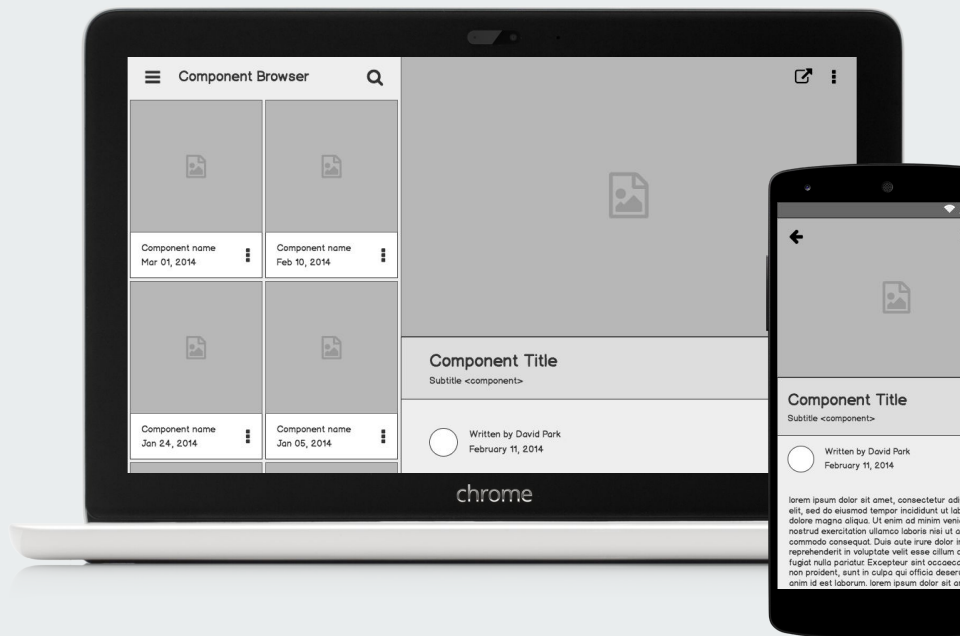


Identification and Classification of Offensive Tweets (SemEval19: OffensEval)

Archit Kumar
Arnav Kapoor
Neeraj Barthwal
Shruti Chandra



Problem

- Sub-task A: Offensive language identification; [Offensive: OFF, Not Offensive: NOT]
- Sub-task B: Automatic categorization of offense type [Targeted: TIN, Untargeted: UNT]
- Sub-task C: Offense target identification. [Individual: IND, Group: GRP, Other: OTH]

Dataset used

OLID dataset

	id	tweet	subtask_a	subtask_b	subtask_c
0	86426	@USER She should ask a few native Americans wh...	OFF	UNT	NaN
1	90194	@USER @USER Go home you're drunk!!! @USER #MAG...	OFF	TIN	IND
2	16820	Amazon is investigating Chinese employees who ...	NOT	NaN	NaN
3	62688	@USER Someone should'veTaken" this piece of sh...	OFF	UNT	NaN
4	43605	@USER @USER Obama wanted liberals & illeg...	NOT	NaN	NaN

Data Pre-processing

The first step was to clean and process the tweets so as to remove the noise and other unnecessary words, which is needed for any model and would help in providing better results. The steps used were as follows :-

- Decontraction - (can't -> can not , won't -> will not)
- Stop Words Removal
- Case folding
- Tokenization and Lemmatization

Preprocessed data

	id	tweet	new_tweet	user_mentions	n_hash_tags	n_urls	n_emojis	subtask_a	subtask_b	subtask_c	original_tweet_length
0	86426	@USER She should ask a few native Americans what their take on this is.	she ask native americans their take be	1	0	0	0	OFF	UNT	NaN	14
1	90194	@USER @USER Go home you're drunk!!! @USER #MAGA #Trump2020 🇺🇸🇺🇸 URL	go home you drink maga trump2020	3	2	1	1	OFF	TIN	IND	11
2	16820	Amazon is investigating Chinese employees who are selling internal data to third-party sellers looking for an edge i...	amazon investigate chinese employees sell internal data third party sellers look edge competitive marketplace amazon...	0	5	1	0	NOT	NaN	NaN	27
3	62688	@USER Someone should'veTaken" this piece of shit to a volcano. 🤡"	someone should vetaken piece shit volcano	1	0	0	1	OFF	UNT	NaN	11
4	43605	@USER @USER Obama wanted liberals & illegals to move into red states	obama want liberals amp illegals move red state	2	0	0	0	NOT	NaN	NaN	12

Feature Engineering

We added new columns i.e extracting features from the original data to improve the prediction of our models. The features added were as follows :

- Number of Hashtags
- Number of User Mentions
- Length of Tweet
- Number of URLs and Emojis.

