Findings of Shared Task on Offensive Language Identification in Tamil and Malayalam

Prasanna Kumar Kumaresan
Indian Institute of Information Technology and Management-Kerala, Kerala, India
prasanna.mi20@iiitmk.ac.in

Premjith B
Amrita Vishwa Vidyapeetham,Coimbatore, Tamil Nadu, India
b_premjith@cb.amrita.edu

Ratnasingam Sakuntharaj
Eastern University
Sri Lanka
sakuntharaj@esn.ac.lk

Sajeetha Thavareesan
Eastern University
Sri Lanka
sajeethas@esn.ac.lk

Subalalitha Chinnaudayar Navaneethakrishnan
SRM Institute of Science and Technology, Chennai, Tamil Nadu, India
subalalitha@gmail.com

Anand Kumar Madasamy
National Institute of Technology Karnataka Surathkal, Karnataka, India
m_anandkumar@nitk.edu.in

Bharathi Raja Chakravarthi
Insight SFI Research Centre for Data Analytics, National University of Ireland
Galway, Ireland
bharathi.raja@insight-centre.org

John P. McCrae
Insight SFI Research Centre for Data Analytics, National University of Ireland
Galway, Ireland
john.mccrae@insight-centre.org

ABSTRACT

We present the results of HASOC-Dravidian-CodeMix shared task[1] held at FIRE 2021, a track on offensive language identification for Dravidian languages in Code-Mixed Text in this paper. This paper will detail the task, its organisation, and the submitted systems. The identification of offensive language was viewed as a classification task. For this, 16 teams participated in identifying offensive language from Tamil-English code mixed data, 11 teams for Malayalam-English code mixed data and 14 teams for Tamil data. The teams detected offensive language using various machine learning and deep learning classification models. This paper has analysed those benchmark systems to find out how well they accommodate a code-mixed scenario in Dravidian languages, focusing on Tamil and Malayalam.

CCS CONCEPTS

• **Information systems** → **Clustering and classification**; • **Computing methodologies** → *Machine learning algorithms.*

KEYWORDS

Hate speech, datasets, evaluation, deep learning

[1]https://dravidian-codemix.github.io/HASOC-2021/index.html

FIRE 2021, December 13–17, 2021, Virtual Conference, India

ACM ISBN 978-1-4503-9596-0/21/12.
https://doi.org/10.1145/3503162.3503179

ACM Reference Format:
Prasanna Kumar Kumaresan, Premjith B, Ratnasingam Sakuntharaj, Sajeetha Thavareesan, Subalalitha Chinnaudayar Navaneethakrishnan, Anand Kumar Madasamy, Bharathi Raja Chakravarthi, and John P. McCrae. 2021. Findings of Shared Task on Offensive Language Identification in Tamil and Malayalam. In *Forum for Information Retrieval Evaluation (FIRE 2021), December 13–17, 2021, Virtual Conference, India.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3503162.3503179

1 INTRODUCTION

Advancements in technology have aimed to ease peoples' lives and have attracted many users towards digitization, particularly younger generations [9, 10]. As a result, the number of people using social media to express their opinions and beliefs has increased dramatically [25]. However, the lack of regulation gives individuals the freedom to post offensive content. There is also no mechanism to regulate the posting of hateful content in under-resourced languages [1, 6, 13].

Tamil is a Dravidian language spoken primarily in Sri Lanka, India, Malaysia, and Singapore [21–23]. It is an agglutinative language with a rich morphological structure [20]. Tamil has 247 letters comprising of 12 vowels, 18 consonants, 216 composite letters combining each consonant with each vowel, and one special letter known as "Ayutha eluththu". Malayalam is also a Dravidian language spoken in Kerala, India [26–28]. Malayalam also has its own script for writing; however, social media users use Latin script or mix languages when commenting or posting online [11, 14].

