EXAMPLE OF DRAFT 0

Abstract. Public violence is a major health problem in the United States. Incidents involving violent crimes are often not reported to law enforcement (LE). The Cardiff Model is a violence prevention program developed in the UK that combines violent injury information from Emergency Rooms (ER) and LE. The model is now in use in several major cities in the US to reduce violence. Las Vegas has seen a significant increase in public violence since the 2017 Route 91 Harvest music festival shooting. As a result, the Southern Nevada Health District and UNLV researchers believe the Cardiff Model is a viable solution to address this public health crisis. This research explores natural language processing and machine learning models to extract violence injury location information from ER records in preparation for implementing the Cardiff Violence Prevention Model in Clark County, Las Vegas.

1 Introduction

Public Violence Prevention is a major area of research because of the increase in violent injuries in recent years. The city of Las Vegas has seen a significant increase in public violence since 2017 and considers it a public health crisis.

The Cardiff Model is a violence prevention program that was developed by an ER physician in the UK to combine violent injury information from ERs and LE to make improvements in the community to reduce violence. The Cardiff Model has been replicated in the U.S. in Atlanta, Milwaukee, Philadelphia, and other cities. Researchers in Las Vegas believe the Cardiff Model is a viable solution to address this public health crisis.

In previous implementations of the Cardiff Model, nurses were trained to collect injury information including the location of the injury. They input the data into special Cardiff Model Screening Tools (CMST). This requires staff training for nurses and ongoing efforts to collect the information. It is problematic because the CMSTs are not integrated with the hospital electronic medical records.

This research aims to use Natural Language Processing (NLP) to solve the problem of training and system integration by using existing ER records to automate the identification of injury locations and hotspots.

2 Literature Review

The literature review focuses on four principal areas: The Cardiff Model, NLP methods in clinical settings, Named Entity Recognition (NER) in clinical data, and location-based NER in non-clinical data.

2.1 The Cardiff Model

Violence is a major problem in the United States. An estimated 50% of violent injuries are never reported to law enforcement. The Cardiff Model is a solution developed in the UK for enhancing data collection and sharing between ERs and local LE to identify community improvements that could result in the reduction of violent injuries (Kollar et al., 2020). Two cities in the UK, Cardiff and Merseyside, showed significant reductions of violent injuries by 42% and 36% respectively after implementing the Cardiff Model. Several studies have reviewed other implementations. Kollar et al evaluated the implementation of the model in Atlanta, GA . Boyle et al explored the implementation of the model in Cambridge, England (Boyle et al., 2013).

The Atlanta study focused on how the model was implemented, the impact on hospital staff to collect the data, and the results of sharing the data with local LE. The study identified one area in Atlanta to make community improvements. It identified businesses and public spaces to improve lighting, add security cameras, add patrols, and support youth programs. The study did not assess whether the changes identified in the communities were made or whether it had an impact of reducing violence.

The Cambridge study focused on data collection, data sharing, and following the results over several years to see if data sharing resulted in fewer violent injuries. They found that there were fewer injuries reported to LE, but not a statistically significant reduction in violent injuries admissions to the ER. Even though there was not a targeted region to make community improvements or a specific action plan, the data did inform various community decisions. For example, a liquor license was denied for an area of Cambridge that had a homeless shelter and a high number of alcohol-related violent injuries.

Both studies involved training hospital staff and required system upgrades to support the collection of data. The Atlanta study used nurses to collect the data while the Cambridge study utilized receptionists. Both implementations collected the date/time of the injury, location of the injury, and the type of assault and the weapon used.

The Cardiff Model provides an opportunity to make data-driven community improvements. Previous implementations have required special training of staff and system enhancements to support the collection of data. Both are potential barriers for hospitals and health agencies that want to take advantage of the Cardiff Model but do not have the resources to change existing processes. This research aims to remove both barriers by collecting data in an automated way.

