Intro to ML

December 15th, 2021

教學助理
教學意見調查

WE NEED YOUR FEEDBACK

即日起
~1/2止

濃濃的聖誕氛圍
預告著學期即將步入尾聲
你，是否有些話想要對TA說呢?
不論是想要表達你的感謝
或是要給認真的TA一個愛的鼓勵
還是有其他的建議想提供給TA
採無記名方式填卷

就是現在!!

❶ 登入「校務資訊系統」。

校務資訊系統 Academic Information Systems
https://www.ccep.nthu.edu.tw/ccep/INQUIRE/

❷ 點選「教學助理評量問卷」即可填寫。

問卷填寫時間：110/12/10-111/1/2止。

完成全部問卷填寫的同學，就有機會抽到NTHU限量小禮唷!
同學現在就拿起手機掃描QR Code前往填寫問卷!

主辦單位：教務處教學發展中心 | 聯絡資訊：03-5716200 | http://ctld.nthu.edu.tw/

教務處課務組

110上期末教學意見調查
懶人包

2021.12.10~2022.01.10上午9:00

抽獎獎品有哪些?

品名	數量
iPad Air 64GB Wi-Fi	1
Apple Watch S7 41mm搭配運動錶帶	4
BIRDYEDGE G3 (輕量)電動滑板車	4

填卷截止後一周內，以電腦亂數抽出獲獎者並公布於課務組網頁。

求關注幫高調

懇請幫忙
分享宣傳

《班級鼓勵獎》獎金有多少?

填卷率高的班級發給鼓勵金(獲獎率約10%)。

首獎：共1班 獎金 5,000元
貳獎：共1班 獎金 4,000元
參獎：共1班 獎金 3,000元
特別獎：共15班 獎金各2,000元

★首獎、貳獎、參獎全校限各1班。

去哪查排名?

課務組網頁
每個工作天更新
最新排名

比賽得獎金！ 加碼抽大獎！

CHAPTER 20:

Design and Analysis of ML Experiments

Algorithm Preference

- Criteria (Application-dependent):
  - Misclassification error, or risk (loss functions)
  - Training time/space complexity
  - Testing time/space complexity
  - Interpretability
  - Easy programmability
- Cost-sensitive learning

Guidelines for ML experiments

A. Aim of the study
B. Selection of the response variable
C. Choice of factors and levels
D. Choice of experimental design
E. Performing the experiment
F. Statistical Analysis of the Data
G. Conclusions and Recommendations

Resampling and *K*-Fold Cross-Validation

- The need for multiple training/validation sets

  $\{X_i,V_i\}_i$: Training/validation sets of fold i
- K-fold cross-validation: Divide X into k, X_i,i=1,...,K

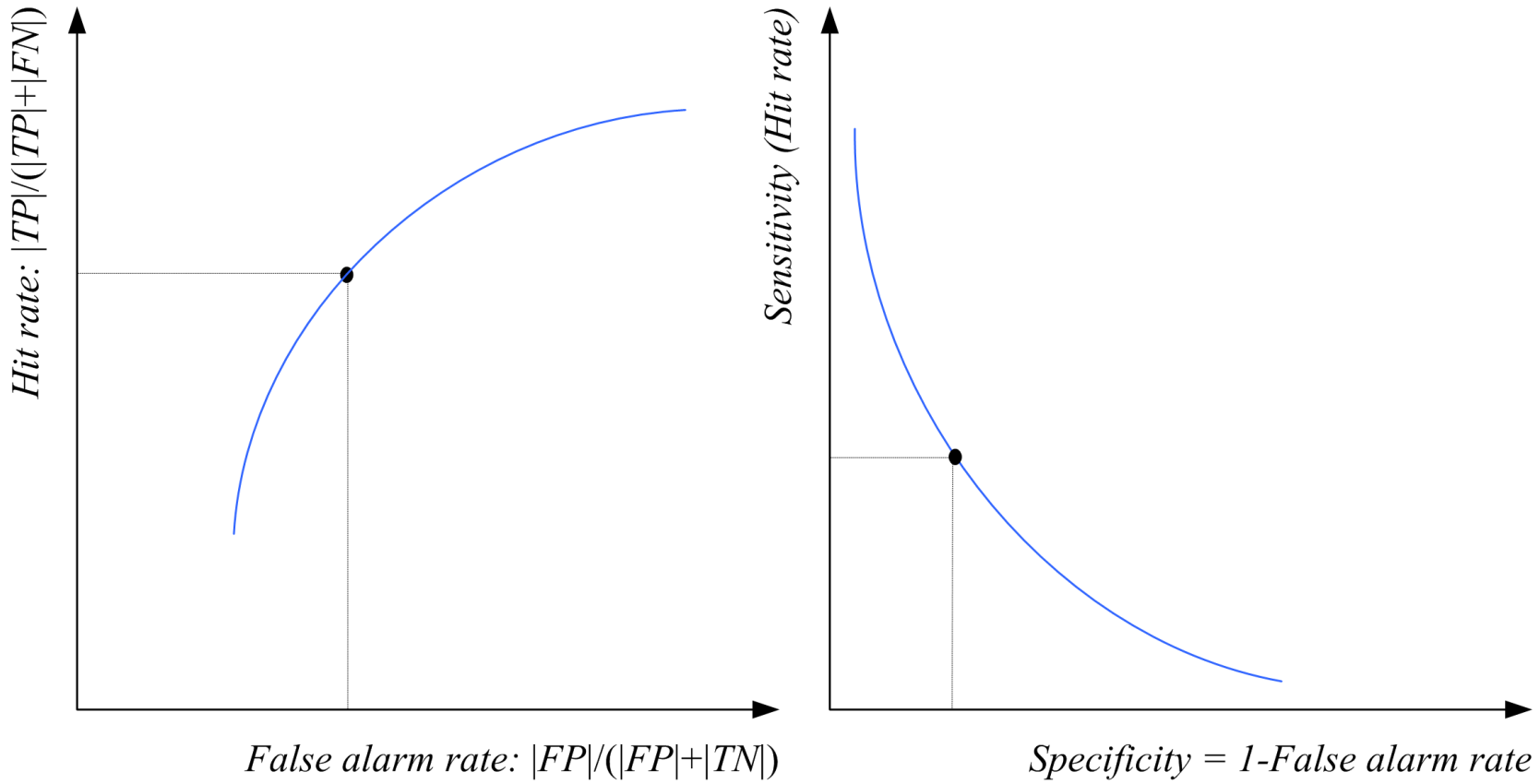
$$\begin{array}{ll} \mathcal{V}_1 = \mathcal{X}_1 & \mathcal{T}_1 = \mathcal{X}_2 \cup \mathcal{X}_3 \cup \cdots \cup \mathcal{X}_K \\ \mathcal{V}_2 = \mathcal{X}_2 & \mathcal{T}_2 = \mathcal{X}_1 \cup \mathcal{X}_3 \cup \cdots \cup \mathcal{X}_K \\ \vdots & \\ \mathcal{V}_K = \mathcal{X}_K & \mathcal{T}_K = \mathcal{X}_1 \cup \mathcal{X}_2 \cup \cdots \cup \mathcal{X}_{K-1} \end{array}$$

