CHAPTER-1

ABOUT THE COMPANY

1.1 History of the Organisation

Prinston Smart Engineers traces its roots back to 2004 in New Delhi and since then have never looked back. With Dozens of successful projects under our belt, they can proudly say that we are one of the most trusted Engineering, Maintenance &Training Services in Delhi.

For more than 13 years, Smart Engineers has offered a wide range of Maintenance & Services in Delhi, many other cities of India they strive to maintain the highest standards while exceeding client's expectations at all levels. They not only honor commitments, but are known for meeting tough deadlines while delivering nothing but the best. They aim to create a responsive client relationship that allows us to meet and even exceed the goals of each of our projects.

Smart Engineers is well known for its innovation and by collaborating successfully with customers, designers, sub-contractors, consultants as well as suppliers; they have been able to provide more specialized level of services. They are continuously evolving and understand the different aspects of delivering high value construction and complex projects with ease.

In 2020 Prinston Smart Engineers collaborated with Wedir-Tech Trading Contracting & Services W.L.L, Doha, Qatar for mutual benefit and better service.

1.2 Ownership

Prinston Smart Engineers is wholly owned by Mr. Asif Akhtar. He is the sole investor and there are no shareholders or shares as this organization is neither registered Indian national market nor Bombay exchange market.

1.3 Sector

The Prinston Smart Engineers are involved in a wide range of sectors as mentioned below:

1. Operation and Maintenance

2. Skill development and training

3. Interior design

4. Energy saving

1.4   Customers of the Company

1. Honeywell

2. Elf

3. Shipra group

4. Deepali designs

5. Ascenso management and consulting services pvt. Ltd

6. Eta

7. India Islamic cultural center

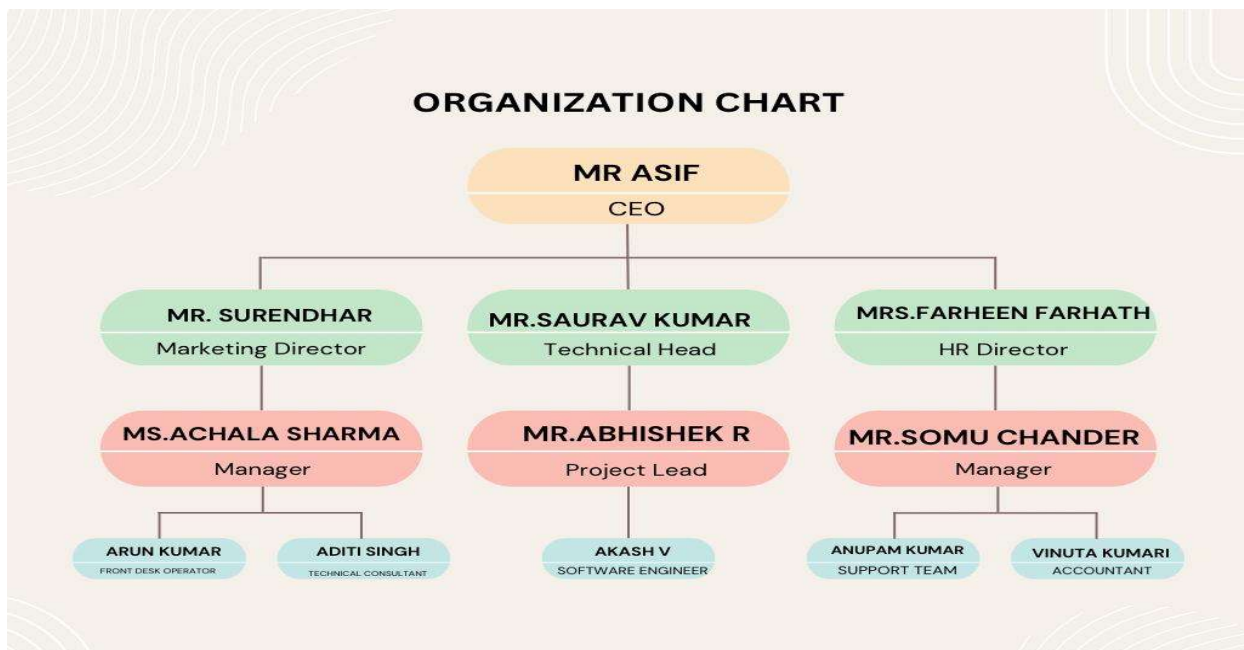## 1.5   Structure of the Company

Figure 1.1 Organization Chart

An organization structure shown in Figure 1.1 is used by the organization to provide an overall

hierarchy and reporting structure for each job within the organization, and defines how an organization will operate to meet its goals. For small businesses, one of several basic types of organizational structure can provide a platform to allow for future growth without making extensive changes to job responsibilities, the management structure or financial budgets for each area. Bureaucratic structures have many levels of management ranging from senior executives to regional managers, all the way to department store managers. Since there are many levels, decision-making authority has to pass through more layers than flatter organizations. A bureaucratic organization has rigid and tight procedures, policies and constraints.

