EAI Endorsed Transactions
on Scalable Information Systems Research Article **EAI.EU**

Event Extraction with Spectrum Estimation Using Neural Networks Linear Methods

Vuyyuru Madhavi[1]

[1]Department of Computer Science, Eswar College of Engg, Narsaraopeta, Palnadu (Dist.), Andrra Pradesh, India

Abstract

The timely extraction of event information from enormous amounts of textual data is a critical research problem. Event extraction using deep learning technology has gained academic attention as a result of the fast growth of deep learning. Event extraction requires costly, expert-level, high-quality human annotations. As a result, developing a data-efficient event extraction model that can be trained using only a few labelled samples has emerged as a key difficulty. Existing research work focuses mainly on the structured data with supervised models. The proposed work focuses on Movie Scene Event Extraction, a practical media analysis problem that seeks to extract structured events from unstructured movie screenplays. We suggest using the correlation between various argument roles in situations where different argument roles in a movie scene share similar qualities. This can be beneficial to the Movie Scene Argument Extraction (argument classification and identification) and film scene trigger extraction (Trigger recognition and classifying). In order to represent the relation between different roles in argument and their respective roles, we propose a Superior Concept of Role (SRC) as a top-level idea of a role that is based on the classic argument role, as well as an SRC-based Graph Attention System (GAT). To assess the efficacy of the model we designed, we constructed the dataset MovieSceneEvent to extract movies' scene-related events. Additionally, we conducted tests on an existing dataset in order to compare results with different models. Results from the experiments like extraction of words, aspect keywords from the documents indicate that our model does better than other models. Furthermore, the information on the relationship between the argument roles helps improve the effectiveness of film scene extraction of events.

Keywords: Spectrum estimation, Event extraction, Movie scene arguments extraction, Scene trigger extraction, Graph attention system

Received on 31 September 2023, accepted on 06 November 2023, published on 10 November 2023

DOI: 10.4108/eetsis.4355

Corresponding Author Email: mvuyyuru9@gmail.com

1. Introduction

EVENT Extraction (EE) is a vital and challenging aspect of the field of information extraction. A single event can be described as a event that takes place in a specific date and time, that involves several people and often referred to as a change of condition. [1]. The aim of the task of Event Extraction is to convert this information from plain text that is unstructured into a structured format which primarily explains events that occurred "who", "where" and "when" of incidents that took place [2]. It makes it simpler to gather event data and study the behaviour of people which can be used to develop applications for information retrieval, suggestions, smart question-answering, information graph building as well as different tasks that are related to events. The task is to make the event data more structured and readable [3].

Close-domain event extraction as well as open-domain event extraction are the two types that deal with event removal. The events are usually seen in the framework of a pre-defined event schema in which specific individuals or things are involved in a certain moment and place [5]. Close-domain event extraction is a search for words that have a connection to a specific events schema that indicates the state or action which occurs. Its extracting goals are people, time, place as well as activities. Events are seen as a collection of connected descriptions of a subject in the open domain event extraction task, which may be expressed as a classification or clustering job. Open-domain event extraction is the process of gathering a number of events that are connected to a single subject and are often made up of many events. The goal of event extraction, whether it be a close-domain or open-domain activity, is to identify the relevant event types from a large body of texts and provide the core arguments of events in a systematic manner [6].

The research taxonomy for deep learning event extraction on the general domain has seen significant development. When an event is mentioned, it is found in phrases that include one or more triggers and arguments. It then extracts the events that contain those triggers and arguments. It is necessary to identify the event, categorize the sort of event, locate the argument, and evaluate the argument's contribution to the event [7]. The job of classifying each event's kind using several classifications is known as ‘Trigger Classification’. The role classification job establishes the role link between any two triggers and entities in a phrase. It is a multi-class classification challenge based on word pairs. Technically speaking, relation extraction, Named Entity Recognition (NER), and semantic parsing are some of the core Natural Language Processing (NLP) activities that Event extraction might rely on [8].

