DATA MINING

Introduction

1) What is Data Mining?

- **Ans) The process of extracting knowledge or insights from large amounts of data using various statistical and computational techniques.**
- **The data can be structured, semi-structured or unstructured, and can be stored in various forms such as databases, data warehouses, etc.**
- **The primary goal of data mining is to discover hidden patterns and relationships in the data that can be used to make informed decisions or predictions.**

2) Applications of Data Mining.

- **Ans) Marketing: Data mining can be used to identify customer segments and target marketing campaigns.**
- **Healthcare: It can be used to identify risk factors for diseases and develop personalized treatment plans.**

3) What is Business Intelligence?

- **Ans) A set of concepts and methodologies to improve decision-making in business through the use of facts and fact-based systems.**
- **The Goal of Business Intelligence is to improve decision-making in business ideas and analysis.**
- **BI (Business Intelligence) uses a set of processes, technologies, and tools (such as Informatica/IBM) to transform raw data into meaningful information and then transform information to provide knowledge.**

4) Explain the Types of Users of Business Intelligence.

Ans) Types of Users of Business Intelligence:

(i) Analyst (Data Analyst or Business Analyst): They are the statistician of the company, they use BI on the basis of historical data priorly stored in the system.

(ii) Head or Manager of the Company: Used to increase the profitability of their company by increasing the efficiency in their decisions on the basis of all the knowledge they discovered.

(iii) IT Engineer: For his company.

(iv) Small Business Owners: Can be used by a small businessman because it is quite affordable too.

(v)Government Officials: In the decision-making of the government.

5) Explain the Advantages of Data Mining.

Ans) Advantages of Data Mining:

- Data mining can uncover hidden patterns and relationships in data that are not immediately apparent.
- Data mining can be used to identify new opportunities and improve business processes.
- Data mining can help organizations predict future outcomes based on historical data.

6) Explain the Disadvantages of Data Mining.

Ans) Disadvantages of Data Mining:

- Data mining requires significant technical expertise and specialized software.
- Data mining can be computationally intensive and may require significant computing resources.
- Data mining can produce false positives or overfitting if not carefully validated.

7) Explain the similarities between Business Intelligence (BI) and Data Mining.

Ans) Similarities between Business Intelligence (BI) and Data Mining:

- Both BI and data mining use data to provide insights and support decision-making.
- Both technologies require a significant amount of data to be effective.
- Both BI and data mining can be used to identify patterns and trends in data.

DATA MINING

DATA PREPROCESSING

1) What is Knowledge Discovery in Databases (KDD) Process?

Ans) Knowledge Discovery in Databases (KDD) Process:

A method of finding, transforming, and refining meaningful data and patterns from a raw database in order to be utilized in different domains or applications.

2) Write the Steps in typical KDD Process.

Ans) Steps in typical KDD Process:

(i) Goal-Setting and Application Understanding: Requires prior understanding and knowledge of the field to be applied in.

This is where we decide how the transformed data and the patterns arrived at by data mining will be used to extract knowledge.

(ii) Data Selection and Integration: The data collected needs to be selected and segregated into meaningful sets based on availability, accessibility importance and quality.

(iii) Data Cleaning and Preprocessing: It involves searching for missing data and removing noisy, redundant and low-quality data from the data set in order to improve the reliability of the data and its effectiveness.

Certain algorithms are used for searching and eliminating unwanted data based on attributes specific to the application.

(iv) Data Transformation: This step prepares the data to be fed to the data mining algorithms.

Hence, the data needs to be in consolidated and aggregate forms. The data is consolidated on the basis of functions, attributes, features etc.

(v) Data Mining: This is the root or backbone process of the whole KDD.

This is where algorithms are used to extract meaningful patterns from the transformed data, which help in prediction models.

It is an analytical tool which helps in discovering trends from a data set using techniques such as artificial intelligence, advanced numerical and statistical methods and specialized algorithms.

(vi) Pattern Evaluation/Interpretation: Once the trend and patterns have been obtained from various data mining methods and iterations, these patterns need to be represented in discrete forms such as bar graphs, pie charts, etc. to study the impact of data collected and transformed during previous steps.

This also helps in evaluating the effectiveness of a particular data model in view of the domain.

3) What is Data Cleaning? Explain.

Ans) DATA CLEANING: Defines to clean the data by filling in the missing values, smoothing noisy data, analyzing and removing outliers, and removing inconsistencies in the data.

Sometimes data at multiple levels of detail can be different from what is required.

Ex: It can need the age ranges of 20-30, 30-40, 40-50, and the imported data includes birth date. The data can be cleaned by splitting the data into appropriate types.

4) Types of Data Cleaning.

Ans) Types of data cleaning:

(i) Missing Values: Missing values are filled with appropriate values. There are the following approaches to fill the values.

The tuple is ignored when it includes several attributes with missing values.

The values are filled manually for the missing value.

The same global constant can fill the values.

The attribute mean can fill the missing values.

The most probable value can fill the missing values.

(ii) Noisy Data: Noise is a random error or variance in a measured variable.

(iii) Smoothing methods to handle noise:

• Binning – These methods smooth out a arrange data value by consulting its "neighborhood," especially, the values around the noisy information.

• The arranged values are distributed into multiple buckets or bins because binning methods consult the neighborhood of values, they implement local smoothing.

(iv) Regression:

• Data can be smoothed by fitting the information to a function, including with regression.

• Linear regression: Contains finding the "best" line to fit two attributes (or variables) so that one attribute can be used to forecast the other.

