Research Review
DOI: https://doi.org/10.52868/RR/2022-3-3-4
RRJ 3 (3), 781-788 (2022)
Received 12 Oct 2022 | Revised 21 Oct 2022 | Accepted 07 Nov 2022 | Online Available 12 Nov 2022

Research Review
MEERP

OPEN ACCESS

ISSN (O) 2693-5007

ARTICLE

Transformer Model for Bomb Blast Prediction

Shweta Khera[1] | Moayed D. Daneshyari[1 *]

Abstract:
Global terrorism has increased in the last 15 years, with bomb blast incidents accounting for the majority of the attacks. Bomb explosion strikes accounted for more than 45 percent of all terrorist activity worldwide from 1970 to 2019. Several experts believe that socioeconomic conditions of a country influence the level of terrorism. This characteristic is not accounted in previous machine learning models designed to forecast terrorism. Decision Trees, Random Forest, Nave Bayes (NB), K-Nearest Neighbour (KNN), Tree Induction (C4.5), Iterative Dichotomiser (ID3), and Support Vector Machine (SVM) have all been used in the past to create terrorist predicting models. These forecasts are deterministic, so the models' outputs are best guesses, and no model can guarantee 100% accuracy due to unknown and unpredictable variables. The transformer model has been shown to perform effectively with considerably longer sequences and is capable of learning complex relationships between each piece of incoming data. This work offers a transformer-based time series forecasting technique that learns from socioeconomic data alongside bomb explosion incidents. We also use probabilistic forecasting to narrow down the options, which can prove to be valuable in averting future terrorist attacks.
Keywords: GTD, transformer models, probabilistic forecasting, random forest, deep learning

1 | INTRODUCTION

Terrorists have successfully launched the explosions all throughout the world, with long-reaching implications that go well beyond those who have been harmed. By far the most prevalent cause of terrorism-related deaths is explosives. Between 1970 and 2019, explosives were employed in 45.8% of the 209,451 verified terrorist attacks throughout the world, killing over 93,000 people. Several methods are used to evaluate prospective hazards based on historical data. There is sufficient evidence to support the hypothesis that a country's socioeconomic characteristics are linked to its level of terrorization. Figure 1 shows all Global Terrorism Data (GDT) plot from 1970 to 2019. To increase the accuracy of future bombing predictions, we intended to put this into practice by creating a forecasting model that is related to socioeconomic data.

Even though much research shows that there is a connection between a country's economic status and terrorism, the terrorism predicting models today do not take these factors into equation for forecasting the data. To forecast the GTD time

1 Department of Computer Science, California State University East Bay, Hayward, CA, USA

series data, a number of algorithms have been devised. The statis-tical and machine learning algorithms have been used to understand trends and patterns in this data. Auto-regression (AR), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) were the three prominent approaches. GTD data has also been modelled using deep learning algorithms based on convolutional and recurrent neural networks.

The introduction of Self-Attention based models is intended to increase deep learning models' capacity to make more accurate predictions. Transformer models, which were first proposed in the paper 'Attention is All You Need,' have quickly become the state-of-the-art in NLP. We created a new time series forecasting technique based on Transformer architecture in this work. Transformer-based models offer the capacity to simulate complicated time series data dynamics that are difficult to model using sequence models. In this paper, we show that, when applied to the problem of times series forecasting, a Transformer-based model can produce good results and beats several existing forecasting models, taking into account nations' monetary and developmental positions.

2 | LITERATURE REVIEW

Globalization and national socio-economic conditions have been widely linked to the predictability of terrorism activities. In 2000, Brynjar and Annika analysed and concluded that there are no systemic factors that clearly showing that there will be more terrorism in the future. However, another research three years later by Brock, Gregory, and Akila deduced that groups unsatisfied with the current economic status quo, yet unable to bring about drastic institutional changes, may find it rational to engage in terrorist activities. Later in 2011, findings by Andreas, Jens, Daniel, and Friedrich also implied that countries benefit from economic development and growth in terms of reduction in terrorism. A paper by Seung-Whan Choi in 2014 studied that even though economic growth is not a cure-all solution, healthy economic conditions are, without doubt, beneficial to the war on terrorism. Another study (Tench and Fry, 2016) shows that terrorist attacks often follow a general pattern that can be modelled and predicted using math. Therefore, socio-economic factors should be accounted for while making predictive models for future terrorism. Following are the studies that built predictive models:

1. Which countries will experience more terror attacks
   a. (Predicting Terrorism: A Machine Learning Approach
      i. Date range: 1970 – 2014
      ii. Algorithms: classical regression, Poisson regression, artificial. Neural network, regression tree, boot- strap aggregating, boosting, and random forest.
2. Type and location of the attack
   a. Machine Learning Techniques to Visualize and Predict Terrorist Attacks Worldwide using the Global Terrorism Database
      i. Date range: 1970 - 2018
      ii. Algorithms: Decision trees and Random Forest
   b. Suicide Bomb Attack Identification and Analytics through Data Mining Techniques
      i. Date range: 1995 - 2017
      ii. Algorithms: Naïve Bayes, ID3 and J48 algorithms, K-means algorithm, A priori algorithm
   c. Future Terrorist Attack Prediction using Machine Learning Techniques
      i. Algorithm: Ensemble Learning, random forest classification, and random-forest regression
3. Responsible perpetrators
   a. Using Global Terrorism Database (GTD) and Machine Learning Algorithms to Predict Terrorism and Threat
      i. Date range: 1970 - 2017
      ii. Algorithms: KNN algorithm and random forest algorithm

b. An Experimental Study of Classification Algorithms for Terrorism Prediction
    i. Date range: 1970 - 2013
    ii. Algorithms: Naïve Bayes (NB), K-Nearest Neighbour (KNN), Tree Induction (C4.5), Iterative Dichotomi-ser (ID3), and Support Vector Machine (SVM)
c. TGPM: Terrorist Group Prediction Model for Counter-Terrorism
    i. Date range: 1998 - 2008
    ii. Models: Crime Prediction Model [16, 17], Group Detection Model (GDM) [18] and Of-fender Group Detection Model (OGDM) [18, 19]
d. Terrorism Prediction Using Artificial Neural Network
    i. Date range: 1996 - 2017
    ii. Algorithms: Feedforward neural networks
4. Similar Terrorist Events
a. Detection of Similar Terrorist Events
    i. Technologies: NLP
b. Spatio-temporal patterns of IED usage by the Provisional Irish Republican Army
    i. Date range: 1970 – 1998
5. Casualty based on bombing type
a. 40 years of terrorist bombings – A meta-analysis of the casualty and injury profile
    i. Date range: 1970 – 2014
b. Mass Casualty Terrorist Bombings: A Comparison of Outcomes by Bombing Type
    i. Date range: 1966 – 2002
6. Identifying deceptive behaviour
a. A Neural Network for Counter-Terrorism
    i. Algorithm: feed-forward backpropa-gation network
b. Pattern classification in social network analysis: a case study
    i. Algorithm: Multivariate Bayesian classifiers
c. Big data-based prediction of terrorist attacks
    i. Date range: 1970 - 2014
    ii. Algorithm: KNN, C4.5, bagging, and SVM, Hybrid classifier

Although much research exists around terrorist activities forecast, only a few focus on explosive attacks and rarely consider external factors. Most of the predictive models' input data only con-sist of variables directly related to terrorism such as location, attack type, target, attack weapon. Along with forecasting methods, external socio-economic factors should factor into the equation as the most published findings show its considerable impact on terrorist activities.

3 | TRANSFORMER MODEL FOR TIME SERIES

In the paper "Attention Is All You Need," a new architecture called Transformer is introduced. Transformer is an architecture that uses two parts to transform one sequence into another: Encoder and Decoder (Figure 2). On the left is the encoder, and on the right is the decoder. Encoder and Decoder are both made up of modules that may be layered on top of one another. Although the encoders are structurally identical, they do not share weights.

