Preeti Harkanth - 800845749

Prakash Wagle - 800889950

Siva Krishna Sirigineedi - 800887277

Vineet Gahlout - 800866880

Project Report

ITCS 6162 KNOWLEDGE DISCOVERY IN DATABASES (KDD)

Implementing Action Reducts & Action Rules

**Concepts**

**Action:** Controlling or changing some of the attribute values in an information system to achieve desired result is generally defined as Action. We can decide on how and which attribute to be changed using the concept action reduct and use the action reduct to formulate necessary actions

The action reduct has following properties:

* They are obtained from objects which have favorable decision values.
* Mostly has distinct set of values not found in the other group, the group not having favorable decision values.
* It is the minimal set of differences between separate groups of objects.

There are also two types of attributes

Stable- An attribute is called stable if the values assigned to objects do not change over time. Eg. Area of a Country

Flexible- An attribute is called Flexible if the values assigned to objects changes over time.

Eg. GDP of a Country

Advantage of using this algorithm is that we do not have to calculate Support & Confidence.

**Input:**

1. Downloaded dataset in excel sheet from <http://fsi.fundforpeace.org/rankings-2014> and named it as **project dataset.**
2. Downloaded population of 2014 of all countries from <http://data.worldbank.org/indicator/SP.POP.TOTL> format)
3. Discretize the flexible values integers by replacing n.k by n+1 if k>5 and by n if k < 6. Since decimal is only one place n.5, n.4, n.3, n.2, n.1 is replaced by n .n.6, n.7, n.8, n.9 will be replaced by n+1
4. Inserted population column in dataset and named it as **project dataset with population**
5. After round off I saved it in project dataset with population AND ROUND OFF.xls,

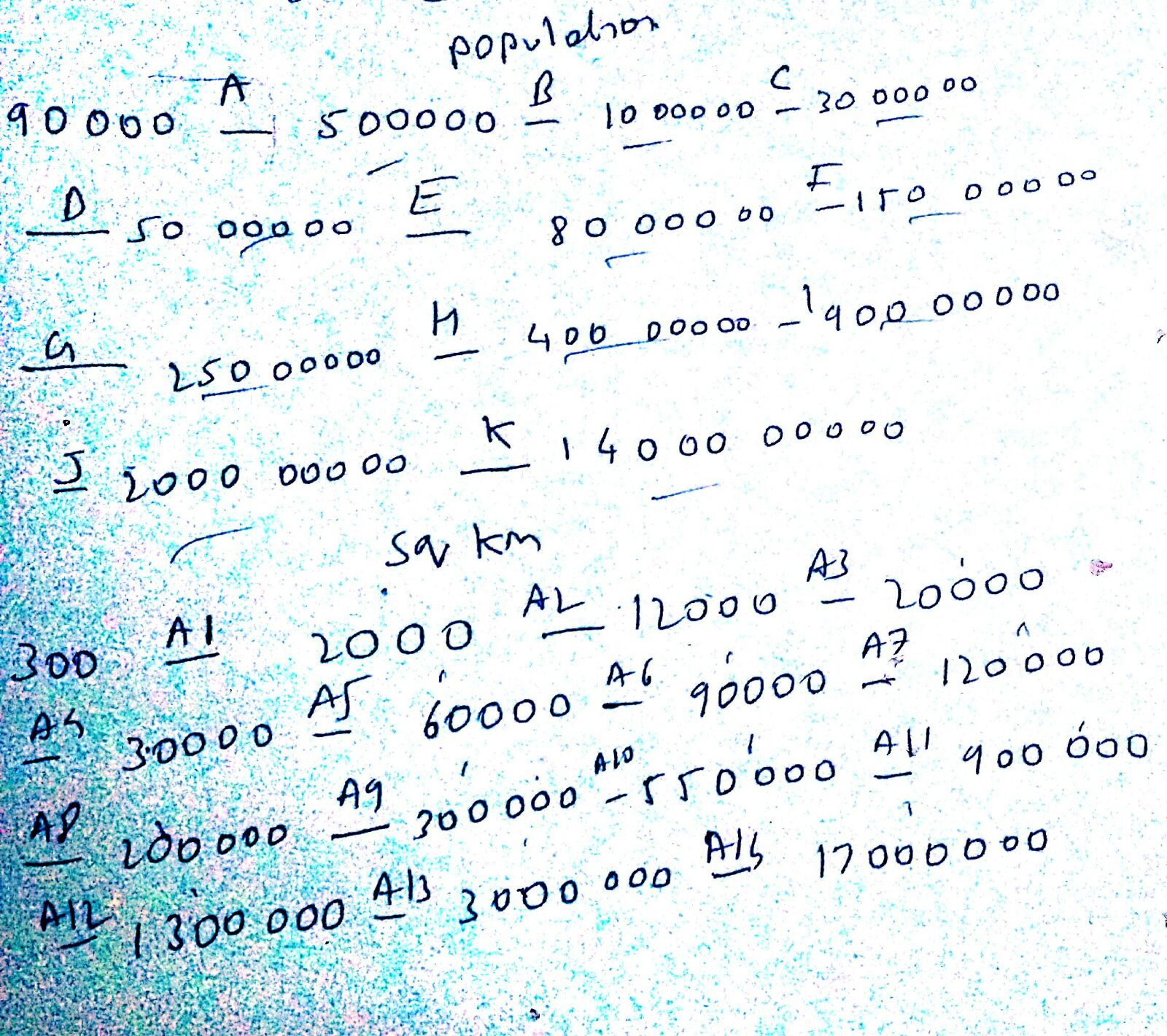
**Note: after rounding index values average are again calculated, total value is also rounded off from original dataset.**

