Stance Detection in Tweets

W266 Natural Language Processing - Final Project

Alex Dessouky & Tim Spittle

December 7, 2019

Abstract

TBD - Abstract goes here

1 Introduction

Text

1.1 Background

More text

1.2 Objectives

More text

2 Related Work

3 Data, Task, & Evaluation

SemEval 2016 Task 6 (see Mohammad et al. 2016)

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character
```

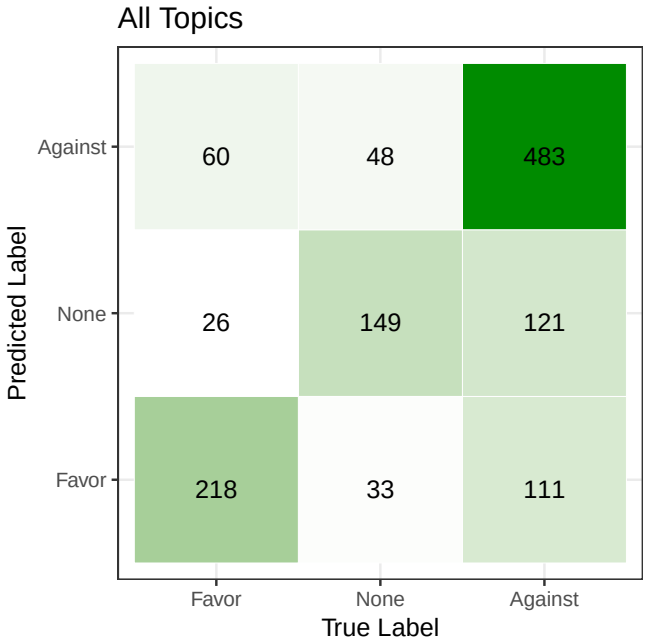
```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```



	Precision	Recall	F1.Score
Against	81.73	67.55	73.97
None	50.34	64.78	56.65
Favor	60.22	71.71	65.47

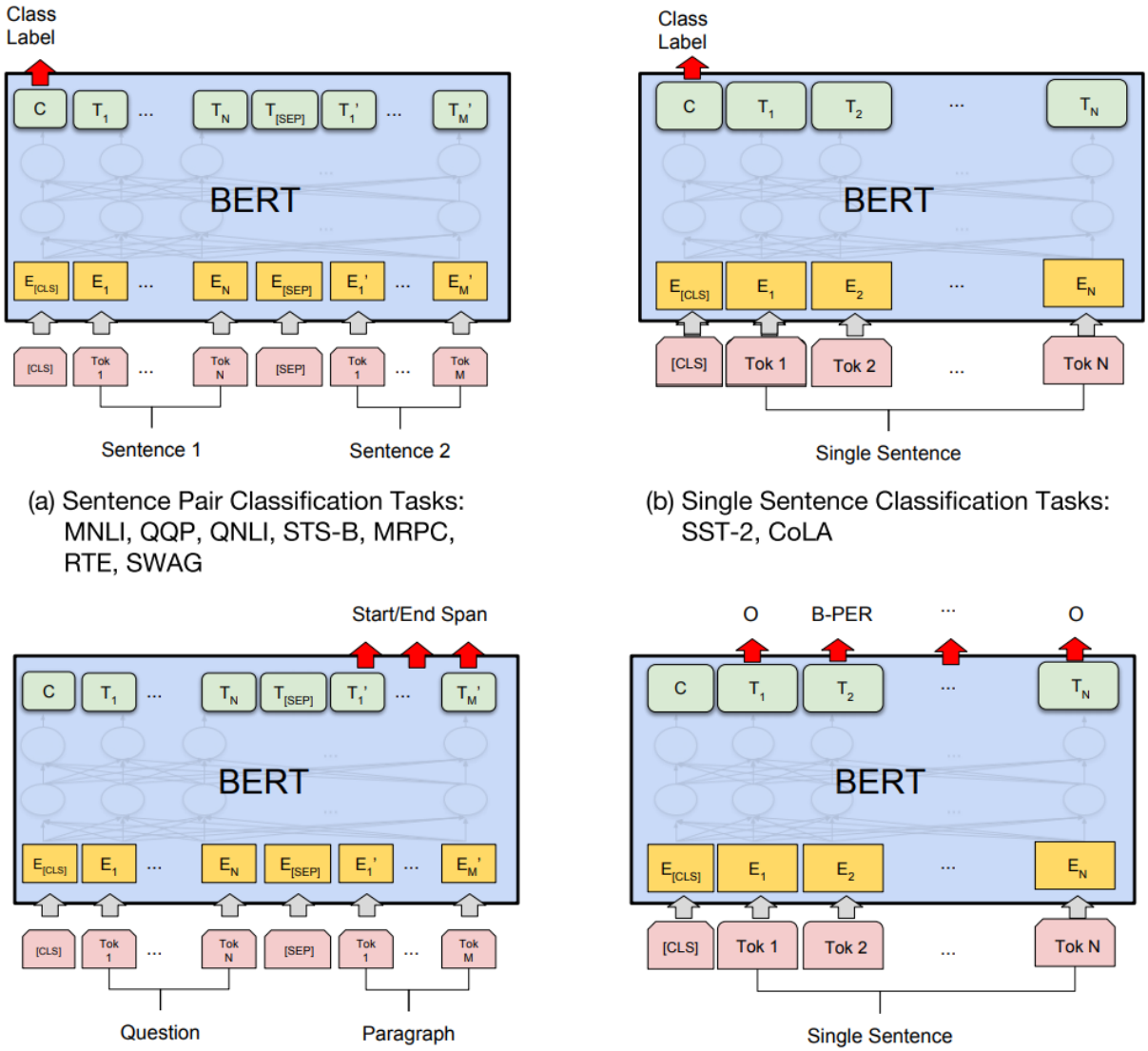
	F1.Score
F−micro	69.72
F−macro	60.88
Atheism	69.66
Climate Change	48.47
Feminism	57.76
Hillary Clinton	64.48
Abortion	64.06

