POS with NetworkX in articles

CYSHIELD

ADS Departments

Description

- Write a pipeline coding of part of speech (POS) tagging that can be represented by network graph .
- Do pre-processing on the data just to be able to run it.
- Do optimization time for article to represent every word in article .
- Feel free to use any technique can help to solve this problem .

✓ U can use Deep Learning Methods and show the diff and which one the best.(optional)

What to turn in:

1) Project documentation (details given below)
2) Report any resource you use
3) Your source code Jupyter Notebook with comment

POS with NetworkX in articles Use Case

Project documentation:

Please make sure that the documentation is easy to read and follow. Python output that is presented in the form of a screen dump with no formatting is not acceptable. Do not include code in the documentation. You must provide enough details so that the reader of your document can understand what you are trying to do and re-produce your results or something close to those results. Statements like "I've used a Weka filter", without specifying which filter you are referring to, are not acceptable. It will even be better, if you abstract away from implementation details and stick to the technicalities, so for example, if you've used an attribute selection filter configured to carry out feature selection using information gain, you can simply state that you used information gain for feature selection.

The structure of your document should be as follows:

- Introduction: State the problem and what you are trying to accomplish
- Data description: Provide a description of the used dataset (number of instances, features, division between training and testing, etc).
- Baseline experiments: Here you must state what the goal of the experiments is and after you have presented those experiments, what your conclusion is.
- Other experiments: each experiment must have a goal, a set of steps, results and a conclusion.
- Overall conclusion

You documentation should also include:

- A list of tools that you have used (you may use any tool or programming language you are comfortable with)
- A list of any external resources that you have used When including tables, make sure they have a caption that illustrates what you are presenting, if its f-score than state, if its accuracy than also state. Since the datasets you are using have been benchmarked, compare your result to previously obtained ones.

After you are done with your experiments answer the following questions:

1) What was the biggest challenge you faced when carrying out this project?
2) What do you think you have learned from the project?