

BT5152 Tutorial 2

AY 2018/19, Semester 1, Week 4

Lu Wei

Key Concepts Revision

Confusion Matrix

		Actual class		
		Cat	Dog	Rabbit
Predicted class	Cat	5	2	0
	Dog	3	3	2
	Rabbit	0	1	11

https://en.wikipedia.org/wiki/Confusion_matrix

Terminology and derivations from a confusion matrix

condition positive (P)

the number of real positive cases in the data

condition negative (N)

the number of real negative cases in the data

true positive (TP)

eqv. with hit

true negative (TN)

eqv. with correct rejection

false positive (FP)

eqv. with false alarm, Type I error

false negative (FN)

eqv. with miss, Type II error

https://en.wikipedia.org/wiki/Precision_and_recall

sensitivity, recall, hit rate, or true positive rate (TPR)

$$\text{TPR} = \frac{\text{TP}}{P} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

specificity, selectivity or true negative rate (TNR)

$$\text{TNR} = \frac{\text{TN}}{N} = \frac{\text{TN}}{\text{TN} + \text{FP}} = 1 - \text{FPR}$$

precision or positive predictive value (PPV)

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

accuracy (ACC)

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{P + N} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

F1 score

is the harmonic mean of precision and sensitivity

$$F_1 = 2 \cdot \frac{\text{PPV} \cdot \text{TPR}}{\text{PPV} + \text{TPR}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

Cohen's Kappa

Same accuracy, which model is better?

		Actual	
		Yes	No
Predicted	Yes	45	15
	No	25	15

70 30
Model Left

$$\kappa = \frac{0.6 - 0.54}{1 - 0.54}$$

60
40

		Actual	
		Yes	No
Predicted	Yes	25	35
	No	5	35

30 70
Model Right

$$p(\epsilon) = \frac{45}{60} \times \frac{45}{70}$$

$$0.6 \times 0.7 + 0.4 \times 0.3 = 0.42 + 0.12 = 0.54$$

Kappa vs. F Measure

Is this a good model?

		Actual	
		Yes	No
Predicted	Yes	899	99
	No	1	1

900 100

998

2

$$\text{accuracy} = \frac{900}{1000} = 0.9$$
$$\text{Kappa} = 0.9 -$$

$$0.9 \times 0.998 + 0.002 \times 0.1$$

Kappa Fail?

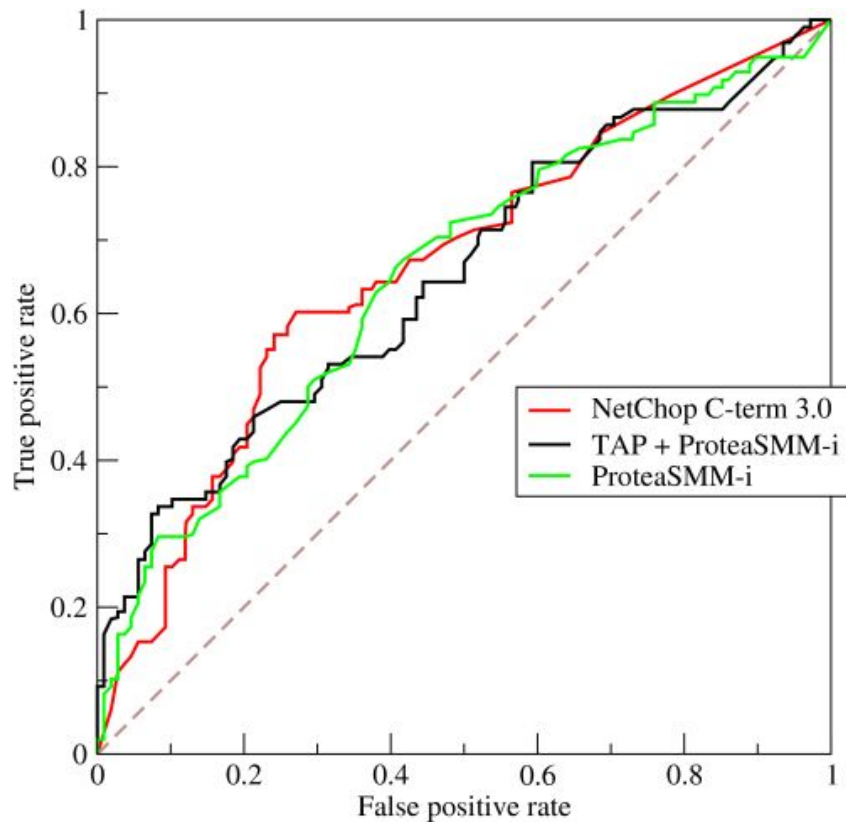
What's Kappa stats in this case? What about precision, recall, and F-score?

		Actual	
		Yes	No
Predicted	Yes	0	1
	No	1	14

K-Fold Cross Validation

- Simplified vs. leave-one-out
- Stratification
- How does the value of K affect program running speed?
- Which folds do we use to build the final classifier?

Receiver Operating Characteristic (ROC)



Assignment 1

- Upload 2 files (*.R and *.pdf) into IVLE BT5152 workbin under the folder: Student Submission > A1
- Assume dataset files are in the same directory as the R script,
i.e. `train_data <- read.csv("loan_train.csv", stringsAsFactors = TRUE)`
- Make sure your R file is runnable and has all the dependency packages imported e.g. `library(C50)`
- I will fail any submission with an R file that's not executable
- You may revise and submit as many times before deadline. Make sure to remove any old version that you don't wish to be graded.
- Plagiarism: 0 grade + reported to the department

Tutorial Exercises:

RStudio > Console:

```
# install.packages("swirl")  
library(swirl)  
# delete_progress('your name')  
install_course_github('weilu', 'BT5152', multi=TRUE)  
swirl()
```

1: Performance Metrics

2: ROC and AUC

3: Model Selection