**CERTIFICATE**

Certified that **Sanya Rastogi (2100290140116), Akshi Vishnoi (2100290140018), Ayush Pathak (2100290140046), Ayush Jha (2100290140045)** have carried out the project work having “**Visualizing and Forecasting Stocks using Dash**” for Master of Computer Applications from Dr. A.P.J. Abdul Kalam Technical University (AKTU**)** (formerly UPTU), Technical University, Lucknow under my supervision. The project report embodies original work, and studies are carried out by the student himself and the contents of the project report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

**Date:**

**Sanya Rastogi (2100290140116)**

**Akshi Vishnoi (2100290140018)**

**Ayush Pathak (2100290140046)**

**Ayush Jha (2100290140045)**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

**Mr. Ankit Verma**

**Assistant Professor**

**Department of Computer Applications**

**KIET Group of Institutions, Ghaziabad**

**Signature of Internal Examiner Signature of External Examiner**

**Dr. Arun Tripathi**

**Head, Department of Computer Applications**

**KIET Group of Institutions, Ghaziabad**

**ABSTRACT**

In this project we attempt to implement machine learning approach to predict stock prices. Machine learning is effectively implemented in forecasting stock prices. The objective is to predict the stock prices in order to make more informed and accurate investment decisions. We propose a stock price prediction system that integrates mathematical functions, machine learning, and other external factors for the purpose of achieving better stock prediction accuracy and issuing profitable trades. There are two types of stocks. You may know of intraday trading by the commonly used term "day trading." Interday traders hold securities positions from at least one day to the next and often for several days to weeks or months. LSTMs are very powerful in sequence prediction problems because they’re able to store past information. This is important in our case because the previous price of a stock is crucial in predicting its future price. While predicting the actual price of a stock is an uphill climb, we can build a model that will predict whether the price will go up or down.

**ACKNOWLEDGEMENT**

Success in life is never attained single handedly. My deepest gratitude goes to my thesis supervisor, **Mr. Ankit Verma** for his guidance, help and encouragement throughout my research work. Their enlightening ideas, comments, and suggestions.

Words are not enough to express my gratitude to Dr Arun Tripathi, Professor and Head, Department of Computer Applications, for his insightful comments and administrative help at various occasions.

Fortunately, I have many understanding friends, who have helped me a lot on many critical conditions.

Finally, my sincere thanks go to my family members and all those who have directly and indirectly provided me moral support and other kind of help. Without their support, completion of this work would not have been possible in time. They keep my life filled with enjoyment and happiness.

**Sanya Rastogi (2100290140116)**

**Akshi Vishnoi (2100290140018)**

**Ayush Pathak (2100290140046)**

**Ayush Jha (2100290140045)**

**List of Chapters**

**Chapter 1 – Introduction**

1.1 Project Description

1.2 Project Scope

1.3 Hardware / Software used in Project

**Chapter 2 – Literature Survey**

**Chapter 3 – Feasibility Study**

3.1 Operational Feasibility

3.2 Technical Feasibility

3.3 Economical Feasibility

**Chapter 4 – Database Design**

4.1 Structure Chart

4.2 Use Case Diagram

4.3 Sequence Diagram

4.4 Activity Diagram

4.5 Flow Chart Diagram

4.6 Component Diagram

4.7 E-R Diagram

**Chapter 5 – Form Design**

5.1 Input / Output Form (Screenshot)

**Chapter 6 – Coding**

6.1 Module Wise Code

**Chapter 7 – Testing**

7.1 Test Case-1

7.2 Test Case-2

**Chapter 8 – Conclusion and Future Work**

**Bibliography and References**

**Chapter 1**

**Introduction**

The financial market is a dynamic and composite system where people can buy and sell currencies, stocks, equities and derivatives over virtual platforms supported by brokers. The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets. This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career. Stock markets are affected by many factors causing the uncertainty and high volatility in the market. Although humans can take orders and submit them to the market, automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human. However, to evaluate and control the performance of ATSs, the implementation of risk strategies and safety measures applied based on human judgements are required. Many factors are incorporated and considered when developing an ATS, for instance, trading strategy to be adopted, complex mathematical functions that reflect the state of a specific stock, machine learning algorithms that enable the prediction of the future stock value, and specific news related to the stock being analyzed. Time-series prediction is a common technique widely used in many real-world applications such as weather forecasting and financial market prediction. It uses the continuous data in a period to predict the result in the next time unit. Many timeseries prediction algorithms have shown their effectiveness in practice. The most common algorithms now are based on Recurrent Neural Networks (RNN), as well as its special type - Long-short Term Memory (LSTM) and Gated Recurrent Unit (GRU). Stock market is a typical area that presents time-series data and many researchers’ studies on it and proposed various models. In this project, LSTM model is used to predict the stock price.

