\_premier\_combined\_df\_2022 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2022]

dv\_premier\_combined\_df\_2023 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2023]

dv\_premier\_combined\_df\_2024 = dv\_premier\_combined\_sorted\_df[dv\_premier\_combined\_sorted\_df["Year"] == 2024]

*#plotting domestic violence in Newcastle per month by year on graph*

plt.figure(figsize=(8, 5))

plt.plot(dv\_premier\_combined\_df\_2019['Month'], dv\_premier\_combined\_df\_2019['ReportedInstances'], marker='o', color='pink', label='2019')

plt.plot(dv\_premier\_combined\_df\_2020['Month'], dv\_premier\_combined\_df\_2020['ReportedInstances'], marker='o', color='blue', label='2020')

plt.plot(dv\_premier\_combined\_df\_2021['Month'], dv\_premier\_combined\_df\_2021['ReportedInstances'], marker='o', color='green', label='2021')

plt.plot(dv\_premier\_combined\_df\_2022['Month'], dv\_premier\_combined\_df\_2022['ReportedInstances'], marker='o', color='orange',label='2022')

plt.plot(dv\_premier\_combined\_df\_2023['Month'], dv\_premier\_combined\_df\_2023['ReportedInstances'], marker='o', color='red', label='2023')

plt.plot(dv\_premier\_combined\_df\_2024['Month'], dv\_premier\_combined\_df\_2024['ReportedInstances'], marker='o', color='purple', label='2024')

*#adding titles and labels*

plt.title('Domestic Violence in Newcastle Per Month')

plt.xlabel('Month')

plt.ylabel('Reported Instances')

plt.xticks(rotation=90)

plt.legend()

*#displaying the plot*

plt.show()

*#plotting number of domestic violence instances in Newcastle by amounts of matches*

plt.figure(figsize=(8, 5))

plt.scatter(dv\_premier\_combined\_df\_2019['ReportedInstances'], dv\_premier\_combined\_df\_2019['NoOfMatches'], marker='o', color='pink', s=100, label='2019')

plt.scatter(dv\_premier\_combined\_df\_2020['ReportedInstances'], dv\_premier\_combined\_df\_2020['NoOfMatches'], marker='o', color='blue', s=100, label='2020')

plt.scatter(dv\_premier\_combined\_df\_2021['ReportedInstances'], dv\_premier\_combined\_df\_2021['NoOfMatches'], marker='o', color='green', s=100, label='2021')

plt.scatter(dv\_premier\_combined\_df\_2022['ReportedInstances'], dv\_premier\_combined\_df\_2022['NoOfMatches'], marker='o', color='orange', s=100, label='2022')

plt.scatter(dv\_premier\_combined\_df\_2023['ReportedInstances'], dv\_premier\_combined\_df\_2023['NoOfMatches'], marker='o', color='red', s=100, label='2023')

plt.scatter(dv\_premier\_combined\_df\_2024['ReportedInstances'], dv\_premier\_combined\_df\_2024['NoOfMatches'], marker='o', color='purple', s=100, label='2024')

*#adding title and labels*

plt.title('Domestic Violence in Newcastle By Match No')

plt.xlabel('Reported Instances')

plt.ylabel('Matches')

plt.legend()

*#displaying the plot*

plt.show()

*#checking dataframe types*

print(dv\_premier\_combined\_sorted\_df.dtypes)

print(dv\_premier\_combined\_sorted\_df.head())

*#dropping NA from premier league combined dataframe*

dv\_premier\_combined\_sorted\_df.replace(0,np.nan,inplace=True)

dv\_premier\_combined\_sorted\_df.dropna()

*#sorting months and then sorting dataframe into months*

*#mapping the months*

month\_map = {

'January': 1, 'February': 2, 'March': 3, 'April': 4, 'May': 5, 'June': 6,

'July': 7, 'August': 8, 'September': 9, 'October': 10, 'November': 11, 'December': 12

