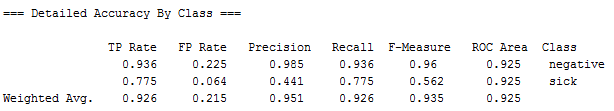
**2.1 Exercise A: Apply the Discretize filter to the ‘Sick’ dataset**

**Task A1:**

7 Numeric Attributes

1 age, 18 TSH, 20 T3, 22 TT4, 24 T4U, 26 FTI, 28 TBG

**Task A2:**



**Task A3:**

How are they different?

1 age: 3 distinct ranges have been created

18 TSH: 1 distinct ranges have been created

20 T3: 2 distinct ranges have been created

22 TT4: 2 distinct ranges have been created

24 T4U: 4 distinct ranges have been created

26 FTI: 1 distinct ranges have been created

28 TBG: 0 distinct ranges have been created

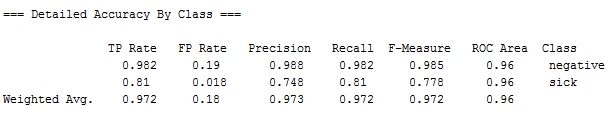
**Task A4:**

Bin = 5

Bin = 10

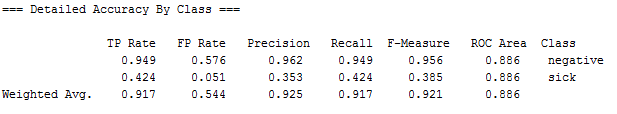
‘Bins’ setting affects …

**Task A5:**

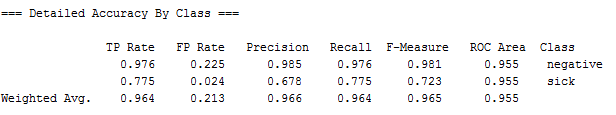


**Task A6:**

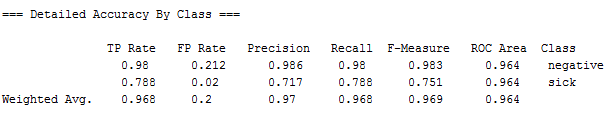
Bin = 5



Bin = 10



Bin = 20



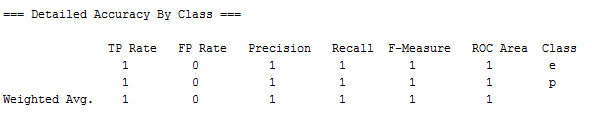
**Task A7:**

The highest accuracy rate is from Task 5 result. Chose FilteredClassifier and set its classifier to Naïve Bayes and its filter to (supervised) Discrete. Select cross validation from the test oprions.

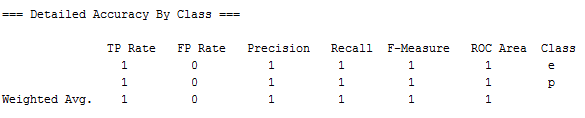
**3.1 Exercise B: Use CfsSubsetEval to improve the J48, IBk and NB classification algorithms**

**Task B1:**

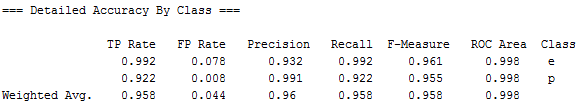
Classifier: J48



Classifier: IBk



Clissifier: NB

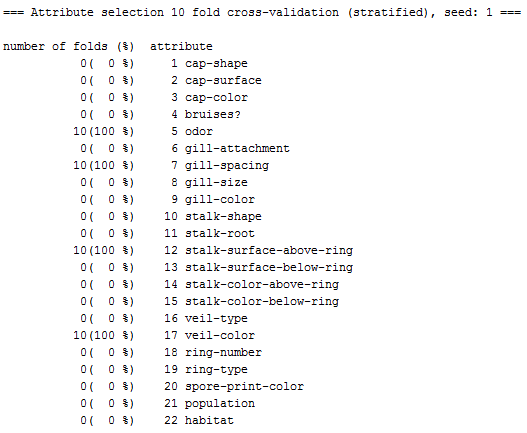


**Task B2:**

What do the results say about the attributes with a score of 0%:

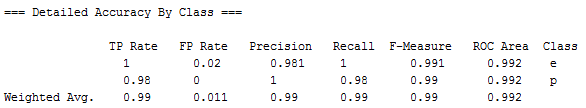
The attributes with a score of 0% are not relevant attributes

