Memotion 3: Dataset on Sentiment and Emotion Analysis of codemixed Hindi-English Memes

Shreyash Mishra*[1], S Suryavardan*[1], Parth Patwa[2], Megha Chakraborty[3], Anku Rani[3], Aishwarya Reganti[4], Aman Chadha†[5,6], Amitava Das[3], Amit Sheth[3], Manoj Chinnakotla[7], Asif Ekbal[8] and Srijan Kumar[9]

[1]*IIIT Sri City, India*
[2]*UCLA, USA*
[3]*University of South Carolina, USA*
[4]*Carnegie Mellon University, USA*
[5]*Stanford University, USA*
[6]*Amazon AI, USA*
[7]*Microsoft, USA*
[8]*IIT Patna, India*
[9]*Georgia Tech, USA*

Abstract

Memes are the new-age conveyance mechanism for humor on social media sites. Memes often include an image and some text. Memes can be used to promote disinformation or hatred, thus it is crucial to investigate in details. We introduce Memotion 3, a new dataset with 10,000 annotated memes. Unlike other prevalent datasets in the domain, including prior iterations of Memotion, Memotion 3 introduces Hindi-English Codemixed memes while prior works in the area were limited to only the English memes. We describe the Memotion task, the data collection and the dataset creation methodologies. We also provide a baseline for the task. The baseline code and dataset will be made available at https://github.com/Shreyashm16/Memotion-3.0.

Keywords

Memes, Hindi-English, Multimodality, Dataset, Machine Learning, Entailment

1. Introduction

With the rise of social media platforms as a conduit for users to communicate their thoughts and interact with one another, the amount of hate online has also parallely proliferated. The power of free uncensored speech however can cause considerable angst in the online community by demeaning other people. A popular form of producing such harmful content is the creation

*Equal contribution.
†Work does not relate to position at Amazon.

De-Factify 2: 2nd Workshop on Multimodal Fact Checking and Hate Speech Detection, co-located with AAAI 2023. 2023 Washington, DC, USA

shreyash.m19@iiits.in (S. Mishra*); suryavardan.s19@iiits.in (S. Suryavardan*); amitava@mailbox.sc.edu (A. Das)

of memes. Memes generally consist of popular images and texts associated with them that intend to spark humor among the readers. A popular definition of memes, now widely used in the field, describes them as "a group of texts with shared characteristics, with a shared core of content, form, and stance". Broadly, "content" refers to ideas and ideologies, "form" refers to our sensory experiences such as audio or visual, and "stance" refers to the tone or style, structures for participation, and communicative functions of the meme [1]. The artistic use of images and text makes the content relatable and viral. Although initially used for comic purposes only, memes have quickly evolved as a mechanism used to taunt and demean certain sections of the society. They are also used to spread misinformation and fake news. Memes are a language in themselves, with a capacity to transcend cultures and construct collective identities between people. These shareable visual jokes can also be powerful tools for self-expression, connection, social influence and even political subversion [2].

Social media platforms have many initiatives to moderate this kind of content, but memes have managed to hold their relevance despite these efforts. Detecting hate-speech and aggression on social media platforms is a popular research field, both in academia and industry, however, memes are continuously evolving and outpace contemporary hate-classification systems because (i) they can be multi-modal in nature, (ii) they might not use explicit hate content/words but more subtler forms of aggression like satire or sarcasm, and (iii) they can contain code-mixed content (languages like Hindi, Telugu, etc. written in Latin script) which is harder to parse and detect. Code-mixed content is especially prevalent in multilingual societies.

The previous iterations of Memotion each curated 10k multi-modal memes from various social media websites like Reddit, Facebook, Imgur, and Instagram, and proposed emotion and sentiment classification tasks on these datasets. In the current iteration- Memotion 3, we add an additional layer of complexity by introducing memes that are Hindi-English code-mixed. This addition ensures that models will see data that is more current and prevalent on social media and hence improve robustness. The rest of the paper is organised as follows: we describe the related work and the task in Section 2 and Section 3, respectively. Section 4 contains the details of the dataset we collected for memotion analysis: Memotion 3; followed by a brief description of baseline models 5 and their results in Section 6. We conclude with the mention of future work and limitations, in Section 7.