}

dv\_premier\_combined\_sorted\_df['Month\_Num'] = dv\_premier\_combined\_sorted\_df['Month'].map(month\_map)

*#converting month\_num to Numeric data type and then dropping original month column*

dv\_premier\_combined\_sorted\_df['Month\_Num'] = pd.to\_numeric(dv\_premier\_combined\_sorted\_df['Month\_Num'])

dv\_premier\_combined\_sorted\_df = dv\_premier\_combined\_sorted\_df.drop(columns=['Month'])

*#finding Pearson correlation coefficient*

corr\_coeff = dv\_premier\_combined\_sorted\_df.corr().loc['ReportedInstances', 'NoOfMatches']

print(f'Pearson correlation coefficient: {corr\_coeff:.2f}')

*#plotting Pearson correlation coefficient in a scatter plot using Seaborn library*

plt.figure(figsize=(8,6))

sns.regplot(data=dv\_premier\_combined\_sorted\_df, x='ReportedInstances', y='NoOfMatches', scatter\_kws={'color': 'blue'}, line\_kws={'color': 'red'})

*#annotating the Pearson correlation coefficient*

plt.text(

x=dv\_premier\_combined\_sorted\_df['ReportedInstances'].max() \* 0.7,

y=dv\_premier\_combined\_sorted\_df['NoOfMatches'].max() \* 0.9,

s=f'Pearson Correlation: {corr\_coeff:.2f}',

fontsize=12,

bbox=dict(facecolor='white', alpha=0.5)

)

*#adding labels and titles*

plt.xlabel("Reported Instances")

plt.ylabel("Number of Matches")

plt.title("Scatter Plot with Pearson Correlation Coefficient")

*#displaying the plot*

plt.show()

*#filtering police dataframe to Sunderland*

fb\_championship\_df\_filtered = fb\_championship\_df\_cleaned[(fb\_championship\_df\_cleaned['HomeTeam'] == "Sunderland AFC") | (fb\_championship\_df\_cleaned['AwayTeam'] == "Sunderland AFC")].reset\_index()

fb\_championship\_df\_filtered.head()

*#transforming Date of Match to correct format and grouping by Year, Month*

fb\_championship\_df\_filtered['DateofMatch'].astype(str)

fb\_championship\_df\_filtered['DateofMatch'] = pd.to\_datetime(fb\_championship\_df\_filtered['DateofMatch'], format='%d/%m/%Y')

fb\_championship\_df\_filtered['MonthNumber'] = fb\_championship\_df\_filtered['DateofMatch'].dt.month

fb\_championship\_df\_filtered['Year'] = fb\_championship\_df\_filtered['DateofMatch'].dt.year

fb\_championship\_df\_filtered['Month'] = fb\_championship\_df\_filtered['DateofMatch'].dt.month\_name()

fb\_championship\_df\_filtered = fb\_championship\_df\_filtered.groupby(['Year','Month']).size().reset\_index(name='NoOfMatches')

*#filling empty datapoints with 0s*

zero\_matches = [

{"Year": 2022, "Month": "January", "NoOfMatches": 0},

{"Year": 2022, "Month": "February", "NoOfMatches": 0},

{"Year": 2022, "Month": "March", "NoOfMatches": 0},

{"Year": 2022, "Month": "April", "NoOfMatches": 0},

{"Year": 2022, "Month": "May", "NoOfMatches": 0},

{"Year": 2022, "Month": "June", "NoOfMatches": 0},

{"Year": 2022, "Month": "July", "NoOfMatches": 0},

{"Year": 2023, "Month": "June", "NoOfMatches": 0},

{"Year": 2023, "Month": "July", "NoOfMatches": 0},

{"Year": 2024, "Month": "June", "NoOfMatches": 0},

{"Year": 2024, "Month": "July", "NoOfMatches": 0},

]

