

Stance Detection in Hinglish Data using the BART-large-MNLI Integration Model

Somasekhar Giddaluru

Department of Computer Science and Engineering, GITAM (Deemed to be University), Hyderabad, India
giddalurisomasekhar@gmail.com (corresponding author)

Sreerama Murty Maturi

Department of Computer Science and Engineering, GITAM (Deemed to be University), Hyderabad, India
sreeramssit@gmail.com

Obulesu Ooruchintala

Department of CSE (Data Science), G. Narayanamma Institute of Technology & Science, Hyderabad, India
obulesh194@gmail.com

Mahendra Munirathnam

Department of CSE (Data Science), G. Narayanamma Institute of Technology & Science, Hyderabad, India
mahendrame405@gmail.com

Received: 4 May 2024 | Revised: 16 May 2024 | Accepted: 21 May 2024

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.7741>

ABSTRACT

Real-time stance detection can be used in a wide range of applications like debate analysis, sentiment analysis, system feedback, etc. This study focuses on stance detection in political speeches, discerning whether the speaker is in favor, against, neutral, or lacking any stance on a given topic. The problem with this type of speeches dwells in the modification of the existing methods of stance detection to allow for the fine distinctions of Hinglish, a language mixture blending Hindi and English, as conveyed in human-edited texts. The proposed method utilizes the Bidirectional Auto-Regressive Transformers Multi-Genre Natural Language Inference (BART-large-MNLI) model with zero-Shot, few-Shot and N-Shot learning approaches. The proposed model is compared with the existing models of stance detection on Hinglish texts. For the pre-trained BART-based models, a limited number of labeled examples are utilized to determine the labels of test instances. For the other models, the train-test split method is adopted to get accurate results. The results indicate that the model surpasses the previous models.

Keyword-stance detection; analysis; BART; Hinglish

I. INTRODUCTION

Stance detection is a substantial issue to handle in the fields of Machine Learning (ML), Artificial Intelligence (AI) and Natural Language Processing (NLP). A stance recognition method can be applied to get community sentiments and viewpoints by scanning social media platforms addressed to diverse subjects, such as movies, sports, and finance. The former can be utilized to grasp the impression of common people, recognizing trends and action plans in marketing or public relations. It can be also employed for the appraisal of communal policies, political discourses, and deliberations to

comprehend the general stance and feeling towards anticipated initiatives or statutory variations, assisting therefore legislators, representatives and officials in generating knowledgeable decisions and determining policy guidelines that agree with civic judgment and opinions.

This paper emphasizes the improvement of stance detection approaches for Hindi speech articulated in English over several subjects. This involves the utilization of vigorous models to catch and categorize nuanced language patterns in dissimilar contexts. Present stance discovery methods have limits when employed for Hinglish text analysis, since they may not work while capturing the complexities existing in Hinglish speech

patterns. The same applies for the prevailing mechanisms, which may lack bendiness to the different language expressions and vocabularies. The unavailability of the categorized datasets for stance discovery through divergent fields generates a huge challenge for model preparation and performance, probably causing classification problems and inexact outcomes.

An expert model for immediate stance detection, which precisely labels political discourses conversed in Hinglish, is proposed. It exhibits a simultaneous monitoring scheme or platform for uninterrupted analysis.

