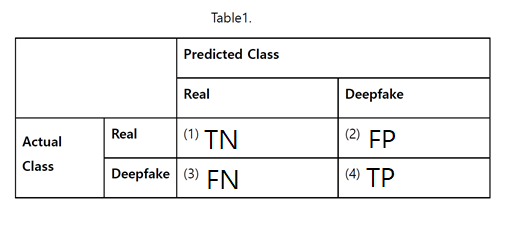
**Data Science Application  
ADS5032\_01(우사이먼성일 교수님)**

**2020713016**

**심우인**

**HW#3 (Data Analysis).**

**Q1. You are given the following, where you are trying to detect Deepfake images (anomalies) from real images (normal data):**

****

**a. What is TP in above cell? Explain and indicate in Table 1.**

True Positive, Deepfake을 Deepfake로 예측. TP is the number of positive predictions that the model correctly classified as positive for example, given class is Deepfake class and the classifier has been correctly predicted it as Deepfake class.

**b. What is FP in above cell? Explain and indicate in Table 1.**

False Positive, Real를 Deepfake로 예측. 제1종 오류. FP is the number of incorrect predictions that an example is positive which means negative class incorrectly identified as positive. For example, given class is Real class however, the classifier has been incorrectly predicted it as Deepfake class. It also referred to as Type 1 errors.

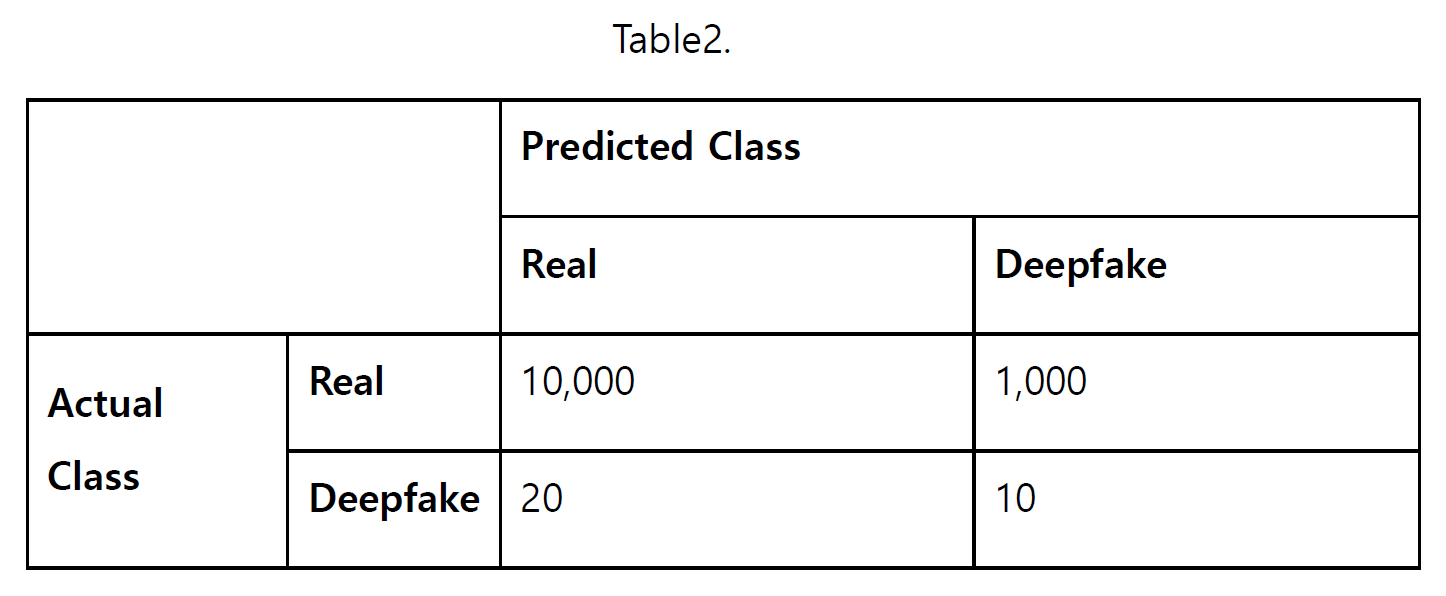
**c. What is TN in above cell? Explain and indicate in Table 1.**

True Negative, Real를 Real로 예측. TN is the number of correct predictions that an example is negative which means negative class correctly identified as negative for example, given class is Real class and the classifier has been correctly predicted it as Real class.