1. Downloaded area of all countries from <http://data.worldbank.org/indicator/AG.LND.TOTL.K2>
2. Inserted area column in the dataset according to the countries.
3. POPULATION DISCRETIZATION FORMULA

**=IF(T2<500000,"1",IF(T2<1000000,"2",IF(T2<3000000,"3",IF(T2<5000000,"4",IF(T2<8000000,"5",IF(T2<15000000,"6",IF(T2<25000000,"7",IF(T2<40000000,"8",IF(T2<90000000,"9",IF(T2<200000000,"10",IF(T2<1400000000,"11")))))))))))**

1. SQKM DISCRETIZATION FORMULA

**=IF(S2<2000,"1",IF(S2<12000,"2",IF(S2<20000,"3",IF(S2<30000,"4",IF(S2<60000,"5",IF(S2<90000,"6",IF(S2<120000,"7",IF(S2<200000,"8",IF(S2<300000,"9",IF(S2<550000,"10",IF(S2<900000,"11",IF(S2<1300000,"12",IF(S2<3000000,"13",IF(S2<17000000,"14"))))))))))))))**

****

For sqkm A1,A2,A3,A4… are replaced with numbers 1,2,3,4,5,6,7,8,…

For population A,B.C,D,E… are replaced with numbers 1,2,3,4,5,6…

1. PROJECT DTATASET FINAL VER 4.0 has been saved
2. **Country name, Population, and Area are stable in our dataset**
3. The excel file has been saved as csv files with no header ,but just with column names and saved as **project.csv**

**Output:**

This program finds all the specialized rules from action reducts.

Example: (B:b2->B:b1)(C:c2->C:c2)=>(d1->d2)

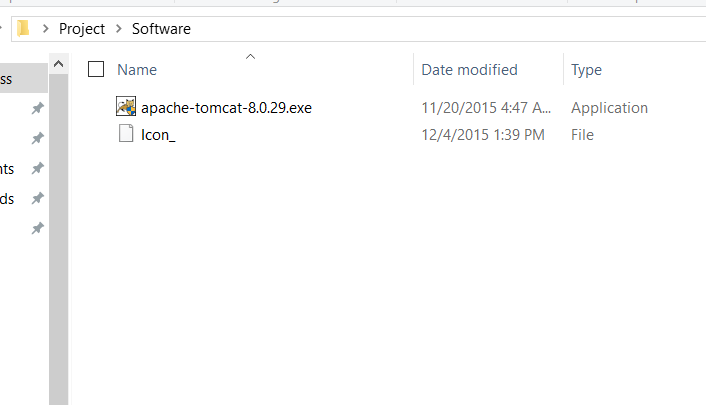
Example: (DemographicPressure:10) --> (DemographicPressure: 3)(Refugees:10) --> (Refugees : 5) ==> Very High Alert --> Stable

**Algorithm:**

The aim of developing this algorithm is to provide a list of attribute values which can help to make changes in few of the objects to convert the unfavorable decision value to more favorable value. To achieve this we use reduct which is a minimal set of attributes which help in distinguishing an object from all other objects in the information system.

