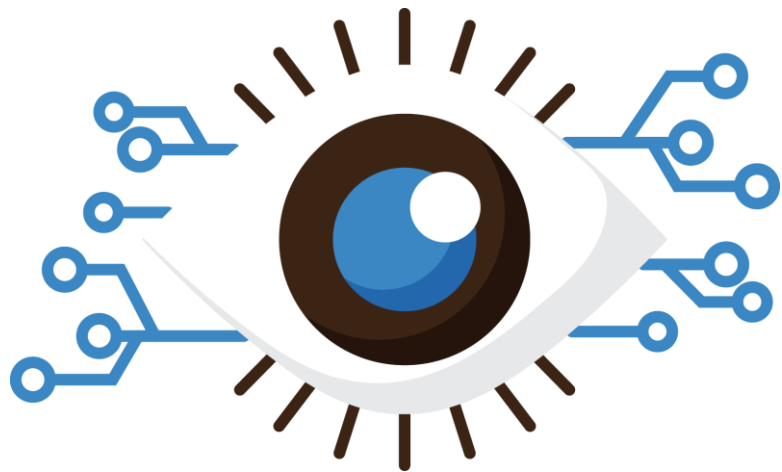# Machine Learning Beginner Course

4.Evaluation Metrics

4. مقاييس التقييم

دورة تعليم الالة للمبتدئين

Salam again!

In the last lessons, we learned about the supervised ML, in which we trained models on small datasets, So after we develop a model, we want to find out how good it is, that's the evaluation phase.

What's measured improves

Evaluation Metric

Building Machine learning Model is based on the principle of continuous feedback. The Machine learning Models are built and model performance is evaluated further Models are improved continuously and continue until you achieve a desirable accuracy. Model Evaluation metrics are used to explain the performance of metrics. Model Performance metrics aim to discriminate among the model results.

Example

Country	Age	Salary	Purchased
Algeria	44	72000	0
Tunis	27	48000	1
Morocco	30	54000	0
Tunis	38	61000	0
Morocco	40	60333.33	1
Algeria	35	58000	1
Tunis	37.333	52000	0
Algeria	48	79000	1
Morocco	50	83000	0
Algeria	37	67000	1

$$X \qquad\qquad Y$$

The output can be in two main type, either discrete or continuous, if it is the discrete case we say it's a problem of **classification** otherwise it's a problem of **Regression**

Example

Healthy and Sick
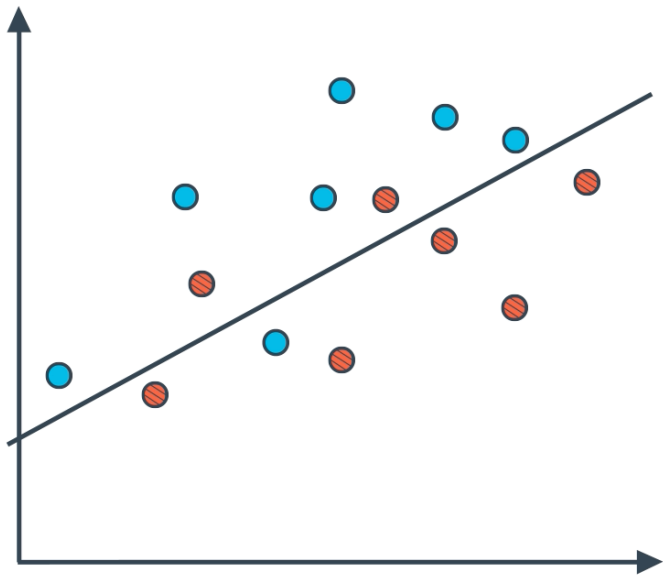
Spam or Not

1.
Classification Metrics

Confusion Matrix

	Diagnosed Sick	Diagnosed Healthy
SICK	True Positive	False Negative
HEALTHY	False Positive	True Negative

Confusion Matrix Example

10,000 Patient	Diagnosed Sick	Diagnosed Healthy
SICK	1000	200
HEALTHY	800	8000

Exercise

	Guessed Positive	Guessed Negative
Positive	True Positives	False Negatives
Negative	False Positives	True Negatives

Confusion Matrix Example

14 Example	Guessed Positive	Guessed Negative
Positive	6	1
Negative	2	5

Type I and type II Error

Accuracy - الدقة

Out of all patient, how many did we classify correctly?

10,000 Patient	Diagnosed Sick	Diagnosed Healthy
SICK	1000	200
HEALTHY	800	8000

$$Accuracy = \frac{1000 + 8000}{10000}$$

$$= 90\%$$

Exercise

1000 Email	Spam Folder	Inbox
Spam	100	170
Not Spam	30	700

Accuracy ?