Performance Measures

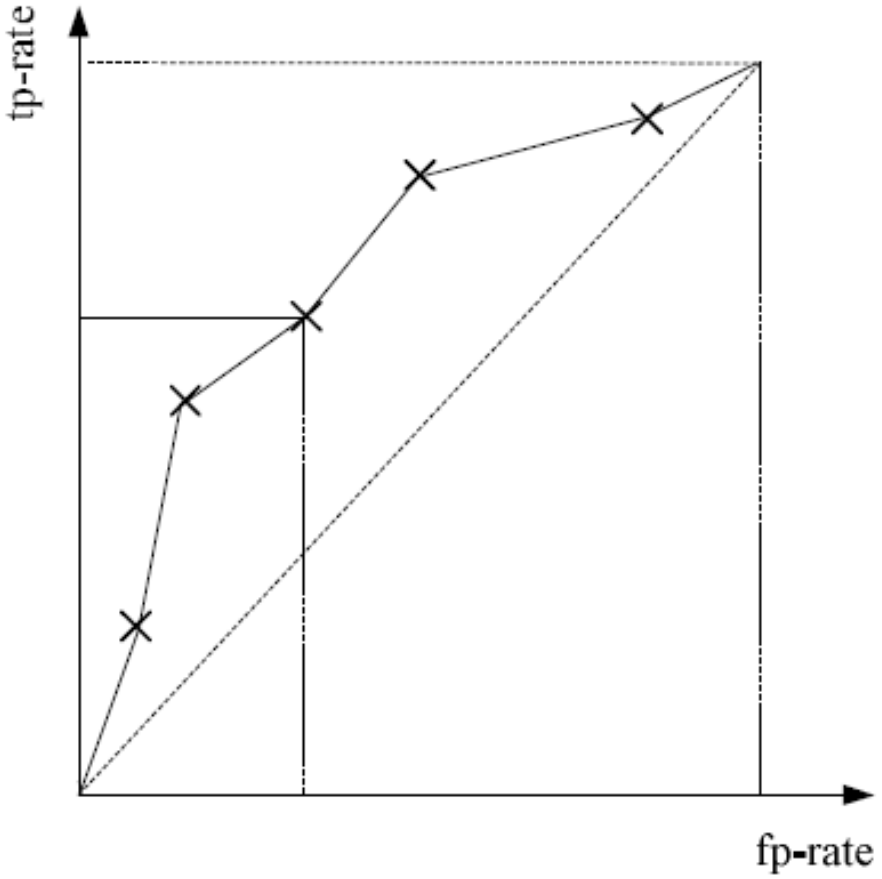
	Predicted class	
True Class	Yes	No
Yes	TP: True Positive	FN: False Negative
No	FP: False Positive	TN: True Negative