2. Related Work

Analysis of data to extract the sentiment and emotion has gained a lot of traction in recent years. This has been majorly focused on the large amount of data generated every second, thanks to social media. Most research in this area are focused on textual modality with some inclusion of multi-lingual data aimed at determining the polarity of the given data.

Many of the existing sentiment analysis datasets are of textual in nature and are in English [3, 4] with negative/positive or neutral categories. This also includes the works on hate speech detection in English from platforms such as Twitter [5, 6, 7] that classifies tweets based on the detected racism, sexism etc. Further works in this area shed light on multilingual or code-mixed data with inclusion of languages, such as Hindi [8], Hindi-English [9, 10, 11], Spanish-English [9, 12], Malayalam-English [13] and more [14, 15, 16, 17]. Recent approaches

to solve this problem mostly involve the use of deep learning [18, 19, 20] and large language models [21, 22, 23, 24].

However, social media is a multi-modal platform, as a result a combination of textual and visual data is vital to capture the context and analyse the data. Text-image pairs can be used for image captioning, sentiment analysis, hate speech detection and mitigate cyberbullying as shown by existing research [25, 26, 27, 28, 29]. Moreover, research has also been done towards sentiment and emotion analysis of video based data [30, 31].

One of the most commonly occurring formats of multi-modal data in social media is a meme. Although there has only been limited work, specifically towards memes, research in this area has grown in recent times. MultiOFF [32] is binary classification dataset that aims to detect whether memes are offensive or not. The Hateful memes dataset from Facebook [33] provides memes collected from USA based social media groups along with some manually reconstructed memes annotated for both uni-modal and multi-modal hate speech. The previous iterations of Memotion i.e. Memotion 1 [34] and Memotion 2[35, 36] drew attention to analysis of English memes that covered several categories, such as hatefulness, motivation, humour, sarcasm and overall sentiment. TamilMemes [37] is a also meme classification dataset that categorizes memes as being trolls or not, however this is one of the few datasets not in English. Some approaches toward this include [38, 39, 40, 41].

With Memotion 3, we present the first code-mixed Hinglish (Hindi-English) meme analysis dataset with 10k memes annotated for the aforementioned categories in the previous iterations of Memotion.

3. Memotion 3 task

A dataset of 10,000 annotated Hindi-english memes is made available. Each data point has a label for each sub-task as well as an accompanying image and text. Similar to Memotion 1 [34] and Memotion 2 [35], we consider sentiment, emotions, and their intensities. Unlike previous works, however, this iteration of the Memotion challenge focuses on Hinglish language memes. Our subtasks are as follows:

- **Task A: Sentiment Analysis** - Classify a meme as positive, negative, or neutral. Figure 1 explains the potential negative connotations of a specific meme.
- **Task B: Emotion Classification** - Classify a meme into humorous, sarcastic, offensive, or inspirational. More than one category can apply to a meme. Tasks B can be clearly understood by looking at the meme in Figure 2.
- **Task C: Scales/Intensity of Emotion Classes** - Calculate the degree to which a given emotion is being conveyed is the third task. The intensity of each emotion is shown in Figure 2.

4. Data

In this section we describe the data collection, annotation and data analysis.

Figure 1: Example for Task A. People found this meme to have a negative sentiment.

Figure 2: Example for Task B and C. Majority of annotators found this meme's humour intensity as very funny, sarcasm as twisted meaning, offensive as not offensive and motivational as not motivational. The corresponding labels for Task B will be funny, sarcastic, not offensive and not motivational.

4.1. Data Collection

We downloaded the memes after on topics of interest, such as politics, sports etc. We also collected memes using a Selenium-based web crawler. All memes are gathered from public websites Reddit and Google images. We cleaned the data to remove redundancies and performed

