Overview of the FIRE 2020 EDNIL Track: Event Detection from News in Indian Languages

Bhargav Dave[a], Surupendu Gangopadhyay[a], Prasenjit Majumder[a], Pushpak Bhattacharya[b], Sudeshna Sarkar[c] and Sobha Lalitha Devi[d]

[a] *Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India*
[b] *Indian Institute of Technology Bombay, Mumbai, India*
[c] *Indian Institute of Technology Kharagpur, Kharagpur, India*
[d] *AU-KBC Research Centre,MIT Campus of Anna University, Chennai, India*

Abstract

The goal of FIRE 2020 EDNIL track was to create a framework which could be used to detect events from news articles in English, Hindi, Bengali, Marathi and Tamil. The track consisted of two tasks: (i) Identifying a piece of text from news articles that contains an event (Event Identification). (ii) Creating an event frame from the news article (Event Frame Extraction). The events that were identified in Event Identification task were Man-made Disaster and Natural Disaster. In Event Frame Extraction task the event frame consists of Event type, Casualties, Time, Place, Reason.

Keywords

Multilingual Event Detection, Event Identification, Event Frame Extraction,

1. Introduction

An event is defined as an occurrence happening in a certain place during a particular interval of time with or without the participation of human agents. It may be part of a chain of occurrences or an outcome or effect of preceding occurrence or a cause of succeeding occurrences. An event can occur naturally or it can be because of human actions. An event can have a location, time, agents involved (causing agent and on which the effect of the event is felt) etc.

This paper gives the description of FIRE 2020 shared task:Event Detection from News in Indian Languages (EDNIL). We give a short description of the sub-tasks, the multilingual dataset that was used in the subtasks and the results that were obtained in the subtasks. Two tasks were proposed in the track: (1) Identifying a piece of text from news articles that contains an event (Event Identification). (2) Creating an event frame from the news article (Event Frame Extraction). In both the tasks news articles of five Indian languages: English, Hindi, Bengali, Marathi and Tamil were used as dataset.

Forum for Information Retrieval Evaluation, 16-20 December 2020, Hyderabad, India
bhargavdave1@gmail.com (B. Dave); surupendu.g@gmail.com (S. Gangopadhyay); prasenjit.majumder@gmail.com (P. Majumder); pushpakbh@gmail.com (P. Bhattacharya); shudeshna@gmail.com (S. Sarkar); sobha@au-kbc.org (S. L. Devi)
0000-0003-2742-480X (B. Dave)

1.1. Task 1: Event Identification

In this task the participants had to identify a event given a news article. The events were of two type: Natural disaster and Manmade disaster.

1.2. Task 2: Event Frame Extraction

In this task the participants had to form an event frame given a news article.The event frame consists of the following fields:

1. Type: Detect the type of the event. There are two type of events
   a) Natural disaster
   b) Manmade disaster
2. Subtype: It is the event which is subtype of Natural or Manmade disaster.
   The subtypes of Natural disaster are forest fire, hurricane, cold wave, tornado, storm, hail storms, blizzard, avalanches, heat wave, cyclone, drought, heavy rainfall, limnic erruptions, floods, tsunami, land slide, volcano, earthquake, rock fall, seismic risk, famine, epidemic and pandemic.
   The subtypes of Manmade disaster are crime, riots, aviation hazard, accidents, train collision, vehicular collision, transport hazards, industrial accident, fire, normal bombing, terrorist attack, miscellaneous, shoot out, surgical strikes, suicide attack and armed conflicts.
3. Casualties: Number of people injured or killed and Damage to properties.
4. Time: When the event took place
5. Place: Where the event took place
6. Reason: Why and how the event took place

Shared tasks on event detection have also been proposed earlier, such as TAC-KBP 2016 Event Nugget track [1] where the task was to detect an event and then link the words that refer to that event from English, Spanish and Chinese articles, FIRE 2018 EventXtract-IL [2] where the task was to detect an event and also extract arguments like location, cause, effect from Hindi and Tamil news articles. CLEF 2019 Lab ProtestNews [3] where the task was to detect protest news and form an event frame (Event, Participant, Target, Place, Time) from English news articles.

The contribution of EDNIL is that it provides an annotated dataset for event detection from five Indian languages i.e. English, Hindi, Bengali, Marathi and Tamil.

2. Dataset

The dataset was created as part of the project "A Platform for Cross-lingual and Multilingual Event Monitoring in Indian Languages" [1]. The dataset consists of news articles in English, Hindi, Bengali, Marathi and Tamil languages which have been collected from different news agencies. The statistics of the dataset documents is shown in Table 1.

