Patel Dipam J.

18011011 6036

D. J. Patel

Date

Page

Q.6 Bayes' Rule :-

→ Bayes' theorem has been called the most powerful rule of probability and statistics.

→ It describes the probability of an event, based on prior knowledge of conditions that might be related to the event.

$$\underbrace{P(A/B)}_{\text{posterior}} = \frac{\overbrace{P(B/A)}^{\text{Like Hood}} \cdot \overbrace{P(A)}^{\text{prior (probability of A occuring)}}}{\underbrace{P(B)}_{\text{marginalization. (probability of B occuring)}}}$$

for example, If a disease is related to age, then using Bayes' theorem, a person's age can be used to more accurately assess the probability that they have the disease, compared to the assessment of the probability of disease made without knowledge of the person's age.

→ It is a powerful law of probability that brings concept of 'subjectivity' or 'the degree of belief' into the cold, hard statistical modeling. Baye's Rule is the only mechanism that can be used to gradually the probability of an event as the evidence or data gathered sequentially.

Patel Dipam J.
180110116036
D.J.Patel

→ It is a powerful law of probability that brings in the concept of 'subjectivity' or 'the degree of belief' into the cold, hard statistical modeling.

→ Bayesian statistics and modeling have had a recent resurgence with the global rise of AI and data driven machine learning system.

→ formula for Bayes' theorem

$$\boxed{P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A).\,P(B/A)}{P(B)}}$$

P(A) = The probability of A occuring
P(B) = The probability of B occurring
P(A/B) = The probability of A given B
P(B/A) = The probability of B given A
P(A∩B) = The probability of both A and B occurring.

Patel Dipam J.
180110116036
D.J.patel

Date
Page

Q-1 CROSS - Validation :-

→ The main drawback of random sampling is, it does Not have control over the No. of times each tuple is used for training and testing.

→ There are two variations in the cross validation method:

① K-fold cross - validation
② N-fold cross - validation.

① K-fold cross - validation :-

→ Data set consisting of N tuples is divided into K equal, mutually exclusive parts, and if Not divisible by K, then the last part will have fewer tuples then other (K-1) parts.

→ A series of K runs in carried out with this decomposition, and in i^{th} iteration D_i is used as test Data and other folds as tranning data.

② N-fold cross validation :-

→ N-fold cross - validation is an extreme case of K-fold often known as "have one out" cross - validation.

→ Here, dataset is divided into as many folders are there are instances. Thus, all most each tuple forming a training set, building N classifiers.

Scanned with CamScanner

Patel Dipam J.

180110116036

Date
Page

D. J. Patel

→ In this method, therefore N classifiers are built from N-1 instances and each type is used to classify a single test instances

(*) Advantages of cross-validation:

1. more accurate estimate of out of sample accuracy.
2. more "efficient" as of data as every observation is used for both traning and testing.

Patel Dipam J.
180110116036
D.J.Patel

Q-5 why overfitting happens ? Explain it with example.

⇒ Model is too "complex" and fits irrelevent characteristic (noise) in the data.

→ overfitting happens when a model learns the detail and noice in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the ~~training~~ learning data is picked up and learned as concepts by the model. the problem is that these concept by the model. The problem is that these concepts do not apply to new data and negative impact the model ability to generalize.

→ overfiting is more likely with non-parametric and non-linear models that have more flexibility when learning a target function as such, many non-parametric or technique to limit and constrain how much detail the model learns

→ for example, decision trees are a non-parametric machine learning algorithm that is very flexible and is subject to overfitting training data this problem can be addressed by pruning a tree after it has learned in order to remove some of the detail it has picked up.

Patel Dipam J.
180110116036
D.J.Patel

Date
Page

Q-4 Inductive machine learning :-

→ From the perspective of inductive machine learning, we are given input sample(x) and output sample (f(x)) and problem is to estimate the function f. specifically, the problem is to generalize from the samples and the mapping to be useful to estimate the output for new sample in the futher.

→ In practive it is almost always too hard to estimate the function, so we are looking for very good approximation of the function.

ex (1) credit risk assesment

→ The x is properties of the customer.

→ The f(x) is credit approved or Not.

(2) face recognition.

→ The x are bitmakes & people faces.

→ The f(x) is to assign a name to the face.

(3) Automatic steering:

→ The x are bit map images from a camera in front of the car.

→ The f(x) is the degree the steering wheel should be turned.

(4) Disease diagonosis:

→ The x are the properties of the patient.

→ The f(x) is the disease they suffer from.