*#concatenating zero data points into main dataframe*

zero\_matches\_df\_champ = pd.DataFrame(zero\_matches)

fb\_championship\_df\_filtered = pd.concat([fb\_championship\_df\_filtered, zero\_matches\_df\_champ], ignore\_index=True)

*#filtering by location*

dv\_police\_df\_filtered\_sunderland = dv\_police\_df\_cleaned[(dv\_police\_df\_cleaned["Location"] == "Sunderland")].reset\_index()

dv\_police\_df\_filtered\_sunderland\_group\_by\_month = dv\_police\_df\_filtered\_sunderland.groupby(['Year', 'Month'])['ReportedInstances'].sum().reset\_index()

*#merging dataframes - sunderland domestic violence by month and championship*

dv\_championship\_combined\_df = pd.merge(fb\_championship\_df\_filtered, dv\_police\_df\_filtered\_sunderland\_group\_by\_month, on=['Year', 'Month'], how="inner")

*#ordering by month column*

month\_order = [

'January', 'February', 'March', 'April', 'May', 'June',

'July', 'August', 'September', 'October', 'November', 'December'

]

*#sorting dataframe by year and month*

dv\_championship\_combined\_df['Month'] = pd.Categorical(dv\_championship\_combined\_df['Month'], categories=month\_order, ordered=True)

dv\_championship\_combined\_sorted\_df = dv\_championship\_combined\_df.sort\_values(by=['Year','Month'])

dv\_championship\_combined\_sorted\_df.head(80)

*#filtering data by year 2022 to 2024 (years Sunderland were in Championship)*

dv\_championship\_combined\_df\_2022 = dv\_championship\_combined\_sorted\_df[dv\_championship\_combined\_sorted\_df['Year'] == 2022]

dv\_championship\_combined\_df\_2023 = dv\_championship\_combined\_sorted\_df[dv\_championship\_combined\_sorted\_df['Year'] == 2023]

dv\_championship\_combined\_df\_2024 = dv\_championship\_combined\_sorted\_df[dv\_championship\_combined\_sorted\_df['Year'] == 2024]

*#plotting data in graph for dataframe*

plt.figure(figsize=(8, 5))

plt.plot(dv\_championship\_combined\_df\_2022['Month'], dv\_championship\_combined\_df\_2022['ReportedInstances'], marker='o', color='orange',label='2022')

plt.plot(dv\_championship\_combined\_df\_2023['Month'], dv\_championship\_combined\_df\_2023['ReportedInstances'], marker='o', color='red', label='2023')

plt.plot(dv\_championship\_combined\_df\_2024['Month'], dv\_championship\_combined\_df\_2024['ReportedInstances'], marker='o', color='purple', label='2024')

*#adding titles and labels*

plt.title('Domestic Violence in Sunderland Per Month')

plt.xlabel('Month')

plt.ylabel('Reported Instances')

plt.xticks(rotation=90)

plt.legend()

*#displaying the plot*

plt.show()

*#plotting data in a scatter graph for dataframe*

plt.figure(figsize=(8, 5))

plt.scatter(dv\_championship\_combined\_df\_2022['ReportedInstances'], dv\_championship\_combined\_df\_2022['NoOfMatches'], marker='o', color='orange', s=100, label='2022')

plt.scatter(dv\_championship\_combined\_df\_2023['ReportedInstances'], dv\_championship\_combined\_df\_2023['NoOfMatches'], marker='o', color='red', s=100, label='2023')

plt.scatter(dv\_championship\_combined\_df\_2024['ReportedInstances'], dv\_championship\_combined\_df\_2024['NoOfMatches'], marker='o', color='purple', s=100, label='2024')

*#adding titles and labels*

plt.title('Domestic Violence in Sunderland By Match No')

plt.xlabel('Reported Instances')

plt.ylabel('Matches')

plt.legend()

*#displaying the plot*

plt.show()