**d. What is FN in above cell? Explain and indicate in Table 1.**

False Negative, Deepfake을 Real로 예측. 제2종 오류. FN is the number of negative predictions that the model incorrectly identified as negative for example, given class is Deepfake class however, the classifier has been incorrectly predicted it as Real class. It also referred to as Type 2 errors.

**Q2. You have the following data and results from your ML algorithm.**



**a. Calculate the accuracy.**

**b. Calculate the precision.**

**c. Calculate the recall.**

**d. Calculate the F1.**

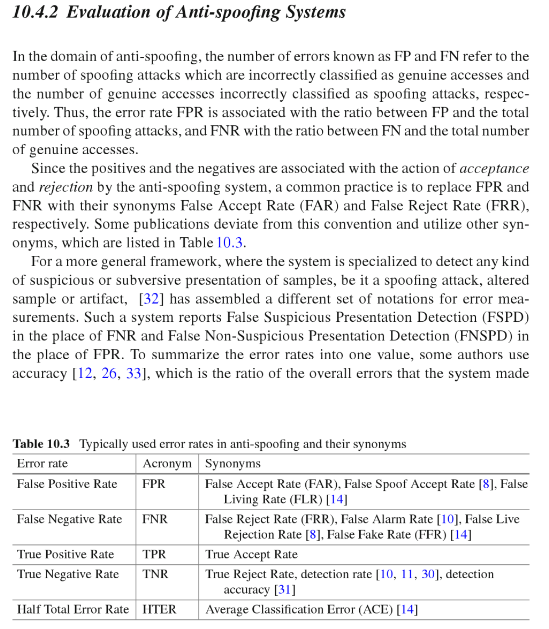
**e. Calculate the accuracy of Real class.**

**f. Calculate the accuracy of Deepfake class.**

**Q3. Calculate the False Acceptance Rate (FAR) (you can search the definition on the Internet).**

생체인식에서 오수락률(FAR)은 등록되지 않은 사용자를 등록자로 잘못 인식하여 수락할 확률이다. 오수락률은 일반적으로 허가 받지 않은 사용자가 문자 그대로 시스템에 접근할 수 있도록 허용함으로 생체인식 오류 중 가장 심각한 것이다.

<https://books.google.ca/books?id=Go4kBAAAQBAJ&pg=PA195&lpg=PA195&dq=FRR+vs+FNR&source=bl&ots=wZQadPKSIM&sig=fXrSks9EKc_ebkMaDuuXBMMqugM&hl=en&sa=X&ved=0ahUKEwjd9dDkvJrTAhXC5YMKHS1LAIIQ6AEITzAJ#v=onepage&q=FRR%20vs%20FNR&f=false>



FAR = FPR = FP/(FP + TN)

FRR = FNR = FN/(FN + TP)

where

FP: False positive

FN: False Negative

TN: True Negative TP: True Positive

**Q4. Calculate the False Rejection Rate (FRR) (you can search the definition on the Internet).**

생체인식에서 오거부율(FRR)은 등록된 사용자를 거부할 확률이다. 즉, 생체인식시스템에 등록된 사용자의 접근시도를 식별하지 못하고 잘못 받아드릴 가능성을 측정한 것이다. 오거부율(FRR)은 일반적으로 접근거부 횟수를 접근시도 횟수로 나눈 비율이다.

FAR과 FRR은 서로 독립적이고 반비례한다. FAR이 낮을수록 해당 FRR이 높고, 반대로 FRR이 낮을수록 해당 FAR은 높다. 그런 의미에서 FAR과 FRR은 상호의존적이 아니라 반비례 관계에 있다.

즉, FAR이 높아지면 FRR이 낮아짐에 따라 인가된(등록된) 사용자를 거부하는 상황이 초래되는 것이고, FRR이 높아지면 FAR이 낮아져 인가되지 않은 일반 사용자를 받아드릴 확률이 높아짐을 의미한다. 즉, 보안적인 측면만 고려한다면 FAR을 낮추고 FRR을 높이는 것이 안전하지만 인가된 사용자가 거부되는 상황을 염두에 두어야 한다.

