Natural Language Processing Final: Day 2

Date: 23rd Dec 2014

Time: 1:30 pm to 5:30 pm IST

Problem: For the information extraction system develop a Query Generator

In this project you will build a query generator that converts the data structure obtained from the relationship extractor to a MongoDb query-able data structure. This may be implemented in 2 stages:

1. In stage 1 convert the output of relationship extractor to a high level command structure
2. In stage 2 convert the high level command structure to a DB query using the functions provided.
3. The backend database has the following structure and schema. You don't need to access this directly but need to use the functions provided.

   Db Name = "devices"

   Collections

   brands

   products

   specs

4. You will be provided with web service functions to query the data. You don't need to write MongoDb queries to get the backend data but you will need to process the data from the web service that will be returned in JSON form.

Backend Database access functions:

Please try executing the functions below and observe the schema of the database:

ner.get_products("Samsung")

ner.get_brand_product_bigrams_dict()

get_spec(brand = "Samsung", product = "None")

get_spec(brand = "Samsung", product = "E1195")

You should use the above functions, NER and Relationship extractor to build the query generation and backend access. Please make the file names all lowercase letters as in Linux file names are case sensitive.

1. qgen.py – query generator
2. backend.py – backend query handling
3. After training the classifier create a pickle file with the name: rer.p

Upload all the above files using ner_client.upload()

Deliverables:

1. Source code of all your py modules (use ner_client.upload())

Best wishes from your faculty ☺