**Detection of Cyber Attacks Using Artificial Intelligence**

**OJECTIVE:**

The primary goal of this project is to determine the cyber-attack whether there will be attack or not and to know this we have used the Support Vector , Decision Tree, Random forest, Extra Tree Classifier and ad boost and neural network classifier classification techniques.

**ABSTRACT:**

Cyber-physical systems(cps) have made significant progress in many dynamic applications due to the integration between physical processes, computational resources, and communication capabilities. However, cyber-attacks are a major threat to these systems. Unlike faults that occurs by accidents cyber-physical systems, cyber-attacks occur intelligently and stealthy. Some of these attacks which are called deception attacks, inject false data from sensors or controllers, and also by compromising with some cyber components, corrupt data, or enter misinformation into the system. If the system is unaware of the existence of these attacks, it won’t be able to detect them, and performance may be disrupted or disabled altogether. Therefore, it is necessary to adapt algorithms to identify these types of attacks in these systems. It should be noted that the data generated in these systems is produced in very large number, with so much variety, and high speed, so it is important to use machine learning algorithms to facilitate the analysis and evaluation of data and to identify hidden patterns. In this research, the CPS is modeled as a network of agents that move in union with each other, and one agent is considered as a leader, and the other agents are commanded by the leader. The proposed method in this study is to use the structure of deep neural networks for the detection phase, which should inform the system of the existence of the attack in the initial moments of the attack. The use of resilient control algorithms in the network to isolate the misbehave agent in the leader-follower mechanism has been investigated. In the presented control method, after the attack detection phase with the use of a deep neural network, the control system uses the reputation algorithm to isolate the misbehave agent. Experimental analysis shows us that deep learning algorithms can detect attacks with higher performance that usual methods and can make cyber security simpler, more proactive, less expensive and far more effective.

**INTRODUCTION**

Recent advances in technology have led to the introduction of cyber-physical systems, which due to their better computational and communicational ability and integration between physical and cyber-components, has led to significant advances in many dynamic applications. But this improvement comes at the cost of being vulnerable to cyber-attacks. Cyber-physical systems are made up of logical elements and embedded computers, which communicate with communication channels such as the Internet of Things(IoT). More specifically, these systems include digital or cyber components, analog components, physical devices and humans that designed to operate between physical and cyber parts. In other words, a cyber-physical system is any system that includes cyber and physical components and humans, and has the ability to trade between the physical and cyber parts. In cyber-physical systems, the security of these types of systems becomes more important due to the addition of the physical part.

Physical components including sensors, which receive data from the physical environment, maybe attacked and be injected incorrect data into the system. One of the most important challenges of a cyber-physical system, in its physical part is the presence of a large number of sensors in the environment, which collect so much data, with so much variety, and at high speed. Also, the connection between the sensors and the necessary calculations and the analysis of the obtained data will be among the main challenges. Therefore, one of the most important features of a cyber-physical system is to communicate between these sensors, compute and control the system

The security of cyber-physical systems to detect cyber-attacks is an important issue in these systems . It should be noted that cyber-attacks occur in irregular ways, and it is not possible to describe these attacks in a regular and orderly manner. In general, cyber attacks in cyber-physical systems are divided into two main types: denial of service(Dos) and deception attacks. In denial of service, the attacker prevents communication between network nodes and communication channels. However, in the deception attacks that inject false data to system, which are carried out by abusing system components , such as sensors or controllers and it can corrupt data or enter incorrect information into the system and cause misbehaving.

These attacks can be detected by system monitoring in the system. But if the attacker can plan a high-level attack to prevent himself from being identified, these attacks are called stealthy deception attacks, and other common methods of counteracting such attacks will not work. Therefore, it is important to be aware of the attacks that occur in order to respond in a timely manner to attackers. In other words, the security system must be aware of the attack, otherwise it will not be able to detect and control the attack. Cyber defense can be improved by using security analytic to search for hidden patterns and how to deceive.

**LITERATURE SURVEY**

**[1] Kwon, Cheolhyeon, Weiyi Liu, and Inseok Hwang. ”Security analysis for cyber-physical systems against stealthy deception attacks.” In 2013 American control conference, IEEE (2013): 3344-3349**

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drastically

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drastically

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

Mobile app distribution platforms such as Google Play Store allow

users to share their feedback about downloaded apps in the form

of a review comment and a corresponding star rating. Typically,

the star rating ranges from one to ve stars, with one star denoting

a high sense of dissatisfaction with the app and ve stars denoting

a high sense of satisfaction.

Unfortunately, due to a variety of reasons, oen the star rating

provided by a user is inconsistent with the opinion expressed in

the review. For example, consider the following review for the

Facebook App on Android; “Awesome App”. One would reasonably

expect the rating for this review to be ve stars, but the actual

rating is one star!

Such inconsistent ratings can lead to a deated (or inated)

overall average rating of an app which can aect user downloads,

as typically users look at the average star ratings while making a

decision on downloading an app. Also, the app developers receive

a biased feedback about the application that does not represent

ground reality. is is especially signicant for small apps with a

few thousand downloads as even a small number of mismatched

reviews can bring down the average rating drasticall

The security issue in the state estimation problem is investigated for a networked control system (NCS). The communication channels between the sensors and the remote estimator in the NCS are vulnerable to attacks from malicious adversaries. The false data injection attacks are considered. The aim of this paper to find the so-called insecurity conditions under which the estimation system is insecure in the sense that there exist malicious attacks that can bypass the anomaly detector but still lead to unbounded estimation errors. In particular, a new necessary and sufficient condition for the insecurity is derived in the case that all communication channels are compromised by the adversary. Moreover, a specific algorithm is proposed for generating attacks with which the estimation system is insecure. Furthermore, for the insecure system, a system protection scheme through which only a few (rather than all) communication channels require protection against false data injection attacks is proposed. A simulation example is utilized to demonstrate the effectiveness of the proposed conditions/algorithms in the secure estimation problem for a flight vehicle.

**[2] Pajic, Miroslav, James Weimer, Nicola Bezzo, Oleg Sokolsky, George J. Pappas, and Insup Lee. ”Design and implementation of attack-resilient cyberphysical systems: With a focus on attack-resilient state estimators.” IEEE Control Systems Magazine 37, no. 2 (2017): 66-81.**

Recent years have witnessed a significant increase in the number of security-related incidents in control systems. These include high-profile attacks in a wide range of application domains, from attacks on critical infrastructure, as in the case of the Maroochy Water breach [1], and industrial systems (such as the StuxNet virus attack on an industrial supervisory control and data acquisition system [2], [3] and the German Steel Mill cyberattack [4], [5]), to attacks on modern vehicles [6]-[8]. Even high-assurance military systems were shown to be vulnerable to attacks, as illustrated in the highly publicized downing of the RQ-170 Sentinel U.S. drone [9]-[11]. These incidents have greatly raised awareness of the need for security in cyberphysical systems (CPSs), which feature tight coupling of computation and communication substrates with sensing and actuation components. However, the complexity and heterogeneity of this next generation of safety-critical, networked, and embedded control systems have challenged the existing design methods in which security is usually consider as an afterthought.

**[3]** **Sheng, Long, Ya-Jun Pan, and Xiang Gong. ”Consensus formation control for a class of networked multiple mobile robot systems.” Journal of Control Science and Engineering 2012 (2012).**

environments and requirements. Tracking and understanding

changes in modules and relationships in a software project is

difﬁcult, but even more so when the software goes through several

types of changes. The typical complexity and size of software

also makes it harder to grasp software evolution patterns. In

this paper, we present an interactive matrix-based visualization

technique that, combined with animation, depicts how software

designs evolve. For example, it shows which new modules

and couplings are added and removed over time. Our generic

visualization supports dynamic and weighted digraphs and is

applied in the context of software evolution. Analyzing source

code changes is important to determine the software’s structural

organization and identify quality issues over time. To demonstrate

our approach, we explore open-source repositories and discuss

some of our ﬁndings regarding these evolving software designs

Embedded computational resources in autonomous robotic vehicles are becoming more abundant and have enabled improved operational effectiveness of cooperative robotic systems in civilian and military applications. Compared to autonomous robotic vehicles that operate single tasks, cooperative teamwork has greater efficiency and operational capability. Multirobotic vehicle systems have many potential applications, such as platooning of vehicles in urban transportation, the operation of the multiple robots, autonomous underwater vehicles, and formation of aircrafts in military affairs [1–3]. The study of group behaviors for multirobot systems is the main objective of the work. Group cooperative behavior signifies that individuals in the group share a common objective and action according to the interest of the whole group. Group cooperation can be efficient if individuals in the group coordinate their actions well. Each individual can coordinate with other individuals in the group to facilitate group cooperative behavior in two ways, named local coordination and global coordination. For the local coordination, individuals react only to other individuals that are close, such as fish engaged in a school.

**[4]** **Zeng, Wente, and Mo-Yuen Chow. ”Resilient distributed control in the presence of misbehaving agents in networked control systems.” IEEE transactions on cybernetics 44, no. 11 (2014): 2038-2049.**

In this paper, we study the problem of reaching a consensus among all the agents in the networked control systems (NCS) in the presence of misbehaving agents. A reputation-based resilient distributed control algorithm is first proposed for the leader-follower consensus network. The proposed algorithm embeds a resilience mechanism that includes four phases (detection, mitigation, identification, and update), into the control process in a distributed manner. At each phase, every agent only uses local and one-hop neighbors' information to identify and isolate the misbehaving agents, and even compensate their effect on the system. We then extend the proposed algorithm to the leaderless consensus network by introducing and adding two recovery schemes (rollback and excitation recovery) into the current framework to guarantee the accurate convergence of the well-behaving agents in NCS. The effectiveness of the proposed method is demonstrated through case studies in multirobot formation control and wireless sensor networks.

**[5]** **Sun, Hongtao, Chen Peng, Taicheng Yang, Hao Zhang, and Wangli He. ”Resilient control of networked control systems with stochastic denial of service attacks.” Neurocomputing 270 (2017): 170-177.**

This paper focuses on resilient control of networked control systems (NCSs) under the denial of service (DoS) attacks which is characterized by a Markov process. Firstly, the packets dropout are modeled as Markov process according to the game between attack strategies and defense strategies. Then, an NCS under such game results is modeled as a Markovian jump linear system and four theorems are proved for the system stability analysis and controller design. Finally, a numerical example is used to illustrative the application of these theorems. Networked control systems (NCSs) have received an increasing attention in the past decades. Now, NCSs have been widely applied in industrial processes, electric power networks, intelligent transportation and so on . With the growing of the NCSs, network, as a critical element in an NCS, is vulnerable to cyber-threats which can menace the control systems.

**SYSTEM ANALYSIS & FEASIBILITY STUDY**

**Existing Method:**

In the existing system, implementation of machine learning algorithms is bit complex to build due to the lack of information about the data visualization. Mathematical calculations are used in existing system for model building this may takes the lot of time and complexity. To overcome all this, we use machine learning packages available in the scikit-learn library.

**Disadvantages:**

* High complexity.
* Time consuming.

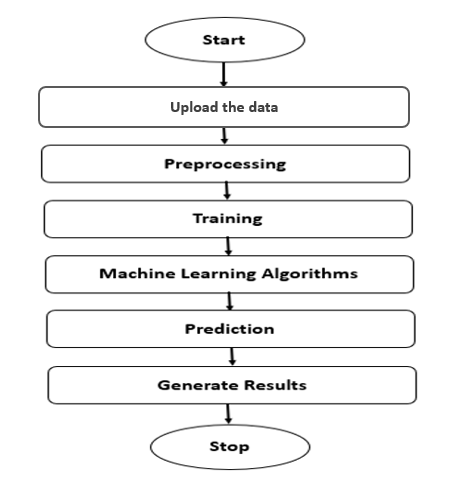
**Proposed System:**

Proposed several machine learning models to classify whether there will be a cyber-attack or not, but none have adequately addressed this misdiagnosis problem. Also, similar studies that have proposed models for evaluation of such performance classification mostly do not consider the heterogeneity and the size of the data Therefore, we propose a Support Vector , Decision Tree, Random forest, Extra Tree Classifier and ad boost and neural network classifier classification techniques.

**Advantages**:

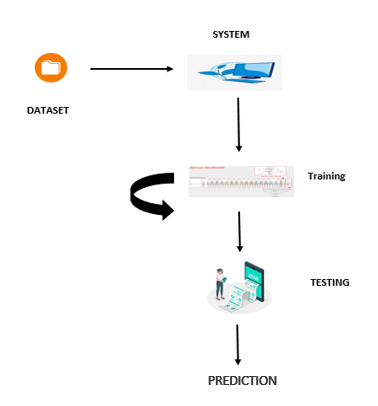
* Highest accuracy
* Reduces time complexity.

