Using ensemble models with structural information in social media to aid rumour stance classification

Chen Wang, July 5th 2021

Supervisors: Anne Dirkson, Suzan Verberne, Stephan Raaijmakers.

Universiteit Leiden
The Netherlands

Overview

- Introduction
- Single models
  1. Method
  2. Experiment
- Ensemble models
  1. Method
  2. Experiment
- Discussion
- Future work

Introduction

- Motivation of rumour verification
- Pipeline of rumour verification

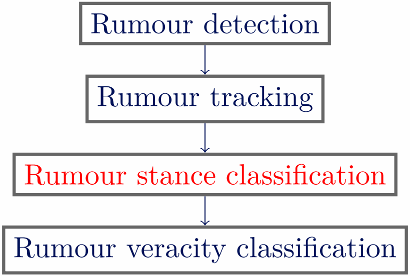


Figure: Pipeline of the rumour verification.

Introduction

- Stances: Support, Comment, Deny, Query
- Rumour stance classification in tree-structure social media

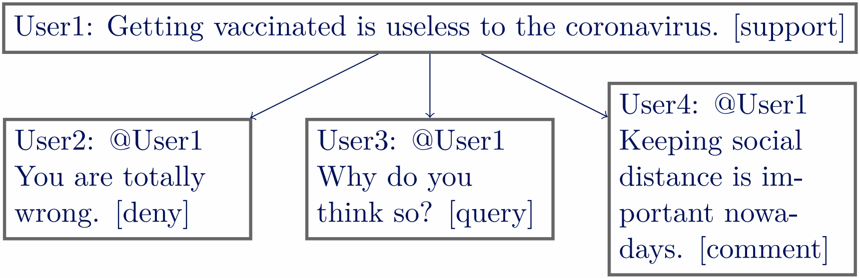


Figure: Example of a tree-structure Twitter conversation.

Introduction

- Previous research of utilizing structural information in stance classification
  1. Linear-CRF, Tree-CRF
  2. BranchLSTM, TreeLSTM
- Motivation of this research
  1. Blossoming of pre-trained language models
  2. Exploring the benefit of constructing ensemble models with sequential and non-sequential models

Introduction

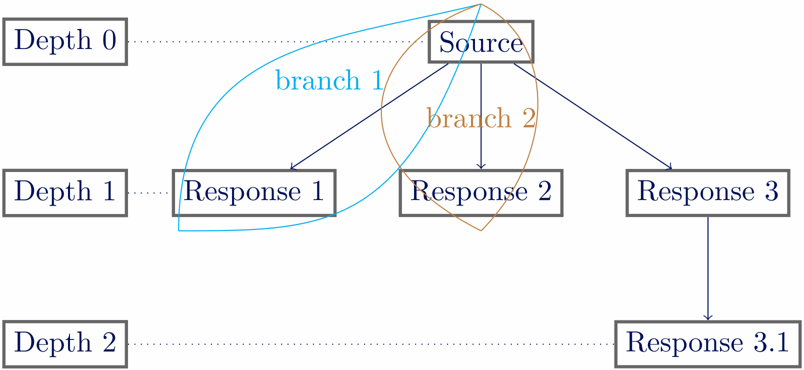
- Research questions:
  1. How does DistilBERT compare to the sequential models in rumour stance classification?
  2. How the combination of DistilBERT and sequential model help to rumour stance classification?