Multi-Head Attention and Feed Forward layers make up the majority of the encoder. Be-cause strings cannot be utilized directly, the inputs and outputs are first embedded in an n-dimensional space. Since there are no recurrent networks that can recall how sequences are fed into a model, positions are added to each data in the sequence's embedded representation (n-dimensional vector). Only the bottom-most encoder is used for embedding. These embeddings pass via a self-attention layer first, which allows the encoder to look at other words in the input sequence while encoding a single information. The outputs of the self-attention layer are fed into a feed-forward neural network. The feed-forward neural network's output is passed on to the next encoder.

The output of the top encoder is then turned into a set of attention vectors. Each encoder has a residual link surrounding each sub-layer, which is then normalized. The decoder has self-attention, encoder-decoder attention, and feed forward layers. Each de-coder's "encoder-decoder attention" layer uses the attention vectors from the topmost encoder to assist the decoder in focusing on relevant points in the input sequence. In the following time step, the output of each step is passed to the bottom decoder, and the decoders, like the encoders, bubble up their decoding results. We embed and add positional encoding to those decoder inputs, just like we did with the encoder inputs, to indicate where each information is located. A float vector is produced by the decoder stack.

The Linear layer is a basic fully connected neural network that projects the vector produced by the stack of decoders into a much, much bigger vector known as a logits vector. Assume our machine has learnt 10,000 unique English words from its training dataset. This would make the logits vector 10,000 cells wide, with each cell representing the score of a distinct word. That is how we interpret the output of the model, which is followed by the linear layer. Following that, the softmax layer converts those scores into probabilities. The cell with the highest probability is chosen, and the time step's output is the information linked with it. "Non-sequential," "Self-Attention," and "Positional embeddings" are the primary qualities of a transformer. Because the input sequences are non-sequential, they are analysed as a whole rather than piece by piece in the sequence. Self-Attention is used to calculate similarity scores between sequence segments. The "Non-Sequential' property is the primary reason why trans-formers do not have lengthy dependency concerns. The original transformers process an input sequence as a whole, rather than relying on prior hidden states to capture relationships with earlier data. As a result, there is no chance of losing historical data. Furthermore, both multi-head attention and positional embeddings convey information about the connection between data. And all this happens in parallel (non-recurrent), which makes both training/inferences much faster.

4 | METHODOLOGY

4.1 Libraries

Here is the list of libraries adopted:

- Darts: It is a Python library for easy manipulation and forecasting of time series. Darts supports both univariate and multivariate time series and models. Some of the models offer a rich support for probabilistic forecasting.
- Numpy: NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
- Matplotlib: Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.
- Pandas: pandas is a Python package that provides fast, flexible, and expressive data structures designed to make working with "relational" or "labelled" data both easy and intuitive.
- Seaborn: Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
- PyTorch lightning: PyTorch Lightning is an open-source Python library that provides a high-level interface for PyTorch, a popular deep learning framework.

4.2 Data Collection

The terrorism dataset was collected from the global terrorism database. Additional Features:

- hdi2019: The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. (Sources: PopulationData.net, Wikipedia.org, hdr.undp.org)
- country_type
  – Threshold to consider country as 'developed': hdi2019 0.8
  – Threshold to consider country as 'developing': 0.5 hdi2019 < 0.8

– Threshold to consider country as 'underdeveloped': hdi2019 < 0.5

- Unemployment rate: The unemployment rate represents the number of unemployed people as a percentage of the labour force (Sources: OurWorldInData.org, Statsama.com, UN Data, Ma-crotrends.net)
- gdp: Gross domestic product is a monetary measure of the market value of all the final goods and services produced in a specific time period by countries (Source: OurWorldInDa-ta.org).

4.3 Analysing Quantity and Quality of Data

There was total 95395 samples and 135 features in global terrorism dataset. Table 1 shows analysis of the collected data. Table 1 (top section) shows that out of 81 categorical features, there were 24 binary features, meaning they had 0 or 1 as their values. Also, among 81 categorical features, there were 28 pairs of features which were in direct correlation. Table 1 (bottom section) shows that although from 135 features, 77 features did not have more than 50% of the data, 35 fields did not lack any data. 12 features' missing values ranged between 0.1% to 11% and 11 features were missing 21-36% data

4.4 Data Cleaning

Out of 30 textual features, following five were converted to categorical data:

1. gname (Perpetrator Group Name)
2. Corp1 (The name of the corporate entity or government agency that was targeted)
3. Target1 (This is the specific person, building, installation, etc., that was targeted and/or victimized and is a part of the entity named corp1)
4. provstate (This variable records the name (at the time of event) of the 1st order subnational administrative region in which the event occurs)
5. City (This field contains the name of the city, village, or town in which the incident occurred).