Extraction of movie scene events is a useful media analysis job that aids in narrative comprehension for viewers. After separating the event trigger as well as its argument from the unstructured text of the movie scene's event extraction process, the trigger and the argument are provided with a defined event type as well as a specific purpose, in turn.

Conventional event extraction models incorporated kernel-based techniques and handcrafted characteristics [9]. The use of distributional representation-based approaches, however, grew in popularity and produced greater performance as deep learning received more and more attention [8,9]. The question-answering (QA) framework, for example, has been implemented in several recent research that have evolved ways to further enhance event extraction [10,11]. Argument Extraction as well as Trigger Extraction both benefit by the connection of information between the different roles of argument..

Objective:

To extract the event / keywords from the movie scene which triggers the argument that defines the specific purpose of the movie.

Motivation:

There hasn't been a fully functional movie text extraction system created, according to a review of the literature. Furthermore, it is clear that no one method is reliable for detecting a wide range of text in the video. The majority of techniques were created to extract text from intricate color photographs and were then expanded for use with movie data. These techniques, however, do not benefit from the temporal redundancy in the movie.

The rest of this work is structured as follows. We offer similar work on event extraction in Section 2 of this article. We describe our suggested relation extraction model in Section 3 of this paper. The experimental results of our model are then presented in Section 4 along with an analysis of the findings. Section 5 concludes with our paper's findings and introduces our next work.

2. Literature Survey

A 3D Convolutional Neural Networks (3D CNN) model for action recognition was presented by Ji et al. (2013). By using 3D convolutions, this model pulls features from both the spatial and the temporal dimensions, therefore capturing the motion information stored in several neighbouring frames. From the input frames, the generated model creates numerous channels of information, and the final feature representation incorporates data from all channels. Regularization is used to combine predictions from many models and add high level characteristics to the outputs. When compared to baseline approaches, the generated models perform better at identifying human behaviors in the real-world setting of airport surveillance footage.

Mixture of Gaussians (MOG) and Principal Component Analysis (PCA) were developed by Abdallah et al. (2016) for the identification and classification of aberrant events in movies. Both locally inside pixel blocks and globally at the level of the picture, complementary models are constructed. Three aspects are analyzed: (1) motion temporal derivatives, (2) optical flow, and (3) the spatiotemporal development of binary motion, where foreground pixels are identified using an improved background removal algorithm that maintains track of momentarily static pixels. For each block and each flow characteristic at the local level, a normalcy MOG model is created and compressed using PCA. A collection of compact hybrid histograms that contain optical flow direction, temporal gradient orientation, and foreground statistics are used to qualitatively characterize the activity. These many elements are combined into a compact binary signature with a maximum size of 13 bits for event categorization.

By collecting spatiotemporal data from video sequences, Fang et al. (2016) presented a Multi scale Histogram Optical Flow (MHOF) and Principal Component Analysis Network (PCANet) for anomalous event identification. Since aberrant occurrences in video sequences often catch people's attention, saliency information (SI), which is present in video frames, is first extracted as a feature representation in the spatial domain. In the temporal domain, optical flow (OF) is thought to be a significant aspect of video sequences. Through OF, it is possible to collect the precise motion information that MHOF extracts. MHOF and SI are combined to improve the spatiotemporal characteristics of video frames. Finally, high level characteristics are extracted for anomalous event identification using a deep learning network called PCANet.

Using training samples that only included normal events, Bouindour et al. (2017) proposed a Convolutional Neural Networks (CNN) for the identification and localization of spatial and temporal anomalous events in surveillance films. This work is separated into two steps. The first stage involves extracting features from the first two convolution layers of a pre-trained CNN for each patch of the input picture. One class SVM is trained using the output features in the second step. With regard to the existence of outliers in the training dataset, SVM classifier enables quick and reliable anomalous identification.