**Stock Price Prediction**

Due to the high profit of the stock market, it is one of the most popular investments. People investigated for methods and tools that would increase their gains while minimizing the risk, as the level of trading and investing grew. Two stock exchanges namely- the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE), which are the most of the trading in Indian Stock Market takes place. Sensex and Nifty are the two prominent Indian Market Indexes. Since the prices in the stock market are dynamic, the stock market prediction is complicated.

From gradually the very past years some forecasting models are developed for this kind of purpose and they had been applied to money market prediction. Generally, this classification is done by:

1. Time series analysis
2. Fundamental analysis
3. Technical analysis

**Time Series Analysis**

The definition of forecasting can be like this the valuation of some upcoming result or results by analyzing the past data. It extents different areas like industry and business, economics and finance, environmental science. Forecasting problems can be classified as follows:

**Long term forecasting (estimation beyond 2 years)**

**Medium-term forecasting (estimation for 1 to 2 years)**

Short term forecasting (estimation for weeks or months, days, minutes, few seconds) the analysis [1] of time consist of several forecasting problems. The designation of a time series is a linear classification of observations for a selected variable. The variable of the stock price in our case. Which can weather multivariate or univariate? Only particular stock is included in the univariate data while more than one company for various instances of time is added in multivariate. For investigating trends, patterns and cycle or periods the analysis of time series advantages in the present data. In spending money wisely an early data of the bullish or bearish in the case of the stock market. Also, for categorizing the best-performing companies the analysis of patterns plays its role for a specific period. This makes forecasting as well as time series analysis an important research area.

**Fundamental Analysis**

Fundamental Analysts are concerned with the business that reasons the stock itself. They assess a company's historical performance as well as the reliability of its accounts. Different performance shares are created that aid the fundamental forecaster with calculating the validity of a stock, such as the P/E ratio. Warren Buffett is probably the foremost renowned of all Fundamental Analysts.

What fundamental analysis within the stock market is making an attempt to reach, is organizing the true value of a stock, that then will be matched with the worth it is being listed on stock markets and so finding out whether or not the stock on the market is undervalued or not. Find out the correct value will be completed by numerous strategies with primarily a similar principle. The principle is that an organization is price all of its future profits. Those future profits has to be discounted to their current value. This principle goes on the theory that a business is all about profits and nothing else. Differing to technical analysis, the fundamental analysis is assumed as further as a long approach.

Fundamental analysis is created on conviction that hominoid society desires capital to make progress and if the company works well, than it should be rewarded with an additional capital and outcome in a surge in stock price. Fundamental analysis is usually used by the fund managers as it is the maximum sensible, objective and prepared from openly existing data like financial statement analysis.

One more meaning of fundamental analysis is on the far side bottom-up business analysis, it discusses the top-down analysis since initial analysing the world economy, followed by country analysis and also sector analysis, and last the company level analysis.

**Technical Analysis**

Chartists or the technical analysts are not involved with any other of the fundamentals of the company. The long run price of a stock based generally exclusively on the trends of the past value (a form of time series analysis) that is set by them. The head and shoulders or cup and saucer are various numerous patterns that are employed. Also the techniques, patterns are used just like the oscillators, exponential moving average (EMA), support and momentum and volume indicators. Candlestick patterns, believed to have been initial developed by Japanese rice merchants, are nowadays widely used by technical analysts. For the short-term approaches, the technical analysis is used compare to long-run ones. So, in commodities and forex markets it is more predominant wherever traders target short-term price movements. There are basic rules are used in this analysis, first all significant about a company is already priced into the stock, another being that the value changes in trends and finally that history (of prices) tends to repeat itself that is especially due to the market science.

**Objectives**

A stock market prediction is described as an action of attempting to classify the future value of the company stock or other financial investment traded on the stock exchange. The forthcoming price of a stock of the successful estimation is called the Yield significant profit. This helps you to invest wisely for making good profits.

Stock market prediction aims to determine the future movement of the stock value of a financial exchange. The accurate prediction of share price movement will lead to more profit investors can make.

Stock Price Prediction using machine learning helps you discover the future value of company stock and other financial assets traded on an exchange. The entire idea of predicting stock prices is to gain significant profits. Predicting how the stock market will perform is a hard task to do.

* 1. **Project Description**

We will be creating a single-page web application using Dash (a python framework) and some machine learning models which will show company information (logo, registered name and description) and stock plots based on the stock code given by the user. Also, the ML model will enable the user to get predicted stock prices for the date inputted by the user. Stock investments provide one of the highest returns in the market. Even though they are volatile in nature, one can visualize share prices and other statistical factors which helps the keen investors carefully decide on which company they want to spend their earnings on. Developing this simple project idea using the Dash library (of Python), we can make dynamic plots of the financial data of a specific company by using the tabular data provided by yfinance python library. On top of it, we can use a machine learning algorithm to predict the upcoming stock prices. This project is a good start for beginners in python/data science and a good refresher for professionals who have dabbled in python / ML before. This web application can be applied to any company (whose stock code is available) of one's choosing, so feel free to explore!