[1]https://imprint-india.org/knowledge-portal-5592-a-platform-for-crosslingual-and-multilingual-event-monitoring-in-indian-languages

Table 1
Statistics of Train and Test Data

Language	Train	Test	Total
English	828	206	1034
Hindi	828	194	1022
Bengali	800	204	1004
Tamil	1013	257	1270
Marathi	1035	265	1300
Total	4504	1126	5630

Table 2
Statistics of Annotation Tags in the Dataset

Tag	English		Hindi		Bengali		Tamil		Marathi	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
MAN MADE EVENT	3774	891	2185	544	4233	966	3255	997	2571	530
NATUARAL EVENT	1078	103	2279	531	887	310	1185	333	2259	275
CASUALTIES_ARG	2708	633	2166	484	3480	859	2247	746	2364	353
TIME_ARG	1454	315	1579	395	2600	645	842	259	1435	311
PLACE_ARG	2324	455	4045	952	3176	863	2335	753	4021	645
REASON_ARG	562	125	285	71	364	93	426	90	434	85

News article of each language is annotated manually by annotators from IIT Kharagpur (Bengali), IIT Bombay (Marathi), IIT Patna (Hindi), AU-KBC (English and Tamil). The annotation has been done at word level and the news articles after annotation are stored in XML format. The description of the XML tags are given below and the statistics of the XML tags is shown in Table 2.

```
<MAN_MADE_EVENT ID="number" TYPE="subtype">
Event Trigger
</MAN_MADE_EVENT>

<NATURAL_EVENT ID="number" TYPE="subtype">
Event Trigger
</NATURAL_EVENT>
```

Here MAN_MADE_EVENT and NATURAL_EVENT tag is related to Manmade disaster and Natural disaster event respectively, contains the event trigger and has the following attributes:

1. ID : A number which is unique for each event/tag in a given document.
2. TYPE : Represents subtype of the particular event (Manmade disaster or Natural disaster).

The event Manmade disaster has subtypes crime, riots, aviation hazard, accidents, train collision, vehicular collision, transport hazards, industrial accident, fire, normal bombing, terrorist attack, miscellaneous, shoot out, surgical strikes, suicide attack and armed conflicts. Language wise details statistics of subtypes of man made event XML tag shown in Table 3. The event Natural Disaster has subtypes forest fire, hurricane, cold wave, tornado, storm, hail

Table 3
Statistics of subtypes of Manmade disaster XML tag

Subtype	English		Hindi		Bengali		Tamil		Marathi	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
CRIME	98	64	0	0	0	0	818	0	0	0
RIOTS	15	6	143	22	144	23	32	53	54	27
AVIATION HAZARD	78	33	94	27	84	42	76	40	118	5
ACCIDENTS	735	310	0	0	0	0	317	0	0	0
TRAIN COLLISION	109	9	139	41	44	4	24	19	40	25
VEHICULAR COLLISION	643	116	250	63	688	162	329	113	402	53
TRANSPORT HAZARDS	323	9	132	37	210	49	137	0	40	66
INDUSTRIAL ACCIDENT	120	12	194	58	90	5	21	25	390	6
FIRE	806	99	229	72	384	82	313	199	279	42
NORMAL BOMBING	153	20	61	5	916	174	210	191	241	120
TERRORIST ATTACK	67	0	299	72	285	84	117	96	252	77
MISCELLANEOUS	59	85	0	0	0	0	0	0	0	0
SHOOT OUT	341	43	282	65	495	138	497	138	287	85
SURGICAL STRIKES	106	15	0	76	170	32	188	78	68	1
SUICIDE ATTACK	110	4	326	76	386	87	51	45	125	4
ARMED CONFLICTS	11	1	36	5	337	84	125	0	193	19

storms, blizzard, avalanches, heat wave, cyclone, drought, heavy rainfall, limnic erruptions, floods, tsunami, land slide, volcano, earthquake, rock fall, seismic risk, famine, epidemic and pandemic. Language wise details statistics of subtypes of natural disaster event XML tag shown in Table 4.

The event arguments are casualties, reason, time of occurrence of event and location of event. The XML tags wrt each event argument is given below:

1. <CAUSUALITIES-ARG> : This tag contains the words that are casualties that have occurred due to an event.
2. <TIME-ARG> : This tag contains the words that are time at which the event has occurred.
3. <PLACE-ARG> : This tag contains the words that is the place at which the event has occurred.
4. <REASON-ARG> : This tag contains the words that are the reason due to which the event has occurred.

For example, the “casualties” attribute of an event is annotated as follows:

```
<CASUALTIES-ARG ID="number">
casualties
<CASUALTIES-ARG>
```

Each argument tag of an event has the attribute “ID,” which is an unique number for each tag in a given news article.