The HASOC-DravidianCodeMix shared task 2021 aims to provide a new gold standard corpus for offensive language identification of code-mixed text in Dravidian languages (Tamil-English and Malayalam-English). Code-mixed content online results from

No.	TeamName	Precision	Recall	F1-Score	Rank
1	AIML [19]	0.776	0.762	0.766	1
2	MUCIC [3]	0.764	0.760	0.762	2
3	HSU [4]	0.744	0.730	0.735	3
4	IIIT Surat [7]	0.752	0.727	0.734	4
5	IRLab [24]	0.754	0.705	0.714	5
6	IIITD-ShankarB [8]	0.715	0.693	0.700	6
7	SSNCSE_NLP [2]	0.692	0.678	0.683	7
8	Pegasus [18]	0.708	0.660	0.670	8
9	CEN_NLP	0.652	0.635	0.641	9
10	MUM	0.628	0.637	0.632	10
11	JBTTM	0.577	0.584	0.580	11

Table 1: Rank list based on weighted average F1-score along with other evaluation metrics (Precision and Recall) for Task 2: Malayalam track

people mixing multiple languages, especially their native language and another commonly spoken language while expressing their views [16]. Offensive language often comprises of hate speech, such as racism, ageism, homophobia, transphobia, ableism and any hate-promoting content against an individual or group [31]. It has been an active area of research in both academia and industry for the past two decades [12]. There is an increasing demand for the identification of offensive language in code-mixed social media texts [30].

2 TASK DESCRIPTION

The task aims to identify offensive language content of the code-mixed comments/posts in Dravidian Languages (Tamil, Tamil-English and Malayalam-English) collected from social media. The comment/post may contain more than one sentence, but the average sentence length in the corpora is one. Each comment/post is annotated at the comment/post level. This dataset also exhibits class imbalance problems that mirrors real-world scenarios.

- **Task 1**
  Task 1 focuses on offensive language identification from Tamil text. Task 1 is a coarse-grained binary classification where each participating system has to classify YouTube comments in Tamil into two classes: Offensive and Not-offensive.
  - Not-Offensive – The comments does not contain offensive language. Example:
  - Offensive - The comments contain hate, offensive or profane content.
- **Task 2**
  Task 2 focus on offensive language identification in code-mixed Malayalam-English and Tamil-English comments.

3 METHODOLOGY

We received fourteen, sixteen and eleven submissions for Task 1: Tamil track, Task 2: Tamil track and Task 2: Malayalam track, respectively. The submissions were evaluated based on weighted average F1-score, and rank lists were prepared accordingly. Table 2 shows the rank list of teams that participated in Task 1: Tamil track. Tables 3 and 1 show the rank lists of the teams that competed in Task 2: Tamil track and Task 2: Malayalam track, respectively.

No.	TeamName	Precision	Recall	F1-Score	Rank
1	SSN_NLP	0.856	0.864	0.859	1
2	MUCIC [3]	0.850	0.861	0.852	2
3	SSN_NLP_MLRG [17]	0.841	0.847	0.844	3
4	IRLab [24]	0.839	0.835	0.837	4
5	BITS Pilani [29]	0.831	0.846	0.835	5
6	AIML [19]	0.823	0.843	0.825	6
7	Pegasus [18]	0.812	0.807	0.810	7
8	KonguCSE	0.749	0.797	0.764	8
9	Jusgowithurs	0.750	0.817	0.750	9
10	Gothainayaki.A	0.855	0.824	0.749	10
11	MUM	0.853	0.821	0.742	11
12	SSNCSE_NLP [2]	0.747	0.725	0.735	12
13	AI_ML NIT Patna	0.710	0.717	0.714	13
14	Saahil Raj	0.706	0.547	0.599	14

Table 2: Rank list based on weighted average F1-score along with other evaluation metrics (Precision and Recall) for Task 1: Tamil track

No.	TeamName	Precision	Recall	F1-Score	Rank
1	MUCIC [3]	0.679	0.685	0.678	1
2	AIML [19]	0.670	0.670	0.670	2
3	SSN_IT_NLP [15]	0.685	0.688	0.668	3
4	ZYBank AI	0.671	0.676	0.654	4
5	IRLab [24]	0.654	0.662	0.650	5
6	HSU [4]	0.655	0.664	0.649	6
7	IIITSurat [7]	0.679	0.673	0.636	7
8	Team Pegasus [18]	0.633	0.644	0.612	8
9	PSG [5]	0.614	0.609	0.611	9
10	SSNCSE_NLP [2]	0.615	0.607	0.610	10
11	IIITD-shanker [8]	0.599	0.568	0.573	11
12	CEN_NLP	0.596	0.540	0.539	12
13	RameshKannan	0.524	0.526	0.525	13
14	MUM	0.591	0.527	0.522	14
15	AI_ML_NIT_Patna	0.539	0.509	0.515	15
16	JBTTM	0.537	0.483	0.503	16

Table 3: Rank list based on weighted average F1-score along with other evaluation metrics (Precision and Recall) for Task 2: Tamil track

Tables 2, 3 and 1 show the precision, recall and weighted average F1-score of all the participating teams on test data. The teams used a variety of machine learning and deep learning classification models to identify offensive language.