*#checking datatype columns*

print(dv\_championship\_combined\_sorted\_df.dtypes)

print(dv\_championship\_combined\_sorted\_df.head())

*#replacing 0s with NaN*

dv\_championship\_combined\_sorted\_df.replace(0,np.nan,inplace=True)

dv\_championship\_combined\_sorted\_df.dropna()

*#mapping the months*

month\_map = {

'January': 1, 'February': 2, 'March': 3, 'April': 4, 'May': 5, 'June': 6,

'July': 7, 'August': 8, 'September': 9, 'October': 10, 'November': 11, 'December': 12

}

*#mapping month names to number and converting to numeric*

dv\_championship\_combined\_sorted\_df['Month\_Num'] = dv\_championship\_combined\_sorted\_df['Month'].map(month\_map)

dv\_championship\_combined\_sorted\_df['Month\_Num'] = pd.to\_numeric(dv\_championship\_combined\_sorted\_df['Month\_Num'])

*#dropping month column*

dv\_championship\_combined\_sorted\_df = dv\_championship\_combined\_sorted\_df.drop(columns=['Month'])

*#calculating pearson correlation coefficent*

corr\_coeff\_sun = dv\_championship\_combined\_sorted\_df.corr().loc['ReportedInstances', 'NoOfMatches']

print(f'Pearson correlation coefficient: {corr\_coeff\_sun:.2f}')

*#creating scatter plot*

plt.figure(figsize=(8,6))

sns.regplot(data=dv\_championship\_combined\_sorted\_df, x='ReportedInstances', y='NoOfMatches', scatter\_kws={'color': 'blue'}, line\_kws={'color': 'red'})

*#annotate the Pearson correlation coefficient*

plt.text(

x=dv\_championship\_combined\_sorted\_df['ReportedInstances'].max() \* 0.7,

y=dv\_championship\_combined\_sorted\_df['NoOfMatches'].max() \* 0.9,

s=f'Pearson Correlation: {corr\_coeff\_sun:.2f}',

fontsize=12,

bbox=dict(facecolor='white', alpha=0.5)

)

*#adding labels and title*

plt.xlabel("Reported Instances")

plt.ylabel("Number of Matches")

plt.title("Scatter Plot with Pearson Correlation Coefficient")

*#displaying the plot*

plt.show()

*#add a new column 'Location' to the dv\_championship\_combined\_sorted\_df DataFrame and set the value to 'Sunderland'*

dv\_championship\_combined\_sorted\_df["Location"] = "Sunderland"

*#add a new column 'Location' to the dv\_premier\_combined\_sorted\_df DataFrame and set the value to 'Newcastle'*

dv\_premier\_combined\_sorted\_df["Location"] = "Newcastle"

*#combine the dv\_championship\_combined\_sorted\_df (Sunderland) and dv\_premier\_combined\_sorted\_df (Newcastle) DataFrames vertically.*

*#the 'ignore\_index=True' ensures that the index is reset after concatenation, so it's continuous across the new combined DataFrame.*

newcastle\_sunderland\_combined\_df = pd.concat([dv\_championship\_combined\_sorted\_df, dv\_premier\_combined\_sorted\_df], ignore\_index=True)

*#display the first 50 rows of the newly combined DataFrame to inspect the result*

newcastle\_sunderland\_combined\_df.head(50)

*#select all columns in the newcastle\_sunderland\_combined\_df DataFrame that have the data type 'object' (categorical columns)*

categorical\_cols = newcastle\_sunderland\_combined\_df.select\_dtypes(include=['object']).columns

print("Categorical Columns:", categorical\_cols)

*#converting categorical columns to dummy/indicator variables using pd.get\_dummies.*

newcastle\_sunderland\_combined\_df\_encoded = pd.get\_dummies(newcastle\_sunderland\_combined\_df, columns=categorical\_cols)