Word Representations

We need to form a representation of the tweets (that is the mapping from textual data to real valued vectors) before sending it to our model. The various representations used were as follows :

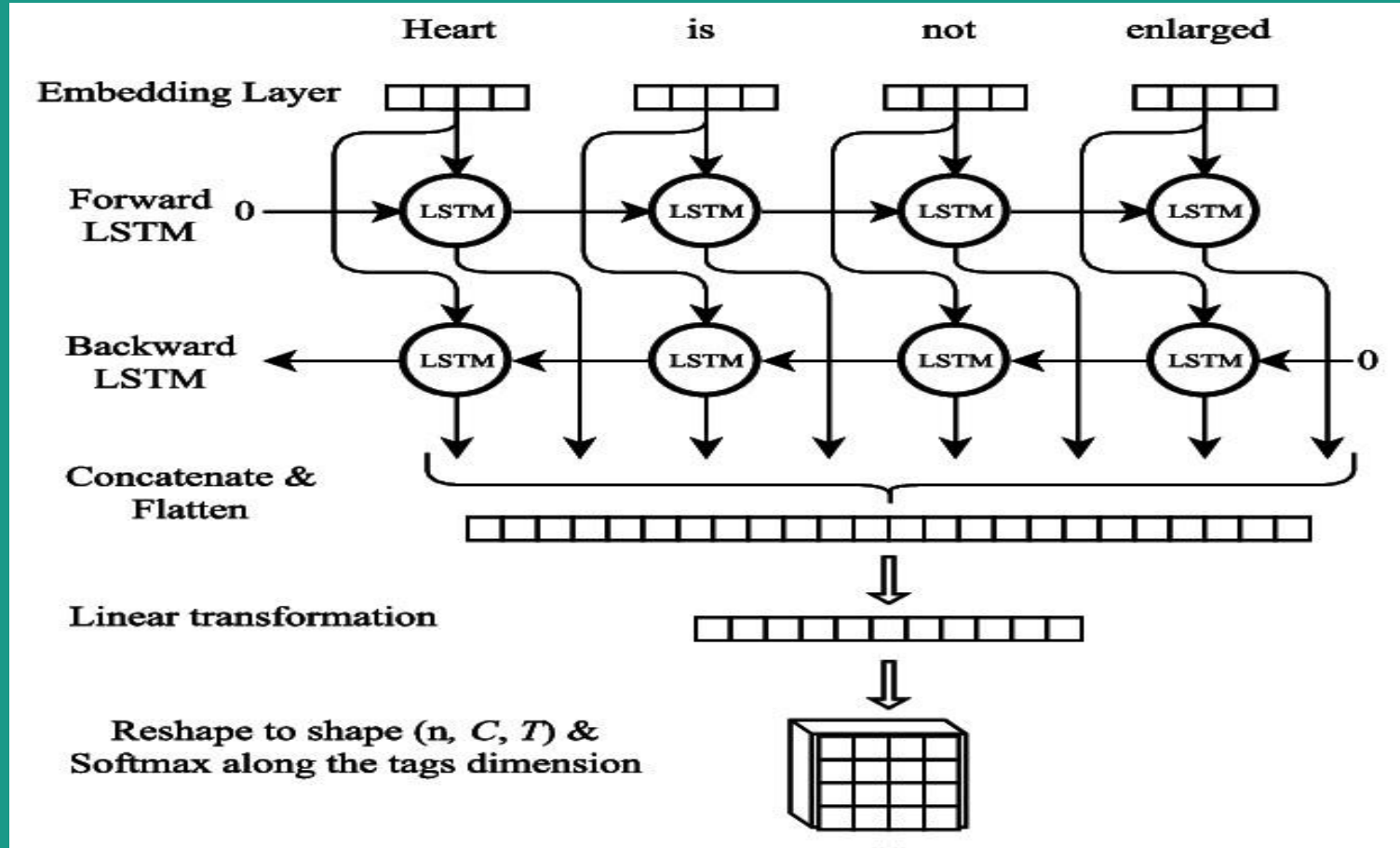
- Bag of Words
- TF-IDF
- GloVe Vectors (300 dimensions)

