

Readme

Tools used:

- Stanford CoreNLP - to install visit <https://stanfordnlp.github.io/CoreNLP/index.html#download>
- Eclipse - IDE for Java code
- To use Stanford CoreNLP on Eclipse using Maven- [http://www.sfs.uni-tuebingen.de/~keberle/NLPTools/presentations/CoreNLP/NLP%20Tools%20-%20Stanford%20CoreNLP%20-%20Installation%20\(1\).pdf](http://www.sfs.uni-tuebingen.de/~keberle/NLPTools/presentations/CoreNLP/NLP%20Tools%20-%20Stanford%20CoreNLP%20-%20Installation%20(1).pdf)

Running the Project in Eclipse IDE:

Load App.java and run it(it contains the main method). Located at PA2/src/main/java/PA2NLP/PA2/App.java

Required files from folder: readFile.java, Preprocess.java, Merge.Java, Matrix.java, KNN.java, pom.xml

Dependency files:

```
<dependencies>
  <dependency>
    <groupId>junit</groupId>
    <artifactId>junit</artifactId>
    <version>3.8.1</version>
    <scope>test</scope>
  </dependency>
  <dependency>
    <groupId>edu.stanford.nlp</groupId> <artifactId>stanford-corenlp</artifactId>
    <version>3.8.0</version>
  </dependency>
  <dependency>
    <groupId>edu.stanford.nlp</groupId> <artifactId>stanford-corenlp</artifactId>
    <version>3.8.0</version>
    <classifier>models-english</classifier>
  </dependency>
  <dependency>
    <groupId>org.apache.cassandra</groupId>
    <artifactId>cassandra-all</artifactId>
    <version>0.8.1</version>
  </dependency>
  <!-- https://mvnrepository.com/artifact/ch.qos.logback/logback-classic -->
  <dependency>
    <groupId>ch.qos.logback</groupId>
    <artifactId>logback-classic</artifactId>
    <version>1.0.13</version>
    <scope>test</scope>
  </dependency>
```

PS: don't forget the log4j.properties file

Supplying New Parameters

Adding new files:

In class main in App.java there is a string called testFiles, add the name of the text file without the txt. Add the required file in the data folder PA2/data.

Changing parameter values

K value in KNN : Change parameter k to any integer greater than 0