• Multiple linear regression: A development of linear regression, where more than two attributes are contained and the data are fit to a multidimensional area.

(v) Clustering:

• Clustering supports in identifying the outliers.

• The same values are organized into clusters and those values which fall outside the cluster are known as outliers.

5) TERMS RELATED TO DATA.

Ans) TERMS RELATED TO DATA:

• ATTRIBUTE: Property or characteristic of an object. Also called as variable, field, characteristic or feature. Ex: Eye color of a person, temperature, etc.

• INSTANCE: Collection of attributes describing an object. Also known as record, point, case, sample or entity.

6) TYPES OF ATTRIBUTES.

Ans) TYPES OF ATTRIBUTES:

(i) Nominal: This type of data is also referred to as categorical data that cannot be measured or compared with numbers. Ex: Gender, race, religion, and occupation.

(ii) Ordinal: Qualitative data that can be ranked in a particular order. Ex: Ranks (scale of 0 to 10), Grades, Height (tall, short, medium), etc.

(iii) Interval: Quantitative data with equal intervals between consecutive values. Ex: Time, dates, temperature, etc.

(iv) Ratio: Similar to interval data, but with an absolute zero point. Ex: height, weight, and income.

7) Properties of Attribute Values.

Ans) Properties of Attribute Values:

Type of an attribute depends on which of the following properties it

possesses:

Distinctness: =, ≠

Order: <, >

Addition: +, -

Multiplication: *

Nominal Attribute: Distinctness.

Ordinal Attribute: Distinctness and Order.

Interval Attribute: Distinctness, Order and Addition.

Ratio Attribute: All 4 attributes.

8) Discrete and Continuous Attributes.

Ans) Discrete and Continuous Attributes:

- **Discrete: Number of words in a given sentence.**
- **Continuous: Temperature, height, etc.**
- **We an convert continuous to discrete by dividing the continuous values in intervals. This is called Discretization.**

Information Security

1) State and explain principals of security.

Ans i) Confidentiality: Information is not disclosed to unauthorized individuals, entities and process. Ex: Gmail password

(ii) Integrity: Maintaining accuracy and completeness of data. Data cannot be edited in an unauthorized way. Ex: If person leaves the job.

(iii) Availability: Information must be available when needed. Ex: To access information of a particular employee.

(iv) Non-repudiation: One party cannot deny receiving a message or a transaction nor can the other party deny sending a message or a transaction.

•   Ex: In cryptography it is sufficient to show that message matches the digital signature signed with sender's private key and that sender could have a sent a message and nobody else could have altered it in transit.

(v) Authenticity: Verifying that users are who they say they are and that each input arriving at destination is from a trusted source.

Ex: If sender sends the message along with digital signature which was generated using the hash value of message and private key. Now at the receiver side this digital signature is decrypted using the public key generating a hash value and message is again hashed to generate the hash value. If the 2 value matches then it is known as valid transmission with the authentic or we say genuine message received at the recipient side.

(vi) Accountability: It should be possible to trace actions of an entity uniquely to that entity. Ex: Not every employee should be allowed to do changes in other employees data. For this there is a separate department in an organization that is responsible for making such changes. Approval from higher authority.

2) What is non-repudiation?

Ans) Non-repudiation: One party cannot deny receiving a message or a transaction nor can the other party deny sending a message or a transaction.

• Ex: In cryptography it is sufficient to show that message matches the digital signature signed with sender's private key and that sender could have a sent a message and nobody else could have altered it in transit.

3) State some advantages and disadvantages of Information Security.

Ans) Advantages:

(i) Improved security: By identifying and classifying sensitive information, organizations can better protect their most critical assets from unauthorized access or disclosure.

(ii) Compliance: Many regulatory and industry standards require organizations to implement information classification and data protection measures.

(iii) Improved efficiency: By clearly identifying and labelling information, employees can quickly and easily determine the appropriate handling and access requirements for different types of data.

(iv) Better risk management: By understanding the potential impact of a data breach or unauthorized disclosure, organizations can prioritize resources and develop more incident response plans.

Disadvantages:

(i) Complexity: Developing and maintaining an information classification system can be complex and time-consuming, especially for large organizations with a diverse range of data types.

(ii) Cost: Implementing and maintaining an information classification system can be costly, especially if it requires new hardware or software.

(iii) Resistance to change: Some employees may resist the implementation of an information classification system, especially if it requires them to change their usual work habits.

(iv) Lack of flexibility: Information classification systems can be rigid and inflexible, making it difficult to adapt to changing business needs or new types of data.

4) What are different types of criminal attacks?

Ans) TYPES OF CRIMINAL ATTACKS:

(i) FRAUD: Credit cards, ATMs, stock certificates, etc.

(ii) SCAMS: Nigeria Scam.

(iii) DESTRUCTION: Unhappy employees.

(iv) IDENTITY THEFT: Stealing password.

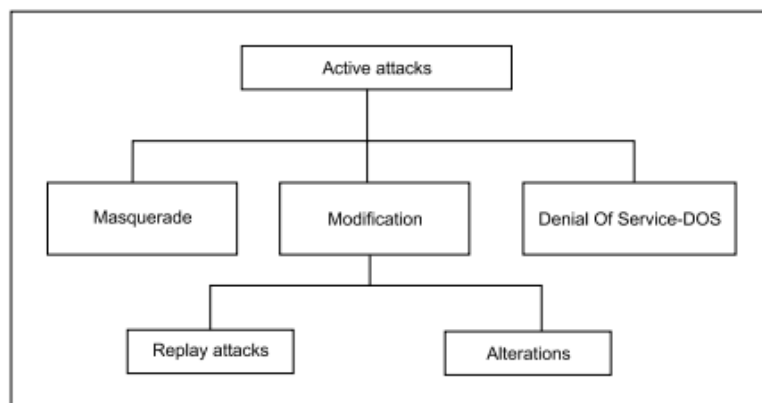**(v) INTELLECTUAL PROPERTY THEFT: Digital videos, music, software, sounds, etc.**

(vi) BRAND THEFT: Setting up fake website.