Zero-shot learning was boosted by the explicit attributes of learning and optimization in [1]. Certain meta-learning techniques were used to strengthen the Few-Shot learning in [2]. A new library called LibFewShot was generated to improve the performance of Few-Shot learning in [3]. A novel generative mixup of networks with semantic graph alignment was engaged to enhance Zero-Shot learning in [4]. A task-adaptive classifier predictor was proposed to boost Few-Shot learning in [5]. The N-Shot learning performance was ameliorated via a new auto-augmentation method in [6]. A Deep Learning (DL) structure promised to recognize stances in code-mixed social media records, highlighting Multi-Task Learning (MTL) was elucidated in [7]. The usefulness of CNN-BiLSTM in apprehending miscellaneous patterns of feelings in code-mixed social media transcript was debated in [8]. Few-shot learning built on sentiments for pre-training of stance discovery was recommended in [9]. A framework based on conditional generations for stance discovery by the implementation of pre-demarcated patterns was provided and a supplementary job termed joint target prediction was accomplished in [10]. Several practices and complications in zero-shot stance discovery were pondered in [11]. A self-supervised learning method consuming data without labels to identify stances was recommended in [12]. Hindi-English code-mixed memes were utilized to differentiate feelings and views in [13]. A DL process called hybrid fine-tuned Smith algorithm combined with Adam optimizer to catch intuitions from multi-lingual social media text was introduced in [14]. An innovative technique using gated Multilayer Perceptron (gMLP) and a prompt learning method appropriate for zero-shot stance discovery was established in [15]. Bilingual machine translation was discussed in [16]. GPT and LLaMA-2 models were analyzed in [17]. A review on machine translation was conducted in [18].

II. THE PROPOSED MODEL

The BART-large-MNLI model is assimilated with zero-Shot, few-Shot and N-Shot learning approaches to discover stances towards Hinglish speech covering diverse subjects. A Hinglish dataset and an IIIT code-mixed dataset for model participation and appraisal were considered, warranting compatibility and orientation of stance labels. The trained learning method was followed to envisage stances on Hinglish data, using its discernment of language semantics and context. The performance of the model was gauged by associating the anticipated stances with ground truth labels from the IIIT code-mixed dataset, computing quality measures. The model deals with an exact and yielding solution to uncover stances through diverse domains in Hinglish data. It was adjusted for proficient

implication and response times, lessening computational resources. The assessment pipeline was modernized to enable fast reiteration and model enhancement without depending merely on IIIT code-mixed dataset. The assimilation with few-Shot and N-Shot learning allows the fruitful generalization even in situations with inadequate categorized data. By using a smaller number of categorized instances from every class, the model is able to adjust to fresh themes, increasing its performance and generalization competences.

BART combines elements from both auto-regressive models like GPT and bidirectional models like BERT to achieve state-of-the-art performance in various NLP tasks. It consists of two main components: the encoder and the decoder, each composed of multiple layers of transformer blocks. The encoder processes the input sequence, generating hidden representations capturing contextual information. Each encoder layer typically consists of 12 transformer blocks, each with 16 attention heads and a feed forward neural network with a hidden layer with size of 4096. The hidden layer size of the encoder output is typically 1024. Figure 1 portrays the BART architecture. Figure 2 depicts noise transformations. Figure 3 illustrates the comparison between BERT, GPT and BART.

The decoder generates the output sequence based on the representations learned by the encoder. Similarly to the encoder, each decoder layer comprises 12 transformer blocks. The decoder incorporates masked self-attention and cross-attention mechanisms. The hidden size of the decoder is also typically set to 1024.

Figure 4 demonstrates BART’s variants. Figures 5, 6, and 7 show the Zero-Shot, Few-Shot and N-Shot learning architectures, respectively.

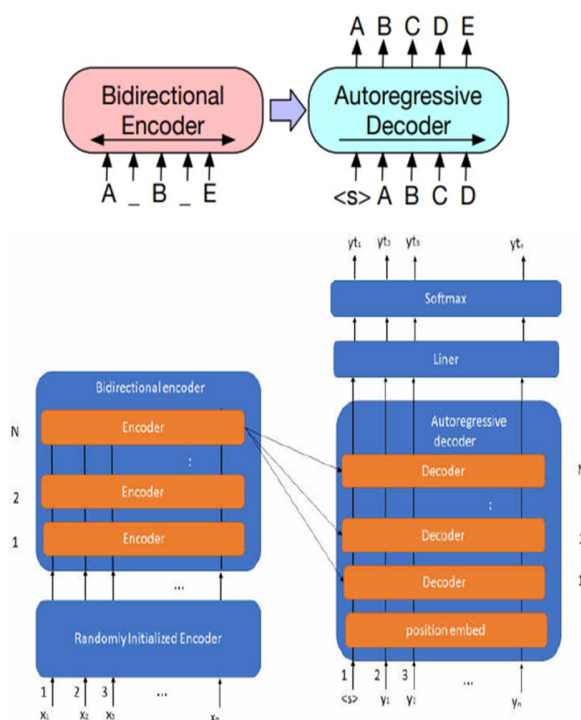


Fig. 1. Architecture of BART.

