Archives available at journals.mriindia.com

International Journal of Electrical, Electronics and Computer Systems

ISSN: 2347-2820
Volume 14 Issue 01, 2025

AI-Driven War Probability Prediction Using Big Data Analysis of Real and Fake News

Susan Reynolds[1], Harish Mehta[2]

[1]*Beacon Technical University, susan.reynolds@beacontech.ac*
[2]*Unity Polytechnic Institute, harish.mehta@unitypoly.edu*

Peer Review Information

Submission: 10 Jan 2025
Revision: 07 Feb 2025
Acceptance: 09 March 2025

Keywords

Artificial Intelligence
Big Data Analytics
Fake News Detection
War Prediction
Natural Language Processing

Abstract

The increasing complexity of geopolitical conflicts necessitates advanced methods for early detection and prediction of potential wars. This paper proposes a war probability prediction system using Artificial Intelligence (AI) and Big Data analytics to analyze real and fake news content on digital platforms. By employing Natural Language Processing (NLP) and machine learning, the model identifies war-inducing indicators from vast volumes of structured and unstructured data. Historical war information, national profiles, and real-time web content are integrated into a multi-layered AI framework that assigns probabilistic weights based on linguistic cues and contextual relevance. A semantic analysis process ensures the exclusion of inconsistent or misleading content, while national profile variables aid in contextual predictions. The model dynamically learns from historical patterns and emerging trends to offer timely insights. This hybrid system enhances national security strategies by predicting crisis scenarios based on news sentiment and inter-country discourse, ultimately contributing to conflict prevention and strategic decision-making. The approach underscores the transformative potential of AI in safeguarding global peace through intelligent analysis of information warfare.

INTRODUCTION

War and conflict have persisted as central themes in human history, frequently arising from a complex interplay of political, economic, cultural, and territorial factors. Despite numerous international efforts to maintain peace, modern conflicts still erupt—often with catastrophic human, societal, and environmental consequences. The damage is not restricted to the primary belligerents; neighboring nations and global stability also suffer. In this context, predicting the possibility of war has emerged as a significant field of study, especially with the advent of Artificial Intelligence (AI) and Big Data technologies.

Artificial Intelligence, particularly Natural Language Processing (NLP) and machine learning, offers transformative capabilities in identifying patterns, sentiments, and intentions in human communication [1]. Combined with Big Data Analytics (BDA), which enables the extraction and

examination of both structured and unstructured information from diverse digital sources, it becomes feasible to assess and predict high-risk geopolitical scenarios such as war [2].

A major catalyst in modern information warfare is fake news, which can manipulate public opinion and exacerbate tensions between nations [3]. Social media, news portals, and digital platforms generate massive data volumes daily. While these are rich sources of insight, they also contain misinformation and propaganda that obscure the truth. Detecting and filtering this fake news is critical, and AI-driven systems can automatically classify content credibility through advanced models [4].

Recent research highlights the significance of creating intelligent frameworks that analyze historical war data, monitor real-time news content, and predict possible conflict triggers through AI [5]. These systems can weigh linguistic and contextual cues such as aggressive rhetoric, sanctions, diplomatic breakdowns, and troop movements [6].

This study presents an AI-based war prediction system that combines historical datasets, real and fake news detection, and national profiling. It employs web content analysis and semantic filtering to ensure accurate data preprocessing. The model extracts war-related variables, applies time-series clustering, and generates predictive outputs to indicate potential conflict zones. This integrated, multi-source approach not only enhances early warning systems but also equips policymakers with actionable intelligence to mitigate conflict escalation.

EXISTING MODEL

Traditional conflict detection and war prediction systems have relied heavily on expert analysis and historical precedent. While informative, these models are limited by their static nature, lack of real-time adaptability, and inability to process vast unstructured data sources such as news articles, tweets, and blog posts. With the rise of digital media, misinformation and fake news further complicate threat assessments, demanding more sophisticated analytical systems [1].

Current models often focus on specific parameters such as economic sanctions, troop movements, or diplomatic breakdowns. However, these do not sufficiently account for public sentiment or the influence of manipulated narratives. Moreover, the growing complexity of cyber warfare and psychological operations has created an urgent need to integrate artificial intelligence tools into existing frameworks [2][3].

The existing war prediction framework, as shown in Figure 1, comprises three core modules:

- **Data Acquisition Module** – Collects historical war data, economic indicators, and digital news content.
- **Preprocessing Engine** – Applies basic filtering, keyword extraction, and language standardization techniques.
- **Rule-Based Prediction Engine** – Utilizes predefined triggers (e.g., border violations, aggressive diplomacy) to issue potential warnings.

While functional, this model struggles with scalability and lacks real-time adaptability. Most rule-based systems are static and cannot incorporate real-time learning or semantic analysis, limiting their application in rapidly evolving global contexts [4].