**Block Diagram:**



**Fig**: Block Diagram

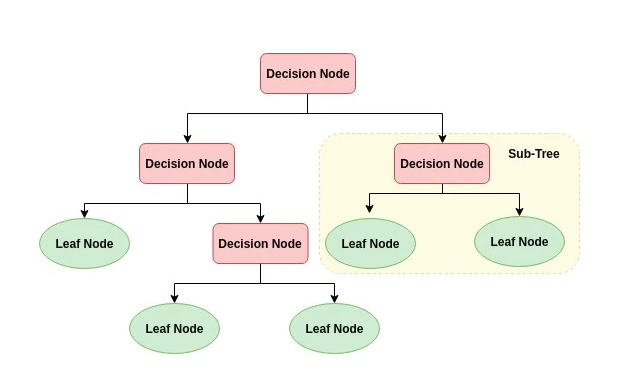
**Architecture:**



**METHODOLOGY AND ALGORITHMS:**

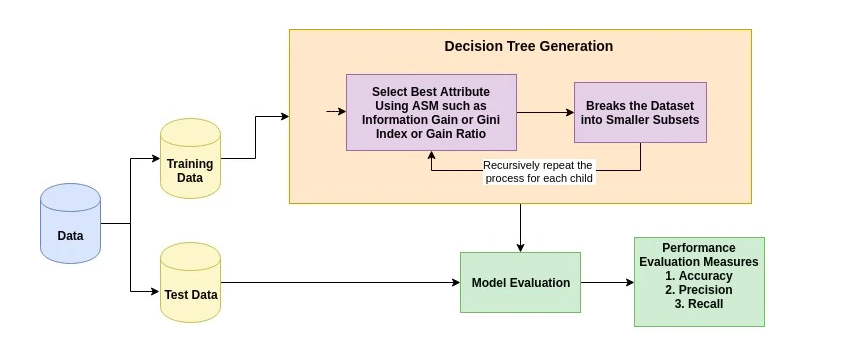
**1. DECISION TREE:**

Decision tree is a flowchart-like tree structure where an internal node represents feature(or attribute), the branch represents a decision rule, and each leaf node represents the outcome. The topmost node in a decision tree is known as the root node. It learns to partition on the basis of the attribute value. It partitions the tree in recursively manner call recursive partitioning. This flowchart-like structure helps you in decision making. It's visualization like a flowchart diagram which easily mimics the human level thinking. That is why decision trees are easy to understand and interpret.

The basic idea behind any decision tree algorithm is as follows:

1. Select the best attribute using Attribute Selection Measures (ASM) to split the records.
2. Make that attribute a decision node and breaks the dataset into smaller subsets.
3. Starts tree building by repeating this process recursively for each child until one of the conditions will match:

* All the tuples belong to the same attribute value.
* There are no more remaining attributes.
* There are no more instances.



**2. Extra tree classifier:**

ExtraTreesClassifier is an ensemble learning method fundamentally based on decision trees. ExtraTreesClassifier, like RandomForest, randomizes certain decisions and subsets of data to minimize over-learning from the data and overfitting. Let’s look at some ensemble methods ordered from high to low variance, ending in ExtraTreesClassifier.

**Decision Tree (High Variance)**

A single decision tree is usually overfits the data it is learning from because it learn from only one pathway of decisions. Predictions from a single decision tree usually don’t make accurate predictions on new data.

## **Random Forest (Medium Variance)**

Random forest models reduce the risk of overfitting by introducing randomness by:

* building multiple trees (n\_estimators)
* drawing observations with replacement (i.e., a bootstrapped sample)
* splitting nodes on the best split among a **random subset**of the features selected at every node

## Extra Trees (Low Variance)

Extra Trees is like Random Forest, in that it builds multiple trees and splits nodes using random subsets of features, but with two key differences: it does not bootstrap observations (meaning it samples without replacement), and nodes are split on random splits, not best splits. So, in summary, ExtraTrees:

* builds multiple trees with **bootstrap** = **False**by default, which means it samples without replacement
* nodes are split based on **random** splits among a **random subset**of the features selected at every node

In Extra Trees, randomness doesn’t come from bootstrapping of data, but rather comes from the random splits of all observations.

ExtraTrees is named for (Extremely Randomized Trees).

**3. Random Forest Classifier:**

A random forest is a machine learning technique that’s used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

A random forest algorithm consists of many decision trees. The ‘forest’ generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.

The (random forest) algorithm establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome.

A random forest eradicates the limitations of a decision tree algorithm. It reduces the over fitting of datasets and increases precision. It generates predictions without requiring many configurations in packages (like [Scikit-learn](https://en.wikipedia.org/wiki/Scikit-learn)).

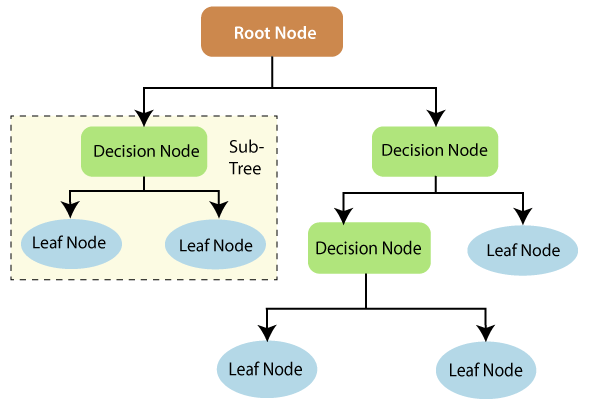
Features of a Random Forest Algorithm:

* It’s more accurate than the decision tree algorithm.
* It provides an effective way of handling missing data.
* It can produce a reasonable prediction without hyper-parameter tuning.
* It solves the issue of over fitting in decision trees.
* In every random forest tree, a subset of features is selected randomly at the node’s splitting point.

Decision trees are the building blocks of a random forest algorithm. A decision tree is a decision support technique that forms a tree-like structure. An overview of decision trees will help us understand how random forest algorithms work.

A decision tree consists of three components: decision nodes, leaf nodes, and a root node. A decision tree algorithm divides a training dataset into branches, which further segregate into other branches. This sequence continues until a leaf node is attained. The leaf node cannot be segregated further.

The nodes in the decision tree represent attributes that are used for predicting the outcome. Decision nodes provide a link to the leaves. The following diagram shows the three types of nodes in a decision tree.



The information theory can provide more information on how decision trees work. Entropy and information gain are the building blocks of decision trees. An overview of these fundamental concepts will improve our understanding of how decision trees are built.

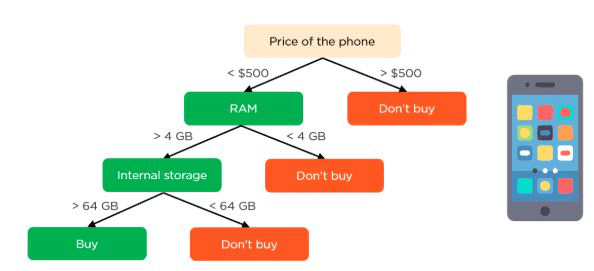
Entropy is a metric for calculating uncertainty. Information gain is a measure of how uncertainty in the target variable is reduced, given a set of independent variables.

The information gain concept involves using independent variables (features) to gain information about a target variable (class). The entropy of the target variable (Y) and the [conditional entropy](https://en.wikipedia.org/wiki/Conditional_entropy) of Y (given X) are used to estimate the information gain. In this case, the conditional entropy is subtracted from the entropy of Y.

Information gain is used in the training of decision trees. It helps in reducing uncertainty in these trees. A high information gain means that a high degree of uncertainty (information entropy) has been removed. Entropy and information gain are important in splitting branches, which is an important activity in the construction of decision trees.

Let’s take a simple example of how a decision tree works. Suppose we want to predict if a customer will purchase a mobile phone or not. The features of the phone form the basis of his decision. This analysis can be presented in a decision tree diagram.

The root node and decision nodes of the decision represent the features of the phone mentioned above. The leaf node represents the final output, either *buying* or *not buying*. The main features that determine the choice include the price, internal storage, and Random Access Memory (RAM). The decision tree will appear as follows.



Applying decision trees in random forest

The main difference between the decision tree algorithm and the random forest algorithm is that establishing root nodes and segregating nodes is done randomly in the latter. The random forest employs the bagging method to generate the required prediction.

Bagging involves using different samples of data (training data) rather than just one sample. A training dataset comprises observations and features that are used for making predictions. The decision trees produce different outputs, depending on the training data fed to the random forest algorithm. These outputs will be ranked, and the highest will be selected as the final output.

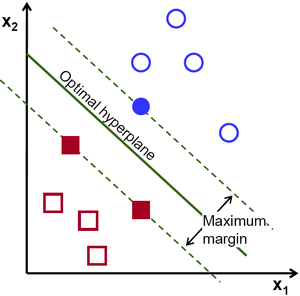
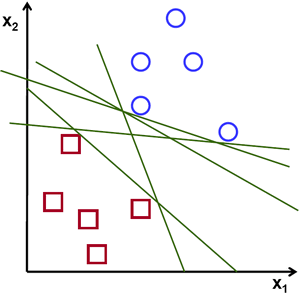
Our first example can still be used to explain how random forests work. Instead of having a single decision tree, the random forest will have many decision trees. Let’s assume we have only four decision trees. In this case, the training data comprising the phone’s observations and features will be divided into four root nodes.

The root nodes could represent four features that could influence the customer’s choice (price, internal storage, camera, and RAM). The random forest will split the nodes by selecting features randomly. The final prediction will be selected based on the outcome of the four trees.

The outcome chosen by most decision trees will be the final choice. If three trees predict *buying*, and one tree predicts *not buying*, then the final prediction will be *buying*. In this case, it’s predicted that the customer will buy the phone.

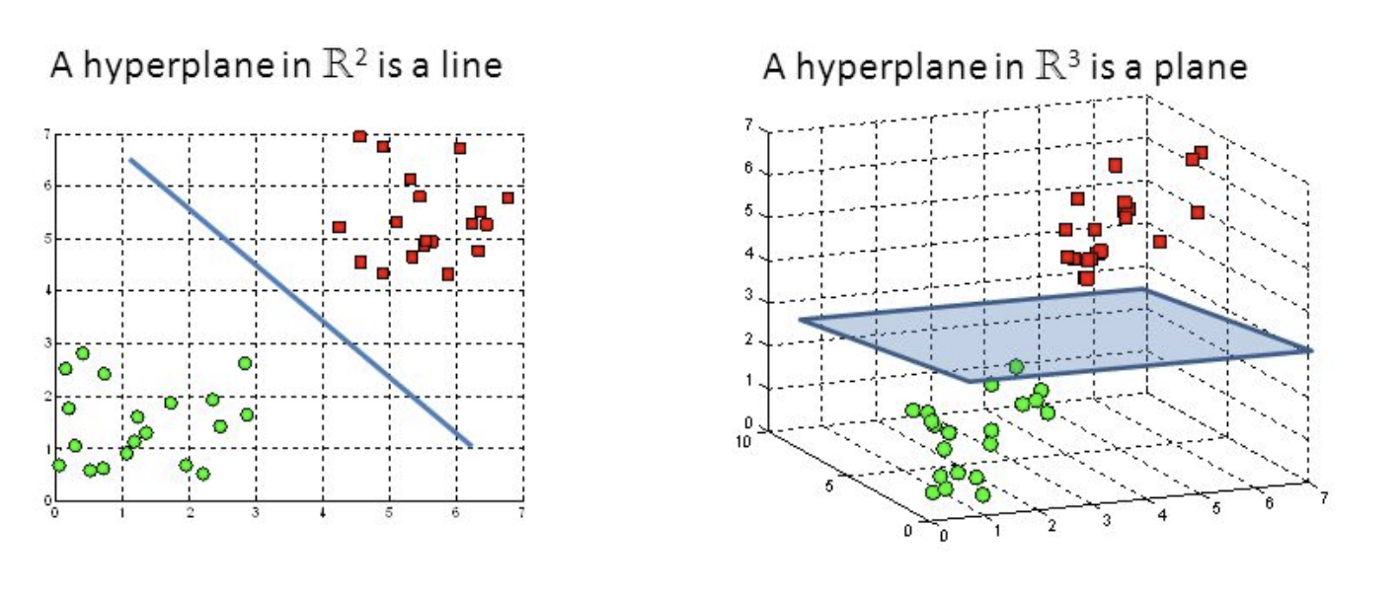
**4. SUPPORT VECTOR MACHINES:**

The objective of the support vector machine algorithm is to find a hyper plane in an N-dimensional space (N — the number of features) that distinctly classifies the data points.

**Possible hyper planes :**

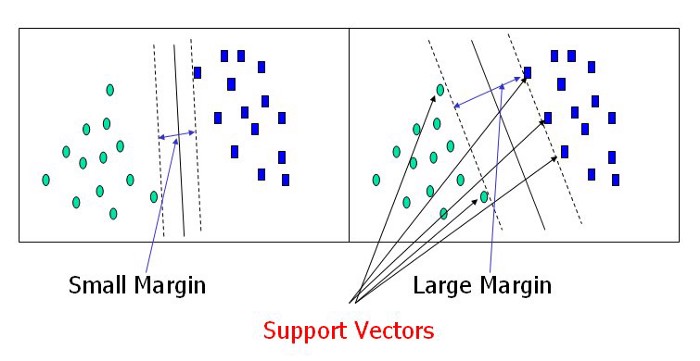
To separate the two classes of data points, there are many possible Hyper planes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e. the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.