In order to identify anomalous occurrences in public surveillance systems, Amraee et al. (2018) presented the Histogram of Oriented Gradients based Local Binary Pattern (HOG-LBP) and Histogram of Optical Flow (HOF). The duplicate information is removed from candidate areas in the first phase of the newly proposed approach. Using the appearance and motion of the extracted areas, HOG-LBP and HOF are computed for each region. Finally, two unique one class SVM models are used to identify aberrant occurrences.

A Self Organizing Map (SOM) was presented by Kumar et al. (2018) to find and list the events in multi-view surveillance films. A key component of multimedia surveillance systems might be action recognition in multiple view surveillance recordings. This approach counts the fundamental components of feature extraction, structuring, mapping, and visualization, while subjecting video skimming as action identification of the activities to event detection and summarization. For real-time applications like surveillance, security systems and other systems, this approach may perform better. The performance of the SOM-based summarization approach and cutting-edge models are compared using both qualitative and quantitative evaluation.

A modified 3D residual Convolutional Neural Network (3D CNN) was developed by Bouindour et al. (2019) to identify unusual occurrences in films. To extract spatiotemporal characteristics and create a reliable classifier based on the choice of vectors of interest, a modified version of a pretrained 3D CNN was used. It has the capacity to learn the model of typical behavior and recognize potentially harmful anomalous occurrences. Since it reduces redundant information and adapts to the advent of new normal events that occur during the testing phase, this unsupervised technique eliminates marginalization of normal events that occur seldom during the training phase.

A real-time target object event detection method using Iterative Clustering based Segmentation (ICS) was presented by Gupta et al. in 2020. One event is found in each cycle of the analysis of the test picture. The following steps are included in each cycle: (1) image segmentation using a modified K-Means clustering method; (2) elimination of segments that are empty or have no events based on statistical analysis of each segment; (3) merging segments that overlap correspond to same event; and (4) selection of the strongest event. Until all of the occurrences have been recognized, the four procedures are repeated.

A 3D CNN for video activity recognition and segmentation was presented by Hou et al. in 2021. Based on 3D convolution characteristics, this architecture is a unified deep network that can identify and localize activity. A video is first segmented into equal-length chunks, and then a collection of tube recommendations based on 3D CNN characteristics is created for each clip. Finally, spatiotemporal action identification is carried out utilizing the connected video suggestions from the tube proposals of various clips. The performance of this top-down action detection method clearly depends on a strong collection of tube suggestions, and training the bounding box regression often requires a large number of annotated examples. Each frame's foreground areas are divided into segments, and the bounding boxes are created using the segmented foreground maps. Through the use of segmentation's pixel-by-pixel annotations, this bottom-up strategy successfully avoids the development of tube proposals.

For the purpose of detecting video events while using underlying semantic clues, Huang et al. (2022) presented the hierarchical unified model and max margin latent ideas. Intentionally combining the processes of underlying semantics discovery and event modelling from video data, the model for video event detection is a single entity. The unified model not only makes it possible to represent videos in a discriminative and descriptive way, but it also solves the issue of error propagation from video representation to event modelling that plagued earlier approaches. The model is learned with the help of a max margin framework.

3. Proposed Model

In this section, we provide a thorough overview of our method for extracting movie scene events. The complete movie scene extraction procedure is illustrated schematically in Figure 1. This shows the three steps that comprise the entire process. First, you need to obtain the SRC embedding which is focused on arguments through the judicious higher-level module for role. It is the second step to incorporate argument based SRC embedding to facilitate trigger extraction into the GAT framework. Before we present GAT as a GAT structure, the rest of this article explains how to set up the argument based SRC embedding through an extremely high-level role model. We also demonstrate how you can extract arguments as well as triggers within GAT with the Argument-oriented SRC embedding.