5) State and explain active attacks.

OR

Write a short note on active attacks.

Ans) Active attacks: Based on the modification of the original message in some manner, or in the creation of false message.

- **Masquerade Attack: When an unauthorized entity pretends to be another entity. The attack may involve capturing the user's authentication sequence (e.g. user ID and password). Later, those details can be replayed to gain illegal access to the computer system.**
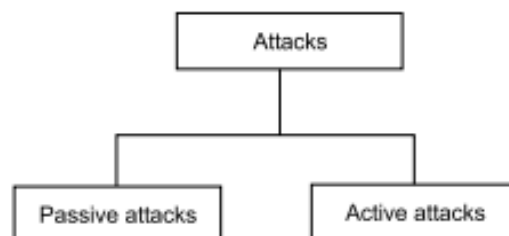
- **Modification:**

  - ➤ **Replay Attack: A user captures a sequence of events, or some data units, and re-sends them.**

➢ **Alteration of messages: Involves some change to the original message.**

• **Denial Of Service (DOS): Attacks make an attempt to prevent legitimate users from accessing some services, which they are eligible for.**

➢ **For instance, an unauthorized user might send too many login requests to a server using random user ids in quick succession, so as to flood the network and deny other legitimate users to use the network facilities.**
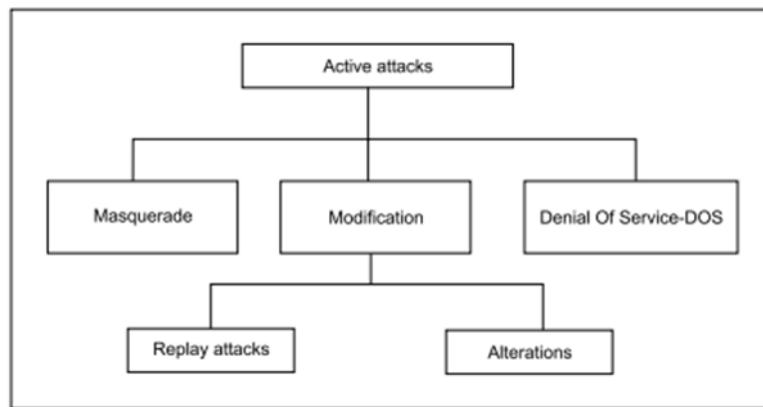
6) Differentiate between active attacks and passive attacks.

Ans)

(i) Passive Attacks: Passive attacks are those wherein the attacker indulges in monitoring of data transmission. Passive attacks do not involve any modifications to the contents of an original message.
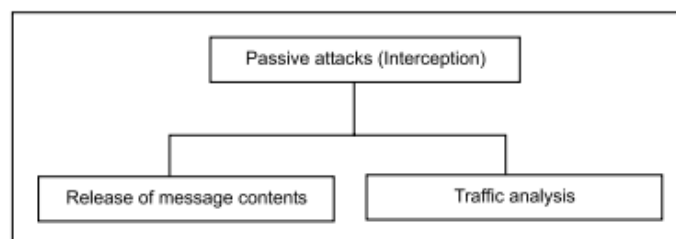

• **Release of message contents: When the contents of the message is confidential, only the sender and the receiver should have an access. Release of message to someone else refers to Release of message contents. We can encode messages to prevent such problem.**

• **Traffic-analysis attack attempts of analysing (encoded) messages. A passive attacker could try to come up with some sort of pattern that provides some clues regarding the communication that is taking place.**

(ii) Active attacks: Based on the modification of the original message in some manner, or in the creation of false message.

• **Masquerade Attack: When an unauthorized entity pretends to be another entity. The attack may involve capturing the user's authentication sequence (e.g. user ID and password). Later, those details can be replayed to gain illegal access to the computer system.**

• **Modification:**

➢ **Replay Attack: A user captures a sequence of events, or some data units, and re-sends them.**

➢ **Alteration of messages: Involves some change to the original message.**

• **Denial Of Service (DOS): Attacks make an attempt to prevent legitimate users from accessing some services, which they are eligible for.**

➢ **For instance, an unauthorized user might send too many login requests to a server using random user ids in quick succession, so as to flood the network and deny other legitimate users to use the network facilities.**

**7) Explain different types of programs that attacks.**

Ans i) Virus: A virus is a computer program that attaches itself to another legitimate program, and causes damage to the computer system or to the network.
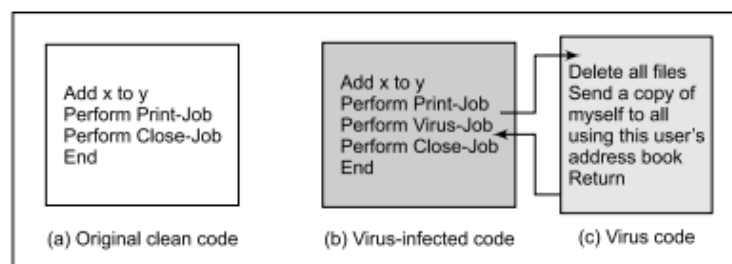
(ii) Worm: A worm does not perform any destructive actions, and instead, only consumes system resources to bring it down.