Mr. Asif being the CEO of the company, manages the entire organization. Mr. Surendar being the Marketing director, manages and coordinates the activities between the departments. Mr. Saurav being the technical head guides Mr. Abhishek who is the project lead. Mr. Somu is the manager who executes all the managerial duties. Working under the manager we have Anupam and Vinuta who are the support team and accountant respectively, also Akash is a software engineer working under the project lead.

1.6   Services of the Company

The various services offered include:

Approaches:

1.   The company develops the best talent at every level of our business

2.   Provide best services to attract new client, retaining and expanding the relationship with the existing clients.

3.   Deliver leading quality and innovative services with maximum efficiency.

4.   Constantly update and adapt latest technology emerging in the control market.

5.   Managing risk while protecting our business values and brand.

6.   Thinking long term and act responsibly and strategically.

Values:

People-oriented: The company offers people the opportunity to accelerate morerapidly than is possible elsewhere. The company will continue to drive the greater dimension length to identify and recruit the very best person for every position. The company perceives the importance of individual creativity, but a team effort produces the best results.

Client-focused: The company are big enough to solve clients problems, yet substantialenough to sustain the loyalty, intimacy and culture that all treasure and which contributes greatly to company success. The client's interest always comes first. The experience shows that if the company treats there clients well, success will pursue.

Thecompany always seek aggressively to widen up, their client relationships but they will never put down competitors in this pursuit. Extensive Integrity in everything companydo Integrity and honesty are at the heart of business. Company expect people to maintain high ethical standards in everything they do, both in their work for the company and in their personal lives.

Innovative: The company assets are their people, technology, intellectual capital, culture of innovation and reputation. If any of this is ever lost, the last is the most difficult to regain. The company covet culture of innovation and as such the company stress creativity and imagination in all of its work. The company are and will continueto be thought leaders in chosen industries. The company take great honor in the professional quality of its work.

The company have an uncompromising determinationto achieve excellence in everything they do. Company would rather not pursue opportunities than deliver solutions that do not represent the highest level of quality.

CHAPTER-2

ABOUT THE DEPARTMENT

There are departments on various domain like:

1. Embedded systems and iot

2. Blockchain

3. Security and ethical hacking

4. Artificial intelligence and machine learning

5. Augmented and virtual reality

2.1 About the Working Department

Today's highly increasing competitiveness over the industry demands high quality and most consistent products with a competitive price. To address this challenge number of industries considering various new product designs and integrated manufacturing techniques in parallel with the use of automated devices. Automation takes a step further mechanization that uses a particular machinery mechanism aided human operators for performing a task. Mechanizationis the manual operation of a task using powered machinery that depends on human decision making. One of the remarkable and influential moves for getting the solutions of above-mentioned challenge is the industrial automation. Industrial Automation facilitates to increasethe product quality, reliability and production rate while reducing production and design cost by adopting new, innovative and integrated technologies and service

2.2 Objectives of the Department

1. Providing complete knowledge of ai and ml

2. Able to program, testing and commissioning of hardware

3. Able to program, test, and debug of software

2.3   Organisational Procedures

The organizational procedure include:

Analysis:

1. Pre-Project Analysis – Techno-Logics will perform a complete analysis.

Proposal:

1. Accurate quotes that include detailed preliminary schedules.

2. Comprehensive risk analysis.

 Design:

1. Designing a complete network structure, schematics.

2. Prepare detailed sequence of operation.

3. Detail a complete panel layout, wiring diagram, loop diagrams etc

Execution:

1. Detailed Scope of Work / Description of Operation.

2. Action item list.

Controlling/monitoring:

1. Timely project status updates.

Summary/closing:

1. Lessons learned.

CHAPTER 3

TASKS ASSIGNED AND PERFORMED

3.1 Technical Activity Performed in Company

As an Intern, I was assigned to work in the Machine Learning department. The Department was mainly focused on developing python project using appropriate tools and platform. As an internI was trained on selected libraries in jupyter notebook

3.2 Requirement Specification

System Requirements specify the hardware and software requirements that are necessary to begin with performing the given tasks accurately which are shown in Table 3.1 and Table 3.2.

3.2.1 Hardware Requirements

Processor	Intel Core i3 with clock speed of 2.4 GHz or above
RAM	2GB or above
Hard disk	40 GB or above
Input device	Keyboard or mouse or compatible pointing devices
Display	XGA (1024*768 pixels) or higher resolution monitor with 32-bit color settings
Miscellaneous	USB Interface, Power adapter, etc.