4 CONCLUSION

This paper gives an overview of the HASOC- Dravidian-CodeMix shared task at FIRE 2021. The shared task consisted of three subtasks for Tamil, CodeMix Tamil and Malayalam languages. The purpose of this study is to examine such benchmark systems in order to determine how well they fit a code-mixed scenario in Dravidian languages, with a particular emphasis on Tamil and Malayalam.

REFERENCES

[1] Judith Jeyafreeda Andrew. 2021. JudithJeyafreedaAndrew@DravidianLangTech-EACL2021:Offensive language detection for Dravidian Code-mixed YouTube

comments. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics, Kyiv, 169–174. https://aclanthology.org/2021.dravidianlangtech-1.22

[2] Nitin Nikamanth Appiah Balaji, Bharathi B, and Bhuvana J. 2020. SSNCSE_NLP@Dravidian-CodeMix-FIRE2020: Sentiment Analysis for Dravidian Languages in Code-Mixed Text. In *FIRE (Working Notes)*.

[3] F. Balouchzahi, S. Bashang, G. Sidorov, and H. L. Shashirekha. 2021. CoMaTa OLI- Code-mixed Malayalam and Tamil Offensive Language Identification. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[4] Sai Naga Viswa Chaitanya Basava and Anjali Poornima Karri. 2021. Transformer Ensemble System for Detection of Offensive Content in Dravidian Languages. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[5] Sean Benhur J and Kanchana S. 2021. Pretrained Transformers for Offensive Language Identification in Tanglish. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[6] B Bharathi and Agnusimmaculate Silvia A. 2021. SSNCSE_NLP@DravidianLangTech-EACL2021: Offensive Language Identification on Multilingual Code Mixing Text. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics, Kyiv, 313–318. https://aclanthology.org/2021.dravidianlangtech-1.45

[7] Snehaan Bhawal, Pradeep Roy, and Abhinav Kumar. 2021. Hate Speech and Offensive Language Identification on Multilingual Code Mixed Text using BERT. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[8] Shankar Biradar, Sunil Saumya, and Arun Chauhan. 2021. mBERT based model for identification of offensive content in south Indian languages. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[9] Bharathi Raja Chakravarthi. 2020. HopeEDI: A Multilingual Hope Speech Detection Dataset for Equality, Diversity, and Inclusion. In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media*. Association for Computational Linguistics, Barcelona, Spain (Online), 41–53. https://aclanthology.org/2020.peoples-1.5

[10] Bharathi Raja Chakravarthi and Vigneshwaran Muralidaran. 2021. Findings of the Shared Task on Hope Speech Detection for Equality, Diversity, and Inclusion. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*. Association for Computational Linguistics, Kyiv, 61–72. https://aclanthology.org/2021.ltedi-1.8

[11] Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Navya Jose, Shardul Suryawanshi, Elizabeth Sherly, and John P McCrae. 2021. DravidianCodeMix: Sentiment Analysis and Offensive Language Identification Dataset for Dravidian Languages in Code-Mixed Text. *arXiv preprint arXiv:2106.09460* (2021).

[12] Bharathi Raja Chakravarthi, R. Priyadharshini, V. Muralidaran, Shardul Suryawanshi, Navya Jose, E. Sherly, and John P. McCrae. 2020. Overview of the track on Sentiment Analysis for Dravidian Languages in Code-Mixed Text. *Forum for Information Retrieval Evaluation* (2020).

[13] Bharathi Raja Chakravarthi, Ruba Priyadharshini, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kayalvizhi Sampath, Durairaj Thenmozhi, Sathiyaraj Thangasamy, Rajendran Nallathambi, and John Phillip McCrae. 2021. Dataset for Identification of Homophobia and Transophobia in Multilingual YouTube Comments. *arXiv preprint arXiv:2109.00227* (2021).