## Hyper planes and Support Vectors



**Hyper planes in 2D and 3D feature space**

Hyper planes are decision boundaries that help classify the data points. Data points falling on either side of the hyper plane can be attributed to different classes. Also, the dimension of the hyper plane depends upon the number of features. If the number of input features is 2, then the hyper plane is just a line. If the number of input features is 3, then the hyper plane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3.



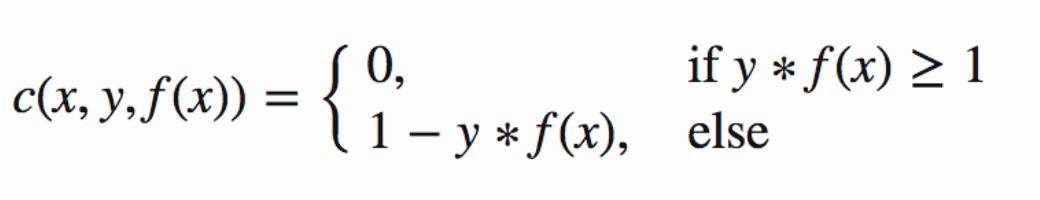
**Support Vectors**

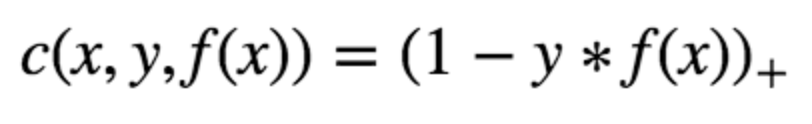
Support vectors are data points that are closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build our SVM.

**Large Margin Intuition**

In logistic regression, we take the output of the linear function and squash the value within the range of [0,1] using the sigmoid function. If the squashed value is greater than a threshold value (0.5) we assign it a label 1, else we assign it a label 0. In SVM, we take the output of the linear function and if that output is greater than 1, we identify it with one class and if the output is -1, we identify is with another class. Since the threshold values are changed to 1 and -1 in SVM, we obtain this reinforcement range of values ([-1, 1]) which acts as margin.

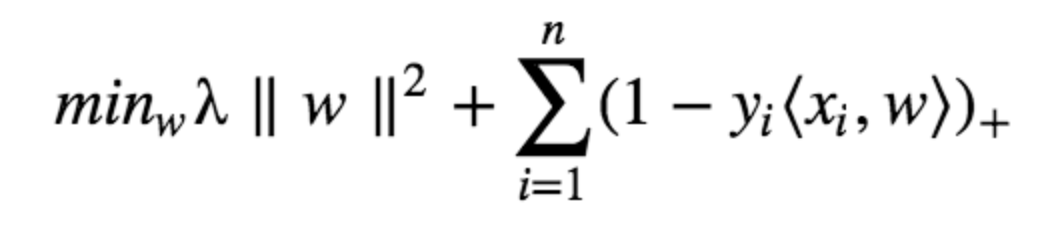
**Cost Function and Gradient Updates**

In the SVM algorithm, we are looking to maximize the margin between the data points and the hyper plane. The loss function that helps maximize the margin is hinge loss.



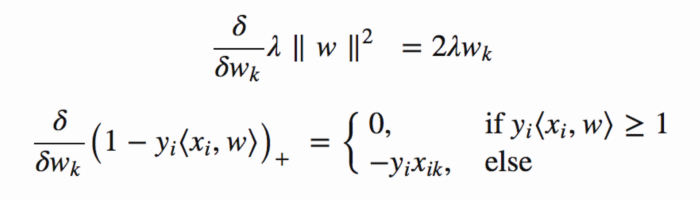
Hinge loss function (function on left can be represented as a function on the right)

The cost is 0 if the predicted value and the actual value are of the same sign. If they are not, we then calculate the loss value. We also add a regularization parameter the cost function. The objective of the regularization parameter is to balance the margin maximization and loss. After adding the regularization parameter, the cost functions looks as below.



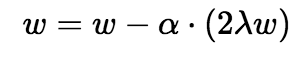
**Loss function for SVM**

Now that we have the loss function, we take partial derivatives with respect to the weights to find the gradients. Using the gradients, we can update our weights.



Gradients

When there is no misclassification, i.e. our model correctly predicts the class of our data point, we only have to update the gradient from the regularization parameter.



Gradient Update — No misclassification

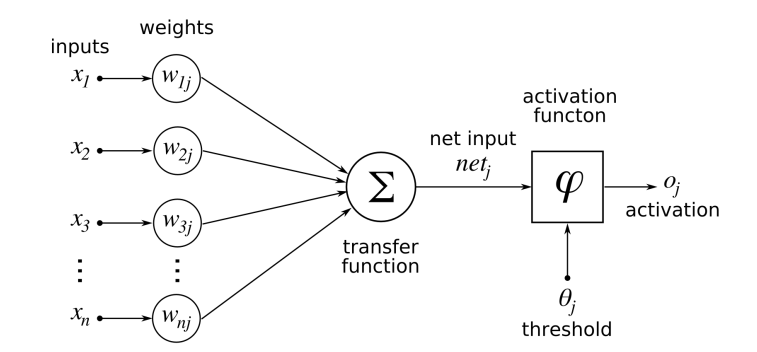
When there is a misclassification, i.e. our model make a mistake on the prediction of the class of our data point, we include the loss along with the regularization parameter to perform gradient update.

**5. Neural Network:**

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. It is the foundation of artificial intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards. ANNs have self-learning capabilities that enable them to produce better results as more data becomes available.

An ANN has hundreds or thousands of artificial neurons called processing units, which are interconnected by nodes. These processing units are made up of input and output units. The input units receive various forms and structures of information based on an internal weighting system, and the neural network attempts to learn about the information presented to produce one output report. Just like humans need rules and guidelines to come up with a result or output, ANNs also use a set of learning rules called backpropagation, an abbreviation for backward propagation of error, to perfect their output results.

An ANN initially goes through a training phase where it learns to recognize patterns in data, whether visually, aurally, or textually. During this supervised phase, the network compares its actual output produced with what it was meant to produce—the desired output. The difference between both outcomes is adjusted using backpropagation. This means that the network works backward, going from the output unit to the input units to adjust the weight of its connections between the units until the difference between the actual and desired outcome produces the lowest possible error.



Whenever we increase the layers in our ANN then it is nothing but our Deep Neural Networks.

A **deep neural network** (DNN) is an artificial **neural network** (ANN) with multiple layers between the input and output layers. There are different types of **neural networks** but they always consist of the same components: neurons, synapses, weights, biases, and functions.

**SOFTWARE DEVELOPMENT LIFE CYCLE – SDLC:**

In our project we use waterfall model as our software development cycle because of its step-by-step procedure while implementing.



**Fig1**: Waterfall Model

* **Requirement Gathering and analysis** − All possible requirements of the system to be developed are captured in this phase and documented in a requirement specification document.
* **System Design** − the requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture.
* **Implementation** − with inputs from the system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality, which is referred to as Unit Testing.
* **Integration and Testing** − All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.
* **Deployment of system** − Once the functional and non-functional testing is done; the product is deployed in the customer environment or released into the market.
* **Maintenance** − There are some issues which come up in the client environment. To fix those issues, patches are released. Also, to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment.

**FEASIBILITY STUDY**

The feasibility of the project is analysed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

* ECONOMICAL FEASIBILITY
* TECHNICAL FEASIBILITY
* SOCIAL FEASIBILITY

**Economic feasibility:**

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

### Technical feasibility:

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

**Social feasibility:**

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

**SYSTEM REQUIREMENTS SPECIFICATION**

**Functional and non-functional requirements:**

Requirement’s analysis is very critical process that enables the success of a system or software project to be assessed. Requirements are generally split into two types: Functional and non-functional requirements.

**Functional Requirements**: These are the requirements that the end user specifically demands as basic facilities that the system should offer. All these functionalities need to be necessarily incorporated into the system as a part of the contract. These are represented or stated in the form of input to be given to the system, the operation performed and the output expected. They are basically the requirements stated by the user which one can see directly in the final product, unlike the non-functional requirements.

Examples of functional requirements:

1. Authentication of user whenever he/she logs into the system
2. System shutdown in case of a cyber-attack
3. A verification email is sent to user whenever he/she register for the first time on some software system.

**Non-functional requirements**: These are basically the quality constraints that the system must satisfy according to the project contract. The priority or extent to which these factors are implemented varies from one project to other. They are also called non-behavioral requirements.  
They basically deal with issues like:

* Portability
* Security
* Maintainability
* Reliability
* Scalability
* Performance
* Reusability
* Flexibility

Examples of non-functional requirements:

1. Emails should be sent with a latency of no greater than 12 hours from such an activity.
2. The processing of each request should be done within 10 seconds
3. The site should load in 3 seconds whenever of simultaneous users are > 10000

**SOFTWARE AND HARDWARE REQUIREMENTS:**

**Hardware:**

Operating system : Windows 7 or 7+

RAM : 8 GB

Hard disc or SSD : More than 500 GB

Processor : Intel 3rd generation or high or Ryzen with 8 GB Ram

**Software:**

Software’s : Python 3.6 or high version

IDE : PyCharm.

Framework : Flask

**SYSTEM DESIGN:**

## **Input Design:**

In an information system, input is the raw data that is processed to produce output. During the input design, the developers must consider the input devices such as PC, MICR, OMR, etc.

Therefore, the quality of system input determines the quality of system output. Well-designed input forms and screens have following properties −

* It should serve specific purpose effectively such as storing, recording, and retrieving the information.
* It ensures proper completion with accuracy.
* It should be easy to fill and straightforward.
* It should focus on user’s attention, consistency, and simplicity.
* All these objectives are obtained using the knowledge of basic design principles regarding −
  + What are the inputs needed for the system?
  + How end users respond to different elements of forms and screens.

### Objectives for Input Design:

The objectives of input design are −

* To design data entry and input procedures
* To reduce input volume
* To design source documents for data capture or devise other data capture methods
* To design input data records, data entry screens, user interface screens, etc.
* To use validation checks and develop effective input controls.

**Output Design:**

The design of output is the most important task of any system. During output design, developers identify the type of outputs needed, and consider the necessary output controls and prototype report layouts.

### Objectives of Output Design:

The objectives of input design are:

* To develop output design that serves the intended purpose and eliminates the production of unwanted output.
* To develop the output design that meets the end user’s requirements.
* To deliver the appropriate quantity of output.
* To form the output in appropriate format and direct it to the right person.
* To make the output available on time for making good decisions.

**MODULES:**

1. **User**:
   1. **View Home page:**

Here user view the home page of the cyber-attack web application.

* 1. **View about page:**

In the about page, users can learn more about the poverty classification.

* 1. **Input Model:**

The user must provide input values for the certain fields in order to get results.

* 1. **View Results:**

User view’s the generated results from the model.

* 1. **View score:**

Here user have ability to view the score in %

1. **System**
   1. **Working on dataset:**

System checks for data whether it is available or not and load the data in csv files.

* 1. **Pre-processing:**

Data need to be pre-processed according the models it helps to increase the accuracy of the model and better information about the data.

* 1. **Training the data:**

After pre-processing the data will split into two parts as train and test data before training with the given algorithms.

* 1. **Model Building**

To create a model that predicts the personality with better accuracy, this module will help user.

* 1. **Generated Score:**
  2. Here user view the score in %
  3. **Generate Results:**

We train the machine learning algorithm and predict the cyber-attack detection.

**UML DIAGRAMS**

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object-oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artefacts of software system, as well as for business modelling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems.