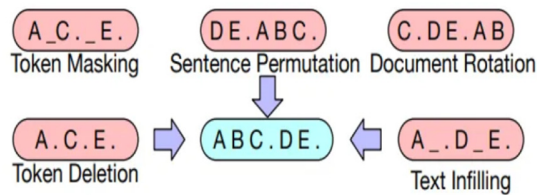


Fig. 2. Noise transformations.

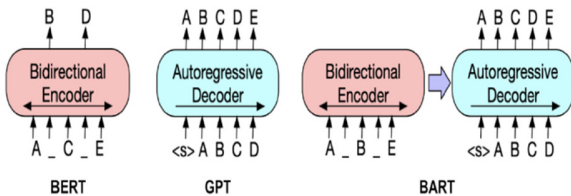


Fig. 3. Comparison between BERT, GPT and BART.

Model	Description	# params
bart.base	BART model with 6 encoder and decoder layers	140M
bart.large	BART model with 12 encoder and decoder layers	400M
bart.large.mnli	bart.large finetuned on MNLI	400M
bart.large.cnn	bart.large finetuned on CNN-DM	400M
bart.large.xsum	bart.large finetuned on Xsum	400M

Fig. 4. BART's variants.

Transfer Learning vs Zero-Shot Learning

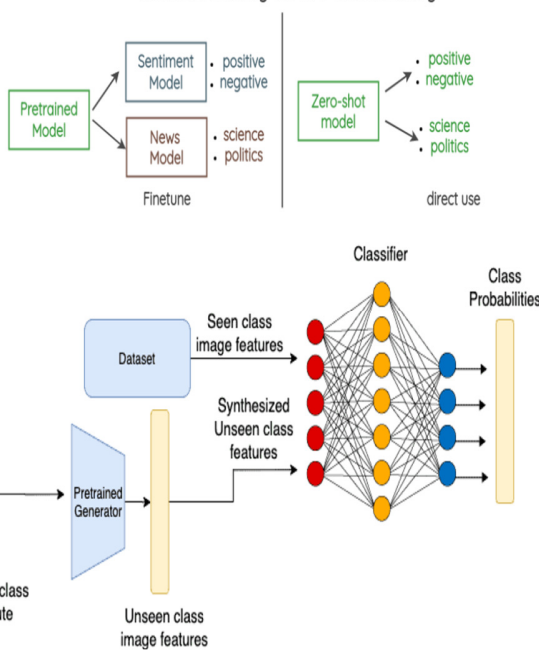


Fig. 5. Zero-Shot architecture.

BART-large-MNLI is a variant of the BART model finetuned specifically for the MNLI task. It has been fine-tuned on the MNLI dataset, comprising approximately 393,000 training examples across various genres. During fine-tuning, model parameters are adjusted to optimize performance specifically

for the natural language inference task defined by MNLI. It is a "large" variant, usually consisting of nearly 406 million features. Its large size permits it to catch complex language forms and associations. Owing to the wide-ranged pre-training and refinement, it can be utilized in many NLP tasks. In addition to natural language inference, it exhibits up-to-date performance in data recapitulation, question answering, and interpretation.