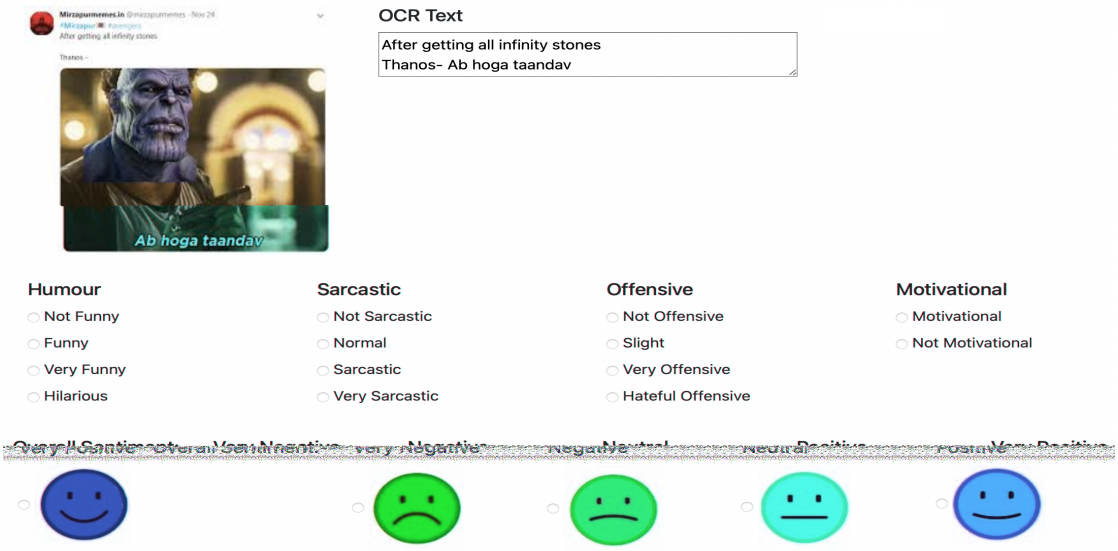
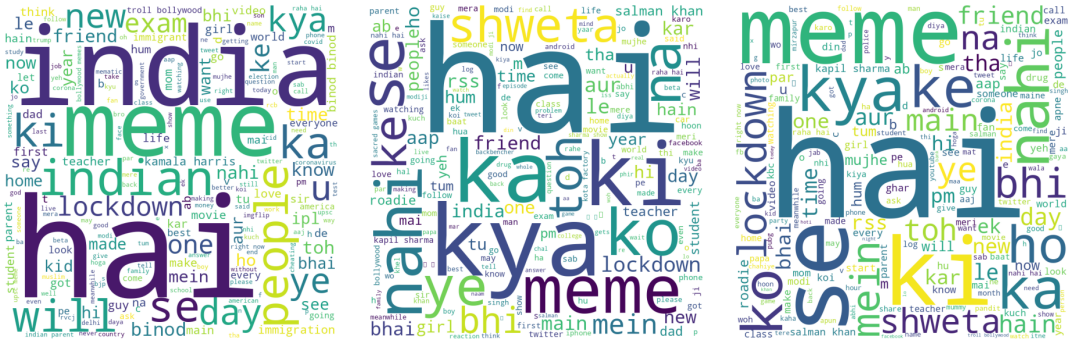


Figure 3: Annotator Interface. The annotators see a meme and have to mark the sentiment and emotion intensities of the meme. They also have to correct the OCR extracted using the Google Vision API, if there are any discrepancies.

random manual quality check. The memes are release along with the source URLs and OCR text. For OCR, we utilised the Google Vision API[1].



(a) (b) (c)

Figure 4: Word clouds indicating top words used for the (a) train, (b) validation and (c) test sets.

4.2. Data Annotation

we recruited Undergraduate student proficient in English, Hindi and meme knowledge. For annotation, they use an interface built by us, as shown in Figure 3. The annotators were asked to assess whether the meme's creator intended it to be positive, negative, or neutral in Task A.