Table 2 shows how these five textual features were given different numerical values for different textual entries. The missing values in the dataset were set to '-999' to indicate that the value for that cell is unknown.

4.5 Feature Engineering

The target features 'date_num' and 'latlon' were built by feature engineering. This process is shown at Table 3. For date_num, all dates between 1970-2019 are indexed starting from 1. The dataset had few missing month and date values, so an index is also assigned for 'yyyy-00-00' and 'yyyy-mm-00'. For latlon, all combinations of latitude (-90 to 90 degree) and longitude (-180 to 180 degree) with 0.1 distance be-tween points are indexed starting from 1.

4.6 Scaler Transformation

MinMaxScaler with 'feature_range' set to (0, 1) to features used for predicting 'date_num' and (-1, 1) to features used for predicting 'latlon'. fit_transform () method is applied to the training data whose output is then used to transform the training and testing sets of data.

4.7 Feature Selection Results

Three different approaches were used to select features for predicting 'date_num' and 'latlon'. For 'date_num', feature selection of random forest is used to finalize the features whereas for 'latlon', the recursive feature elimination technique is used.

4.7.1 Feature Selection Score

The resulting feature selection scores are listed in Table 4.

4.7.2 Recursive Feature Elimination Output

- date_num (indexed dates for years 1970-2019)
- City (This field contains the name of the city, village, or town in which the incident occurred)
- target1 (This is the specific person, building, installation, etc., that was targeted and/or victimized and is a part of the entity named corp1

4.7.3 Pearson Correlation

Pearson Correlation Plot was implemented for both 'date_num' and 'latlon'. Table 5 shows the Pearson correlation data:

Figure 3 shows the Pearson Correlation Plot calculated with respect to “date_num". Similarly, Figure 4 is the Pearson Correlation Plot calculated with respect to “latlon".

4.8 Hyper parameter Tuning and Resulting MAPE

For transformer predicting “date_num"(date index), the hyperparameters and the resulting MAPE are shown in Table 6.

For transformer predicting “latlon"(latitude longitude index), the hyperparameters and the resulting MAPE are shown in Table 7.

4.9 Training Parameters and Results

The training parameters including number of features, epochs, input size, number of heads, number of encoder/decoders, feedforward network, the batch size, activation function, learning rate, dropout, output size, and quantiles for both transformer models are shown in Table 8.

The results of MAPE and RMSE for the transformer model for prediction of “date_num” is shown in Table 9 for different probability, while the results for transformer model for prediction of “latlon” is shown in Table 10. Furthermore,

Furthermore, the Date Index and latitude-longitude index, corresponding to the time and location of terrorist attack, have been plotted for actual date vs. predicted date (with various probability) in Figures 5 and 6 respectively.

5 | SUMMARY

This study offers a Transformer-based technique to forecast bomb attack events. While training the data, extra parameters such as unemployment, human development index, GDP, and so on in addition to the assault event data had been added. After completing the data collecting procedure, we examined the completeness of the data and converted a few key attributes from textual to categorical. Using the data’s year, month, and date properties, we were able to create a new "date_num" field to provide an index to that date. Similarly, we added an index to latitude and longitude variables that were concatenated. Next, we selected essential characteristics that would be utilized in training the model using the Random Forest’s feature selection and recursive feature elimination procedures. Following the completion of the appropriate scaling adjustments, the process of training the model and adjusting hyperparameters began. Our prediction error values for prediction probability=0.5 for "date_num" were 3.52% and 7.98% for "latlon" using the best model.