*#dropping any rows with missing values (NaN) in the encoded DataFrame*

newcastle\_sunderland\_combined\_df\_encoded.dropna(inplace=True)

newcastle\_sunderland\_combined\_df\_encoded.head()

*#selecting the feature columns for the model (X)*

X = newcastle\_sunderland\_combined\_df\_encoded[["Year", "Month\_Num", "NoOfMatches", "Location\_Sunderland", "Location\_Newcastle"]].values

*#selecting the target variable (y).*

y = newcastle\_sunderland\_combined\_df\_encoded['ReportedInstances'].values

*#splitting the data into training and testing sets.*

X\_MLP\_train, X\_MLP\_test, y\_MLP\_train, y\_MLP\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

*#initialising a StandardScaler object to scale (normalize) the feature data (X)and target variable (y).*

scaler\_X\_MLP = StandardScaler()

scaler\_y\_MLP = StandardScaler()

*#fitting the scaler onto each bit of data*

X\_train\_MLP\_scaled = scaler\_X\_MLP.fit\_transform(X\_MLP\_train)

X\_test\_MLP\_scaled = scaler\_X\_MLP.transform(X\_MLP\_test)

*#reshaping the data for training*

y\_train\_MLP\_scaled = scaler\_y\_MLP.fit\_transform(y\_MLP\_train.reshape(-1, 1))

y\_test\_MLP\_scaled = scaler\_y\_MLP.transform(y\_MLP\_test.reshape(-1, 1))

*#converting the test data to a 32-bit floating point format to ensure data type matches expected input format*

X\_MLP\_test = X\_MLP\_test.astype(np.float32)

*#defining MLP model*

model = keras.Sequential([

keras.layers.Input(shape=(X\_train\_MLP\_scaled.shape[1],)),

keras.layers.Dense(256, activation='relu', input\_shape=(X\_train\_MLP\_scaled.shape[1],)), # More neurons

keras.layers.Dense(256, activation='relu', kernel\_regularizer=tf.keras.regularizers.l2(0.01)),

keras.layers.BatchNormalization(),

keras.layers.Dense(128, activation='relu'),

keras.layers.BatchNormalization(),

keras.layers.Dense(64, activation='relu'),

keras.layers.Dropout(0.3), # Prevent overfitting

keras.layers.Dense(32, activation='relu'),

keras.layers.Dense(1, activation='linear') # Output single value

])

early\_stopping = tf.keras.callbacks.EarlyStopping(monitor='val\_mae', patience=15, restore\_best\_weights=True)

*#compiling model*

model.compile(optimizer=tf.keras.optimizers.Adam(learning\_rate=0.00001), loss='huber', metrics=['mae'])

*#training model*

history\_MLP = model.fit(X\_train\_MLP\_scaled, y\_train\_MLP\_scaled, epochs=500, batch\_size=16, validation\_data=(X\_test\_MLP\_scaled, y\_test\_MLP\_scaled), callbacks=[early\_stopping])

y\_pred = model.predict(X\_test\_MLP\_scaled)

y\_pred\_original = scaler\_y\_MLP.inverse\_transform(y\_pred.reshape(-1, 1))

*#inverse transforming the scaled test data back to original values*

y\_test\_original = scaler\_y\_MLP.inverse\_transform(y\_test\_MLP\_scaled)

*#inverse transform the scaled predictions back to original values*

y\_pred\_original = scaler\_y\_MLP.inverse\_transform(y\_pred.reshape(-1, 1))

*#checking y\_pred\_original is correct*

print("Predicted Reported Instances:", y\_pred\_original[:5].flatten())

*#calculating the Mean Absolute Error (MAE) between the original test values (y\_test\_original) and printing*

mae\_original = mean\_absolute\_error(y\_test\_original, y\_pred\_original)

print(f"Mean Absolute Error: {mae\_original:.2f}")

*#printing the minimum and maximum values of the actual target values (y\_CNN\_test) in the test set*

print("Actual y range:", y\_test\_original.min(), "to", y\_test\_original.max())