Exercise
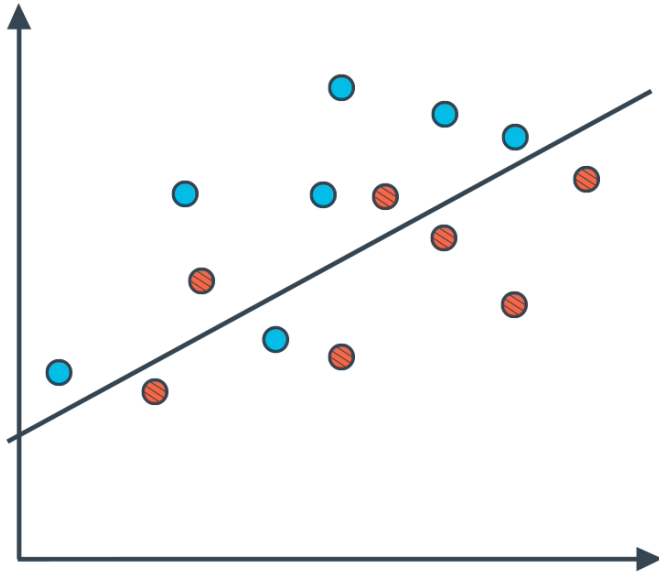
Accuracy ?

Accuracy Won't Work Each Time

✅ 284,335 Non Fraud Card

❌ 472 Fraud Card

We have a model that says all the card are non Fraud.

$$Accuracy = \frac{284,335}{284,887}$$

Accuracy = 99.83%

Precision and Recall

Precision : It is the number of correct positive results divided by the number of positive results predicted by the classifier.

Recall : It is the number of correct positive results divided by the number of *all* relevant samples (all samples that should have been identified as positive).

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

F1 Score

The F1 score is the harmonic mean of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0.

$$F1 = \frac{2 * (Precison * Recall)}{(Precison + Recall)}$$

2.
Regression Metrics

Mean Absolute Error

Mean Absolute Error is the average of the difference between the Original Values and the Predicted Values. It gives us the measure of how far the predictions were from the actual output. However, they don't gives us any idea of the direction of the error i.e. whether we are under predicting the data or over predicting the data. Mathematically, it is represented as :

$$MAE = \frac{1}{N}\sum_{j=1}^{N}\left|y_j - \widehat{y}_j\right|$$

Mean Squared Error

Mean Squared Error(MSE) is quite similar to Mean Absolute Error, the only difference being that MSE takes the average of the square of the difference between the original values and the predicted values. The advantage of MSE being that it is easier to compute the gradient, whereas Mean Absolute Error requires complicated linear programming tools to compute the gradient. As, we take square of the error, the effect of larger errors become more pronounced then smaller error, hence the model can now focus more on the larger errors.

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left(y_j - \hat{y}_j\right)^2$$

Root Mean Squared Error

RMSE is the most popular evaluation metric used in regression problems.

$$RMSE = \sqrt{MSE}$$

Practical Time

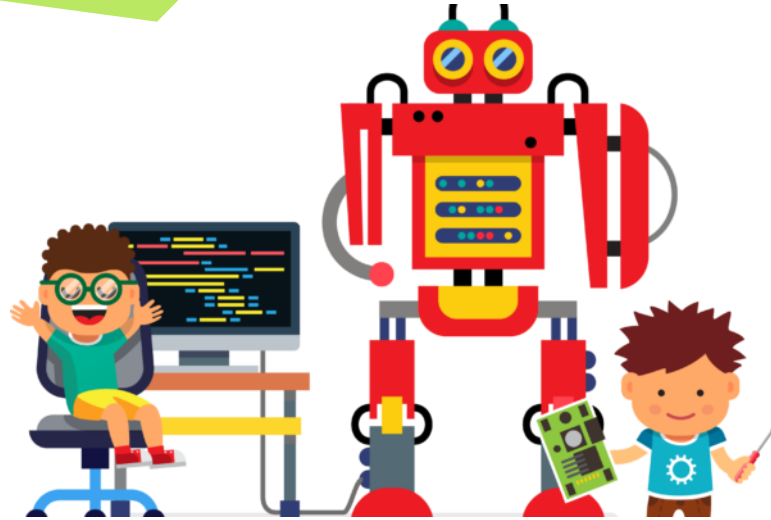*Open up your PC, launch your anaconda and let's create some regression model.*

Coding Interview

Given a string with repeated characters, rearrange the string so that no two adjacent characters are the same. If this is not possible, return None.

For example, given "aaabbc", you could return "ababac". Given "aaab", return None.

شكرا لحضوركم

Thanks for Assisting!