

# Decision Tree Implementation

Assignment Report

**Machine Learning**

*To the*

**The University Of Texas at Dallas**



*Submitted by:*

**Prashant Prakash**

**(Net id: pxp141730)**

The University of Texas at Dallas

Dallas – 75252

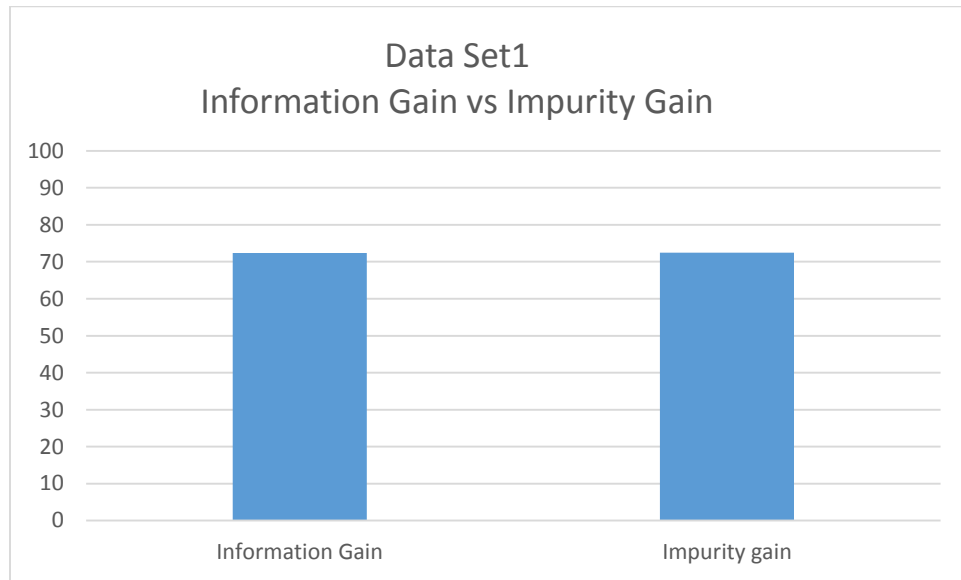
February, 2015

1. Accuracy Reported for Information Gain and Impurity Gain Heuristic :

- DataSet1

Information Gain Heuristic: 75.85 %

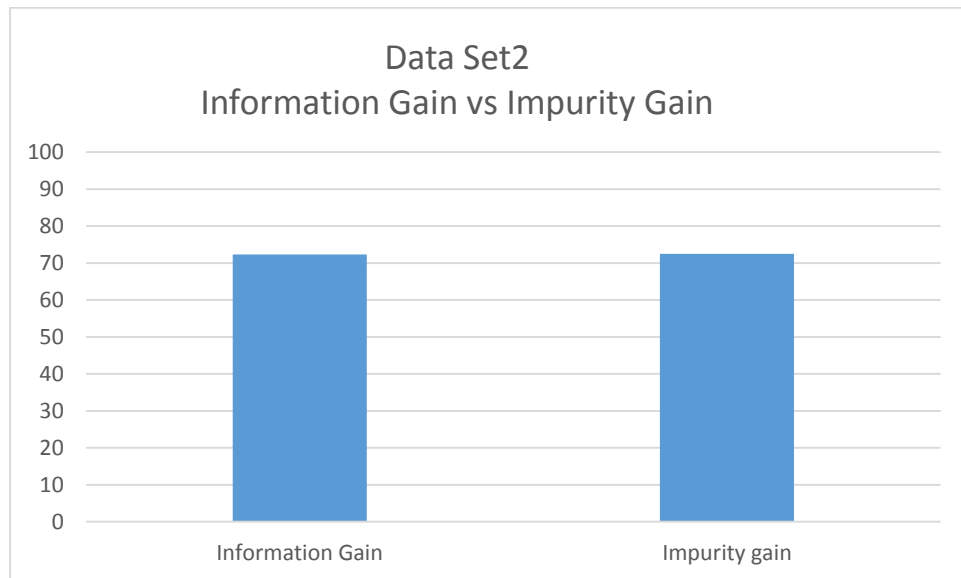
Impurity Gain Heuristic: 76.8%



- DataSet2

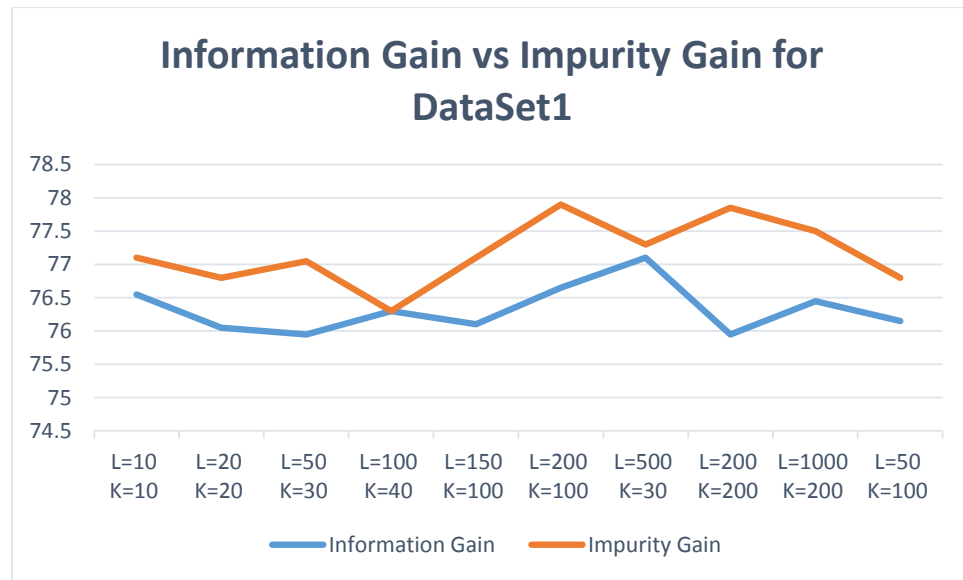
Information Gain Heuristic: 72.33333333333334%

Impurity Gain Heuristic: 72.5%



2. Accuracy Reported for Information Gain and Impurity gain for Different Values of L and K after pruning
  - For Data Set1

L and K Values	L=10 K=10	L=20 K=20	L=50 K=30	L=100 K=40	L=150 K=100	L=200 K=100	L=500 K=30	L=200 K=200	L=1000 K=200	L=50 K=100
Information Gain Heuristic	76.55	76.05	75.95	76.3	76.1	76.65	77.1	75.95	76.45	76.15
Impurity Gain Heuristic	77.1	76.8	77.05	76.3	77.1	77.9	77.3	77.85	77.5	76.8



- For Dataset2

L and K Values	L=10 K=10	L=20 K=20	L=50 K=30	L=100 K=40	L=150 K=100	L=200 K=100	L=500 K=30	L=200 K=200	L=1000 K=200	L=50 K=100
Information Gain Heuristic	72.67	73	73.66	74.17	71.34	73.34	73	72.83	75.67	72.16
Impurity Gain Heuristic	73.67	75.16	75.67	75.67	73.67	74.83	74.5	73.17	74.83	74