[1]https://cloud.google.com/vision

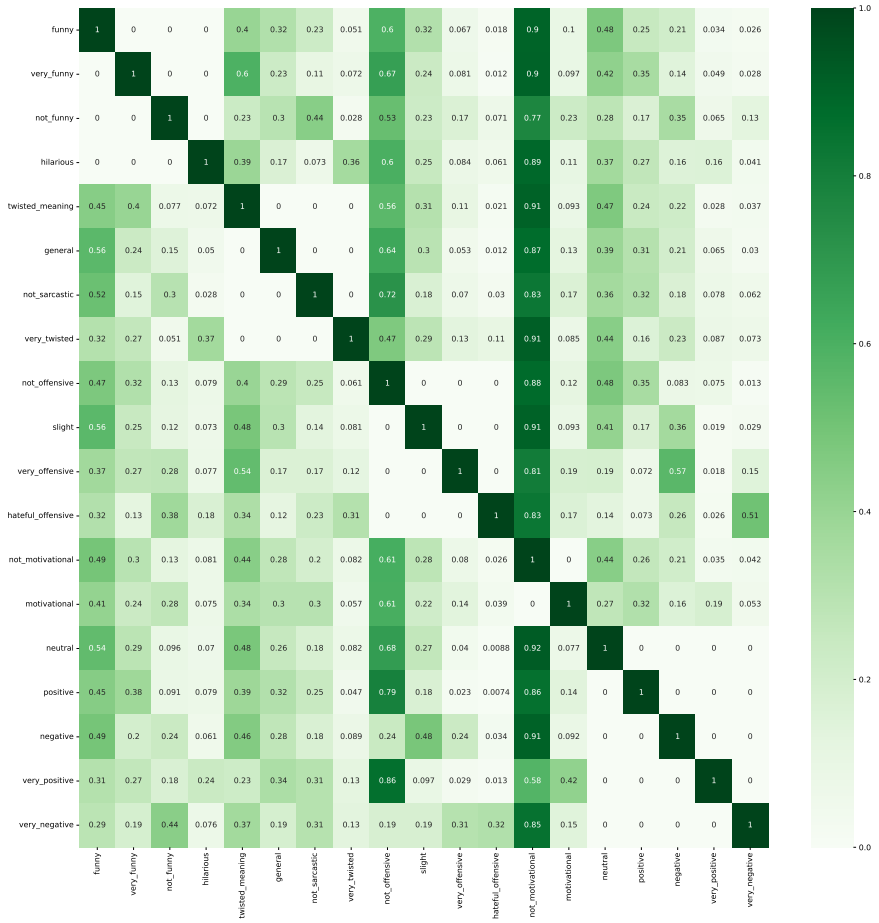


Figure 5: Overall distribution of the dataset showing overlap between all 20 labels

The annotators were asked to provide their thoughts on the emotion of the meme for Tasks B and C. The perception of a meme and societal elements may vary from person to person. Each meme is annotated by three separate annotators. The decision of the final annotations is made using a majority voting system.

4.3. Data Distribution and Analysis

The dataset consists of 10,000 memes, which are split into train, validation and test sets of size 8500,1500 and 1500 respectively. Annotations for each meme include its overall sentiment (positive, neutral, or negative), emotion (humour, sarcasm, offence, or motivation), and scale of

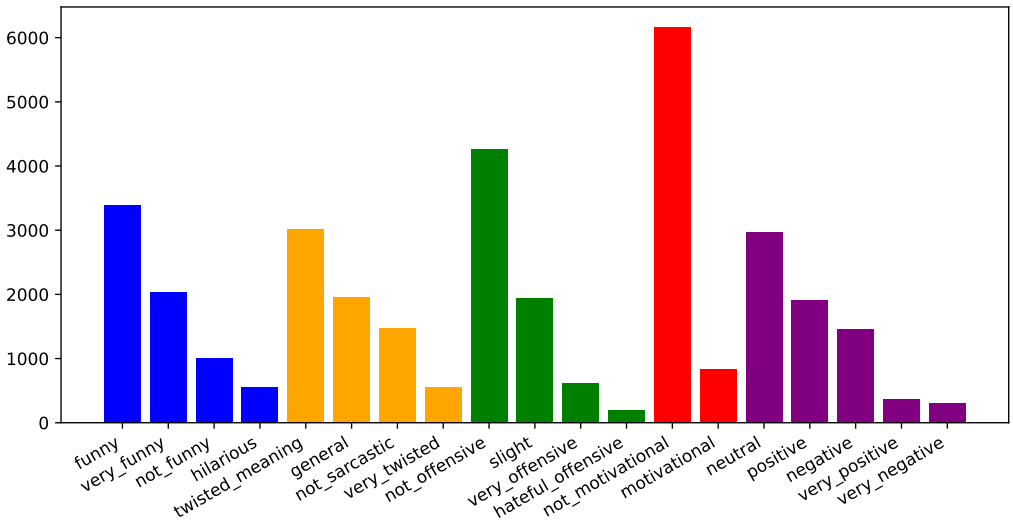


Figure 6: Distribution of the samples across all labels.

