SemEval-2022

The 16th International Workshop on Semantic Evaluation

IISERB Brains at SemEval 2022 Task 6: A Deep-learning Framework to Identify Intended Sarcasm in English

Tanuj Singh Shekhawat[1], Manoj Kumar[2], Udaybhan Rathore[3], Aditya Joshi[4], Jasabanta Patro[5]

[1235]Indian Institute of Science Education and Research, Bhopal, [4]SEEK, Australia

Introduction.

Sarcasm in spoken or written form is a **type of verbal** irony that indicates the **difference between the literal and intended meanings** of an utterance.

A large portion of the web and social media text is sarcastic, which creates a challenge for traditional natural language processing (NLP) tasks like sentiment classification, opinion mining, harassment detection, author profiling.

The **SemEval-2022 Task 6** identifies some of the challenges persisting till now, particularly in English and Arabic texts.

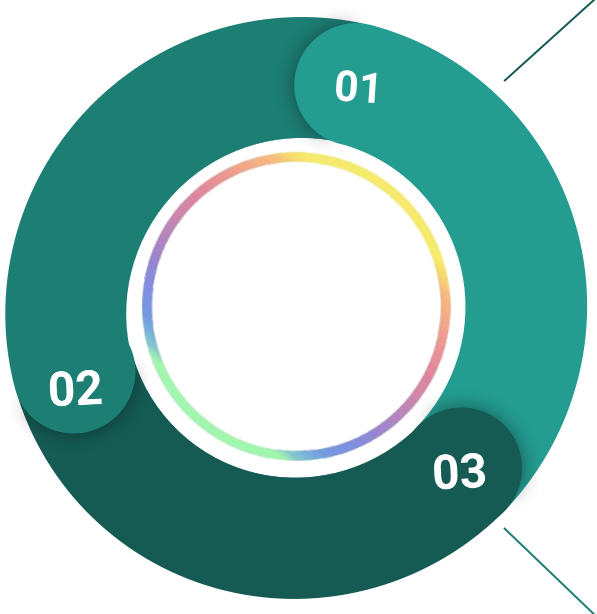
Task Details

Sub-task A:

Given a text, determine whether it is sarcastic or non-sarcastic

Sub-task B:

This sub-task is designed for particularly English dataset. It is a binary multi-label classification task. Here, given a text, we have to determine which ironic speech category it belongs to, if any.