*#calculating and printing the Mean Absolute Error (MAE) as a percentage of the range of actual target values*

print("MAE as % of y range:", (51.68 / (y\_test\_original.max() - y\_test\_original.min())) \* 100, "%")

*#plotting loss & MAE curves*

plt.figure(figsize=(12, 4))

plt.subplot(1, 2, 1)

plt.plot(history\_MLP.history['loss'], label='Train Loss')

plt.plot(history\_MLP.history['val\_loss'], label='Validation Loss')

plt.legend()

plt.title("Loss Over Epochs")

plt.subplot(1, 2, 2)

plt.plot(history\_MLP.history['mae'], label='Train MAE')

plt.plot(history\_MLP.history['val\_mae'], label='Validation MAE')

plt.legend()

plt.title("MAE Over Epochs")

plt.show()

*#creating an index for each data point (to keep x-axis simple)*

indices = np.arange(len(y\_test\_original))

*#scatter plot for actual values*

plt.figure(figsize=(10, 6))

plt.scatter(indices, y\_test\_original, color='blue', alpha=0.6, label='Actual Values')

*#scatter plot for predicted values*

plt.scatter(indices, y\_pred\_original, color='red', alpha=0.6, label='Predicted Values')

*#labels and title*

plt.xlabel('Data Point Index')

plt.ylabel('Values')

plt.title('Actual vs Predicted Values (MLP Model)')

plt.legend()

plt.grid(True)

*#showing plot*

plt.show()

*#initialising the StandardScaler to scale the input features for the CNN model*

scaler\_x\_CNN = StandardScaler()

*#reshaping the input data (X) to ensure it's 2D for scaling and applying scaling to the input features (X) and store the scaled data*

X = X.reshape(X.shape[0], -1)

X\_scaled = scaler\_x\_CNN.fit\_transform(X)

*#reshaping the scaled input data into a 3D format (samples, features, 1) for CNN input*

X\_reshaped\_CNN = X\_scaled.reshape((X\_scaled.shape[0], X\_scaled.shape[1], 1))

*#initialising the StandardScaler to scale the target values for the CNN model*

scaler\_y\_CNN = StandardScaler()

*#applying scaling to the target values (y) and flatten the result to make it 1D*

y\_scaled\_CNN = scaler\_y\_CNN.fit\_transform(y.reshape(-1, 1)).flatten()

*#splitting the data into training and testing sets (80% train, 20% test) for both features and target values*

x\_CNN\_train, x\_CNN\_test, y\_CNN\_train, y\_CNN\_test = train\_test\_split(X\_reshaped\_CNN, y\_scaled\_CNN, test\_size=0.2, random\_state=42)

*#defining the 1D CNN model*

model = keras.Sequential([

layers.Conv1D(32, kernel\_size=1, activation='relu', input\_shape=(X\_reshaped\_CNN.shape[1], 1)),

layers.MaxPooling1D(pool\_size=1), *#reduce pooling*

layers.Conv1D(64, kernel\_size=1, activation='relu'),

layers.Flatten(),

layers.Dense(64, activation='relu'),

layers.Dropout(0.5),

layers.Dense(1, activation='linear')

])

*#compiling the model*

model.compile(optimizer='adam', loss='mse', metrics=['mae'])

*#training the model*

history\_CNN = model.fit(x\_CNN\_train, y\_CNN\_train, epochs=200, validation\_data=(x\_CNN\_test, y\_CNN\_test))

*#using the trained CNN model to make predictions on the test set (x\_CNN\_test) and then inversing/transforming the scaled predictions*

y\_pred\_CNN = model.predict(x\_CNN\_test)

y\_pred\_CNN = scaler\_y\_CNN.inverse\_transform(y\_pred\_CNN.reshape(-1, 1))