Worm
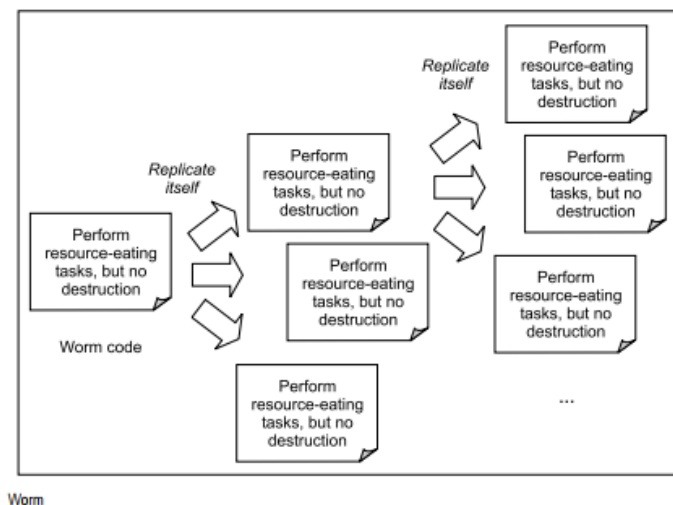
(iii) Trojan Horse: A Trojan horse is a hidden piece of code. a trojan horse attempts to reveal confidential information to an attacker.

Trojan horse
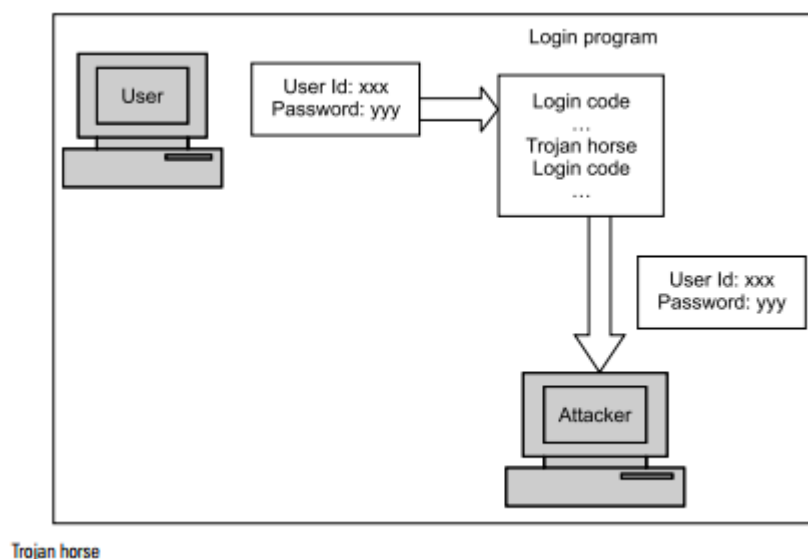
8) What do you mean by sniffing and spoofing?

Ans i) Sniffing and Spoofing: On the Internet, computers exchange messages with each other in the form of small groups of data, called packets.

A packet, like a postal envelope contains the actual data to be sent, and the addressing information.

Attackers target these packets, as they travel from the source computer to the destination computer over the Internet.

(ii) Packet Sniffing: Packet sniffing is a passive attack on an ongoing conversation.

An attacker need not hijack a conversation, but instead, can simply observe (i.e. sniff) packets as they pass by.

(iii) Packet Spoofing: In this technique, an attacker sends packets with an incorrect source address.

9) Define cryptography and crypt-analysis with the help of a neat diagram.

Ans) Cryptography is the art of achieving security by encoding messages to make them non-readable.
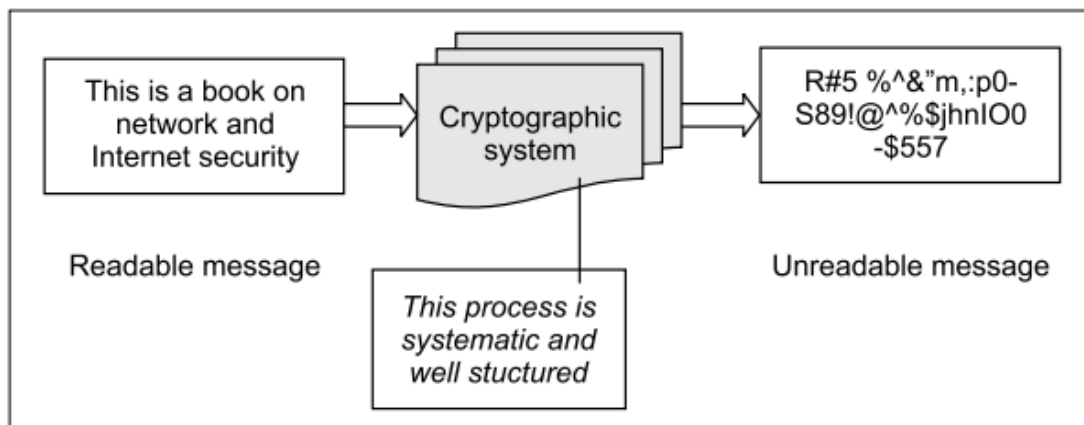
Cryptanalysis is the technique of decoding messages from a non-readable format back to a readable format without knowing how they were initially converted from readable format to non-readable format.
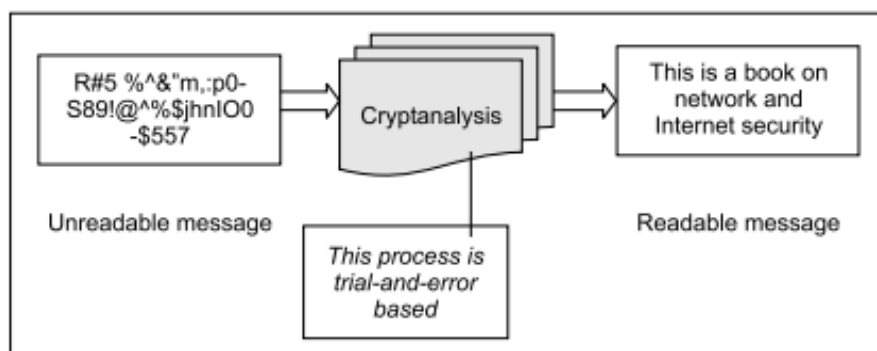
Cryptology is a combination of cryptography and cryptanalysis.

