A Model for Conflicts' Prediction using Deep Neural Network

Olabanji B. Olaide
Department of Computer Science
University of Ibadan, Ibadan

Adebola K. Ojo
Department of Computer Science
University of Ibadan, Ibadan

ABSTRACT

Conflict is part of human social interaction, which may occur from a mere misunderstanding among groups of settlers. In recent times, advanced Machine Learning (ML) techniques have been applied to conflict prediction. Strategic frameworks for improving ML settings in conflict research are emerging and are being tested with new algorithm-based approaches. These developments have given rise to the need to develop a Deep Neural Network model that predicts conflicts. Hence, in this study, two Artificial Neural Network models were developed, the dataset which was extracted from https://www.data.worlduploaded by the Armed Conflict Location and Event Data Project (ACLED), in four separate CSV files (January 2015 to December 2018). The dataset for the year 2015 has 2697 instances and 28 features, for 2016 was 2233 with the same feature, for 2017 has 2669 instances with the same features, and 2018 has 1651 instances. After the development of the models: the baseline Artificial Neural Network achieved an accuracy of 95% and a loss of 5% on the training data and an accuracy of 90% and 10% loss on the test set. The Deep Neural Network Model achieved 98% accuracy and 2% loss on the training set, with 89% accuracy and 11% loss on the test set. It was concluded that to further improve the prediction of conflict, there is a need to address the issue of the dataset, in developing a better and more robust model.

Keywords

Artificial Neural Network, Conflict, Deep Neural Network, Multi-Class Target Label,Prediction, Model.

1. INTRODUCTION

Every event that occurs on this planet and beyond generates data.But they are intangible if information cannot be derived from the data; become useless, ambiguous, sometimes not what keeps. Well, when informative insights can be generated from these data in dealing with pressing matters that pertain to securing life and properties, then they become what a nation can spend half of Its budgets on in getting those insights for better decision-making. Conflict is a relationship dispute between an individual or a group of individuals, which may be due to sharing of resources or differences in ideology and thinking. When the population of people increased, different characters and behaviours spawn up, which in turnincrease clashes inrecent decades. The outcome of the conflict psychological features: constant grasp of fear, feeling powerless, experiencing fear and anger, and unbearable aggression toward particular ethnic or religious groups.

Nigeria with the highest population of black people in the world makes her prone to extremists. Starting from the dawn of the twenty-first century, Nigeria has experienced a wave of extremist attacks that have claimed the lives of its citizens. As stated in [1], "the United Nations estimated that 1.8 million people are internally displaced from Boko Haram's insurgency and 27,000 people have been killed, and thousands of women and girls abducted" as of 2019. In the last twenty years, the world leaders and the United Nations specifically, have looked into the way of developing strong, reliable, and effective early alert systems for conflict prevention [2]. Also, [3]stated that "Both peace and sustainable natural resource management are an integral part of the interlinked Sustainable Development Goals of Agenda 2030".

However, since it is imperative to tackle these psychological issues, the need to develop a model that predicts these conflicts is paramount to the physical and emotional well-being of society.[4] stated,that "Identification of conflict ideologies and prediction of future conflict attacks have been proven to be of great importance and time-consuming process.

Machine learning algorithms have been used recently to study the different factors of conflicts", but not yielding maximum results when they are faced with a large volume of data.

Artificial neural networks and deep neural networks are going mainstream because of the advancement in computational power and the large volume of label data available. These predictions are significant in understanding ways to implements conflict resolution strategies. They can make government and security agencies look for how conflicts can be resolved and protect the lives of the citizens.

2. RELATED WORKS

When studying conflict,many issues increase the likelihood that an event is going to happen, ranging from political opposition conflicting ideology, change in climatic condition of that particular geolocation, presence of natural resource and solid mineral, etc. [3].

