CRIME ANALYSIS AND PREDICTION BASED ON ONLINE MEDIA NEWS USING DEEP LEARNING ALGORITHM

THESIS

A Partial Requirement To Fulfill For Master Degree In Computer Science

HERI SUPRIHATIN

NIM : 20200130009

Supervisor:

Prof. Ir.Teddy Mantoro, MSc., PhD., SMIEEE

Prof. Ir. Media Anugerah Ayu, MSc., PhD., SMIEEE

SCHOOL OF COMPUTER SCIENCE

NUSA PUTRA UNIVERSITY

2023

AUTHOR'S STATEMENT

Tittle	:	Crime Analysis And Prediction Based On Online Media News Using Deep Learning Algorithm
Name	:	Heri Suprihatin
ID of student	:	20200130009

"I solemnly declare and assume responsibility that this thesis is my own work, except for excerpts and summaries, each of which I have explained the source of. If at a later time there are other parties who claim that this Thesis is his work, which is accompanied by sufficient evidence, then i am willing to cancel my Master of Computer Degree along with all rights and obligations attached to the title.".

Sukabumi, 16th July 2023

Stamp

Heri Suprihatin

Writer

APPROVAL OF THESIS

Tittle	:	Crime Analysis And Prediction Based On Online Media News Using Deep Learning Algorithm
Name	:	Heri Suprihatin
ID of student	:	20200130009

This thesis has been reviewed and approved

Sukabumi, 16th July 2023

Head of Study Program,	Supervisor,
Prof. Ir. Teddy Mantoro, M.Sc., PhD	Prof. Ir. Teddy Mantoro, M.Sc., PhD
NIP.	NIP.

THESIS APPROVAL

The Tittle of Thesis	:	Crime Analysis And Prediction Based On Online Media News Using Neural Deep Learning Algorithm
Name	:	Heri Suprihatin
ID of student	:	20200130009

This Thesis has been tested and defended in front of the Board of Examiners in Thesis session on 16 July 2023. In our review, this Thesis adequate in terms of quality for the purpose of awarding the Master of Computer Degree.

Sukabumi, 16th July 2023

Supervisor I	Examiner 1
Prof. Ir. Teddy Mantoro, M.Sc., PhD NIP.	Dr. Rahmadya Trias Handayanto, Ph.D NIP.
Supervisor 2	Examiner 2
Prof. Ir. Media Anugerah Ayu, M.Sc., PhD NIP.	Dr. Haris Al Qodri Maarif, Ph.D NIP.

PUBLICATION APPROVAL

As a member of the academic community of Nusa Putra University, i undersigned:

Name	:	Heri Suprihatin
ID of student	:	20200130009
Study Program	:	Computer Science
Type of Work	:	Thesis

For the sake of scientific development, agree to grant to the University of Nusa Putra the Non-Exclusive Royalty-Free Right for my scientific work entitled:

Crime Analysis And Prediction Based On Online Media News Using Deep Learning Algorithm.

Along with existing devices (if needed) With this non-exclusive royalty-free right, Nusa Putra University has the right to store, transfer media/formats, process in the form of a database, maintain and publish my thesis as long as I keep my name as the author/creator and as the copyright owner.

This statement I made in truth.

Made in : Sukabumi

At the Date of : 16th July 2023

That States

Heri Suprihatin

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FOREWORD

Praise and gratitude the author prays to God Almighty, because only with His blessings and grace can the author complete this thesis. Writing this thesis is one of the requirements to achieve a Master's degree in Computers at Nusa Putra University. I realize that, without the help and guidance of various parties, from the lecture period to the preparation of this thesis, it is very difficult for the author to complete this thesis. Therefore, I would like to thank:

1. Dr. Kurniawan, ST., M.Sc., MM as Chancellor of Nusa Putra University;
2. Anggy Pradiftha Junfitrana, MT as Vice Chancellor 1 for Academic Affairs;
3. Prof. Ir. Teddy Mantoro, M.Sc., PhD as Head of School Computer Science Nusa Putra University and Supervisor;
4. Prof. Ir. Teddy Mantoro, M.Sc., PhD and Prof. Ir. Media Anugerah Ayu, M.Sc., PhD as Supervisor;
5. All Masters Of Computer Science Lecturers who have provided very useful knowledge during lectures;
6. Parents for their services, patience, prayers and never getting tired of educating and giving sincere love to the author since childhood, both material and non-material;
7. Fellow comrades in Master of Computer Science 2020 who always give encouragement and always accompany from the beginning of the lecture until now;
8. All parties who have helped the author in writing this thesis.