Table 3.1 Hardware Requirements

3.2.2 Software Requirements

Programming language	Python
Development environment	Visual Studio Code

Table 3.2 Software Requirements

3.3　Tasks Assigned and Performed

SL	Week	Tasks Performed
1	Week-1 (11th Aug 2023–16th Aug 2024)	1. Joining the Internship Program. 2. Getting to know about the Company and Mentors. 3. Understanding the basics of AI/ML.
2	Week-2 (17th Aug 2023 –23nd Aug 2024)	1. Learn about the usage of AI/ML in the real world and how various business problems are solved using it.
3	Week-3 (24th Aug 2023 –30th Sep 2023)	1. Studying Unsupervised Learning: K-Means Clustering, Hierarchical Clustering, D imensionality Reduction. 2. Handson project with supervised and uns upervised learning algorithms.
4	Week-4 (31st Sep 2021 – 7th Sep 2021)	1. Got assigned a Task –"Medical cost prediction" to be performed. 2. Started working on the allotted Project by t he Company. 3. Modify the project accordingly.

Table 3.3 Internship Timeline

3.3.1 Swot Analysis

A study undertaken by an organization to identify its internal strengths and weaknesses,as well as its external opportunities and threats as shown in the Figure 3.1.

Figure 3.1 SWOT Analysis

3.4 Artificial Intelligence

Swot Analysis

Strength:

1   The company enjoys good human resources.

2   The company is financially good.

3   The company has adopted the latest technology for operation.

4   The greatest strength is its highly skilled and committed employees.

5   Strong research and development.

Weakness:

1.   Sometimes Lack of resources which is on time at the project delivery in company.

2.   Lack of demand from Corporate.

3.   Lack of aggressive marketing and dependence on the market.

4.   Lack of financial resources

Opportunities:

1. Low-cost skilled labor for Graduate freshers provides an opportunity for the company for producing quality goods economically and efficiently.
2. Company follows same MNC rules as it leads to opportunities for branding it for CMM
3. level 5 company list.

Threats:

1. Threat from spurious component manufactures.

2. Technological inequality among foundries.

3. Industries heavily rely on non-renewable resource.

4. Increasing competition

5. Research and development initiatives

3.4.1 What is Artificial Intelligence?

In today's world, technology is growing very fast, and are getting in touch with different new technologies day by day. Here, one of the booming technologies of computer science is Artificial Intelligence which is ready to create a new revolution in the world by making intelligent machines. Artificial Intelligence is now all round us. It is currently working with a variety of subfields, ranging from general to specific, such as self-driving cars.

3.4.2 Advantages of Artificial Intelligence

High Accuracy with less errors:

AI machines or systems are prone to less errors andhigh accuracy as it takes decisions as per pre-experience or information.

High-Speed:

AI systems can be of very high-speed and fast decision making, becauseof that AI systems can beat a chess champion in the Chess game.

High reliability:

AI machines are highly reliable and can perform the same action multiple times with high accuracy.

Useful for risky areas:

AI machines can be helpful in situations such as defusing a bomb, exploring the ocean floor, where to employ a human can be risky.

Digital Assistant:

AI can be very useful to provide digital assistant to the users such asAI technology is currently used by various E-commerce websites to show the productsas per customer requirement.

3.5   Python

Python is a high-level, interpreted, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation.

Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly procedural), object-oriented and functional programming. It is often described as a "batteries included" language due to its comprehensivestandard library.

Python consistently ranks as one of the most popular programming languages. It is used by many organizations and companies. Pixar, Disney, Instagram and the developers of the LinuxKernel are among many of its high-profile users, which includes many developers of Free andOpen-source software.

3.5.1 Advantages of Python

Easy to Read, Learn and Write:

Python is a high-level programming language that has English-like syntax. This makesit easier to read and understand the code. Python is really easy to pick up and learn, that is whya lot of people recommend Python to beginners. You need less lines of code to perform the same task as compared to other major languages like C/C++ and Java.

Improved Productivity:

Python is a very productive language. Due to the simplicity of Python, developers can focus on solving the problem. They don't need to spend too much time in understanding the syntax or behavior of the programming language. You write less code and get more things done.

Interpreted Language:

Python is an interpreted language which means that Python directly executes the code line by line. In case of any error, it stops further execution and reports back the error which hasoccurred. Python shows only one error even if the program has multiple errors. This makes debugging easier.

Dynamically Typed:

Python doesn't know the type of variable until the code is run. It automatically assignsthe data type during execution. The programmer doesn't need to worry about declaring variables and their data types.

Free and Open-Source:

Python comes under the OSI approved open-source license. This makes it free to use and distribute. You can download the source code, modify it and even distribute your version of Python. This is useful for organizations that want to modify some specific behavior and usetheir version for development.

Vast Libraries Support:

The standard library of Python is huge, you can find almost all the functions needed foryour task. So, you don't have to depend on external libraries. But even if you do, a Python package manager (pip) makes things easier to import other great packages from the Python package index (PyPi). It consists of over 200,000 packages.

Portability:

In many languages like C/C++, there is a need to change code to run the program on different platforms. That is not the same in python. Python follows write your code once and run it

anywhere. However, should be careful not to include any system-dependent features.