* 1. **Project Scope**

This project aims are to predict public listed companies stock price listed in open market all over the world. Because the prediction tool is using data mining method, historical data of the price is very important. The project will use existing historical data that distributed freely in internet and provided by trustable provider to make sure the desired result is fully trusted and reliability. This project could be adapted into large capital market around the world. However, certain of limitation is necessary to make sure that the scope will not be large and become time consuming and decided to only focus on 10 stocks of the highest market capitalization of companies publicly listed in Bursa Malaysia. Furthermore, the implementation of several methods needs to be tested into different environment where the excellent desired output will be chosen as the object of this project. The historical data is downloaded from search engine and historical data provider ‘Yahoo Financial’. Several developer and time series experts have used yahoo finance historical data in order to calculate the price prediction with different approaches. Other than that, the data would also be tested and analyze to ensure the accuracy and reliability of data. The historical data that will be used in this project would be naturally come from ‘closed price’ without investigate volume of transaction and other variable in the historical data. This is because the historical closed price is already adequate to find the pattern or plot and to predict the next pattern of future stock price. The scope of the study would further assist in creating the accuracy and reliability that the system proposes. The historical data downloaded from not native provider such as ‘Yahoo Financial’ will always be investigate the real data provided by the regulators in particular country. The make sure accuracy of data, this project will also try to find the best model approach that can be used to provide an output that become solution for large communities. **1.3 Hardware & Software Requirement:**

**Hardware Interfaces**

Minimum Hardware requirement

* Processor: Pentium IV or Higher
* RAM: 4 GB or Higher
* Operating System: windows 7 and above or Linux based OS or MAC OS
* System Type: 32-bit operating system
* Hard disk: 512 GB
* CPU: 2 GHz or faster
* Architecture: 32-bit or 64-bit

**Software Interfaces**

* Minimum Software requirement
* Anaconda Navigator
* Visual Studio Code
* Python 3.5 in Google Colab

All these types of software automatic configure inside operating system after installation it is having Jupyter Notebook, Apache and operating system base configuration file, it doesn’t need to configure manually.

**Chapter 2. Literature Survey**

#### Stock Price Forecasting Using Data From Yahoo Finance and Analysing Seasonal and Nonseasonal Trend:

#### Publication Year: 2018

**Author:** Jai Jagwani, Hardik Sachdeva, Manav Gupta, Alka Singhal

**Journal Name:** 2018 IEEE

**Summary:** To identify the [2] relationship between different existing time series algorithms namely ARIMA and Holt Winter and the stock prices is the main objective of the proposed work, for the investments a good risk-free range of stock prices are analyzed and therefore better accuracy of the model can be seen. To find distinguished results for shares in the stock market, the combination of two different time series analysis models is opted by producing a range of prices to the consumer of the stocks. Not complex in nature and estimation of values which are purely based on the past stock prices for non-seasonal or seasonal is the main advantage of these models. In this experiment, some limitations are, the work that never takes into consideration and other circumstances like news about any new market strategy or media release relevant to any company which may get affected by the prices of stocks.

#### Stock Market Prediction Using Machine Learning:

#### Publication Year: 2018

**Author:** Ishita Parmar, Ridam Arora, Lokesh Chouhan, Navanshu Agarwal, Shikhin Gupta, Sheirsh Saxena, Himanshu Dhiman

**Journal Name:** 2018 IEEE

**Summary:** In this paper studies, the use [3] of Regression and LSTM based Machine learning to forecast stock prices. Factors measured are open, close, low, high and volume. This paper was an attempt to determine the future prices of the stocks of a company with improved accuracy and reliability using machine learning techniques. LSTM algorithm resulted in a positive outcome with more accuracy in predicting stock prices.

#### Multi-Category Events Driven Stock Price Trends Prediction: Publication Year: 2018

**Author:** Youxun Lei, Kaiyue Zhou, Yuchen Liu

**Journal Name:** 2018 IEEE

**Summary:** In this paper, [4] multi-category news events are used as features to develop stock price trend prediction, model. The multi-category events are based on already defined feature word dictionary. And we have employed both neural networks and SVM models to analyse the relationship between stock price movements and specific multi-category news. Experimental results showed that the predefined multi-category news events are more improved than the baseline bag-of-words feature to predict stock price trend. As compared to long term prediction, short term prediction is better based on this study.

