Information Retrieval
Term Project Report

Mentor: Rajdeep Mukherjee

Group 17

Jalend Bantupalli	18EC10023
Ganesh Shiridi Balaji Udayagiri	18EC35010
Chapala Chinnikrishna Naidu	18EC10010
Kandi Vishnu Vardhan Reddy	18CS30022

Common task

Problem Statement:

Incorporate *stance classification* from “All-in-one: Multi-task Learning for Rumour Verification” into Tree LSTM-based *rumour detection* from “Going Beyond Content Richness: Verified Information Aware Summarization of Crisis-Related Microblogs” and obtain the results on PHEME-RNR dataset

Implementation done so far:

The individual papers are run the datasets given. Verified Information Aware Summarization of Crisis-Related Microblogs on PHEME-RNR dataset and All-in-on: Multitasking for Rumour Verification on already preprocessed data.

The stance labels are added to the trees generated as a result of generate_trees.py

Results:

Verified summarisation:

Accuracies reported in the paper

Model	Charliehebdo		Germanwings		Ottawa		Sydney	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
CRF	0.715	0.570	0.678	0.696	0.718	0.721	0.717	0.640
RNN	0.694	0.621	0.617	0.642	0.711	0.729	0.704	0.639
CSI	0.758	0.605	0.502	0.612	0.515	0.649	0.594	0.625
LDA-TL	0.734	0.523	0.561	0.668	0.715	0.553	0.672	0.603
CETM-RNN	0.776	0.613	0.704	0.716	0.696	0.739	0.741	0.645
CETM-TL	0.787	0.656	**0.716**	0.721	**0.755**	**0.752**	0.737	0.686
CETM-RS-TL	**0.804**	**0.686**	0.702	**0.728**	0.745	0.740	**0.744**	**0.698**

Accuracies obtained from experiments

K	T	numItr
30	10	1000

IN_FEATURES	OUT_FEATURES	NUM_ITERATIONS	BATCH_SIZE	HIDDEN_UNITS	LEARNING_RATE
40	2	10	100	128	0.001

Eval/Data	charliehebdo	germanwings-crash	ottawashooting	sydneysiege
Accuracy	0.700400801603	0.538461538461	0.498245614035	0.610921501706
F1-score	0.584557446660	0.421834505738	0.403997237681	0.492577105811

IN_FEATURES	OUT_FEATURES	NUM_ITERATIONS	BATCH_SIZE	HIDDEN_UNITS	LEARNING_RATE
40	2	10	50	128	0.001

Eval/Data	charliehebdo	germanwings-crash	ottawashooting	sydneysiege
Accuracy	0.748997995991	0.501240694788	0.466666666666	0.574232081911
F1-score	0.595323485312	0.338259441707	0.318181818182	0.364769647696

Multitask Learning
Accuracies reported in the paper

RumourEval						
	Majority (True)	**NileTMRG***	**branchLSTM**	**MTL2 Veracity+Stance**	**MTL2 Veracity+Detection**	**MTL3**
Macro F	0.148	0.539	0.491	0.558	-	-
Accuracy	0.286	0.570	0.500	0.571	-	-
PHEME 5 events						
	Majority (True)	**NileTMRG***	**branchLSTM**	**MTL2 Veracity+Stance**	**MTL2 Veracity+Detection**	**MTL3**
Macro F	0.226	0.339	0.336	0.376	0.373	0.396
Accuracy	0.511	0.438	0.454	0.441	0.410	0.492
PHEME 9 events						
	Majority (True)	**NileTMRG***	**branchLSTM**	**MTL2 Veracity+Stance**	**MTL2 Veracity+Detection**	**MTL3**
Macro F	0.205	0.297	0.259	0.318	0.345	0.405
Accuracy	0.444	0.360	0.314	0.357	0.397	0.405

Results obtained

```
python outer.py --model='mtl2stance' --data='RumEval' --
search=True --ntrials=10 --params="output/bestparams.txt"
```

Rumour Eval - MLT2STANCE - accuracy: 0.5741811175337187,0.3387889270262222


```
{'Params': {'batchsize': 32, 'l2reg': 0.001, 'learn_rate': 0.001,
'num_dense_layers': 2, 'num_dense_units': 300, 'num_epochs': 50,
'num_lstm_layers': 1, 'num_lstm_units': 200
    }, 'TaskA': {'accuracy': 0.5741811175337187, 'Macro':
{'Macro_Precision': 0.3387889270262222, 'Macro_Recall':
0.3704994066402397, 'Macro_F_score': 0.3371820265926949
        }, 'Micro': {'Micro_Precision': 0.5741811175337187,
'Micro_Recall': 0.5741811175337187, 'Micro_F_score':
0.5741811175337187
        }, 'Per_class': {'Pclass_Precision': array([
                0.20283019,
                0.78475336,
                0.15789474,
                0.20967742
            ]), 'Pclass_Recall': array([
                0.46236559,
                0.68270481,
                0.21428571,
```

```
                0.12264151
            ]), 'Pclass_F_score': array([
                0.28196721,
                0.73018081,
                0.18181818,
                0.1547619
            ])
        }
    }, 'TaskB': {'accuracy': 0.4444444444444444, 'Macro':
{'Macro_Precision': 0.4682539682539682, 'Macro_Recall':
0.4545454545454546, 'Macro_F_score': 0.4481481481481482
        }, 'Micro': {'Micro_Precision': 0.4444444444444444,
'Micro_Recall': 0.4444444444444444, 'Micro_F_score':
0.4444444444444444
        }, 'Per_class': {'Pclass_Precision': array([
                0.5,
                0.57142857,
                0.33333333
            ]), 'Pclass_Recall': array([
                0.5,
                0.36363636,
                0.5
            ]), 'Pclass_F_score': array([
                0.5,
                0.44444444,
                0.4
            ])
        }
    },
```


