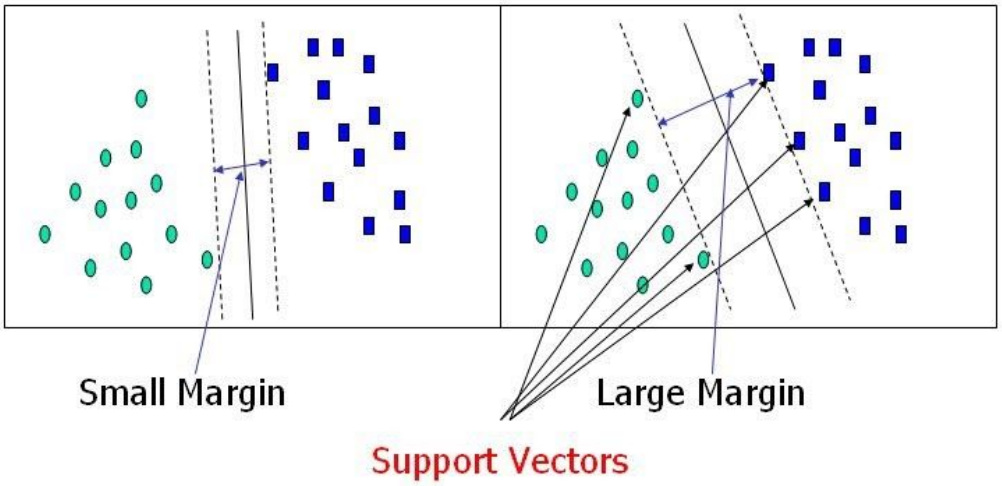
Day 9 of 100 Data Science Interview Questions Series!!

Q 41.) Describe how the support vector machine (SVM) algorithm works, or any other algorithm that you've used.

The objective of the support vector machine algorithm is to find a hyperplane in N-dimensional space(N — the number of features) that distinctly classify the data points.

SVM attempt to find a hyperplane that separates classes by maximizing the margin.



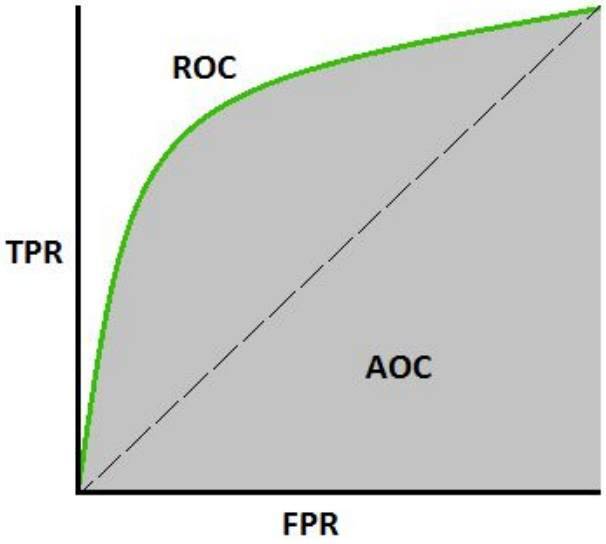
The Edge points in this diagram are the support vectors, against the decision hyperplane. These are the extreme values that represent the data and thus are used to do classification. They in a way support the data, thus known as support vector machine.

Here we show linear classification, but SVMs can perform nonlinear classification. SVMs can employ the kernel trick which can map linear non-separable inputs into a higher dimension where they become more easily separable.

Q 42.) How and when can you use ROC Curve?

The ROC curve is a graphical representation of the contrast between true positive rates and false-positive rates at various thresholds. It is often used as a proxy for the trade-off between the sensitivity(true positive rate) and the false-positive rate. It tells how much the model

is capable of distinguishing between classes. Higher the AUC(area under the curve of ROC), the better the model is at predicting 0s as 0s and 1s as 1s.



Intuitively, in a logistic regression we can have many thresholds, thus what we can do is check the model's performance on every threshold to see which works best. Calculate ROC at every threshold and plot it, this will give you a good measure of how your model is performing.

Q 43.) Give one scenario where false positive is more imp than false negative, and vice versa.

A false positive is an incorrect identification of the presence of a condition when it's absent.

A false negative is an incorrect identification of the absence of a condition when it's actually present.

An example of when false negatives are more important than false positives is when screening for cancer. It's much worse to say that someone doesn't have cancer when they do, instead of saying that someone does and later realizing that they don't.
This is a subjective argument, but false positives can be worse than false negatives from a psychological point of view. For example, a false positive for winning the lottery could be a worse outcome than a false negative because people normally don't expect to win the lottery anyway.

Q 44.) Why we generally use Softmax non-linearity function in last layer but ReLU in rest? Can we switch?

We use Softmax because it takes in a vector of real numbers and returns a probability distribution between 0 and 1, which is useful when we want to do classification.

We use ReLU in all other layers because it keeps the original value and removes all the -ve, max(0,x). This performs better in general but not in every case and can easily be replaced by any other activation function such as tanh, sigmoid, etc.

Q 45.) What do you understand by p-value?

When you perform a hypothesis test in statistics, a p-value can help you determine the strength of your results. p-value is a number between 0 and 1. Based on the value it will denote the strength of the results. The claim which is on trial is called the Null Hypothesis.

Low p-value (≤ 0.05) indicates strength against the null hypothesis which means we can reject the null Hypothesis. A high p-value (≥ 0.05) indicates strength for the null hypothesis which means we can accept the null Hypothesis p-value of 0.05 indicates the Hypothesis could go either way. To put it in another way,

High P values: your data are likely with a true null. Low P values: your data are unlikely with a true null.

- Alaap Dhall

Follow Alaap Dhall on LinkedIn for more insights in Data Science and Deep Learning!!

Visit https://www.aiunquote.com for a 100 project series in Deep Learning.