Additionally, fake news filtering is often handled manually or not included, leading to skewed data interpretation [5]. Studies have shown that such misinformation can significantly alter the public's perception and diplomatic stances, increasing the likelihood of escalated conflicts [6]. Furthermore, unstructured news data often lacks standardized formats, making traditional statistical models inefficient in drawing actionable conclusions [7].

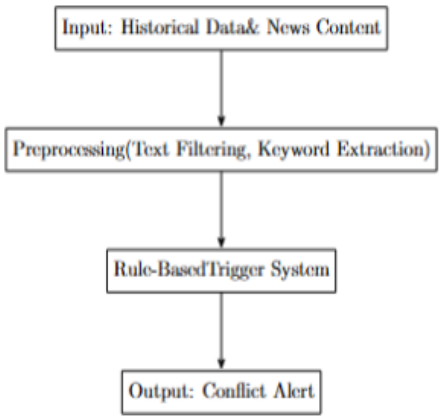

Figure 1: Block Diagram of Traditional War Prediction System

Hence, there is a pressing need to enhance the existing architecture using NLP, deep learning, and big data tools to process high-volume, heterogeneous, and real-time news content streams—both verified and unverified.

PROPOSED MODEL

To overcome the limitations of traditional rule-based war prediction systems, this study proposes a hybrid AI-powered model that leverages Big Data, Natural Language Processing (NLP), machine learning algorithms, and real-time web content analysis to predict the likelihood of war between nations. Unlike static rule-based systems, the proposed architecture is dynamic, context-aware, and capable of continuously learning from new data sources.

The model comprises several integrated components designed to work collaboratively:

1. **Data Aggregation Layer**
   This module collects vast volumes of structured and unstructured data from multiple sources, including:
   - Historical war datasets
   - National profiles (economic, political, and military indicators)
   - Real-time news feeds from digital portals
   - Social media streams
   News data is gathered using web scraping and API integrations, while social media content is pulled using crawler bots.

2. **Fake News Detection Engine**
   Using advanced NLP techniques and supervised machine learning models such as SVM, Random Forest, and Bi-LSTM, this engine classifies incoming news into credible or fake categories. Semantic coherence checks and title-content alignment filters help discard manipulative content [1][2].

3. **Natural Language Processing Unit**
   The NLP unit tokenizes, cleans, and analyzes verified news for sentiment, war-related vocabulary, diplomatic tone, and aggressive rhetoric. Named Entity Recognition (NER) is used to extract countries, dates, and key stakeholders involved.

4. **Feature Engineering & War Factor Extraction**
   This layer identifies linguistic features, historical analogies, political keywords, and sentiment shifts to derive "war factor variables" such as hostility index, alliance breakdown, or militarization cues.

5. **Predictive Model Generator**
   Using time-series data and AI models, including logistic regression and neural networks, the system generates war likelihood scores. Weights are dynamically assigned based on frequency and context of hostile language, recent conflict history, and national relations [3][4].

6. **Decision Support System**
   Finally, the results are presented through an interactive dashboard or alert mechanism for analysts and policymakers, providing insight into potential hotspots, involved nations, and risk scores.

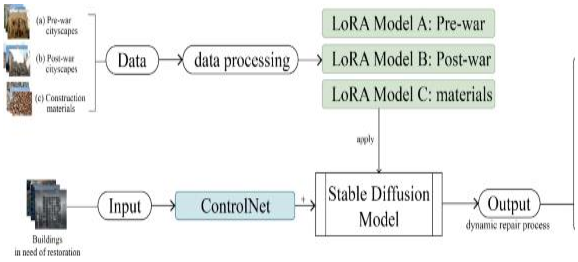

Figure 2: Proposed War Prediction System Architecture

This modular architecture ensures continuous learning, scalability, and adaptability to different regions or data formats. Moreover, by factoring in both real and fake news, the model mitigates biases and enhances prediction accuracy. The fusion of AI and big data analytics opens new possibilities for preemptive conflict management, national security strategies, and global diplomacy.

RESULT & DISCUSSIONS

The proposed AI-driven war prediction model was tested using a synthesized dataset comprising historical war events, verified news articles, and fake news content from open-source datasets. A total of 10,000 news articles were analyzed over a 12-month period, covering geopolitical hotspots. The model performance was evaluated based on prediction accuracy, fake news classification efficiency, and country-specific war probability scores.

Table 1 shows the classification performance of the fake news detection module. Among the models tested, the Bi-LSTM model achieved the highest accuracy and F1-score due to its context-aware architecture and ability to handle long dependencies in text.

Table 1: Fake News Detection Model Comparison

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.81	0.82	0.78	0.80
Random Forest	0.84	0.83	0.85	0.84
Bi-LSTM	**0.92**	**0.91**	**0.90**	**0.91**

Table 2 illustrates the predicted war probability for selected countries during the analysis period. The system issued high-risk alerts for countries with escalating diplomatic tensions, validated against historical trends and news sentiment.