Relevant attributes (high score):  
5 odor, 7 gill-spacing, 12 stalk-surface-above-ring, 17 veil-color

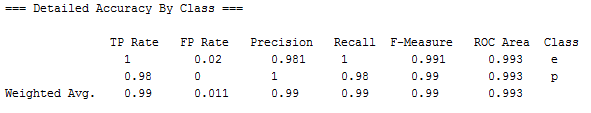


**Task B3:**

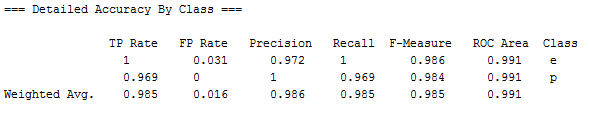
Classifier: J48



Classifier: IBk



Classifier: NB

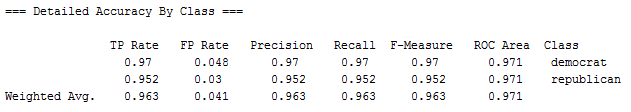


**Task B4:**

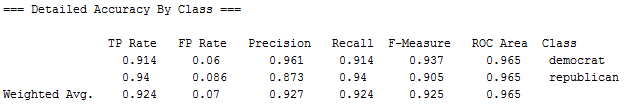
**3.2 Exercise C: Use WrapperSubsetEval to improve the J48, IBk and NB algorithms**

**Task C1:**

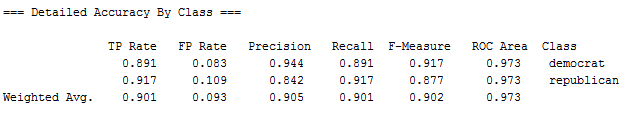
Classifier: J48



Classifier: IBk



Classifier: NB



--------------------------------------------------------------------------------------------------------------------------------------------------

**Task C2:**

Record the attribute numbers of attributes that are irrelevant: 2, 10, 16

=== Run information ===

Evaluator: weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.trees.J48 -F 5 -T 0.01 -R 1 -- -C 0.25 -M 2

Search:weka.attributeSelection.RankSearch -S 1 -R 0 -A weka.attributeSelection.InfoGainAttributeEval --

Relation: vote

Instances: 435

Attributes: 17

handicapped-infants

water-project-cost-sharing

adoption-of-the-budget-resolution

physician-fee-freeze

el-salvador-aid

religious-groups-in-schools

anti-satellite-test-ban

aid-to-nicaraguan-contras

mx-missile

immigration

synfuels-corporation-cutback

education-spending

superfund-right-to-sue

crime

duty-free-exports

export-administration-act-south-africa

Class

Evaluation mode:10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute

9( 90 %) 1 handicapped-infants

0( 0 %) 2 water-project-cost-sharing

10(100 %) 3 adoption-of-the-budget-resolution

10(100 %) 4 physician-fee-freeze

10(100 %) 5 el-salvador-aid

10(100 %) 6 religious-groups-in-schools

10(100 %) 7 anti-satellite-test-ban

10(100 %) 8 aid-to-nicaraguan-contras

10(100 %) 9 mx-missile

3( 30 %) 10 immigration

10(100 %) 11 synfuels-corporation-cutback

10(100 %) 12 education-spending

10(100 %) 13 superfund-right-to-sue

10(100 %) 14 crime

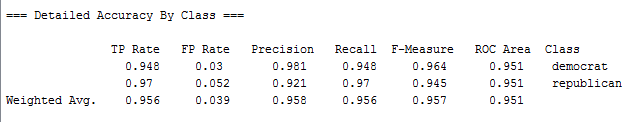
10(100 %) 15 duty-free-exports

3( 30 %) 16 export-administration-act-south-africa

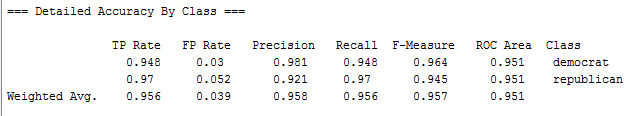
--------------------------------------------------------------------------------------------------------------------------------------------------

**Task C3:**

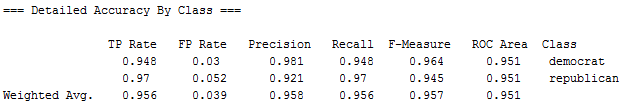
Classifier: J48



Classifier: IBk



Classifier: NB



--------------------------------------------------------------------------------------------------------------------------------------------------

**Task C4:**