What We've Got

- Pretty good DSE extraction results. [1].
- Better than [2]'s Agent extraction results.
- Strange and very bad Target extraction results.

DSE

We've tried OCLASS weight, drop out and used different word vector settings and we have similar(slightly better?) results than that in [1].

Agent

The agent extraction is great! We have better results on agent extraction than that in [2].

Target

1 Using the Original RNN Structure

We can get the best results from training with the original RNN by tuning the OCLASS weight fairly low, but in this case, the results on validation set can be better than the training set.

Example sentence when Oclass weight is low shows that the mistake usually happens when the RNN thinks a whole span

	OCLASS WEIGHT = 0.09		OCLASS WEIGHT = 0.25	
	Prop	Bin	Prop	Bin
P	0.274956	0.324164	0.375136	0.411554
R	0.702289	0.808863	0.324453	0.467323
F1	0.39519	0.462838	0.347958	0.437669

Table 1: *Target Extraction using RNN*

is a target but actually none of the part is included in the target. Sometimes the span do seem look like a named entity. The original results are in oclassLOWres.out.
All the parameter settings and the output format are in README. Also the full results when Oclass weight is high are in oclassHIGHres.out.

2 Merge the Input Labels

We tried to combine all the labels into one 7-dimensional vector. So the original output is the 3-dimensional vector indicating the probability of the label of a word being B, I or O. Now is the 7-dimensional vector indicating the probability being None, Begining of target/agent/dse or part of them.
Since there are overlapping of the tags in the MPQA database, we tried two different ways to generate the input we want. These and the results can be seen in Louis's report. But in general, the target extration is worse than the simpler way we tried above.

3 Multitasking

We also tried to do multitasking. The network structure is shown below. But the result is very bad. The good thing about this multitasking network is that we can assign different oclass weight to the three labels, but once these weights are assigned. My guess is there should be some extra tuning with the loss function regarding different label classes but every settings I tried just make things worse. Here's an example of the multitasking results. And the original one is in multi.out

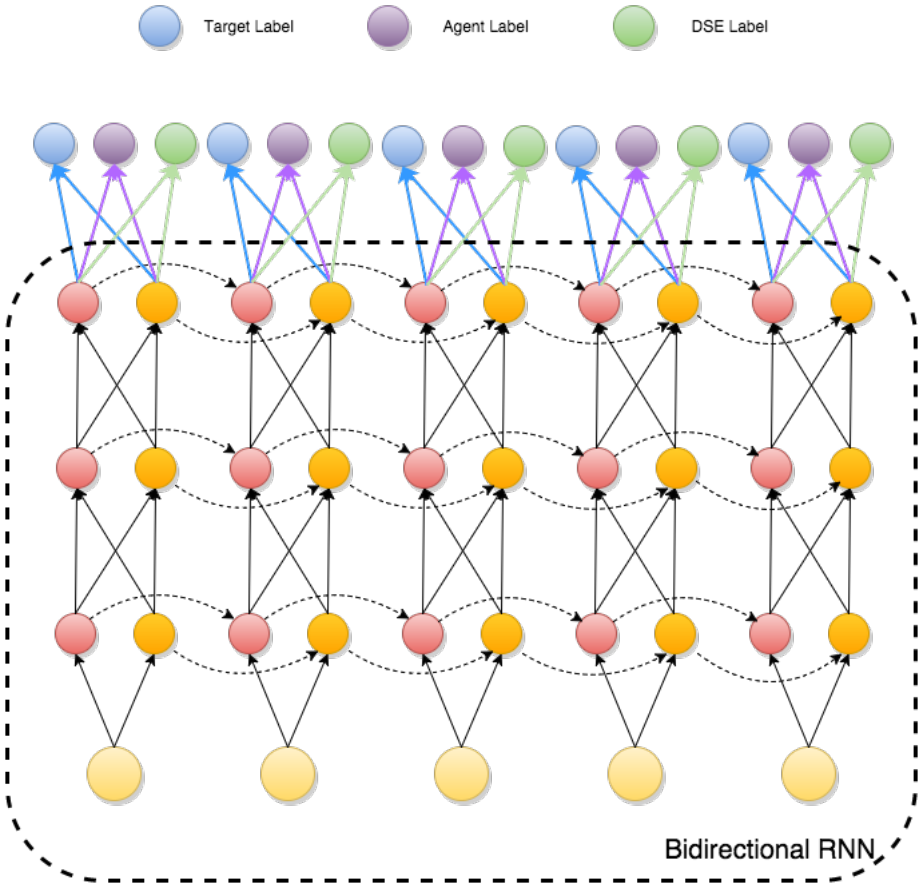

Figure 1: *Multitask RNN*

	Target Extraction		Agent Extraction		DSE Extraction	
	Prop	Bin	Prop	Bin	Prop	Bin
P	0.242472	0.242902	0.538143	0.54434	0.552941	0.552941
R	0.311916	0.727395	0.424234	0.535906	0.0304725	0.0342566
F1	0.272845	0.364189	0.474447	0.54009	0.0577618	0.0645161

Table 2: *Multitasking Results*

Bibliography

[1] Ozan Irsoy and Claire Cardie, Opinion Mining with Deep Recurrent Neural Networks, In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 720–728, 2014, Doha, Qatar, http://aclweb.org/anthology/D14-1080 1

[2] Bishan Yang and Claire Cardie, Joint inference for fine-grained opinion extraction, 2013, In *Proceedings of ACL* 1