The UML is a very important part of developing objects-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

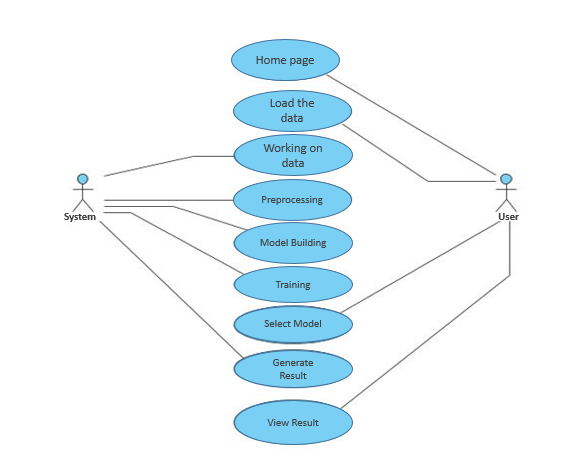
**GOALS:**

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modelling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modelling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

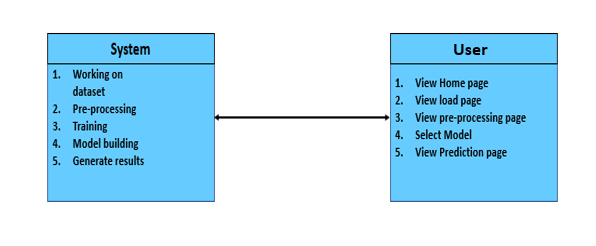
**USE CASE DIAGRAM**

* A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis.
* Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases.
* The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

****

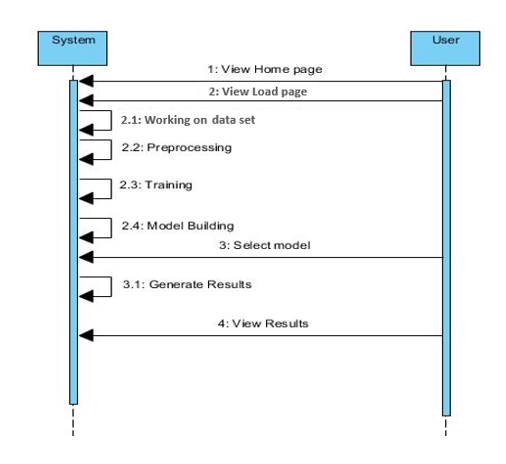
**CLASS DIAGRAM**

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information



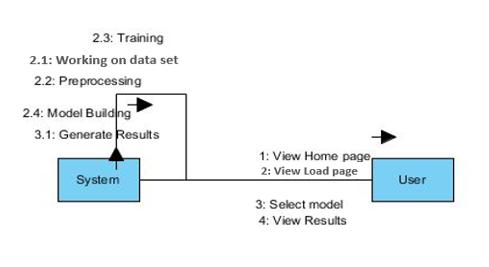
**SEQUENCE DIAGRAM**

* A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order.
* It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams



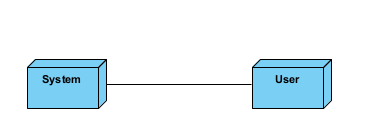
**COLLABORATION DIAGRAM:**

In collaboration diagram the method call sequence is indicated by some numbering technique as shown below. The number indicates how the methods are called one after another. We have taken the same order management system to describe the collaboration diagram. The method calls are similar to that of a sequence diagram. But the difference is that the sequence diagram does not describe the object organization whereas the collaboration diagram shows the object organization.



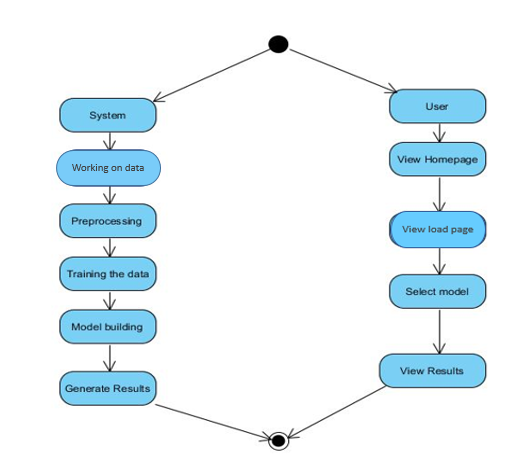
**DEPLOYMENT DIAGRAM**

Deployment diagram represents the deployment view of a system. It is related to the component diagram. Because the components are deployed using the deployment diagrams. A deployment diagram consists of nodes. Nodes are nothing but physical hardware’s used to deploy the application.



**ACTIVITY DIAGRAM:**

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.



**COMPONENT DIAGRAM**:

A component diagram, also known as a UML component diagram, describes the organization and wiring of the physical **c**omponents in a system. Component diagrams are often drawn to help model implementation details and double-check that every aspect of the system's required function is covered by planned development.



**ER DIAGRAM:**

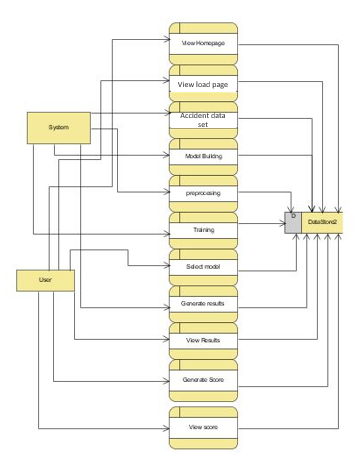
An Entity–relationship model (ER model) describes the structure of a database with the help of a diagram, which is known as Entity Relationship Diagram (ER Diagram). An ER model is a design or blueprint of a database that can later be implemented as a database. The main components of E-R model are: entity set and relationship set.

An ER diagram shows the relationship among entity sets. An entity set is a group of similar entities and these entities can have attributes. In terms of DBMS, an entity is a table or attribute of a table in database, so by showing relationship among tables and their attributes, ER diagram shows the complete logical structure of a database. Let’s have a look at a simple ER diagram to understand this concept.

# 

**DFD DIAGRAM:**

A Data Flow Diagram (DFD) is a traditional way to visualize the information flows within a system. A neat and clear DFD can depict a good amount of the system requirements graphically. It can be manual, automated, or a combination of both. It shows how information enters and leaves the system, what changes the information and where information is stored. The purpose of a DFD is to show the scope and boundaries of a system as a whole. It may be used as a communications tool between a systems analyst and any person who plays a part in the system that acts as the starting point for redesigning a system.

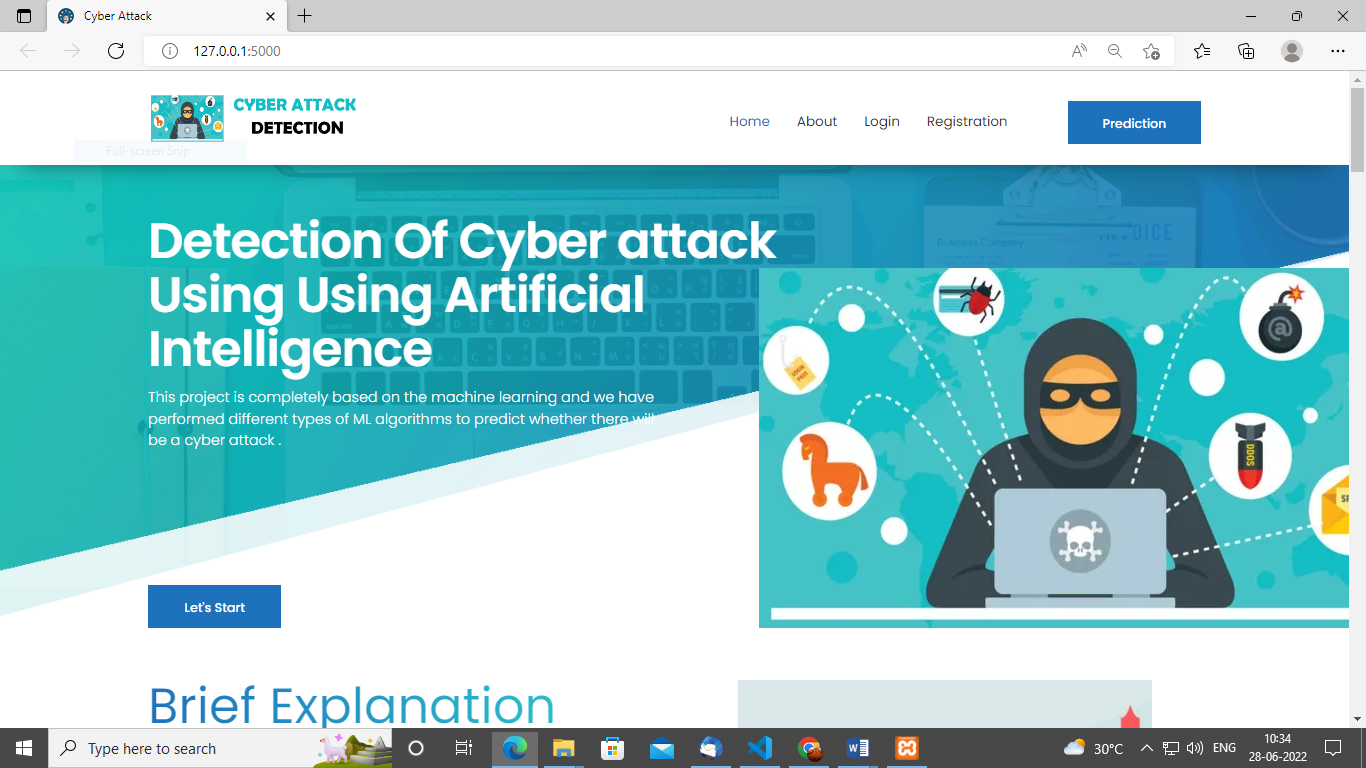
sss



**OUTPUT SCREEN SHOTS WITH DESCRIPTION.**

**Home Page:**

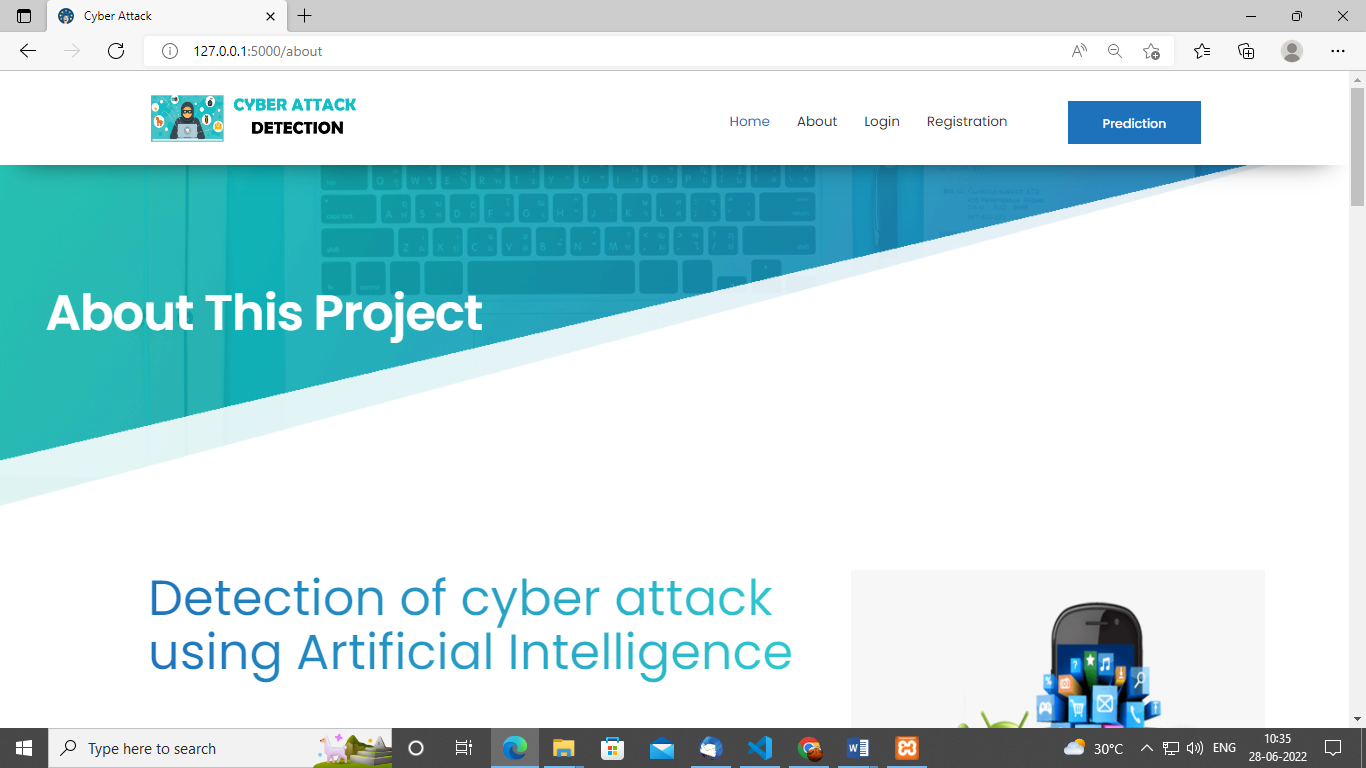
Here user view the home page of cyber-attack detection web application.