Methods

- Single-model method
  1. Adapted BranchLSTM
  2. DistilBERT
- Ensemble-model method
  1. one-step voting classifier
  2. two-step voting classifier

Single-model Method

- Baseline BranchLSTM



Single-model Method

- Adapted BranchLSTM
    1. Modified maximum input branch length
    2. Able to receive various sizes of conversations as input

Single-model Method

- DistilBERT
  1. Distilled version of BERT
  2. Treat each tweet as an input unit

Dataset split

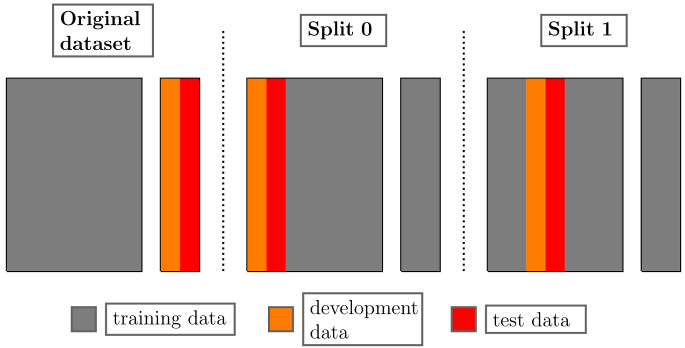


Figure: The split of dataset

Single-model Experiment

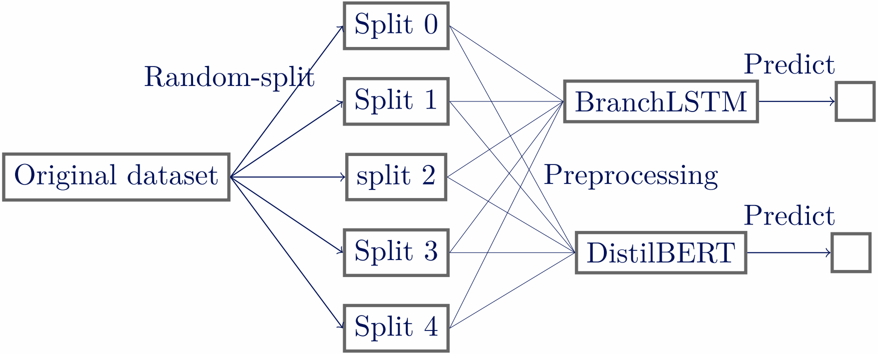


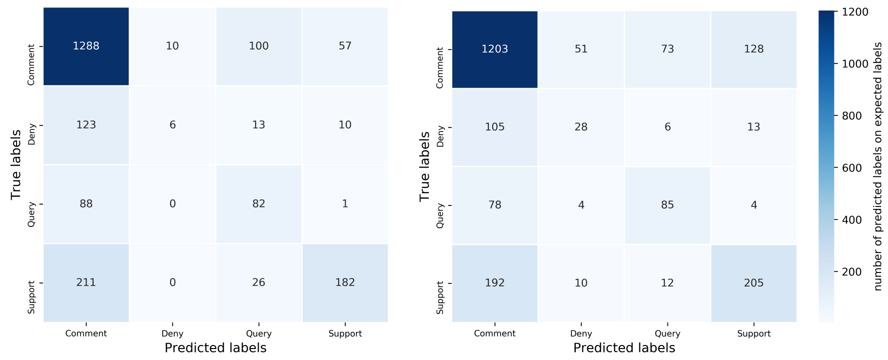
Figure: The single method classification pipeline.

Single-model Experiment

Model \ Metrics	Accuracy	Macro-average		
		Precision	Recall	F-score
BranchLSTM	0.704	0.525	0.459	0.457
DistilBERT	0.688	0.530	0.497	0.505
RumourEval2017	0.784			0.434

Table: BranchLSTM and DistilBERT average performance.

Single-model Experiment



(a) BranchLSTM Confusion matrix (b) DistilBERT Confusion matrix

Figure: CBranchLSTM's and DistilBERT's confusion matrix.