An event or gerund is often used to start an event, which is the occurrence of an action or a change in status. It comprises the main plot elements, such as the setting, the characters, and the period. The orderly display of events that users are interested in from unstructured texts is made possible by the use of event extraction technologies. Briefly stated, event extraction, as shown in Fig. 1, locates events based on their categorization and extracts the primary arguments from the text. The events in a text, the triggers, the arguments that are connected to every single event, as well as the reason for each argument could be discovered using an extraction method. Locating the two occasions (Die as well as Attack) that are caused by the terms "died" and "fired," respectively this is the initial step of the process for extracting events. For the Die category of events, "Baghdad", "cameraman", "cameraman" and "American tank" play the role as Place and Victim, Place and Instrument respectively. Attack gives the role of an event's argument Place and Instrument "Baghdad" and "American tank" for example. In addition, "cameraman" and "Palestine Hotel" take on the roles of the argument for events Target. Many cutting-edge fields, including machine learning, pattern matching, and NLP, are involved in event extraction [14]. Likewise, event extraction in a variety of sectors may speed up job turnaround times, enhance quantitative analysis technically, and assist relevant workers in swiftly locating pertinent material among a wealth of information. Event extraction hence offers a wide range of potential applications in numerous industries. The phrase "event extraction task" in Automatic Content Extraction (ACE) often refers to the following terminologies:

• Entity: An item or set of objects in a semantic category constitutes an entity. Entity mostly refers to individuals, groups, locations, occasions, items, etc. Entity includes the terms "Baghdad," "cameraman," "American tank," and "Palestine Hotel."
• Mentions of the event: The trigger and pertinent justifications are included in the words or sentences that describe the incident.

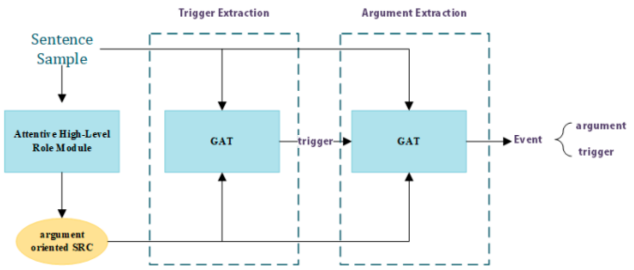

Figure 1. Movie scene event extraction procedure

Semantic Level Contexts

Semantic interactions between event entities are reflected in contexts. Because an event's semantic level context includes both the object and the person, two of the event's most important constituents and because it shows the connections between the three, the events are presented in a symbolic setting made up of scenes, acts, and ideas connected to physical things. Events with a variety of attributes may be represented using this area [15].

Described as a semantic K-dimensional space with each dimension by K ideas C = C1, C2, C3,...... Cn. A collection of k functions is provided that assign a value to a video that shows the confidence in the existence of a given notion to one that is embedded in a K-dimensional space. The definition of determines the applicability. Also take note that the idea detector is not required. If the idea detector uses the full video as a single input, both may be handled equally. If a sliding window that splits a video into input windows and generates W outputs is used to expand the detector to a video, for combining the W outputs from the idea detector into a single esteem value, the max function has been developed. As a result, the K-dimensional semantic space is extended by the function set to include a video as a vector (C1..Ck). A space cluster is made up of movies with comparable semantic content, and it is from this space cluster that the training classifier does event recognition. Additionally, the classification of events is divided into stages, a video in idea space is provided, and events are identified using derived characteristics. The statistics of the concept score technique explain the maximum detection score, but it is inappropriate for various uses, thus scoring distribution is required to represent a specific event. A pack of ideas feature calculates the frequency of each phrase across the full video clip in line with the bag of concepts descriptors used in visual word-like features [16].

The attentive high-level role module's structure is shown in Figure 2, and we'll go into more depth about it below.

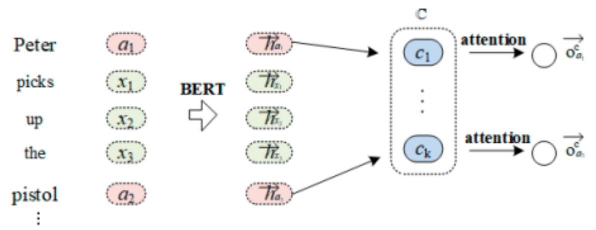

Figure 2. Framework for high-level role modules that are attentive.