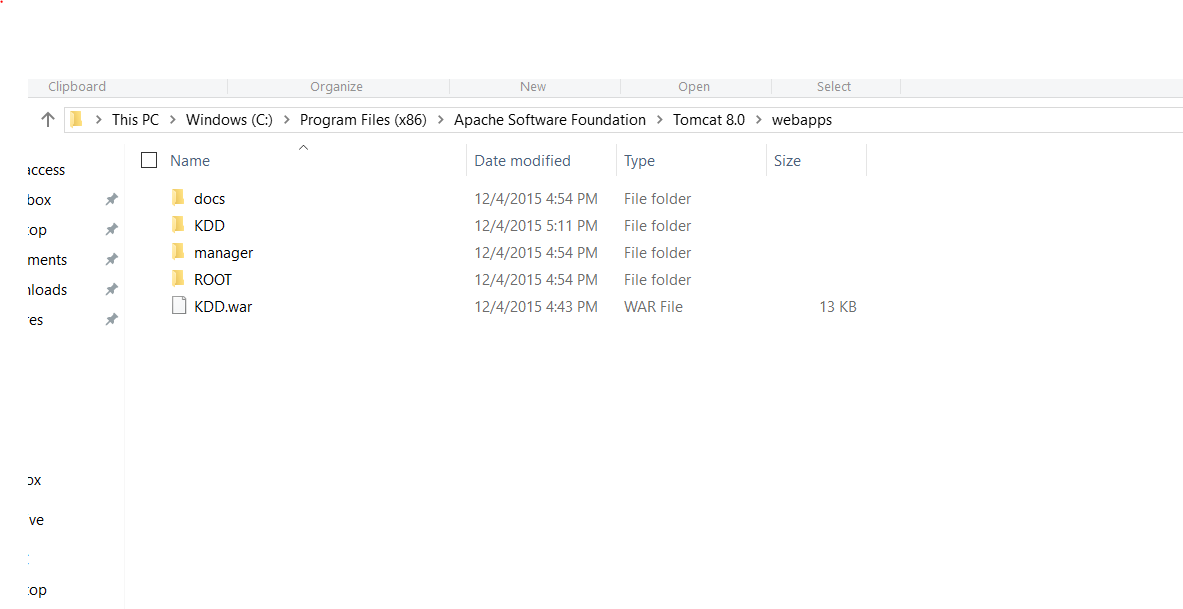
**Implementation Instructions:**

1. Make sure JRE and Apache Tomcat is installed on the PC.(ApacheTomcat is in software folder)

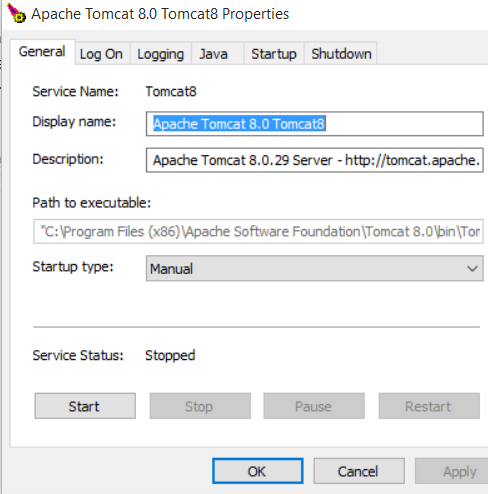


1. Input Dataset is present in the Input Folder.(*filepath***\Project\Input\project.csv**)
2. Download the KDD.war file and place it in webapps folder inside Apache Tomcat