현재 학계나 업계에서는 이러한 확률을 근거로 생체인식기술을 발전시켜 나가고 있으며 균형 있게 FRR와 FAR 비율을 맞추는 것을 권고하고 있지만, 이러한 사항을 준수하지 않아 발생한 피해가 있다.

[1] <http://oaji.net/articles/2017/786-1509075793.pdf>

[2] https://[www.bayometric.com/increasing-importance-of-biometric-security/](http://www.bayometric.com/increasing-importance-of-biometric-security/)

[3] https://[www.gemalto.com/govt/inspired/biometrics](http://www.gemalto.com/govt/inspired/biometrics)

[4] https://heimdalsecurity.com/blog/biometric-authentication/

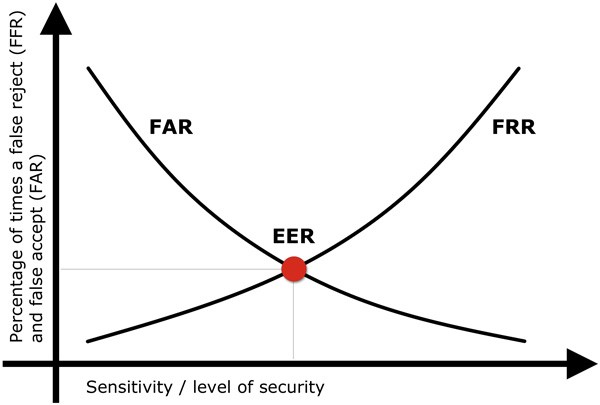
[5] https://[www.esecurityplanet.com/trends/biometric-authentication-how-it-works.html](http://www.esecurityplanet.com/trends/biometric-authentication-how-it-works.html)

**Q5. (Extra Credit) What is Equal Error Rate (EER)?**

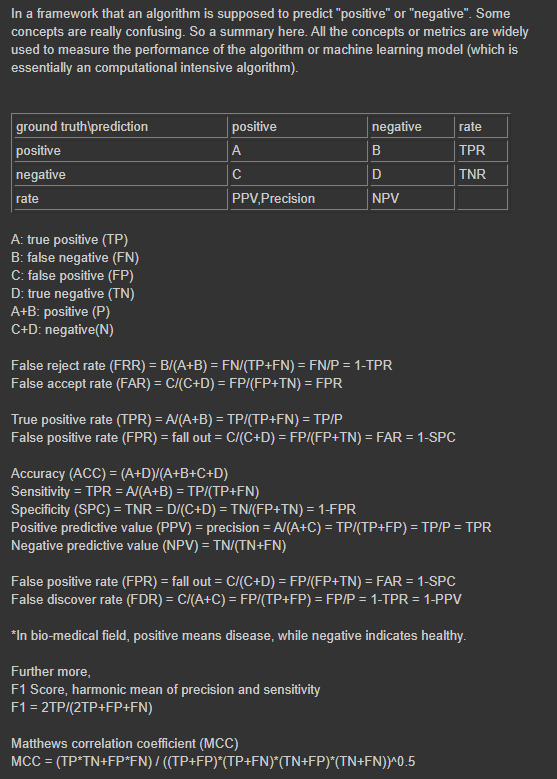
**What is EER with Table2, you can make assumptions if needed?**