An example, of annotation of man-made event news ”The accident occurred around 6.30 pm at Manathoor Church junction on the Pala-Thodupuzha State Highway.” is shown in Fig. 1 and an example annotation of natural event news ”An earthquake measuring 5.5 on the Richter

Table 4
Statistics of subtypes of Natural disaster XML tag

Subtype	English		Hindi		Bengali		Tamil		Marathi	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
FOREST FIRE	57	0	114	35	9	0	5	12	63	5
HURRICANE	35	0	132	35	7	0	0	15	0	0
COLD WAVE	23	0	101	15	9	8	0	0	117	7
TORNADO	52	13	113	30	0	11	0	0	0	0
STORM	104	11	401	100	107	18	9	29	71	2
HAIL STORMS	23	3	106	23	0	0	0	0	119	1
BLIZZARD	10	0	74	10	18	2	0	6	214	0
AVALANCHES	34	4	135	31	1	0	0	7	91	0
HEAT WAVE	15	4	185	29	48	5	4	3	72	5
CYCLONE	87	4	142	40	28	0	223	14	415	2
DROUGHT	7	0	5	0	3	11	0	0	23	7
HEAVY RAINFALL	1	0	0	0	0	0	0	0	0	0
LIMNIC ERRUPTIONS	2	0	0	0	0	0	0	5	0	0
FLOODS	158	0	173	40	27	9	343	31	178	78
TSUNAMI	11	1	9	1	28	15	10	11	159	39
LAND SLIDE	65	5	157	44	20	11	123	37	129	38
VOLCANO	88	0	96	21	4	0	9	3	139	2
EARTHQUAKE	256	58	336	77	203	112	320	146	411	88
ROCK FALL	3	0	0	0	0	0	0	0	57	1
SEISMIC RISK	0	0	0	0	0	1	1	0	1	0
FAMINE	1	0	0	0	0	0	3	10	0	0
EPIDEMIC	46	0	0	0	150	34	104	0	0	0
PANDEMIC	0	0	0	0	225	73	31	4	0	0

Scale rattled the north-east coast of Japan's Amami Oshima Island on Wednesday." is shown in Fig. 2.

3. Evaluation

In both task 1 and task 2 the evaluation metrics that was used was F1-score. The F1-score was calculated separately for all the five languages in both Task 1 and Task 2. For Task 2 the F1 score was calculated separately for each argument in the event frame and then the score was averaged out. While evaluating the arguments in the event frame only exact string match of the values was considered. Eg: If the PLACE argument in test article is New Delhi and the output of the PLACE argument for test article given by the participant's method is Delhi then it was not be considered as a match.

```
<P>
    <W> The </W>
    <MAN_MADE_EVENT ID="7" TYPE="ACCIDENTS">
        <W> accident </W>
    </MAN_MADE_EVENT>
    <W> occurred </W>
    <W> around </W>
    <TIME-ARG ID="8">
        <W> 6.30 </W>
        <W> pm </W>
    </TIME-ARG>
    <W> at </W>
    <PLACE-ARG ID="9">
        <W> Manathoor </W>
        <W> Church </W>
        <W> junction </W>
        <W> on </W>
        <W> the </W>
        <W> Pala-Thodupuzha </W>
    </PLACE-ARG>
    <W> State </W>
    <W> Highway.</W>
</P>
```

Figure 1: Sample Annotation of manmade event news "The accident occurred around 6.30 pm at Manathoor Church junction on the Pala-Thodupuzha State Highway. "

4. Results

For the first task of Event Identification in English language, we received seven runs from five teams. For Hindi language we received five runs from three teams. For Bengali language we received six runs from four teams. In Marathi and Tamil language, for each we received two runs from two teams.

For the second task of Event Frame Extraction in English language, we received three runs from three teams. In case of Hindi, Bengali, Marathi and Tamil languages for each language we received one run from one team. The submission statistics are shown in Table 5. The results for all the five languages shown from Tables 6,7,8,9.