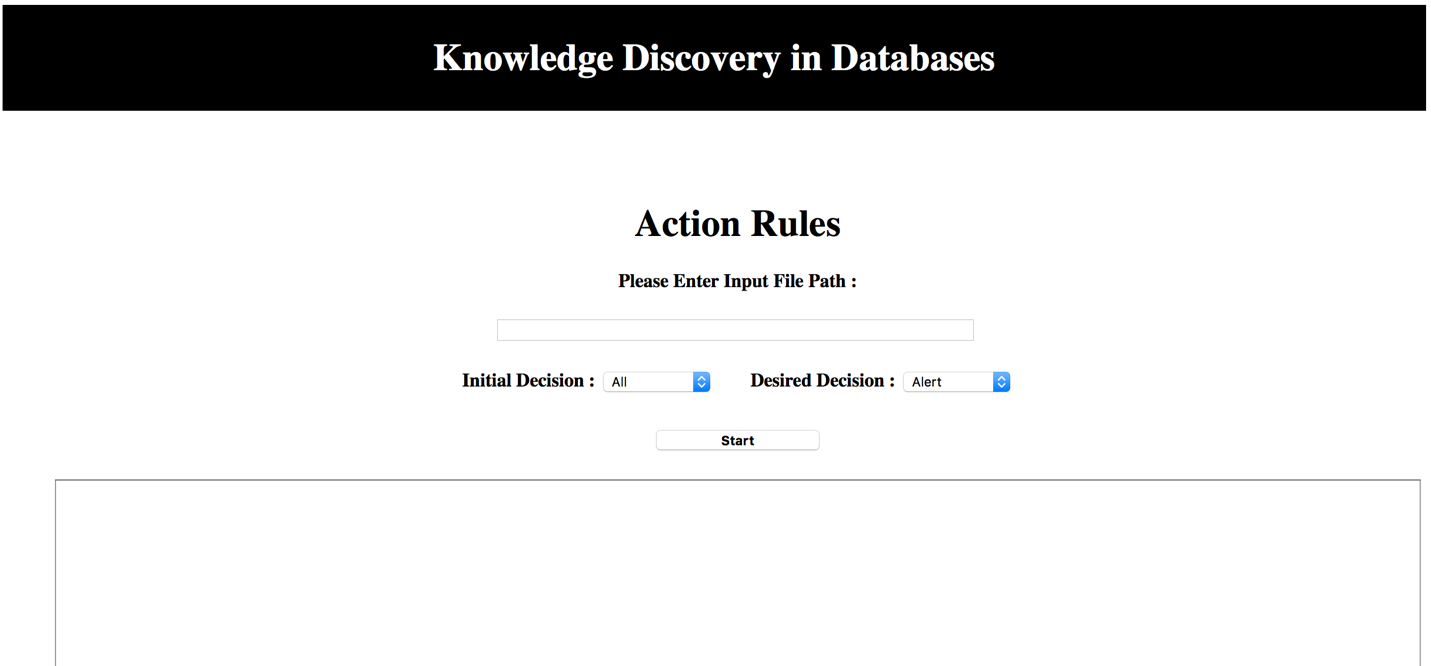
(Eg.C:\Program Files (x86)\Apache Software Foundation\Tomcat 8.0\webapps)