Event Trigger Extraction

An illustration of the whole event extraction procedure as well as the specifics of trigger extraction are shown in Figure 3. We outline the specifics of the latter procedure in this section.

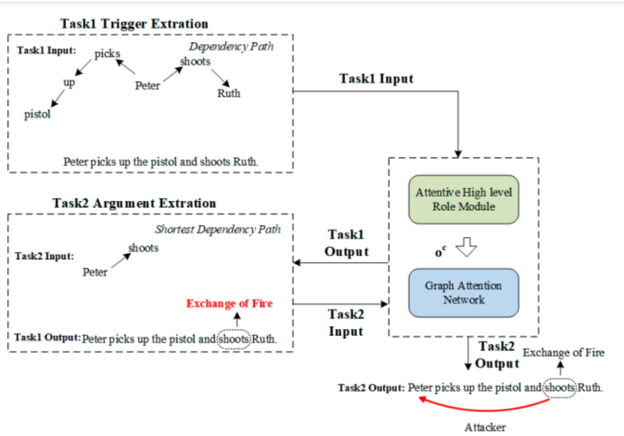

Figure 3. shows the general structure of the whole event extraction model in (a) as well as the specifics of trigger extraction in (b).

When compared to traditional sequential models, it exhibits gains in capturing semantic characteristics. One may consider GAT to be an expansion of memory networks.

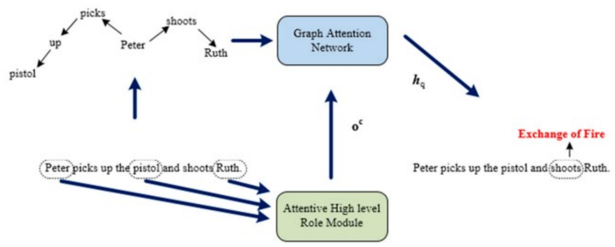

Figure 4. The weighted sum of the hidden states of each node in GAT's graph structure is calculated using the attention unit to determine each node's hidden states.

One node in the GAT is shown operating in Figure 4 with regard to every node in the dependency tree Image 3. (1) The general structure of the whole event extraction model; (2) Trigger extraction's specifics. The Network for Graph Attention as its name implies demonstrates advancements in capture.

4. Experiments

4.1. Experimental Setup

Datasets:

1. **MovieSceneEvent:** For this proposed work, we created a movie scene event extraction dataset called MovieSceneEvent. Based on study requirements and recommendations from experts in the film industry, we initially summarized 12 typical categories of events to create a movie-scene-specific event extraction dataset. After that, we selected terms from the screenplay text which were relevant to the circumstances. A variety of genres that are well-liked by film fans were selected to select the scripts for these movies (including comedy, romance, action films, war etc.). Prior to manually filtering these terms into the appropriate events, we employed the manual template to approximately separate all the text that is linked to the specific event type, the script's text. These phrases were then manually given additional labels. Each sample was assigned a label by two annotators. This outcome was applied to the sample if its labelling was reliable. If not, an additional annotator was employed to ensure that the labeling was correct. The movie scene's event extraction dataset, consisting of 486 samples for

testing, contains twelve event categories as well as 18 arguments roles.

2. ACE2005:

As with previous studies [3,8], we used ACE2005, an ace event extraction dataset to test the efficiency of our methodology. The ACE2005 data set includes 13,672 words and 599 documents and have been classified according to eight types of events along with 33 types of events and 35 arguments that have a given function. The ACE2005 data set has been divided into data into 529, 30, and 40 pages that can be utilized to develop, train, and testing as per the [4,6].

Hyper parameters, in order to obtain the embeddings hidden and to use the BERT-BASE-CASED paradigm, we used BERT to encoded sequences. The following hyper parameters were evaluated: learning rate is 0.0025, 0.005, 0.01. The batch size was 15,25,50. The results of the experiment were based on various parameters available on the MovieSceneEvent dataset. The rate of learning and the batch size that we agreed on were 0.005 and 25 respectively. Table 1 gives a listing of the parameters that were used. The operating environment was comprised of two NVIDIA K40 systems, and Table 1 contains information about the training of the model.

