

**Gebze Technical University
Computer Engineering**

CSE 222 - 2018 Spring

HOMEWORK 6 REPORT

**AHMET DENİZLİ
161044020**

Course Assistant: Ayşe Şerbetçi Turan

1 INTRODUCTION

1.1 Problem Definition

In this homework we implementing two different HashMap classes to perform basic Natural Language Processing operations, which was Bi-grams and TFIDF. Using HashMap for this structure for efficient. Word_Map has Node table, which Node has word and File_Map values. File_Map has ArrayLists of filenames and ArrayLists of List of indexes. File_Map has ArrayLists of filenames and ArrayLists of List of indexes.

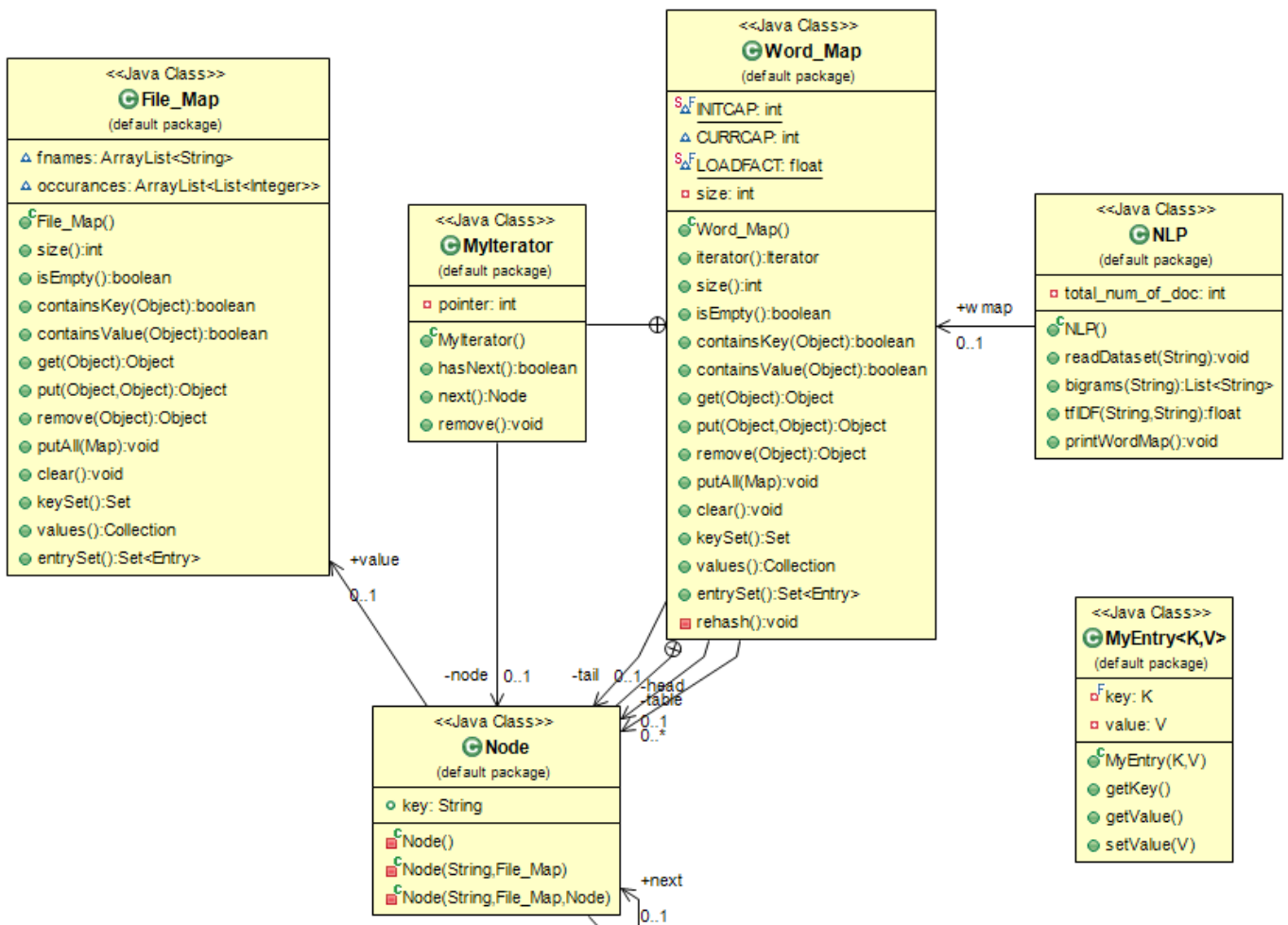
1.2 System Requirements

JDK 1.8 or greater release must be installed to compile this program. The compiled version of program will work any jvm installed environment.

Program will work at least 64 MB of physical RAM and 98 MB of free disk space for jvm.