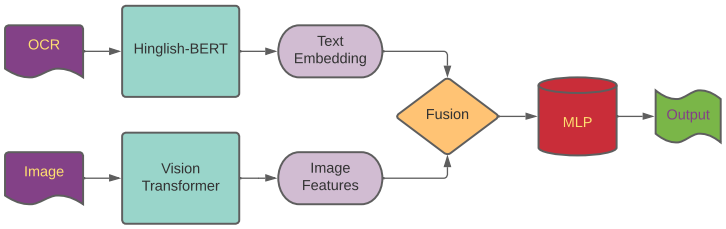


Figure 7: Baseline model architecture. It combines the text and image features for final classification. The hinglish BERT accounts for the code-mixing in data.

emotion (0-4 levels). Fig. 6 shows the distribution of memes across all the labels. Fig. 4 displays the word occurrence in the dataset. From the wordcloud we can see that lot of code mixed words like *nahi*, *kya* are prominent in the dataset.

from the statistical features in Fig. 5, we can conclude that the emotions in memes overlap, demonstrating the difficulty of the tasks. A number of intriguing facts are revealed, including the fact that many offending memes are humorous. Additionally, a lot of the memes are humorous and lack inspiration, as can be seen. On average, the Code Mixed Index (CMI) [42] for the training, validation and test set is 14.94, 20.19 and 20.06 respectively.

5. Baseline model

The importance of considering both the visual and textual features is vital for multi-modal data, especially in the case of memes where the context can only be captured using a combination of

both components. Attention models are exceptional at representing text with respect to context and a widely known model with strong performance is BERT [43]. As the dataset is not in English but instead in Hindi-English, we use a multilingual variant of BERT, specifically Hinglish-BERT from Verloop [44], with both the backbone and linear layers of the LM finetuned. The model is implemented using *BERT-base-multilingual-cased*, which is fine-tuned on Hinglish data. The visual features are obtained from the pre-trained Vision transformer model (ViT) [45]. The ViT model can outperform normal CNNs computationally and by accuracy, thanks to the positional embedding of image patches done by ViT. The pooled output from the ViT model is concatenated with the Hinglish-BERT embedding. The combined features are then classified after being passed through a MLP. With changes to the MLP, the multi-modal features are used for all three sub-tasks. The model architecture is displayed in Figure 7. The results for each task are provided in Table 1. The codes will be made available at https://github.com/Shreyashm16/Memotion-3.0.

6. Results

Baseline results in Table 1 show Weighted F1 scores for each task and sub-task. Using ViT for extracting visual features and Hinglish-BERT for the textual features, the baseline models scores 33.28% for Task A, 74.74% for Task B and 52.27% for Task C.

This dataset will be made public, and we leave it to future research to develop more sophisticated systems that go deeper into Memotion Analysis.

Task	Class	Weighted F1 score
Task-A	Sentiment	33.28%
Task-B	Humour	84.55%
	Sarcasm	74.82%
	Offensive	48.84%
	Motivation	90.78%
	Average	74.74%
Task-C	Humour	43.03%
	Sarcasm	32.89%
	Offensive	42.40%
	Motivation	90.78%
	Average	52.27%

Table 1
Baseline scores (Weighted F1) of the baseline model on Memotion Analysis tasks.

7. Conclusion and Future Work

In this study, we present a hindi-english dataset for the challenge of sentiment in a multimodal environment. This is the first significant multimodal dataset for Hindi code-mixed meme categorization that we are aware of. We provide annotated data for three tasks, namely sentiment analysis, emotion classification, and strength of emotion, in order to provide a fine-grained and thorough analysis of memes. By combining the image features extracted from the ViT model

and the textual features using Hinglish-BERT, and then passing these joint embeddings to a simple MLP, we design the baseline for the tasks. It should be mentioned that our models are preliminary and that more creative approaches will enhance performance much more. In the future, we intend to extend our work by designing a single model for all languages, instead of creating separate models for memes of different languages. We could also work on generating memes for the task, instead of collection to customize the dataset.

References

[1] L. Shifman, Memes in a Digital World: Reconciling with a Conceptual Troublemaker, Journal of Computer-Mediated Communication 18 (2013) 362–377. URL: https://doi.org/10.1111/jcc4.12013. doi:10.1111/jcc4.12013. arXiv:https://academic.oup.com/jcmc/article-pdf/18/3/362/19492245/jjcmcom0362.pdf.