Cryptography + Cryptanalysis = Cryptology.

10) Define encryption and decryption with the help of a neat diagram.

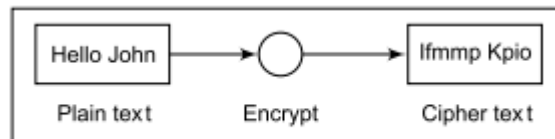**Ans i) ENCRYPTION: The process of encoding plaintext messages into cipher text messages.**

Fig. 2.43   Encryption

(ii) DECRYPTION: The reverse process of transforming cipher-text messages back to plain text messages.
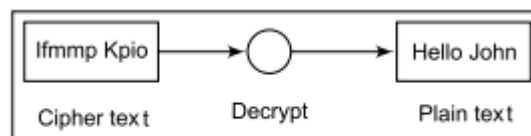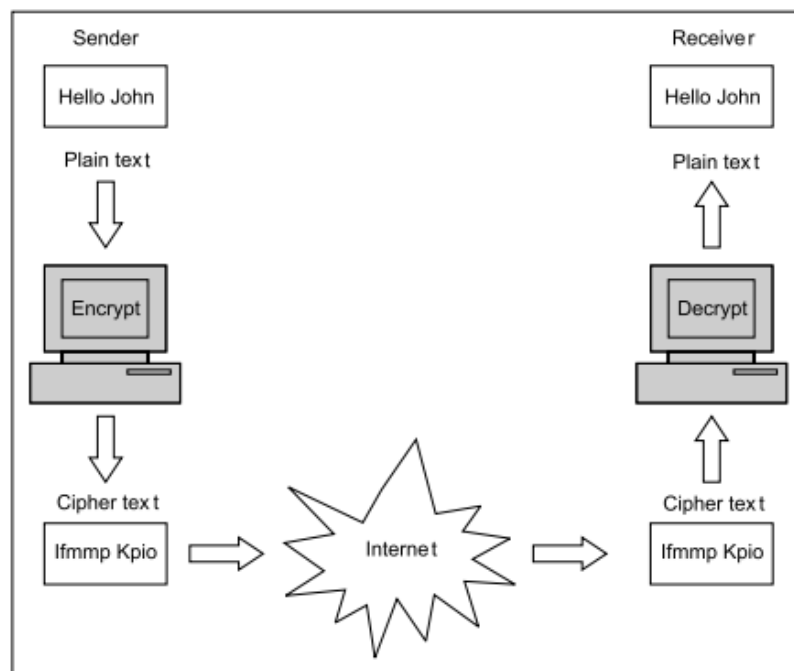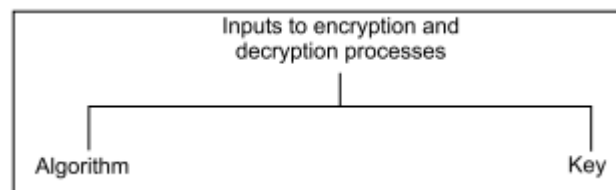

Fig. 2.44   Decryption

ɟ. 2.45   Encryption and decryption in the real world

11) What do you mean by a algorithm and a key?

Ans) Algorithm and Key are the aspects of Encryption and Decryption.

(i) How to open the lock is an algorithm (pieces of public knowledge).

(ii) However, the actual value of the key required for opening a specific lock is the key.
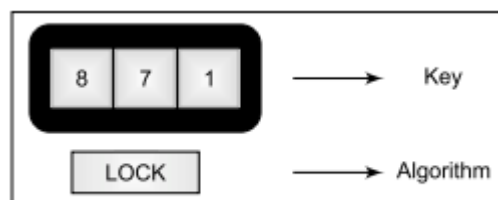
Fig. 2.47 Combination lock

12) Define symmetric and asymmetric key cryptography.

Ans)
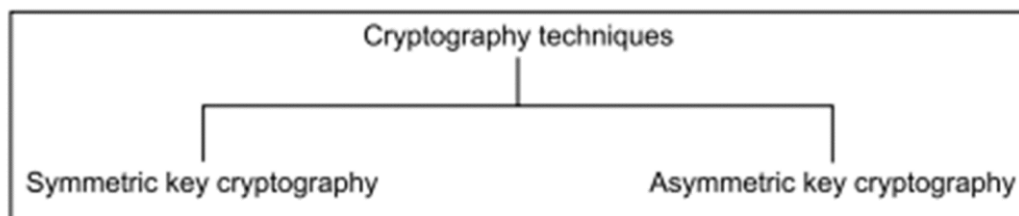
(i) Symmetric key cryptography: If the same key is used for encryption and decryption. It is also called as Private key.