3.6   Basic Training

The various modules that we learned during the training are: -

NumPy:

NumPy is a very popular python library for large multi-dimensional array and matrix processing, with the help of a large collection of high-level mathematical functions. It is very useful for fundamental scientific computations in Machine Learning.

T5:

The T5 (Text-To-Text Transfer Transformer) is a powerful deep learning model for natural language processing (NLP) developed by Google AI Language. It is based on the Transformer architecture and uses transfer learning to improve its performance on variousNLP tasks.

The t5 module is a Python library that provides an easy-to-use interface for working with the T5 model. It allows you to fine-tune the pre-trained T5 model on your own specific NLP task, such as text classification, question answering, summarization, translation, and more.

The T5 model is trained in a text-to-text setting, which means that it learns to map one textsequence to another text sequence. The t5 module also provides a pre-trained T5 model that can be used out-of-the-box forvarious NLP tasks.

Google-trans:

The googletrans module is a Python library that provides a simple interface for using the Google Translate API. It allows you to translate text from one language to another,detect the language of a given text, and get a list of supported languages.

The googletrans module is based on the Google Translate API, which is a paid serviceprovided by Google. However, the googletrans module provides a free and open-sourcealternative to the official API, making it a popular choice for small-scale projects and personal use. The googletrans module supports a wide range of languages, and its translations are generally accurate and reliable.

3.7   Machine Learning

3.7.1 What is Machine Learning?

Machine learning is an application of Artificial Intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

3.7.2 Need for Machine Learning

Machine learning is a greater tool to analyze the data, find hidden data patterns and relationships and extract the information from the data to enable the information driven decisions and provide insights. There is no human intervention needed (Automation), Handlingmulti-dimensional and multi-variety data where in it has s wide applications. A Machine Learning Model is the learned program that maps inputs to predictions. This can be set of weights for a linear model or for a neural network.

3.7.3 Machine Learning Algorithm

Logistic Regression:

Logistic regression, despite its name, is a classification algorithm rather than regressionalgorithm. Based on a given set of independent variables, it is used to estimate discretevalue (0 or 1, yes/no, true/false). It is also called logit or MaxEnt Classifier. Basically, it measures the relationship between the categorical dependent variable and one or more independent variables by estimating the probability of occurrence of an event using its logistics function.

Naive Bayes:

Naïve Bayes methods are a set of supervised learning algorithms based on applying Bayes' theorem with a strong assumption that all the predictors are independent to each other i.e. the presence of a feature in a class is independent to the presence of any other feature inthe same class. This is naïve assumption that is why these methods are called Naïve Bayesmethods.

Bayes theorem states the following relationship in order to find the posterior probability ofclass i.e. the probability of a label and some observed features, $P(Y | features)$.

Support Vector Machine:

Support Vector Machines (SVMs) are powerful yet flexible supervised machinelearning methods used for classification, regression, and, outliers' detection. SVMs are very efficient in high dimensional spaces and generally are used in classification problems.

Random Forest:

The Random Forest classifier creates a set of decision trees from a randomly selected subset of the training set. It is basically a set of decision trees from a randomly selected subsetof the training set and then It collects the votes from different decision trees to decide the final prediction. Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The basic idea behind thisis to combine multiple decision trees in determining the final output rather than relying on individual decision trees. Random Forest has multiple decision trees as base learning models.

Decision trees:

Decision trees are a versatile supervised learning algorithm used for both classification and regression tasks. They partition the feature space into regions and make predictions by following a tree-like structure of decision rules. Each internal node represents a "decision" based on a feature, and each leaf node represents the outcome or prediction.

Unsupervised Learning:

Unsupervised learning is a type of machine learning where the model learns patterns from input data without any explicit supervision or labeled responses. It's often used for tasks such as clustering, dimensionality reduction, and anomaly detection.

K-Means Clustering:

K-means clustering is a popular unsupervised learning algorithm used for clustering data points

into a pre-defined number of clusters. The algorithm works by iteratively assigning each data point to the nearest cluster centroid and then recalculating the centroids based on the mean of the data points assigned to each cluster. The process continues until the centroids no longer change significantly, or a specified number of iterations is reached.

Supervised Learning:

Train a supervised learning model (e.g., logistic regression, decision tree, or random forest classifier) to predict customer churn based on historical data. Later, we Split the data into training and testing sets, train the model on the training set, and evaluate its performance on the testing set using metrics like accuracy, precision, recall, and F1-score.

Linear Regression:

Linear regression is a supervised learning algorithm used for predicting the relationship between two continuous variables. It assumes a linear relationship between the independent variables (features) and the dependent variable (target). The goal is to find the best-fitting line that minimizes the error between predicted and actual values.

Dimensionality Reduction:

Dimensionality reduction techniques are used to reduce the number of features (dimensions) in a dataset while preserving its important information. Principal Component Analysis (PCA) is a widely used dimensionality reduction technique that projects the data onto a lower-dimensional subspace by finding the principal components that capture the maximum variance in the data.