ML Models

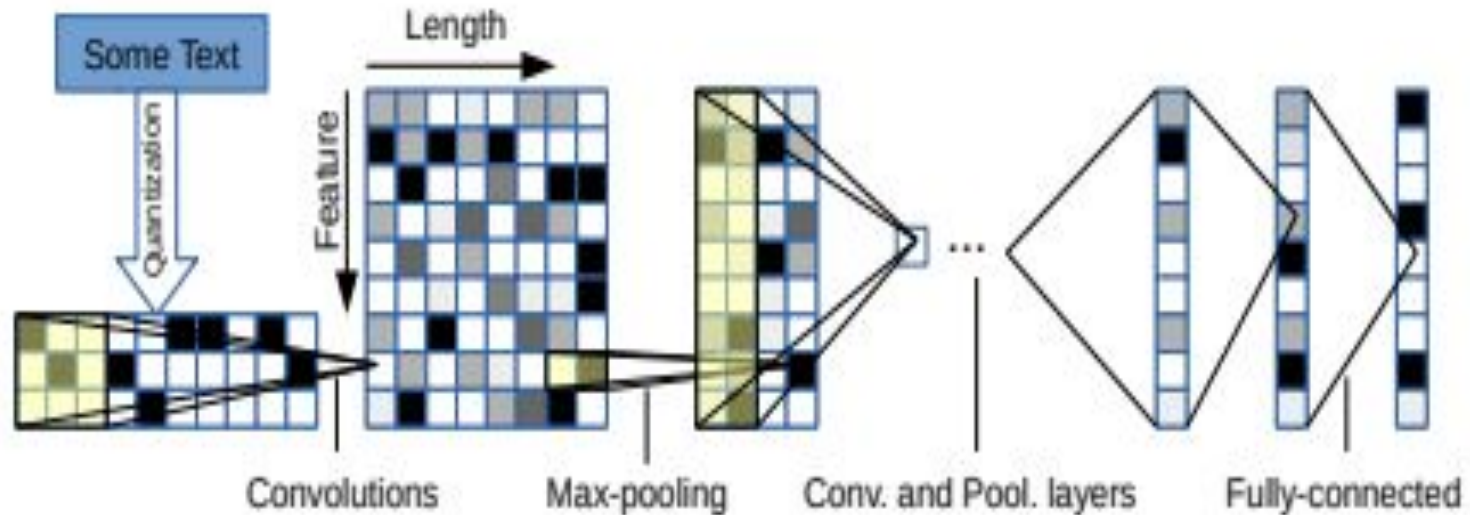
The models used for the classification tasks included the basic machine learning models (LR, Naive Bayes and SVM), and deep learning models (CNN and RNN).

- Logistic Regression
- Naive Bayes
- SVM
- **CNN**
- **RNN (LSTM and BiLSTM)**

BiLSTM Architecture



Char CNN Architecture



Handling Class Imbalance with weights

	Predicted label class 1	Predicted label class 2
True label class 1	correct true positive for class 1	wrong false positive for class 2
True label class 2	wrong false positive for class 1	correct true positive for class 2

$$\text{accuracy} = \frac{\text{orange} + \text{blue}}{\text{orange} + \text{yellow} + \text{blue} + \text{green}}$$

$$\text{class 1 precision} = \frac{\text{orange}}{\text{orange} + \text{yellow}}$$

$$\text{class 2 precision} = \frac{\text{blue}}{\text{blue} + \text{green}}$$

$$\text{class 1 recall} = \frac{\text{orange}}{\text{orange} + \text{green}}$$

$$\text{class 2 recall} = \frac{\text{blue}}{\text{blue} + \text{yellow}}$$

Results

Task A

Offensive language identification

F1- Macro Scores

	Bag of Words	Tf-Idf	Avg Word2Vec
Logistic Regression	0.712	0.710	0.649
Naive Bayes	0.705	0.658	0.591
SVM	0.700	0.702	0.643
CNN	0.612		
RNN	0.67		

Task B

Automatic categorization of
offense type

F1- Macro Scores

	Bag of Words	Tf-Idf	Avg Word2Vec
Logistic Regression	0.53	0.5613	0.565
Naive Bayes	0.542	0.597	
SVM	0.56091	0.55884	0.538198
CNN	0.49		
RNN	0.534		

Task C

Offensive Target Identification

F1- Macro Scores

	Bag of Words	Tf-Idf	Avg Word2Vec
Logistic Regression	0.70991522	0.72871	0.6000737
Naive Bayes	0.6603611	0.6595012	0.62080825
SVM	Linear: 0.60705 Rbf: 0.67411	Linear: 0.63628 Rbf: 0.59587	Linear : 0.5932 Rbf: 0.5666
CNN	0.599		
RNN (BiLSTM)	0.621		

Thank You!