(ii) Asymmetric key cryptography: If two different keys are used in a cryptographic mechanism, wherein one key is used for encryption, and another, different key is used for decryption. It is also called as Public key.

13) Differentiate between private and public key.

Ans)

Private Key	Public Key
(i) In this, the same key (secret key) and algorithm are used to encrypt and decrypt the message.	(i) In public-key cryptography, two keys are used, one key is used for encryption, and the other is used for decryption.
(ii) In private key cryptography, the key is kept a secret.	(ii) In public-key cryptography, one of the two keys is kept a secret.
(iii) It is Symmetrical because there is only one key that is called a secret key.	(iii) It is Asymmetrical because there are two types of keys: private and public keys.
(iv) In this cryptography, the sender and receiver need to share the same key.	(iv) In this cryptography, the sender and receiver do not need to share the same key.
(v) It is an efficient technology.	(v) It is an inefficient technology.

14) Explain block cryptography with the help of a neat diagram.

Ans) BLOCK ENCRYPTION (BLOCK CIPHER): The block-cipher technique involves encryption of one block of text at a time. Decryption also takes one block of encrypted text at a time.
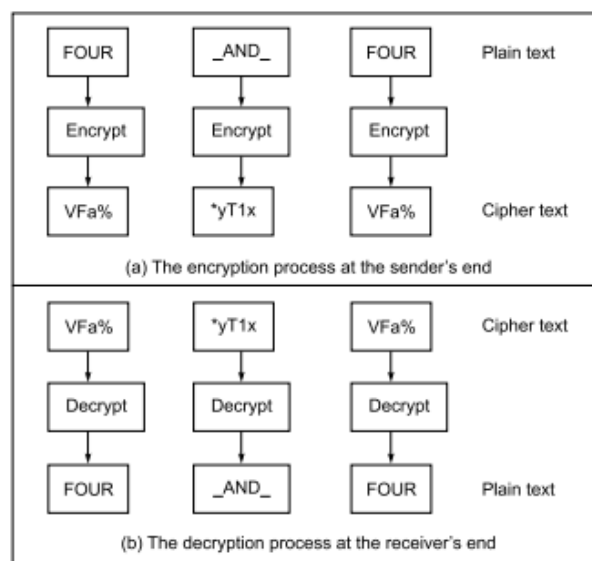
SOFTWARE ENGENEERING PART 2

1) What is SDLC?

- Ans) Software Development Life Cycle (SDLC) is the application of standard business practices to building software applications.
- The SDLC aims to produce a high-quality software that meets or exceeds customer expectations, reaches completion within times and cost estimates. It's typically divided into six to eight steps.
- It is also called as Software Development Process.
- SDLC is a framework defining tasks performed at each step in the software development process.
- ISO/IEC 12207 is an international standard for software life-cycle processes.

2) What are steps involved in SDLC?

Ans i) Planning: Project leaders evaluate the terms of the project.

- Planning can also include feedback from stakeholders.
- Planning should clearly define the scope and purpose of the application.
- It plots the course and provisions the team to effectively create the software.
- It also sets boundaries to help keep the project from expanding or shifting from its original purpose.

(ii) Define Requirements: Defining requirements is considered part of planning to determine what the application is supposed to do and its requirements. Requirements also include defining the resources needed to build the project.

(iii) Design and Prototyping: The Design phase models the way a software application will work. Some aspects of the design include: Architecture, User Interface, Platforms, Programming, Communications, Security.

(iv) Software Development: This is the actual writing of the program.

- A small project might be written by a single developer, while a large project might be broken up and worked by several teams.
- An Access Control or Source Code Management application is used in this phase. These systems help developers track changes to the code. They also help ensure compatibility between different team projects and to make sure target goals are being met.

(v) Testing: It's critical to test an application before making it available to users.

- Testing should ensure that each function works correctly.
- Different parts of the application should also be tested to work seamlessly together—performance test, to reduce any hangs or lags in processing.
- The testing phase helps reduce the number of bugs and glitches that users encounter. This leads to a higher user satisfaction and a better usage rate.

(vi) Deployment: In the deployment phase, the application is made available to users. Deployment can be complex and may take time and effort.

(vii) Operations and Maintenance: In this phase, users discover bugs that weren't found during testing.

These errors need to be resolved, which can spawn new development cycles.

3) Explain the phases of waterfall model with advantages and disadvantages

Ans) Phases of Waterfall Model:

(i) Feasibility Study

(ii) Requirements analysis and specification:

    a) Requirement gathering and analysis:

    b) Requirement specification:

(iii) Design

(iv) Coding and Unit testing

(v) Integration and System testing

(vi) Maintenance

Advantages:

Simplicity: The linear and sequential nature of the Waterfall model makes it easy to understand and use. Each phase has specific goals and deliverables.

Drawbacks:

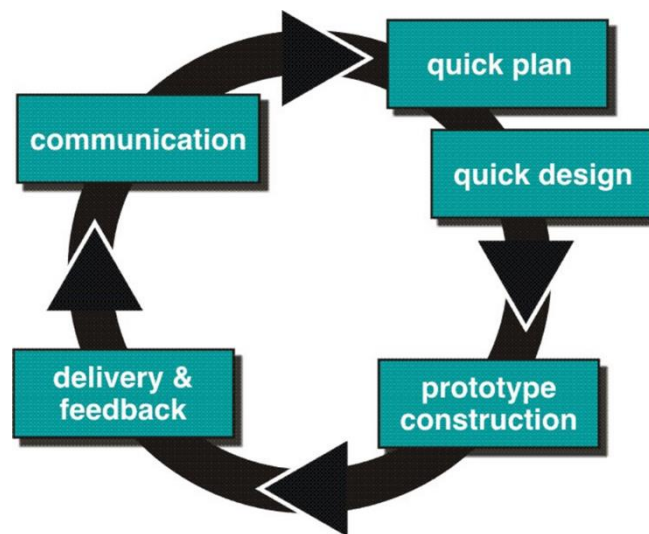**Difficult to incorporate change requests.**

Incremental delivery not supported.