The impact of conflict on the global economyis estimated to be 12.4 percent of world GDP ($14.6 trillion in purchasing power) as of 2017 [5]. [6] stated, identifying conflict areas in a precise and less expensive way is important to prevent or limit the economic and social impact of conflict. Moreover, predicting conflict depends on accessibility, correctness, volume, and cost of data collection; the selection of the best algorithm also affects the prediction[2]. Some algorithms give better performance on some problem over the other, selecting the best algorithm is imperatively important in other to achieve good prediction [3].[7] explored if adding geography can aid prediction of political violence more accurate using a simulated approach, of the predictive accuracy of the spatial-temporal model

which the results, compared to a standard regression model that only incorporated time lags. [8] worked on the how binary choice model performance as compared to supervised machine learning algorithm in predicting conflict events in sub-SaharanAfrica. These enable the policymaker to leverage the best algorithm for a particular event of a conflict. [3] worked on the impact of natural resources as the risk of violent conflicts with socio-economic conditions through the use of machine learning techniques. They compared four modelling approaches on four different datasets, from the four models they observed that natural resource features act as an early warning in violent conflict.

3. MATERIALS AND METHODS

The methodology for the proposed Deep Neural Network Model waspresented in this section. This is illustrated in Figure 1. In this section, data collection, data analysis and preprocessing, programming tools and implementation were discussed.
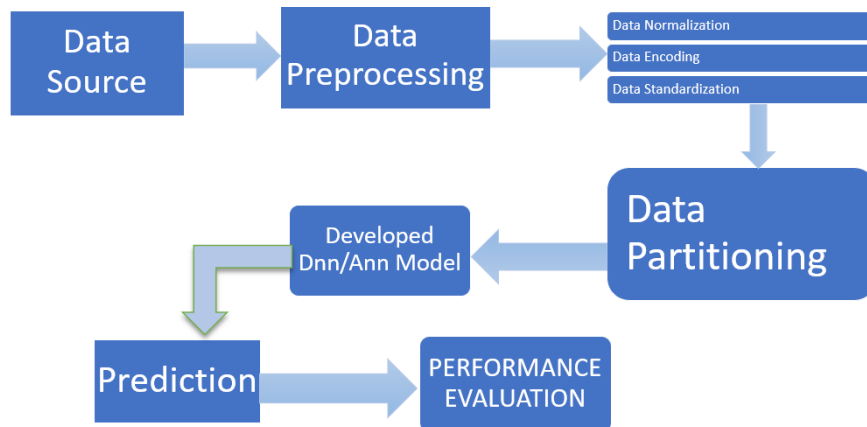

Figure 1: Developed Deep Neural Network and Artificial Neural Network Model

3.1 Data Collection

The dataset for this study was extracted from https://data.world. These were uploaded by the Armed Conflict Location and Event Data Project (ACLED), in four separate CSV files which were for four years between January 2015 to December 2018.

The dataset for the year 2015 has 2697 instances and 28 features, 2016 has 2233 with the same feature, 2017 has 2669 instances with the same features, while 2018 has 1651 instances. The total number of instances was 9258 rows. Table 1 described the features that are present in the combined dataset.

Table 1: Description of Dataset Used for the Developed Model

Features	Description
Data_Id	The identification for the data instance
Iso	The iso number
Event_Id_Cnty	An individual identifier by number and country acronym
Event_Id_No_Cnty	An individual numeric identifier
Event_Date	date of conflict, DD/MM/YYYY
Year	Year event occurred
Time Precision	A numeric code indicating the level of certainty of the date code for the event
Event Type	Event occurrence, string e.g., "Riots against police"
Actor1	Name of the first actor
Assoc_Actor_1	The accomplice of the event actor
Inter1	A numeric code indicating the type of ACTOR1
Interaction	A numeric code indicating the interaction between types of ACTOR1 and ACTOR2
Region	The region of the world that the event took place
Country	country of conflict
Admin1	The largest sub-national administrative region in which the event took place
Location	The location where the event occurred
Latitude	The latitude of the location
Longitude	The longitude of the location
Geo_Precision	A numeric code indicating the level of certainty of the geocode for the event
Source	Source of conflict information
Source_Scale	The coverage of the source
Notes	additional notes, string
Fatalities	The integer value of fatalities that occurred, as reported by the source
Timestamp	The timestamp for the event
Iso3	The country iso number
Location_2	Combination of both the latitude and longitude
Admin2	The second-largest sub-national administrative region in which the event took place
Admin3	The third-largest sub-national administrative region in which the event took place

