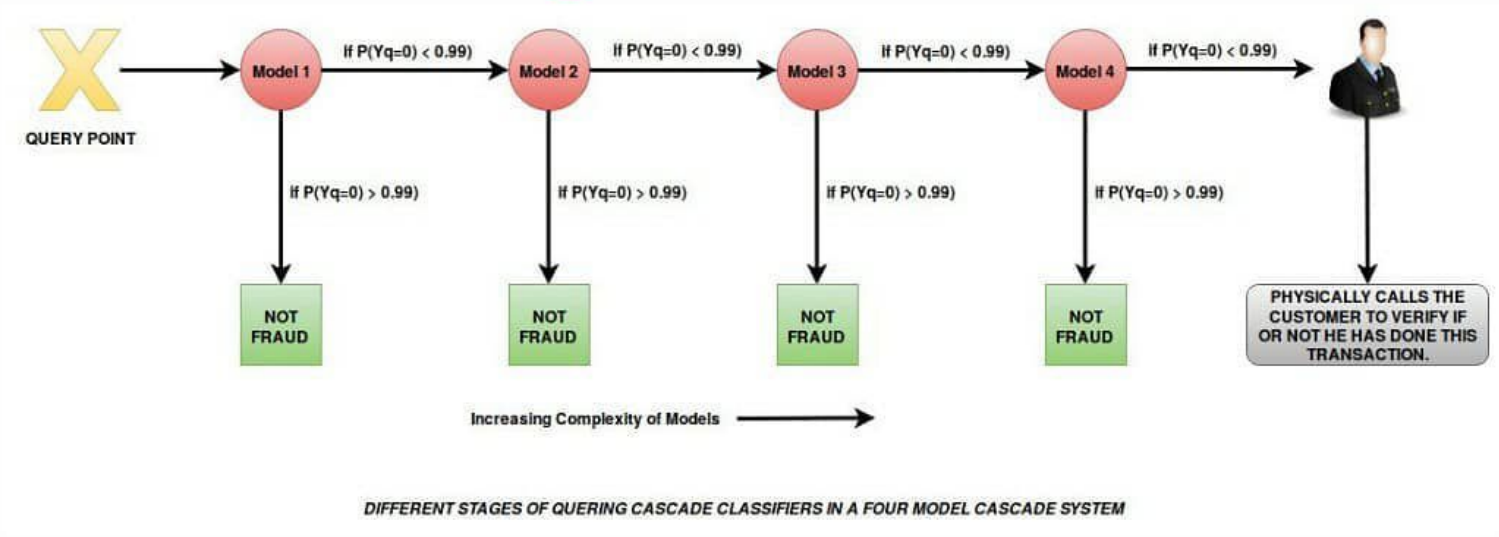
Cascading Ensemble method

- When you completely care about the accuracy (precision and recall)then cascading helps.
- Cascade models are mostly used when the cost of making a mistake is very very high.
- what we do is build a sequence of models to be absolutely sure about the fact that the class is correctly classified or not(binary classification example).
- Here there will be a series of models where you will be sending the data points to the next model whey you are not sure about that data point(prediction class by the previous model).
- At last, if you are still not confident about the data point then you may send this to the domain expert to check.
- We will try to explain this with an example taken from a medium blog by Saugata Paul.

Cascading Ensemble method



DIFFERENT STAGES OF QUERING CASCADE CLASSIFIERS IN A FOUR MODEL CASCADE SYSTEM

- Look at the above diagram. Given that we have a transaction query point Xq, we will feed it to Model 1. Model 1 can be anything a random forest, or a logistic regression model or maybe a support vector machine.
- It can be anything! Basically what Model 1 does is that it predicts class probabilities to determine to which class do a given query point has higher chances of belonging to.