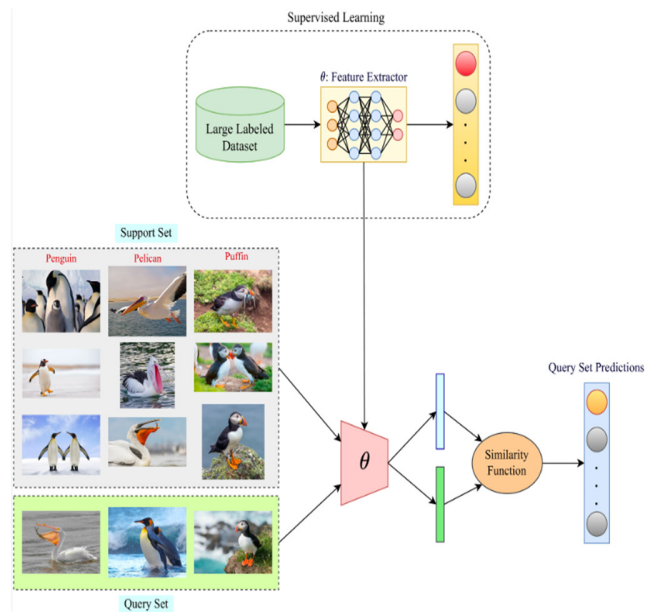


Fig. 6. Few-Shot architecture.

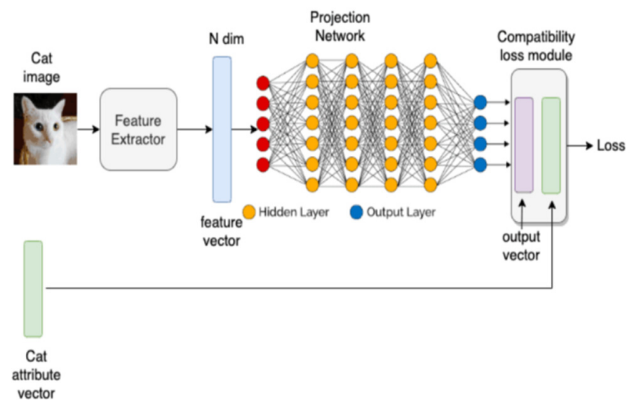


Fig. 7. N-Shot architecture.

N-Shot learning is an ML method intended for improvement of model performance using a lesser number of labeled instances (N) from every class through training. This approach allows the successful generalization through extensive datasets, even in circumstances with inadequate labeled records. The design of N-shot learning incorporates numerous key constituents that expedite its efficiency to improve the generalization and performance of the model.

A. N-Shot Learning Key Constituents

- **Support Set:** It comprises a lower number of labeled instances (N) from every class, acting as the principal training records for the model. These labeled samples give the necessary material to the model to acquire the properties and structures related to every class.
- **Query Set:** It contains supplementary samples, usually unlabeled, which can be employed by the model for interpretation and assessment. These samples are not utilized throughout training but are decisive to evaluate the performance and generalization competence of the model.
- **N-Shot Learning Algorithm:** It is liable to train the model by the support set and adjusting its factors to improve performance on a particular task. It adjusts the parameters of the model deploying the existing labeled samples, which empowers it for competent generalization through miscellaneous jobs and fields.
- **Feature Extraction:** It is engaged to excerpt appropriate patterns from the inputs, permitting the model to catch necessary forms and features related to every class. These structures act as the foundation for model classification and management.
- **Classifier:** It is accountable for allocating class labels to the input instances using the learned characteristics and features of the model. It utilizes the mined patterns to create learned predictions and categorize samples into classes.

B. Emphasis on Zero-Shot Learning

The major characteristic of N-shot learning is its capability to generalize outside the classes realized throughout training, which is termed as zero-Shot learning. The latter allows models to distinguish and categorize classes which have never met before, reflecting the human talent to generalize and recognize novel perceptions deprived of clear supervision. With this capability, the model becomes more flexible and handy. With nominal supervision, it can be useful for multiple purposes and tasks.