3.2 Data Analysis and Preprocessing

After vivid looking into the dataset, it was discovered that the dataset has some missing values, features with string data type instead of nominal data type, multi-class target variable, unbalance data instances and inappropriate time-stamp format. These were reified by using python (fillna and encoding libraries) after which the dataset has reduced the features to 11 from 28 features. The data set was standardized, normalized for a faster and better model. It was partitioned into 70% and 30% for training and testing respectively. The event type which was a multi-class categorical column also, being the target column was encoded as binary vectors for prediction. The partitioned set was used to build a baseline ANN and the Deep Neural Network model as compared with the performance of the Models.

3.3 Programming Tool and Libraries

The following tools were used in the analysis of the dataset, selection of features, development and evaluation of the deep learning model: Python programming language via the google cloud computing services (research.colab.google.com) for all development and implementation. The Jupyter notebook is the IDE of choice. The libraries and packages used are: Pandas, Numpy, Tensorflow, Keras, Sci-kit learn, Seaborn and Matplotlib.

The Deep Neural Network Architecture was made up of the following:

- Data source from https://data.world by ACLED.
- Data preprocessing for data cleaning.
- Data transformation (encoding, normalization and standardization).
- Data partitioning (train and test set).
- Building the Deep Neural Network model.
- Backpropagation of error.
- Predicting multi-class instances of conflicts.
- Model Evaluation.

4. RESULTS AND FINDINGS

The model was able to achieve an accuracy of 98% during the learning which is shown in Figure 2 after running 1000 epochs while Figure 3 presents the loss of the deep neural network model showing training and validation sets.

4.1 Performance Evaluation

The metrics used in evaluating the performance of the model are the accuracy and the categorical loss function. The accuracy shows how the model is performing on prediction and the categorical loss function was able to show how well the model was, by effectively penalizingit in the event of a wrong decision.

In figure 2 the graph shows the performance of the Deep Neural Network Model accuracies with the training dataset and testing dataset after running the model for 1000 iteration, the accuracy for the two phases begins to rise after around 100 iterations and stabilized around 200 iterations for the test dataset but the accuracy in the training dataset still climbing high till the last iteration.

In figure 3 the graph shows the performance of the Deep Neural Network Model losses with the training dataset and test dataset after 1000 iterations, the loss for the two phases begins to step down around 150 iterations but for the training set the loss decrease on the last iteration, while the loss in the test dataset, on the other hand, begins to rise

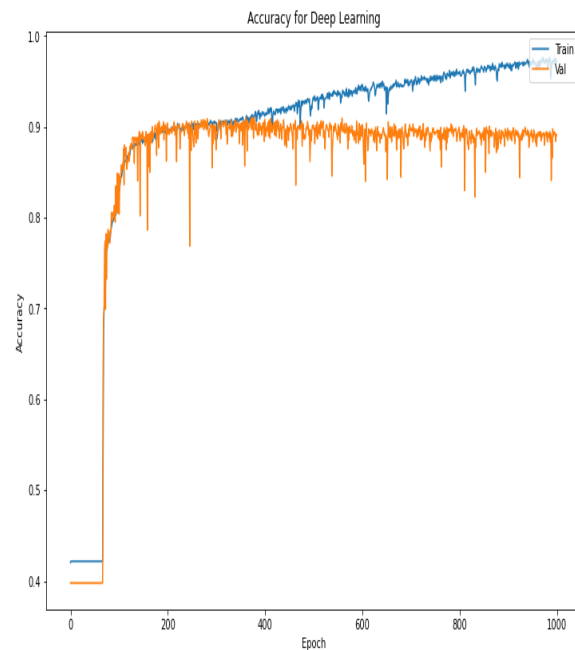again. This shows that the best iteration for this model should be around 200 iterations.