Figure 1: Confusion Matrix

4 Methodology

4.1 Other Work

BERT (see *Figure 2* from Devlin et al. 2018)



(a) Sentence Pair Classification Tasks: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks: SST-2, CoLA

(c) Question Answering Tasks: SQuAD v1.1

(d) Single Sentence Tagging Tasks: CoNLL-2003 NER

Figure 2: BERT by Task (Source: @devlin2018bert)

5 Model

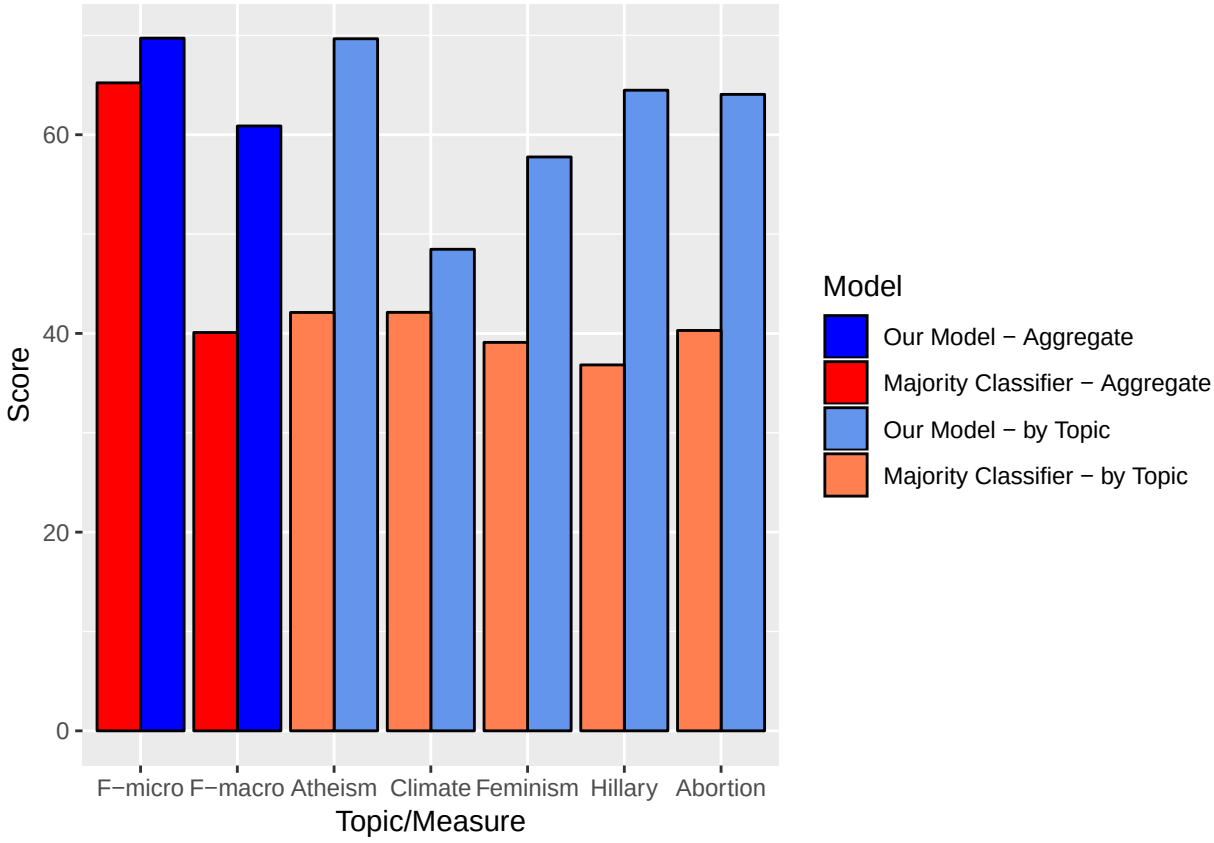
5.1 Infrastructure

5.2 Hyperparameter Tuning

5.3 Training

Hyperparameter	Final Value
Learning Rate	0.001
Dropout Rate	50%
Batch Size	32
# of Bert Fine Tune Layers	6 layers
Tweet Pre-processing	None

6 Results



Topic	F1 Score	SemEval-2016 Rank	Overall Rank
Overall Macro F1	60.88	1	2
Overall Micro F1	69.72	1	3
Atheism	69.66	1	2
Hillary Clinton	64.48	2	3
Abortion	64.06	1	3

Topic	F1 Score	SemEval-2016 Rank	Overall Rank
Climate Change	48.47	3	5
Feminism	57.76	2	3

Team	Overall (F-micro)	Overall (F-macro)	Atheism	Climate	Feminism	Hillary	Abortion
Muhammed et. al	70.32	59.01	69.19	43.80	58.72	61.74	66.91
Sun	69.79	61.00	70.53	49.56	57.50	61.84	66.16
Dessouky & Spittle	69.72	60.88	69.66	48.47	57.76	64.48	64.06
Du	68.79	59.56	59.77	53.59	55.77	65.38	63.72
MITRE	67.82	56.03	61.47	41.63	62.09	57.67	57.28
pkudblab	67.33	58.57	63.34	52.69	51.33	64.41	61.09
TakeLab	66.83	58.00	67.25	41.25	53.01	67.12	61.38
Majority Class	65.22	40.09	42.11	42.12	39.10	36.83	40.30
DeepStance	63.54	52.86	52.90	40.40	52.34	55.35	63.32
IDI@NTNU	62.47	55.08	59.59	54.86	48.59	57.89	54.47

7 Conclusion

7.1 Discussion

8 Limitations

8.1 Error Analysis

```
