

- XAMPP is a free and open-source cross-platform software package that provides a local web server environment for developing and testing dynamic web applications. The name "XAMPP" stands for the software components that are included in the package: X (cross-platform), Apache (web server), MySQL (database), PHP (scripting language), and Perl (scripting language).
- PHP (Hypertext Preprocessor) is a popular server-side scripting language used primarily for web development. It is an open-source, general-purpose programming language that is designed to create dynamic web pages and applications. PHP code can be embedded in HTML code, and it is executed on the server-side, generating HTML code that is sent to the client's web browser.
- Asymptotic notation is a mathematical notation used to describe the time or space complexity of an algorithm in computer science. It is used to describe the performance of an algorithm as the size of the input increases, without specifying the exact amount of time or space required.
- The brute force approach is a straightforward method of problem-solving that involves trying all possible solutions systematically until the correct one is found. In computer science, it often refers to a method for solving a problem by checking all possible combinations of input values.
- Regression is a statistical modeling technique used to analyze the relationship between a dependent variable and one or more independent variables. It is commonly employed to make predictions or estimate the value of the dependent variable based on the values of the independent variables.
- Probably Approximately Correct (PAC) learning is a framework in computational learning theory that addresses the problem of learning from data in a probabilistic setting. It provides a theoretical foundation for understanding the trade-off between the sample complexity (the number of training examples needed) and the generalization error (the error on unseen data) of a learning algorithm.
- The GEMINI algorithm is a parallel and distributed algorithm used for similarity search in large-scale metric spaces. It is designed to efficiently perform k-nearest neighbor (k-NN) search and range search operations on high-dimensional datasets. The algorithm was proposed by Zhan et al. in their paper titled "GEMINI: A Novel Grid-Based Multi-Index for Efficient Nearest Neighbor Search in High-Dimensional Data" (2018).
- Boolean search is a search technique that allows users to combine keywords using Boolean operators to perform precise and targeted searches. The Boolean operators, such as AND, OR, and NOT, are used to create logical relationships between keywords or search terms. Here's how each operator works:
- A one-dimensional time series refers to a sequence of data points that are collected and recorded at successive points in time. It represents the variation of a single variable over time. The data points in a one-dimensional time series are typically recorded at regular intervals, such as seconds, minutes, hours, days, months, or years.
- MULTOS (Multiple Application Operating System) is an operating system designed for smart cards and secure embedded devices. It is an open and independent platform that provides a secure environment for running multiple applications on a single smart card or embedded device. Here are some key features and aspects of MULTOS:
- Huffman coding is a variable-length prefix coding technique that assigns shorter codes to frequently occurring symbols and longer codes to less frequent symbols. It relies on constructing a binary tree called a Huffman tree based on the frequency distribution of symbols in the input data. The steps involved in Huffman coding are as follows:

1. Frequency Analysis: The frequency of occurrence of each symbol in the input data is analyzed to determine its relative importance.
 2. Building the Huffman Tree: A Huffman tree is constructed based on the frequency distribution. This tree is a binary tree where each leaf node represents a symbol, and the path from the root to each leaf node corresponds to the binary code for that symbol. The most frequent symbols have shorter codes, reducing the overall size of the encoded data.
 3. Generating the Huffman Codes: The Huffman codes are assigned to each symbol by traversing the Huffman tree. The codes are obtained by concatenating the binary values of the edges along the path from the root to the corresponding leaf node.
 4. Encoding and Decoding: The input data is encoded by replacing each symbol with its corresponding Huffman code. The encoded data is then decoded using the same Huffman tree to recover the original data.
- Arithmetic coding is another variable-length coding technique that assigns non-uniform length codes to symbols based on their probabilities. Unlike Huffman coding, arithmetic coding encodes the entire message as a single code, resulting in a more compact representation. The steps involved in arithmetic coding are as follows:
 1. Interval Mapping: Each symbol is mapped to an interval on the real number line based on its cumulative probability. The size of each interval is proportional to the probability of the corresponding symbol.
 2. Subdividing the Interval: The interval representing the entire message is successively subdivided based on the probabilities of the symbols. The subdivision process narrows down the interval to a specific subinterval representing the encoded message.
 3. Encoding and Decoding: The subinterval representing the encoded message is converted into a binary representation, which becomes the compressed output. Decoding involves reversing the process to recover the original message.
 - Harvest is an open-source distributed web crawler and data extraction framework developed by the Common Crawl Foundation. It follows a distributed architecture to efficiently crawl and collect web data at a large scale. The architecture of Harvest consists of several components working together to achieve high-performance web crawling and data extraction. Here is an overview of the key components of the Harvest distributed architecture:
 - S/MIME stands for "Secure/Multipurpose Internet Mail Extensions." It is a widely used security protocol that provides a way to secure and authenticate email messages. S/MIME is primarily used for encrypting and digitally signing email messages to ensure the confidentiality, integrity, and authenticity of the message content. Here are the key components and functions of S/MIME