Team 3Idiots [4] ranked first for both Task 1 and Task 2 across all languages. They used n-gram and regex based features for representing the news articles. And then used these features in a CRF model for doing Task 1 and Task 2. For each language the CRF model was trained separately.

```
<P>
    <W> An </W>
    <NATURAL_EVENT ID="3" TYPE="EARTHQUAKE">
        <W> earthquake </W>
    </NATURAL_EVENT>
    <W> measuring </W>
    <W> 5.5 </W>
    <W> on </W>
    <W> the </W>
    <W> Richter </W>
    <W> Scale </W>
    <W> rattled </W>
    <W> the </W>
    <PLACE-ARG ID="5">
        <W> north-east </W>
        <W> coast </W>
        <W> of </W>
        <W> Japan's </W>
        <W> Amami </W>
        <W> Oshima </W>
        <W> Island </W>
    </PLACE-ARG>
    <W> on </W>
    <TIME-ARG ID="6">
        <W> Wednesday. </W>
    </TIME-ARG>
</P>
```

Figure 2: Sample Annotation of natural event news "An earthquake measuring 5.5 on the Richter Scale rattled the north-east coast of Japan's Amami Oshima Island on Wednesday. "

Team BUDDI_SAP [2] ranked second in both task in English language. They used DistillBERT based word embedding, POS tags based embeddings and character level embeddings which were then concatenated together to represent a word. This was then passed through Bi-LSTM the output of which passed through fully connected layer which was used to predict the words associated with an argument. Two separate models were trained for Task 1 and Task 2.

Run number 3,2 and 1 of team ComMA [5] were ranked second,third and fourth respectively for Task 1 in Hindi and Bengali languages. And third, fourth and fifth for Task 1 in English language. In run number 3 XLM RoBERTa was used for text representation of all three languages mentioned earlier, which was then fine tuned for Task 1, in run number 2 DistillBERT was used

[2]Anand Subramanian, Praveen Kumar Suresh, Sharafath Mohamed were not able to submit a paper due to prior commitments but gave a presentation in FIRE 2020

Table 5
Submission Statistics for all languages

Submission	Task 1	Task 2
English	7	3
Hindi	5	1
Bengali	6	1
Tamil	2	1
Marathi	2	1

Table 6
Results of Task 1 and Task 2 for English

Task1

SR NO.	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.7925170068	0.7032193159	0.7452025586	N-Gram & Regex + CRF
2	BUDDI_SAP	1	0.6110581506	0.6448692153	0.6275085658	DistilBERT, POS tag & Character level embedding + Bi-LSTM
3	ComMA	3	0.5911885246	0.5834175935	0.5872773537	XLM RoBERTa
4	ComMA	2	0.5846774194	0.587639311	0.5861546235	DistilBERT
5	ComMA	1	0.5800395257	0.5905432596	0.5852442672	BERT
6	MUCS	1	0.3066255778	0.4004024145	0.3472949389	N-gram,Suffix & Prifix + Linear SVC
7	NLP@ISI	1	0.3109475621	0.3400402414	0.3248438251	Bag of Word

Task2

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.5038099507	0.4469184891	0.4736620312	N-Gram & Regex + CRF
2	BUDDI_SAP	1	0.2008368201	0.248111332	0.2219850587	DistilBERT, POS tag & Character level embedding + Bi-LSTM
3	NLP@ISI	1	0.1128436602	0.1093439364	0.1110662359	Bag of Word

for text representation of all three languages, which was then fine tuned for Task 1.And in run number 3 BERT was used for text representation of all three languages, which was then fine tuned for Task 1.

Team MUCS [6] ranked second in Task 1 in Marathi and Tamil languages, ranked fifth in Task 1 in Hindi and Bengali languages and ranked sixth in Task 1 in English language. They used Linear SVC based on char n-grams, suffix and prefix features of tokens for all the five language of Task 1.

Team NLP@ISI [7] ranked sixth and seventh for Bengali and English language respectively in Task 1 and ranked third in Task 2 in English language. They used bag-of-words approach to

Table 7
Results of Task 1 and Task 2 for Hindi

Task1

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.6851612903	0.5691318328	0.6217798595	N-Gram & Regex + CRF
2	ComMA	3	0.5046641791	0.5167144222	0.5106182161	XLM RoBERTa
3	ComMA	2	0.4963167587	0.5133333333	0.5046816479	DistilBERT
4	ComMA	1	0.4776785714	0.5095238095	0.4930875576	BERT
5	MUCS	1	0.1981491562	0.3453510436	0.2518159806	N-gram,Suffix & Prifix + Linear SVC

Task2

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.4722369117	0.3405797101	0.3957456238	N-Gram & Regex + CRF

Table 8
Results of Task 1 and Task 2 for Bengali

Task1

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.7045226131	0.5532754538	0.6198054819	N-Gram & Regex + CRF
2	ComMA	3	0.3788343558	0.3914421553	0.385035074	XLM RoBERTa
3	ComMA	2	0.3902654867	0.3505564388	0.3693467337	DistilBERT
4	ComMA	1	0.3457804332	0.3668779715	0.3560169166	BERT
5	MUCS	1	0.1732625483	0.2833464878	0.2150344414	N-gram,Suffix & Prifix + Linear SVC
6	NLP@ISI	1	0.09563994374	0.1073401736	0.1011528449	Bag of Word

Task2

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.5476017442	0.410626703	0.4693241981	N-Gram & Regex + CRF

identify the disaster event and used string based keyword matching to identify the arguments like Casualty, Reason.