6| CONCLUSIONS AND FUTURE WORK

As terrorism has grown in the last several years, we need a strong approach to prevent these acts from happening. The forecasting models used to anticipate crimes only evaluate the variables surrounding the attack and do not take into account socioeconomic circumstances which clearly have a great impact as it can be seen in several research studies and also in our work. Furthermore, because models like LSTM and RNN can only access previously seen states of the model, these models forecast the future threat based on the prior occurrence. The transformer model takes different relationships between input sequences into view while predicting the data with high accuracy. Also, to improve the predictions, probabilistic forecasting gives us the power to narrow down the possibilities which can be very useful in preventing future terrorism at-tacks. The future work should focus on improving the model by hyper parameter tuning and predicting different features such as ‘city’, ’target’, ’country’, etc. Also, in this work we built two different models for predicting two different features. Future work can also focus on building a model which predicts multiple features in single model with high accuracy. Other types of terrorism can use similar approach to predict and prevent the attacks from occurring.

Appendix A: Github Code Links

1. Random Forest (Feature Selection and Recursive Feature Elimination): https://github.com/Shweta Khera/GTD_Transformer/blob/main/GTD _RandomForest.ipynb

2. AutoCorrelation Plot and Principle Component Analysis: https://github.com/ShwetaKhera/GTD_ Transformer/blob/main/gtd_autoCorrelatio n.ipynb
3. Transformer Model for predicting "date_num" (date index): https://github. Com/ShwetaKhera/GTD_ Transformer/blob/main/DateNum_Transfo

   rmer.ipynb

4. Transformer Model for predicting "latlon" (latitude-longitude index): https://github.com/Shweta

Khera/GTD_Transformer/blob/main/Latlon_tr ansformer.ipynb.

REFERENCES

1. Choi, S.W., 2015. Economic growth and terrorism: domestic, international, and suicide. Oxford Economic Papers, 67(1), pp.157-181.
2. Lia, B. and Hansen, A., 2000. Globalisation and the future of terrorism-patterns and predictions.
3. Bang, J., Basuchoudhary, A., David, J. and Mitra, A., 2018. Predicting terrorism: a machine learning approach.
4. Blomberg, S.B., Hess, G.D. and Weerapana, A., 2004. Economic conditions and terrorism. European Journal of Political Economy, 20(2), pp.463-478.
5. Huamaní, E.L., Alicia, A.M. and Roman-Gonzalez, A., 2020. Machine learning techniques to visualize and predict terrorist attacks worldwide using the global terrorism database. Machine Learning, 11(4).
6. Kalaiarasi, S., Mehta, A., Bordia, D. and Sanskar, 2019. Using global terrorism database (GTD) and machine learning algorithms to predict terrorism and threat. International Journal of Engineering and Advanced Technology, 9(1), pp.5995-6000.
7. Tolan, G.M. and Soliman, O.S., 2015. An experimental study of classification algorithms for terrorism prediction. International Journal of Knowledge Engineering, 1(2), pp.107-112.
8. Sachan, A. and Roy, D., 2012. TGPM: Terrorist group prediction model for counter terrorism. International Journal of Computer Applications, 44(10), pp.49-52.
9. Soliman, G.M. and Abou-El-Enien, T.H., 2019. Terrorism Prediction Using Artificial Neural Network. Rev. d'Intelligence Artif., 33(2), pp.81-87.
10. Ferooz, F., Hassan, M.T., Awan, M.J., Nobanee, H., Kamal, M., Yasin, A. and Zain, A.M., 2021. Suicide bomb attack identification and analytics through data mining techniques. Electronics, 10(19), p.2398.
11. Cozza, V. and Rubino, M., 2014. Detection of Similar Terrorist Events. In IIR (pp. 28-33).
12. Tench, S., Fry, H. and Gill, P., 2016. Spatio-temporal patterns of IED usage by the Provisional Irish Republican Army. European Journal of Applied Mathematics, 27(3), pp.377-402.
13. Edwards, D.S., Mcmenemy, L., Stapley, S.A., Patel, H.D.L. and Clasper, J.C., 2016. 40 years of terrorist bombings–a meta-analysis of the casualty and injury profile. Injury, 47(3), pp.646-652.
14. Arnold, J.L., Halpern, P., Tsai, M.C. and Smithline, H., 2004. Mass casualty terrorist bombings: a comparison of outcomes by bombing type. Annals of emergency medicine, 43(2), pp.263-273.
15. Freytag, A., Krüger, J.J., Meierrieks, D. and Schneider, F., 2011. The origins of terrorism: Cross-country estimates of socio-economic determinants of terrorism. European Journal of Political Economy, 27, pp.S5-S16.
16. Malathi, A. and Baboo, S.S., 2011. Evolving data mining algorithms on the prevailing crime trend–an intelligent crime prediction model. Int J Sci Eng Res, 2(6).