2.2 NLP in Medical Records

Medical records often contain unstructured data in the form of medical notes. NLP can be utilized to extract meaningful information from medical records. One study reviewed numerous research papers on using NLP on clinical data (Spasic & Nenadic, 2020). Many of the studies used text classification for prognosis, care improvement, resource management, and surveillance. Of the 110 studies reviewed, the main problem identified was the need for manually labeling data. Most studies were limited to hundreds or thousands of records due to the manual effort required to label the data. Also, because of the sensitive nature of health care data, the models trained usually only involved one hospital or facility and did not generalize well when testing models on other sources. The authors concluded that more exploration can be done in data augmentation, transfer learning, and distance learning to address the annotation problem. Unsupervised models could avoid the labeling problem altogether.

Violence is a major health concern resulting in the loss of life. It also has a significant economic impact costing $671 billion per year. The Centers for Disease Control and Prevention (CDC) has acknowledged that data science, especially NLP, is a growing area that could help reduce or prevent injuries and violence (Ballesteros et al., 2020). Ballesteros et al review approaches that the CDC will take to implement data science to reduce violence. The researchers note that the geographic prediction of violent crime and injuries is a critical area for identifying health threats in communities. However, system limitations often result in analyses on stale data. They also point out that manual efforts to label data have been a barrier in the past. NLP provides an opportunity to overcome both issues.

2.3 NER for Clinical Text

Conditional Random Fields (CRFs) and other supervised learning models like Support Vector Machines (SVMs), Structural Support Vector Machines (SSVMs), and Hidden Markov Models (HMMs) have been commonly used for NER in clinical texts. The primary downside of these models is that they require human feature engineering. Recurrent Neural Networks (RNNs) is a deep learning methodology that captures long-term dependences in sequence data. They can provide an alternative for NER on clinical texts without the manual effort of feature engineering. Several studies of NER on clinical text have shown that RNN outperforms other models in accuracy and F1 scores.

The research of Liu et al (Liu et al., 2017) and Wu et al (Wu et al., 2017) focused on entity recognition using Long Short-Term Memory (LSTM), a variant of RNN. For both studies, the models produced labels for clinical entities such as disease names, types of lab tests, treatments, and medication names. Liu's study also identified protected health information (PHI). In both studies, the LSTM models performed better than other models tested.

Liu's research compared several LSTM models with CRFs and HMMs using clinical notes from i2b2 datasets. One model used only a word-level input layer. Another model used both word- and character-level inputs. Using both inputs resulted in higher scores. The best LSTM model contained three layers: an input layer consisting of a representation of every word at the token and character level; an LSTM output layer with the context of each word; and an inference layer that outputs a label. The token-based input layer uses a continuous bag-of-words (CBOW) and skip-grams while the character-based input uses a bi-directional LSTM that captures the past and future contexts of words. Within the LSTM layer there are three propagative gates: an input gate, a forget gate, and an output gate. The main function of these gates is to control the proportion of information transferred to a memory cell. The inference layer uses a CRF to estimate a label from a sequence of context resemblances.

Wu et al also used an i2b2 dataset containing medical notes for discharge, radiology, electrocardiogram (ECG), and echocardiograms (ECHO). A Convolutional Neural Network (CNN) and an RNN model with word embeddings were tested against the more traditional CRF model. The word embeddings were pre-trained on the MIMIC II dataset. The RNN with bi-directional LSTM had the best overall F1 score. The researchers noted that word embeddings inherent with deep learning models are better at identifying related concepts than CBOW models because related concepts often do not contain overlapping words (e.g., "mildly dilated right atrium" and "somewhat enlarged left ventricle").

Deep learning is a promising method for NER for clinical texts. It offers several advantages including replacing CBOWs with word embeddings, eliminating manual feature engineering, and addressing long-term dependencies. While this model outperformed traditional state-of-the-art methods in the clinical domain, it does not always perform better in other domains. It also may not perform as well for identifying location-based entities, which is the focus of the current research.