5. Concluding Discussions

The FIRE 2020 EDNIL track was successful in releasing a multilingual dataset of Indian languages for event detection. As can be observed from the result tables for Task 1 barring English there is still lot of scope to improve the F1 scores for other languages. And for Task 2 there is still a huge scope for improvement in all languages. In the future we plan to extend the task by introducing event linking which will link one event to another if they are related to each other.

Table 9
Results of Task 1 and Task 2 for Marathi

Task1

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.6092362345	0.4336283186	0.5066469719	N-Gram & Regex + CRF
2	MUCS	1	0.1239203905	0.417193426	0.1910828025	N-gram,Suffix & Prifix + Linear SVC

Task2

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.3871382637	0.2784458834	0.3239171375	N-Gram & Regex + CRF

Table 10
Results of Task 1 and Task 2 for Tamil

Task1

SR NO	Team Name	Run	Precision	Recall	F1-Score	Method Summary
1	3Idiots	1	0.6921296296	0.6764705882	0.6842105263	N-Gram & Regex + CRF
2	MUCS	1	0.1383417316	0.2277526395	0.1721288116	N-gram,Suffix & Prifix + Linear SVC

Task2

SR NO	Team Name	Run	Precision	Recall	F1-Score	
1	3Idiots	1	0.505633322	0.4688192466	0.4865308804	N-Gram & Regex + CRF

For evaluation we intend to evaluate partial matching strings along with full matching strings. We also plan to introduce a summarization of event task wherein a summary of events within a particular time period will be generated and a short description of the events will be generated. However for this task annotators will be required who can create a gold standard dataset of event based summaries, which may require significant amount of time.

Acknowledgments

The track organizers thank all the participants for their interest in this track. We also thank the FIRE 2020 organizers for their support in organizing the track. We thank the Principal Investigator, Co-Principal Investigators and Host Institute (IIT Kharagpur) of "A Platform for Crosslingual and Multilingual Event Monitoring in Indian Languages" for providing us with this opportunity of using the dataset in the track. We also thank Ministry of Electronics and Information Technology (MeitY) and Ministry of Human Resource Development, Government of India for providing this opportunity to develop the dataset and other resources.

References

[1] Y. Zeng, B. Luo, Y. Feng, D. Zhao, Wip event detection system at tac kbp 2016 event nugget track, TAC (2016).

[2] P. R. K. Rao, S. L. Devi, Eventxtract-il: Event extraction from newswires and social media text in indian languages @ FIRE 2018 - an overview, in: P. Mehta, P. Rosso, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2018 - Forum for Information Retrieval Evaluation, Gandhinagar, India, December 6-9, 2018, volume 2266 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018, pp. 282–290. URL: http://ceur-ws.org/Vol-2266/T5-1.pdf.

[3] A. Hürriyetoğlu, E. Yörük, D. Yüret, Ç. Yoltar, B. Gürel, F. Duruşan, O. Mutlu, A. Akdemir, Overview of clef 2019 lab protestnews: Extracting protests from news in a cross-context setting, in: F. Crestani, M. Braschler, J. Savoy, A. Rauber, H. Müller, D. E. Losada, G. Heinatz Bürki, L. Cappellato, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction, Springer International Publishing, Cham, 2019, pp. 425–432. doi:10.1007/978-3-030-28577-7_32.

[4] S. Mishra, Non-neural Structured Prediction for Event Detection from News in Indian Languages, in: P. Mehta, T. Mandl, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020, CEUR Workshop Proceedings, CEUR-WS.org, 2020.

[5] B. L. Ritesh Kumar, A. Ojha, CoMA@FIRE 2020: Exploring Multilingual Joint Training across different Classification Tasks, in: P. Mehta, T. Mandl, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020, CEUR Workshop Proceedings, CEUR-WS.org, 2020.

[6] F. Balouchzahi, H. Shashirekha, An Approach for Event Detection from News in Indian Languages using Linear SVC, in: P. Mehta, T. Mandl, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020, CEUR Workshop Proceedings, CEUR-WS.org, 2020.

[7] S. Basak, Event Detection from News in Indian Languages Using Similarity Based Pattern Finding Approach, in: P. Mehta, T. Mandl, P. Majumder, M. Mitra (Eds.), Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020, CEUR Workshop Proceedings, CEUR-WS.org, 2020.