*#evaluating the model's performance on the test set (x\_CNN\_test, y\_CNN\_test) and finding the MAE*

test\_loss, test\_mae = model.evaluate(x\_CNN\_test, y\_CNN\_test, verbose=2)

print(f'\n Mean Absolute Error: {test\_mae:.4f}')

*#printing the minimum and maximum values of the actual target values (y\_CNN\_test) in the test set*

print("Actual y range:", y\_CNN\_test.min(), "to", y\_CNN\_test.max())

*#calculating and printing the Mean Absolute Error (MAE) as a percentage of the range of actual target values*

print("MAE as % of y range:", (0.2929 / (y\_CNN\_test.max() - y\_CNN\_test.min())) \* 100, "%")

*#plotting loss & MAE curves*

plt.figure(figsize=(12, 4))

plt.subplot(1, 2, 1)

plt.plot(history\_CNN.history['loss'], label='Train Loss')

plt.plot(history\_CNN.history['val\_loss'], label='Validation Loss')

plt.legend()

plt.title("Loss Over Epochs")

plt.subplot(1, 2, 2)

plt.plot(history\_CNN.history['mae'], label='Train MAE')

plt.plot(history\_CNN.history['val\_mae'], label='Validation MAE')

plt.legend()

plt.title("MAE Over Epochs")

plt.show()

*#reverting y values back to actuals*

y\_CNN\_test\_original = scaler\_y\_CNN.inverse\_transform(y\_CNN\_test.reshape(-1, 1))

*#creating an index for each data point (to keep x-axis simple)*

indices = np.arange(len(y\_CNN\_test))

*#scatter plot for actual values*

plt.figure(figsize=(10, 6))

plt.scatter(range(len(y\_CNN\_test\_original)), y\_CNN\_test\_original, color='blue', label='Original y\_CNN\_test')

*#scatter plot for predicted values*

plt.scatter(indices, y\_pred\_CNN, color='red', alpha=0.6, label='Predicted Values')

*#labels and title*

plt.xlabel('Data Point Index')

plt.ylabel('Values')

plt.title('Actual vs Predicted Values (CNN Model)')

plt.legend()

plt.grid(True)

*#showing plot*

plt.show()

*#initializing MinMaxScaler to scale the input features and target values to a range of [0, 1] for the RNN model*

scaler\_X\_RNN = MinMaxScaler(feature\_range=(0, 1))

scaler\_y\_RNN = MinMaxScaler(feature\_range=(0, 1))

*#applying scaling to the input data X*

X\_scaled\_RNN = scaler\_X\_RNN.fit\_transform(X)

*#getting the number of samples and the number of features from the scaled input data*

num\_samples = X.shape[0]

num\_features = X\_scaled\_RNN.shape[1]

seq\_length = 1

*#checking if the number of samples can be evenly divided by the sequence length*

if num\_samples % seq\_length == 0:

*#reshaping the input data (X\_scaled\_RNN) into 3D shape for the RNN (samples, time steps, features)*

num\_samples = num\_samples // seq\_length

X\_reshaped\_RNN = X\_scaled\_RNN.reshape((num\_samples, seq\_length, num\_features))

else:

*#raising an error if the data cannot be reshaped evenly according to the sequence length*

raise ValueError(f"Cannot reshape array of size {X\_scaled\_RNN.size} into ({num\_samples}, {seq\_length}, {num\_features}). Adjust seq\_length.")

*#scaling the target values (y) and flatten the result to make it 1D*

y\_scaled\_RNN = scaler\_y\_RNN.fit\_transform(y.reshape(-1, 1)).flatten()

*#splitting the dataset into training and test sets, using 80% of data for training*

train\_size = int(0.8 \* num\_samples)

X\_RNN\_train, X\_RNN\_test, y\_RNN\_train, y\_RNN\_test = train\_test\_split(

X\_reshaped\_RNN, y\_scaled\_RNN,

test\_size=0.2,

random\_state=42,

)