NER has been widely used to analyze medical records due to its unstructured text while other ordinary statistical tools have failed. A combination of feature engineering and standard ML algorithms such as CRF and SVM needed to be effectively extracted from the information. Having a good feature engineering consumes a lot of manual tasks and time consuming which is very inefficient. A deep learning algorithm, especially RNN has proven to be effectively eliminated these manual tasks by learning the effectiveness of the feature automatically. Inigo et al. compared the two RNN methods: Bidirectional LSTM-CRF and Bidirectional LSTM against other different RNN models and the state-of-the-art systems for Drug Name Recognition (DNR) and Clinical Concept Extraction (CCE). Bidirectional LSTM is using the concatenation of both sides of generated input sentences to produce the final representation whereas Bidirectional LSTM-CRF is the resulting network of joining decoded input sequence in Viterbi-style manner. Both these Bidirectional neural network methods show improvement over the baseline CRF model with high F1-score for DNR and CCE (Inigo et al. 2017). Adding manual handcraft features to these neural network algorithms do not enhance their performance due to the fact the neural networks are able to learn automatically from the pre-trained data which in turn save time and manual tasks of feature engineering.

The current ML-based approaches such CRF and SVM have remarkable success in extracting information from clinical notes in EMR (Electronic Medical Records). However, these ML need many annotated corpora which require manual tasks from domain experts such as nurses or physicians. This effort is time-consuming and not cost-effective. Several active learning strategies have shown success in solving the issue. Active learning models could reduce cost and improve performance when comparing to passive learning approach. The active learning would use NER to select informative sequences from the pool. In order to measure the informativeness of sequences, different methods such as N-best sequence entropy, POS tag, NP chunk and using its words itself to check the similarity between sentences were used. Yukun et al. compare 13 active learning algorithms for clinical NER, six existing Al algorithms and seven new Al algorithms(Yukun et al. 2015). The two newly developed algorithms, uncertainty-based sampling methods: dynamic N-best sequence entropy and entity entropy outperform the baseline methods in term of area under the learning curve (ALC) scores. The dynamic N-best sequence entropy takes only the sum of the probability of the N-best sequence labels that are greater than 0.9(Yukun et al. 2015). The entity entropy sums all the entropies of B-entity words(Yukun et al. 2015). All these active learning algorithms are performing better than passive learning for a clinical NER task.

2.4 Location-Based NER

NER for location extraction is not a common area of research within the clinical domain. To evaluate methods for extracting location data, other types of text corpora have been used.

Location information is crucial during emergency situations or natural disasters. Twitter is one way to track the unfolding of a crisis in real time if the locations of user tweets can be identified. It has been observed that a user's location from Twitter data is often unknown or unreliable. One study predicted exact locations from tweets using a combination of word embeddings from location words, a CNN to extract key features, and a layer for interpreting features (Kumar et al. 2019). Their method produced a high F1 score, which was credited to the n-gram features of the CNN.

Another approach used a deep feedforward neural network to identify locations by checking whether a word or phrase is present in a pre-defined blacklist and whitelist created by Subject Matter Experts (SMEs) (Magge et al. 2018). The F1 score of this model was also extremely high.

In another study, the accuracy of NER to pinpoint locations in Twitter data degraded as the radius of predicted neighborhood increased. Collaboration between urban planners and machine learning models reduced the error of locating the predicted neighborhood. This collaboration, along with cosine similarity between embedded neighborhoods, can improve the accuracy within a 30-miles radius (Dutt et al. 2021).

Some research has looked at a chain or thread of tweets to extract location. While this provides more context and data than a single tweet, having long text may reduce accuracy due to an inability to distinguish and separate multiple entities. Current NER systems such as ANNIE, Stanford NER, NERD-ML, YODIE, and Alchemy API were tested against each other and all of them dropped in accuracy by 30-50% on long tweets (Derczynski et al. 2015). These NER systems were tested on three different datasets to minimize bias results. The NERD-ML system performed the best based on F1 score; the Stanford NER system was the second best.