3.8 Introduction of Project

3.8.1 Problem Statement

The project aims to develop a machine learning model for predicting the potability of water based on its chemical and physical properties. Leveraging a dataset encompassing features like pH levels, hardness, and chemical concentrations, the goal is to train and evaluate multiple classifiers including Logistic Regression, Support Vector Machine, and Random Forest, with the ultimate objective of ensuring access to safe drinking water through accurate potability prediction.

3.8.2 Implementation
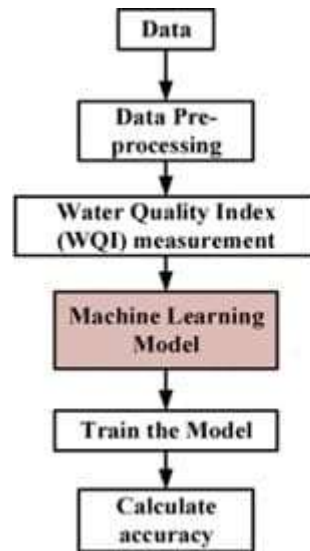
Figure 3.2 Architecture of the Project

Logistic Regression (LR):

Logistic Regression is a popular linear classification algorithm used for binary classification tasks. It models the probability that a given input belongs to a particular class.LR estimates the parameters of a logistic function to predict the probability of the binary outcome. In this implementation, Logistic Regression from sklearn linear model is used.

Support Vector Machine (SVM):

SVM is a versatile supervised learning algorithm capable of performing classification, regression, and outlier detection tasks. It works by finding the hyperplane that best separates classes in the feature space. SVM aims to maximize the margin between classes, making it robust to outliers. In this implementation, the standard SVC (Support Vector Classifier) from sklearn.svm is used for classification.

Random Forest (RF):

Random Forest is an ensemble learning method based on decision trees. It constructs multiple decision trees during training and outputs the mode of the classes (classification) or the mean prediction (regression) of the individual trees. RF is robust to overfitting and performs well on a variety of datasets. In this implementation, RandomForestClassifier from sklearn.ensemble is used with 500 decision trees.

GridSearchCV for Hyperparameter Tuning:

GridSearchCV is a technique used to tune hyperparameters by exhaustively searching through a specified parameter grid and selecting the combination that yields the best performance. It performs cross-validation to evaluate the performance of each parameter combination. In this implementation, GridSearchCV is used to fine-tune hyperparameters of the SVM model (SVC) such as the regularization parameter C, kernel type, degree of polynomial kernel, and gamma parameter.

Importing Libraries:

The necessary libraries such as seaborn, matplotlib, numpy, pandas, and modules from sklearn (Scikit-learn) are imported. Loading the Dataset: The dataset containing information about water quality is loaded using pd.read_csv() function. It seems like some columns are dropped (ph, Sulfate, Trihalomethanes) as they might be irrelevant for the predictive task or have too many missing values.

Data Preprocessing:

The dataset is preprocessed. Columns related to the target variable ('Potability') are separated as the target variable (y), and the remaining columns are considered as features (X). Standard scaling is applied to the features using StandardScaler() from sklearn.preprocessing.

Splitting Data: The dataset is split into training and testing sets using train_test_split() from sklearn.model_selection.

Model Training:

Logistic Regression model is trained using LogisticRegression() from sklearn.linear_model. SVM (Support Vector Machine) classifier is trained using SVC() from sklearn.svm.Random Forest classifier is trained using RandomForestClassifier() from sklearn.ensemble.

Model Evaluation:

The accuracy scores of each model on the test set are printed using accuracy_score() from sklearn.metrics.

Confusion Matrix Visualization:

Confusion matrices for each model are visualized using sns.heatmap() from seaborn and confusion_matrix() from sklearn.metrics.

3.8.3 Code Snippet

Code Snippet:

 # Import necessary libraries

import seaborn as sns

import matplotlib.pyplot as plt

import numpy as np

import pandas as pd

from sklearn.preprocessing import StandardScaler

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LogisticRegression

from sklearn.svm import SVC

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import accuracy_score, confusion_matrix

```
from sklearn.model_selection import GridSearchCV
```

Loading the Dataset

```
data=pd.read_csv('C:/Users/sathv/OneDrive/Desktop/Intrnship_Project/water_potrtabilty.csv'
)
```

Data Preprocessing:

```
data.drop(['ph', 'Sulfate', 'Trihalomethanes'], axis=1, inplace=True)

X = data.drop(['Potability'], axis=1)

y = data['Potability']

scaler = StandardScaler()

X_scaled = scaler.fit_transform(X)

x_train, x_test, y_train, y_test = train_test_split(X_scaled, y, train_size=0.85,
random_state=42)
```

Model Training:

```
logistic_regression = LogisticRegression(random_state=0).fit(x_train, y_train)

svm_classifier = SVC().fit(x_train, y_train)

random_forest = RandomForestClassifier(n_estimators=500, oob_score=True,
random_state=100).fit(x_train, y_train)
```