C. Uses of Few-Shot Learning

The main objective of Few-Shot learning is to decrease the model's dependence on labeled records and allow for effective generalization through numerous jobs and fields. It is operative when the attainment of large size labeled data sets is expensive and impractical.

III. RESULTS AND DISCUSSION

A. Dataset

The International Institute of Information Technology (IIITH) Code Mixed dataset, is a treasured source for examining Hinglish code-mixed data. Hinglish is a crossbreed language that blends features of Hindi and English, frequently used in casual speaking, mainly on social media. The dataset comprises tweets recorded in Hinglish. It facilitates researchers with a quality resource of language text as they can get clear perceptions on linguistic dissimilarity, opinion mining, and trends.

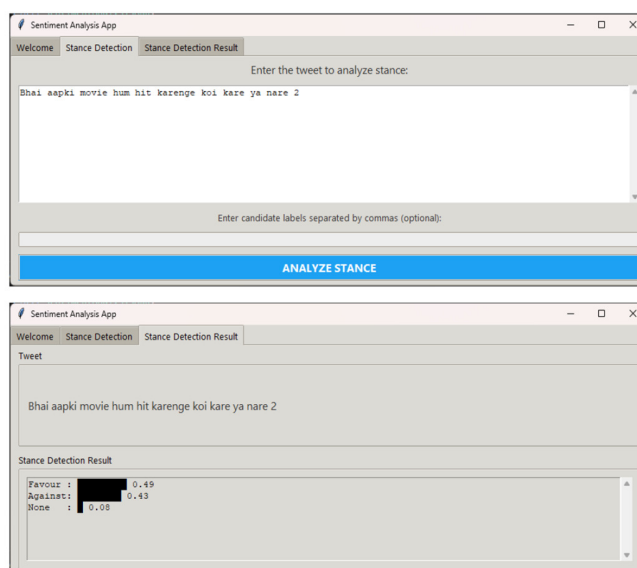


Fig. 8. Sample stance detection result for a tweet in Hinglish

3	Love u sir love u soo much urs l'ts beautiful vedio ...	2	2
4	Arae sur jee pahelae hamare bharat ke bachho ka parao.... Firr vedesh may jao yaar	1	1
5	Wahl! Jitni sundar geet ke bhao hain utnihi sundar aur sureeli aawaz hai .	2	2
8	Sundar ekdam sahi Gaya Hua gana.chhotisi gudiyi ko naman	2	2
9	Wao lata manglekar	2	2
10	Aatma aur mun ki pavitrata ki jhalak hai is beti ki aawaz mei, khush raho beti.	2	2
11	Joy hind joy hindMeri bhatan tuje salamModi ji aapko salam	2	2
12	Kash Modiji Aap Manmohan Singh se phly aa gy hoty....aj 2015 humara Desh kha phuch gya hota	2	2
13	India me student pe atyachaar karra hai result nahi dikhaijara hai aur yaha ye inke aankho me dhul jhokra hai	0	0
37	bharat ko vishwa guru banne se koi nahi rok sakta.....vande matram	2	2
38	Sir yeh tho sirf aap hi kar sakte hai. Great sir	2	2
39	Modi jigood preminister hia	2	2
40	Jai sri ram nama namo	1	1
41	Jan Jan ka viswas kam ho na jai hind	1	1
42	modi ji aage badho pura desh aapke ssath hai.	2	2
43	RAJ KAPOOR KI FILMO KE BAARE ME SUNA THA KI USSR ME PASAND KI JAATI HALLAGTA HAI VAH AAJ BHI KAYAM HAY.	2	2
45	we are proud our pm ; ab lg rha h ki hmara desh ek achchi rah pr chal para hai jaha se hm puri world me apna jhanda gaar skte hai	2	2
46	modi ji aap ki soch bahut sahi lgti hai progres n hona hi sabki wajah hai is liye sabse zaroori hai ki desh ka vikas ho	2	2

Fig. 9. Samples of Hinglish code-mixed tweets in the dataset.

