

Independent Study Report

Vedant Puri

May 2019

1 Introduction

The task assigned to me was to perform supervised learning on scientific procedural texts. This is also known as shallow semantic parsing and in simple terms, my task was to implement code that learnt the label between a predicate and an argument in scientific texts from the training data and “successfully” and competitively predicted a label for a predicate argument pair in the testing/dev set.

2 Data

Helpfully enough, my mentor during this project, Sheshera Mysore, helped with the splitting of data into train, development and testing sets. The data was in the form of annotated sentences that can also been seen in a GUI at (ADD MIT LINK HERE).

- what data is Explain using example

- How big it is [sentences]

Data Split	# of sentences
Training	715
Development	211
Test	814

- How many predicates present

Data Split	# of predicates
Training	235
Development	127
Test	305

- Label Distribution

Label	Count
Recipe_Target	412
Recipe_Precursor	1029
Solvent_Material	524
Participant_Material	1955
Condition_Of	2041
Apparatus_Of	511
Atmospheric_Material	206

3 Task Definition

Mathematically define the task

4 Method

4.1 Model Description

Mathematically say model

4.2 Training

4.3 Hyper parameters

5 Experiments

6 Results

Train Loss graph

Dev f1 score graph

Comparing to baselines

Model	Precision	Recall	F1
Baseline 1			
Baseline 2			
My Method			

Label Wise metrics

Label	Precision	Recall	F1
Recipe_Target			
Recipe_Precursor			
Solvent_Material			
Participant_Material			
Condition_Of			
Apparatus_Of			
Atmospheric_Material			

Other Metrics

Model	Train Set	Dev Set	Test Set
Precision			
Recall			
F1			