Standard NER is not enough for geography NLP tasks due to the geographic ambiguity of toponyms (name of places). Geoparsing is a method of translating free-text toponyms into geographic coordinates. It is a basic component of Geographic Information Retrieval (GIR), Geographic Information Extraction (GIE), and Geographic Information Analysis (GIA) to extract the topography of a document. A pragmatic taxonomy is used to evaluate geoparsing. This taxonomy of toponyms uses two different taxonomy types: literal (where something is physically located) and associative (an association with a location) (Milan et al. 2020). It outperformed Google Cloud NLP and Space NLP in terms of F1 score. This is an improvement upon existing NER taggers which are unable to extract locations due to their inability to extract and classify the pragmatic types of toponyms.

Several methods have been used successfully to identify location information in Twitter data. These methods can be explored in medical text in the current research to determine the location of violent injuries.

2.2 Citations -READ ME

The list of references is headed "References" and is not assigned a number. The list should be set in small print and placed at the end of your contribution, in front of the appendix, if one exists. Please do not insert a page break before the list of references if the page is not completely filled. An example is given at the end of this information sheet. For citations in the text please use square brackets and consecutive numbers: [1], [2], [3], etc. Use **APA format** in the reference section. You can choose to either have it alphabetical order or order of which it is shown in the paper.

Annotated Bibliography

Hypothesis at the end of your literature Review

NLP and machine learning models can be used to extract and identify violent injury location information from the textual notes in medical records.

3 Methods

A. Data
   i. Where are you getting the data? Or where are you thinking you can find the data?

   - The data source is ER records from a system called Essence, which will be provided by the project sponsors at the Southern Nevada Health District (SNHD) and UNLV.

- Spatial geolocation APIs may also be used to identify latitude, longitude, neighborhood, and zip code data.

B. Methods plan to use

- Use NLP and machine learning models to extract violent injury location information from ER records.
- The data will help identify hotspots and locations for making community improvements and strengthening LE personnel in Las Vegas and help facilitate collaboration between the ER, public health agencies, and LE.

4 Results

A. What you hope to find in your research? Accept or reject the hypothesis
**This Section is for statistical jargon and tables/Figures. Results are facts.

This research aims to identify specific locations that violent injuries occur in Las Vegas based on ER hospital records. The goal is to share the information with local LE so that they can make community improvements to reduce violence.

5 Discussion

***Do not add New Results. This section is to apply and interpretate the results into lay terms.
*** Write questions you hope to answer in your research.

- Can location information be consistently extracted from ER records without requiring extra intervention, training, or processing by hospital staff?
- What location information is most relevant to LE in their efforts to reduce violence? There are different levels of granularity for location information (e.g., street name, full address, zip code, specific business name, etc.) and this research aims to identify the most useful information to LE.
- What level of aggregation is useful to LE for identifying hot spots? For example, how many violent injuries occur within a 1-mile radius?
- What is the best way to extract accurate locations when there are misspellings and location ambiguity in the data. Examples of where mismatches may occur include: "avenue" vs. "street", "3010 Awesome

Way" vs. "3001 Awesome Way", parking lot next to the gas station vs. parking lot next to 7-Eleven. What if the 7-eleven is out of business?

- Can other data sources like an index of street names in Las Vegas improve the accuracy/performance metrics?

A. Interpretations: What do the results mean?
B. Implications: Why do the results matter? How should the reader apply these findings?
C. What stood out as interesting/unique/unexpected?
D. Limitations
   a. What challenges occurred during analysis?
E. Ethics
   - Data used in this research will not contain any Personally Identifiable Information (PII). In cases where a violent injury location is identified to be a residential address, a method will be created to not disclose the address (e.g., zip code or neighborhood will be used as a proxy for address).