B. Results

Python 3.12.0 and its important ML and DL libraries were deployed for the experiments. The proposed model is compared with all the known methods employed for stance discovery on Hinglish texts [7, 8, 10] (Table I). The results (Table I) demonstrate that the model surpasses the other approaches.

The proposed model gives 8.5% error rate. Due to the scarcely available limited size of Hinglish data sets, it fails to capture the semantics of some Hinglish text to label few instances. In future, an attempt to collect or develop large size Hinglish data sets will be made so that the model can be perfectly trained to further reduce the error rate.

BART-based conditional generation model [10] has certain limitations if it is to be used for social media text. CNN+MTL model [7] has not been tested on high quality data sets. A large corpus is required for other existing models (Table I) to perform well. However, the proposed model overcomes the the existing models. The former can be also used for multi-class

text classification, code summarization, emotion recognition, and text summarization, while it can be tested on different language mixtures depending on the input data set.

Figure 8 exhibits a sample stance detection result for a tweet in Hinglish. Figure 9 presents sample Hinglish code-mixed tweets in the data set.

TABLE I. ACCURACY RESULTS

Model	Zero-Shot	Few-Shot	N-Shot
CNN + MTL	72.1	73.4	74.6
LSTM	73.2	73.8	74.2
Bi-LSTM	75.6	75.9	76.3
CNN-LSTM	77.4	77.8	78.5
CNN-Bi-LSTM	78.9	79.4	80.7
BERT-GCN	80.3	81.1	81.8
CKE-Net	81.6	81.9	82.5
WS-BERT	82.3	82.7	83.4
BART-based conditional generation model	83.1	83.5	84.1
BART-large-MNLI integration model	83.4	87.3	91.5

The proposed model was effectively assimilated into both an application and a webpage, permitting consumers to input their opinions simultaneously and obtain precise classification outcomes. The particular model offers a comprehensible interface for customers to express their thoughts in Hinglish, contributing to a profound understanding of opinion mining within this language. As research studies on Hinglish text are very limited in number, stance discovery in this area should be encouraged by government bodies.

IV. CONCLUSION

Further research and development attempts could concentrate on taming the model's understanding of Hinglish language tones and enhancing accuracy. Moreover, escalating the dataset and including further innovative methods, such as refinement of pre-trained models, could lead to even better performance and extensive usage of the model introduced. As a whole, the current work sets a concrete basis for the upcoming developments in stance discovery within Hinglish data and proves the capable strength of N-shot learning in natural language processing tasks. In future, massive high quality multi-lingual data sets could be collected and used for a wider application of the model. The model's accuracy could be improved further by deploying various methods to boost the performance of Zero-Shot [1, 4], Few-Shot [2, 3, 5] and N-Shot [6] learning techniques.

REFERENCES

[1] Y. Du, M. Shi, F. Wei, and G. Li, "Boosting Zero-shot Learning via Contrastive Optimization of Attribute Representations." arXiv, Jul. 18, 2023, <https://doi.org/10.48550/arXiv.2207.03824>.

[2] G. K. Nayak, R. Rawal, I. Khatri, and A. Chakraborty, "Robust Few-shot Learning Without Using any Adversarial Samples." arXiv, Nov. 03, 2022, <https://doi.org/10.48550/arXiv.2211.01598>.

[3] W. Li *et al.*, "LibFewShot: A Comprehensive Library for Few-Shot Learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 12, pp. 14938–14955, Dec. 2023, <https://doi.org/10.1109/TPAMI.2023.3312125>.

[4] B. Xu, Z. Zeng, C. Lian, and Z. Ding, "Generative Mixup Networks for Zero-Shot Learning," *IEEE Transactions on Neural Networks and*

Learning Systems, pp. 1–12, 2022, <https://doi.org/10.1109/TNNLS.2022.3142181>.

