

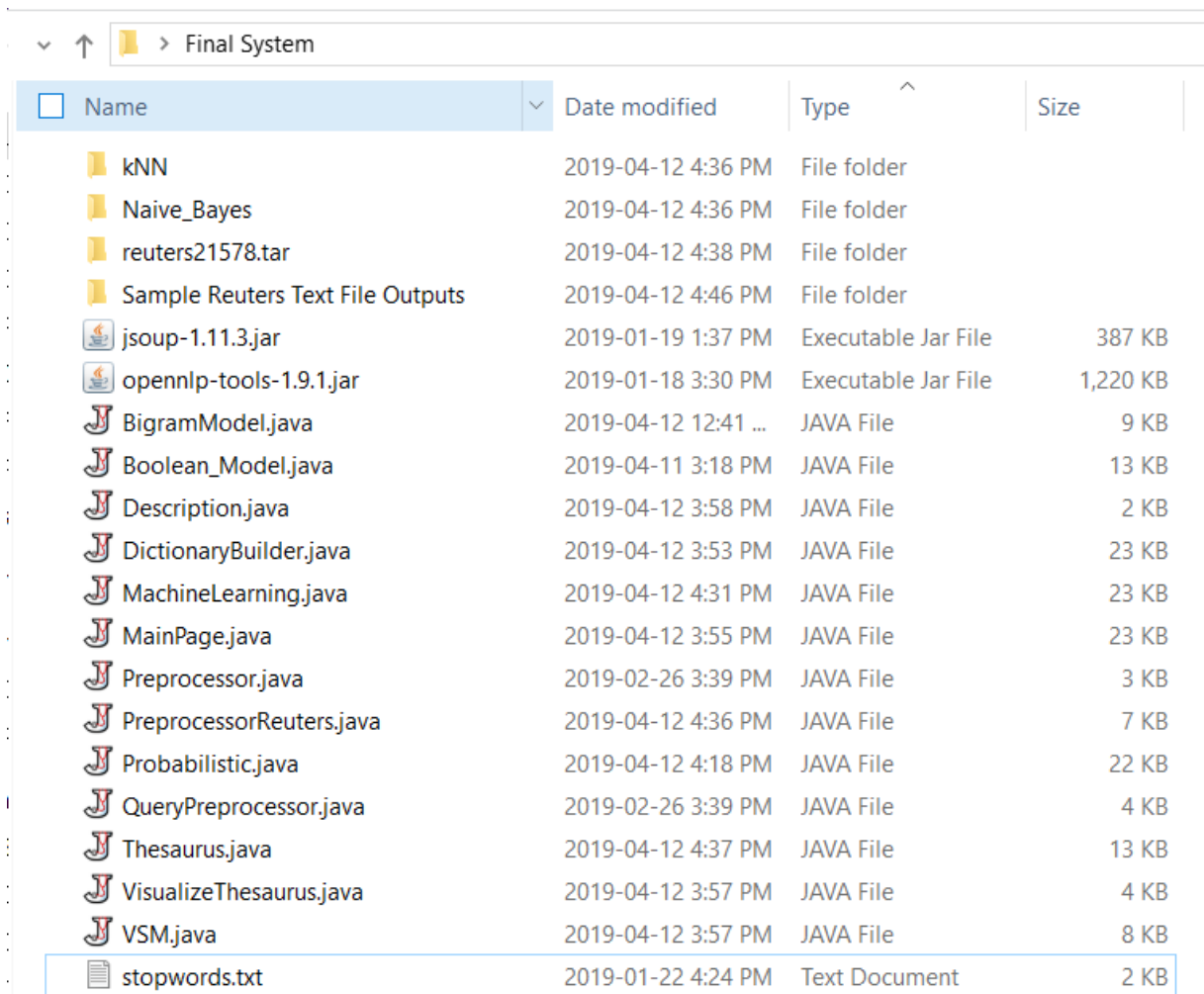
Instructions to run program with Command Prompt:

Student 1: Catherine DesOrmeaux, 7746139

Student 2: NamChi Nguyen, 7236760

Notes:

- For convenience, please have the JAR files and reuters21578.tar folder in the same folder as the Java classes as shown in the screenshot below. The JARs can be found here:
<https://www.dropbox.com/s/unwz0bqlcr2amhd/jsoup-1.11.3.jar?dl=0>
<https://www.dropbox.com/s/ddpw9bcawim7f4w/opennlp-tools-1.9.1.jar?dl=0>
- We have provided the required external text files (for kNN and Naïve_Bayes) after assigning topics and building the dictionary, so the preprocessing/machine learning step can be skipped. But, PreprocessorReuters will require the reuters21578.tar folder in the same directory to compile and start the program without error.
- We have only tested the code on Windows 10 using Eclipse with Java SE 8 and up



