Sequence to Sequence in SPARQL

Anand Panchbhai—Computer Science and Engineering IIT Bhilai, India, anandp@iitbhilai.ac.in

Abstract

With booming amount of information being continuously added to the internet orgaising the facts becomes a very difficult task, currently wikipedia hosts billions of such data points and corresponding relation in the form of RDFs.

RDF are a form of linked data and requiers the use of syntax strict language to extract information from such large corpuse. Today we use SPARQL to extract information from the database. Extracting data requires a query to be made in SPARQL and the response in another link that contains the information pertaining to the answer.

Today accesing data is difficult for lay user, who doesn't know how to make a query, this report tries to built upon a sytem built by — to further test the reaches of such system. The report also points out the issues faced during this research and suggest and request the author of the repository to consider them.

I. Introduction

This document contains the information regarding the warm up task for the project **A Neural QA Model for DBpedia** done by made and corresponding inference made in each case. Some of predicted reference results don't run properly on the portal (https://dbpedia.org/sparql).

The sections are divided into 5 sections:

1) Initial Impressions
2) Paper Survey
3) Download edit a sample template and train a Neural SPARQL Machine model on a DBpedia class.
4) Read the blog of last year's student to understand what has been achieved so far and reproduce the experiments.
5) My proposal and direction of the projects.

II. Initial Impressions

A. Issues faced and corresponding fixes

The repository had a few issues that i tried to fix by creating a pull request, as i worked there were more subtle mistakes in the repository that i will try to correct in the coming months. The pull request made by me can be found **here**.

While using the nmt submodule I came across a lot of errors, to overcome them I cloned the nmt repository separately and learnt how to use it. The results i obtained on doing so are mentioned in detail in the coming sections.

III. Paper Survey

From the initial survey i came across the following new poits that might help the project:

1) 1
2) 2
3) 3
4) 4

IV. Reproduce the experiments

On reproducing the experiment using the monuments and testing the results the performance recorded was as folows:

	perplexity	BLEU
dev	1.26	87.3
test	1.04	96.6

Note : *The whole model was divided into 3 sets: train, dev and test. The table above depects the results of the dev and the test set after 12000 iterations. Runtime log can be found at Click here.*

V. Download edit a sample template and train a Neural SPARQL Machine model on a DBpedia class.

On changing the template the final results obtained were as shown below, instead of monuments, **person** was used.

VI. Proposal and direction of work.

The direction which i would like to move on is : **Perform experiments on compositionality for complex QA and improve the performance of the current simple QA.** . Followed by train the NSpM model on DBpedia; evaluate the model against the QALD benchmark.;

1) A syntax verification and fixing system so that the predicted results always run.
2)