2 METHOD

2.1 Class Diagrams



2.2 Methods Complexities

File_Map Class :

Size() -> O(1);
isEmpty() -> O(1);
containsKey () -> O(n);
containsValue () -> O(n);
get() -> O(n);
put() -> O(n);
clear() -> O(1);

Word_Map Class :

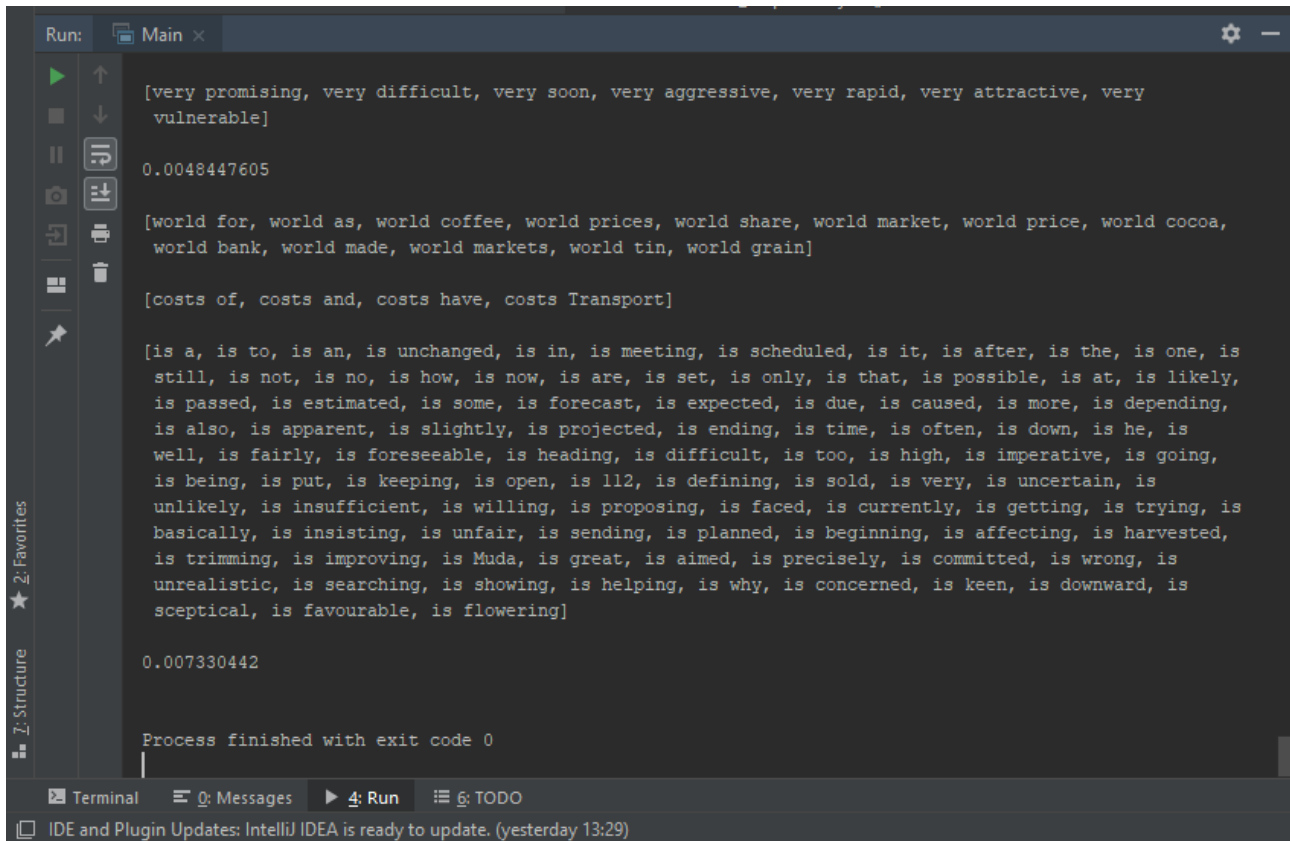
Size() -> O(1);
isEmpty() -> O(1);
containsKey () -> O(1);
containsValue () -> O(1);
get() -> O(1);
put() -> O(1);
clear() -> O(1);
rehash()->O(n)

3 RESULT

3.1 Test Cases

Each class methods tested inside main with given dataset files. All errors handled with exceptions.