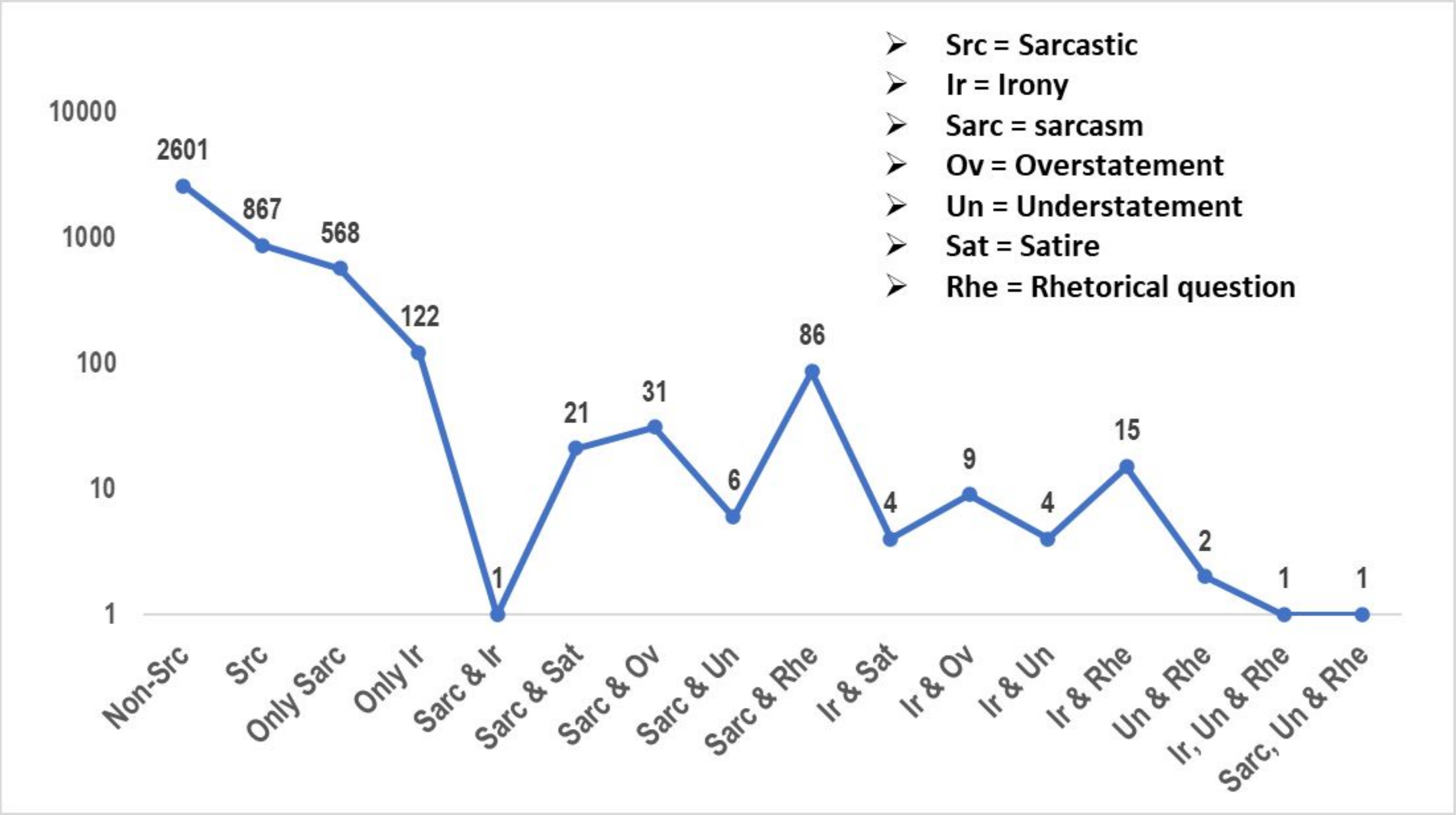
Sub-task C:

Given a sarcastic text and its non- sarcastic rephrase, i.e. two texts that convey the same meaning, determine which of the two is the sarcastic.