all_topics_errors = all_topics %>%
  filter(predicted_label != true_label) %>%
  select(obs_num, topic, tweet, predicted_label_f, true_label_f)

write.csv(all_topics_errors, file = "./final_outputs/all_topics_errors.csv")

### Abortion
## Potential negation?
# #pregnancyforall seems like a pro-choice tag but the "I have a right/power of choice" rhetoric is pro
print(all_topics_errors %>% filter(obs_num == 6 & topic == "abort") %>% select(tweet) %>% as.character(
```

```
## [1] "I have a right to identify as pregnant and have an abortion. Having an abortion confirms my powe
```

```
# There is no way to tall from the existing tweet
print(all_topics_errors %>% filter(obs_num == 10 & topic == "abort") %>% select(tweet) %>% as.character
```

```
## [1] "The government has given no explanation of why the law was changed #macedonia #HRCtte #SemST"
```

```
## Just wrong true label
# There is really no ambiguity here - this person is pro-abortion, should be FAVOR, labeled as AGAINST
print(all_topics_errors %>% filter(obs_num == 13 & topic == "abort") %>% select(tweet) %>% as.character
```

```
## [1] "Those who deny women who've been raped abortion are the same ppl who tell rape victims they aske
```

8.2 Further Work

References

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding."

Mohammad, Saif, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. "SemEval-2016 Task 6: Detecting Stance in Tweets." In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 31–41. San Diego, California: Association for Computational Linguistics. doi:10.18653/v1/S16-1003.