

# **Hate/Offensive Text Detection in Hindi**

**HASOC- HOF Detection using Word Embeddings**

**Manish Kumar & Sivashankaran S**

# **Introduction: Problem Statement**

## **Sentiment Analysis: Hate/Offensive Text Detection**

- This project aims to create a binary classification model for detecting offensive language in social media(Twitter) content written in Hindi.
- Inspired by recent advancements in natural language processing, we seek to explore the effectiveness of incorporating sentence embeddings built on top of word embeddings in the initial layers of the model to capture subtle linguistic cues and provide a nuanced score for the offensiveness of the text.

# Resources and Data

## Dataset and Libraries

- Dataset: Hindi dataset from HASOC 2020 (Hate Speech and Offensive Content) competition
- Embeddings: FastText pre-trained word embeddings trained on Hindi Wikipedia
- Libraries:
  - Torch: PyTorch deep learning library
  - FastText: Library for efficient learning of word representations
- Hardware: Standard laptop with CPU for experimentation

text_id	text	task_1	task_2	task_3
hasoc_hi_5556	बांग्लादेश की शानदार वापसी, भारत को 314 रन पर रोका #INDvBAN #CWC19	NOT	NONE	NONE
hasoc_hi_5648	सब रंडी नाच देखने मे व्यस्त जैसे ही कोई #शांतीदूत के साथ कुछ होगा सब #रंडीरोना शुरू कर देगे	HOF	PRFN	UNT
hasoc_hi_164	तुम जैसे हरामियों के लिए बस जुतों की कमी है शुक कर अभी तुम्हारी लिचिंग हुई नहीं है हिंदुओं के जागने की देर है सच में होगी अभी तो तुम जैसे हगमी सुवर झामा बनाएं हो .सुवर कहीं का मौलाना। तुम जैसे हरामियों कुले की मौत मारना चाहिए सुवर जैसी शक्ल रंडी की औलाद सुवर व	HOF	PRFN	TIN
hasoc_hi_3530	बीजेपी MLA आकाश विजयवर्गीय जेल से रिहा, जमानत मिलने की खुशी में एक समर्थक ने इंदौर में हर्ष फायरिंग की #AkashVijayvargiya https://abpnews.abplive.in/india-news/celebratory-firing-outside-bjp-mia-akash-vijayvargiya-office-in-indore-1	NOT	NONE	NONE
hasoc_hi_5206	घमकी बुछार: विधानसभा परिसर में आरजेडी का प्रदर्शन, तेजस्वी यादव नदारद #biharencephallitisdeaths https://abpnews.abplive.in/bihar-news/aes-deaths-rjd-protest-in-vidhan-sabha-campus-but-tejashwi-yadav-was-not-present-11587	NOT	NONE	NONE
hasoc_hi_5121	मुंबई में बारिश से लोगों को काफी समस्या हो रही है	NOT	NONE	NONE
hasoc_hi_7142	Ahmed's dad:-- beta aaj teri mammy kyu nahi baat kr rhi h. Ahmed....	NOT	NONE	NONE
hasoc_hi_4321	5 लाख मुसलमान उत्त में, अजमेर की दरगाह पर आते हैं, सिर्फ 300 पुलिस वालों के भरोसे, 2 लाख हिंदुओं की अमरनाथ यात्रा के लिए, 80 हजार कमांडो, पैरामिलिट्री फोर्स, तथा करोड़ों के उपकरण लगाए जा रहे हैं #खतरे में #कौन है ?	NOT	NONE	NONE
hasoc_hi_4674	Do mahashaktiyan mili hain, charo taraf khushi ki leher hai, khus hone wale khus hi rhe hain aur bhakton ko taklif ho rhi hai, khair honi bhi chahiye.	NOT	NONE	NONE
hasoc_hi_1637	Chants of 'Jai Sri Ram' as Owaisi takes oath: AIMIM chief responds with 'Jai Bhim' @thenewsminute @asadowaisi	NOT	NONE	NONE
hasoc_hi_5771	नृत्य होगा..। संगीत होगा..। जीवन मे उत्सव होगा..। 'यह' आतुर है प्राणों में उतरने को.. 'तुम' इदय दूवार तो खोलो..। जगह तो खाली करो..। #हिंदी_शब्द	NOT	NONE	NONE
hasoc_hi_288	गाये के नाम पर जुल्म करने वालों तुम्हारी औकात है उन देश के दस बड़े गाये का मांस निर्यात करने वाले हिंदुओं से हिसाब लेने की..	NOT	NONE	NONE
hasoc_hi_1803	डॉ. मुखर्जी ने एक देश में दो विधान दो प्रधान और दो निशान का विरोध करते हुए कश्मीर को भारत का अभिन अंग मानकर परमिट राज को खत्म करने की लड़ाई लड़ी, उनके बलिदान से ही परमिट राज खत्म हुआ- श्री @ManojTiwarIMP	NOT	NONE	NONE
hasoc_hi_5900	धार्मिक पदों पर बैठे कुछ लोग अपने को भगवान समझ लेते हैं,कुछ लोग नत मस्तक हो उन्हें भगवान का दर्जा दिलाने के लिए अपना समय बर्बाद करते हैं...	NOT	NONE	NONE
hasoc_hi_5951	सिक्का मोहब्बत का उछाल रक्खा है ~ हाँ मगर खुद को सम्माल रक्खा है॥ अकेले तुम ही नहीं जो परेशान हो ~ उसने बहुतों को उलझन में डाल रक्खा है॥ ~ शाहरुख सिद्दीकी @Rekhta #shahruksiddiqui @srspoet	NOT	NONE	NONE
hasoc_hi_1723	कौन कहता है आँखों में प्यार नहीं होता। जो डूब गया फिर उसमे से पार नहीं होता। ये इश्क के बाजार का दस्तूर है जमाने में। जो लूट गया वह कभी खरीदार नहीं होता। नईम #बज़्म #हिंदी_शब्द @Aayatkhan_ @10398108Ritesh @Shaheen31304594 @queen_j	NOT	NONE	NONE