For further improvement, suggestions and constructive criticism will be gladly accepted. Finally, only to Allah SWT the author submits everything, hopefully it can be useful especially for writers in general for all of us.

Sukabumi, 16th July 2023

Writer,

ABSTRACT

Theft, murder, assault, violence, rape, obscenity, kidnapping, vandalism, arson, incarceration, narcotics, fraud, corruption, embezzlement, and public order are all forms of crimes that can induce psychiatric illnesses in anybody. These are all natural phenomena and events that can have an impact on social circumstances and human existence. As a result, criminal activities performed by law enforcement agents must be prevented, and perpetrators must be apprehended and punished in accordance with their crimes.

The goal of this project is to use supervised learning to do analysis and prediction in order to monitor and detect all probable crimes in Indonesia. The neural network method is provided and a deep learning algorithm approach is used in this study to display crime data in order to obtain optimum crime data accuracy. Using this method, we can extract previously unknown and valuable information from unstructured data. Data processing activities and approaches include convolutional neural networks and recurrent neural networks.

Keywords: supervised learning, deep learning algorithm, convolutional neural network and recurrent neural network, neural network.

CHAPTER I

INTRODUCTION

1.1 Background

According to police data, the number of criminal events (total crime) in 2018 was 294,281, reduced to 269,324 in 2019, and decreased to 247,218 in 2020. The number of incidences of crime or criminal acts in Indonesia fluctuated between 2018 and 2020 (*Statistik Kriminal 2021*, n.d.)

The previous year's data also revealed a significant increase in the number of criminal incidents, particularly in the last three years, but also a modest decrease. The table below shows crime incidence data generated from BPS data for the years 2018 to 2020.

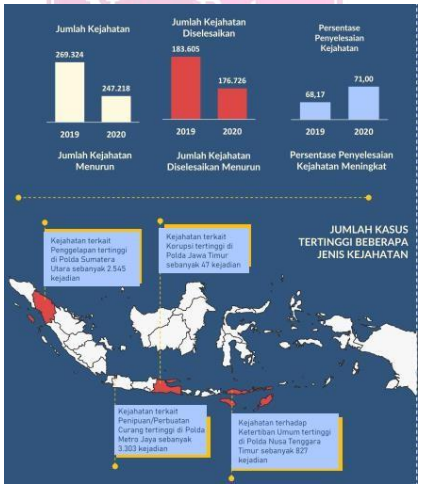

Figure 1.1. Crime incident in indonesia (*Statistik Kriminal 2021*, n.d.)

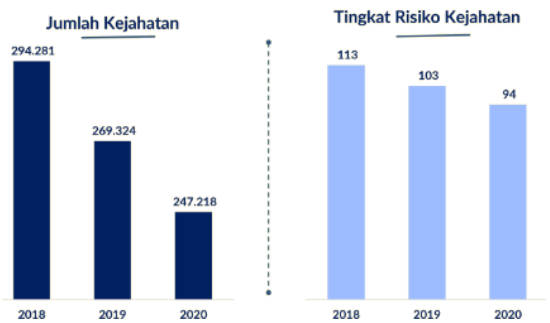

Figure 1.2. Crime rates and crime risk levels from 2018 to 2020

The total number of crime occurrences in 2019 was 269,324; this figure fell further in 2020 and 2021, to 247,218 incidents and 239,481 incidents, respectively. The crime rate per 100,000 residents has also continued to fall between 2019 and 2021, reaching 90 in 2021. This number fell from 94 in 2020 to 103 in 2019 (Pusat Statistik, n.d.). Based on the number of crimes seen to fall year after year, it was determined that the year 2018-2023 was the year when the crime rate decreased dramatically. More accurate research in the analysis and prediction of future crimes is required to speed the reduction in the number of crimes.