Problems faced and contributions:

Jalend

- **Creating corpus file from PHEME9 dataset**- The corpus file was not provided but the steps on how to obtain it are mentioned. This took quite a time due to large dataset preprocessing and POS tagging. The given raw data json files had to be split according to tweet ids and tweet text for source - rumour, non-rumour

and reactions - rumour, non-rumour. Then the text had to be POS tagged and all the files had to be merged according to the format mentioned.

- We created the corpus file from scratch using the raw data.

- It was later realised (after going through the accuracies) that some parent tweets of reactions are missing in the raw dataset due to which the accuracies reduced.

- **Topic model** - Beta function used wasn't available in our devices. So we had to use another beta function with a different mean, variance. The accuracies are not affected much. We compared them to the other group's accuracies where the original beta function was working.

- Made necessary changes to the **code - 2 (multitask learning)** to run it on the dataset (saved_data) provided .

Balaji and Chinnikrishna

- Made necessary modifications to the **code - 1 (Verified Summarization)** to run it for the PHEME dataset

- **Generate trees + stance** - Generate tree python file was initially run for the corpus created. It did not have the stance component in each node. So using stance.json file (merged from three json files available) , we added a new field in the node named "stance" whose data is one hot encoding of the available stances (4 in number + 1 extra no stance label for unverified (no stance).

- **Tree lstm** - Tree lstm was run for the trees not including the stance label and the accuracies obtained are mentioned in the above results section

Plan to complete remaining task:

- Since stance was labelled in every node, the plan is to use this for training the tree lstm. One more loss is to be incorporated in the function and the prediction for the stance part should be outputted at every node.
- The lstm is to be modified to provide 4 outputs (one for each stance) and pass it through a softmax layer for every node.

Individual task

Problem Statement:

Run the assigned paper “Cascade-LSTM: A Tree-Structured Neural Classifier for Detecting Misinformation Cascades” on PHEME-RNR dataset

Implementation done so far:

The code was run for the provided dataset. - FalseNews_Code_Data_

PHEME-RNR dataset is preprocessed to suit the needs.

Raw_data_anon.csv	PHEME-RNR
tid	id
veracity	True/False/Unverified
cascade_id	Each source node’s folder is numbered as a cascade id
rumor_id	Each rumour is given an id
rumor_category	Depending on one of the events, manually should assign - {Politics, War/Terrorism/Shootings, Viral}
parent_tid	parent_tweetid
tweet_date	created_at
user_account_age	Put to a constant value of 21 for all
user_verified	TRUE/FALSE
user_followers	followers_count
user_followees	Made equal to user_followers
user_engagement	Put to a constant value of 8.39449955042 for all
cascade_root_tid	Tweet id of the root id
was_retweeted	Retweet_count > =1

emotions_anon.csv	PHEME-RNR
tweet_id	id
sadness	Random value
anticipation	Random value
disgust	Random value
surprise	Random value
anger	Random value
joy	Random value
fear	Random value
trust	Random value
misc	Random value

The green rows indicate both columns are presented in the dataset used for this paper and PHEME-RNR dataset.

Yellow rows indicated the assumption made to process PHEME-RNR dataset into the required csv format.

Results:

Paper results:

Model	AUC	Precision
Cascade LSTM	0.741	0.848

Experiment results (upto 14 epochs)

Started experiment 10_12_2021__20_20_22__903103
Model 10_12_2021__20_20_22__903103 saved with test AUC 0.6042 | train AUC 0.5401 at epoch 0
Model 10_12_2021__20_20_22__903103 saved with test AUC 0.6379 | train AUC 0.5656 at epoch 1
Model 10_12_2021__20_20_22__903103 saved with test AUC 0.6441 | train AUC 0.5945 at epoch 3

Model 10_12_2021__20_20_22__903103 saved with test AUC 0.6464 | train AUC 0.6135 at epoch 5
Experiment 10_12_2021__20_20_22__903103 terminated with **test AUC 0.6464** at epoch 14

Precision - **0.830208233445699**

Problems faced and contributions:

Jalend and Vishnu

- **DGL library** had outdated functions - so functions had to be updated and **brought to the latest version** accordingly which took some time
- Made necessary changes in the code **(Cascade LSTM)** to run it on the given dataset.
- **PHEME-RNR doesn't have the required features** - as mentioned in the observations above.
- Assumption that reactions are retweets - csv is created based on the assumptions made mentioned above
- Due to this assumption, we found out there was a loop in the graph created -
- dgl._ffi.base.DGLError: [14:08:47] /tmp/dgl_src/src/array/cpu/./traversal.h:222: **Error in topological traversal: loop detected in the given graph.**

Plan to complete remaining task:

- Try to remove the features not there in PHEME data set, probably mask them
- Try to remove the loops in the cascade
- Experiment a different approach to avoid loops