#### Share Price Prediction using Machine Learning Technique:

#### Publication Year: 2018

**Author:** Jeevan B, Naresh E, Vijaya kumar B P, Prashanth Kambli

**Journal Name:** 2018 IEEE

**Summary:** This paper is mostly [5] based on the approach of predicting the share price using Long Short Term Memory (LSTM) and Recurrent Neural Networks (RNN) to forecast the stock value on NSE data using various factors such as current market price, price to earnings ratio, base value and other anonymous events. The efficiency of the model is analyzed by comparing the true data and the predicted data using an RNN graph. Machine learning to predict stock price as see the model is able to predict the stock price very close to the actual price where this model captures the detailed feature and uses different strategies to make a prediction. The model train for all the NSE data from the internet and recognize the input and group them and provide input according to the user configuration this RNN based architecture proved very efficient in forecasting the stock price by changing the configuration accordingly which also use back propagation mechanism while gathering and grouping data to avoid mixing of data.

#### Stock Market Prediction Using Machine Learning Techniques: Publication Year: 2016

**Author:** Mehak Usmani, Syed Hasan Adil, Kamran Raza, Syed Saad Azhar Ali

**Journal Name:** 2016 IEEE

**Summary:** The prominent aim of this study is to [6] forecast the market performance of the Karachi Stock Exchange (KSE) on day closing using machine learning algorithms. A variety of attributes as an input and forecasts market as Positive & Negative is predicted by using the predictions model. The features employed in the model are contains Oil rates, Gold & Silver rates, Interest rate, Foreign Exchange (FEX) rate, NEWS and social media feed. The machine learning algorithms including Single Layer Perceptron (SLP), Multi-Layer Perceptron (MLP), Radial Basis Function (RBF) and Support Vector Machine (SVM) are compared. The algorithm MLP that is multi-layer perceptron performed best as compared to different methods. The foremost helpful feature in predicting the market was the oil rate attribute. The end results of this research confirm that machine learning techniques have the ability to predict the stock market performance. The Multi-Layer Perceptron algorithm of machine learning predicted 70% correct market performance.

#### Forecasting stock price in two ways based on LSTM neural network: Publication Year: 2019

**Author:** Jingyi Du, Qingli Liu, Kang Chen, Jiacheng Wang

**Journal Name:** 2019 IEEE

**Summary:** The [7] LSTM neural network is used to predict Apple stocks by consuming single feature input variables and multi-feature input variables to verify the forecast effect of the model on stock time series. The experimental results show that the model has a high accuracy of 0.033 for the multivariate input and is accurate, that is in line with the actual demand. For the univariate feature input, the predicted squared absolute error is 0.155, which is inferior to the multi-feature variable input.

#### Share Price Trend Prediction Using CRNN with LSTM Structure: Publication Year: 2018

**Author:** Shao-En Gao , Bo-Sheng Lin ,Chuin-Mu Wang

**Journal Name:** 2018 IEEE

**Summary:** The [8] entire financial market majorly runs by the stock market and one of the most attractive research issues is predicting stock price volatility. The information of historical stocks for assuming the future stock price as well deep learning method is applied to find approximate trend value of stock prices which are mentioned in this paper. This paper not only stores the data of historical stock with the time scale but also estimates prices of the future stock by a designed neural network, this is due to the fact that the trend of stocks is usually connected to the previous information of stock price. In this paper, the design of the neural network proposed then with the memory performance the convolutional recurrent neural network (CRNN) and for improving the long-term dependency of traditional RNN the Long Short-term memory (LSTM) are the major components. Also to enhance the accuracy as well as stability of prediction of the RNN LSTM architecture is put. This paper accumulates a total of ten stock historic data to test and accomplish an average error rate of 3.449 RMSE. [3]

#### Applying Long Short Term Memory Neural Networks for Predicting Stock Closing Price:

#### Publication Year: 2017

**Author:** Tingwei Gao, Yueting Chai, Yi Liu

**Journal Name:** 2017 IEEE

**Summary:** To [9] assess the scheme that merges RNNs with informative input variables which can give an improved and effective method to forecast the next-day market is the main objective of this paper. The stock prediction model analyses using long-short memory (LSTM) and stock basic trading data. On Standard & Poor’s (S&P500) and NASDAQ, the case study relies. The stock closing price is more precisely predicted using their forecasting system for the next day, which outperforms the comparison models. This is the main discovery of the case study. Five various models namely – moving average (MA), exponential moving average (EMA), support vector machine (SVM) and LSTM are tested by them to demonstrate the utility of the system. The closing value of the next day is the predicting target.