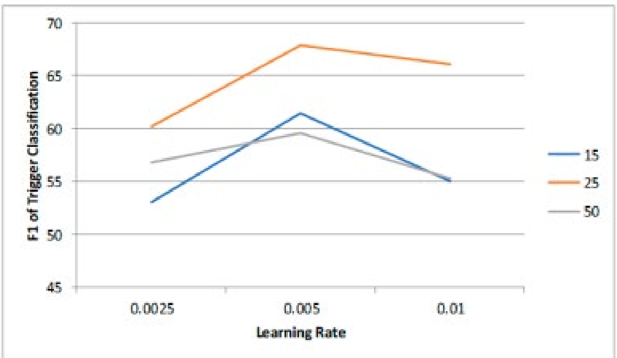

Figure 6. The impact of learning rate and batch size on categorization of arguments.

Table 1

Parameter	Value
Word Embedding size d	766
Batch size	25
Word count	3
Max count	5
Learning Rate	0.001
Epoch	100
Optimiser	Adam

Following is what we may deduce from the experimental findings:

(1) The trigger extraction as well as the argument extraction in movie scene events extraction as well as Open Domain Event Extraction routinely outperform other models. This suggests that the SRC data could aid in the trigger extraction as well as argument extraction for event extraction.

(2) The method we propose is superior to earlier studies in the field of argument extraction possibly due to that the SRC information is in a more direct relation to the arguments role.

(3) It is crucial to be aware that the F1 gap between the classification of the roles of arguments and trigger events, and between trigger identification and classification, and trigger identification and classification is lower than it was in previous studies. A greater amount of semantic information is preserved between the identification and classification by the aid of SRC data.

(4) The model's performance improvement is considerably lower when it comes to the performance of open domain datasets. This could be due to the fact that it's much easier to expand the argument role composition of movie scene extraction, which can be incorporated into several better role concepts. In the end, SRC data has more influence.

5. Discussion

To determine the events of a movie scene, this model has been presented in the form of an attention graph with argument correlation in this study. The research evaluated and contrasted the effectiveness of five different models: JOINTFEATURE, JOINTFEATURE DbRNN Joint3EE, BS, and Text2Event with two different data sets: MovieSceneEvent and ACE2005 to test if the model was effective. model. Our findings indicate that the proposed model is more efficient than any other model within the domain of movie scene events extraction and also free domain events with regard to the extraction of argument and triggers. The efficiency of our model for movie scene extract results is much higher contrasted to the results of the open domain data collection. This could be due to the fact that it's considerably easier to extend the argument-role composition for movie scene extraction and to incorporate a range of better role concepts. In addition, the research in ablation within the Section shows how beneficial the data on correlation across different roles is in triggers and argument extraction. In addition, research on the effects of the size of the dataset indicates that when the data set is significantly reduced the component can have an impact on effectiveness. The impact of increasing the size of the training sets in terms of outcomes decreases when the data expands in dimensions.

6. Conclusions.

Since the existing open-domain method of event extraction haven't utilized the full relationship information from arguments, which constitutes important implicit semantic data in the process of determining movie scene events and have not paid enough particular attention to the facts in the domain, we offer an argument correlation-based and information-based movie scene extraction model in our research paper. The GAT is built on SRC to capture this hidden information, and then combine the information about correlation of the argument roles with semantic aspects to make use of this important implicit semantic information for enhancing film scene events extraction. Dependency tree structures permits to the GAT module to gather information about semantics while integrating the SRC data with the secret embedding of nodes. We developed a film scene extracting dataset in order to assess the effectiveness of our model. According to experimental findings, the SRC aids in enhancing the effectiveness of both argument and event trigger extraction. Our model may considerably enhance movie scene event extraction performance. Additionally, it performs well while extracting open domain events. We will make more use of the impact of outside data on event extraction in the future. We'll also make an effort to combine our model with several lately used frameworks, such a QA framework.

