An N-Gram Frequency Builder

Introduction

When the application is run on Runner.java, the user is given a prompt to select an option from 1-5. Option 1 prompts the user to "Specify Text File Directory". Option 2 prompts the user to "Specify n-Gram Size". Option 3 prompts the user to " Specify Output File Location for file ngramOutputG00411386.csv to be saved ". Option 4 creates the csv file and option 5 quits the application.

Features

nGramFreqOutputCreation – Writes the n-gram frequency output to the specified output file location. For example, if the user inputted the output file location to be "./textfiles". The outputted file will be saved to "./textfiles/ngramOutputG00411386.csv".

numberOfChunksArray – Removes all the white space between characters, changes all text to lower case and removes all special characters in the text file. Splits the string into chunks of the n-gram frequency. HashMap is used to identify the frequency of each chunk.

readStringData – Reads the text file from the specified file directory. A string is returned of the text file.

getTextDirectory – Prompts the user to "Specify Text File Directory" when option 1 is selected. The input is stored to variable filePath.

getSize – Prompts the user to "Specify n-Gram Size" when option 2 is selected. The input is stored to variable nGramFreqNumber. Method readStringData is executed with parameter filePath. It returns a string of the textfile. Method numberOfChunksArray is executed to return the output of the n-gram frequency.

getOutputFile - – Prompts the user to "Specify Output File Location for file ngramOutputG00411386.csv to be stored " when option 3 is selected. The input is stored to variable outputPath.

build - Prompts the user that the n-gram frequency application is being built when option 4 is selected. Method nGramFreqOutputCreation is executed which creates the csv file in the specified output file path.

Start – Starts application.