NAÏVE BAYES for discrete data

data={'Sunny','Hot','High','Weak','No';'Sunny','Hot','High','Strong','No';'Overcast','Hot','High','Weak','yes';'Rain','Mild','High','Weak','Yes';'Rain','cool','Normal','Weak','Yes';'Rain','Cool','Normal','Strong','No';'Overcast','Cool','Normal','Strong','Yes';'Sunny','Mild','High','Weak','No';'Sunny','Cool','Normal','Weak','Yes';'Rain','Mild','Normal','Weak','Yes';'Sunny','Mild','Normal','Strong','Yes';'Overcast','Mild','High','Strong','Yes';'Overcast','Hot','Normal','Weak','Yes';'Rain','Mild','High','Strong','No'};

data=lower(data);

dyes={};

dno={};

u=1;

o=1;

for i=1:length(data)

if(strcmp(data{i,5},'yes'))

dyes{o,1}=data{i,1};

dyes{o,2}=data{i,2};

dyes{o,3}=data{i,3};

dyes{o,4}=data{i,4};

dyes{o,5}=data{i,5};

o=o+1;

else

dno{u,1}=data{i,1};

dno{u,2}=data{i,2};

dno{u,3}=data{i,3};

dno{u,4}=data{i,4};

dno{u,5}=data{i,5};

u=u+1;

end

end

id=strfind(dyes,'overcast');

id=find(not(cellfun('isempty',id)));

Pyes\_overcast=length(id);

id=strfind(dno,'overcast');

id=find(not(cellfun('isempty',id)));

Pno\_overcast=length(id);

id=strfind(dyes,'sunny');

id=find(not(cellfun('isempty',id)));

Pyes\_sunny=length(id);

id=strfind(dno,'sunny');

id=find(not(cellfun('isempty',id)));

Pno\_sunny=length(id);

id=strfind(dyes,'rain');

id=find(not(cellfun('isempty',id)));

Pyes\_rain=length(id);

id=strfind(dno,'rain');

id=find(not(cellfun('isempty',id)));

Pno\_rain=length(id);

Pyes\_outlook=[Pyes\_overcast;Pyes\_sunny;Pyes\_rain]

Pno\_outlook=[Pno\_overcast;Pno\_sunny;Pno\_rain]

Pyes\_outlook=Pyes\_outlook./length(dyes)

Pno\_outlook=Pno\_outlook./length(dno)

id=strfind(dyes,'hot');

id=find(not(cellfun('isempty',id)));

Pyes\_hot=length(id);

id=strfind(dno,'hot');

id=find(not(cellfun('isempty',id)));

Pno\_hot=length(id);

id=strfind(dyes,'mild');

id=find(not(cellfun('isempty',id)));

Pyes\_mild=length(id);

id=strfind(dno,'mild');

id=find(not(cellfun('isempty',id)));

Pno\_mild=length(id);

id=strfind(dyes,'cool');

id=find(not(cellfun('isempty',id)));

Pyes\_cool=length(id);

id=strfind(dno,'cool');

id=find(not(cellfun('isempty',id)));

Pno\_cool=length(id);

Pyes\_temp=[Pyes\_hot;Pyes\_mild;Pyes\_cool]

Pno\_temp=[Pno\_hot;Pno\_mild;Pno\_cool]

Pyes\_temp=Pyes\_temp./length(dyes)

Pno\_temp=Pno\_temp./length(dno)

id=strfind(dyes,'high');

id=find(not(cellfun('isempty',id)));

Pyes\_high=length(id);

id=strfind(dno,'high');

id=find(not(cellfun('isempty',id)));

Pno\_high=length(id);

id=strfind(dyes,'normal');

id=find(not(cellfun('isempty',id)));

Pyes\_normal=length(id);

id=strfind(dno,'normal');

id=find(not(cellfun('isempty',id)));

Pno\_normal=length(id);

Pyes\_humidity=[Pyes\_high;Pyes\_normal]

Pno\_humidity=[Pno\_high;Pno\_normal]

Pyes\_humidity=Pyes\_humidity./length(dyes)

Pno\_humidity=Pno\_humidity./length(dno)

id=strfind(dyes,'strong');

id=find(not(cellfun('isempty',id)));

Pyes\_strong=length(id);

id=strfind(dno,'strong');

id=find(not(cellfun('isempty',id)));

Pno\_strong=length(id);

id=strfind(dyes,'weak');

id=find(not(cellfun('isempty',id)));

Pyes\_weak=length(id);

id=strfind(dno,'weak');

id=find(not(cellfun('isempty',id)));

Pno\_weak=length(id);

Pyes\_wind=[Pyes\_strong;Pyes\_weak]

Pno\_wind=[Pno\_strong;Pno\_weak]

Pyes\_wind=Pyes\_wind./length(dyes)

Pno\_wind=Pno\_wind./length(dno)

Prioryes=length(dyes)/length(data)

Priorno=length(dno)/length(data)

disp('X = (Outlook = Sunny, Temperature = Cool, Humidity = High, Wind = Strong)')

X={'Sunny','Cool','High','Strong'};

X=lower(X)

p=[length(Pno\_outlook);length(Pno\_temp);length(Pno\_humidity);length(Pno\_wind)];

m=1;

for i=1:length(X)

id=strfind(dyes,X{i});

id=find(not(cellfun('isempty',id)));

Pyes=length(id);

Pyes=(Pyes+(m\*(1/p(i))))/(length(dyes)+m);

Prioryes=Prioryes\*Pyes;

id=strfind(dno,X{i});

id=find(not(cellfun('isempty',id)));

Pno=length(id);

Pno=(Pno+(m\*(1/p(i))))/(length(dno)+m);

Priorno=Priorno\*Pno;

end

Probability\_yes=Prioryes;

Probability\_no=Priorno;

Probability\_yes

Probability\_no

if(Prioryes>Priorno)

disp('we can play tennis')

else

disp('we cannot play tennis')

end

OUTPUT:

Probability\_yes =

0.0061

Probability\_no =

0.0193

we cannot play tennis