### OG Data

Unnamed: 0	text	preprocessed_text	task_1	HOF	NOT
2509	ये साला धर्म है या फिर रंडी भड़वों का कोठा।	ये साला धर्म है या फिर रंडी भड़वों का कोठा।	HOF	1	0
332	एक जवान जो अपना घर, बीवी, बच्चे छोड़ के सीमापर...	एक जवान जो अपना घर, बीवी, बच्चे छोड़ के सीमापर...	NOT	0	1
1113	LIVE: PM Shri @narendramodi addresses the Indi...	1113	NOT	0	1
751	रंडी अंजनवा ओम मोदी नॉट कश्यप ,भड़वे सारदाना, ...	रंडी अंजनवा ओम मोदी नॉट कश्यप ,भड़वे सारदाना, ...	HOF	1	0
720	ये इतनी बडी रंडी है की इसको कितना भी गाली दो इ...	ये इतनी बडी रंडी है की इसको कितना भी गाली दो इ...	HOF	1	0

### Preprocessed Data

```
import fasttext
import fasttext.util
ft = fasttext.load_model('wiki.hi.bin')
word = "नृत्य"
print("Embedding Shape is {}".format(ft.get_word_vector(word).shape))
print("Nearest Neighbors to {} are:".format(word))
ft.get_nearest_neighbors(word)
```

Embedding Shape is (300,)

Nearest Neighbors to नृत्य are:

[(0.8913929462432861, 'नृत्या'),  
(0.8440190553665161, 'नृत्यगान'),  
(0.8374733924865723, 'नृत्यगीत'),  
(0.8336297869682312, 'नृत्यों'),  
(0.8265783190727234, 'नृत्यरत'),  
(0.7971948385238647, 'नृत्यकला'),  
(0.7879464626312256, 'नृत'),  
(0.7682990431785583, 'नृतक'),  
(0.7622954845428467, 'नृत्यरचना'),  
(0.7602956295013428, 'नृत्यग्राम')]

# Experiment

## Sentence Embeddings

- Sentence embeddings capture semantic information of sentences and are crucial for various natural language processing tasks such as sentiment analysis, text classification, and machine translation.
- Classic techniques
  - Bag-of-words
  - TF-IDF vectors
- Traditional approaches for representing text data. Effective for documents but less effective for sentences. Lack of sequence information in BoW. TF-IDF may not capture semantic similarities between words (context)
- Unsupervised document embedding techniques
  - Averaging word embeddings
  - Sent2Vec
- Averaging word embeddings to generate sentence embeddings provides a robust representation of sentence semantics.

# Experiment

## Model and Training

- Model and Training:
  - fully connected neural network
  - Train-test split: 80% training, 20% validation
  - Optimiser: Adam optimiser with a learning rate of 0.001
  - Epochs: 5
- Evaluation Metrics:
  - Accuracy: Proportion of correctly classified samples
  - Precision: Proportion of true positive predictions out of all positive predictions
  - Recall: Proportion of true positive predictions out of all actual positives

Layer (type:depth-idx)	Param #
└Sequential: 1-1	--
└Linear: 2-1	60,200
└Linear: 2-2	20,100
└Linear: 2-3	1,010
└Linear: 2-4	11
└BCEWithLogitsLoss: 1-2	--
Total params: 81,321	
Trainable params: 81,321	
Non-trainable params: 0	

# Results and Discussion

## Model Performance

Accuracy: 0.7787913340935005

Precision: 0.7991452991452992

Recall: 0.7890295358649789