[2] N. Akhther, Internet memes as form of cultural discourse: A rhetorical analysis on facebook, 2018. doi:10.31234/osf.io/sx6t7.

[3] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, C. Potts, Learning word vectors for sentiment analysis, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Portland, Oregon, USA, 2011, pp. 142–150. URL: http://www.aclweb.org/anthology/P11-1015.

[4] A. Go, R. Bhayani, L. Huang, Twitter sentiment classification using distant supervision, Processing 150 (2009).

[5] Z. Waseem, D. Hovy, Hateful symbols or hateful people? predictive features for hate speech detection on twitter, 2016, pp. 88–93. doi:10.18653/v1/N16-2013.

[6] Z. Waseem, Are you a racist or am i seeing things? annotator influence on hate speech detection on twitter, in: NLP+CSS@EMNLP, 2016.

[7] P. Burnap, M. Williams, Us and them: identifying cyber hate on twitter across multiple protected characteristics, EPJ Data Science 5 (2016). doi:10.1140/epjds/s13688-016-0072-6.

[8] P. Patwa, M. Bhardwaj, V. Guptha, G. Kumari, S. Sharma, S. PYKL, A. Das, A. Ekbal, S. Akhtar, T. Chakraborty, Overview of constraint 2021 shared tasks: Detecting english covid-19 fake news and hindi hostile posts, in: Proceedings of the First Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation (CONSTRAINT), Springer, 2021.

[9] P. Patwa, G. Aguilar, S. Kar, S. Pandey, S. PYKL, B. Gambäck, T. Chakraborty, T. Solorio, A. Das, SemEval-2020 task 9: Overview of sentiment analysis of code-mixed tweets, in: Proceedings of the Fourteenth Workshop on Semantic Evaluation, International Committee for Computational Linguistics, Barcelona (online), 2020. URL: https://aclanthology.org/2020.semeval-1.100.

[10] A. Joshi, A. Prabhu, M. Shrivastava, V. Varma, Towards sub-word level compositions for sentiment analysis of Hindi-English code mixed text, in: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, The COLING 2016 Organizing Committee, Osaka, Japan, 2016, pp. 2482–2491. URL: https://aclanthology.org/C16-1234.

[11] R. Kumar, A. N. Reganti, A. Bhatia, T. Maheshwari, Aggression-annotated Corpus of Hindi-English Code-mixed Data, in: N. C. C. chair), K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, T. Tokunaga (Eds.), Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), European Language Resources Association (ELRA), Miyazaki, Japan, 2018.

[12] V. Basile, C. Bosco, E. Fersini, D. Nozza, V. Patti, F. M. Rangel Pardo, P. Rosso, M. Sanguinetti, SemEval-2019 task 5: Multilingual detection of hate speech against immigrants and women in Twitter, in: Proceedings of the 13th International Workshop on Semantic Evaluation, Association for Computational Linguistics, Minneapolis, Minnesota, USA, 2019, pp. 54–63. URL: https://aclanthology.org/S19-2007. doi:10.18653/v1/S19-2007.

[13] B. R. Chakravarthi, N. Jose, S. Suryawanshi, E. Sherly, J. P. McCrae, A sentiment analysis dataset for code-mixed malayalam-english, CoRR abs/2006.00210 (2020). URL: https://arxiv.org/abs/2006.00210. arXiv:2006.00210.

[14] S. Bhattacharya, S. Singh, R. Kumar, A. Bansal, A. Bhagat, Y. Dawer, B. Lahiri, A. K. Ojha, Developing a multilingual annotated corpus of misogyny and aggression, in: Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying, European Language Resources Association (ELRA), Marseille, France, 2020, pp. 158–168. URL: https://www.aclweb.org/anthology/2020.trac2-1.25.

[15] B. R. Chakravarthi, V. Muralidaran, R. Priyadharshini, J. P. McCrae, Corpus creation for sentiment analysis in code-mixed tamil-english text, CoRR abs/2006.00206 (2020). URL: https://arxiv.org/abs/2006.00206. arXiv:2006.00206.