Figure 2: The accuracy of the deep neural network model showing training and validation set
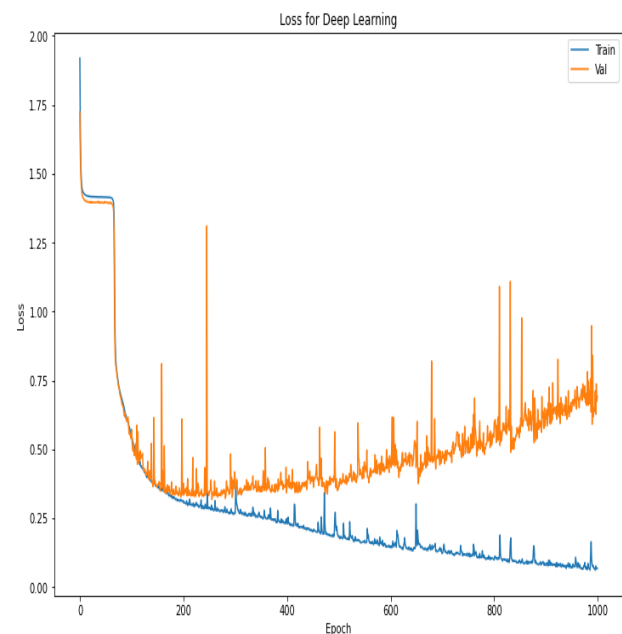

Figure 3: The loss of the deep neural network model showing training and validation set

4.2 Model Evaluation

The model has only one hidden layer. To see how our model is performing, a baseline artificial neural model was built which shows the significance of the dataset on the performance of the deep neural network model.

Table 2: Performance of the Artificial Neural Network Model

Data	Accuracy	Loss
TrainingData-ta	0.9485368132591248	0.051463186740875244
Test Data	0.901706874370575	0.09829312562942505

Table 3: Performance of the Deep Learning Model

Data	Accuracy	Loss
Training Data	0.979313850402832	0.02068614959716797
Test Data	0.8911124467849731	0.10888755321502686

Table 2 gives the interpretation of the baseline Artificial Neural Network Model, here the accuracy in the training phase was 94.85 per cent and the loss was 5.14 per cent respectively which give confidence that it will make a fine prediction when used for predicting new instance. As we can see in the testing phase the model was able to achieve 90.17 per cent accuracy and a loss of 9.8 per cent. This gives a remarkable result since the data set used are a multi-categorical dataset.

Likewise, Table 3 shows the performance of the Deep Neural Network Model, here the accuracy for the training phase was 97.93 per cent and loss of 2.06 per cent. Moreover, the result when tested with the new dataset was 89.11 per cent time accurate and then 10.88 per cent mistake. This shows that table 2 result performed better than table 3 when trained with a small dataset.
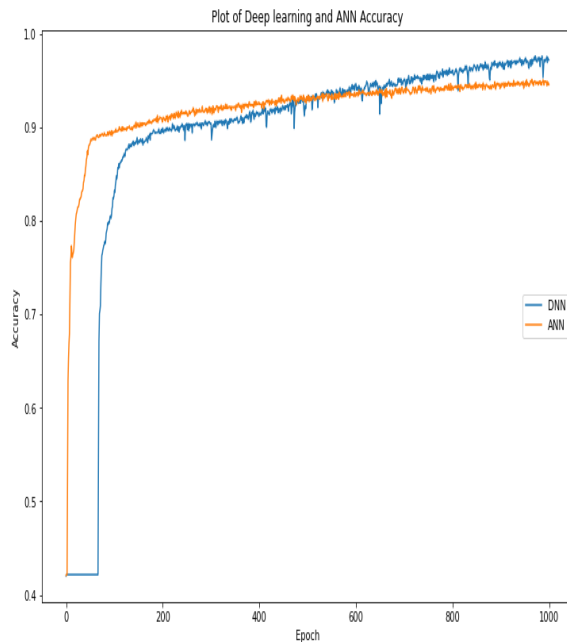