Model Evaluation:

```
print("Logistic Regression Accuracy:", logistic_regression.score(x_test, y_test))

print("SVM Accuracy:", accuracy_score(y_test, svm_classifier.predict(x_test)))

print("Random Forest Accuracy:", accuracy_score(y_test, random_forest.predict(x_test)))
```

Confusion Matrix for Logistic Regression:

```
plt.title("Logistic Regression Confusion Matrix")

sns.heatmap(confusion_matrix(y_test, logistic_regression.predict(x_test)), annot=True,
cbar=False)

plt.show()
```

 # Confusion Matrix for SVM:

plt.title("SVM Confusion Matrix")

sns.heatmap(confusion_matrix(y_test, svm_classifier.predict(x_test)), annot=True, cbar=False)

plt.show()

Hyperparameter Tuning for SVM:

param_grid = {'C': [1.2, 1.5, 2.2, 3.5, 3.2, 4.1], 'kernel': ['linear', 'poly', 'rbf', 'sigmoid'],

'degree': [1, 2, 4, 8, 10], 'gamma': ['scale', 'auto']}

gridsearch = GridSearchCV(SVC(), param_grid=param_grid, n_jobs=-1, verbose=4, cv=3)

gridsearch.fit(x_train, y_train)

Test Confusion Matrix:

y_pred = gridsearch.predict(x_test)

conf_matrix = confusion_matrix(y_true=y_test, y_pred=y_pred)

plt.figure(figsize=(15, 8))

sns.set(font_scale=1.4)

sns.heatmap(conf_matrix, annot=True, annot_kws={"size": 16}, cbar=False, linewidths=1)

plt.title("Test Confusion Matrix")

plt.xlabel("Predicted class")

plt.ylabel("Actual class")

plt.show()

3.8.4 Dataset

Figure 3.3 Dataset

3.8.5 Results

Figure 3.4 Dataframe and Dataset information

Figure 3.5 Model Evaluation

```
In [12]: # Confusion matrix for Logistic Regression
         plt.title("Logistic Regression Confusion Matrix")
         sns.heatmap(confusion_matrix(y_test, logistic_regression.predict(x_test)), annot=True, cbar=False)
         plt.show()
```

Logistic Regression Confusion Matrix

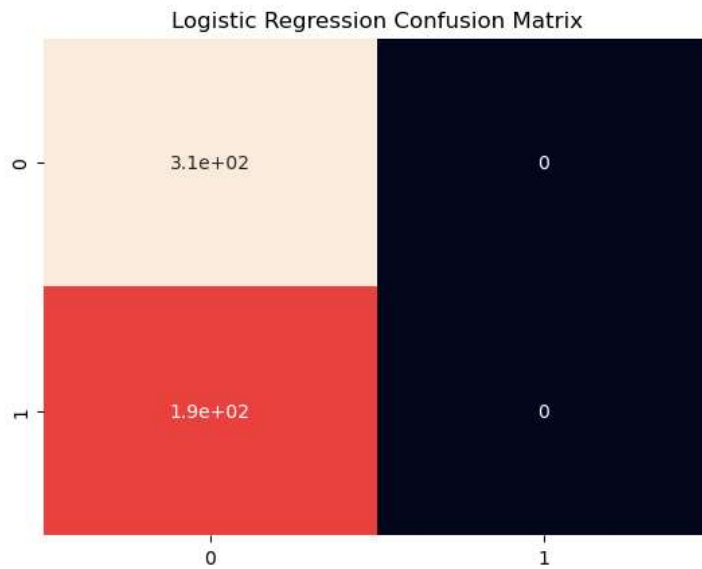|	0	1
0	3.1e+02	0
1	1.9e+02	0

Figure 3.6 Logistic Regression Confusion Matrix

```
In [13]: # Confusion matrix for SVM
         plt.title("SVM Confusion Matrix")
         sns.heatmap(confusion_matrix(y_test, svm_classifier.predict(x_test)), annot=True, cbar=False)
         plt.show()
```