#### Developing a Prediction Model for Stock Analysis:

#### Publication Year: 2017

**Author:** R. Yamini Nivetha, Dr. C. Dhaya

**Journal Name:** 2017 IEEE

**Summary:** A [10] relative study of the three algorithms namely - Multiple Linear Regression (MLR), Support Vector Machine (SVM) and Artificial Neural Network (ANN) is the main aim of this study. To predict the coming day market price, the prediction will be determined by monthly prediction and daily prediction. Sentiment analysis with the best prediction algorithm forecast the stock price. The less-developed algorithm is the Multiple Linear Regression algorithm which calculates the correlation between volume and the stock price. The result of the study shows that deep learning algorithms are more developed than MLR algorithms and SVM algorithm.

#### Stock Price Prediction Based on Information Entropy and Artificial Neural Network: Publication Year: 2019

**Author:** Zang Yeze, Wang Yiying

**Journal Name:** 2019 IEEE

**Summary:** One of the most important components of the financial system is the stock market.

[11] For supporting the activity and evolvement, money is directed by the investors of the associated frim. Along with information theory and Artificial Neural Network (ANN) the combination of machine learning framework is formed. Information entropy for non-linear causality and stock relevance also to facilitate ANN time series modelling are creatively used by this method. The feasibility of this machine learning framework is analysed with Amazon, Apple, Google and Facebook prices. A time series analysis method based on information theory as well as LSTM to model the stock price dynamics are outlined in this paper. The transfer entropy between relevant variables to help LSTM time series prediction is merged in this modelling infrastructure, thus the accuracy of the assumption outcome is broadly granted. Modelled and real stock price is highly correlated while differ slightly in terms of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) which are investigated by the outcomes.

#### Summary of Literature Survey:

Here, I have reviewed various approaches for Stock price prediction. All approaches have their own advantages and disadvantages. CNN & LSTM is a most popular algorithm to prediction the stock price but there are some challenges in this method like use to need a lot of training data, High computational cost, without GPU data quite slow to train, depend on any previous information for prediction. A hybrid approach can be used to overcome these issues. While machine learning is able to provide highly accurate prediction result using standards tools and also outperforms all standard prediction methods.

**Chapter 3**

**Feasibility Study**

This project will be developed on computer, so first check whether the technology is technically available or not. Now a day’s computer interaction with any job becomes common for any kind of job or work.

And because of increasing usage of Computer, Computer is also available with a variety of hardware. Vendors can fulfill any type of hardware requirement. The whole project is developed by some special tools or by using languages and databases, which are also available in a variety.

Preliminary investigation of a system examines the feasibility of a system that is useful to an organization. It is the first phase of system development.

The main objective of this phase is to identify the current deficiencies in the user’s environment and to determine which existing problem are going to be solve in proposed system and also which new function needs to be added in proposed system.

An important outcome of such preliminary investigation is to determine whether the system that will meet all needed requirements.

Thus, three tests are carried out on the system namely operation, technical and economical.

Any project is beneficial if and only satisfies the organization requirement. For any new system setup, it only meets to be communicated and work the other supporting system.

The new system meets all existing operations since it provides right information at a right time to the right user. A Leigh man can easily operate with the system.



**3.1 Operational Feasibility**

Operational feasibility is the ability to utilize, support and perform the necessary tasks of a system or program. It includes everyone who creates, operates or uses the system. To be operationally feasible, the system must fulfill a need required by the business.

Programs that reduce costs without reducing the quality of a product are an example of operational feasibility. Studies are performed to make sure that programs can be initiated in the current production facility, without the need for more equipment or staff. If there is a need for more room, machinery or personnel, the system needs to make an improvement in the way the product is perceived by consumers. This allows the added expense to manufacture the product because of the anticipated increase in sales and revenue. Economic feasibility is one portion of operational feasibility. Everyone involved in the manufacturing and use of the system, product or program needs to understand it and be able to use it in the manner it was designed to be used. Feasibility studies are done not only in the private business sector, but by all governmental organizations. If the changes made to a product or program are not understood by the end users, it is not considered feasible.

**3.2 Technical Feasibility**

Technical Feasibility examines whether the technology needed is available and if it is available then it feasible to carry out all project activities.

The technical needs of a system include:

* The facility to produce outputs in each time.
* Ability to process large number of transactions at a particular speed.
* Giving response to users under certain conditions.

The technology needed for our system is mainly:

* Latest version of browsers.
* Any operating system.

These technologies are available which helps to carry out the system efficiently.

**3.3 Economical Feasibility**

Economic feasibility of a system examines whether the finance is available for implementing the new system and whether the money spent is recoverable the satisfaction.

The cost involves is in designing and developing a good investment for the organization.

Thus, hardware requirements used for proposed system are very standard. Moreover, by making use of proposed system to carry out the work speedily will increase and saves the valuable time of an organization.