- Error rate = # of errors / # of instances = (FN+FP) / N
- Recall = # of found positives / # of positives
  = TP / (TP+FN) = sensitivity = hit rate
- Precision = # of found positives / # of found
  = TP / (TP+FP)
- Specificity = TN / (TN+FP)
- False alarm rate = FP / (FP+TN) = 1 - Specificity

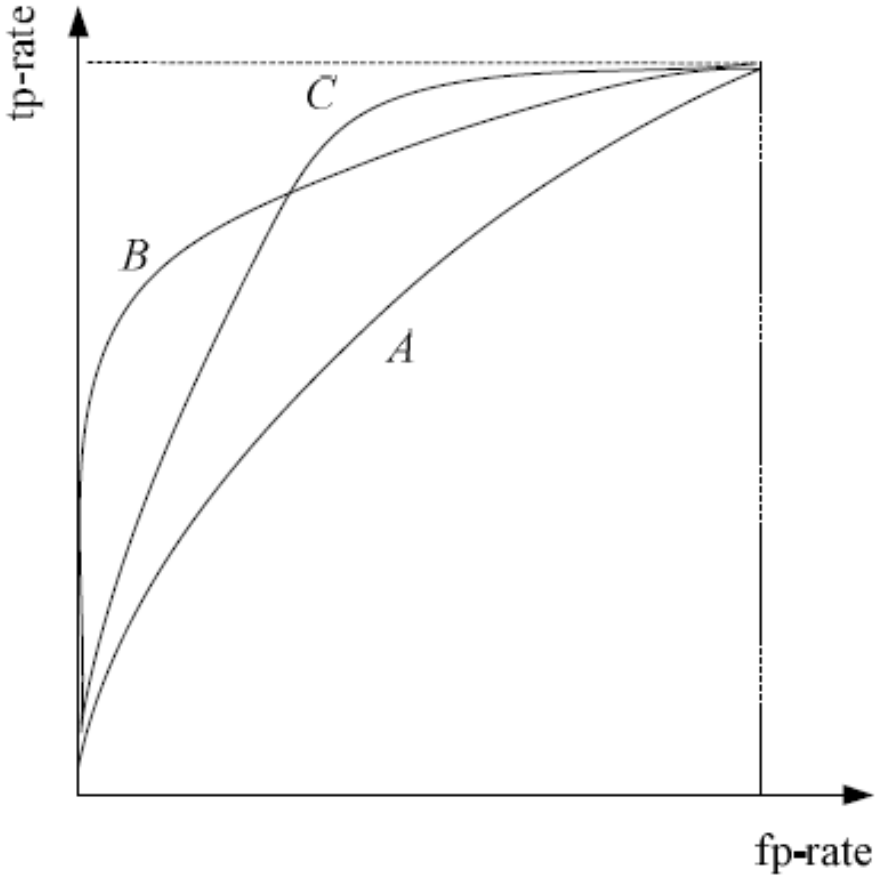
Plotting over Different Decision Thresholds



ROC curve

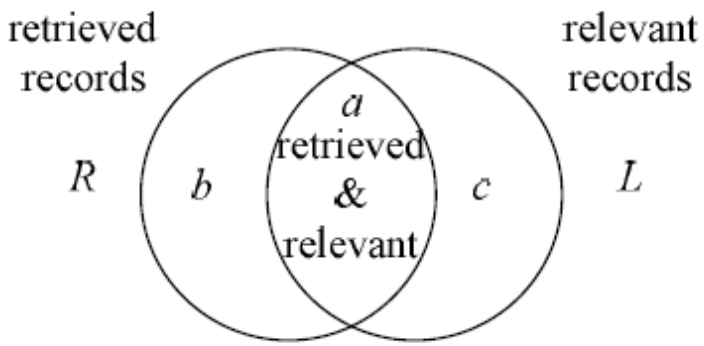


(a) Example ROC curve



(b) Different ROC curves for different classifiers

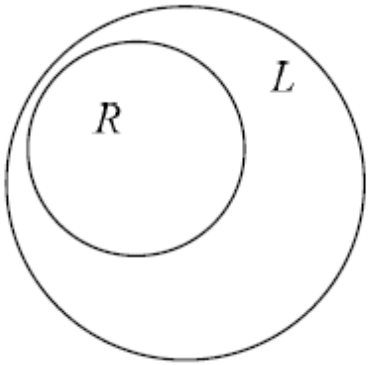
Precision and Recall



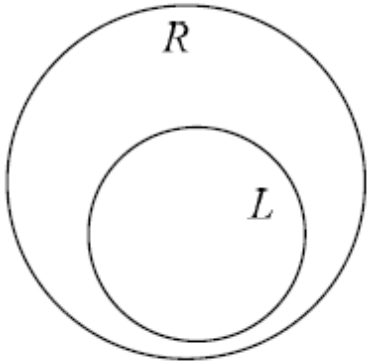
$$\text{Precision: } \frac{a}{a + b}$$

$$\text{Recall: } \frac{a}{a + c}$$

(a) Precision and recall



(b) Precision = 1



(c) Recall = 1