Overlapping of phases not supported.

Risk handling not supported.

Limited customer interactions.

4) Write a short note on prototyping
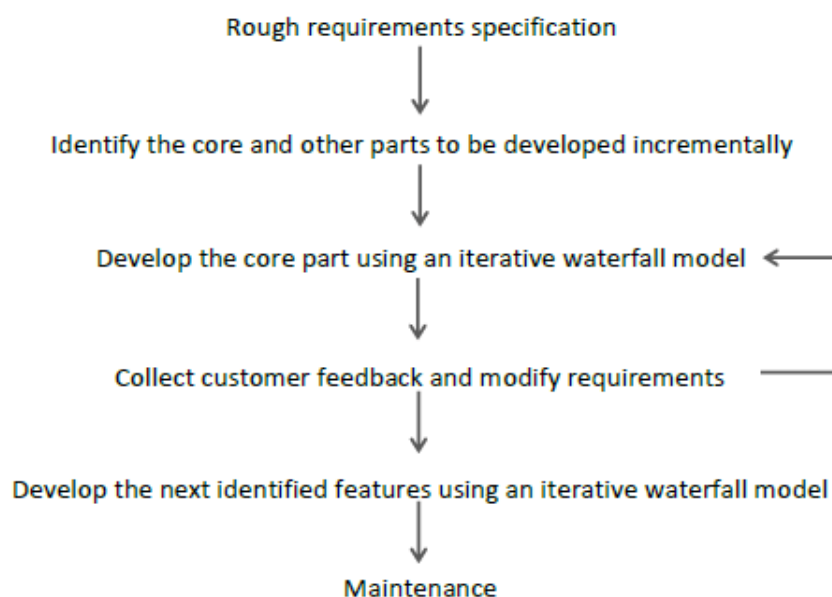
Ans

Best approach when:

- **Objectives defined by customer are general but does not have details like input, processing, or output requirement.**
- **Developer may be unsure of the efficiency of an algorithm, O.S., or the form that human machine interaction should take.**
- **It can be used as standalone process model.**
- **Model assist software engineer and customer to better understand what is to be built when requirement are fuzzy.**
- **Prototyping start with communication, between a customer and software engineer to define overall objective, identify requirements and make a boundary.**
- **Going ahead, quick planning and modeling (software layout visible to the customers/end-user) occurs.**
- **Quick design leads to prototype construction.**
- **Prototype is deployed and evaluated by the customer/user.**
- **Feedback from customer/end user will refine requirement and that is how iteration occurs during prototype to satisfy the needs of the customer.**
- **Prototype can be serve as "the first system".**

- **Both customers and developers like the prototyping paradigm. Both must be agree that the prototype is built to serve as a mechanism for defining requirement.**
- **Customer/End user gets a feel for the actual system**
- **Developer get to build something immediately.**

5) Explain evolutionary model with its advantages and disadvantages

Ans) Evolutionary model is also referred to as the successive versions model and sometimes as the incremental model.

- **In Evolutionary model, the software requirement is first broken down into several modules that can be incrementally constructed and delivered**
- **The development first develops the core modules of the system. The core modules are those that do not need services from the other modules. The initial product skeleton is refined into increasing levels of capability by adding new functionalities in successive versions.**
- **Each evolutionary model may be developed using an iterative waterfall model of development.**
- **Is normally useful for very large products, where it is easier to find modules for incremental implementation.**
- **Is used when the customer prefers to receive the product in increments so that he can start using the different features as and when they are developed rather than waiting all the time for the full product to be developed and delivered.**

Rough requirements specification

↓

Identify the core and other parts to be developed incrementally

↓

Develop the core part using an iterative waterfall model ←

↓

Collect customer feedback and modify requirements

↓

Develop the next identified features using an iterative waterfall model

↓

Maintenance

➤ **Advantages of Evolutionary Models:**

• **Normally useful for very large products.**

• **User gets a chance to experiment with a partially developed software much before the complete version of the system is released.**

• **Helps to accurately elicit user requirements during the delivery of different versions of the software.**

• **The core modules get tested thoroughly, thereby reducing the chances of errors in the core modules of the final products.**

• **Avoids the need to commit large resources in one go for development of the system.**

➤ **Disadvantage of Evolutionary Model:**

• **Difficult to divide the problem into several versions that would be acceptable to the customer and which can be incrementally implemented and delivered.**

6) Explain phases of agile model.

Ans)

(i) Requirements gathering.

(ii) Design the requirements.

(iii) Construction/ iteration.

(iv) Testing/ Quality assurance.

(v) Deployment.

(vi) Feedback.

7) Write shot note on agile model

Ans) Based on iterative development.

• **Agile methods break tasks into smaller iterations, or parts do not directly involve long term planning.**

• **The project scope and requirements are laid down at the beginning of the development process.**

• **Plans regarding the number of iterations, the duration and the scope of each iteration are clearly defined in advance.**

- Each iteration is considered as a short time "frame" in the Agile process model, which typically lasts from one to four weeks.
- The division of the entire project into smaller parts helps to minimize the project risk and to reduce the overall project delivery time requirements.
- Each iteration involves a team working through a full software development life cycle including planning, requirements analysis, design, coding, and testing before a working product is demonstrated to the client.