17. Ozgul, F., Erdem, Z. and Bowerman, C., 2009, April. Prediction of unsolved terrorist attacks using group detection algorithms. In Pacific-Asia Workshop on Intelligence and Security Informatics (pp. 25-30). Springer, Berlin, Heidelberg.
18. Ozgul, F., Bondy, J. and Aksoy, H., 2007, December. Mining for offender group detection and story of a police operation. In Proceedings of the sixth Australasian conference on Data mining and analytics-Volume 70 (pp. 189-193).
19. Ozgul, F., Erdem, Z. and Aksoy, H., 2008, June. Comparing two models for terrorist group detection: Gdm or ogdm? In International Conference on Intelligence and Security Informatics (pp. 149-160). Springer, Berlin, Heidelberg.
20. Dixon, S.J., Dixon, M.B., Elliott, J., Guest, E. and Mullier, D.J., 2011, December. A neural network for counter-terrorism. In International Conference on Innovative Techniques and Applications of Artificial Intelligence (pp. 229-234). Springer, London.
21. Coffman, T.R. and Marcus, S.E., 2004, March. Pattern classification in social network analysis: A case study. In 2004 IEEE Aerospace Conference Proceedings (IEEE Cat. No. 04TH8720) (Vol. 5, pp. 3162-3175). IEEE.
22. Saha, Snehanshu & Aladi, Harsha & Kurian, Abu & Basu, Aparna. (2017). Future Terrorist Attack Prediction using Machine Learning Techniques. 10.13140/RG.2.2.17157.96488.
23. [Meng, X., Nie, L. and Song, J., 2019. Big data-based prediction of terrorist attacks. Computers & Electrical Engineering, 77, pp.120-127.
24. Alammar, Jay.“The Illustrated Transformer.” Jalammar.github.io,jalammar.github.io/illustrated-transformer/.
25. Vaswani, Ashish, et al. ‘Attention Is All You Need’. ArXiv: 1706.03762 [Cs], Dec. 2017. arXiv.org, http://arxiv.org/ abs/1706.03762.
26. ‘Machine Learning Mastery’. Machine Learning Mastery, https://machinelearningmastery.com/.
27. Onnen, Heiko. ‘Transformer Unleashed: Deep Forecasting of Multivariate Time Series in Python’. Medium, 20 Feb. 2022, https://towardsdatascience.com/transformer-unleashed-deep-forecasting-of-multivariate-time-series-in- python-9ca729dac019.
28. START (National Consortium for the Study of Terrorism and Responses to Terrorism). (2021). Global Terrorism Database (GTD) [Data set]. University of Maryland. https://www.start.umd.edu/gtd
29. START (National Consortium for the Study of Terrorism and Responses to Terrorism). (2021, August). Global Terrorism Database codebook: Methodology, inclusion criteria, and variables. University of Maryland. https://www.start.umd.edu/gtd/downloads/Codebook.pdf
30. ‘PopulationData.net’. PopulationData.net, https://www.populationdata.net/
31. Human Development Reports | United Nations Development Programme. https://hdr.undp.org/
32. ‘Our World in Data’. Our World in Data, https://ourworldindata.org
33. Want Statistics? Ask Me Anything :) | Statsama. https://www.statsama.com/
34. UNdata. https://data.un.org/
35. Macrotrends | the Long Term Perspective on Markets. https://www.macrotrends.net.

How to cite this article: Shweta Khera and Moayed D. Daneshyari. Transformer Model for Bomb Blast Prediction Research Review.2022;781-788. https://doi.org/10.52868/RR /2022-3-3-4