**Fig1**: Home Page

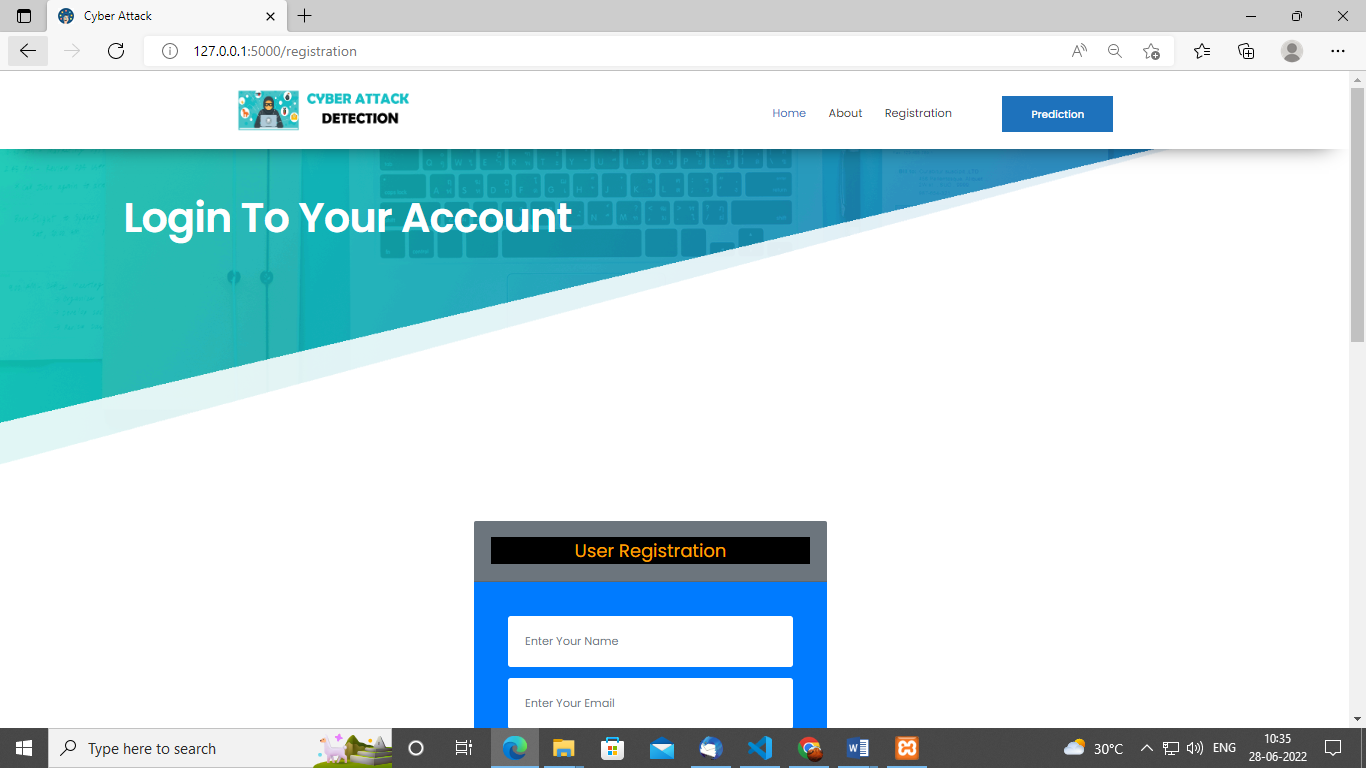
**ABOUT:**

Here we can read about our project.



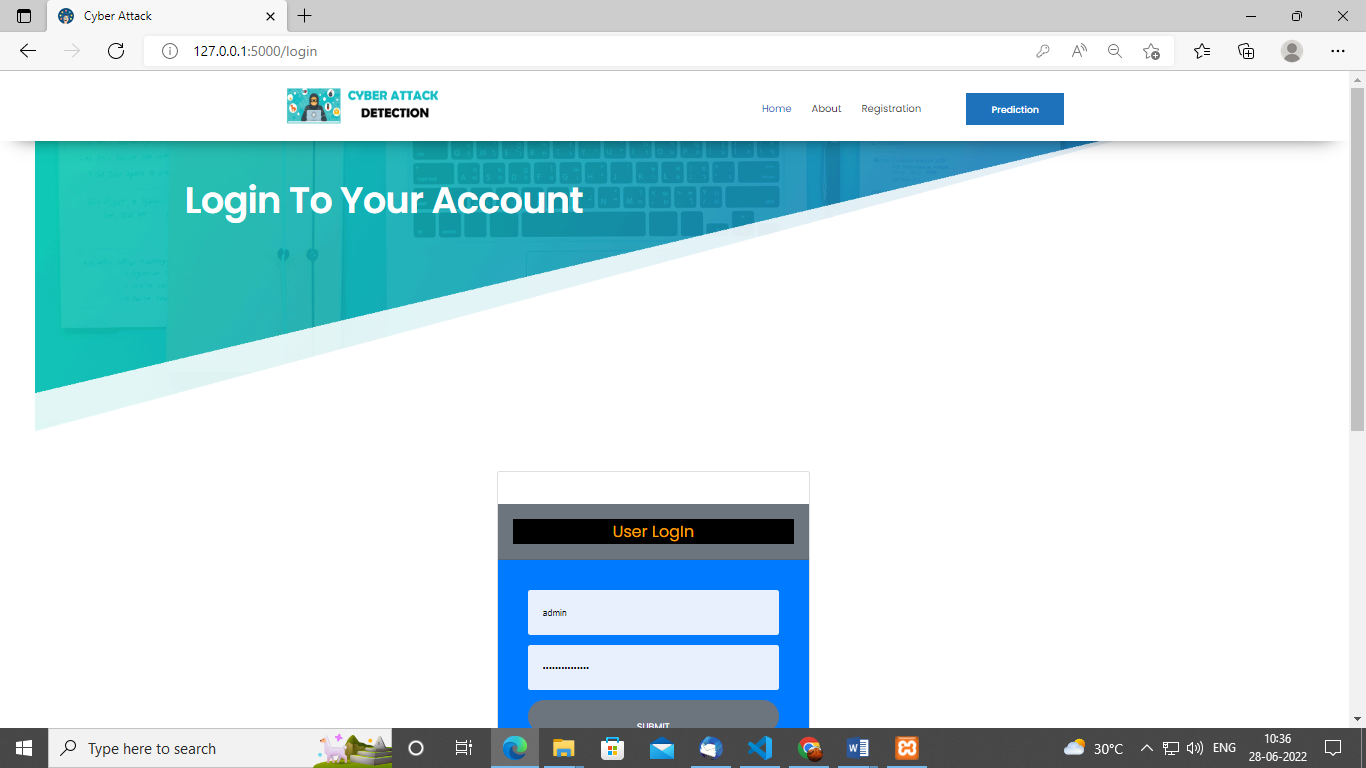
**Register:**

In the page, users need to register by entering his credentials.



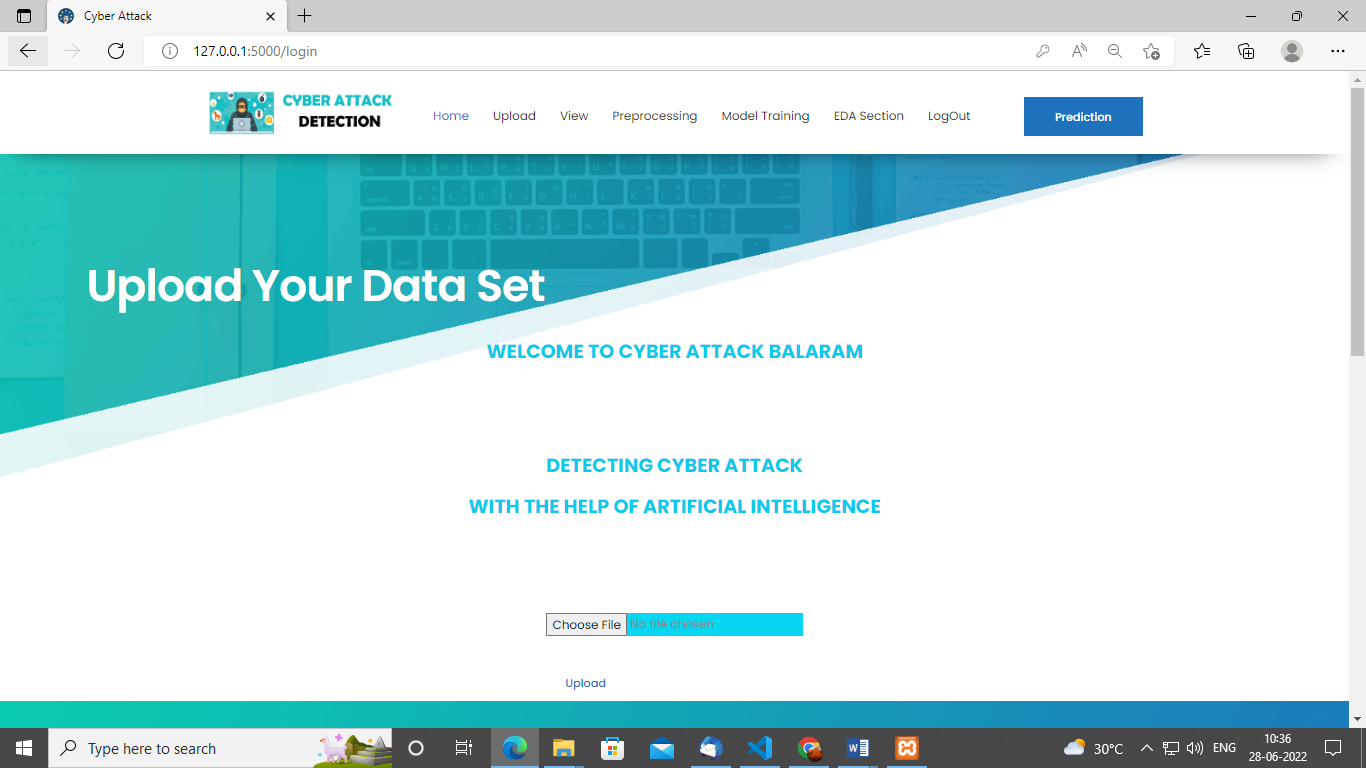
**Log in:**

In the page, users has to enter the credentials to enter into the cyber-attack prediction.



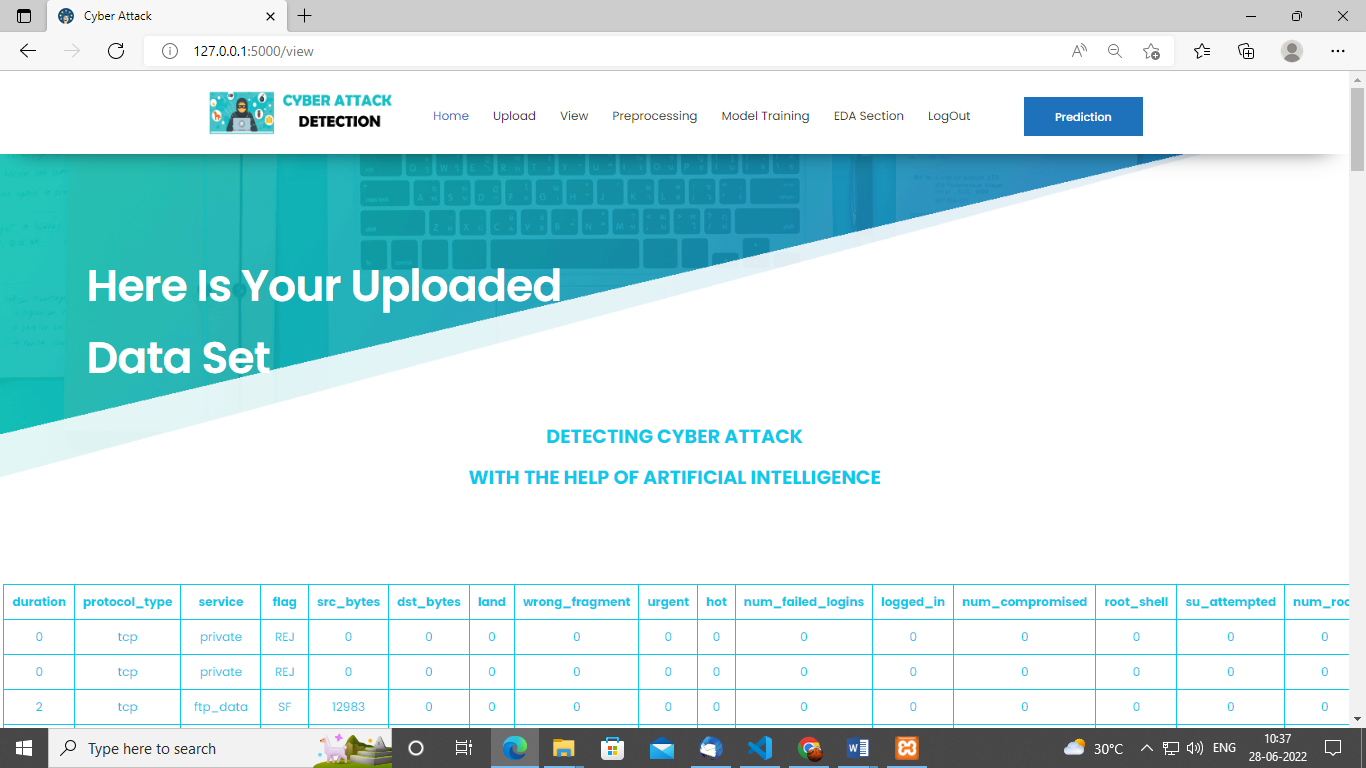
**Load:**

In the load page, users can load the cyber dataset.