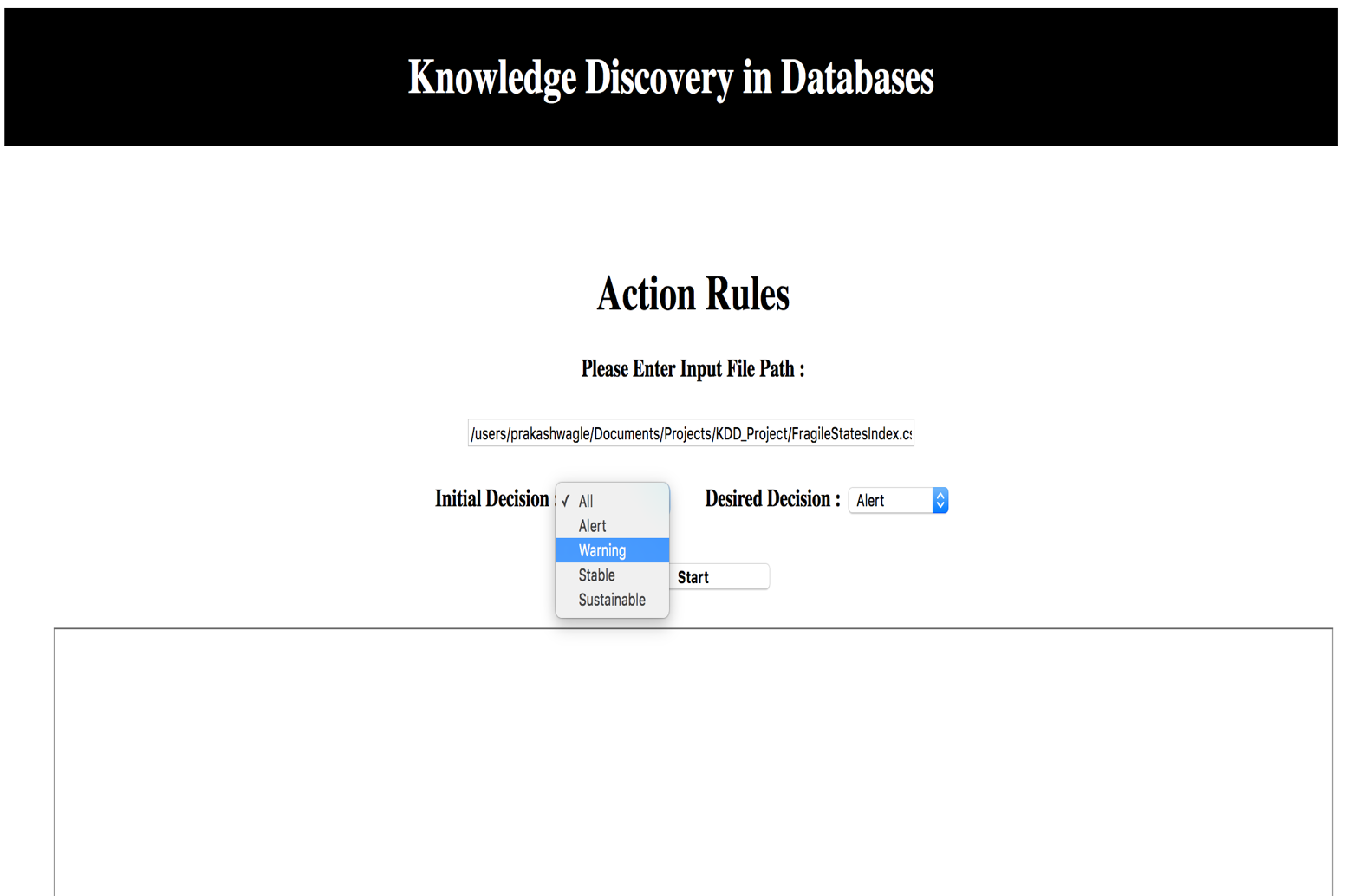
1. Start the Apache Tomcat Server



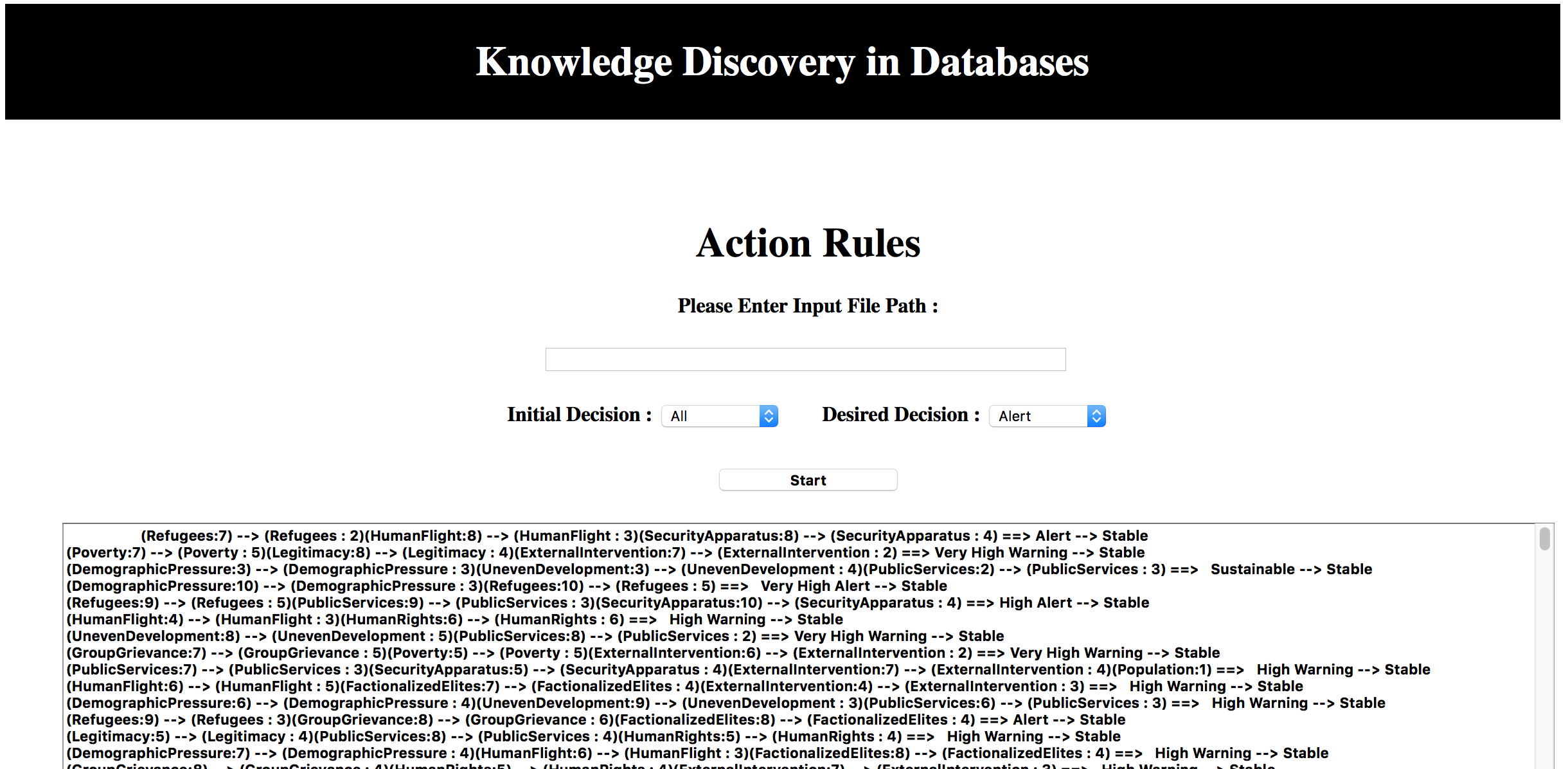
1. Open Browser type **localhost:8080/KDD/** Following Screen will open in your browser