In the proposed system the finance is highly required for the installation of the software’s which can also be recovered by implementing a better system.

**Chapter 4**

**Database Design**

Database design can be generally defined as a collection of tasks or processes that enhance the designing, development, implementation, and maintenance of enterprise data management system. Designing a proper database reduces the maintenance cost thereby improving data consistency and the cost-effective measures are greatly influenced in terms of disk storage space. Therefore, there has to be a brilliant concept of designing a database. The designer should follow the constraints and decide how the elements correlate and what kind of data must be stored.

The main objectives behind database designing are to produce physical and logical design models of the proposed database system. To elaborate this, the logical model is primarily concentrated on the requirements of data and the considerations must be made in terms of monolithic considerations and hence the stored physical data must be stored independent of the physical conditions. On the other hand, the physical database design model includes a translation of the logical design model of the database by keep control of physical media using hardware resources and software systems such as Database Management System (DBMS).

**Why is Database Design important?**

The important consideration that can be taken into account while emphasizing the importance of database design can be explained in terms of the following points given below.

1. Database designs provide the blueprints of how the data is going to be stored in a system. A proper design of a database highly affects the overall performance of any application.
2. The designing principles defined for a database give a clear idea of the behavior of any application and how the requests are processed.
3. Another instance to emphasize the database design is that a proper database design meets all the requirements of users.
4. Lastly, the processing time of an application is greatly reduced if the constraints of designing a highly efficient database are properly implemented.

**Database Design Process**

The process of designing a database carries various conceptual approaches that are needed to be kept in mind. An ideal and well-structured database design must be able to:

1. Save disk space by eliminating redundant data.
2. Maintains data integrity and accuracy.
3. Provides data access in useful ways.
4. Comparing Logical and Physical data models.

**Summary**

Database design is a method of identifying the gaps and opportunities of designing a proper utilization method. It is the main component of a system that gives a blueprint of the data and its behavior inside the system. A proper database design is always kept on priority due to the user requirements being kept excessively high and following up with the constraint practices of designing a database might only stand as a chance to gain the requested efficiency. Moreover, we also learned separately about the different design models that portray the ideal database design along with the limitless discussion on their properties and how to make use of them. Furthermore, we learned how the life-cycle of a database decides the design of the database and how to put the concept of design into the life-cycle methods so that efficient and highly sophisticated databases can be designed based on the user requirements.

* 1. **Structure Chart**

Structure Chart represent hierarchical structure of modules. It breaks down the entire system into lowest functional modules, describe functions and sub-functions of each module of a system to a greater detail. Structure Chart partitions the system into black boxes (functionality of the system is known to the users but inner details are unknown). Inputs are given to the black boxes and appropriate outputs are generated.

Modules at top level called modules at low level. Components are read from top to bottom and left to right. When a module calls another, it views the called module as black box, passing required parameters and receiving results.

Symbols used in construction of structured chart -:

1. **Module**

It represents the process or task of the system. It is of three types.

* Control Module

A control module branches to more than one sub module.

* Sub Module

Sub Module is a module which is the part (Child) of another module.

* Library Module

Library Module are reusable and invokable from any module.

1. **Conditional Call**

It represents that control module can select any of the sub module on the basis of some condition.

1. **Loop (Repetitive call of module)**

It represents the repetitive execution of module by the sub module.

A curved arrow represents loop in the module.

All the sub modules cover by the loop repeat execution of module.

1. **Data Flow**

It represents the flow of data between the modules. It is represented by directed arrow with empty circle at the end.

1. **Control Flow**

It represents the flow of control between the modules. It is represented by directed arrow with filled circle at the end.

Structure Chart -:

Chart, diagram

Description automatically generated

* 1. **Use Case Diagram**

A UML use case diagram is the primary form of system/software requirements for a new software program underdeveloped. Use cases specify the expected behavior (what), and not the exact method of making it happen (how). Use cases once specified can be denoted both textual and visual representation (i.e. use case diagram). A key concept of use case modeling is that it helps us design a system from the end user's perspective. It is an effective technique for communicating system behavior in the user's terms by specifying all externally visible system behavior.

A use case diagram is usually simple. It does not show the detail of the use cases:

It only summarizes some of the relationships between use cases, actors, and systems.

It does not show the order in which steps are performed to achieve the goals of each use case.

As said, a use case diagram should be simple and contains only a few shapes. If yours contain more than 20 use cases, you are probably misusing use case diagram.

**Origin of Use Case**

These days use case modeling is often associated with UML, although it has been introduced before UML existed. Its brief history is as follow:

In 1986, Ivar Jacobson first formulated textual and visual modeling techniques for specifying use cases.

In 1992 his co-authored book Object-Oriented Software Engineering - A Use Case Driven Approach helped to popularize the technique for capturing functional requirements, especially in software development.