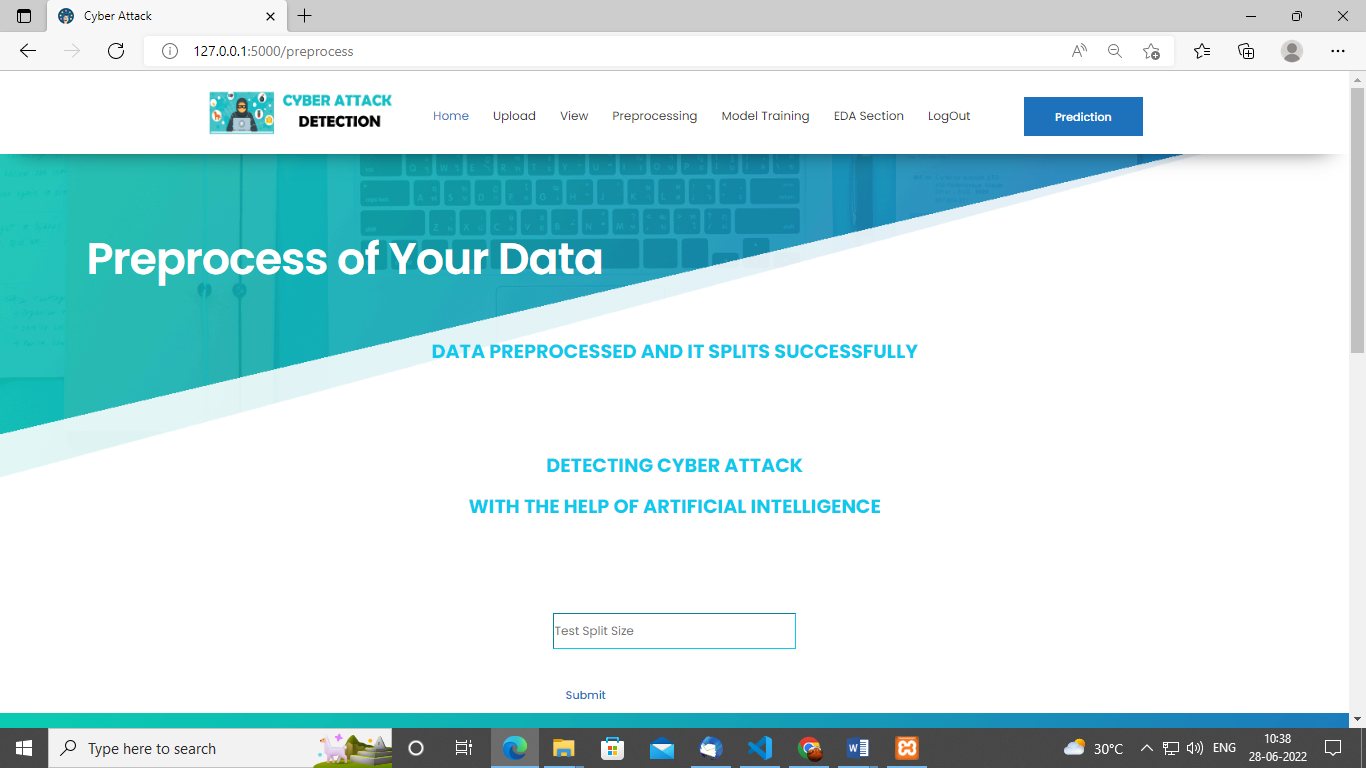
**View:**

Here we can see the uploaded data set.

****

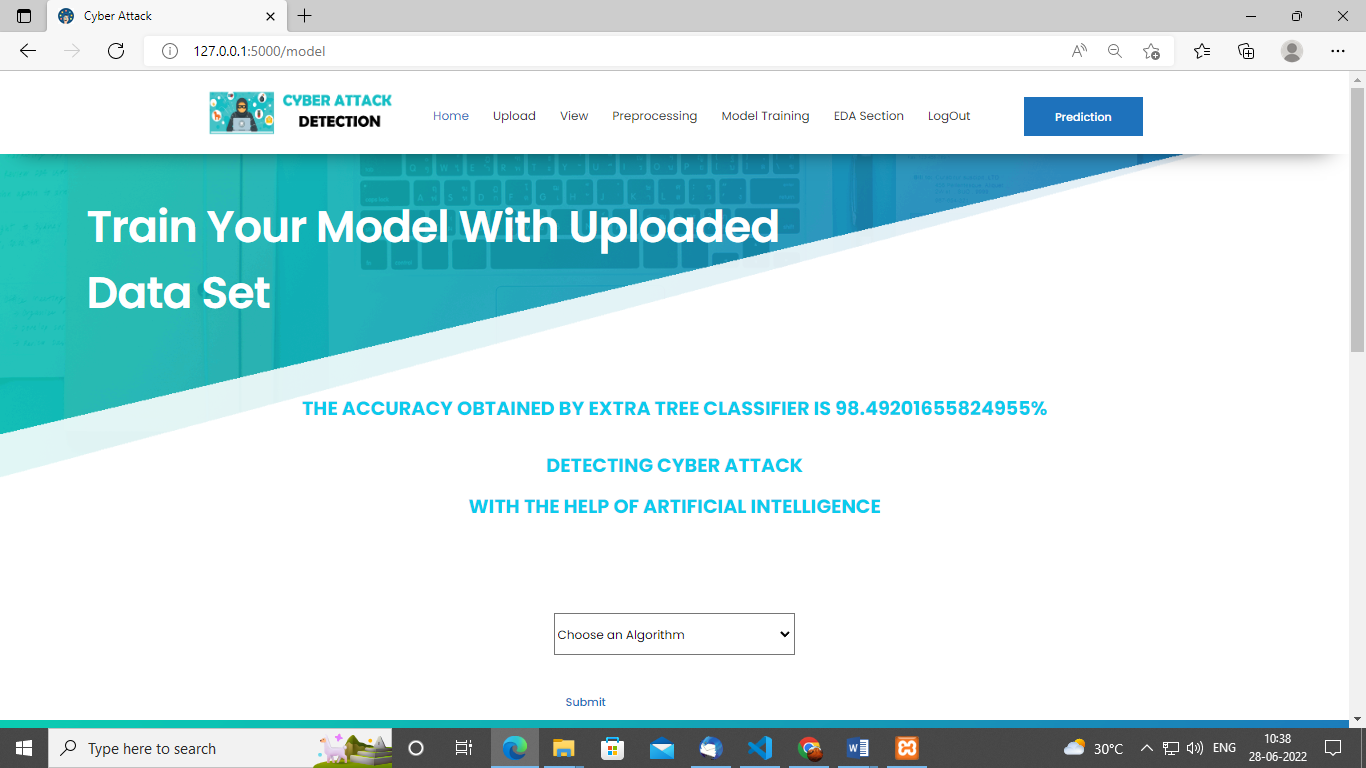
**Pre-process:**

Here we can pre-process and split our data into train and test.

`

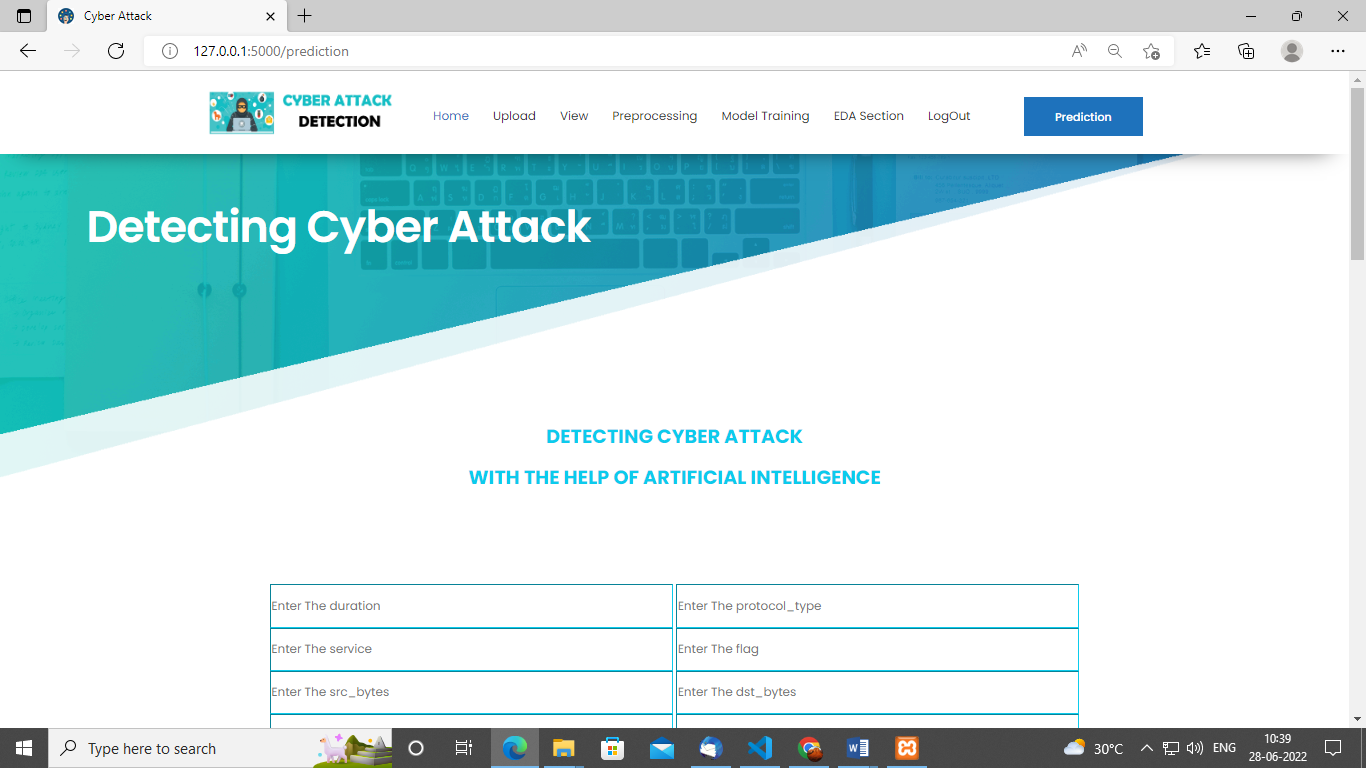
**Model:**

Here we train our data with different ML algorithms.

`

**Prediction:**

This page show the detection result of the cyber-attack detection data.



**CONCLUSION:**

In this study, an attempt was made to use the resilient control consensus method in complex discrete cyber-physical networks with a number of local attacks off. By applying this control method, it was observed that even in the presence of cyber-attacks, the system can remain stable and isolate the attacked node and the performance of the system is not weakened. Using the neural network used in this study, it was observed that with a deep neural network, with 7 hidden layers, the system shows better performance. Also in a recurrent neural network integrated with a deep neural network, a deep layer network with a linear function performs better. Therefore, it can be said that the system has less complexity. So With deep learning method, systems can analyse patterns and learn from them to help prevent similar attacks and respond to changing behaviour. In short, machine learning can make cyber security simpler, more proactive, less expensive and far more effective. After observing the state of the system reported by the neural network, the control system makes decisions based on it and, if there is an attack, detects it and isolates it, so as not to have a detrimental effect on the behaviour of other agents. In future research, more attacks on agents can be considered, also data mining and other machine learning methods, such as support vector machine (SVM) algorithms or other types of neural networks such as recurrent neural networks to evaluate system performance improvements.

**FUTURE SCOPE**

There are quite a few things that can be polished or be added in the future work. • We have opted to use two data mining classifies in this project namely the ID3 and Naive Bayes classifier. There are more classieres such as the Bayesian network classifier, Neural Network classifier and C4.5 classifier. Such classifiers were not included in this paper and could be counted in future to give a more data to be compared with.

**REFERENCES:**

[1] Kwon, Cheolhyeon, Weiyi Liu, and Inseok Hwang. ”Security analysis for cyber-physical systems against stealthy deception attacks.” In 2013 American control conference, IEEE (2013): 3344-3349.

[2] Pajic, Miroslav, James Weimer, Nicola Bezzo, Oleg Sokolsky, George J. Pappas, and Insup Lee. ”Design and implementation of attack-resilient cyberphysical systems: With a focus on attack-resilient state estimators.” IEEE Control Systems Magazine 37, no. 2 (2017): 66-81.

[3] Sheng, Long, Ya-Jun Pan, and Xiang Gong. ”Consensus formation control for a class of networked multiple mobile robot systems.” Journal of Control Science and Engineering 2012 (2012).

[4] Zeng, Wente, and Mo-Yuen Chow. ”Resilient distributed control in the presence of misbehaving agents in networked control systems.” IEEE transactions on cybernetics 44, no. 11 (2014): 2038-2049.

[5] Sun, Hongtao, Chen Peng, Taicheng Yang, Hao Zhang, and Wangli He. ”Resilient control of networked control systems with stochastic denial of service attacks.” Neurocomputing 270 (2017): 170-177.

[6] Zhang, Haotian, and Shreyas Sundaram. ”Robustness of information diffusion algorithms to locally bounded adversaries.” In 2012 American Control Conference (ACC), IEEE (2012): 5855-5861.

[7] Fu, Weiming, Jiahu Qin, Yang Shi, Wei Xing Zheng, and Yu Kang. ”Resilient Consensus of Discrete-Time Complex Cyber-Physical Networks under Deception Attacks.” IEEE Transactions on Industrial Informatics (2019).