****

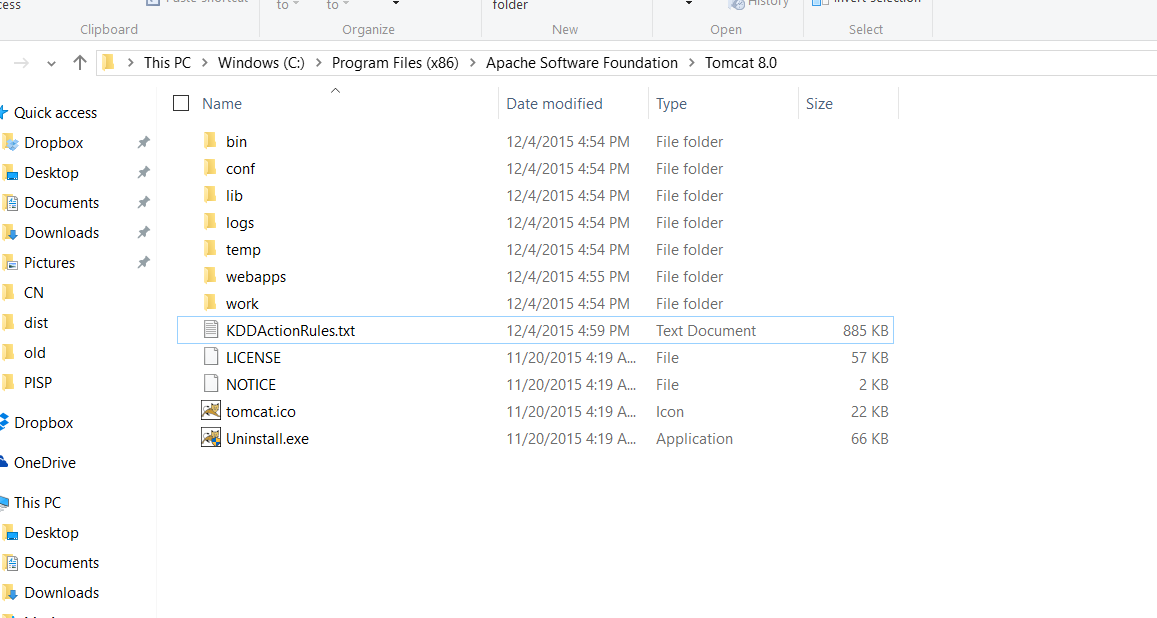
1. Please enter Input File Path as shown below and Select Initial & Desired Decision Value from the drop down menu.

****

1. Press Start Button and **wait for 30 Sec.** Result will be displayed as follow.



1. Also the output will be written in a file called **KDDActionRules.txt** which will be created in apache-tomcat Folder



**References**

1. <http://webpages.uncc.edu/ras/Im-Ras-Tsay-2011.pdf>
2. http://webpages.uncc.edu/ras/KDD-Fall.html