8) Explain the steps and goals for reverse engineering

Ans) Goals of reverse engineering

(i) Cope with Complexity.

(ii) Recover lost information.

(iii) Detect side effects.

(iv) Synthesise higher abstraction.

(v) Facilitate Reuse.

9) Describe objectives of re-engineering.

Ans) Objectives of Re-Engineering:

- To describe a cost-effective option for system evolution.
- To describe the activities involved in the software maintenance process.
- To distinguish between software and data re-engineering and to explain the problems of data re-engineering.

10) Explain the steps involved in re-engineering

Ans) Steps in Re-Engineering:

(i) Recording data flow.

(ii) Recording control flow.

(iii) Review extracted design.

(iv) Generate documentation.

SOFTWARE ENGINEERING
VERSION CONTROL

1) Define Version Control.

Ans) Version control (also known as revision control, source control, or source code management) is a class of systems responsible for managing changes to computer programs, documents, large web sites, or other collections of information.

2) Define Version Control System.

Ans) Version control systems are software tools that help software teams manage changes to source code over time.

3) What is the purpose of version control?

- Ans) Version control enables multiple people to simultaneously work on a single project.
- Each person edits his or her own copy of the files and chooses when to share those changes with the rest of the team. Thus, temporary or partial edits by one person do not interfere with another person's work.
- Version control also enables one person to use multiple computers to work on a project, so it is valuable even if you are working by yourself.
- Version control integrates work done simultaneously by different team members. In most cases, edits to different files or even the same file can be combined without losing any work. In rare cases, when two people make conflicting edits to the same line of a file, then the version control system requests human assistance in deciding what to do.
- Version control gives access to historical versions of your project. This is insurance against computer crashes or data lossage. If you make a mistake, you can roll back to a previous version. You can reproduce and understand a bug report on a past version of your software. You can also undo specific edits without losing all the work that was done in the meanwhile. For any part of a file, you can determine when, why, and by whom it was ever edited.

4) Explain the process and procedure of Version Control.

- Ans) Version control uses a repository (a database of changes) and a working copy where you do your work.
- Your working copy (sometimes called a checkout) is your personal copy of all the files in the project. You make arbitrary edits to this copy, without affecting your teammates. When you are happy with your edits, you commit your changes to a repository.
- A repository is a database of all the edits to, and/or historical versions (snapshots) of, your project.

5) Explain the concepts of versioning.

Ans i) Check-in/Check-out:

- To check in is to upload a file to the repository.
- To check out is to create a local working copy from the repository.
- A user may specify a specific revision or obtain the latest.
- The term 'checkout' can also be used as a noun to describe the working copy.
- When a file has been checked out from a shared file server, it cannot be edited by other users. Think of it like a hotel, when you check out, you no longer have access to its amenities.

(ii) Cloning: Cloning implies obtaining your own working copy of the project.

(iii) Commit: Commit is an operation which sends the latest changes made to the source code to the repository, making these changes part of the head revision of the repository.

(iv) Branching:

- Branching is the duplication of an object under version control such as a source code file or a directory tree.
- Each object can thereafter be modified separately and in parallel so that the objects become different.
- In this context the objects are called branches. The users of the version control system can branch any branch.

(v) Merging:

- Merging (also called integration) is a fundamental operation that reconciles multiple changes made to a version-controlled collection of files.
- Most often, it is necessary when a file is modified on two independent branches and subsequently merged.
- The result is a single collection of files that contains both sets of changes.

(vi) Synchronization: To synchronize your project means that the project files are brought into sync with the project as it looks on the remote server.

(vii) Conflicts:

- A conflict occurs when two different users make simultaneous, different changes to the same line of a file.
- In this case, the version control system cannot automatically decide which of the two edits to use (or a combination of them, or neither!). Manual intervention is required to resolve the conflict.

6) Write Short note on:

(i) Check-in & Check-out.

- Ans) To check in is to upload a file to the repository.
- To check out is to create a local working copy from the repository.
- A user may specify a specific revision or obtain the latest.
- The term 'checkout' can also be used as a noun to describe the working copy.
- When a file has been checked out from a shared file server, it cannot be edited by other users. Think of it like a hotel, when you check out, you no longer have access to its amenities.

(ii) Branching.

- Ans) Branching is the duplication of an object under version control such as a source code file or a directory tree.
- Each object can thereafter be modified separately and in parallel so that the objects become different.
- In this context the objects are called branches. The users of the version control system can branch any branch.

(iii) Merging.

- Ans) Merging (also called integration) is a fundamental operation that reconciles multiple changes made to a version-controlled collection of files.
- Most often, it is necessary when a file is modified on two independent branches and subsequently merged.
- The result is a single collection of files that contains both sets of changes.

7) Define Cloning.

Ans) Cloning implies obtaining your own working copy of the project.

8) Define commit in Version Control.

Ans) Commit is an operation which sends the latest changes made to the source code to the repository, making these changes part of the head revision of the repository.

9) What do you mean by Synchronization?

Ans) To synchronize your project means that the project files are brought into sync with the project as it looks on the remote server.

10) Short note on Conflicts in Versioning.

Ans) Conflicts:

- A conflict occurs when two different users make simultaneous, different changes to the same line of a file.
- In this case, the version control system cannot automatically decide which of the two edits to use (or a combination of them, or neither!). Manual intervention is required to resolve the conflict.