# ASSIGNMENT 03

# PROBABILITY AND STATISTICS/PROBABILITY METHODS IN ENGG.

# [CLO1, PLO1]

**Q1. Describe at least 2 applications of Bayesian Probability in Computer Science / Software Engineering.**

## Applications of Bayesian Statistics:

In Bayesian Statistics the probability is updated for a hypothesis as more evidence or Information becomes available. In simple words, all the data including old and new is used to make a better prediction. This can be used in Computer Software too such as given below two applications

1. **Google Search (Auto Complete) Prediction:**

It predicts the search that user wants to make using the words they have already typed in the search bar.

2. **Image Recognition:**

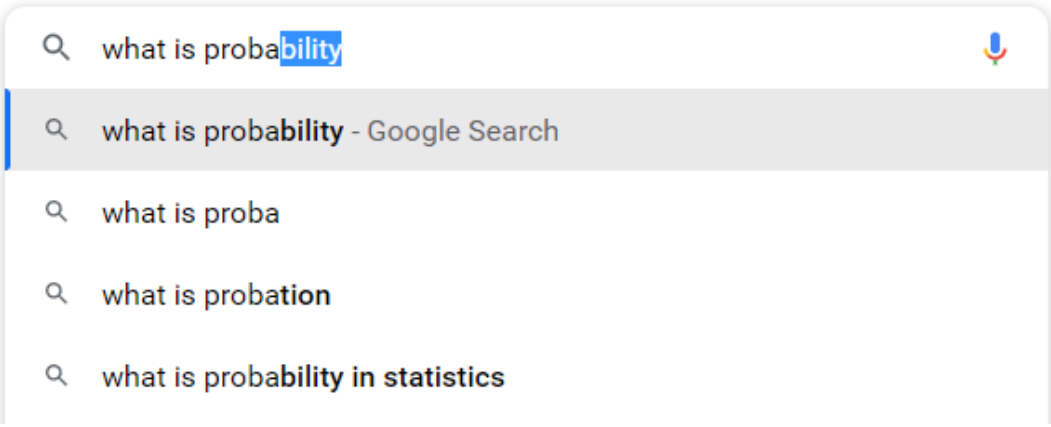
It is used to improve Image Recognition by making a better prediction.

**Q2. Describe the mechanism which is used to apply Bayesian Probability in your stated applications in Q1.**

## Mechanism of Bayesian Statistics:

### Google Search (Auto Complete) Prediction:

Whenever we type a word, the next word with highest probability is shown in a list below. The Prediction is made using the "Most Searches made following the word just typed in" of all the users. It also takes into account the searches made by the current user and gives his search history more priority.
Example:
The user typed in "What is prob". The probability according to most searches following these words are




Since more priority is given to the searches of current user, they are displayed on top what he want.

## Image Recognition:

Image is a 2d array of pixel. Each pixel is translated by computer by giving each pixel a different value “Red, Green, Blue”. Furthermore, each image goes through filtering to reduce noise in the image. Once the image is translated in a compute readable format then prediction
Is made of similar images.

Bayesian Statistics can improve the prediction of images by making the machine learn from previous experiences. Different images are feed to the system. If the result of the image is incorrect then the correct result is shown. In future if the similar image is shown to the system then it will know from its previous experience the correct result.

It should be noted that this type of machine learning is very expensive since it takes a lot of time and processing power.
Example:

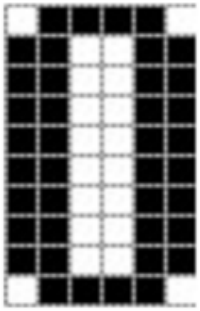


**FIGURE 2.** Computer Pixel Schematic Diagram

## BAYESIAN STATISTICAL THEORY

### Bayesian formula

Assume two events as A and B, respectively, and satisfy:

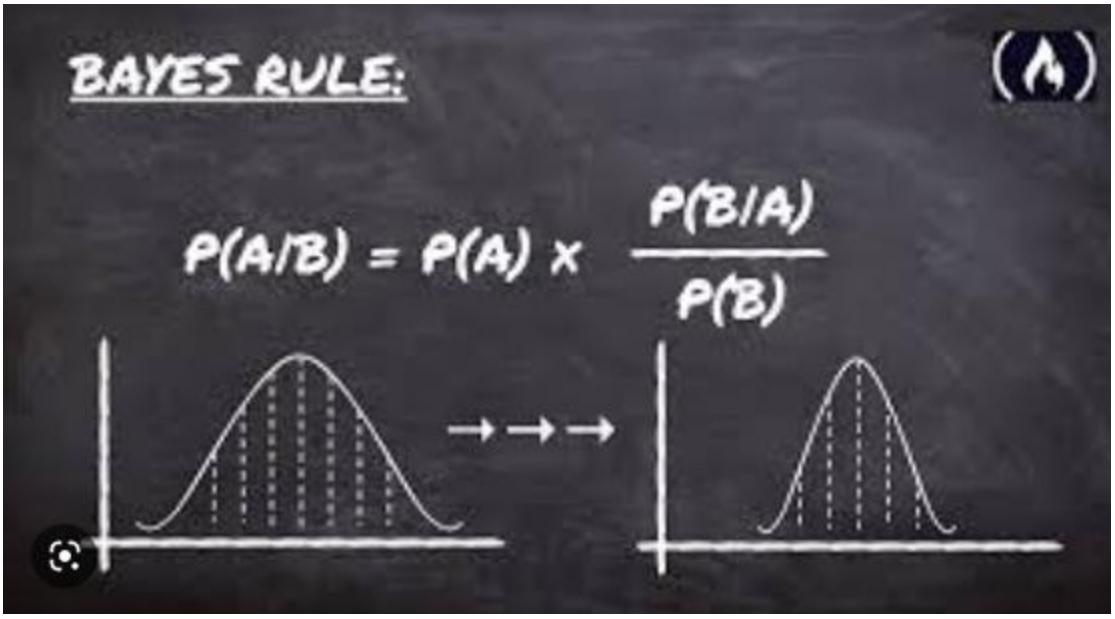
$A = A_1 \cup \cdots \cup A_n, A_i \cap A_j = \varnothing, \forall i \neq j,$

Under condition B, the conditional probability of event Ai is:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}, \quad (1)$$

Formula (1) is called the Bayesian formula [5].
More generally, the formula can be written as,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \propto P(B|A)P(A) \quad (2)$$



Bayesian decision-making theory is a basic method in statistical pattern recognition. Given a sample set containing two pattern classes and the statistical distribution of each sample in n-dimensional feature space, there is a need to solve the problem of how to classify the recognized samples. The key idea of Bayesian decision theory is to use the Bayesian formula to convert the conditional probability density function and prior probability into posterior probability and make decisive classification according to the posterior probability or loss function size. The x vector is the sample, ωi = 1, 2, 3…c.
From the known prior probability P (ωi) and the class conditional probability density P (x|ωi), the posterior probability P (ωi|x) can be obtained. From above it can be seen that the rule of Bayesian decision based on the minimum error rate is:

$$P(\omega_i|x) = \max_{j=1,2,\cdots,c} P(\omega_i|x), \qquad x \in \omega_i$$