*#defining the RNN model*

model = keras.Sequential([

layers.SimpleRNN(100, activation='relu', return\_sequences=True, input\_shape=(seq\_length, num\_features)),

layers.Dropout(0.2),

layers.SimpleRNN(100, activation='relu'),

layers.Dropout(0.2),

layers.Dense(1)

])

*#compiling the model*

model.compile(optimizer=keras.optimizers.Adam(learning\_rate=0.0005), loss='mse', metrics=['mae'])

*#training the model*

history\_RNN = model.fit(X\_RNN\_train, y\_RNN\_train, epochs=50, batch\_size=8, validation\_data=(X\_RNN\_test, y\_RNN\_test))

*#evaluating the model's performance on the test set (x\_CNN\_test, y\_CNN\_test) and finding the MAE*

test\_loss, test\_mae = model.evaluate(X\_RNN\_test, y\_RNN\_test, verbose=2)

print(f"\nMean Absolute Error: {test\_mae:.4f}")

*#inverse transforming the scaled target values (y\_RNN\_test) and predicted values (y\_pred\_scaled\_RNN) back to the original scale*

y\_RNN\_test\_original = scaler\_y\_RNN.inverse\_transform(y\_RNN\_test.reshape(-1, 1))

y\_pred\_RNN\_original = scaler\_y\_RNN.inverse\_transform(y\_pred\_scaled\_RNN.reshape(-1, 1))

*#printing the minimum and maximum values of the actual target values (y\_CNN\_test) in the test set*

print("Actual y range:", y\_RNN\_test\_original.min(), "to", y\_RNN\_test\_original.max())

*#calculating and printing the Mean Absolute Error (MAE) as a percentage of the range of actual target values*

print("MAE as % of y range:", (0.2235 / (y\_RNN\_test\_original.max() - y\_RNN\_test\_original.min())) \* 100, "%")

*#using the trained RNN model to make predictions on the test set (X\_RNN\_test)*

y\_pred\_scaled\_RNN = model.predict(X\_RNN\_test)

*#inverse transform the scaled predicted values (y\_pred\_scaled\_RNN) back to their original scale and reshaping to 2d*

y\_pred\_RNN = scaler\_y\_RNN.inverse\_transform(y\_pred\_scaled\_RNN.reshape(-1, 1))

*#inverse transform the scaled actual target values (y\_RNN\_test) back to their original scale*

y\_actual\_RNN = scaler\_y\_RNN.inverse\_transform(y\_RNN\_test.reshape(-1, 1))

*#printing the first 5 actual and predicted values, formatted to 4 decimal places for comparison*

for i in range(5):

print(f"Actual: {y\_actual\_RNN[i][0]:.4f}, Predicted: {y\_pred\_RNN[i][0]:.4f}")

*#plotting loss & MAE curves*

plt.figure(figsize=(12, 4))

plt.subplot(1, 2, 1)

plt.plot(history\_RNN.history['loss'], label='Train Loss')

plt.plot(history\_RNN.history['val\_loss'], label='Validation Loss')

plt.legend()

plt.title("Loss Over Epochs")

plt.subplot(1, 2, 2)

plt.plot(history\_RNN.history['mae'], label='Train MAE')

plt.plot(history\_RNN.history['val\_mae'], label='Validation MAE')

plt.legend()

plt.title("MAE Over Epochs")

plt.show()

*#plotting the actual and predicted values*

plt.figure(figsize=(10, 6))

plt.scatter(range(len(y\_RNN\_test\_original)), y\_RNN\_test\_original, color='blue', label='Actual y\_RNN\_test', alpha=0.6)

plt.scatter(range(len(y\_pred\_RNN\_original)), y\_pred\_RNN\_original, color='red', label='Predicted y\_pred\_RNN', alpha=0.6)

plt.title("Actual vs Predicted Values (RNN Model)")

plt.xlabel("Index")

plt.ylabel("Value")

plt.legend()

plt.show()