# Source Code:

**#include<bits/stdc++.h>**

**using namespace std;**

**int main()**

**{ map< int , int > hm;**

**int n,k;**

**double maxentropychange,t,test;**

**cout<<"\n\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* OUTLIER DETECTION USING GREEDY ALGORITHM \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\n\n";**

**cout<<"\t1.Enter the size of the database (number of vectors) : ";**

**cin>>n;**

**vector<int> database;**

**int index,gg;**

**for(int i=0;i<n;i++)**

**{ int temp;**

**cout<<"\n Enter the enter the value of first attribute for V"<<i+1<<" Of the database :";**

**cin>>temp;**

**database.push\_back(temp);**

**}**

**//sort(database.begin(),database.end());**

**int sum=0;**

**vector<int> frequency;**

**for(int i=0;i<n;i++)**

**{ hm[database.at(i)]++; }**

**vector< pair < double,bool> > entropy(n);**

**double tentropy=0;**

**for(int i=0;i<n;i++)**

**{**

**double result;**

**result=hm[database[i]]/(double)n\*(log(hm[database[i]]/(double)n));**

**entropy[i].first=-1\*result;**

**tentropy+=-1\*result;**

**entropy[i].second=false;**

**}**

**cout<<"\n\tEnter the number of Outliers to be detected : ";**

**cin>>k;**

**if(k>n)**

**{**

**cout<<"\n\tOutliers cant be more than the number of elemnts itself.....abort";**

**exit(0);**

**}**

**vector<int> outlierset;**

**int counter =0;**

**while(counter!=k)**

**{ int ii=0;**

**counter++;**

**maxentropychange=0;**

**while(ii!=n)**

**{**

**test=entropy[ii].first;**

**if(entropy[ii].second==false)**

**{**

**if(abs(tentropy-test)>maxentropychange)**

**{**

**maxentropychange=fabs(tentropy-test);**

**t=test;**

**}**

**}**

**ii++;**

**}**

**for(int i=0;i<n;i++)**

**{**

**if(entropy[i].first==t&&entropy[i].second==false)**

**{**

**// cout<<"for counter= "<<counter<<endl;**

**// cout<<i+1<<endl;**

**entropy[i].second=true;**

**break;**

**}**

**}**

**}**

**for(int i=0;i<n;i++)**

**{**

**if(entropy[i].second==true)**

**{**

**cout<<"\n\t\tOutlier detected at V"<<i+1<<" "<<database.at(i);**

**}**

**}**

**cout<<endl;**

**cout<<"\n\n\n THE ABOVE PROGRAM CALCULATED OUTLIERS IN A TEXT FILE WITH HELP OF A FAST GREEDY ALGORITHM.. \n\n";**

**return 0;**

**}**