

Animesh Prasad

School of Computing, National University of Singapore

❖ Introduction

➤ Tasks:

- Task1a: Given citance, i.e. citation context from a Citing Paper (CP) referring to a Reference Paper (RP), identify the text in RP which is referred by the CP i.e. provenance
- Task1b: Typing the identified text in one or many of the possible discourse facet

➤ Prior Work

- Almost all prior systems deploy TF-IDF to capture surface level closeness and different textual similarity features for semantic closeness

➤ Challenges

- Identifying correct reference text among huge noisy signal
- Prior work doesn't capture instances of implied/interpreted provenance with less surface level similarity

❖ Proposed Technique

$$\text{Score} = \alpha_1 (\text{surface level closeness score}) + \alpha_2 (\text{learned semantic relation score})$$

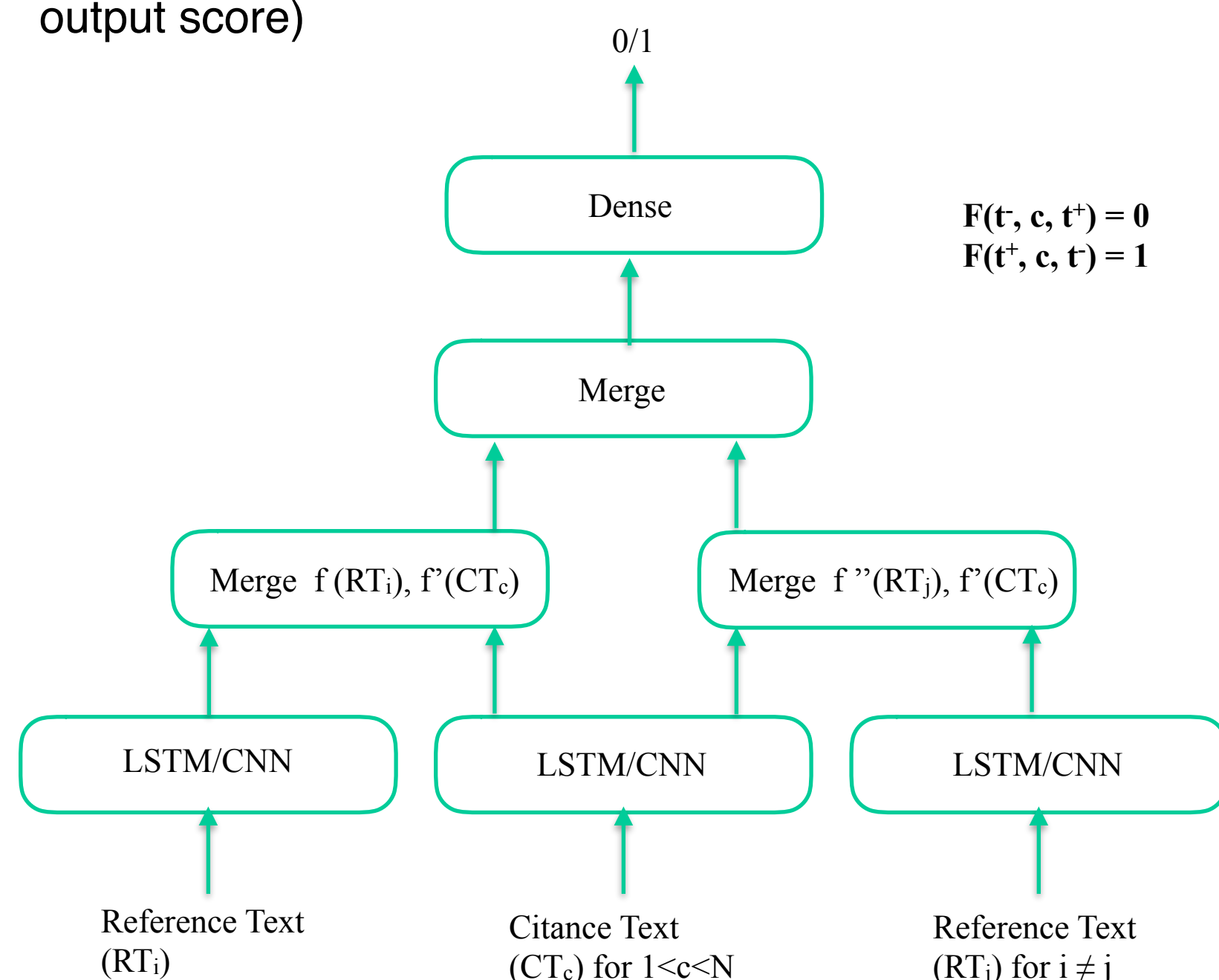
$$\text{surface level closeness score} = \alpha'_1 (\text{TF-IDF score}) + \alpha'_2 (\text{LCS score}) + \alpha'_3 (\text{Jaccard similarity}) + \dots$$