SVM Confusion Matrix

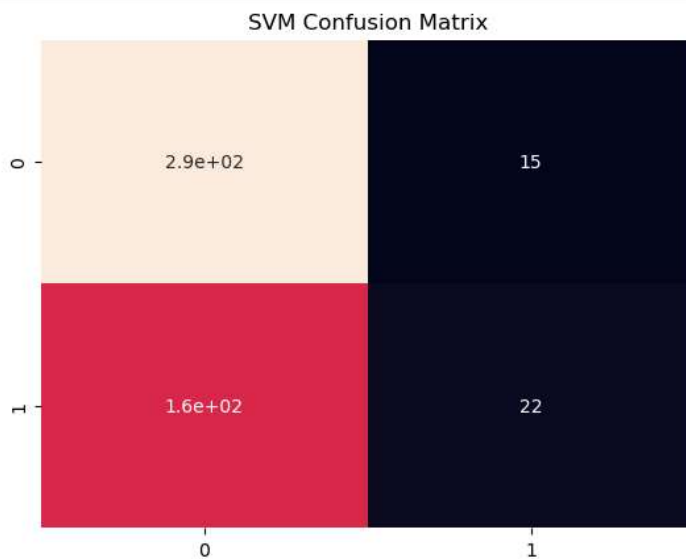|	0	1
0	2.9e+02	15
1	1.6e+02	22

Figure 3.7 SVM Confusion Matrix

```
In [7]:  # Confusion matrix for Random Forest
         plt.title("Random Forest Confusion Matrix")
         sns.heatmap(confusion_matrix(y_test, random_forest.predict(x_test)), annot=True, cbar=False)
         plt.show()
```
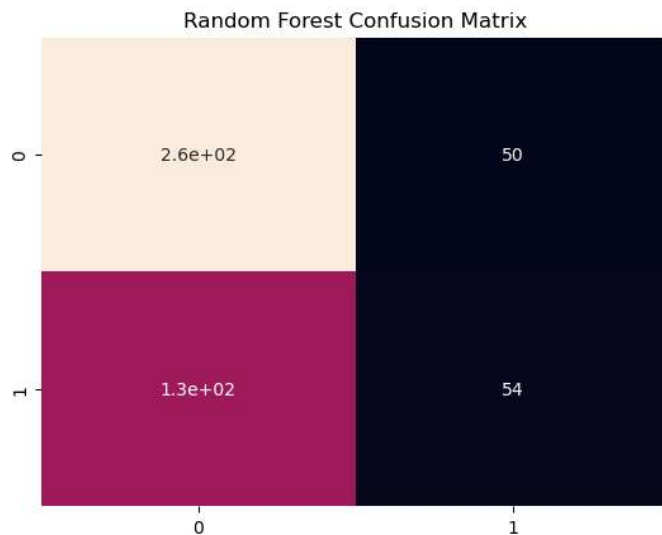

Figure 3.8 Random Forest Confusion Matrix

```
In [14]:  # Hyperparameter tuning for SVM
          # Hyperparameter grid with different values for SVM
          param_grid = {'C': [1.2, 1.5, 2.2, 3.5, 3.2, 4.1],'kernel': ['linear', 'poly', 'rbf', 'sigmoid'],'degree': [1, 2, 4, 8, 10],'gamm

          # Create GridSearchCV object
          gridsearch = GridSearchCV(SVC(), param_grid=param_grid, n_jobs=-1, verbose=4, cv=3)

          # Fit the grid search to the data
          gridsearch.fit(x_train, y_train)
```

```
Fitting 3 folds for each of 240 candidates, totalling 720 fits
```

```
Out[14]:  ▸ GridSearchCV
          ▾ estimator: SVC
          SVC()
                  ▾ SVC
                  SVC()
```