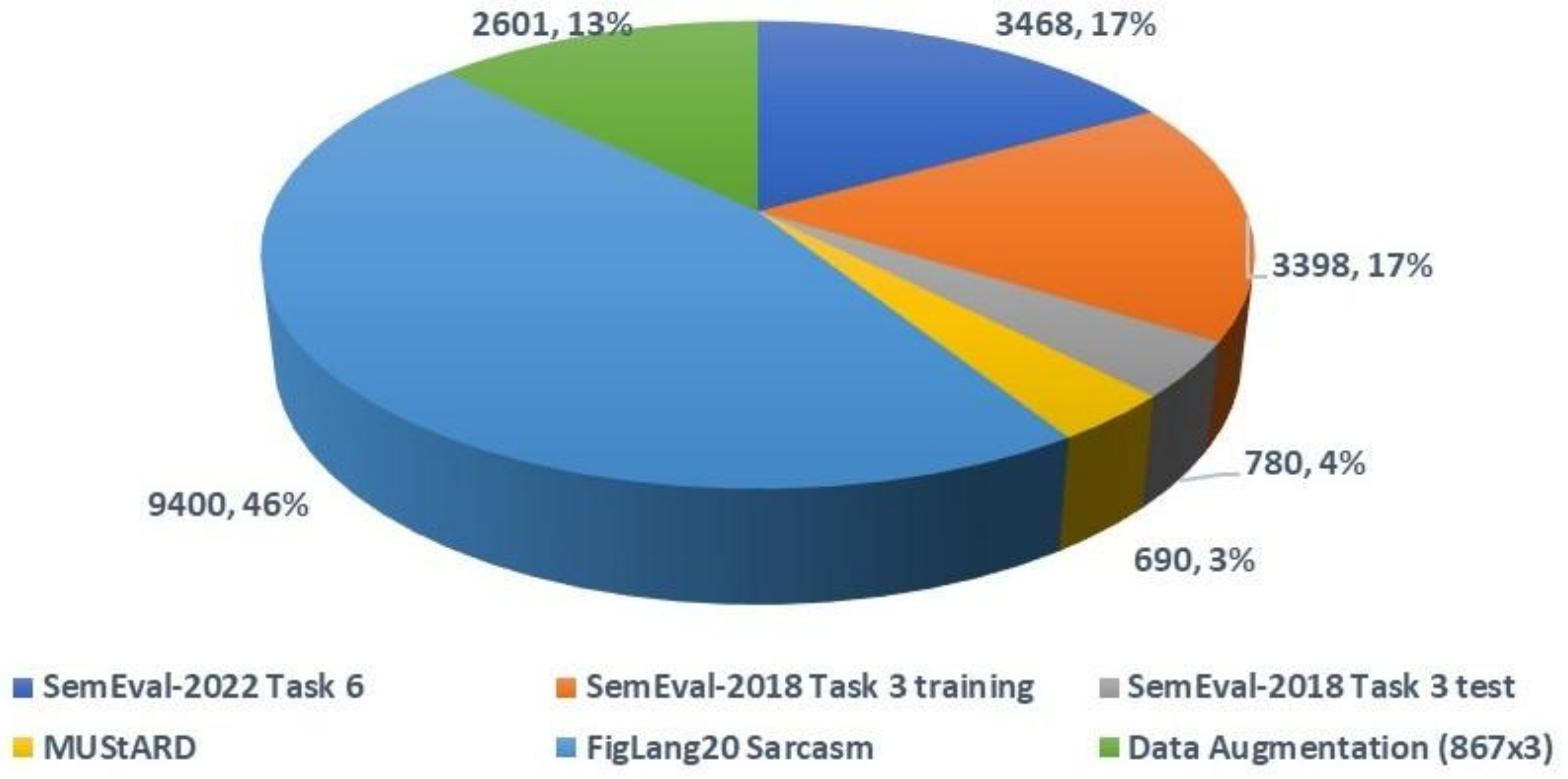
Dataset

- For the subtasks, we used the English dataset provided by the task organizers. It has 3468 English tweets which included 867 sarcastic and 2601 non-sarcastic tweets.

- For each sarcastic tweet, the organizers have also provided the **ironic sub-classes** to which the tweet belongs. The sub-classes are sarcasm, irony, satire, understatement, overstatement, and rhetorical question.



Src = Sarcastic
Ir = Irony
Sarc = sarcasm
Ov = Overstatement
Un = Understatement
Sat = Satire
Rhe = Rhetorical question
10000
1000
100
10
1
2601
867
568
122
1
21
31
6
86
4
9
4
15
2
1
1
Non-Src
Src
Only Sarc
Only Ir
Sarc & Ir
Sarc & Sat
Sarc & Ov
Sarc & Un
Sarc & Rhe
Ir & Sat
Ir & Ov
Ir & Un
Ir & Rhe
Un & Rhe
Ir, Un & Rhe
Sarc, Un & Rhe



For the sub-task A and C, we considered additional publicly available dataset as shown above.

We have used data augmentation for Sarcastic Class of SemEval-2022 Task 6 English Dataset.