- [5] N. Lai, M. Kan, C. Han, X. Song, and S. Shan, "Learning to Learn Adaptive Classifier–Predictor for Few-Shot Learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 8, pp. 3458–3470, Aug. 2021, <https://doi.org/10.1109/TNNLS.2020.3011526>.
- [6] A. Naghizadeh, D. N. Metaxas, and D. Liu, "Greedy auto-augmentation for n-shot learning using deep neural networks," *Neural Networks*, vol. 135, pp. 68–77, Mar. 2021, <https://doi.org/10.1016/j.neunet.2020.11.015>.
- [7] S. R. Sane, S. Tripathi, K. R. Sane, and R. Mamidi, "Stance Detection in Code-Mixed Hindi-English Social Media Data using Multi-Task Learning," in *Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Minneapolis, MN, USA, Mar. 2019, pp. 1–5, <https://doi.org/10.18653/v1/W19-1301>.
- [8] T. T. Sasidhar, B. Premjith, and K. P. Soman, "Emotion Detection in Hinglish(Hindi+English) Code-Mixed Social Media Text," *Procedia Computer Science*, vol. 171, pp. 1346–1352, Jan. 2020, <https://doi.org/10.1016/j.procs.2020.04.144>.
- [9] M. Hardalov, A. Arora, P. Nakov, and I. Augenstein, "Few-Shot Cross-Lingual Stance Detection with Sentiment-Based Pre-training," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 10, pp. 10729–10737, Jun. 2022, <https://doi.org/10.1609/aaai.v36i10.21318>.
- [10] H. Wen and A. Hauptmann, "Zero-Shot and Few-Shot Stance Detection on Varied Topics via Conditional Generation," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Toronto, Canada, Apr. 2023, pp. 1491–1499, <https://doi.org/10.18653/v1/2023.acl-short.127>.
- [11] E. Allaway and K. McKeown, "Zero-shot stance detection: Paradigms and challenges," *Frontiers in Artificial Intelligence*, vol. 5, Jan. 2023, <https://doi.org/10.3389/fraci.2022.1070429>.
- [12] V. Barriere and A. Balahur, "Multilingual Multi-Target Stance Recognition in Online Public Consultations," *Mathematics*, vol. 11, no. 9, Jan. 2023, Art. no. 2161, <https://doi.org/10.3390/math11092161>.
- [13] S. Mishra *et al.*, "Overview of Memotion 3: Sentiment and Emotion Analysis of Codemixed Hinglish Memes." arXiv, Sep. 12, 2023, <https://doi.org/10.48550/arXiv.2309.06517>.
- [14] A. Shahade, K. Walse, V. M. Thakare, and M. Atique, "Multi-lingual opinion mining for social media discourses: an approach using deep learning based hybrid fine-tuned smith algorithm with adam optimizer," *International Journal of Information Management Data Insights*, vol. 3, Nov. 2023, Art. no. 100182, <https://doi.org/10.1016/j.jjime.2023.100182>.
- [15] Z. Yao, W. Yang, and F. Wei, "Enhancing Zero-Shot Stance Detection with Contrastive and Prompt Learning," *Entropy*, vol. 26, no. 4, Apr. 2024, Art. no. 325, <https://doi.org/10.3390/e26040325>.
- [16] A. K. Barman, J. Sarmah, S. Basumatary, and A. Nag, "Word Sense Disambiguation applied to Assamese-Hindi Bilingual Statistical Machine Translation," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12581–12586, Feb. 2024, <https://doi.org/10.48084/etasr.6342>.
- [17] A. Kumar, R. Sharma, and P. Bedi, "Towards Optimal NLP Solutions: Analyzing GPT and LLaMA-2 Models Across Model Scale, Dataset Size, and Task Diversity," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14219–14224, Jun. 2024, <https://doi.org/10.48084/etasr.7200>.
- [18] D. Chopra, N. Joshi, and I. Mathur, "A Review on Machine Translation in Indian Languages," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3475–3478, Oct. 2018, <https://doi.org/10.48084/etasr.2288>.