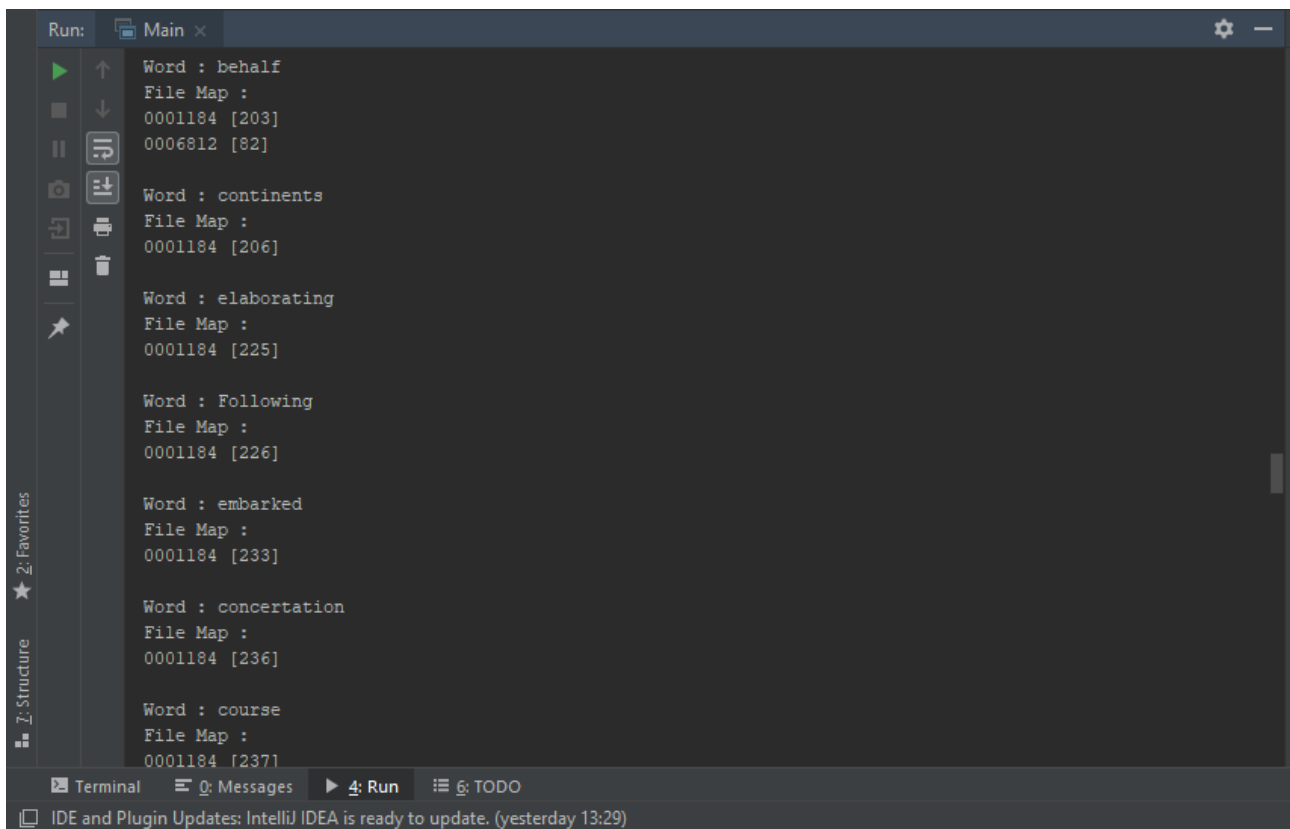
3.2 Running Results



The screenshot shows the IntelliJ IDEA Run window for a test named 'Main'. The output is as follows:

```
[very promising, very difficult, very soon, very aggressive, very rapid, very attractive, very vulnerable]
0.0048447605
[world for, world as, world coffee, world prices, world share, world market, world price, world cocoa, world bank, world made, world markets, world tin, world grain]
[costs of, costs and, costs have, costs Transport]
[is a, is to, is an, is unchanged, is in, is meeting, is scheduled, is it, is after, is the, is one, is still, is not, is no, is how, is now, is are, is set, is only, is that, is possible, is at, is likely, is passed, is estimated, is some, is forecast, is expected, is due, is caused, is more, is depending, is also, is apparent, is slightly, is projected, is ending, is time, is often, is down, is he, is well, is fairly, is foreseeable, is heading, is difficult, is too, is high, is imperative, is going, is being, is put, is keeping, is open, is ll2, is defining, is sold, is very, is uncertain, is unlikely, is insufficient, is willing, is proposing, is faced, is currently, is getting, is trying, is basically, is insisting, is unfair, is sending, is planned, is beginning, is affecting, is harvested, is trimming, is improving, is Muda, is great, is aimed, is precisely, is committed, is wrong, is unrealistic, is searching, is showing, is helping, is why, is concerned, is keen, is downward, is sceptical, is favourable, is flowering]
0.007330442
Process finished with exit code 0
```

The bottom of the window shows the IDE status bar with tabs for Terminal, Messages, Run, and TODO. A notification at the bottom states: 'IDE and Plugin Updates: IntelliJ IDEA is ready to update. (yesterday 13:29)'.



The screenshot shows the IntelliJ IDEA Run window for a test named 'Main'. The output is as follows:

```
Word : behalf
File Map :
0001184 [203]
0006812 [82]

Word : continents
File Map :
0001184 [206]

Word : elaborating
File Map :
0001184 [225]

Word : Following
File Map :
0001184 [226]

Word : embarked
File Map :
0001184 [233]

Word : concertation
File Map :
0001184 [236]

Word : course
File Map :
0001184 [237]
```

The bottom of the window shows the IDE status bar with tabs for Terminal, Messages, Run, and TODO. A notification at the bottom states: 'IDE and Plugin Updates: IntelliJ IDEA is ready to update. (yesterday 13:29)'.