Data Pre-processing

- Addition of Special Tokens
- Data Cleaning
  - Stop-word removal
  - Include removal of links, punctuation marks, etc
- Case conversion

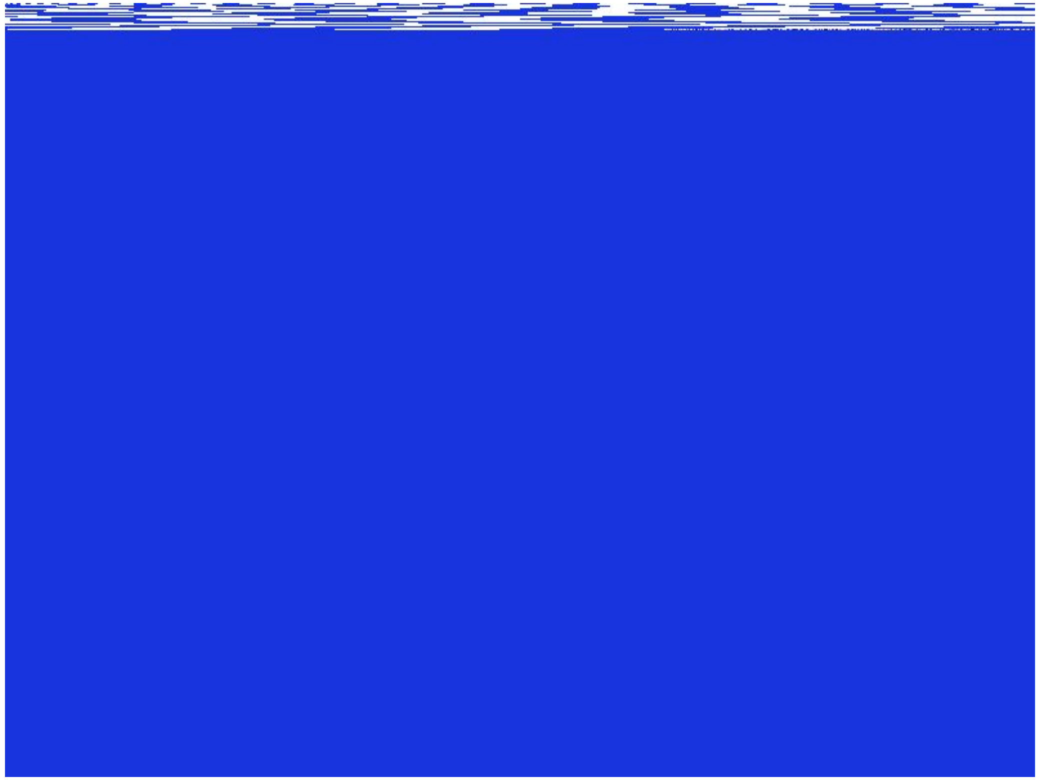
Methodology

- We relied on transformer based architectures to design our models for all sub-tasks.

- We built our models using the hugging face transformer library. They support generic transformer based architectures with the ability to seamlessly initialize the tokens with different pre-trained embeddings.

- In addition we use Data Augmentation using the python **nlpaug library**. For increasing the instances labeled with subcategories in the train data library.