**If cannot be calculated in Table 2, then explain the reason.**

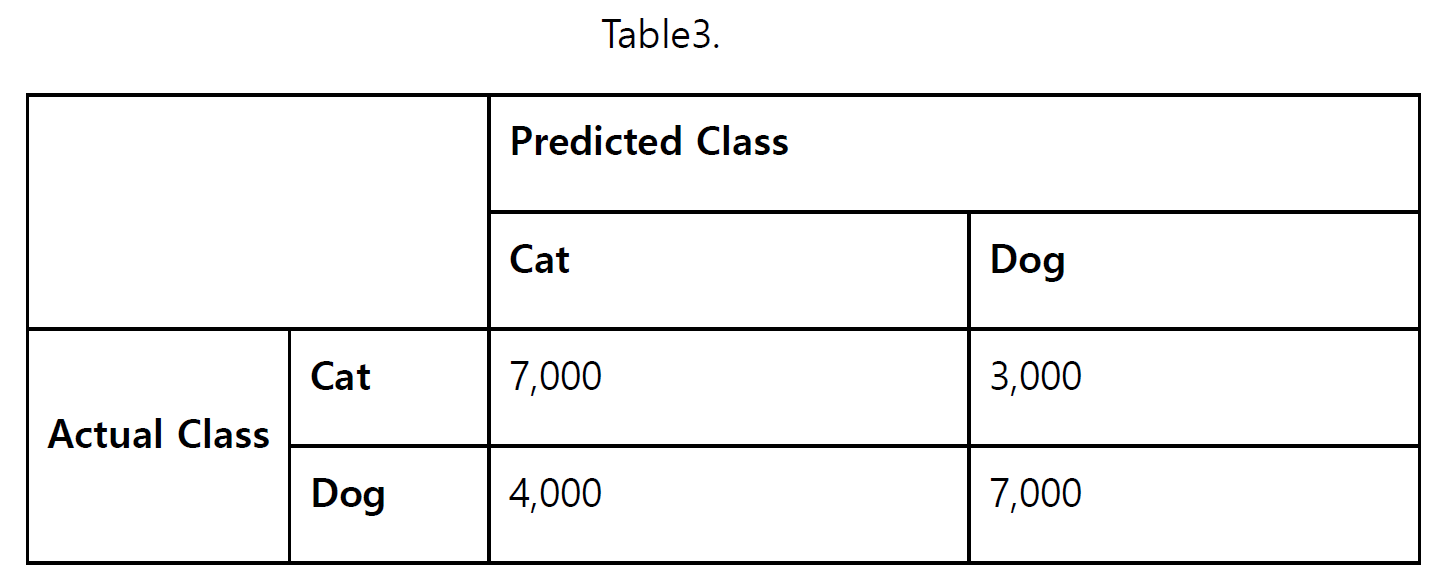
As the number of false acceptances (FAR) goes down, the number of false rejections (FRR) will go up and vice versa (see the figure below). The point at which the lines intersect also has a name: the Equal Error Rate (EER). This is where the percentage of false acceptances and false rejections is the same.

****

**https://www.recogtech.com/en/knowledge-base/security-level-versus-user-conveniencehttps://www.recogtech.com/en/knowledge-base/security-level-versus-user-convenience**

****

**Q6. You have the following results from your ML algorithm for classifying cats and dogs:**



**a. Calculate the accuracy.**

**b. Calculate the precision.**

**c. Calculate the recall.**

**d. Calculate the F1.**

**e. Calculate the accuracy of Cat class.**

**f. Calculate the accuracy of Dog class.**

**Q7. Compare the results between Q2 and Q6, argue whether precision, recall, F1, and accuracy are the reasonable metrics to measure the performance.**

In the case of'Table 2 of Problem 2', the case of the real class and the Deepfake class are imblanced. Our goal is to accurately classify Deepfake class as Deepfake class, but predicting Real class as Real is not very important. In this case, in the case of accuracy, since we try to predict more classes with higher weights, the accuracy can be increased. For example, if the values ​​of 1 and 0 each have a ratio of 9:1 in 10 data distributions, even if all values ​​are predicted as 1, the accuracy is 90%. In this dataset, the accuracy is high at 90.75%, but most of the real classes are classified as real, so it is better to use precision, recall, or f1 score, not accuracy. At this time, you can use the evaluation index that meets the problem requirements. If a big problem occurs when the actual positive data is incorrectly judged as negative, it is recommended to use Recall, and if a big problem occurs when the actual negative data is incorrectly classified as a positive, it is recommended to use Precision.

On the other hand, in the case of'Table 3 of Problem 6', the data of the Cat class and the Dog class are uniform. Therefore, accuracy as well as Precion, Recall, and F1 scores can be used, and evaluation indexes that fit the problem requirements are used.

**Q8. Suggest better ways or metrics to measure the performance of highly imbalanced dataset such as in Q2. Feel free to research things on the Internet or papers, but do cite them, and write things in your own words (Merely copying and pasting, and/or not citing the sources will be given F in the course, because of the violation of Plagiarism)**

**https://machinelearningmastery.com/tour-of-evaluation-metrics-for-imbalanced-classification/**