[8] Ozay, Mete, Inaki Esnaola, Fatos Tunay Yarman Vural, Sanjeev R. Kulkarni, and H. Vincent Poor. ”Machine learning methods for attack detection in the smart grid.” IEEE transactions on neural networks and learning systems 27, no. 8 (2015): 1773-1786.

[9] Tianfield, Huaglory. ”Data mining based cyber-attack detection.” System simulation technology 13, no. 2 (2017): 90-104.

[10] Pasqualetti, Fabio, Florian Dorfler, and Francesco Bullo. ”Attack detection and ¨ identification in cyber-physical systems.” IEEE Transactions on Automatic Control 58, no. 11 (2013): 2715-2729.

**BIBLIOGRAPHY:**

**SOFTWARE INSTALLATION FOR MACHINE LEARNING PROJECTS:**

**Installing Python:**

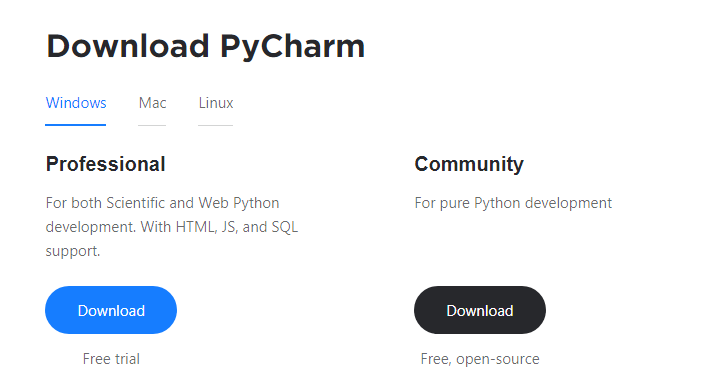
1. To download and install Python visit the official website of Python <https://www.python.org/downloads/> and choose your version.



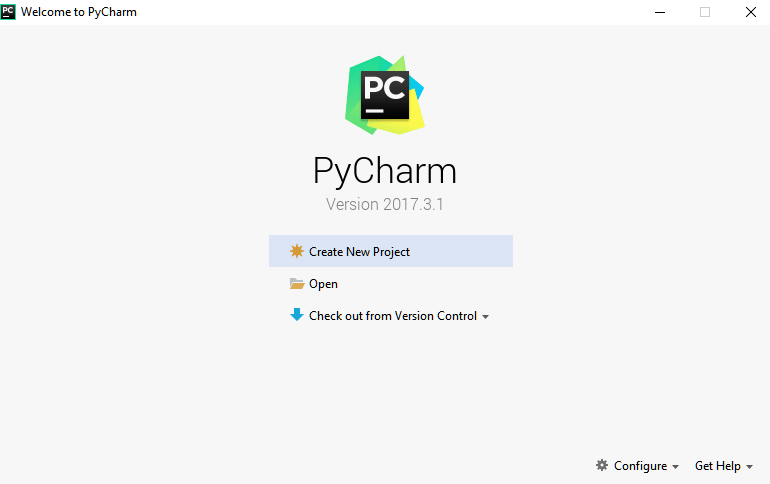
1. Once the download is complete, run the exe for install Python. Now click on Install Now.
2. You can see Python installing at this point.
3. When it finishes, you can see a screen that says the Setup was successful. Now click on "Close".

**Installing PyCharm:**

1. To download PyCharm visit the website <https://www.jetbrains.com/pycharm/download/> and click the "DOWNLOAD" link under the Community Section.



1. Once the download is complete, run the exe for install PyCharm. The setup wizard should have started. Click “Next”.
2. On the next screen, Change the installation path if required. Click “Next”.
3. On the next screen, you can create a desktop shortcut if you want and click on “Next”.
4. Choose the start menu folder. Keep selected JetBrains and click on “Install”.
5. Wait for the installation to finish.
6. Once installation finished, you should receive a message screen that PyCharm is installed. If you want to go ahead and run it, click the “Run PyCharm Community Edition” box first and click “Finish”.
7. After you click on "Finish," the Following screen will appear.

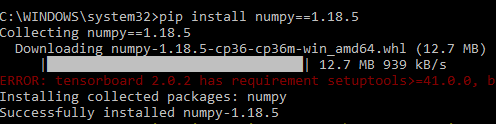


9. You need to install some packages to execute your project in a proper way.

10. Open the command prompt/ anaconda prompt or terminal as administrator.

11. The prompt will get open, with specified path, type “pip install package name” which you want to install (like NumPy, pandas, sea born, scikit-learn, Matplotlib, Pyplot)

Ex: Pip install NumPy



# **INTRODUCTION TO PYTHON**

* Python

### What Is a Script?

Up to this point, I have concentrated on the interactive programming capability of Python.  This is a very useful capability that allows you to type in a program and to have it executed immediately in an interactive mode

Scripts are reusable

Basically, a script is a text file containing the statements that comprise a Python program.  Once you have created the script, you can execute it over and over without having to retype it each time.

Scripts are editable

Perhaps, more importantly, you can make different versions of the script by modifying the statements from one file to the next using a text editor.  Then you can execute each of the individual versions.  In this way, it is easy to create different programs with a minimum amount of typing.

You will need a text editor

Just about any text editor will suffice for creating Python script files.

You can use *Microsoft Notepad, Microsoft WordPad, Microsoft Word,*or just about any word processor if you want to.

Difference between a script and a program

Script:

Scripts are distinct from the core code of the application, which is usually written in a different language, and are often created or at least modified by the end-user. Scripts are often interpreted from source code or byte code, whereas the applications they control are traditionally compiled to native machine code.

Program:

The program has an executable form that the computer can use directly to execute the instructions.

The same program in its human-readable source code form, from which executable programs are derived (e.g., compiled)

Python

What is Python? Chances you are asking yourself this. You may have found this book because you want to learn to program but don’t know anything about programming languages. Or you may have heard of programming languages like C, C++, C#, or Java and want to know what Python is and how it compares to “big name” languages. Hopefully I can explain it for you.

Python concepts

If you’re not interested in the how’s and whys of Python, feel free to skip to the next chapter. In this chapter I will try to explain to the reader why I think Python is one of the best languages available and why it’s a great one to start programming with.

• Open-source general-purpose language.

• Object Oriented, Procedural, Functional

• Easy to interface with C/ObjC/Java/Fortran

• Easy-is to interface with C++ (via SWIG)

• Great interactive environment

• Great interactive environment

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

* Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
* Python is Interactive − you can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
* Python is Object-Oriented − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.
* Python is a Beginner's Language − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

History of Python

Python was developed by Guido van Possum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands.

Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, Smalltalk, and UNIX shell and other scripting languages.

Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL).

Python is now maintained by a core development team at the institute, although Guido van Possum still holds a vital role in directing its progress.

Python Features

Python's features include −

* Easy-to-learn − Python has few keywords, simple structure, and a clearly defined syntax. This allows the student to pick up the language quickly.
* Easy-to-read − Python code is more clearly defined and visible to the eyes.
* Easy-to-maintain − Python's source code is fairly easy-to-maintained.
* A broad standard library − Python's bulk of the library is very portable and cross-platform compatible on UNIX, Windows, and Macintosh.
* Interactive Mode − Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.
* Portable − Python can run on a wide variety of hardware platforms and has the same interface on all platforms.
* Extendable − you can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.
* Databases − Python provides interfaces to all major commercial databases.
* GUI Programming − Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix.
* Scalable − Python provides a better structure and support for large programs than shell scripting.

Apart from the above-mentioned features, Python has a big list of good features, few are listed below −

* It supports functional and structured programming methods as well as OOP.
* It can be used as a scripting language or can be compiled to byte-code for building large applications.
* It provides very high-level dynamic data types and supports dynamic type checking.
* IT supports automatic garbage collection.
* It can be easily integrated with C, C++, COM, ActiveX, CORBA, and Java.

Dynamic vs. Static

Types Python is a dynamic-typed language. Many other languages are static typed, such as C/C++ and Java. A static typed language requires the programmer to explicitly tell the computer what type of “thing” each data value is.

For example, in C if you had a variable that was to contain the price of something, you would have to declare the variable as a “float” type.

This tells the compiler that the only data that can be used for that variable must be a floating point number, i.e. a number with a decimal point.

If any other data value was assigned to that variable, the compiler would give an error when trying to compile the program.

Python, however, doesn’t require this. You simply give your variables names and assign values to them. The interpreter takes care of keeping track of what kinds of objects your program is using. This also means that you can change the size of the values as you develop the program. Say you have another decimal number (a.k.a. a floating point number) you need in your program.

With a static typed language, you have to decide the memory size the variable can take when you first initialize that variable. A double is a floating point value that can handle a much larger number than a normal float (the actual memory sizes depend on the operating environment).

If you declare a variable to be a float but later on assign a value that is too big to it, your program will fail; you will have to go back and change that variable to be a double.

With Python, it doesn’t matter. You simply give it whatever number you want and Python will take care of manipulating it as needed. It even works for derived values.

For example, say you are dividing two numbers. One is a floating point number and one is an integer. Python realizes that it’s more accurate to keep track of decimals so it automatically calculates the result as a floating point number

Variables

Variables are nothing but reserved memory locations to store values. This means that when you create a variable you reserve some space in memory.

Based on the data type of a variable, the interpreter allocates memory and decides what can be stored in the reserved memory. Therefore, by assigning different data types to variables, you can store integers, decimals or characters in these variables.

Standard Data Types

The data stored in memory can be of many types. For example, a person's age is stored as a numeric value and his or her address is stored as alphanumeric characters. Python has various standard data types that are used to define the operations possible on them and the storage method for each of them.

Python has five standard data types −

* Numbers
* String
* List
* Tuple
* Dictionary

## Python Numbers

Number data types store numeric values. Number objects are created when you assign a value to them

## Python Strings

Strings in Python are identified as a contiguous set of characters represented in the quotation marks. Python allows for either pairs of single or double quotes. Subsets of strings can be taken using the slice operator ([ ] and [:]) with indexes starting at 0 in the beginning of the string and working their way from -1 at the end.

## Python Lists

Lists are the most versatile of Python's compound data types. A list contains items separated by commas and enclosed within square brackets ([]). To some extent, lists are similar to arrays in C. One difference between them is that all the items belonging to a list can be of different data type.

The values stored in a list can be accessed using the slice operator ([ ] and [:]) with indexes starting at 0 in the beginning of the list and working their way to end -1. The plus (+) sign is the list concatenation operator, and the asterisk (\*) is the repetition operator.

## Python Tuples

A tuple is another sequence data type that is similar to the list. A tuple consists of a number of values separated by commas. Unlike lists, however, tuples are enclosed within parentheses.

The main differences between lists and tuples are: Lists are enclosed in brackets ([ ]) and their elements and size can be changed, while tuples are enclosed in parentheses (( )) and cannot be updated. Tuples can be thought of as read-only lists.

## Python Dictionary

Python's dictionaries are kind of hash table type. They work like associative arrays or hashes found in Perl and consist of key-value pairs. A dictionary key can be almost any Python type, but are usually numbers or strings. Values, on the other hand, can be any arbitrary Python object.

Dictionaries are enclosed by curly braces ({ }) and values can be assigned and accessed using square braces ([]).

Different modes in python

Python has two basic modes: normal and interactive.

The normal mode is the mode where the scripted and finished .pie files are run in the Python interpreter.

Interactive mode is a command line shell which gives immediate feedback for each statement, while running previously fed statements in active memory. As new lines are fed into the interpreter, the fed program is evaluated both in part and in whole

# 20 Python libraries

1. Requests. The most famous http library written by Kenneth remits. It’s a must have for every python developer.

2. Scrappy. If you are involved in web scraping then this is a must have library for you. After using this library you won’t use any other.

3. Python. A guy toolkit for python. I have primarily used it in place of tinder. You will really love it.

4. Pillow. A friendly fork of PIL (Python Imaging Library). It is more user friendly than PIL and is a must have for anyone who works with images.

5. SQL Alchemy. A database library. Many love it and many hate it. The choice is yours.

6. Beautiful Soup. I know it’s slow but this xml and html parsing library is very useful for beginners.

7. Twisted. The most important tool for any network application developer. It has a very beautiful ape and is used by a lot of famous python developers.

8. Numbly. How can we leave this very important library? It provides some advance math functionalities to python.

9. Skippy. When we talk about numbly then we have to talk about spicy. It is a library of algorithms and mathematical tools for python and has caused many scientists to switch from ruby to python.

10. Matplotlib. A numerical plotting library. It is very useful for any data scientist or any data analyser.

11. Pygmy. Which developer does not like to play games and develop them? This library will help you achieve your goal of 2d game development.

12. Piglet. A 3d animation and game creation engine. This is the engine in which the famous [python port](https://github.com/fogleman/Minecraft) of mine craft was made

13. Pit. A GUI toolkit for python. It is my second choice after python for developing GUI’s for my python scripts.

14. Pit. Another python GUI library. It is the same library in which the famous Bit torrent client is created.

15. Scaly. A packet sniffer and analyser for python made in python.

16. Pywin32. A python library which provides some useful methods and classes for interacting with windows.

17. Notch. Natural Language Toolkit – I realize most people won’t be using this one, but it’s generic enough. It is a very useful library if you want to manipulate strings. But its capacity is beyond that. Do check it out.