Figure 4: The accuracy of the Deep Neural Network and Artificial Neural Network Model
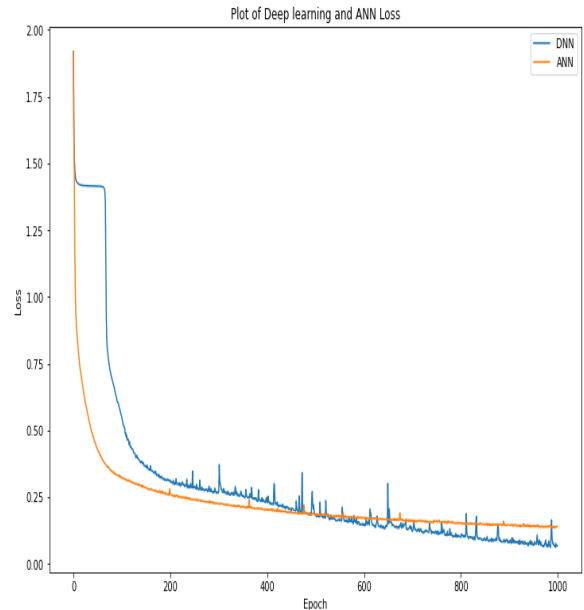

Figure 5: The loss of the Deep Neural Network and Artificial Neural Network Model

In figure 4 shows the accuracy of both Deep Neural Network and Artificial Neural Network Models which the accuracies tend to increase around 50 iterations with a little difference in the value until the final iteration, but the fluctuation in accuracies in-between iteration is minima for the Artificial Neural Network Model as compared to the Deep Neural Network Model.

In figure 5 shows the loss of both Deep Neural Network and Artificial Neural Network Models which the losses tend to step low around 100 iterations until the final iteration, but the fluctuation in losses in-between iteration is minima for the Artificial Neural Network Model as compared to the Deep Neural Network Model.

5. CONCLUSION

In this study, a deep learning model was developed to predict the conflicts' attack in Nigeria. Though the dataset was incomplete, noisy and class imbalanced in predicting the type of conflict attacks. After cleaning and normalizing the multi-class dataset, a quality dataset was obtained which shows a good prediction with convincing accuracy of 98% on the train set and 89% on the test data with losses of 2% and 11% respectively in the Deep Neural Network Model. This shows that the model is good for new predictions for conflict attacks.

Further, the need for a robust dataset is needed in developing a better generalized deep neural network model since the gap between training error and testing error tends to be more which may be due to an imbalanced dataset.

6. REFERENCES

[1] UNHCR., "The UN Refugee Agency: Nigeria Situation," UNHCR, 2017. [Online].

[2] C. Perry, " Machine Learning and Conflict Prediction: A Use Case. Stability:," International Journal of Security and Development, vol. 2, no. 3, p. 56, 2013.

[3] Marie K. S. and Salim B., , "Revisiting the Contested Role of Natural Resources in Violent Conflict Risk through Machine Learning.," 14 August 2020.

[4] I. Pradhan, "Exploratory Data Analysis and Crime Prediction In San Francisco.," 2018.

[5] T. E. V. o. P. (TEVP), 2018.

[6] Musumba, Mark; Fatema, Naureen; Kibriya, Shahriar, "Conflict in sub-Saharan Africa through the lens of supervised classification (prediction) models," 2020.

[7] Nils B. W., and Michael D. W., " Predicting Conflict in Space and Time," Journal of Conflict Resolution, vol. 54, no. 6, pp. 883-901, 2010.

[8] Kumar, V., Mazzara, M., Messina, A., and Lee, J., " A Conjoint Application of Data Mining Techniques for Analysis of Global Terrorist Attacks: Prevention and Prediction for Combating Terrorism.," 2020.