Figure 1.3. Crime rates and crime risk levels from 2019 to 2021

Crime has evolved into a global issue, with intricate linkages with location, time, and the environment (Albo, n.d.). With its ability to collect richer context information, neural network-based models have also been widely employed in charge prediction, and therefore the accuracy of charge prediction has been substantially improved due to neural networks' self-learning and self-adaptability (X. Li et al., 2020). Researchers have focused their efforts mostly on assessing criminal activity from a geographical standpoint. A recurrent neural network is utilized to handle the temporal aspects of crime prediction. A convolutional neural network is fitted for the spatial elements of crime prediction (Meskela et al., 2020).

Deep learning provides an excellent framework for simplifying and speeding up processes. We can design recurrent neural networks and anticipate outcomes using deep learning. Keras may be used to create these neural networks. These neural networks may be

created with Keras (API) and Python (Krishnan et al., n.d.). While RNNs and CNNs capture different aspects of the problem, they can be merged to form a single network that captures both. The most relevant spatial features are set in a grid and sent through the CNN layers one at a time, with the CNN output being merged with the rest of the features. The generated vector is then treated as input by the RNN (Stec & Klabjan, 2018).

1.2 Motivation

Because of their negative impact on human lives, the economy, and safety, criminal activities have become a huge social problem (Rayhan & Hashem, 2020). Criminality is a negative phenomenon that occurs in both industrialized and developing countries around the world (Safat et al., 2021). Crimes are classified into two types: violent crimes and nonviolent crimes (Tariq et al., 2021). For a long time, crime prediction has been a popular academic topic. In recent years, there have been numerous breakthroughs. Deep learning, a new trending field, will also be covered. The most popular trending research subjects are social media posts and data analysis. The section that follows describes key connected works (Lloret Mauri et al., n.d.). Predictioning and forecasting crime serves two purposes. First, preventive actions should be taken, and law enforcement resources should be allocated wisely. Second, to aid the criminal justice system in making individual decisions (Chun et al., 2019a).

For decades, experts have focused on data-driven crime prediction challenges. Existing studies on crime prediction can be classified as follows: I crime rate inference, which predictions a region's crime rate; (ii) crime hotspot discovery, which identifies areas where crimes cluster; and (iii) crime occurrence prediction, which predictions the occurrence of a crime category for a location at a future timestamp (Rayhan & Hashem, 2020). To compensate for the shortcomings of classic approaches, deep learning-based classification models have lately gained popularity. The majority of these models have complicated neural network

topologies. Convolutional neural networks and recurrent neural networks are two classic neural network models (X. Li et al., 2020).

The framework's objective is to analysis the dataset linked to criminal records in various locations and anticipate the possible types of crime that may occur in various areas. Because there is a large amount of crime data available, crime prediction is a critical issue for the police agency. There is a need for technologies that will allow for faster case resolution (Dhanalakshmi & Jacob, 2021). The following are the reasons for doing this research: (1) There have been numerous studies using neural network techniques. Some are employed for research purposes. However, because this is merely an implementation of the algorithm, it cannot be predictioned precisely at this time. (2) For data processing, many research have used neural network techniques. However, researchers are still attempting to anticipate future crimes by comparing multiple algorithms and visualizing crime rates using internet media stories as initial crime data.

1.3 Problem Statement

Artificial intelligence can prevent programs meant to avoid potential crimes that are occurring or will occur. This is a problem that must be addressed soon in order to prevent crimes from occurring. The country of Indonesia has a high rate of crime. There were 294,281 events in 2018, which reduced to 269,324 incidents in 2019 and 247,218 incidents in 2020. The number of incidences of crime or criminal acts in Indonesia fluctuated between 2018 and 2020. The total number of crime occurrences in 2019 was 269,324; this figure fell further in 2020 and 2021, to 247,218 incidents and 239,481 incidents, respectively. The crime rate per 100,000 residents has also continued to fall between 2019 and 2021, reaching 90 in 2021. This number fell from 94 in 2020 to 103 in 2019. It is difficult to assess the effectiveness of crime prevention, and its location remains unknown. In this study, crime data is collected using a deep learning approach via crime news on an internet media site employing grabbing url techniques and a

deep learning algorithm. This study focuses on getting optimal crime data accuracy, such as convolutional neural networks and recurrent neural networks.

