Meeting - 5

19 February 2020 / 7:00PM/BIG DATA ANALYTICS LAB,CC3
Attendees

    Amit Kumar ,Project Instructor
    Harsh Goyal (IIT2018114)
    Aaditya Gadhave (IIT2018144)
    Sourabh Gupta (IIT2018149)
    Meet Singh Gambhir (IIT2018158)
    Tushar Atrey (IIT2018159)

Agenda

1. Main agenda of this meeting was B - T - B metapath
2. Introduction to stemming and lemmatization
3. Instructor introduces us to Word n-gram , character n-gram and topic modeling.
4. Introduction to NLTK toolkit

Notes

- The meeting started with the briefing of the work we did earlier
- The instructor told us how to implement the B - T - B metapath
- Instructor told us that in stemming we remove stop words (is ,are,and ,etc) and in lemmatization we convert all the verbs into present form.
- In word n-gram we took n words subarray from two different text files and compared them . Ex. in word 1-gram we took 1 word subarray from files.
- In character n-gram we took n characters subarray from two different text files and compared them . Ex. in character 3-gram we took 3 characters subarray from files.
- Instructor gave an introduction to topic modeling.
- Tasks to be completed until the next meeting were assigned

Action Items

1. Team is required to find the word n-grams between two different bugs
2. Team is required to find the character n-grams between two different bugs
3. Team is required to learn about the basics of topic modeling.