[16] T. Mandl, S. Modha, A. Kumar M, B. R. Chakravarthi, Overview of the hasoc track at fire 2020: Hate speech and offensive language identification in tamil, malayalam, hindi, english and german, in: Proceedings of the 12th Annual Meeting of the Forum for Information Retrieval Evaluation, FIRE '20, Association for Computing Machinery, New York, NY, USA, 2021, p. 29–32. URL: https://doi.org/10.1145/3441501.3441517. doi:10.1145/3441501.3441517.

[17] B. R. Chakravarthi, R. Priyadharshini, V. Muralidaran, S. Suryawanshi, N. Jose, E. Sherly, J. P. McCrae, Overview of the track on sentiment analysis for dravidian languages in code-mixed text, in: Proceedings of the 12th Annual Meeting of the Forum for Information Retrieval Evaluation, FIRE '20, Association for Computing Machinery, New York, NY, USA, 2021, p. 21–24. URL: https://doi.org/10.1145/3441501.3441515. doi:10.1145/3441501.3441515.

[18] P. Patwa, S. Pykl, A. Das, P. Mukherjee, V. Pulabaigari, Hater-o-genius aggression classification using capsule networks, in: Proceedings of the 17th International Conference on Natural Language Processing (ICON), 2020, pp. 149–154.

[19] S. T. Aroyehun, A. Gelbukh, Aggression detection in social media: Using deep neural networks, data augmentation, and pseudo labeling, in: Proceedings of the First Workshop on Trolling, Aggression and Cyberbullying (TRAC-2018), 2018, pp. 90–97.

[20] A. Ribeiro, N. Silva, INF-HatEval at SemEval-2019 task 5: Convolutional neural networks for hate speech detection against women and immigrants on Twitter, in: Proceedings of the 13th International Workshop on Semantic Evaluation, Association for Computational Linguistics, Minneapolis, Minnesota, USA, 2019, pp. 420–425. URL: https://aclanthology.org/S19-2074. doi:10.18653/v1/S19-2074.

[21] D. Tula, P. Potluri, S. Ms, S. Doddapaneni, P. Sahu, R. Sukumaran, P. Patwa, Bitions@DravidianLangTech-EACL2021: Ensemble of multilingual language models with pseudo labeling for offence detection in Dravidian languages, in: Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages, Association for Computational Linguistics, Kyiv, 2021, pp. 291–299. URL: https://aclanthology.org/2021.dravidianlangtech-1.42.

[22] M. Mozafari, R. Farahbakhsh, N. Crespi, A bert-based transfer learning approach for hate speech detection in online social media, in: Complex Networks and Their Applications VIII: Volume 1 Proceedings of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019 8, Springer, 2020, pp. 928–940.

[23] D. Tula, M. Shreyas, V. Reddy, P. Sahu, S. Doddapaneni, P. Potluri, R. Sukumaran, P. Patwa, Offence detection in dravidian languages using code-mixing index-based focal loss, SN Computer Science 3 (2022) 330.

[24] R. Ali, U. Farooq, U. Arshad, W. Shahzad, M. O. Beg, Hate speech detection on twitter using transfer learning, Computer Speech & Language 74 (2022) 101365.

[25] R. Jha, V. Kaki, V. Kolla, S. Bhagat, P. Patwa, A. Das, S. Pal, Image2tweet: Datasets in Hindi and English for generating tweets from images, in: Proceedings of the 18th International Conference on Natural Language Processing (ICON), NLP Association of India (NLPAI), National Institute of Technology Silchar, Silchar, India, 2021, pp. 670–676. URL: https://aclanthology.org/2021.icon-main.84.

[26] A. Hu, S. Flaxman, Multimodal sentiment analysis to explore the structure of emotions, Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery |& Data Mining (2018). URL: http://dx.doi.org/10.1145/3219819.3219853. doi:10.1145/3219819.3219853.

[27] R. Gomez, J. Gibert, L. Gomez, D. Karatzas, Exploring hate speech detection in multimodal publications, 2019. arXiv:1910.03814.

[28] H. Zhong, H. Li, A. Squicciarini, S. Rajtmajer, C. Griffin, D. Miller, C. Caragea, Content-driven detection of cyberbullying on the instagram social network, 2016.