**Purpose of Use Case Diagram**

Use case diagrams are typically developed in the early stage of development and people often apply use case modeling for the following purposes:

1. Specify the context of a system
2. Capture the requirements of a system
3. Validate a systems architecture
4. Drive implementation and generate test cases
5. Developed by analysts together with domain experts

**Notation Description**

1. **Actor**

* Someone interacts with use case (system function).
* Named by noun.
* Actor plays a role in the business
* Similar to the concept of user, but a user can play different roles
* For example:
* A prof. can be instructor and also researcher
* Plays 2 roles with two systems
* Actor triggers use case(s).
* Actor has a responsibility toward the system (inputs), and Actor has expectations from the system (outputs).

1. **Use Case**

* System function (process - automated or manual)
* Named by verb + Noun (or Noun Phrase).
* i.e. Do something
* Each Actor must be linked to a use case, while some use cases may not be linked to actors.

1. **Communication Link**

* The participation of an actor in a use case is shown by connecting an actor to a use case by a solid link.
* Actors may be connected to use cases by associations, indicating that the actor and the use case communicate with one another using messages.

1. **Boundary of the System**

* The system boundary is potentially the entire system as defined in the requirements document.
* For large and complex systems, each module may be the system boundary.
* For example, for an ERP system for an organization, each of the modules such as personnel, payroll, accounting, etc.
* Can form a system boundary for use cases specific to each of these business functions.
* The entire system can span all of these modules depicting the overall system boundary

**Structuring Use Case Diagram with Relationships**

Use cases share different kinds of relationships. Defining the relationship between two use cases is the decision of the software analysts of the use case diagram. A relationship between two use cases is basically modeling the dependency between the two use cases. The reuse of an existing use case by using different types of relationships reduces the overall effort required in developing a system. Use case relationships are listed as the following:

**Use Case Relationship**

1. **Extends**

* Indicates that an **"Invalid Password"** use case may include (subject to specified in the extension) the behavior specified by base use case **"Login Account"**.
* Depict with a directed arrow having a dotted line. The tip of arrowhead points to the base use case and the child use case is connected at the base of the arrow.
* The stereotype "<<extends>>" identifies as an extend relationship

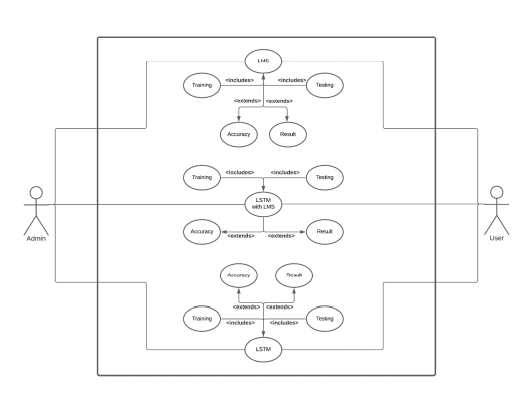
1. **Include**

* When a use case is depicted as using the functionality of another use case, the relationship between the use case is named as include or uses relationship.
* A use case includes the functionality described in another use case as a part of its business process flow.
* A uses relationship from base use case to child use case indicates that an instance of the base use case will include the behaviour as specified in the child use case.
* An include relationship is depicted with a directed arrow having a dotted line. The tip of arrowhead points to the child use case and the parent use case connected at the base of the arrow.
* The stereotype "<<include>>" identifies the relationship as an include relationship.

1. **Generalization**

* A generalization relationship is a parent-child relationship between use cases.
* The child use case is an enhancement of the parent use case.
* Generalization is shown as a directed arrow with a triangle arrowhead.
* The child use case is connected at the base of the arrow. The tip of the arrow is connected to the parent use case.

Use Case Diagram -:



* 1. **Sequence Diagram**

A sequence diagram or system sequence diagram (SSD) shows process interactions arranged in time sequence in the field of software engineering. It depicts the processes involved and the sequence of messages exchanged between the processes needed to carry out the functionality. Sequence diagrams are typically associated with use case realizations in the 4+1 architectural view model of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios.

For a particular scenario of a use case, the diagrams show the events that external actors generate, their order, and possible inter-system events. All systems are treated as a black box; the diagram places emphasis on events that cross the system boundary from actors to systems. A system sequence diagram should be done for the main success scenario of the use case, and frequent or complex alternative scenarios.