1.4 Research Questions

1. To minimize crime rates, learn how to recognize and forecast future crimes.
2. How can you apprehend the perpetrators of crimes that can assist law enforcement authorities in an effective manner by following a pattern and type of crime?
3. How do I discover information about crimes committed via internet media websites?

1.5 Research Objective

The goal of this research is to assess and anticipate crime by collecting crime news from internet media websites from the previous year to the present. This trained model will be used for crime analysis and prediction in future artificial intelligence investigative techniques. Using the proposed technique, this study aims to generate reliable crime predictions in cities and districts. The following are some points to consider about study objectives:

1. To determine the future prediction of crime.
2. In the next crime prevention step, determine the level of criminal accuracy.
3. To evaluate and train the performance of deep learning engineering models, as well as to experiment with criminal data.
4. To compare algorithms and analysis the accuracy of the results of the proposed model by providing the best recent work results.

1.6 Significance of Research

This study compares deep learning algorithms to proposed convolutional neural networks and recurrent neural networks techniques to achieve the best crime prediction accuracy . Knowing the performance of each method, it is implemented as a visualization of crime predictions using a neural network and a PHP-based web application.

1.7 Research Scope

The primary goal of this study is to analysis and prediction crime data utilizing internet news media. Using the feature selection method to assess the strength of numerous crime data highlights in order to identify crimes in online media news This study will address the issue of crime in society. Predictioning and analyzing crimes will provide substantial knowledge into the crime area, which can then be used to take the required steps to prediction crime ratios. This research focuses on crime data analysis and prediction modeling using the Deep learning algorithm, which has been identified as the best effective algorithm for crime recognition in related literature.

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS

CHAPTER V

CONCLUSION AND FUTURE WORK

5.1 Conclusion

The study's goal was to assess existing methods for predictioning index crime, beginning with the basic ways currently utilized by researchers to forecast crime and evaluating each prediction methodology based on accuracy, data needs, neural network requirements, and ease of use. The most essential conclusion from this study is that, while technology has increased our ability to create, preserve, and manage data, much work remains to be done before we can accurately forecast crime trends.

This study intends to expand on the model's explanation by exposing it to other types of data, such as location images, and so on, providing a more advanced experimental aspect in a more efficient test environment using the best learning algorithm that will be specified after an in-depth comparative study.

5.2 Future Work

In future work experiments, algorithms such as LSTM and GIS will be used to generate geographic data about crime and criminals, as well as to integrate data from various sources, to demonstrate the perceived benefits of using time series, particularly since the dataset has a seasonal time series structure.

REFERENCES

Albo, F. (n.d.). *A Survey of Research into Artificial Neural Networks for Crime Prediction*. https://www.researchgate.net/publication/336116157

Alkaff, M., Fatanah Mustamin, N., Aditya, G., & Firdaus, A. (n.d.). International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Prediction of Crime Rate in Banjarmasin City Using RNN-GRU Model. In *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE* (Vol. 2022, Issue 3). www.ijisae.org

Butt, U. M., Letchmunan, S., Hassan, F. H., & Koh, T. W. (2022). Hybrid of deep learning and exponential smoothing for enhancing crime forecasting accuracy. *PLoS ONE*, *17*(9 September). https://doi.org/10.1371/journal.pone.0274172

Chand Bansal, J., Kusum, ·, Nagar, A. K., Goyal, D., Chaturvedi, P., & Purohit, S. D. (n.d.). *Algorithms for Intelligent Systems Series Editors: Proceedings of Second International Conference on Smart Energy and Communication*. http://www.springer.com/series/16171

Chun, S. A., Pathak, R., Paturu, V. A., Atluri, V., Yuan, S., & Adam, N. R. (2019a). Crime Prediction Model using Deep Neural Networks. *ACM International Conference Proceeding Series*, 512–514. https://doi.org/10.1145/3325112.3328221

Chun, S. A., Pathak, R., Paturu, V. A., Atluri, V., Yuan, S., & Adam, N. R. (2019b). Crime Prediction Model using Deep Neural Networks. *ACM International Conference Proceeding Series*, 512–514. https://doi.org/10.1145/3325112.3328221

Dhanalakshmi, R. A., & Jacob, G. (2021). An Exploration of Crime Type and Prediction Using RALASD Feature Selection Algorithm with Deep learning algorithm. In *Turkish Journal of Computer and Mathematics Education* (Vol. 12, Issue 11).

Esquivel, N., Nicolis, O., Peralta, B., & Mateu, J. (2020). Spatio-Temporal Prediction of Baltimore Crime Events Using CLSTM Neural Networks. *IEEE Access*, *8*, 209101–209112. https://doi.org/10.1109/ACCESS.2020.3036715

Gnanasigamani, L. J., & Hari, S. (2019). Deep Learning Based Crime Investigation Framework. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, *8*(11). www.ijstr.org

Hohman, F., Kahng, M., Pienta, R., & Chau, D. H. (2019). Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers. *IEEE Transactions on Visualization and Computer Graphics*, *25*(8), 2674–2693. https://doi.org/10.1109/TVCG.2018.2843369

Hussein, H., & Abdulameer, A. (2022, March 3). *Crime Prediction Using Big Data Analysis*. https://doi.org/10.4108/eai.7-9-2021.2314943

Jeyaboopathiraja, J., & Priscilla, G. M. (n.d.). *J JEYABOOPATHIRAJA AND G MARIA PRISCILLA: OPTIMIZED FEATURES AND DEEP LEARNING BASED CRIME*

TRENDS PREDICTION OPTIMIZED FEATURES AND DEEP LEARNING BASED CRIME TRENDS PREDICTION. https://doi.org/10.21917/ijsc.2021.0328

Kang, H. W., & Kang, H. B. (2017). Prediction of crime occurrence from multimodal data using deep learning. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0176244

Krishnan, A., Sarguru, A., & Shantha Sheela, A. C. (n.d.). *PREDICITVE ANALYSIS OF CRIME DATA USING DEEP LEARNING*. http://www.ijpam.eu

Li, J., Wang, S., Qin, S., Li, X., & Wang, S. (Eds.). (2019). *Advanced Data Mining and Applications* (Vol. 11888). Springer International Publishing. https://doi.org/10.1007/978-3-030-35231-8

Li, X., Kang, X., Wang, C., Dong, L., Yao, H., & Li, S. (2020). A Neural-Network-Based Model of Charge Prediction via the Judicial Interpretation of Crimes. *IEEE Access*, *8*, 101569–101579. https://doi.org/10.1109/ACCESS.2020.2998108

Lira Cortes, A. L., & Fuentes Silva, C. (2021). Artificial Intelligence Models for Crime Prediction in Urban Spaces. *Machine Learning and Applications: An International Journal*, *8*(1), 1–13. https://doi.org/10.5121/mlaij.2021.8101

Lloret Mauri, J., IEEE Communications Society, IEEE Systems, M., Annual IEEE Computer Conference, International Conference on Advances in Computing, C. and I. (ICACCI) 4 2015. 08. 10-13 K., International Symposium on Emerging Topics in Circuits and Systems (SET-CAS) 2015.08.10 Kochi, K., International Symposium on Control, A., International Symposium on Recent Advances in Medical Informatics (RAMI) 4 2015.08.10 Kochi, K., & International Symposium on Natural Language Processing (NLP) 4 2015.08.13 Kochi, K. (n.d.). *International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015 10-13 Aug. 2015, SCMS, Aluva, Kochi, Kerala, India ; [including co-affiliated symposia]*.

MERCon 1. 2015 Moratuwa, Jayasekara, A. G. B. P., University Moratuwa, Institute of Electrical and Electronics Engineers Sri Lanka Section, Moratuwa Engineering Research Conference 1 2015.04.07-08 Moratuwa, & MERCon 1 2015.04.07-08 Moratuwa. (n.d.). *MERCon 2015 Moratuwa Engineering Research Conference : April 7-8, 2015, University of Moratuwa, Sri Lanka*.

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)*, *9*(4), 402–405. https://doi.org/10.35940/ijrte.D5025.119420

Mohamad Zamri, N. F., Md Tahir, N., Megat Ali1, M. S. A., Khirul Ashar, N. D., & Al-misreb, A. A. (2021). Mini-review of Street Crime Prediction and Classification Methods. *Jurnal Kejuruteraan*, *33*(3), 391–401. https://doi.org/10.17576/jkukm-2021-33(3)-02

Pratibha, st, & Lokesh Chouhan, th. (2020). Crime Prediction and Analysis 4 th Suraina Dhiman. In *2nd International Conference on Data, Engineering and Applications (IDEA)*. https://www.overleaf.com/project/5deb178c0230af000196cd16

Pusat Statistik, B. (n.d.). *Statistik Kriminal 2022*.

Raghavendra, R., & Niranjanamurthy, M. (2022). Web Information Extraction methods using Web Content Mining (WCM) for Webapplications. *International Journal of Computing and Digital Systems*, *11*(1), 609–623. https://doi.org/10.12785/ijcds/110149

Rayhan, Y., & Hashem, T. (2020). *AIST: An Interpretable Attention-based Deep Learning Model for Crime Prediction*. http://arxiv.org/abs/2012.08713

Safat, W., Asghar, S., & Gillani, S. A. (2021). Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep learning algorithms. *IEEE Access*, *9*, 70080–70094. https://doi.org/10.1109/ACCESS.2021.3078117

SCAD College of Engineering and Technology, & Institute of Electrical and Electronics Engineers. (n.d.). *Proceedings of the International Conference on Trends in Electronics and Informatics (ICOEI 2019) : 23-25, April 2019*.

Schlegel, U. (2018). *Towards Crime Forecasting Using Deep Learning*. http://nbn-resolving.de/urn:nbn:de:bsz:352-2-hj2u46k5bsms8

Stalidis, P., Semertzidis, T., & Daras, P. (2021). Examining Deep Learning Architectures for Crime Classification and Prediction. *Forecasting*, *3*(4), 741–762. https://doi.org/10.3390/forecast3040046

Statistik Kriminal 2021. (n.d.).

Stec, A., & Klabjan, D. (2018). *Forecasting Crime with Deep Learning*. http://arxiv.org/abs/1806.01486

Tariq, H., Hanif, M. K., Sarwar, M. U., Bari, S., Sarfraz, M. S., & Oskouei, R. J. (2021). Employing deep learning and time series analysis to tackle the accuracy and robustness of the forecasting problem. *Security and Communication Networks*, *2021*. https://doi.org/10.1155/2021/5587511

Tekin, S. F., & Kozat, S. S. (2021). *Crime Prediction with Graph Neural Networks and Multivariate Normal Distributions*. http://arxiv.org/abs/2111.14733

Tony, B., Savarimuthu, R., Arulanandam, R., & Purvis, M. A. (2014). *Extracting crime information from online newspaper articles MASAE (Multi-Agent Self-Adaptive Environments): A Multi-Agent Proactive Approach for the Development of Context Aware Self-Adaptive Systems with Application in Pervasive Computing View project PhD project View project Extracting Crime Information from Online Newspaper Articles*. https://www.researchgate.net/publication/259932789

Vimala Devi, J., & Kavitha, K. S. (n.d.). Automating Time Series Forecasting on Crime Data using RNN-LSTM. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 12, Issue 10). www.ijacsa.thesai.org

Walczak, S. (2021). Predictioning Crime and Other Uses of Neural Networks in Police Decision Making. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.587943

Wang, B., Yin, P., Bertozzi, A. L., Brantingham, P. J., Osher, S. J., & Xin, J. (2017). *Deep Learning for Real-Time Crime Forecasting and its Ternarization*. http://arxiv.org/abs/1711.08833

Wang, B., Zhang, D., Zhang, D., Brantingham, P. J., & Bertozzi, A. L. (2017). *Deep Learning for Real Time Crime Forecasting*. http://arxiv.org/abs/1707.03340