18. Nose. A testing framework for python. It is used by millions of python developers. It is a must have if you do test driven development.

19. Simply. Simply can-do algebraic evaluation, differentiation, expansion, complex numbers, etc. It is contained in a pure Python distribution.

20. I Python. I just can’t stress enough how useful this tool is. It is a python prompt on steroids. It has completion, history, shell capabilities, and a lot more. Make sure that you take a look at it.

NumPy

Humpy’s main object is the homogeneous multidimensional array. It is a table of elements (usually numbers), all of the same type, indexed by a tuple of positive integers. In numbly dimensions are called axes. The number of axes is rank.

• Offers Matlab-ish capabilities within Python

• Fast array operations

• 2D arrays, multi-D arrays, linear algebra etc.

Matplotlib

• High quality plotting library.

Python class and objects

These are the building blocks of OOP. Class creates a new object. This object can be anything, whether an abstract data concept or a model of a physical object, e.g. a chair. Each class has individual characteristics unique to that class, including variables and methods. Classes are very powerful and currently “the big thing” in most programming languages. Hence, there are several chapters dedicated to OOP later in the book.

The class is the most basic component of object-oriented programming. Previously, you learned how to use functions to make your program do something.

Now will move into the big, scary world of Object-Oriented Programming (OOP). To be honest, it took me several months to get a handle on objects.

When I first learned C and C++, I did great; functions just made sense for me.

Having messed around with BASIC in the early ’90s, I realized functions were just like subroutines so there wasn’t much new to learn.

However, when my C++ course started talking about objects, classes, and all the new features of OOP, my grades definitely suffered.

Once you learn OOP, you’ll realize that it’s actually a pretty powerful tool. Plus many Python libraries and APIs use classes, so you should at least be able to understand what the code is doing.

One thing to note about Python and OOP: it’s not mandatory to use objects in your code in a way that works best; maybe you don’t need to have a full-blown class with initialization code and methods to just return a calculation. With Python, you can get as technical as you want.

As you’ve already seen, Python can do just fine with functions. Unlike languages such as Java, you aren’t tied down to a single way of doing things; you can mix functions and classes as necessary in the same program. This lets you build the code

Objects are an encapsulation of variables and functions into a single entity. Objects get their variables and functions from classes. Classes are essentially a template to create your objects.

Here’s a brief list of Python OOP ideas:

• The class statement creates a class object and gives it a name. This creates a new namespace.

• Assignments within the class create class attributes. These attributes are accessed by qualifying the name using dot syntax: ClassName.Attribute.

• Class attributes export the state of an object and its associated behaviour. These attributes are shared by all instances of a class.

• Calling a class (just like a function) creates a new instance of the class.

This is where the multiple copy’s part comes in.

• Each instance gets ("inherits") the default class attributes and gets its own namespace. This prevents instance objects from overlapping and confusing the program.

• Using the term self identifies a particular instance, allowing for per-instance attributes. This allows items such as variables to be associated with a particular instance.

Inheritance

First off, classes allow you to modify a program without really making changes to it.

To elaborate, by sub classing a class, you can change the behaviour of the program by simply adding new components to it rather than rewriting the existing components.

As we’ve seen, an instance of a class inherits the attributes of that class.

However, classes can also inherit attributes from other classes. Hence, a subclass inherits from a superclass allowing you to make a generic superclass that is specialized via subclasses.

The subclasses can override the logic in a superclass, allowing you to change the behaviour of your classes without changing the superclass at all.

Operator Overloads

Operator overloading simply means that objects that you create from classes can respond to actions (operations) that are already defined within Python, such as addition, slicing, printing, etc.

Even though these actions can be implemented via class methods, using overloading ties the behavior closer to Python’s object model and the object interfaces are more consistent to Python’s built-in objects, hence overloading is easier to learn and use.

User-made classes can override nearly all of Python’s built-in operation methods

Exceptions

I’ve talked about exceptions before but now I will talk about them in depth. Essentially, exceptions are events that modify program’s flow, either intentionally or due to errors.

They are special events that can occur due to an error, e.g. trying to open a file that doesn’t exist, or when the program reaches a marker, such as the completion of a loop.

Exceptions, by definition, don’t occur very often; hence, they are the "exception to the rule" and a special class has been created for them. Exceptions are everywhere in Python.

Virtually every module in the standard Python library uses them, and Python itself will raise them in a lot of different circumstances.

Here are just a few examples:

• Accessing a non−existent dictionary key will raise a Key Error exception.

• Searching a list for a non−existent value will raise a Value Error exception

. • Calling a non−existent method will raise an Attribute Error exception.

• Referencing a non−existent variable will raise a Name Error exception.

• Mixing data types without coercion will raise a Type Error exception.

One use of exceptions is to catch a fault and allow the program to continue working; we have seen this before when we talked about files.

This is the most common way to use exceptions. When programming with the Python command line interpreter, you don’t need to worry about catching exceptions.

Your program is usually short enough to not be hurt too much if an exception occurs.

Plus, having the exception occur at the command line is a quick and easy way to tell if your code logic has a problem.

However, if the same error occurred in your real program, it will fail and stop working. Exceptions can be created manually in the code by raising an exception.

It operates exactly as a system-caused exceptions, except that the programmer is doing it on purpose. This can be for a number of reasons. One of the benefits of using exceptions is that, by their nature, they don’t put any overhead on the code processing.

Because exceptions aren’t supposed to happen very often, they aren’t processed until they occur.

Exceptions can be thought of as a special form of the if/elf statements. You can realistically do the same thing with if blocks as you can with exceptions.

However, as already mentioned, exceptions aren’t processed until they occur; if blocks are processed all the time.

Proper use of exceptions can help the performance of your program.

The more infrequent the error might occur, the better off you are to use exceptions; using if blocks requires Python to always test extra conditions before continuing.

Exceptions also make code management easier: if your programming logic is mixed in with error-handling if statements, it can be difficult to read, modify, and debug your program.

User-Defined Exceptions

I won’t spend too much time talking about this, but Python does allow for a programmer to create his own exceptions.

You probably won’t have to do this very often but it’s nice to have the option when necessary.

However, before making your own exceptions, make sure there isn’t one of the built-in exceptions that will work for you.

They have been "tested by fire" over the years and not only work effectively, they have been optimized for performance and are bug-free.

Making your own exceptions involves object-oriented programming, which will be covered in the next chapter

. To make a custom exception, the programmer determines which base exception to use as the class to inherit from, e.g. making an exception for negative numbers or one for imaginary numbers would probably fall under the Arithmetic Error exception class.

To make a custom exception, simply inherit the base exception and define what it will do.

Python modules

Python allows us to store our code in files (also called modules). This is very useful for more serious programming, where we do not want to retype a long function definition from the very beginning just to change one mistake. In doing this, we are essentially defining our own modules, just like the modules defined already in the Python library.

To support this, Python has a way to put definitions in a file and use them in a script or in an interactive instance of the interpreter. Such a file is called a module; definitions from a module can be imported into other modules or into the main module.

Testing code

As indicated above, code is usually developed in a file using an editor.

To test the code, import it into a Python session and try to run it.

Usually there is an error, so you go back to the file, make a correction, and test again.

This process is repeated until you are satisfied that the code works. T

His entire process is known as the development cycle.

There are two types of errors that you will encounter. Syntax errors occur when the form of some command is invalid.

This happens when you make typing errors such as misspellings, or call something by the wrong name, and for many other reasons. Python will always give an error message for a syntax error.

Functions in Python

It is possible, and very useful, to define our own functions in Python. Generally speaking, if you need to do a calculation only once, then use the interpreter. But when you or others have need to perform a certain type of calculation many times, then define a function.

You use functions in programming to bundle a set of instructions that you want to use repeatedly or that, because of their complexity, are better self-contained in a sub-program and called when needed. That means that a function is a piece of code written to carry out a specified task.

## To carry out that specific task, the function might or might not need multiple inputs. When the task is carved out, the function can or cannot return one or more values.

## There are three types of functions in python:

## Help (), min (), print ().

Namespaces in Python are implemented as Python dictionaries, this means it is a mapping from names (keys) to objects (values). The user doesn't have to know this to write a Python program and when using namespaces.

Some namespaces in Python:

* global names of a module
* local names in a function or method invocation
* built-in names: this namespace contains built-in functions (e.g. abs(), camp(), ...) and built-in exception names

Garbage Collection

Garbage Collector exposes the underlying memory management mechanism of Python, the automatic garbage collector. The module includes functions for controlling how the collector operates and to examine the objects known to the system, either pending collection or stuck in reference cycles and unable to be freed.

Python XML Parser

XML is a portable, open source language that allows programmers to develop applications that can be read by other applications, regardless of operating system and/or developmental language.

What is XML? The Extensible Markup Language XML is a markup language much like HTML or SGML.

This is recommended by the World Wide Web Consortium and available as an open standard.

XML is extremely useful for keeping track of small to medium amounts of data without requiring a SQL-based backbone.

XML Parser Architectures and APIs the Python standard library provides a minimal but useful set of interfaces to work with XML.

The two most basic and broadly used APIs to XML data are the SAX and DOM interfaces.

Simple API for XML SAX: Here, you register callbacks for events of interest and then let the parser proceed through the document.

This is useful when your documents are large or you have memory limitations, it parses the file as it reads it from disk and the entire file is never stored in memory.

Document Object Model DOM API : This is a World Wide Web Consortium recommendation wherein the entire file is read into memory and stored in a hierarchical tree − based form to represent all the features of an XML document.

SAX obviously cannot process information as fast as DOM can when working with large files. On the other hand, using DOM exclusively can really kill your resources, especially if used on a lot of small files.

SAX is read-only, while DOM allows changes to the XML file. Since these two different APIs literally complement each other, there is no reason why you cannot use them both for large projects.

Python Web Frameworks

A web framework is a code library that makes a developer's life easier when building reliable, scalable and maintainable web applications.

## Why are web frameworks useful?

Web frameworks encapsulate what developers have learned over the past twenty years while programming sites and applications for the web. Frameworks make it easier to reuse code for common HTTP operations and to structure projects so other developers with knowledge of the framework can quickly build and maintain the application.

Common web framework functionality

Frameworks provide functionality in their code or through extensions to perform common operations required to run web applications. These common operations include:

1. URL routing
2. HTML, XML, JSON, and other output format tinplating
3. Database manipulation
4. Security against Cross-site request forgery (CSRF) and other attacks
5. Session storage and retrieval

Not all web frameworks include code for all of the above functionality. Frameworks fall on the spectrum from executing a single use case to providing every known web framework feature to every developer. Some frameworks take the "batteries-included" approach where everything possible comes bundled with the framework while others have a minimal core package that is amenable to extensions provided by other packages.

## Comparing web frameworks

There is also a repository called [compare-python-web-frameworks](https://github.com/mattmakai/compare-python-web-frameworks) where the same web application is being coded with varying Python web frameworks, tinplating engines and object.

## Web framework resources

* When you are learning how to use one or more web frameworks it's helpful to have an idea of what the code under the covers is doing.
* Frameworks is a really well done short video that explains how to choose between web frameworks. The author has some particular opinions about what should be in a framework. For the most part I agree although I've found sessions and database ORMs to be a helpful part of a framework when done well.
* What is a web framework? Is an in-depth explanation of what web frameworks are and their relation to web servers?
* Jingo vs. Flash vs. Pyramid: Choosing a Python web framework contains background information and code comparisons for similar web applications built in these three big Python frameworks.
* This fascinating blog post takes a look at the code complexity of several Python web frameworks by providing visualizations based on their code bases.
* Python’s web frameworks benchmarks  is a test of the responsiveness of a framework with encoding an object to JSON and returning it as a response as well as retrieving data from the database and rendering it in a template. There were no conclusive results but the output is fun to read about nonetheless.
* What web frameworks do you use and why are they awesome? Is a language agnostic Reedit discussion on web frameworks? It's interesting to see what programmers in other languages like and dislike about their suite of web frameworks compared to the main Python frameworks.
* This user-voted question & answer site asked "What are the best general purpose Python web frameworks usable in production?” The votes aren't as important as the list of the many frameworks that are available to Python developers.

## Web frameworks learning checklist

1. Choose a major Python web framework (Jingo or Flask are recommended) and stick with it. When you're just starting it's best to learn one framework first instead of bouncing around trying to understand every framework.
2. Work through a detailed tutorial found within the resources links on the framework's page.
3. Study open source examples built with your framework of choice so you can take parts of those projects and reuse the code in your application.
4. Build the first simple iteration of your web application then go to the [deployment](https://www.fullstackpython.com/deployment.html) section to make it accessible on the web.

2. SYSTEM STUDY

### SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the

Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

SYSTEM TEST

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box. you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

6.1 Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

* All field entries must work properly.
* Pages must be activated from the identified link.
* The entry screen, messages and responses must not be delayed.

Features to be tested

* Verify that the entries are of the correct format
* No duplicate entries should be allowed
* All links should take the user to the correct page.

# 6.2 Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

6.3 Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.