Data Mining Experiments – Datasets

Experiment 1 – Distance Matrix

Experiment 1 – Distance Matrix

```
@relation distance_example
@attribute Object {A,B,C}
@attribute X numeric
@attribute Y numeric
@data
A,1,2
B,2,4
C,3,5
```

Experiment 2 – K-Means

Experiment 2 – K-Means

```
X1,X2
1,1
1,0
0,2
2,4
3,5
```

Experiment 3 – Preprocessing

Experiment 3 – Data Preprocessing

```
@relation preprocess
@attribute Age numeric
@attribute Income numeric
@attribute Student {Yes,No}
@attribute Credit_Rating {Fair,Excellent}
@attribute Buys_Computer {Yes,No}
@data
25,40,No,Fair,No
28,50,No,Excellent,No
35,60,No,Fair,Yes
45,30,Yes,Fair,Yes
?,20,Yes,Excellent,No
```

Experiment 4 – Naive Bayes (Small)

Experiment 4 – Naive Bayes (Small Dataset)

```
@relation buy
@attribute Age {Young,Adult,Old}
@attribute Student {Yes,No}
@attribute Buys {Yes,No}
@data
Young,Yes,Yes
Young,No,No
Adult,Yes,Yes
Old,Yes,Yes
Old,No,No
```

Experiment 5 – Decision Tree

Experiment 5 – Decision Tree (J48)

```
@relation play_tennis
@attribute Outlook {Sunny,Overcast,Rain}
@attribute Wind {Weak,Strong}
@attribute Play {Yes,No}
@data
Sunny,Weak,No
Sunny,Strong,No
Overcast,Weak,Yes
Rain,Weak,Yes
Rain,Strong,No
```

Experiment 6 – Apriori

Experiment 6 – Apriori (Association Rules)

```
@relation apriori
@attribute Bread {Yes,No}
@attribute Milk {Yes,No}
@attribute Butter {Yes,No}
@attribute Jam {Yes,No}
@data
Yes,Yes,No,No
Yes,No,Yes,No
No,Yes,Yes,Yes
Yes,Yes,Yes,No
```

No,No,Yes,Yes

Experiment 7 – Hierarchical Clustering

Experiment 7 – Agglomerative Hierarchical Clustering

@relation cluster
@attribute X numeric
@attribute Y numeric
@data
1,1
2,1
4,3
5,4

Experiment 8 – FP-Growth

Experiment 8 – FP-Growth

@relation fpgrowth
@attribute Bread {Yes,No}
@attribute Milk {Yes,No}
@attribute Butter {Yes,No}
@attribute Jam {Yes,No}
@data
Yes,Yes,No,No
Yes,No,Yes,No
No,Yes,Yes,Yes
Yes,Yes,Yes,No
No,No,Yes,Yes