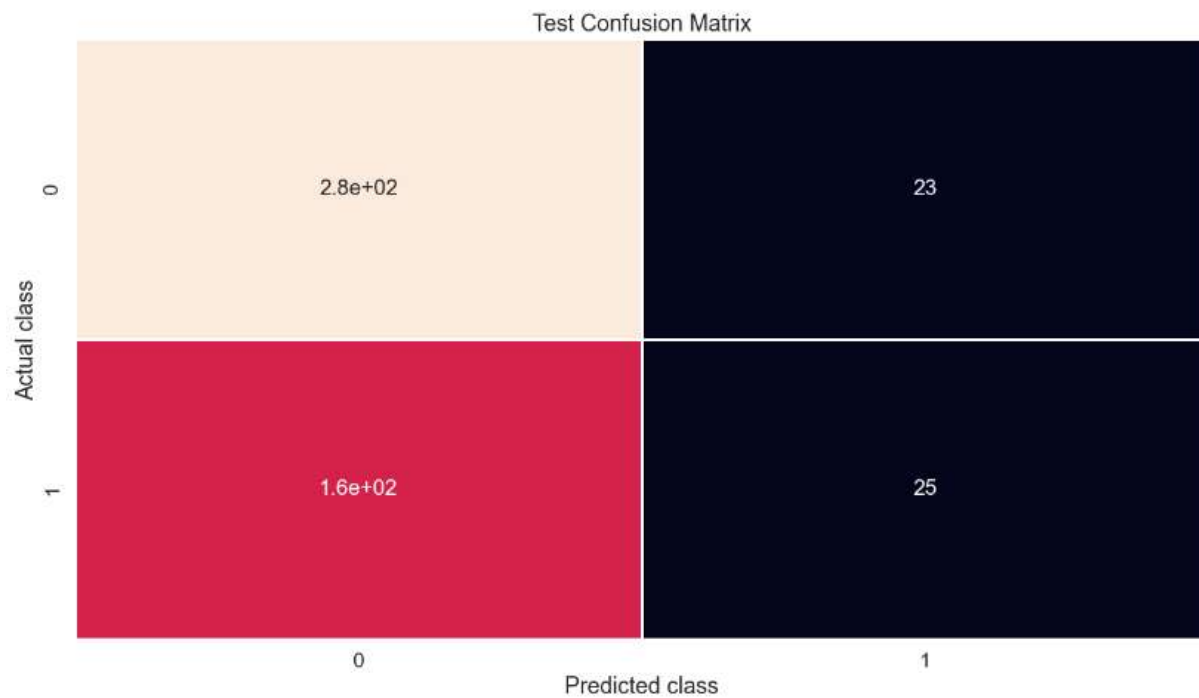Figure 3.9 Hyperparameter tuning for SVM

Figure 3.10 Test Confusion Matrix

CHAPTER 4

REFLECTION NOTES

My technical capabilities as well as soft skills such as interpersonal communication, writing skills, and presenting skills improved as a result of the internship. Internships enabled me to have a better awareness of industry norms and practices. It aided in the application of various technological abilities learned in the classroom to a real-world assignment. It also helped me to improve my technical knowledge.

An internship is a type of experiential learning that combines classroom knowledge and theory with hands-on experience and skill development in professional settings. Internships provide us with the opportunity to get practical experience. They are a great opportunity to obtain valuable work experience and expand your professional network. An internship can help you advance your career by giving you hands-on experience.

We were exposed to a variety of on-going technology, and there is always room for fresh ideas. It is far more crucial to be able to maintain and accomplish activities under pressure. Team meetings and sessions are beneficial in obtaining clarifications or answers to specific difficulties. The internship experience was beneficial since it strengthened both my technical and non-technical skills while also allowing me to gain practical experience in the business while working with seasoned experts who led me through the entire process.

After all, it allows us to demonstrate our talents, build work experience, and to understand institutional organization. Internships also provide an opportunity for us to get to grips with working – meeting deadlines and working in a team. Internships give us and the employer a chance to 'try before you buy' and clarify whether this type of work is really for us.

It's easy to characterize an internship as a "learning experience" or the opposite, but it takes a little more work to figure out what are the positive and negative elements of the experience.

4.1   Technical Outcomes

1. Have a good understanding of the fundamental issues and challenges of Machine-Learning and Artificial Intelligence.

2. Have an understanding of the strengths and weaknesses of many Machine Learning algorithms.

3. Appreciate the underlying mathematical relationships within and across Machine Learning algorithms and the paradigms of supervised and un- supervised learning.

4. Be able to design and implement various machine learning algorithms in a rangeof real-world applications.

5. Ability to integrate machine learning libraries and mathematical and statistical tools with modern technologies.

4.2   Non-Technical Outcomes

This section summarizes the non-technical aspects of the internship. This section summarizes the non-technical aspects of the internship. It includes communication skills, time management, adaptability, and personality development.

4.2.1 Communication Skills

The sharing of specific information with the goal of generating understanding betweenthe sender and the recipient is known as communication. One of the most critical components in the success of a project is effective communication. Verbal or spoken communication, as well as written communication, are examples of communication abilities. All formal communications are made through the internship project. We learned how to communicate by producing documents and submitting them to the guide with updates on the task we had completed. Meetings were also set up on Zoom to learn about the development of the project and to discuss future plans. The use of basic, straightforwardlanguage was really beneficial in aiding comprehension of the activities. Information is passed along with good communication. People understand you lot better if you can communicate properly, and you won't have to put in as much effort to make them grasp whatyou're saying. The mentor's daily face-to-face chats helped me acquire self-esteem.

4.2.2 Time Management

Time management was one of the non-technical skills I had to master and manage during my internship course, according to my availability. Time management is a valuable skill that everyone should learn. It refers to not just how one uses time to achieve a specific activity, but also how one manages time and works efficiently. When a task is assigned, the first step is to prioritize it. The urgent tasks should be completed first than the important work. Taking breaks in between work is not at all a waste of time. Time management may be aided by a range of skills, tools, and techniques used to manage time when accomplishing specific tasks, projects, and goals complying with a due date. Initially, time management referred to just business or work activities, but eventually, the term broadened to include personal activities as well.

4.2.3 Problem Solving

In Technologic corporate, problems of all kinds are sure to arise. It's inevitable. Though you can't predict the nature of the problems you'll encounter in the future, you can prepare by hiring employees who are comfortable solving problems. An employee who can't solve problems on their own will require you to essentially do it for them which will be a burden on your time, your resources, and your patience.

4.2.4 Personality Development

This internship programme has aided me in a variety of ways, including improvements in my attitude, thinking, behavior, and nature. This shift in my personality enabled me to connect more successfully with others, meet new people without trepidation, and tackle problems in the outside world more effectively and efficiently.

REFERNCES

[1]   Technologics-https://technologics.in/

[2]   Machinelearningmastery-https://machinelearningmastery.com

[3]   Kaggle- https://www.kaggle.com/

[4]   Github- https://github.com/

[5]   GeekForGeeks- https://www.geeksforgeeks.com/

[6]   TowardsDataScience- https://towardsdatascience.com/