7, References

[1] [1] Mcclosky, D.; Surdeanu, M.; Manning, C.D. Event Extraction as Dependency Parsing. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; pp. 1626–1635.

[2] [2] Liao, R. Using Document Level Cross-Event Inference to Improve Event Extraction. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; pp. 789–797.

[3] [3] Li, Q.; Ji, H.; Huang, L. Joint Event Extraction via Structured Prediction with Global Features. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Sofia, Bulgaria, 4–9 August 2013; pp. 73–82.

[4] [4] Ji, H.; Grishman, R. Refining Event Extraction through Cross-Document Inference. In Proceedings of the ACL-08: HLT, Columbus, OH, USA, 19–20 June 2008; pp. 254–262.

[5] [5] Huang, R.; Riloff, E. Modeling Textual Cohesion for Event Extraction. Proc. AAAI Conf. Artif. Intell. 2021, 26, 1664–1670. [CrossRef]

[6] [6] Sha, L.; Qian, F.; Chang, B.; Chang, B.; Sui, Z. Jointly Extracting Event Triggers and Arguments by Dependency-Bridge Rnn and Tensor-Based Argument Interaction. Proc. AAAI Conf. Artif. Intell. 2018, 32, 5916–5923. [CrossRef]

[7] [7] Nguyen, T.H.; Grishman, R. Graph Convolutional Networks with Argument-Aware Pooling for Event Detection. Proc. AAAI Conf. Artif. Intell. 2018, 32, 5900–5907. [CrossRef]

[8] [8] Nguyen, T.H.; Grishman, R. Event Detection and Domain Adaptation with Convolutional Neural Networks. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), Beijing, China, 26–31 July 2015; Association for Computational Linguistics: Stroudsburg, PA, USA, 2015.

[9] [9] Chen, Y.; Xu, L.; Liu, K.; Zeng, D.; Zhao, J. Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Beijing, China, 26–31 July 2015; Association for Computational Linguistics: Stroudsburg, PA, USA, 2015.

[10] [10] Chen, Y.; Chen, T.; Ebner, S.; White, A.S.; Van Durme, B. Reading the Manual: Event Extraction as Definition Comprehension.arXiv 2019, arXiv:1912.01586.

[11] [11] G. Ansmann, R. Karnatak, K. Lehnertz, and U. Feudel, Extreme events in excitable systems and mechanisms of their generation, Phys. Rev. E 88, 052911 (2013).

[12] [12] T. Bódai, G. Károlyi, and T. Tél, Driving a conceptual model clim ate by different processes: Snapshot attractors and extreme events, Phys. Rev. E 87, 022822 (2013).

[13] [13] R. Karnatak, G. Ansmann, U. Feudel, and K. Lehnertz, Route to extreme events in excitable systems, Phys. Rev. E 90, 022917 (2014).

[14] [14] C. Mulhern, S. Bialonski, and H. Kantz, Extreme events due to localization of energy, Phys. Rev. E 91, 012918 (2015).
[15] [15] D. Qi and A. J. Majda, Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory, Commun. Math. Sci. 14, 1687 (2016).
[16] [16] M. A. Mohamad and T. P. Sapsis, Sequential sampling strategy for extreme event statistics in nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA 115, 11138 (2018).
[17] [17] C. T. Bolles, K. Speer, and M. N. J. Moore, Anomalous wave statistics induced by abrupt depth change, Phys. Rev. Fluids 4, 011801(R) (2019).
[18] [18] S. Hallerberg, E. G. Altmann, D. Holstein, and H. Kantz, Precursors of extreme increments, Phys. Rev. E 75, 016706 (2007).
[19] [19] R. V. Donner and S. M. Barbosa, Nonlinear time series analysis in the geosciences, Lect. Notes Earth Sci. 112, 37 (2008).
[20] H. L. D. de S. Cavalcante, M. Oriá, D. Sornette, E. Ott, and D. J. Gauthier, Predictability and Suppression of Extreme Events in a Chaotic System, Phys. Rev. Lett. 111, 198701 (2013).