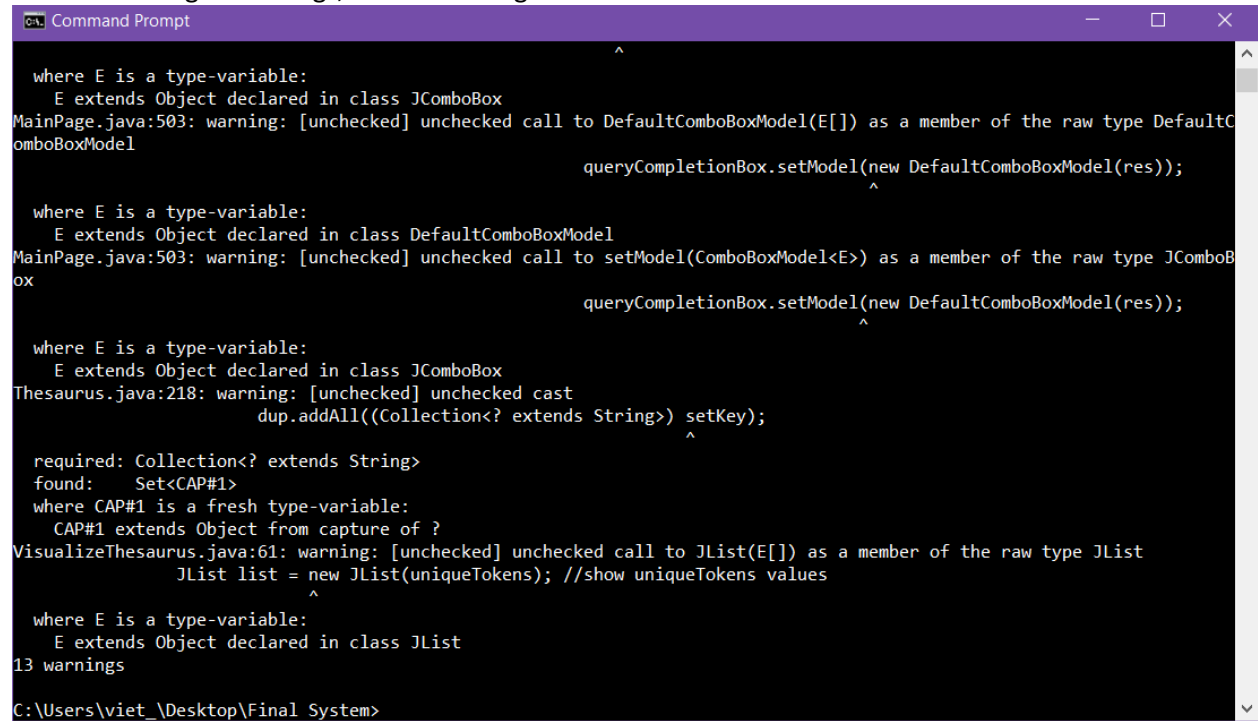
The screenshot shows a Windows File Explorer window with the address bar displaying 'Final System'. The file list includes folders for 'kNN', 'Naive_Bayes', 'reuters21578.tar', and 'Sample Reuters Text File Outputs'. It also lists two JAR files, 'jsoup-1.11.3.jar' and 'opennlp-tools-1.9.1.jar', and a series of Java files including 'BigramModel.java', 'Boolean_Model.java', 'Description.java', 'DictionaryBuilder.java', 'MachineLearning.java', 'MainPage.java', 'Preprocessor.java', 'PreprocessorReuters.java', 'Probabilistic.java', 'QueryPreprocessor.java', 'Thesaurus.java', 'VisualizeThesaurus.java', and 'VSM.java'. A text file named 'stopwords.txt' is at the bottom of the list.

<input type="checkbox"/> Name	Date modified	Type	Size
kNN	2019-04-12 4:36 PM	File folder	
Naive_Bayes	2019-04-12 4:36 PM	File folder	
reuters21578.tar	2019-04-12 4:38 PM	File folder	
Sample Reuters Text File Outputs	2019-04-12 4:46 PM	File folder	
jsoup-1.11.3.jar	2019-01-19 1:37 PM	Executable Jar File	387 KB
opennlp-tools-1.9.1.jar	2019-01-18 3:30 PM	Executable Jar File	1,220 KB
BigramModel.java	2019-04-12 12:41 ...	JAVA File	9 KB
Boolean_Model.java	2019-04-11 3:18 PM	JAVA File	13 KB
Description.java	2019-04-12 3:58 PM	JAVA File	2 KB
DictionaryBuilder.java	2019-04-12 3:53 PM	JAVA File	23 KB
MachineLearning.java	2019-04-12 4:31 PM	JAVA File	23 KB
MainPage.java	2019-04-12 3:55 PM	JAVA File	23 KB
Preprocessor.java	2019-02-26 3:39 PM	JAVA File	3 KB
PreprocessorReuters.java	2019-04-12 4:36 PM	JAVA File	7 KB
Probabilistic.java	2019-04-12 4:18 PM	JAVA File	22 KB
QueryPreprocessor.java	2019-02-26 3:39 PM	JAVA File	4 KB
Thesaurus.java	2019-04-12 4:37 PM	JAVA File	13 KB
VisualizeThesaurus.java	2019-04-12 3:57 PM	JAVA File	4 KB
VSM.java	2019-04-12 3:57 PM	JAVA File	8 KB
stopwords.txt	2019-01-22 4:24 PM	Text Document	2 KB

