Submitted by: Ranji Raj

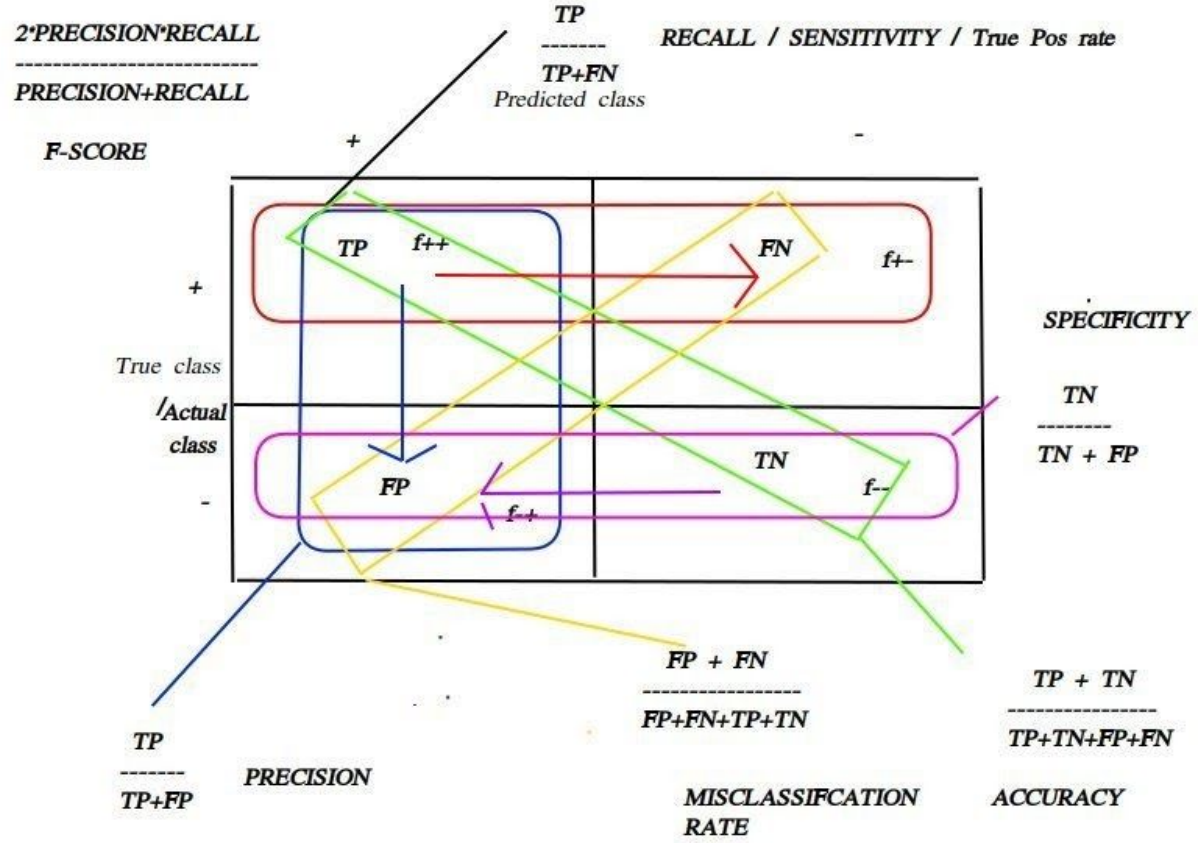
Assignment 1.3

What is a confusion matrix and how does it look like? Accuracy, Precision, Recall, Specificity, the miss rate and the fall-out are measures, that can be derived from it. Show this derivation and explain their meaning! How can these help in selecting a good classifier (e.g. a good concept learner)?

A **confusion matrix** is a table that is often used to describe the performance of a classification model.

Common terms in cm:
TP (f_{++}) - **Positive** examples correctly classified as **positive**
FN (f_{+-}) - **Positive** examples incorrectly classified as **negative**
FP (f_{-+}) - **Negative** examples incorrectly classified as **positive**
TP (f_{--}) - **Negative** examples correctly classified as **negative**

Example of a binary confusion matrix:



Accuracy - defines how often the classifier is correct
Miss error rate - defines how often the classifier is wrong
Recall / Sensitivity - defines when the sample is actually POS how often does it predict as POS.
Specificity - defines when the sample is actually NEG how often does it predict as NEG.
Precision - When it predicts as POS how correct (or accurate) it is.
F-score - weighted average of recall and precision.

A good classifier has higher accuracy and a lower error rate.