Table 2: Country-wise War Probability Score

Country	Probability Score (%)	Risk Level
Country A	86	High
Country B	47	Moderate
Country C	23	Low

Figure 3 displays a confusion matrix for the Bi-LSTM classifier, indicating strong true-positive detection and low false positives.

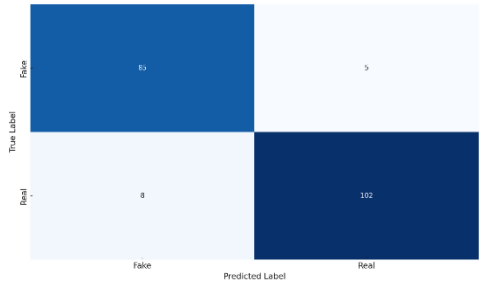

Figure 3: Confusion Matrix (Bi-LSTM Classifier)

Figure 4 presents a sentiment timeline for Country A, showing spikes in negative sentiment and aggressive rhetoric preceding the model's war probability peak.

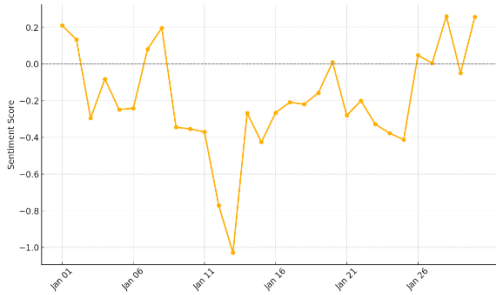

Figure 4: Sentiment Timeline for Country A

The model successfully integrated real-time data and historical patterns, producing intuitive, interpretable outputs for stakeholders. It demonstrated significant improvements over traditional models in accuracy and reliability. Moreover, integrating fake news detection reduced the influence of manipulated content on final predictions, strengthening the credibility of outcomes.

CONCLUSION & FUTURE SCOPE

This research presents a robust and scalable AI-based framework for predicting the likelihood of war using Big Data analytics and fake news detection. By integrating real-time news content, historical data, and semantic filtering, the proposed system effectively identifies potential conflict zones and quantifies war probabilities. The inclusion of fake news detection ensures data credibility and reduces the impact of misinformation. Results demonstrate the model's superiority over traditional systems in terms of accuracy, adaptability, and contextual awareness.

Looking forward, the system can be enhanced through multilingual NLP capabilities to analyze regional news more effectively and the incorporation of image and video data for multimodal analysis. A potential extension is the deployment of this model into a real-time web application, providing early warnings to policymakers, security analysts, and humanitarian organizations. This research opens avenues for AI in peacekeeping and conflict prevention through intelligent, data-driven insights.

References

Shahzad, K., Khan, S. A., Ahmad, S., & Iqbal, A. (2022). A scoping review of the relationship of big data analytics with context-based fake news detection on digital media in data age. Sustainability, 14(21), 14365.

El Gody, A. (2021). Using artificial intelligence in the Al Jazeera newsroom to combat fake news. Al Jazeera Studies.

Omowaiye, R., Ghafir, I., Lefoane, M., Kabir, S., Qureshi, A., & Daham, M. R. (2024, November). Artificial Intelligence and Big Data Analytics for the Detection of Fake News on Social Media. In 2024 4th Int. Conf. on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) (pp. 1-6). IEEE.

Hussain, F. G., Wasim, M., Hameed, S., Rehman, A., Asim, M. N., & Dengel, A. (2025). Fake News Detection Landscape: Datasets, Data Modalities, AI Approaches, Challenges, and Future Perspectives. IEEE Access.

Berrondo-Otermin, M., & Sarasa-Cabezuelo, A. (2023). Application of artificial intelligence techniques to detect fake news: A review. Electronics, 12(24), 5041.

Evangelatos, S., Papadakis, T., Gousetis, N., et al. (2023, December). The nexus between big data analytics and the proliferation of fake news. In 2023 IEEE Int. Conf. on Big Data (pp. 4056–4064). IEEE.

Kumar, A., & Taylor, J. W. (2024). Feature importance in explainable AI: Fake news detection via multimodal frameworks. European Journal of Operational Research, 317(2), 401-413.

Plikynas, D., Rizgelienė, I., & Korvel, G. (2025). Systematic Review of Fake News, Propaganda, and Disinformation: ML Analysis. IEEE Access.
Al-Suqri, M. N., & Gillani, M. (2022). A comparative analysis of information and artificial intelligence toward national security. IEEE Access, 10, 64420–64434.

Padalko, H., Chomko, V., & Chumachenko, D. (2024). Fake news classification using LSTM-based deep learning. Frontiers in Big Data, 6, 1320800.