Models

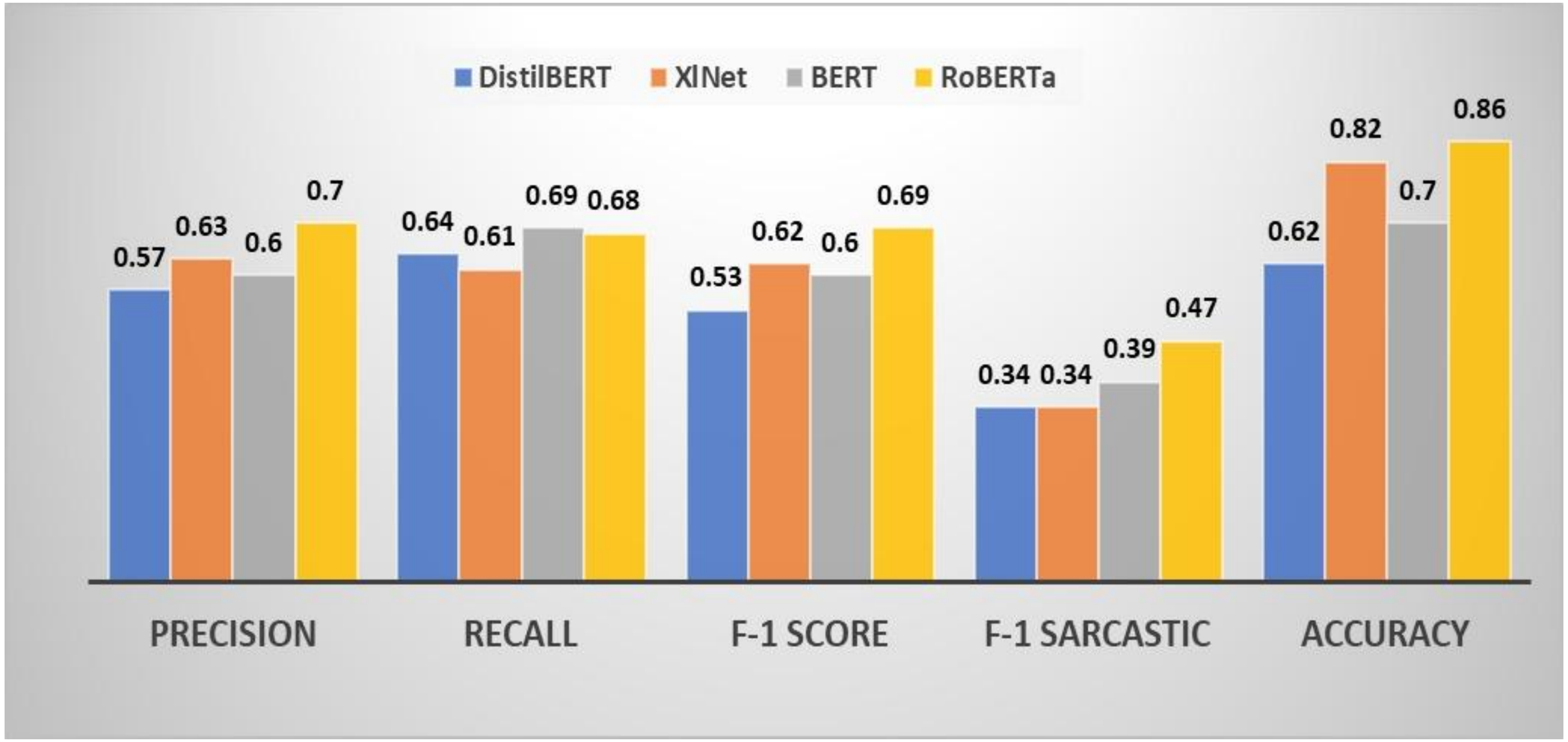


- **Sub-task A:** For this sub-task, we deployed the binary classifier versions of different transformer based architectures.
- **Sub-task B:** Here, instead of using a multilabel classifier, we used **six binary classifier** versions of the transformer based architectures.
- **Sub-task C:** We formulated this sub-task as a **parallel combination** of two sub-task A models

Results

To evaluate our systems we compare our results with the results provided by organizer and calculate F1 score.

Results of each Sub-task are shown alongside.



Results

SubTask A	
Model: BERT	F1=0.34
SubTask B	
Model: BERT	Macro-F1=0.0751
SubTask C	
Model: BERT	Accuracy=0.62

Conclusion

- We applied different transformer-based models to tackle the subtasks..
- **RoBERTa** performs relatively well in our overall experiments.
- Our model achieved the **19th** rank out of 43 teams on sub-task A, **8th** rank out of 22 teams on subtask B and **13th** rank out of 16 teams in sub-task C.

Thank You