Ensemble-model Method

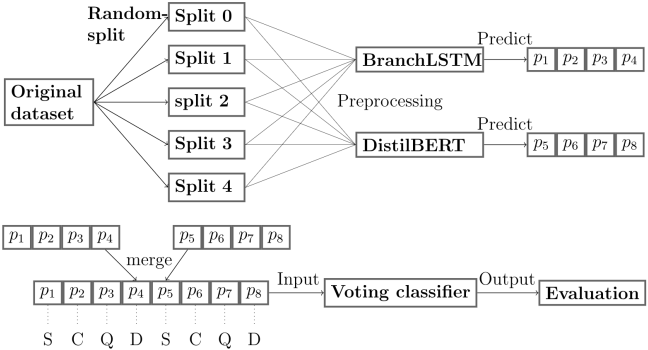


Figure: The ensemble method classification pipeline. S, D, Q, C represent ”Support”, ”Deny”, ”Query” and ”Comment” respectively.

Ensemble-model Method

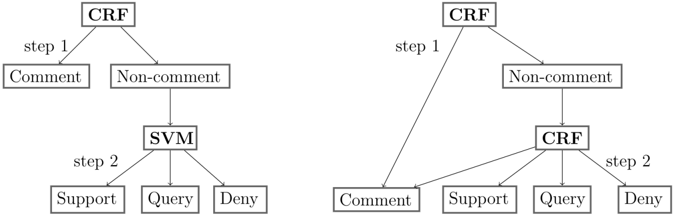


Figure: Two-step voting classification structure.

Ensemble-model Experiment

Model \ Metrics	Macro-average	
	Accuracy	F-score
SVM	0.689	0.489
CRF	0.689	0.488
CRF+SVM	0.678	0.486
CRF+CRF	0.688	0.486

Table: All ensemble models' performance on test dataset

Ensemble-model Experiment

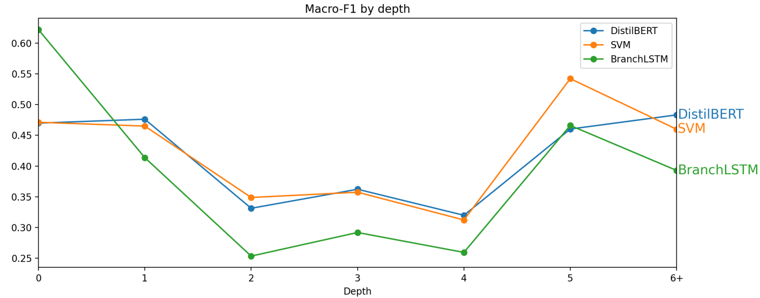


Figure: Single model and one-step SVM ensemble model's Macro-F1 scores by depth.

Ensemble-model Experiment

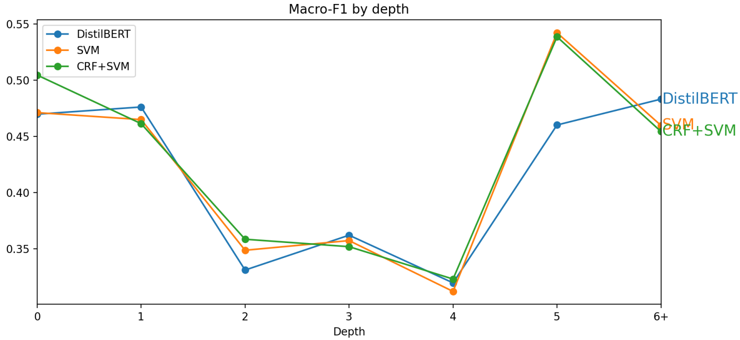


Figure: Macro-F1 scores by depth of a single model, one-step voting ensemble and two-step voting ensemble model.

Discussion

- DistilBERT has advantages in addressing more balanced data comparing to BranchLSTM, but BranchLSTM could perform better in predicting the large categories of imbalanced data.
- Ensemble model could learn features of both single models, but could not keep advantages from both single models.
- Sequential voting classifier could make use of the context but may have less strength in classifying a single simple tweet comparing to the non-sequential classifier.

Futere work

- Try random oversampling to address imbalanced data.
- Change the way of combining probabilities.
- Attempt other ensemble methods.
- Apply Tree-LSTM to the ensemble model.

Q & A