F. Future Research
   a. Are there areas of research where others can pick up and go deeper?

6 Conclusion

2 paragraphs max on the overall findings and summary of the research.

Acknowledgments. The heading should be treated as a 3rd level heading and should not be assigned a number.

References

1. Ballesteros, M. F., Sumner, S. A., Law, R., Wolkin, A., & Jones, C. (2020). Advancing injury and violence prevention through data science. *Journal of Safety Research; J Safety Res, 73*, 189-193. 10.1016/j.jsr.2020.02.018

2. Kumar, A., & Singh, J. P. (2019). Location reference identification from tweets during emergencies: A deep learning approach. *International Journal of Disaster Risk Reduction, 33*, 365-375. 10.1016/j.ijdrr.2018.10.021

3. Magge, A., Weissenbacher, D., Sarker, A., Scotch, M., & Gonzalez-Hernandez, G. (2018). Deep neural networks and distant supervision for geographic location mention extraction. *Bioinformatics; Bioinformatics, 34*(13), i565-i573. 10.1093/bioinformatics/bty273

4. Dutt, F., & Das, S. (2021). Fine-grained Geolocation Prediction of Tweets with Human Machine Collaboration.

5. Spasic, I., & Nenadic, G. (2020). Clinical Text Data in Machine Learning: Systematic Review. *JMIR Medical Informatics; JMIR Med Inform, 8*(3), e17984. 10.2196/17984

6. Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., Petrak, J., & Bontcheva, K. (2015). Analysis of named entity recognition and linking for tweets. Information Processing & Management, 51(2), 32-49. 10.1016/j.ipm.2014.10.006

7. Kollar, L. M. M., Sumner, S. A., Bartholow, B., Wu, D. T., Moore, J. C., Mays, E. W., Atkins, E. V., Fraser, D. A., Flood, C. E., & Shepherd, J. P. (2020). Building Capacity for Injury Prevention: A Process Evaluation of a Replication of the Cardiff Violence Prevention Program in the Southeastern United States. Injury Prevention, 26(3), 221-228. 10.1136/injuryprev-2018-043127

8. Liu, Z., Yang, M., Wang, X., Chen, Q., Tang, B., Wang, Z., & Xu, H. (2017). Entity recognition from clinical texts via recurrent neural network. BMC Medical Informatics and Decision Making; BMC Med Inform Decis Mak, 17, 67. 10.1186/s12911-017-0468-7

9. Wu, Y., Jiang, M., Xu, J., Zhi, D., & Xu, H. (2017). Clinical Named Entity Recognition Using Deep Learning Models. AMIA ...Annual Symposium Proceedings; AMIA Annu Symp Proc, 2017, 1812-1819

10. Milan, G., Pilehvar, M. T., & Nigel, C. (2020). A pragmatic guide to geoparsing evaluation. Language Resources and Evaluation, 54(3), 683-712. 10.1007/s10579-019-09475-3

11. Boyle, A. A., Snelling, K., White, L., Ariel, B., & Ashelford, L. (2013). External validation of the Cardiff model of information sharing to reduce community violence: natural experiment. Emergency Medicine Journal : EMJ; Emerg Med J, 30(12), 1020-1023. 10.1136/emermed-2012-201898

12. Chen, Y., Lasko, T. A., Mei, Q., Denny, J. C., & Xu, H. (2015). A study of active learning methods for named entity recognition in clinical text. Journal of Biomedical Informatics; J Biomed Inform, 58, 11-18. 10.1016/j.jbi.2015.09.010

13. Jauregi Unanue, I., Zare Borzeshi, E., & Piccardi, M. (2017). Recurrent neural networks with specialized word embeddings for health-domain named-entity recognition. Journal of Biomedical Informatics; J Biomed Inform, 76, 102-109. 10.1016/j.jbi.2017.11.007

14.