1.what is bayes theorem?

Bayes' Theorem is a simple mathematical formula used for calculating conditional probabilities. It figures prominently in subjectivist or Bayesian approaches to epistemology, statistics, and inductive logic.

2.Define "the true error of hypothesis".

The true error of hypothesis h with respect to target function f and distribution D is the probability that h will misclassify an instance drawn at random according to D.

3.Interpret "Instance based learning".

In **machine learning**, **instance-based learning** (sometimes called memory-**based learning**) is a family of **learning** algorithms that, instead of performing explicit generalization, compares new problem **instances** with **instances** seen in training, which have been stored in memory.

4.Explain Gaussian function in a few words.

In mathematics, a Gaussian function, often simply referred to as a Gaussian,

for arbitrary real constants a, b and non zero c. It is named after the mathematician Carl Friedrich Gauss. The graph of a Gaussian is a characteristic symmetric "bell curve" shape. The parameter a is the height of the curve's peak, b is the position of the center of the peak and c (the standard deviation, sometimes called the Gaussian RMS width) controls the width of the "bell".

Gaussian functions are often used to represent the probability density function of a normally distributed random variable with expected value $\mu = b$ and variance $\sigma^2 = c^2$. In this case, the

Gaussian functions are widely used in statistics to describe the normal distributions, in signal processing to define Gaussian filters, in image processing where two-dimensional Gaussians are used for Gaussian blurs, and in mathematics to solve heat equations and diffusion equations and to define the Weierstrass transform.

5.What is the Horn clause in first order logic?

A Horn clause is either a definite clause or an integrity constraint. That is, a Horn clause has either false or a normal atom as its head. Integrity constraints allow the system to prove that some conjunction of atoms is false in all models of a knowledge base - that is, to prove disjunctions of negations of atoms.

6.Define Maximum likelihood Hypothesis.

Maximum likelihood estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed.

7.When a concept C is said to be PAC-learnable?

A target concept class C is PAC-learnable by H (the hypothesis class) if there exists an algorithm A which for all concepts $c \in C$, all distributions D, and all $\epsilon > 0, \delta > 0$ takes m = poly($1/\epsilon, 1/\delta, \ldots$) ... The benefit of a PAC learning model is that it actually makes a real connection to learning.

8.Interpret "curse of dimensionality" in K-NN.

The **curse of dimensionality** basically means that the error increases with the increase in the number of features. It refers to the fact that algorithms are harder to design in high dimensions and often have a running time exponential in the dimensions.

9.Explain case-based reasoning in a few words.

As we know Nearest Neighbour classifiers store training tuples as points in Euclidean space. But Case-Based Reasoning classifiers (CBR) use a database of problem solutions to solve new problems. It stores the tuples or cases for problem-solving as complex symbolic descriptions.

10.What is substitution in first order logic?

We want to define and investigate the effect of substituting terms t1,…,tk for variables v1,…,vk in a term or formula. The work here will be entirely syntactical (i.e. manipulations of strings of symbols without meaning attached) but looking ahead to further work we will take care to ensure that our substitutions are potentially meaningful (i.e. can at a later stage be given meaning).

We fix a first order language L with variables from VARL={u},v,w,....and terms TERML.

11.Define least squared error hypothesis.

It's a supervised learning algorithm that takes an input vector, $X^T = (X1, X2, ..., Xp)$, and wants to predict the output Y.

The linear model assumes that the regression function **E(Y|X)** is linear or a reasonable approximation. The **β** are called parameters or coefficients and the variables X can be of different types:

- Quantitative variables that can be transformed with logs, polynomial representations, etc...
    - Numeric or "dummy" variables coding the distinct levels of a qualitative input.
    - Interactions between other variables.

12.Define Minimum description length principle.

The minimum description length (MDL) principle is a powerful method of inductive inference, the basis of statistical modeling, pattern recognition, and machine learning. It holds that the best explanation, given a limited set of observed data, is the one that permits the greatest compression of the data.

13.List radial basis functions.

A radial basis function (RBF) is a function that assigns a real value to each input from its domain (it is a real-value function), and the value produced by the RBF is always an absolute

value; i.e. it is a measure of distance and cannot be negative. ... represents a radial basis function network.

14.what is induction?

Induction is the process of inferring general rules from specific data and is the primary task of **machine learning**. ... We treat **induction** and abduction as two distinct reasoning tasks, but have demonstrated that each can be of direct service to the other in developing AI systems for solving real-world problems.

15.What is the term in first order logic?

There are two key types of well-formed expressions: terms, which intuitively represent objects, and formulas, which intuitively express predicates that can be true or false. The terms and formulas of first-order logic are strings of symbols, where all the symbols together form the alphabet of the language.