[14] Bharathi Raja Chakravarthi, KP Soman, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kingston Pal Thamburaj, John P McCrae, et al. 2021. DravidianMultiModality: A Dataset for Multi-modal Sentiment Analysis in Tamil and Malayalam. *arXiv preprint arXiv:2106.04853* (2021).

[15] S Divya and N Sripriya. 2021. Offensive Content Recognition. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[16] Navya Jose, Bharathi Raja Chakravarthi, Shardul Suryawanshi, Elizabeth Sherly, and John P McCrae. 2020. A survey of current datasets for code-switching research. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 136–141.

[17] A. Kalaivani and D. Thenmozhi. 2020. SSN_NLP_MLRG@Dravidian-CodeMix-FIRE2020: Sentiment Code-Mixed Text Classification in Tamil and Malayalam using ULMFiT. In *FIRE (Working Notes)*.

[18] Pawan Kalyan Jada, Konthala Yasaswini, Karthik Puranik, Anbukkarasi Sampath, Sathiyaraj Thangasamy, and Kingston Pal Thamburaj. 2021. Analyzing Social Media Content for Detection of Offensive Text. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[19] Jyoti Kumari and Abhinav Kumar. 2021. Offensive Language Identification on Multilingual Code Mixing Text. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[20] Nuhman. 2013. *Basic Tamil Grammar*. Readers Association, Kalmunai, Department of Tamil, University of Peradeniya.

[21] Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2016. A novel hybrid approach to detect and correct spelling in Tamil text. In *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*. IEEE, 1–6.

[22] Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2017. Use of a novel hash-table for speeding-up suggestions for misspelt Tamil words. In *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 1–5.

[23] Ratnasingam Sakuntharaj and Sinnathamby Mahesan. 2018. Detecting and correcting real-word errors in Tamil sentences. *Ruhuna Journal of Science* 9, 2 (2018).

[24] Anita Saroj and Sukomal Pal. 2020. IRLab@IIT-BHU@Dravidian-CodeMix-FIRE2020: Sentiment Analysis on Multilingual Code Mixing Text Using BERT-BASE. In *FIRE (Working Notes)*.

[25] Shardul Suryawanshi, Bharathi Raja Chakravarthi, Mihael Arcan, Suzanne Little, and Paul Buitelaar. 2021. TrollsWithOpinion: A Dataset for Predicting Domain-specific Opinion Manipulation in Troll Memes. *arXiv preprint arXiv:2109.03571* (2021).

[26] Sajeetha Thavareesan and Sinnathamby Mahesan. 2019. Sentiment Analysis in Tamil Texts: A Study on Machine Learning Techniques and Feature Representation. In *2019 14th Conference on Industrial and Information Systems (ICIIS)*. 320–325. https://doi.org/10.1109/ICIIS47346.2019.9063341

[27] Sajeetha Thavareesan and Sinnathamby Mahesan. 2020. Sentiment Lexicon Expansion using Word2vec and fastText for Sentiment Prediction in Tamil texts. In *2020 Moratuwa Engineering Research Conference (MERCon)*. 272–276. https://doi.org/10.1109/MERCon50084.2020.9185369

[28] Sajeetha Thavareesan and Sinnathamby Mahesan. 2020. Word embedding-based Part of Speech tagging in Tamil texts. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*. 478–482. https://doi.org/10.1109/ICIIS51140.2020.9342640

[29] Suchismita Tripathy, Ameya Pathak, and Yashvardhan Sharma. 2021. Offensive Language Classification of Code-Mixed Tamil with Keras. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation* (Online). CEUR.

[30] Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, and Ritesh Kumar. 2019. Predicting the Type and Target of Offensive Posts in Social Media. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 1415–1420. https://doi.org/10.18653/v1/N19-1144

[31] Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, and Ritesh Kumar. 2019. SemEval-2019 Task 6: Identifying and Categorizing Offensive Language in Social Media (OffensEval). In *Proceedings of the 13th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, Minneapolis, Minnesota, USA, 75–86. https://doi.org/10.18653/v1/S19-2010