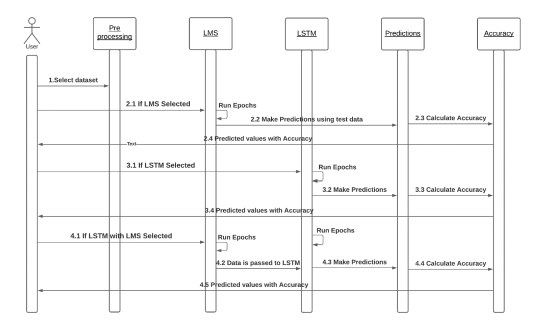
**Key elements of sequence diagram**

A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

A system sequence diagram should specify and show the following:

* 1. External actors
  2. Messages (methods) invoked by these actors
  3. Return values (if any) associated with previous messages
  4. Indication of any loops or iteration area

Sequence Diagram -:



* 1. **Activity Diagram**

Activity diagram is another important diagram in UML to describe the dynamic aspects of the system.

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system.

The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc.

**Purpose of Activity Diagrams**

The basic purposes of activity diagrams is similar to other four diagrams. It captures the dynamic behavior of the system. Other four diagrams are used to show the message flow from one object to another but activity diagram is used to show message flow from one activity to another.

Activity is a particular operation of the system. Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part.

It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single.

The purpose of an activity diagram can be described as –:

* 1. Draw the activity flow of a system.
  2. Describe the sequence from one activity to another.
  3. Describe the parallel, branched and concurrent flow of the system.

Activity Diagram -:

Diagram

Description automatically generated

* 1. **Flow chart Diagram**

A flowchart is a diagram that depicts a process, system or computer algorithm. They are widely used in multiple fields to document, study, plan, improve and communicate often complex processes in clear, easy-to-understand diagrams. Flowcharts, sometimes spelled as flow charts, use rectangles, ovals, diamonds and potentially numerous other shapes to define the type of step, along with connecting arrows to define flow and sequence. They can range from simple, hand-drawn charts to comprehensive computer-drawn diagrams depicting multiple steps and routes. If we consider all the various forms of flowcharts, they are one of the most common diagrams on the planet, used by both technical and non-technical people in numerous fields. Flowcharts are sometimes called by more specialized names such as Process Flowchart, Process Map, Functional Flowchart, Business Process Mapping, Business Process Modeling and Notation (BPMN), or Process Flow Diagram (PFD). They are related to other popular diagrams, such as Data Flow Diagrams (DFDs) and Unified Modeling Language (UML) Activity Diagrams.

**Flowchart symbols**

Here are some of the common flowchart symbols. For a more comprehensive list, see our full flowchart symbols page.

1. Terminal/Terminator
2. Process
3. Decision Box
4. Input / Output Box
5. Data Stored
6. Flow Lines or Flow Arrow
7. Connectors

Flow chart -:

Diagram

Description automatically generated

* 1. **Component Diagram**

A component diagram is used to break down a large object-oriented system into the smaller components, so as to make them more manageable. It models the physical view of a system such as executable, files, libraries, etc. that resides within the node.

It visualizes the relationships as well as the organization between the components present in the system. It helps in forming an executable system. A component is a single unit of the system, which is replaceable and executable. The implementation details of a component are hidden, and it necessitates an interface to execute a function. It is like a black box whose behavior is explained by the provided and required interfaces.

**Notation of a Component Diagram**

1. A Component
2. A Node

**Purpose of a Component Diagram**

Since it is a special kind of a UML diagram, it holds distinct purposes. It describes all the individual components that are used to make the functionalities, but not the functionalities of the system. It visualizes the physical components inside the system. The components can be a library, packages, files, etc.

The component diagram also describes the static view of a system, which includes the organization of components at a particular instant. The collection of component diagrams represents a whole system.

The main purpose of the component diagram are enlisted below:

1. It envisions each component of a system.
2. It constructs the executable by incorporating forward and reverse engineering.
3. It depicts the relationships and organization of components.

Component Diagram -:

Diagram

Description automatically generated

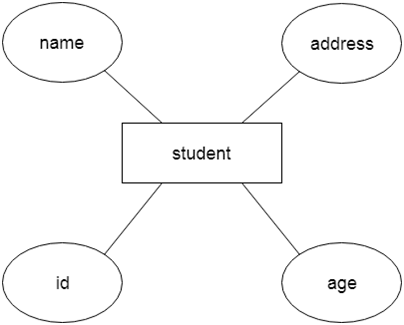
**4.7 E-R Diagram**

ER model stands for an Entity-Relationship model. It is a high-level data model. This model is used to define the data elements and relationship for a specified system.

It develops a conceptual design for the database. It also develops a very simple and easy to design view of data.

In ER modeling, the database structure is portrayed as a diagram called an entity-relationship diagram.

**For example-:** Suppose we design a school database. In this database, the student will be an entity with attributes like address, name, id, age, etc. The address can be another entity with attributes like city, street name, pin code, etc. and there will be a relationship between them.



## **Component of ER Diagram**



E-R Diagram -:

