Breaking the trend: Anomaly detection models for early warning of socio-political unrest

Luca Macis [a,*], Marco Tagliapietra [a], Rosa Meo [b], Paola Pisano [a]

[a] Department of Economics and Statistics Cognetti de Martiis, University of Turin, Italy
[b] Department of Computer Science, University of Turin, Italy

ARTICLE INFO

ABSTRACT

This paper presents an innovative Early Warning System for predicting conflicts and unrest based on Anomaly Detection, identifying sudden and unexpected changes in behavioral patterns that may indicate the potential for these events to occur. This approach draws inspiration from various fields – including industry, such as manufacturing, physics and networking – but its application in the domain of diplomacy is entirely new. The system, tested on three case studies, showcase its ability to enhance open-source intelligence technique in the diplomatic arena. The study provides a fresh perspective on predictive analytics and focuses on examining outbreaks.

1. Introduction

Conflict prevention is critical to international relations and diplomacy, as early intervention can often lead to conflicts being resolved before they are able to escalate into larger and more destructive events. Early warning systems play a crucial role in this context as they identify potential risks and threats and provide decision-makers with timely information to help define policies for mitigating or preventing any conflict.

Early warning systems are often based on data analysis and advanced computational techniques thanks to the large amounts of data gathered in information systems, shared via the Internet, and advancements in the fields of Machine Learning and Artificial Intelligence. Data can be extrapolated from various sources, including open-sources, satellite imagery, and social media. However, knowledge discovered from data is the final step in a much longer process that begins with identifying patterns and anomalies that may indicate a potential conflict. This allows decision-makers to address the underlying causes of conflict proactively and to take action to prevent its emergence.

Prior to these technological advancements taking place, potential conflict was traditionally identified largely by way of individual diplomatic and political knowledge, intuition, and subjective judgment. Although this approach had some success, intuitive approaches to conflict prediction have been criticized for their lack of rigor and strong subjectivity. Human analysts are often biased due to their own experiences, beliefs, and perspectives, leading to inconsistent and unreliable predictions (Gleditsch, 2002), thus entailing a more reactive than a proactive approach. Data and technology have the potential to revolutionize the field of conflict prevention, providing early warning signals of potential unrest. With the advent of big data from multiple sources and machine learning, vast amounts of data can now be analyzed in real time, identifying potential risk factors before they escalate into full-blown conflicts, and standardizing predictions in an objective and accountable way. This can improve the accuracy and timeliness of conflict prevention efforts and give decision-makers the technical support they require in order to act promptly and effectively at the first sign of crises.

Governments and international organizations now have come to acknowledge the importance of using technology and data analysis for conflict prediction in order to prevent or mitigate the imminent risk. This has been the case for several decades, as demonstrated by early studies on preventive diplomacy (George, 1999; Ackermann, 2003) and more recent works, such as the book *Violence, Wars, Peace, Security* (Wallensteen et al., 2018). The use of data and technology may potentially facilitate early peacekeeping interventions, which are a crucial factor in reducing conflict, as indicated in the research by Håvard Hegre, Lisa Hultman, and Håvard Mokleiv Nygård (Hegre et al., 2019b). Their study simulated a scenario in which the United Nations ceased its peacekeeping efforts after 2001. They found that, compared to the actual number of countries with ongoing peacekeeping efforts, three to

---

four more countries would have suffered major conflicts by 2013 in the absence of the UN's efforts.

In response to this, various organizations initiated the development of data analysis and data collection and interpretation systems, such as the Global Conflict Risk Index (Halkia et al., 2020a) by the Joint Research Center of the European Commission and the Violence and Impacts Early-Warning System (VIEWS) (Hegre et al., 2019a) developed by Uppsala University.

This paper outlines the development and outcomes of a new Early Warning System, using Machine Learning techniques. More specifically, it uses an Anomaly Detection approach for the purposes of preventing conflict and unrest. It was chosen to use this approach, which has never been used in this field, and which differs from previous supervised methods, as it is more apt to identify anomalies in a country. The promising results of the three case studies infer that this approach could be used for a wider range of conflicts. The Section 2 and the Section 4 of this paper provide a comprehensive overview of current conflict prevention models and illustrate how paper models could provide a valid alternative to those already in use.

2. Artwork

The conflict prediction field has traditionally used various Machine Learning models, including Logistic Regression (Halkia et al., 2020a), Dynamic Multinomial Logit (Hegre et al., 2012), Random Forest (Muchlinski et al., 2016), naive Bayes classifiers (Perry, 2013), and Neural Networks (Beck et al., 2000). All of these methods are supervised, meaning that they require labeled data, with domain experts providing the correct output for each input. This allows the algorithms to learn patterns and to make predictions based on past data. While supervised methods rely on labeled data to learn patterns and make predictions, unsupervised methods uncover underlying structures in unlabeled data, making them effective for anomaly detection where supervised methods might fall short.

This work, however, adopts an innovative approach based upon Anomaly Detection, which can be treated as an unsupervised method. Unsupervised methods do not require labeled data and instead identify patterns or anomalies within the data themselves. Anomaly Detection has achieved successful results in several domains, such as industry (Siegel, 2020) and physics (Nakazawa and Kulkarni, 2019), as well as with multi-variate time series (Zhang et al., 2021). It is an attractive method of conflict prediction as it does not require domain-expert intervention and it is well-suited to detecting rare events. This is discussed in the book *Understanding Deep Learning* (Ranjan, 2020), which argues that anomaly detection can identify unusual patterns in data that may indicate a rare occurrence, whereas supervised methods struggle to manage rare events due to their reliance on statistical metrics which are better suited to identifying frequent categories and the majority class. Ranjan's book focuses on using an Autoencoder model in Anomaly Detection for rare-event detection.

An Autoencoder is an unsupervised machine learning model which aims to identify patterns or anomalies in data. The first step is to divide the data into two groups: the majority class, which represents the normal state of the data, and the minority class, which represents the rare events. The Autoencoder then compresses and extracts the important information from the majority class, reconstructs it and creates a latent representation (hereinafter, Latent Space) of the majority class. During the last inference step, the Autoencoder compares the reconstructed data with the original data. Once the Latent Space is built, new data are given to the Autoencoder as input, with a view to reconstructing that input. If there is little difference between the new original data and the new reconstructed data, this means that the sample belongs to the majority class, or the normal state. However, if there is a large difference between the two, this suggests that the sample belongs to the minority class and is thus an anomaly.

In the paper entitled *A Variational Autoencoder Solution for Road Tra c*

Forecasting Systems (Boquet et al., 2020), the author shows that anomaly detection utilizing an autoencoder can be a valuable approach to predicting rare events, such as sudden shifts in road traffic patterns; it does so by learning a subspace that captures the underlying characteristics of the traffic data. Boquet proposes an autoencoder model that can extract valuable, compressed information from complex and multidimensional road traffic data sets and assign missing values online, without supervision. This extraction and compression technique improves the accuracy of forecasting systems. The author also highlights that the Latent Space, i.e. the compressed information learned by the model, is suitable for traffic anomaly detection. Similarly, events such as armed conflicts, coups, and unrest can be considered anomalies in a country's history and can be predicted as rare events. In the context of this academic evidence, this paper uses an unsupervised Anomaly Detection Autoencoder model for a conflict prevention Early Warning System. The Sections 3 and 7 below will discuss the databases and the results of this approach.

3. Dataset implementation

Dataset selection is a critical factor in a heavily data-driven model. The datasets adopted for conflict prediction primarily fall into two categories: social and diplomatic.

Many forecasting works have relied on data obtained from social media platforms, particularly Twitter, due to its ease of access. In recent years, numerous examples of conflict prediction using social media have emerged, such as the work by Korkmaz (Korkmaz et al., 2015). That article focuses on the use of multiple data sources, including Twitter, to predict and analyze civil unrest in six Latin American countries, highlighting the importance of using this platform as a dataset in predicting civil unrest. Another example of conflict prediction using Twitter is demonstrated in Ryan Compton's paper *Using Publicly Visible Social Media to Build Detailed Forecasts of Civil Unrest* (Compton et al., 2014). Compton presents a data mining system that generates forecasts of civil unrest incidents in Latin America using public posts on Twitter and Tumblr. The system identifies informational posts through filters applied to tens of millions of posts per day and generates predictions by annotating filtered posts with demographic, spatial, and temporal information.

However, as noted in the paper *Towards Understanding the Use of Telegram by Political Groups in Brazil* (Junior et al., 2021), Twitter is no longer used as a platform for organizing insurrections due to the prohibition imposed on violent threat tweets and on the tracking and persecution of Twitter users by authoritarian regimes. Consequently, a new social messaging platform, Telegram, has taken its place. This study specifically examines the use of Telegram by political communities in Brazil, with the results revealing a significant increase in political mobilization on this platform in recent years. Telegram's large-group structure has proven to be more effective in spreading messages compared to other preceding social media channels.

Telegram, as mentioned in the paper *Telegram: Data Collection, Opportunities and Challenges* (Khaund et al., 2021), remains an unexplored platform for data mining purposes, with only a proposed framework for conducting data mining and no notable applications in the field of conflict prediction.

For these reasons, we discarded social datasets and focused on diplomatic datasets. However, in creating our dataset, we decided to use time-series datasets that were updated regularly and contained daily observations, leading us to consider the Armed Conflict Location and Event Data Project (**ACLED**) (Raleigh et al., 2023) and Global Database of Events, Language, and Tone (**GDELT**) datasets and to discard other diplomatic, economic, and sociopolitical sources, such as the World Bank – which is adopted in several conflict prevention projects, including the Global Conflict Risk Index (Halkia et al., 2020b) – due to its non-daily variables.

The use of both datasets has been demonstrated to be valuable in conflict prediction research, as highlighted by other works (Matina

et al., 2019), as this provides a more comprehensive and diverse set of variables for the model: ACLED offers a more exhaustive view of events such as conflicts, protests, and riots, while GDELT focuses more on Natural Language Processing and sentiment analysis. This combination of perspectives improves the performance of the analysis models compared to the use of either dataset alone.

In addition, this study benefits from both human-coded and machine-coded diplomatic databases in order to overcome the limitations of relying on a single dataset. The human-coded database, ACLED, offers the advantage of accurate information, but with a weekly update frequency. Conversely, the machine-coded database, GDELT, provides a more frequent update cycle – every 15 min – including news from various media sources in over 100 languages, dating back to 1979 (Leetaru and Schrodt, 2013). The merits of merging and combining these two datasets were explored in the paper *Comparison Metrics for Large Scale Political Event Data Sets* (Schrodt and Analytics, 2015).

Data wrangling had to be performed in order to obtain clear and functional variables from the two datasets. The ACLED data are available for download on its platform (ACLED, 2022b) in the form of .csv files, which can be filtered by date range and event type. The types of events recorded in the dataset include battles, explosions and remote violence, violence against civilians, riots, protests, and strategic developments.

One limitation of the ACLED dataset is that its coverage period varies by country and region. For example, while most African regions have data from 1997, some European and Asian regions have data from 2018. An exhaustive country and time coverage list is available on the ACLED website (ACLED, 2022a).

To wrangle the dataset, we initially downloaded separate ones for each event type and then created new copies for each dataset, grouping them by *ISO 3166-1 alpha-3* code and date. *ISO 3166-1 alpha-3* codes are three-letter codes used to identify countries in the ISO standard. This produced a dataset in which each row corresponded to a day in a specific country, with the number of events of a specific type that occurred on that day in that country.

The GDELT dataset, on the other hand, contains many articles, as transpires from its Power Law distribution of the number of mentions per article (as shown in Fig. 1). However, many of these articles are considered noise, lacking significant information relevant to our analysis. To overcome this, the dataset was reduced, retaining only articles that received at least 10 mentions within the first 15 min of publication. This reduction strategy resulted in a dataset size of just 1.4 % of the original. The threshold of 10 mentions was determined using the elbow method, as depicted in Fig. 1. This reduction strategy may lead to the exclusion of important information from countries where freedom of the press is suppressed and international media attention is limited; however, this approach was deemed acceptable as there are relatively few countries of this nature. The Dynamic GCRI – a conflict risk model developed by the European Commission's Joint Research Centre which uses event data to predict conflict – adopted the same approach (Matina et al., 2019) based on the idea that, when an event actually happens, it will be reproduced by more than one media source and in more than one article. Moreover, the Dynamic GCRI highlights that the inclusion of all available information within the GDELT database would lead to greater bias than the exclusion of some information.

Although the results obtained were promising, for the purpose of global implementation, a careful assessment of the freedom of press in each country is recommended by monitoring relevant indices and tools. To this end, we suggest using the *Media Freedom Analyzer* proposed by Laura Schneider (Schneider, 2019) or other established indices such as the *Freedom of the Press Index* by Freedom House and the *Press Freedom Index* by Reporters Without Borders, analyzed by Ewa Sapiezynska and Claudia Lagos (Sapiezynska and Lagos, 2016).

The GDELT dataset includes 62 attributes, which are grouped into the following four categories:

- **Date attributes**, used to identify the globally unique identifier number and the date of the event;
- **Actor attributes**, used to identify the countries actively or passively involved in the event, using a unique 3-letter *ISO 3166-1 alpha-3* code. We use this attribute to identify and associate the observation with a specific country. In the case of the absence of this attribute, the "Event geography" is applied.
- **Event geography**, used to identify geographical information in cases where the actor attributes are null.
- **Event action attributes**, which include the *Average Tone* – a sentiment value obtained through sentiment analysis – the *Number of Mentions* of the event received in the last 15 min, and the *QuadClass* taxonomy. The latter classifies the event type into four primary categories: verbal cooperation, material cooperation, verbal conflict,
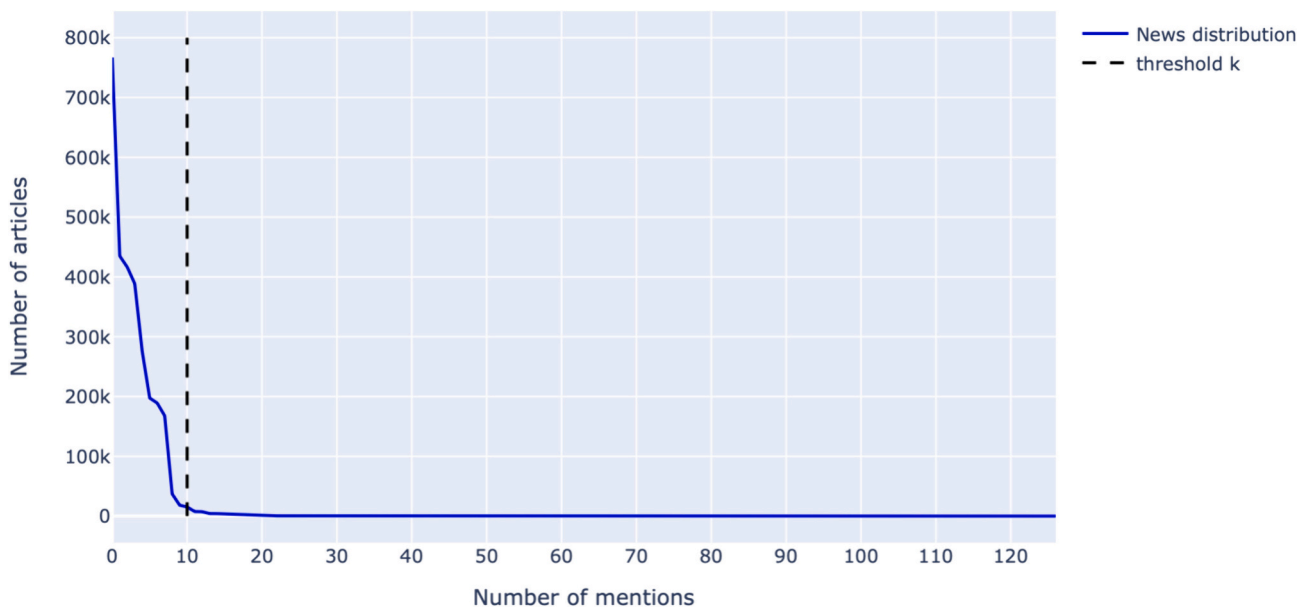

Fig. 1. GDELT mentions distribution from July 1, 2022 to August 1, 2022. The P(*number of mentions* > k) for a single news article follows a Power Law distribution. In this paper, $k = 10$ and P(*number of mentions* > 10) f 1.4 %.

and material conflict. Finally, there is the *Goldstein Scale*, a numeric score between −10 and 10 which captures the event's theoretical potential impact on a country's stability.

Detailed information about the attributes can be found in their documentation (GDELT, 2022).

It is important to note that the GDELT dataset is obtained by semantically analyzing each news item using Natural Language Processing. This analysis identifies the event location, the actors involved, and the other relevant elements embedded in the news. Actors are associated with *ISO 3166-1 alpha-3* codes. News articles lacking information on actors or event locations were removed.

The resulting dataset was then grouped by country and date, as was done for the ACLED dataset. Only the *ISO 3166-1 alpha-3* code associated with the actors actively or passively involved was recorded for each news item, excluding other actors. The country was recorded based on the event location in cases where the news item did not report actors. The dataset included additional attributes, such as the number of news items, the *GoldsteinScale*, and the *Average Tone* associated with each newsgroup belonging to the same *QuadClass*, for each date and country.

In addition to the variables extracted and grouped daily from the two databases, a new variable called **Violence** was created, inspired by the definition of violence given by ACLED (ACLED, 2022c):

Violence rates are calculated using the number of battle events, explosion or remote violence events, violence against civilians, and riot events, as well as the excessive force against protesters, sub-event type of the protest event type, excluding peaceful protest events and protest-with-intervention events.

The *Violence* variable was used as the target variable for the final case study to predict the onset of violence in Sri Lanka on 9 May 2022. In the final case study, the system attempted to predict the sudden increase in unrest events considered in the construction of the variable itself.

4. Methodology

In order to implement Anomaly Detection – a key method for recognizing unusual occurrences within data – we utilized an Autoencoder constructed utilizing an Encoder-Decoder architecture.

The Encoder-Decoder architecture is based on the idea of compressing data into a lower-dimensional representation, known as Latent Space, and then reconstructing the original data from this compressed representation. The Encoder component processes the input data matrix *X* and compresses it into a lower-dimensional representation. The Decoder component is responsible for reconstructing the original data from the compressed representation.

This compressed representation *H* aims to capture the most salient features of the original data and acts as input for the Decoder component, which is responsible for reconstructing the original data (Eq. (1)).

$$f_{enc}(X) = H \tag{1}$$

The decoder component uses the compressed representation, *H*, as input and reconstructs a new matrix, $\widetilde{X}$, which aims to be as similar as possible to the original matrix *X* (Eq. (2)).

$$f_{dec}(H) = \widetilde{X} \tag{2}$$

The squared error *SE* (Eq. (3)) is used as a similarity measure to calculate the difference between the original data and the reconstructed data for each day. Specifically, it computes the dissimilarity between the original data x_i, and the reconstructed data $\widetilde{x}_i$, for the i-th day, where *i* is the index for the day. This provides a scalar value representing the dissimilarity between the original data and the reconstructed data for each day. By minimizing the *SE*, the decoder aims to reconstruct the original data as accurately as possible for each day, while the encoder's goal is to identify the most compact data representation for each day
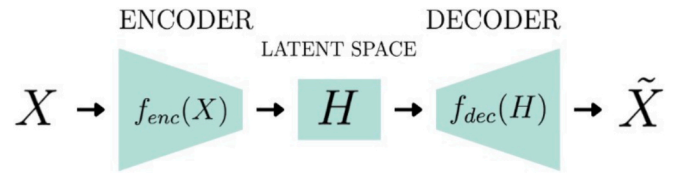

Fig. 2. Illustration of the process of transforming input data matrix *X* into reconstructed matrix $\widetilde{X}$ using an Encoder-Decoder architecture. The Encoder component compresses the input data matrix *X* into a lower-dimensional representation *H* through the encoder function f_{enc}. The compressed representation *H*, depicted in the latent space, is the output of $f_{enc}(X)$. Subsequently, the Decoder component utilizes the compressed representation *H* as input to the decoder function f_{dec} to generate the reconstructed matrix $\widetilde{X}$, aiming to closely resemble the original matrix *X*.

(Fig. 2).

$$SE\left(x_i, \widetilde{x}_i\right) = (x_i - \widetilde{x}_i)^2 \tag{3}$$

Both models were trained with data from "normal" days (i.e. days far away from a conflict) to establish a baseline, or in other words, a Latent Space for what is considered *normal* behavior. During validation and testing, days that deviated significantly from this normal behavior were labeled as anomalous. These anomalies are indicative of potential "pre-unrest" days, which are days leading up to an outbreak of civil unrest or conflict.

To determine the "normal" and "pre-unrest", the data was first filtered to remove actual "unrest" days, namely the days with the highest number of violent occurences and unrest reported in news sources. The remaining data were then analyzed to identify patterns and to classify the days into "normal" and "pre-unrest" categories. The purpose of this process was not to identify "unrest" days, being a trivial task, but rather to distinguish between "normal" days and "pre-unrest" days, which may be characterized by increased tension, signs of social unrest, or other indicators of potential violence.

Two parameters were used to make the distinction between "normal" days and "pre-unrest" days through statistical analysis: *Target Variable Threshold* and *Pre-Unrest Days*. Unlike previous studies which used a fixed threshold to define the state of war and unrest (Halkia et al., 2020a) and a set number of days to predict unrest in the short term, such as, 30 or 90 days (Rustad et al., 2011; Hegre et al., 2019a, 2019b), this study used an optimization system which selects the threshold of "unrest" and the number of "pre-unrest" days; this is unique to each country and has a predefined range.

The *Target Variable Threshold* refers to the value which, if surpassed, defines an event as being in a state of "unrest". The *Pre-Unrest Days* are the *n* days prior to the "unrest" day. These parameters were determined through a Grid Search optimization technique. The Grid Search was used to distinguish effectively between the "normal" days and the "pre-unrest" days by identifying the optimal values for the *Pre-Unrest Days* and *Target Variable Threshold* from a set of possible values. To define an "unrest" day, the value of the target variable – i.e., the variable examined to assess the sociopolitical unrest in a given country – had to be higher than at least than 98 % of the values of the same target variable on all the other days in the country dataset in question, as it had to be a rare event.

The possible values for each case study, which will be analyzed in detail in Section 6, are listed in Table 1.

The selection of pre-unrest days was established within a range of one to three months with the aim of assessing the short-term efficacy of the approach. The lower limit was specifically set at 30 days, representing the minimum prediction threshold reported in prior literature, as mentioned above. For each combination of possible values (i.e., *Thresh-old - Pre-Unrest Days*), the model's performance in separating the "pre-unrest" days from the "normal" days was evaluated during

Table 1
The *Date Range* column includes the start date and end date. The start date is the first day of coverage by ACLED for each country in question; the end date is the day of unrest considered for each case study. (*) 98 % of days in the dataset have a value below the lower threshold; (**) is the highest value of the target variable in the dataset.

Case study	Date range	Target variable	Threshold	Pre-unrest days
Ukraine	Jan. 01, 2018 Feb. 24, 2022	Battle fatalities	5*–52**	30–90
Burkina Faso	Jan. 01, 1997 Jan. 24, 2022	Battle fatalities	6*–147**	30–90
Sri Lanka	Jan. 01, 2010 May 9, 2022	Violence	2*–18**	30–90

validation. The combinations that performed best during validation are listed in Table 2.

The values reported in Table 2 were those ultimately used during testing to evaluate the performance for each case study. A more comprehensive examination of the determination of the parameters and the framework of the data is provided in the Appendix.

The Early Warning System implemented uses two different models based on the Encoder-Decoder architecture: an Encoder-Decoder Principal Component Analysis (PCA) model and a Long Short-Term Memory (LSTM) Autoencoder. The Encoder-Decoder PCA model was selected for its computational efficiency, as it can optimize two key parameters – *Target Variable Threshold* and *Pre-Unrest Days* – quicker than the LSTM Autoencoder. This is due to the fact that the Encoder-Decoder PCA model uses only linear combinations when identifying the Latent Space. Using the computer provided for this research, the Encoder-Decoder PCA model takes approximately 0.01 s to compile, compared to the 2 min taken by the LSTM Autoencoder. However, the LSTM Autoencoder can use non-linear functions, leading to a more accurate representation of the Latent Space and ultimately a better performance. Accordingly, the Encoder-Decoder PCA model was used to identify rapidly the optimal parameters, while the LSTM Autoencoder used those parameters to achieve a better performance.

5. Performance measure

Two plots were used to evaluate the performance of the proposed Early Warning System's performance: the reconstruction error plot and the receiver operating characteristic (ROC) plot. The reconstruction error plot displays the chronological progression of test set days from left to right, up until the day before the unrest event, as shown in Fig. 3.

Each day is represented as a dot, with the height of the dot indicating the squared error (SE) value. The higher the dot, the higher the SE value and, therefore, the more likely the day is considered a "pre-unrest" day. Orange circles represent "normal" days, while coral crosses represent "pre-unrest" days. A horizontal black line shows the threshold for separating "normal" days from "pre-unrest" days.

A ROC plot was also used to assess further the model's ability to distinguish between the two types of days. The ROC plot displays the Area Under the Curve (AUC), which is used as a precision measure to evaluate the model. AUC values close to 1 (100 %) indicate that the model has distinguished all days perfectly, while an AUC value of 0.5 (50 %) indicates that the model cannot distinguish between the two

Table 2
(*) 98.1 % of days in the Ukraine dataset have battle fatalities lower than 6; (⊛) 99.8 % of days in the Burkina Faso dataset have battle fatalities lower than 52; (⊖) 99.8 % of days in the Sri Lanka dataset have violence lower than 15.

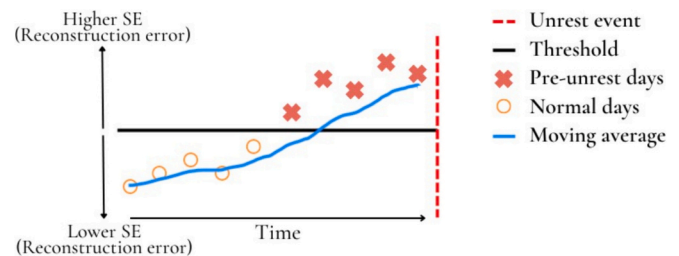| Case study	Target variable	Threshold	Pre-unrest days
Ukraine	Battle fatalities	6 (98.1 %)*	56
Burkina Faso	Battle fatalities	52 (99.8 %)⊛	63
Sri Lanka	Violence	15 (99.8 %)⊖	47

Fig. 3. Reconstruction error plot of the *ideal model*, where all "normal" days, i.e. days far from the unrest event – are represented by orange circles, and have a lower squared error than the "pre-unrest" days, i.e. the days near the unrest event, represented by coral crosses. A clear separation by threshold is possible: the horizontal black line represents the threshold for separating "normal" days from "pre-unrest" days. A moving average – the blue line – had to be included due to the potential for fluctuations in the squared error for a single day in real-world scenarios. This is achieved by computing the squared error over the most recent 15 days, instead of solely relying on the score of a single day. As a result, the moving average provides a more robust indication of potential unrest when it surpasses the threshold, as it considers a larger squared error sample size. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

types of days. The ROC plot examines the relationship between the True Positive Rate (TPR, fraction of true positives) and False Positive Rate (FPR, fraction of false positives) or the relationship between true and false alarms. The closer the AUC in the ROC plot is to 1, the more effective the model is in separating true and false alarms. Therefore, the AUC can be interpreted as the probability of a day being a true alarm (a "pre-unrest" day) if the threshold is crossed, or the probability of being a "normal" day if the threshold is not crossed. This evaluates more accurately the model's ability to distinguish between "normal" days and "pre-unrest" days and, thus, its effectiveness in predicting and preventing potential conflicts.

6. Case studies

In order to evaluate the effectiveness of the models, recent case studies concerning socio-political unrest events were analyzed. For each case, the ACLED database was examined to identify any variables that underwent a dramatic change in trend leading up to the event. The case studies chosen for this analysis were carefully selected to represent a range of potential scenarios in which the Early Warning System may be used.

6.1. Russian invasion of Ukraine

The first case study analyzes a critical event in the continuing Russia-Ukraine conflict that escalated significantly on February 24, 2022. This date marked a dramatic increase in battle-related fatalities, originating from military tensions that began in February 2014. The main objective of this case study is to explore whether this sharp increase in fatalities could have been anticipated earlier through the early warning system.

In the 56 days before the invasion, a total of 1600 events were recorded in Ukraine leading up to the Russian invasion. Despite this, fatalities remained relatively low during this period, with only 31 deaths registered prior to February 24, in stark contrast to the 103 fatalities recorded on the day of the invasion itself. 92 % of the reported events were categorized as Explosions/Remote violence and Battles, indicating prevalent but initially low-lethality conflict activities. Additionally, the geographical distribution of these events was strikingly concentrated, with >92 % occurring in two eastern regions near the Russian border–Donetsk and Luhansk (Fig. 4).
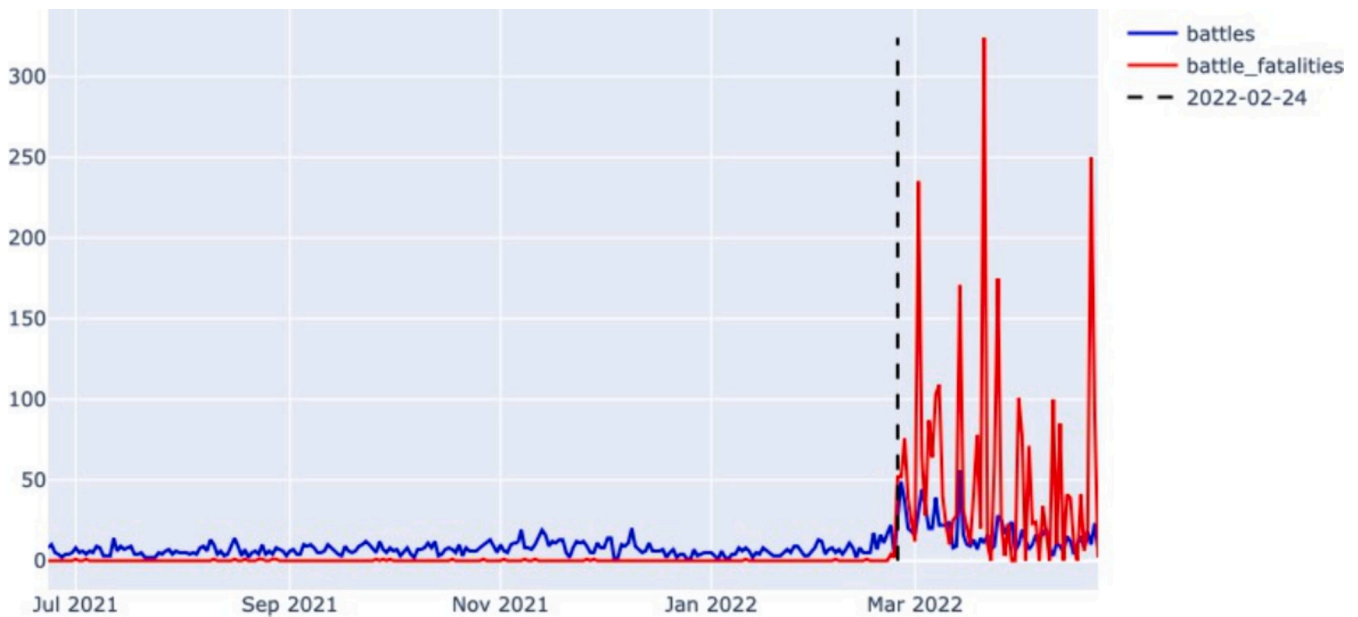
Fig. 4. Evolution of battle fatalities (red) and battles (blue) in Ukraine from Jun. 24, 2021, to Apr. 24, 2022. The black dashed line corresponds to the Russian invasion of Ukraine on Feb. 24, 2022. Source: Armed Conflict Location & Event Data Project (ACLED); www.acleddata.com. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6.2. Burkina Faso coup

The second case study concerns the overthrowing of the democratically-elected government headed by Roch Marc Christian Kaboré on behalf of a fringe army, which occurred between January 23 and 24, 2022. The study primarily explores whether the spike in battle-related fatalities observed just days prior to January 24 could have been anticipated by an early warning system, despite unclear trend changes.

In the 58 days preceding the critical increase in fatalities on January 19, a total of 437 events were reported, leading to 615 fatalities. However, the trend in fatalities did not show a noticeable increase until the five days before the coup. During this period, 41 events occurred with

204 fatalities—an alarming rise in deaths. Most of these events were battles and strategic developments, but there was also significant violence against civilians, riots, and other violent activities. Moreover, from November 22, 2021, to January 24, 2022, the militants involved in the coup engaged in significant property destruction, including sabotage of telecommunications systems, burning buildings and government infrastructures, vandalism, and destruction of schools. These acts of sabotage and destruction underscored the intense escalation leading up to the coup (Fig. 5).
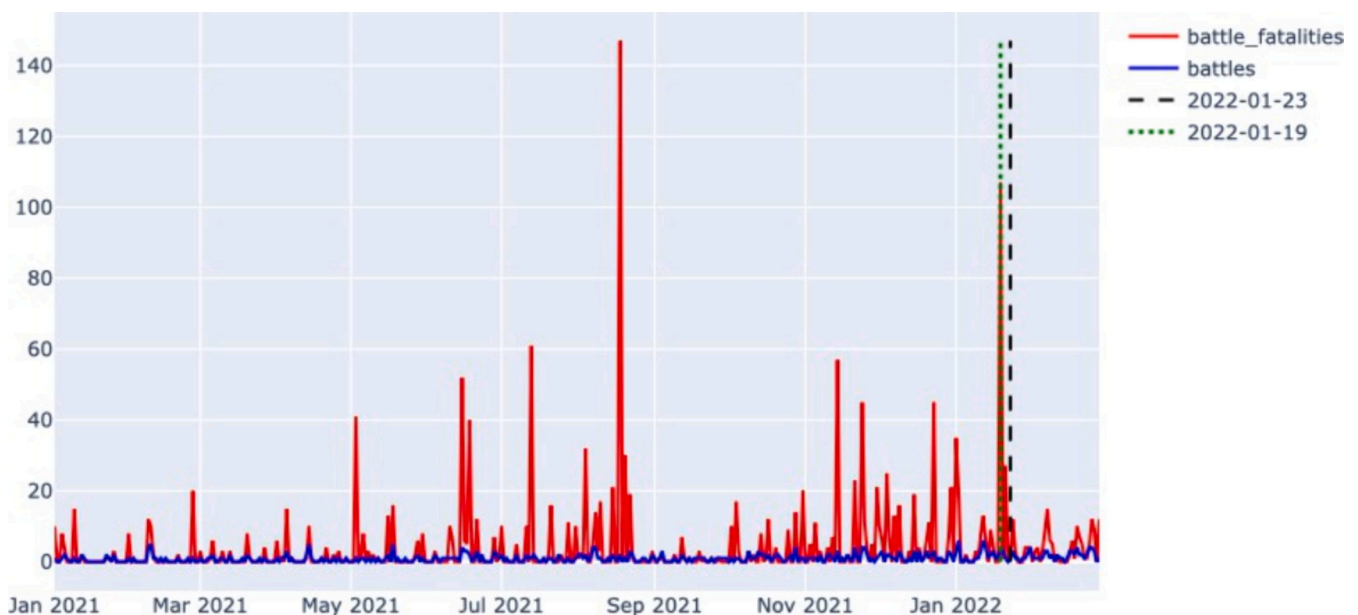

Fig. 5. Evolution of battles (blue) and battle fatalities (red) in Burkina Faso from Jan. 01, 2021, to Feb. 28, 2022. The green dotted line corresponds to the peak of battle fatalities on Jan. 19, 2022; the black dashed line corresponds to the coup on Jan. 23, 2022. Source: Armed Conflict Location & Event Data Project (ACLED); www.acleddata.com. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6.3. Sri Lanka's protests

The last case study focuses on the *violence* variable, derived from the ACLED dataset's disorder factors. Beginning with peaceful protests in late March 2022, the situation quickly escalated after declaring a state of emergency on April 1, witnessing a surge in mass protests that progressively became more violent. Notably, by May 11, the intensity of the unrest compelled the government to deploy the army with shoot-on-sight orders. This study notably investigates if the significant rise in violence from May 9, 2022, was foreseeable. Prior to this peak, several incidents hinted at the escalating tension, starting with inter-group clashes and worsening with responses to the economic downturn, fuel shortages, and the demanding of President Gotabaya Rajapaksa's resignation, culminating in nationwide strikes on April 28. The increasing frequency and intensity of these events painted a clear picture of the unrest leading up to the mentioned surge in violence (Fig. 6).

7. Results

The aim of the first case study was to detect the 56 "pre-unrest" days prior the Russian invasion of Ukraine in 2022 using an LSTM Autoencoder. As already mentioned in the Section 4, the specific parameters used in the study are listed in Table 3.

As noted, the optimal values for the parameters *Battle Fatalities Threshold* and *Pre-Unrest Days* were determined through a Grid Search optimization technique. The case study evaluated the model's ability to detect the 56 "pre-unrest" days preceding the Russian invasion of Ukraine on February 24, 2022. The model was considered successful if it could classify a day as "pre-unrest" even if the number of battle fatalities on that specific day was <6; the requirement was that this day was one of the 56 days leading up to an "unrest" day with >6 battle fatalities.

The analysis of the reconstruction error plot (as shown in Fig. 7) reveals that the days preceding the Russian invasion of Ukraine (February 24, 2022) exhibit a higher reconstruction error than the days deemed *normal* (represented by orange circles in the plot). The threshold for separating "normal" days from "pre-unrest" days, determined during the validation phase in the framework mentioned above, is represented by the black line in the plot. The blue line represents a moving average concerning the reconstruction error of the last 15 days. This moving average was added as an indicator of the trend in the reconstruction

Table 3
[ti]98.1 % of days in the entire dataset have battle fatalities <6 battle fatalities.

Battle fatalities threshold	Pre-unrest days
6 (98.1 %)[ti]	56

error. The day the moving average surpasses the threshold indicates a higher probability of a sudden increase in battle fatalities. The results of this analysis demonstrate the ability of the LSTM Autoencoder model to discriminate "normal" days from "pre-unrest" days effectively. This is further supported by the ROC plot, which shows an AUC very close to 1, at 0.9377.

In the second case study, we aimed to investigate whether it was possible to predict the high number of battle fatalities prior to **the Burkina Faso coup** that occurred on January 23, 2022. The optimal values for the parameters *Battle Fatalities Threshold* and *Pre-unrest days* were once again determined through a Grid Search optimization technique in this case study. The case study evaluated the model's ability to detect the 63 "pre-unrest" days preceding the coup on January 23, 2022. The model was considered successful if it could classify a day as "pre-unrest" even if the number of battle fatalities on that specific day was <52, as long as it was one of the 63 days leading up to an "unrest" day with >52 battle fatalities (Table 4).

As shown in Fig. 8, the LSTM Autoencoder model effectively discriminates the "pre-unrest" days from the "normal" days, as highlighted by the reconstruction error plot. The 15-day moving average of the reconstruction error (the technique adopted in Fig. 7) surpasses the threshold in April 2021 and remains above the threshold throughout the period in question. This indicates a higher probability of approaching a situation of socio-political instability. Additionally, the model's performance can be seen from the ROC plot, with an AUC of 0.8723. It is important to note that while the trend of increasing battle fatalities in Burkina Faso may seem apparent, this case study tested the model's ability to predict such trends rather than sudden breaks in behavior patterns, as seen in the previous case study. Nevertheless, the results can be regarded as excellent given the high AUC value close to the maximum value of 1.

In the final case study, we aimed to predict **the onset of violence in Sri Lanka** that occurred on May 9, 2022, two days before the
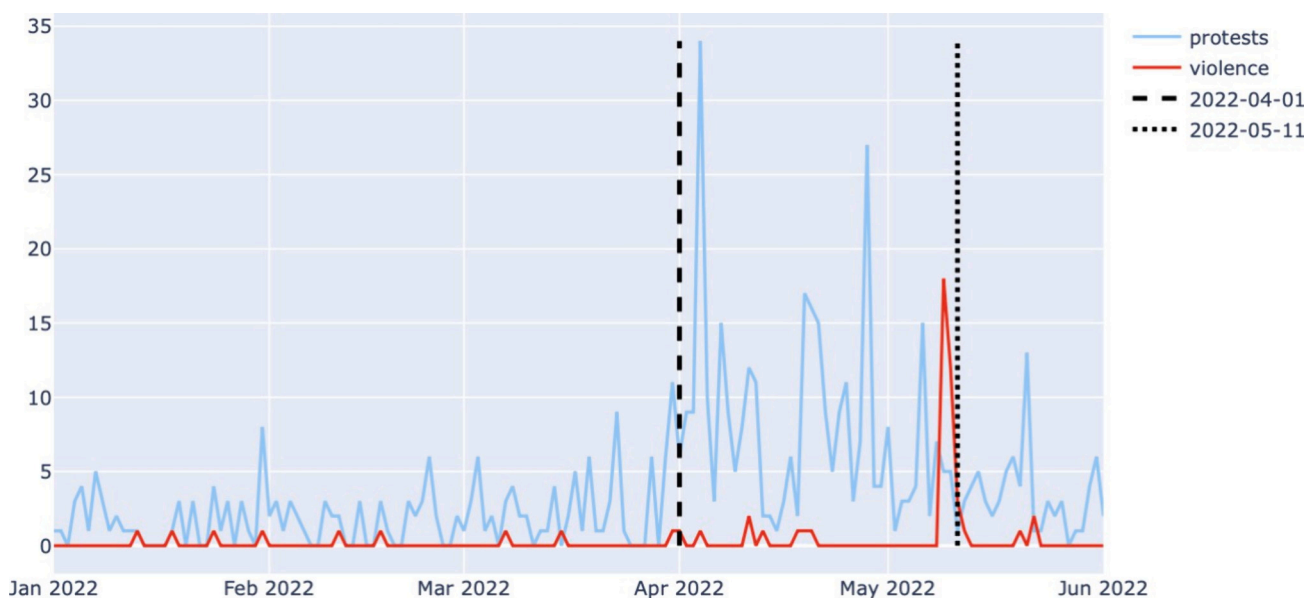

Fig. 6. Evolution of violence (red) and protests (light blue) in Sri Lanka from Jan. 01, 2022, to Jun. 01, 2022. The black dashed line corresponds to the declaration of a state of emergency; the black dotted line corresponds to the deployment of the army by the government. Source: Armed Conflict Location & Event Data Project (ACLED); www.acleddata.com. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
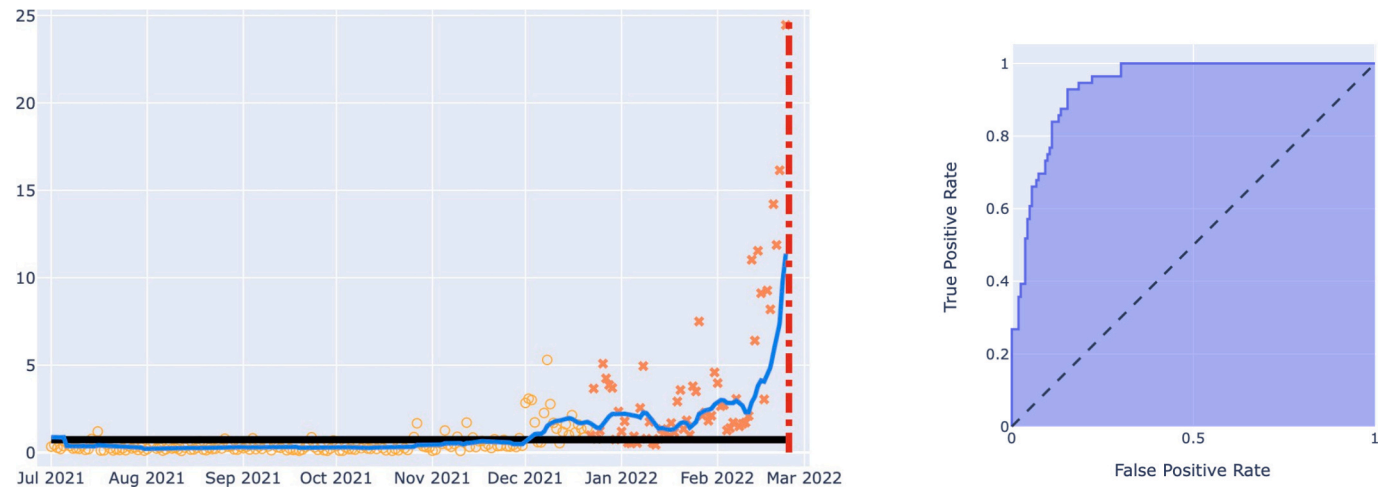
Fig. 7. Reconstruction error and ROC plot, Ukraine. AUC: 0.9377. The circles are "normal" days; the crosses are "pre-unrest" days; the black line is the threshold; the blue line is the moving average concerning the reconstruction error of the last 15 days; the vertical dotted and dashed line represents the date of the Russian invasion of Ukraine. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
[ti]99.8 % of days in the entire dataset have battle fatalities <52.

Battle fatalities threshold	Pre-unrest days
52 (99.8 %)[ti]	63

government's announcement that it was deploying the army with orders to shoot on sight to bring the violence under control. Once again in this final case study, the optimal values for the parameters *Violence Threshold* and *Pre-Unrest Days* were determined through a Grid Search optimization technique. The case study evaluated the model's ability to detect the 47 "pre-unrest" days preceding the onset of violence on May 9, 2022. The model was considered successful if it could classify a day as "pre-unrest" even if the amount of violence on that specific day was <15, as long as it was one of the 47 days leading up to an "unrest" day with >15 violent events (Table 5).

As can be observed from the plots in Fig. 9, the model performed exceptionally well. The reconstruction error plot clearly distinguishes between "normal" and "pre-unrest" days, and the ROC plot demonstrates an AUC of 0.9191. Additionally, this case study is evidence of the model's ability effectively to identify a breaking trend using the variable *violence*, which was derived from other variables in the ACLED dataset.

8. Conclusion

This paper presents an original method for predicting sudden increases in battle fatalities and violence using machine learning models. It should be noted that most studies in this field tend to adopt supervised approaches to conflict prediction, which require labeled data to learn patterns and make predictions. However, it turns out that this approach can be limited in detecting anomalous patterns, especially in contexts where anomalies are rare or ill-defined. We have therefore introduced the unsupervised method of anomaly detection for the first time in the study of conflict early warning. This allows us to explore the underlying structures in the data by analyzing the concept of a conflict anomaly, an unlabelled data, providing an additional tool to identify and address anomalies that may not be captured by traditional supervised approaches.

The anomaly detection approach uses historical data to identify a threshold for the target variable and then applies this threshold to

Table 5
[ti]99.8 % of days in the entire dataset have violence <15.

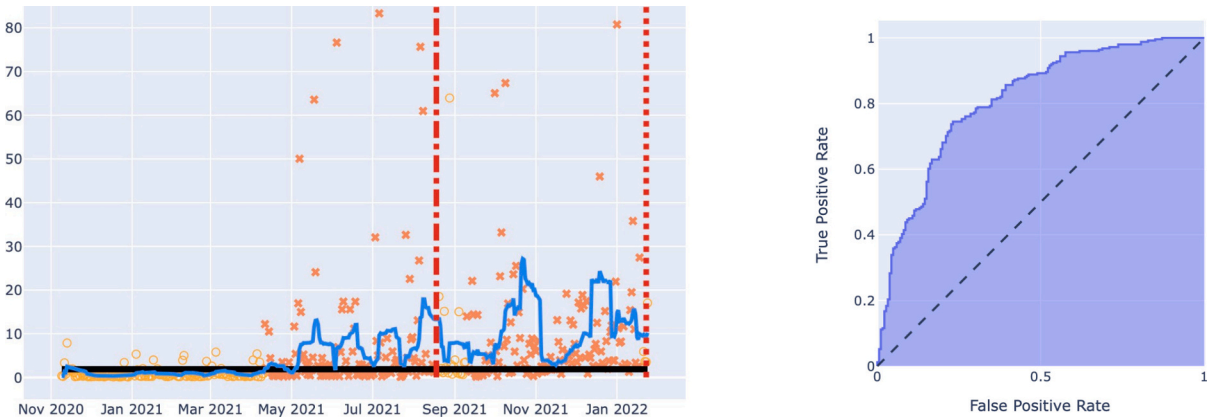| Violence threshold	Pre-unrest days
15 (99.8 %)[ti]	47

Fig. 8. Reconstruction error and ROC plot, Burkina Faso. AUC: 0.8723. The red dashed and dotted line corresponds to the highest peak in battle fatalities ever recorded (147), which occurred on Aug. 18, 2021; the red dotted line corresponds to the high number of battle fatalities that happened on Jan. 19, 2022 (107), 4 days before the coup. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
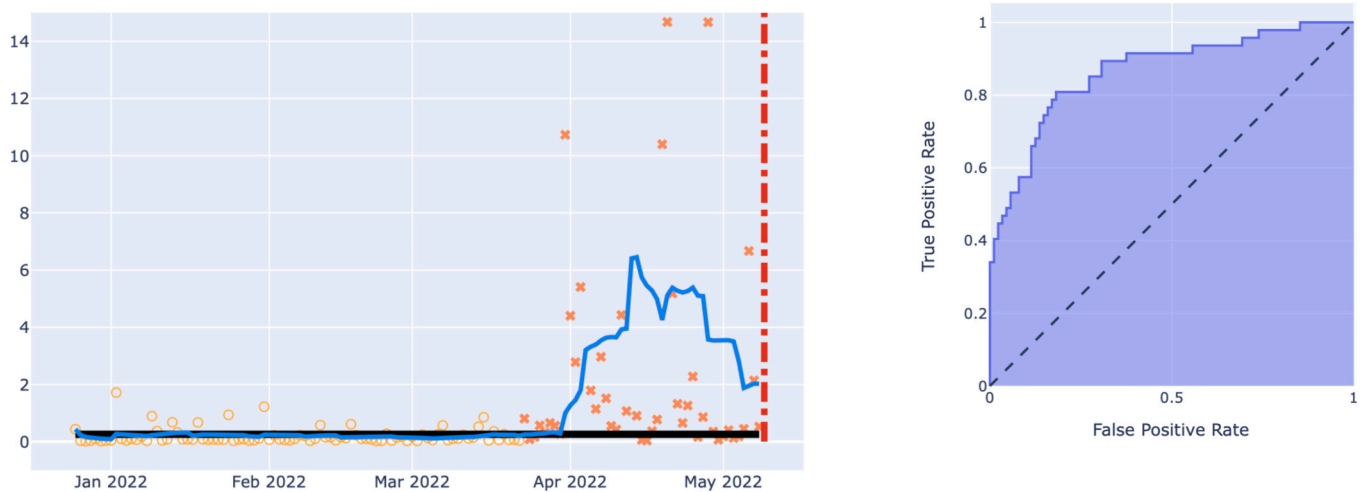
Fig. 9. Reconstruction error and ROC plot, Sri Lanka violence. AUC: 0.8656.

predict the likelihood of an imminent dramatic increase. The results demonstrate that the models developed are highly accurate, with AUC scores ranging from 87.2 % to 93.7 % for predicting battle fatalities and 86.6 % for predicting violence.

The key advantage of this approach is that it prioritizes the identification of sudden outbreaks, rather than predicting the exact number of occurrences. This is crucial for governments and international organizations, as it allows for early intervention and the implementation of preventative measures to mitigate the impact of an unrest event. This method should be used to complement, rather than replace, diplomatic efforts, by providing a quantifiable perspective on potential sociopolitical instability.

There are several avenues for future research that could further improve the accuracy of these models. For instance, creating new independent variables to consider the diplomatic relationships between neighboring countries, testing new variables derived from existing datasets, such as the *violence* variable, the integration with data that is more strictly economic, climatic or otherwise not geopolitical from a human perspective, or incorporating ensemble models that use multiple models instead of just the best-performing one.

Another possible future development in addition to extending it to more countries is to train the model by merging countries in such a way as to predict anomalies occurring in a similar way in other countries, the basic idea behind this possible development being that a crisis occurring for example in Nigeria could be similar to a crisis about to occur in Kenya.

Moreover, the outputs generated by these models can be used as independent variables in forecasting models aimed at predicting the exact number of occurrences. The squared error score assigned to each day can be used as an indicator of potential outbreaks, with a higher score identifying a higher likelihood of occurrence. Incorporating this unsupervised approach as input into classical supervised machine learning models has the potential to enhance the accuracy in predicting the exact number of occurrences (Boquet et al., 2020). In conclusion, the predictive diplomacy method presented in this paper has the potential to assist governments and international organizations in anticipating and preventing sociopolitical instability. Additionally, this method can also be applied in fields other than peacekeeping, such as in the procurement of rare earths, raw materials, and energy sources. For example, if a supplier country is identified as being susceptible to an outbreak, the importing country may be able to pursue the diversification of its sources proactively.

The development and implementation of these models should receive political attention and ought to be integrated into diplomatic efforts to ensure the safety and stability of nations and their citizens worldwide.

It is important to highlight that the conclusions drawn from this study are subject to some limitations. Firstly, we cannot assume that the characteristics of anomalous 60-day periods will repeat exactly at 60-day intervals in subsequent crises, nor can we guarantee that anomalous days will inevitably lead to a crisis. Further studies are needed to better understand the nature of the signal preceding a crisis and to more fully validate the proposed method. Additionally, it is important to note that the presence of an anomalous day does not necessarily guarantee the proximity of a crisis, as days, weeks, or even months may pass before a critical event occurs. Only through rigorous and comprehensive research can we improve our understanding and prediction capabilities of critical events. It is necessary to emphasize that our methodology may not be applicable in all situations or geopolitical contexts, and that customizations and adaptations may be required to make it effective in certain scenarios. Ultimately, while the preliminary results are very promising, it is essential to proceed with caution and continue to develop and enhance these models to ensure their reliability and utility in preventing international crises.

It is also for these reasons that a rigorous evaluation is planned in the coming months to assess this method's efficiency through live forecasting, rather than mere case studies. The ability to support diplomatic action in anticipating and preventing conflicts through data-driven decision-making has the potential to impact greatly the global security, stability, and resilience of all countries.

CRediT authorship contribution statement

Luca Macis: Data curation, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Marco Tagliapietra:** Data curation, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Rosa Meo:** Data curation, Formal analysis, Methodology, Supervision, Writing – original draft. **Paola Pisano:** Investigation, Supervision, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.techfore.2024.123495.

References

Ackermann, Alice, 2003. The idea and practice of conflict prevention. J. Peace Res. 40 (3), 339–347. https://doi.org/10.1177/0022343303040003006.

ACLED, 2022a. ACLED Coverage to Date. Edited by acleddata.com. (Online; accessed 3-October-2022).

ACLED, 2022b. Data Export Tool. Edited by acleddata.com. (Online; accessed 3-October-2022).

ACLED, 2022c. ACLED Trendfinder. Edited by acleddata.com (Online; accessed 10-December-2022).

Beck, Nathaniel, King, Gary, Zeng, Langche, 2000. Improving quantitative studies of international conflict: a conjecture. Am. Polit. Sci. Rev. 94 (March), 21. Accessed December 10, 2019. https://doi.org/10.2307/2586378.

Boquet, Guillem, Morell, Antoni, Serrano, Javier, Vicario, Jose Lopez, 2020. A variational autoencoder solution for road traffic forecasting systems: missing data imputation, dimension reduction, model selection and anomaly detection. Transp. Res. Part C Emerg. Technol. 115, 102622 issn: 0968-090X. https://doi.org/10.1016/j.trc.2020.102622. https://www.sciencedirect.com/science/article/pii/S0968090X19309611.

Compton, Ryan, Lee, Craig, Xu, Jiejun, Luis, Artieda Moncada, Lu, Tsai-Ching, 2014. Using Publicly Visible Social Media to Build Detailed Forecasts of Civil Unrest. https://doi.org/10.1186/s13388-014-0004-6.

GDELT, 2022. Documentation. Edited by gdeltproject.org. (Online; accessed 24-September-2022).

George, Alexander L., 1999. Strategies for preventive diplomacy and conflict resolution: scholarship for policy-making. Coop. Confl. 34 (1), 9–19. https://doi.org/10.1177/00108369921961753.

Gleditsch, Kristian Skrede, 2002. Expanded trade and GDP data. J. Confl. Resolut. 46 (5), 712–724. https://EconPapers.repec.org/RePEc:sae:jocore:v:46:y:2002:i:5:p:712-724.

Halkia, Matina, Ferri, Stefano, Schellens, Marie K., Papazoglou, Michail, Thomakos, Dimitrios, 2020a. The Global Conflict Risk Index: a quantitative tool for policy support on conflict prevention. Prog. Disaster Sci. 6 (April), 100069 https://doi.org/10.1016/j.pdisas.2020.100069.

Halkia, Matina, Ferri, Stefano, Schellens, Marie K., Papazoglou, Michail, Thomakos, Dimitrios, 2020b. The Global Conflict Risk Index: a quantitative tool for policy support on conflict prevention. Prog. Disaster Sci. 6, 100069 issn: 2590-0617. https://doi.org/10.1016/j.pdisas.2020.100069. https://www.sciencedirect.com/science/article/pii/S2590061720300065.

Hegre, Håvard, Karlsen, Joakim, Nygård, Håvard Mokleiv, Strand, Håvard, Urdal, Henrik, 2012. Predicting armed conflict, 2010–2050. Int. Stud. Q. 57 (December), 250–270. https://doi.org/10.1111/isqu.12007.

Hegre, Håvard, Allansson, Marie, Basedau, Matthias, Colaresi, Michael, Croicu, Mihai, Fjelde, Hanne, Hoyles, Frederick, Hultman, Lisa, Högbladh, Stina, Jansen, Remco, et al., 2019a. ViEWS: a political violence early-warning system. J. Peace Res. 56 (2), 155–174.

Hegre, Håvard, Hultman, Lisa, Nygård, Håvard Mokleiv, 2019b. Evaluating the conflict-reducing effect of UN peacekeeping operations. J. Polit. 81 (1), 215–232. https://doi.org/10.1086/700203.

Junior, Manoel, Melo, Philipe, da Silva, Ana Paula Couto, Benevenuto, Fabricio, Almeida, Jussara, 2021. Towards understanding the use of telegram by political groups in Brazil. In: Proceedings of the Brazilian Symposium on Multimedia and the Web, 237–244. WebMedia '21. Association for Computing Machinery, Belo Horizonte, Minas Gerais, Brazil. https://doi.org/10.1145/3470482.3479640 isbn: 9781450386098.

Khaund, Tuja, Hussain, Muhammad Nihal, Shaik, Mainuddin, Agarwal, Nitin, 2021. Telegram: data collection, opportunities and challenges. In: Information Management and Big Data, pp. 513–526. Accessed February 8, 2023. https://doi.org/10.1007/978-3-030-76228-5_37.

Korkmaz, Gizem, Cadena, Jose, Kuhlman, Chris J., Marathe, Achla, Vullikanti, Anil, Ramakrishnan, Naren, 2015. Combining heterogeneous data sources for civil unrest forecasting. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015. ASONAM '15. Association

for Computing Machinery, Paris, France, pp. 258–265. https://doi.org/10.1145/2808797.2808847 isbn: 9781450338547.

Leetaru, Kalev, Schrodt, Philip A., 2013. Gdelt: global data on events, location, and tone, 1979–2012. In: ISA Annual Convention, vol. 2. Citeseer, pp. 1–49, 4.

Matina, Halkia, Ferri, Stefano, Papazoglou, Michail, Van Damme, Marie-Sophie, Jenkinson, Gabriel, Baumann, Kathrin-Manuela, Thomakos, Dimitrios, 2019. Dynamic Global Conflict Risk Index. Report. Joint Research Centre (JRC).

Muchlinski, David, Siroky, David, He, Jingrui, Kocher, Matthew, 2016. Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. Polit. Anal. 24, 87–103. https://doi.org/10.1093/pan/mpv024.

Nakazawa, Takeshi, Kulkarni, Deepak V., 2019. Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder–decoder neural network architectures in semiconductor manufacturing. IEEE Trans. Semicond. Manuf. 32 (2), 250–256. https://doi.org/10.1109/TSM.2019.2897690.

Perry, Chris, 2013. Machine learning and conflict prediction: a use case. Stability: Int. J. Secur. Dev. 2 (October), 56. https://doi.org/10.5334/sta.cr.

Raleigh, Kishi, Linke, 2023. Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. Humanit. Soc. Sci. Commun. 10, 74. https://doi.org/10.1057/s41599-023-01559-4.

Ranjan, Chitta, 2020. Understanding deep learning: application in rare event prediction, 1st. URL: www.understandingdeeplearning.com, Connaissance Publishing. 10.13140/RG.2.2.34297.49765.

Rustad, Siri Camilla, Aas, Halvard Buhaug, Falch, Åshild, Gates, Scott, 2011. All conflict is local: modeling sub-national variation in civil conflict risk. Confl. Manag. Peace Sci. 28 (1), 15–40. https://doi.org/10.1177/0738894210 388122.

Sapiezynska, Ewa, Lagos, Claudia, 2016. Media freedom indexes in democracies: a critical perspective through the cases of Poland and Chile. Int. J. Commun. 10 (January), 22. https://ijoc.org/index.php/ijoc/article/view/3679/1551. (Accessed 30 January 2023).

Schneider, Laura, 2019. Measuring Global Media Freedom: The Media Freedom Analyzer as a New Assessment Tool. Book, pp. 1–252. https://doi.org/10.1007/978-3-658-28095-6.

Schrodt, Philip A., Analytics, Parus, 2015. Comparison Metrics for Large Scale Political Event Data Sets.

Siegel, Barry, 2020. Industrial anomaly detection: a comparison of unsupervised neural network architectures. IEEE Sens. Lett. 4 (8), 1–4. https://doi.org/10.1109/LSENS.2020.3007880.

Wallensteen, Peter, Wieviorka, Michel, Abraham, Itty, Aggestam, Karin, Bellamy, Alexander, Cederman, Lars-Erik, Ferret, Jerôme, et al., 2018. Violence, Wars, Peace, Security, vol. 2. Book chapter, pp. 411–456. https://doi.org/10.1017/9781108399647.003.

Zhang, Ang, Zhao, Xiaoyong, Wang, Lei, 2021. CNN and LSTM based encoder-decoder for anomaly detection in multivariate time series. In: 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), vol. 5, pp. 571–575. https://doi.org/10.1109/ITNEC52019.2021.9587207.

Luca Macis is a Data Scientist, a PhD Candidate in Economics and Statistics, and a Master's in Stochastic and Data Science from the University of Turin. He collaborated with the Italian Ministry of Foreign Affairs on an early warning system for conflict prevention. He has spoken at numerous conferences, including the European Conference on Machine Learning, the Jan Tinbergen European Peace Science Conference, and the ViEWS prediction challenge. Luca has been a visiting professor at the University of Turin and has worked on using eXplainable Artificial Intelligence (XAI) to predict deaths from armed conflicts.

Marco Tagliapietra is a Data Scientist with experience in research and advanced analysis of data. He has a Master's degree in Stochastics and Data Science, and a Bachelor's degree in Theoretical Mathematics. He is currently a PhD student in Economics and Statistics where he is developing a system for early warning and prevention of social and intercultural conflicts in collaboration with the Ministry of Foreign Affairs. He has spoken at numerous conferences, including the European Conference on Machine Learning, the Jan Tinbergen European Peace Science Conference and the ViEWS prediction challenge. Marco has also been a visiting professor at several universities and Master's courses.

Rosa Meo took a Master's Degree in Electronic Engineering and a Ph.D. in Computer Science and Systems Engineering from Politecnico di Torino, Italy. She is an Associate Professor at the Department of Computer Science at the University of Torino. She is active in the scientific research fields of Databases, Data Mining and Machine Learning (ML). She has applied ML methods to juridical problems provided by the Italian Authority for AntiCorruption in Public Administrations. She also worked on techniques for Privacy-Preserving Data Mining. She is the delegate from Italy at OECD in the working group of Experts on AI (ONE AI).

Paola Pisano is a professor of Economics and Business Management at the University of Turin and has published more than 100 articles. A visitng professor at Westminster and Glasgow University, she is now director of the AI and Big Data Lab and the vice director of DISH, a lab on technology and culture at the University of Turin. She was the Minister of Technological Innovation and Digitisation of the Italian Republic during the Conte II government, an advisor to the Minister of Foreign Affairs, and to the Minister of Public Function in the Draghi government, as well as Councillor for Innovation and Smart City of the City of Turin under the Appendino administration. She has worked in Italy and abroad in consultancy companies on technological transformation projects for companies in various sectors, including banking, pharmaceuticals, publishing, and fashion.