Milestone 4

Question Answering System

Pawan Patil (PXP180029)

Poonam Gillurkar (PXG180009)

Writeup Question 1.1: Write 1-2 paragraphs describing your chosen baseline. What

paper was it published in? What machine learning algorithm does it use? What kind of

preprocessing does it require? What are the features?

* This project uses NLP Toolkits such as SpaCy, NLTK to process the texts. “Toward a New Arabic Question Answering System” by Imane Lahbari, Said El Alaoui, and Khalid Zidani ( April 18, 2018 ) published in Journal of Information Technology, Vol. 15, No. 3A is the paper that shows on similar text processing of Arabic data.
* There can be multiple ways to preprocess text data. In this project we perform lemmatization and stop words removal. These techniques help us achieve data in the form the system needs to work on. Features that work for this project are Part of Speech Tagging, dependency parsing and WordNet Relation Extraction.

Writeup Question 1.2: Give step-by-step instructions for how to run your baseline

system. What libraries are needed? What commands are used to run it, with what

arguments, in what order?

* The instructions are:
  + Install SpaCy and load en\_core\_web\_sm
  + Download Elasticsearch from <https://www.elastic.co/downloads/elasticsearch>
  + Install Elastic Search with following command:
  + pip install elasticsearch
  + execute index.py file (this file requires elastic search for indexing)
  + Execute output.py file (to view the result)
* Spacy and nltk libraries provide the right supporting functions to implement the system.
* Commands to run are:
  + Python3 index.py
  + Python3 output.py

Writeup Question 2.1: Write 1-2 paragraphs describing the training process. What

learning scheme did you use (unsupervised, semi-supervised, or supervised)? Where there

any hyperparameters, and if so, how did you tune them? Did you run on a CPU or GPU?

How long did it take?

* As we are working on a test data, there is no training or validation dataset. The system is built to work on unsupervised data. Hence there are no parameters or hyperparameters to tune. Here GPU is not used. CPU does all the computing and processing.
* Time to execute the code:



Writeup Question 3.1: Give step-by-step instructions for how to run evaluate.py.

What arguments does it take?

* The test dataset used in our project doesn’t has any labels. The data contains of 30 documents and it’s in text form. So, even after loading the test data, as the system doesn’t have any labels to compare, the system does not learn from training data. As a result, the comparison of the outputs cannot be performed.

Writeup Question 3.2: Write 1-2 paragraphs describing the evaluation. What was the

performance of your baseline system? Is that good or bad? Why is the metric appropriate

for your task? If there are other metrics commonly used for the task, why did you choose

this one?

* Our project purely deals with text processing and document retrieval, which is unsupervised data for our system, so there isn’t any trained model to implement metric evaluation for the performance of the baseline system.