RESUMEPEDIA PROGRAM Version (1.0)

----------------------------------------------------------------------------------------------------------------------------

Resumepedia is an Information Retrieval and Classification application. It uses LUCENE API of Java to make the query search faster and more efficient using Inverted Indexing and ranked retrieval that uses tf-idf weighting mechanism. The classification of resumes is done using Naive Bayes Classification. The resumes are classified on the basis of the type of resume they belong to like Computers, Management, Lawyers and Miscellaneous. Currently classification and information retrieval are independent of each other.

----------------------------------------------------------------------------------------------------------------------------

FILES ATTACHED

---------------------------------------------------------------------------------------------------------------------------

InformationReterival.java

NBClassifier.java

Web Folder

Data Folder(for the Classifier Training and Testing)

FTI(Documents for information retrieval)

IT(Indexing Directory)

lucene-core-2.3.0.jar

lucene-demos-2.3.0.jar

----------------------------------------------------------------------------------------------------------------------------

INSTALLATION NOTES

----------------------------------------------------------------------------------------------------------------------------

Resumepedia is supported by all the popular Operating Systems like Windows (XP and Above), MacOs and Linux (All Versions). Resumepedia can be tested on any popular IDEs like Eclipse, NetBeans or J- Grasp.

We have included the notes on how to run Resumepedia on Eclipse.

The Resumepedia Program makes use of the LUCENE API of Java. One can download this API from

<https://lucene.apache.org/core/downloads.html>.

Apache LUCENE is a free open source information retrieval library written in Java. While suitable for any application that requires full text indexing and search capability.

This program makes use of the LUCENE 2.3.0.

Follow the steps to get the LUCENE jar files in the Eclipse.

Step 1

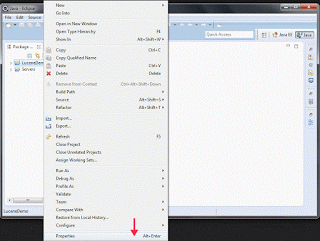
After making a Java project named RESUMEPEDIA in the Eclipse. Add 2 classes to it naming InformationRetrieval.java and NBClassifier.java.

Step 2

Open the InformationRetrieval.java on the Eclipse.

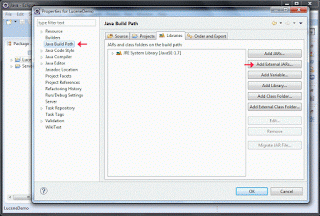
Step 3

Right click on the project (RESUMEPEDIA) and select “Properties” as shown below:



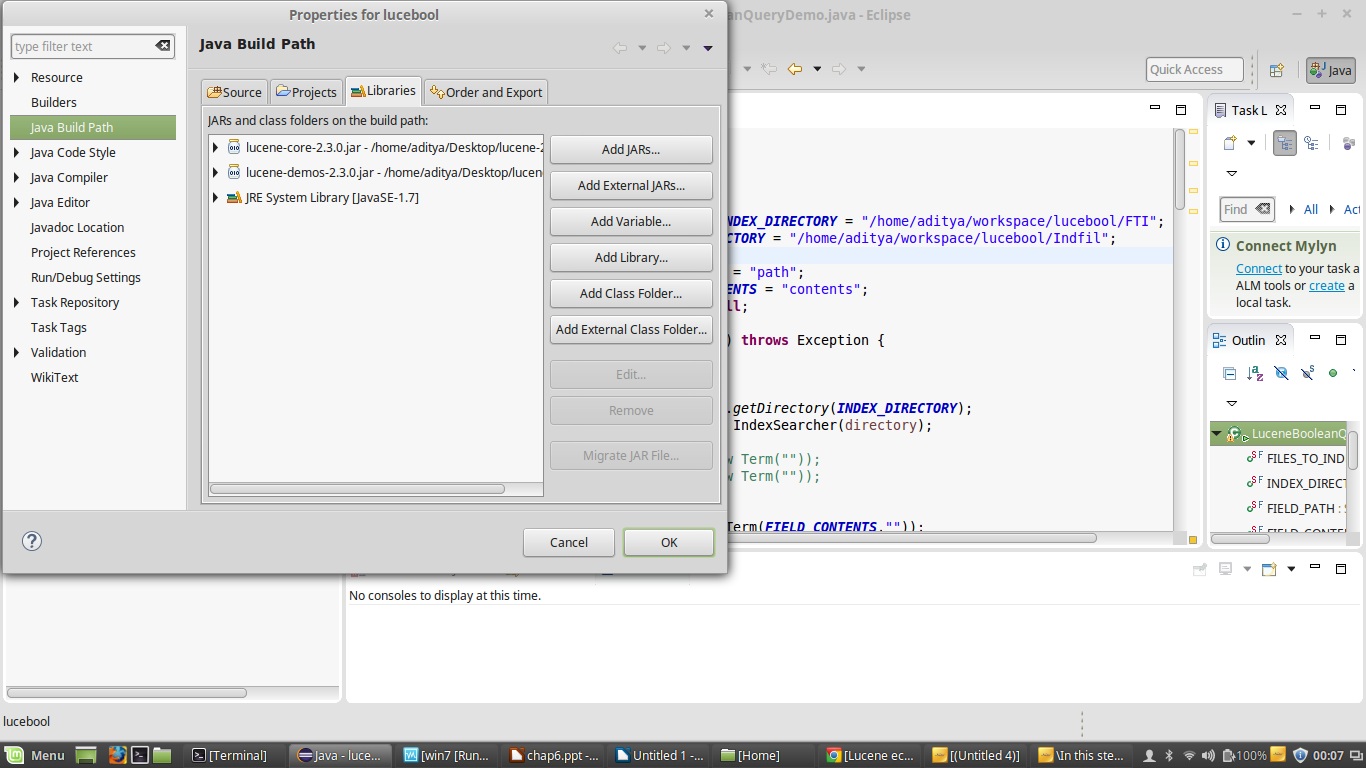
Step 4

Click on “Java Build Path” on the left side and select “Add External JARs…” button on the right as shown below:



Step 5

In this step, we will be adding 2 JAR files to our project (RESUMEPEDIA). The JAR files have been attached with program. After adding the external JAR files, “Java Build Box” will look like this



Step 6

Run the InformationRetrieval.java and give the query for which results are needed.

NOTE: Mac users use the following code to delete .DS\_Store file from your system, which sometimes gives nullpointerException

sudo find / -name ".DS\_Store" -depth -exec rm {} \;

Step 7

Delete the contents of the folder IT and run the program IR again to clean the cache and the previously stored data.

Step 8

The classifier does not have any dependencies except that on the data. Open the classifier program like a regular .java file. Make sure the path of the train and test folder in the main class is changed as the path on the respected local server.

Step 9

The hyperlink generation is not automatic and is manipulated and hence necessary changes will be required. Eg. As the path being returned was not the true path the 1st 62 chars of the string are being deleted to get the proper path in the hyperlink. So on the basis of the location of the file on the sever the path for the hyperlink will have to be modified.

Step 10

Run the entire project as a web application and provide input.

–---------------------------------------------------------------------------------------------------------------------------RESUMEPEDIA CAN BE REACHED AT, in case of any queries or feedback.

Aditya Bisht – [axb9652@rit.edu](mailto:axb9652@rit.edu)

Karan Kothari – [kdk2612@rit.edu](mailto:kdk2612@rit.edu)

Maitri Tolia – [mjt5429@rit.edu](mailto:mjt5429@rit.edu)

----------------------------------------------------------------------------------------------------------------------------

LICENSE INFORMATION

----------------------------------------------------------------------------------------------------------------------------THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS

"AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT

LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS

FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE

COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT,

INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING,

BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES;

LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER

CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT

LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN

ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE

POSSIBILITY OF SUCH DAMAGE.