COSC 2779/2972 (Deep Learning)

ASSIGNMENT 2 MARKING SHEET

	Approach (50%)	
DI	**1) Data exploration leading to well informed approach.**	
	a. Identified class imbalance and target bias.	**Good**
	b. Identified the key properties of the problem like importance of target, tweet, low training data, need for data cleaning etc.	**Good**
	2) Identifying an adequate evaluation framework that is tailored to the problem.	
	a. Clearly identified performance measure (need to discuss suitability): e.g. macro-f1	**Yes**
	b. Identified Target value (Need to justify): e.g. 0.62 f1-score - literature	**Yes**
	c. Clear test/train/val split (or CV for hyper parameter tuning) - before augmentation.	**Yes**
	d. Used the splits appropriately: test not touched till end (ultimate judgment), val for hyper parameter tuning, train to model estimation. If CV no need for val set.	**Yes**
	3) Well justified network architecture and objective.	
	a. Justifications of the base model selection provided. Need to express why the base model is suitable over others that were left out (e.g. based on challenge winners, latest transformer).	**Good**
	b. model is adequate for the task. (RNN with embedding or Transformer, handle imbalanced)	**Good**
	c. Considered using all information - both tweek & target in a single model	**Good**
	d. Transfer learning used & done correctly (e.g freeze base model, train head, unfreeze and finetune.)	**Yes**
	4) Hyper parameters selection strategy.	
	b.Clear evidence of hyper parameter tuning (model improvements) using train/val curve observations (should not do ad-hoc jumps).	**Avarage**
	c. Proper use of validation set (or CV) to tune parameters.	**Yes**
	5) Approach satisfy all the requirements and restrictions.	
	a. Approach achieves reasonable performance (test set)	**No**
	Ultimate Judgment & Analysis (30%)	
DI	**1) Analysis of the model and the outputs using suitable methods.**	
	a. Compared models using independent test data.	**Yes**
	b. Used multiple performance metric to analyse the results (e.g. confusion matrix).	**Yes**
	c. Analysed error cases (different targets & classes considered seperately).	**Good**
	2) Make a clear ultimate Judgment.	
	a. One model selected clearly as the ultimate model (need to be explicit in report)	**Yes**
	3) Rational behind the ultimate model is clear and considers all the aspects.	
	a.Justification provided with performance comparison	**Avarage**
	b.Discussions beyond simple performance comparison in selecting best model	**Avarage**
	c.Considered over/under fitting	**Avarage**
	4) Conduct evaluations with independent test data.	
	a Conduct evaluations with independent test data (self created or obtained from tweets).	**Yes**
	5) Limitations of the model identified.	
	a. Include evidence based discussion of model limitations and potential improvements.	**Avarage**
	Report & Code (20%)	
HD	1) Code is well documented and easy to understand.	**Good**
	2) Code does not contain errors. (e.g. incorrect encording for test/train sets, incorrect scaling)	**Yes**
	3) Code contains evidence of all investigations mentioned in report.	**Yes**
	4) Code is optimal and shows good programming practices.	**Yes**
	5) Report Well structured and easy to read.	**Yes**
	6) Reader can fully understand the rationale for the approach taken	**Yes**

Late Penalty (Days)	0
Final Total (Out of 50)	39.5

Other Comments: 1. Classfication report can be used for the error analysis.