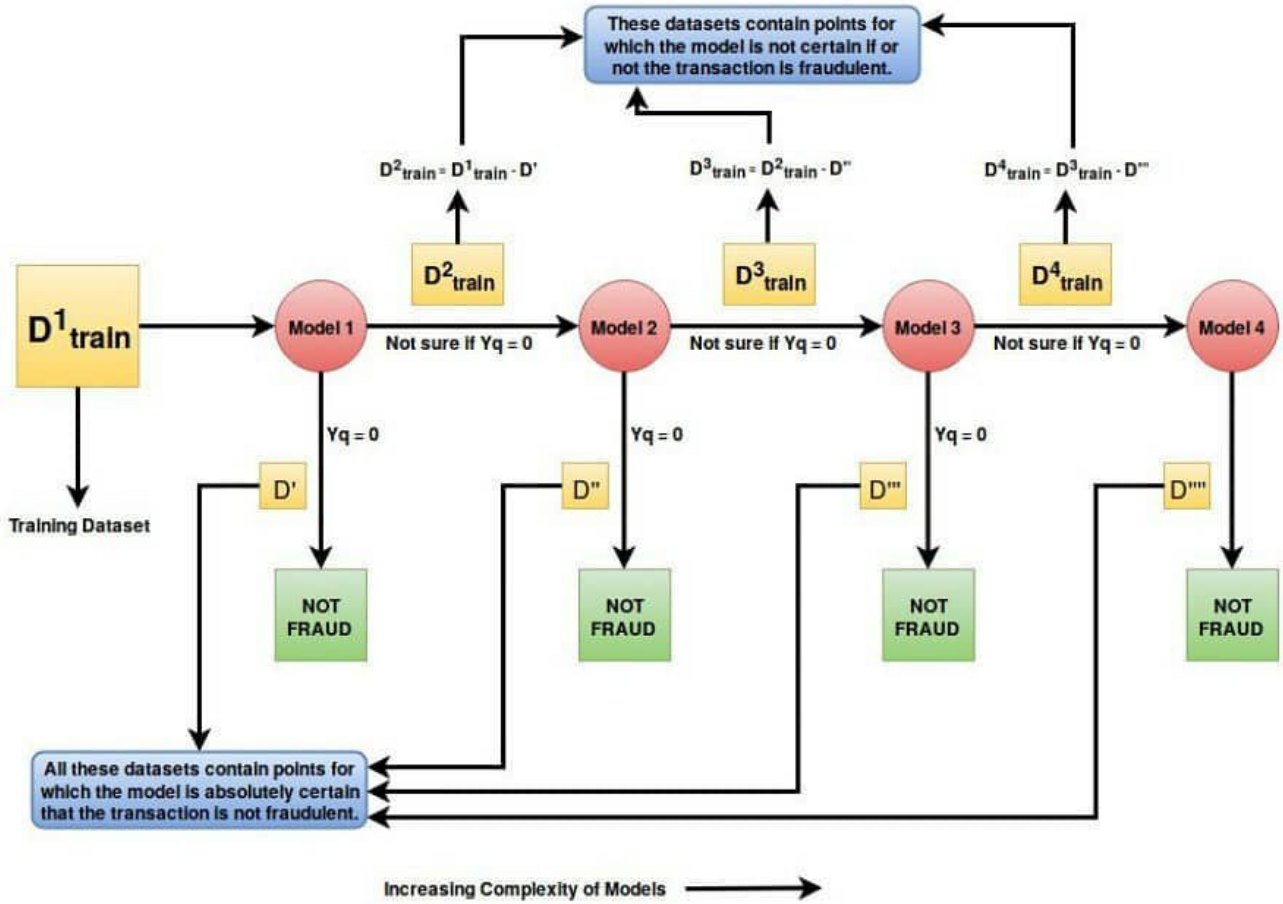
Cascading Ensemble method

- Let's say class label 1 means the transaction is fraudulent, and class label 0 means the transaction is not a fraud. Typically, the predicted probabilities are given by this — P(Yq=0) and P(Yq=1), where Yq is our actual class label. Now let's assume that P(Yq=0), i.e. the probability of the transaction to be not fraudulent is very high. If you think carefully, if P(Yq=0) is extremely high, we will say that the transaction is not a fraud.
- Let's assume we have set a threshold of 99%. It means if and only if P(Yq=0) > 0.99, we will declare the final prediction to be not fraudulent.
- However, if P(Yq=0) < 0.99 we are not very sure if or not it's a fraudulent transaction although though there is a high chance that the transaction is not fraudulent.
- In such a case, when P(Yq=0) < 0.99, we want to be really really sure that the transaction is not fraudulent. We need to be absolutely careful because if our model fails to detect a fraudulent transaction we might lose millions of dollars!

Cascading Ensemble method

- So even when we are slightly unsure, we will train another Model 2. Model 2 does the same thing, it receives the query point and predicts P(Yq=0).
- Just like in stage 1, if P(Yq=0) > 0.99, we will declare the transaction to be not fraudulent and terminate the loop.
- But again if we get P(Yq=0) < 0.99, we aren' sure! Hence, we will pass the query point to another Model 3 in the cascade which does the same thing.
- In a typical cascading system the complexity of models increases as we add more and more models to the cascade.
- Please note that all the models in a cascade are super powerful and has very high accuracy on unseen data. However, it might happen that none of the models can give us a value of P(Yq=0) > 0.99.
- In such a case, typically there is a human being who sits at the end of a cascade. This person will personally call the customer and ask him whether or not he has done the transaction.

Cascading Ensemble method



DIFFERENT STAGES OF TRAINING CASCADE CLASSIFIERS IN A FOUR MODEL CASCADE SYSTEM