[29] H. Hosseinmardi, S. A. Mattson, R. I. Rafiq, R. Han, Q. Lv, S. Mishra, Detection of cyberbullying incidents on the instagram social network, 2015. arXiv:1503.03909.

[30] L.-P. Morency, R. Mihalcea, P. Doshi, Towards Multimodal Sentiment Analysis: Harvesting Opinions from The Web, in: International Conference on Multimodal Interfaces (ICMI 2011), Alicante, Spain, 2011. URL: http://ict.usc.edu/pubs/Towards%20Multimodal%20Sentiment%20Analysis-%20Harvesting%20Opinions%20from%20The%20Web.pdf.

[31] A. Bagher Zadeh, P. P. Liang, S. Poria, E. Cambria, L.-P. Morency, Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 2236–2246. URL: https://aclanthology.org/P18-1208. doi:10.18653/v1/P18-1208.

[32] S. Suryawanshi, B. R. Chakravarthi, M. Arcan, P. Buitelaar, Multimodal meme dataset (MultiOFF) for identifying offensive content in image and text, in: Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying, European Language Resources Association (ELRA), Marseille, France, 2020, pp. 32–41. URL: https://aclanthology.org/2020.trac-1.6.

[33] D. Kiela, H. Firooz, A. Mohan, V. Goswami, A. Singh, P. Ringshia, D. Testuggine, The hateful memes challenge: Detecting hate speech in multimodal memes, 2021. arXiv:2005.04790.

[34] C. Sharma, D. Bhageria, W. Paka, Scott, S. P Y K L, A. Das, T. Chakraborty, V. Pulabaigari, B. Gambäck, SemEval-2020 Task 8: Memotion Analysis-The Visuo-Lingual Metaphor!, in: Proceedings of the 14th International Workshop on Semantic Evaluation (SemEval-2020), Association for Computational Linguistics, Barcelona, Spain, 2020.

[35] S. Ramamoorthy, N. Gunti, S. Mishra, S. Suryavardan, A. Reganti, P. Patwa, A. Das, T. Chakraborty, A. Sheth, A. Ekbal, C. Ahuja, Memotion 2: Dataset on sentiment and emotion analysis of memes, in: Proceedings of Defactify: First workshop on multimodal Fact CHecking and Hatespeech detection, 2022.
[36] P. Patwa, S. Ramamoorthy, N. Gunti, S. Mishra, S. Suryavardan, A. Reganti, A. Das, T. Chakraborty, A. Sheth, A. Ekbal, C. Ahuja, Findings of memotion 2: Sentiment and emotion analysis of memes, in: Proceedings of Defactify: First workshop on multimodal Fact CHecking and Hatespeech detection, 2022.
[37] S. Suryawanshi, B. R. Chakravarthi, P. Verma, M. Arcan, J. P. McCrae, P. Buitelaar, A dataset for troll classification of Tamil memes, in: Proceedings of the 5th Workshop on Indian Language Data Resource and Evaluation (WILDRE-5), European Language Resources Association (ELRA), Marseille, France, 2020.
[38] S. Pramanick, S. Sharma, D. Dimitrov, M. S. Akhtar, P. Nakov, T. Chakraborty, Momenta: A multimodal framework for detecting harmful memes and their targets, arXiv preprint arXiv:2109.05184 (2021).
[39] A.-M. Bucur, A. Cosma, I.-B. Iordache, Blue at memotion 2.0 2022: You have my image, my text and my transformer, arXiv preprint arXiv:2202.07543 (2022).
[40] N. Gunti, S. Ramamoorthy, P. Patwa, A. Das, Memotion analysis through the lens of joint embedding (student abstract), in: Proceedings of the AAAI Conference on Artificial Intelligence, 2022.
[41] G. G. Lee, M. Shen, Amazon pars at memotion 2.0 2022: Multi-modal multi-task learning for memotion 2.0 challenge, Proceedings http://ceur-ws. org ISSN 1613 (2020) 0073.
[42] B. Gambäck, A. Das, On measuring the complexity of code-mixing, in: Proceedings of the 11th international conference on natural language processing, Goa, India, 2014, pp. 1–7.
[43] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
[44] M. Bhange, N. Kasliwal, Hinglishnlp: Fine-tuned language models for hinglish sentiment detection, arXiv preprint arXiv:2008.09820 (2020).
[45] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929 (2020).