learned semantic relation score: Neural Ranking Model

❖ Model

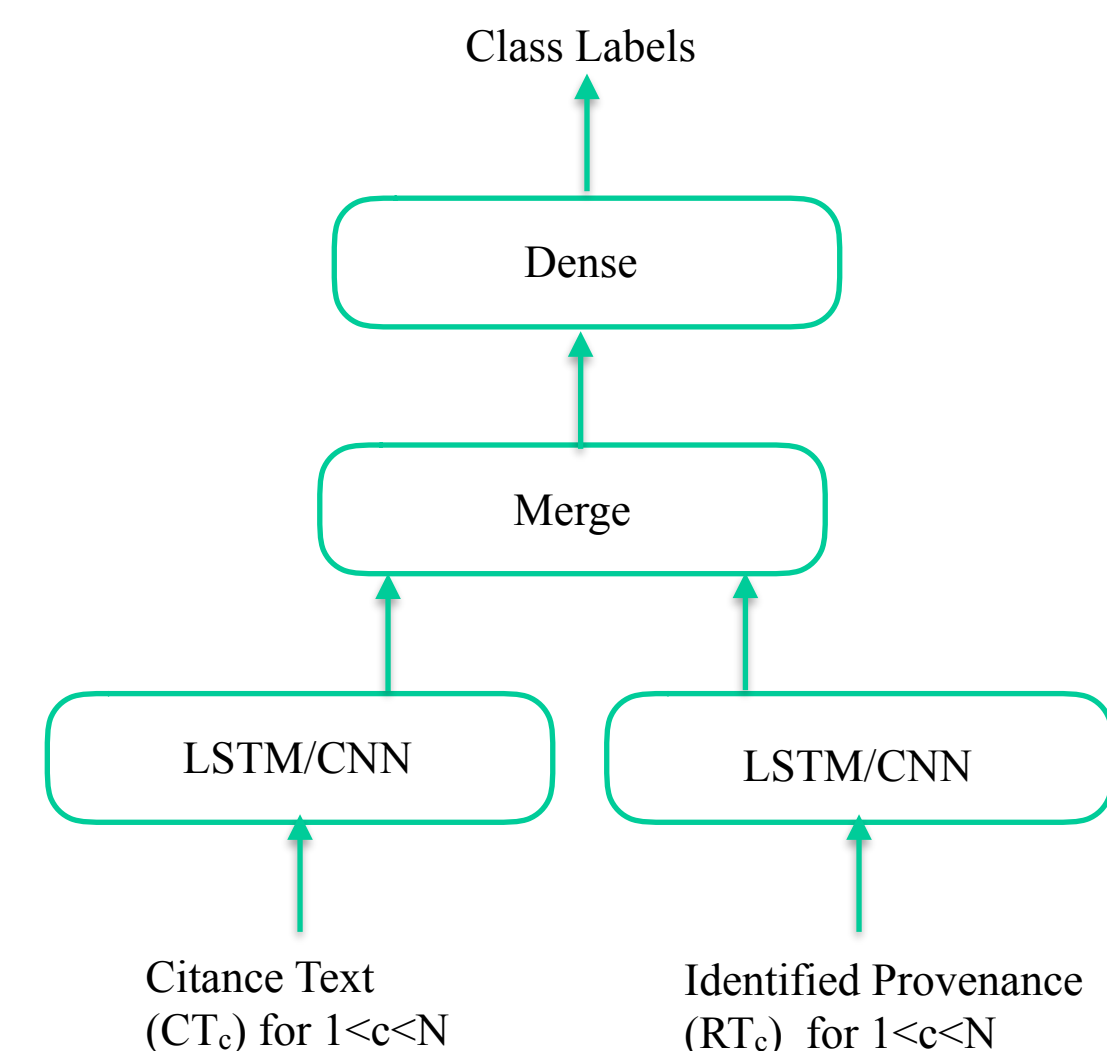
➤ Neural Ranking Model

- Use neural network to capture the underlying relationship between the CP and RP texts
- Find preferential score of all pair of the RP text with respect to the citance
- Use the overall counts of times a RP text is preferred as provenance over other as score (normalised by maximum as output score)



➤ Classification Model

- We classify the identified provenance by using similar architecture as ranking model with log scaled inverse frequency class weights
- Output all labels within δ probability difference of the output class



❖ Results

➤ Scores on Dev Set (CL-SciSumm 2016 Test Set)

➤ Task 1a Development Setups

Setup Type	P	R	F ₁	F ₁ All False
Classification	0.69	0.70	0.69	0.67
Preferential Ranking	0.75	0.73	0.75	0.50

➤ Task 1a Exact Match

	P	R	F ₁
Micro	0.09	0.05	0.07

➤ Task 1b

	P	R	F ₁
Micro	0.06	0.03	0.04

❖ Discussions

➤ Practical Issues

- Very high noise to signal ratio, correct RT difficult to identify even for humans
- OCR errors further exacerbate the problem, more so for neural models
- Facet classification doesn't favour rare class identification as distribution is very skew

➤ Future Directions

- Addressing some of the practical issues
- We haven't learned the α_i in the current approach but fixed it to empirically determined values which need to be trained jointly as parameters
- We only use TF-IDF and LCS, however in all the prior work other textual similarity features have been shown to give improvement, so those can be included to get better results