1. Compile classes:

`javac -cp opennlp-tools-1.9.1.jar;jsoup-1.11.3.jar -Xlint:unchecked *.java`

Note: You will get warnings, but it can be ignored.



```
C:\ Command Prompt

where E is a type-variable:
  E extends Object declared in class JComboBox
MainPage.java:503: warning: [unchecked] unchecked call to DefaultComboBoxModel(E[]) as a member of the raw type DefaultC
omboBoxModel
                                ^
                                queryCompletionBox.setModel(new DefaultComboBoxModel(res));


























where E is a type-variable:
  E extends Object declared in class DefaultComboBoxModel
MainPage.java:503: warning: [unchecked] unchecked call to setModel(ComboBoxModel<E>) as a member of the raw type JComboB
ox
                                ^
                                queryCompletionBox.setModel(new DefaultComboBoxModel(res));

where E is a type-variable:
  E extends Object declared in class JComboBox
Thesaurus.java:218: warning: [unchecked] unchecked cast
                                dup.addAll((Collection<? extends String>) setKey);
                                ^
required: Collection<? extends String>
found:    Set<CAP#1>
where CAP#1 is a fresh type-variable:
  CAP#1 extends Object from capture of ?
VisualizeThesaurus.java:61: warning: [unchecked] unchecked call to JList(E[]) as a member of the raw type JList
                                JList list = new JList(uniqueTokens); //show uniqueTokens values
                                ^

where E is a type-variable:
  E extends Object declared in class JList
13 warnings

C:\Users\viet_\Desktop\Final System>
```

Classes should now be compiled in folder:

Final System			
<input type="checkbox"/> Name	Date modified	Type	Size
 kNN	2019-04-12 4:36 PM	File folder	
 Naive_Bayes	2019-04-12 4:36 PM	File folder	
<input checked="" type="checkbox"/>  reuters21578.tar	2019-04-12 4:38 PM	File folder	
 Sample Reuters Text File Outputs	2019-04-12 4:46 PM	File folder	
 opennlp-tools-1.9.1.jar	2019-01-18 3:30 PM	Executable Jar File	1,220 KB
 jsoup-1.11.3.jar	2019-01-19 1:37 PM	Executable Jar File	387 KB
 stopwords.txt	2019-01-22 4:24 PM	Text Document	2 KB
 Preprocessor.java	2019-02-26 3:39 PM	JAVA File	3 KB
 QueryPreprocessor.java	2019-02-26 3:39 PM	JAVA File	4 KB
 Boolean_Model.java	2019-04-11 3:18 PM	JAVA File	13 KB
 BigramModel.java	2019-04-12 12:41 ...	JAVA File	9 KB
 DictionaryBuilder.java	2019-04-12 3:53 PM	JAVA File	23 KB
 MainPage.java	2019-04-12 3:55 PM	JAVA File	23 KB
 VSM.java	2019-04-12 3:57 PM	JAVA File	8 KB
 VisualizeThesaurus.java	2019-04-12 3:57 PM	JAVA File	4 KB
 Description.java	2019-04-12 3:58 PM	JAVA File	2 KB
 Probabilistic.java	2019-04-12 4:18 PM	JAVA File	22 KB
 MachineLearning.java	2019-04-12 4:31 PM	JAVA File	23 KB
 PreprocessorReuters.java	2019-04-12 4:36 PM	JAVA File	7 KB
 Thesaurus.java	2019-04-12 4:37 PM	JAVA File	13 KB
 BigramModel.class	2019-04-12 4:55 PM	CLASS File	7 KB
 Boolean_Model.class	2019-04-12 4:55 PM	CLASS File	7 KB
 Description\$1.class	2019-04-12 4:55 PM	CLASS File	1 KB
 Description.class	2019-04-12 4:55 PM	CLASS File	2 KB
 DictionaryBuilder.class	2019-04-12 4:55 PM	CLASS File	16 KB

2. Run MainPage:

```
java -cp .;opennlp-tools-1.9.1.jar.;jsoup-1.11.3.jar MainPage
```

Note: It will take a few minutes for the search engine to pop up

