

VISESVARAYA TECHNOLOGICAL UNIVERSITY
“JNANA SANGAMA”, BELAGAVI-590018



THE INTERNSHIP PROJECT REPORT ON
GLASS CLASSIFICATION

Submitted in partial fulfilment of the requirements

for the award of degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

By

YASHWANTH K [1KS18CS121]

Under the guidance of

Mrs. Beena K

Asst. Prof, Dept. Of CS&E

Mr. Prashanth H S

Asst. Prof, Dept. Of CS&E



Department of Computer Science & Engineering
K.S.INSTITUTE OF TECHNOLOGY

#14, Raghuvanahalli, Kanakapura Main Road, Bengaluru-560109

2021-2022

K.S.INSTITUTE OF TECHNOLOGY

#14, Raghuvanahalli, Kanakapura Main Road, Bengaluru-560109

Department of Computer Science & Engineering



CERTIFICATE

Certified that the Internship project entitled “**GLASS CLASSIFICATION**” carried out by Mr. **YASHWANTH K** bearing USN **1KS18CS121** bonafide student of **K.S Institute of Technology** in the partial fulfilment for the award of the Bachelor of Engineering in **Computer Science and Engineering** of the **Visvesvaraya Technological University, Belagavi** during the period of August 2021 to September 2021. The Internship Project Work Report has been approved as it satisfies the academic requirements in respect of internship project work prescribed for the Bachelor of Engineering degree.

Dr. Rekha B. Venkatapur
Prof & HOD, CS&E Department

Dr. Dilip Kumar
Principal, KSIT

Mrs. Beena K
Asst. Prof, Dept. Of CS&E

Mr. Prashanth H S
Asst. Prof, Dept. Of CS&E

ACKNOWLEDGEMENT

I take this opportunity to thank one and all involved in building this project.

Firstly, would like to thank the college for providing us an opportunity to work on the project.

First and foremost, I thank the management of K.S. Institute of Technology for providing all the resources required for the project.

I wish to acknowledge our sincere gratitude to our beloved Principal, Dr. Dilip Kumar for his encouragement and providing all facilities for the accomplishment of this project.

The project would not have been possible but for the unstinted support of our beloved Rekha.B. Venkatapur, professor & HOD, Dept of CS&E. I also, am highly grateful to guides, Mrs. Beena K

Mr. Prashanth H S Asst. Professors, Dept of CS&E who have been very generous in assisting and supporting, to do this project.

I extend my heartfelt gratitude to Mrs. Farheen Farhath and Ms. Achala of Princeton Smart Engineers, for their guidance, constant supervision as well as for providing us the necessary information regarding the project. Without their willing disposition, the spirit of accommodation, frankness, timely clarification and above all faith in me, this project could not have been completed in due time. Their readiness to discuss all important matters at work deserves special attention.

YASHWANT K
(1KS18CS121)

ABSTRACT

The Machine Learning field, which can be briefly defined as enabling computers make successful predictions using past experiences, has exhibited an impressive development recently with the help of the rapid increase in the storage capacity and processing power of computers.

A glass classification is a type of prediction where we have to give which type of glass by using the materials used by the glass. This may help in crime case, glass left in the crime spot can be used as evidence if model is correctly identified.

TABLE OF CONTENTS

CHAPTER 1	
INTRODUCTION	1
CHAPTER 2	
ABOUT THE COMPANY	2-5
2.1 Ownership.	
2.2 The Customers of the Company.	
2.3 Organization Chart.	
2.4 List of Functions Performed.	
CHAPTER 3	
SCOPE OF PROJECT	6
CHAPTER 4	
REQUIREMENTS	7
4.1 Hardware Requirements.	
4.2 Software Requirements.	
4.3 Libraries Required	
CHAPTER 5	
IMPLEMENTATION	8-14
5.1 Overview	
5.2 Software Used	
5.2.1 Anaconda Navigator	
5.3 Python	
5.3.1 Introduction	
5.3.2 What can Python do?	
5.3.3 Why Python?	
5.3.4 Advantages of Python	
5.4 Machine Learning	
5.4.1 What is Machine Learning?	
5.4.2 Phases of Machine Learning	
5.4.2.1 Data Acquisition	
5.4.2.2 Data Preparation	

5.4.2.3	Choosing and Training a Model	
5.4.2.4	Evaluation and Parameter Tuning	
5.4.2.5	Prediction	
5.4.3	Algorithm – K Nearest-Neighbor.	

5.5 Tasks Performed

CHAPTER 6

SNAPSHOTS 15-17

6.1	Loading Dataset
6.2	Overview
6.3	Data Preprocessing
6.4	Outcome

CHAPTER 7

REFLECTION NOTES 18-19

7.1	Learning from the internship program
7.2	Technical Outcomes
7.3	Non-technical Outcomes

CONCLUSION 20

REFERENCES 21

LIST OF FIGURES:

5.4.3 Phases of Machine Learning	12
5.4.2 Working of K-Nearest Neighbor	13
6.1 Loading Dataset	15
6.2 Overview	16
6.3 Data Preprocessing	17
6.4 Outcome	17

CHAPTER 1**INTRODUCTION**

Machine learning involves computer to get trained using a given data set, and use this training to predict the properties of a given new data. For example, we can train computer by feeding it 1000 images of cats and 1000 more images which are not of a cat, and tell each time to computer whether a picture is cat or not.

Then if we show the computer a new image, then from the above training, computer should be able to tell whether this new image is cat or not. Process of training and prediction involves use of specialized algorithms. We feed the training data to an algorithm, and the algorithm uses this training data to give predictions on a new test data. There are various machine learning algorithms like Decision trees, Naive Bayes, Random forest, Support vector machine, K-nearest neighbour, K-means clustering, etc.

Machine Learning is the art (and science) of enabling machines to learn things which are not explicitly programmed. Machine learning is one of the applications of python. Python is widely used general purpose, high level programming language. It was initially designed by Guido van Rossum in 1991 and developed by Python Software Foundation. It was mainly developed for emphasis on code readability, and its syntax allows programmers to express concepts in fewer lines of code.

The most recent major version of Python is Python 3, which we shall be using in this tutorial. Python can be used on a server to create web applications. It can be used along side software to create workflows. It can connect to database system and also read and modify files.

CHAPTER 2**ABOUT THE COMPANY**

COMPANY NAME: PRINSTON SMART ENGINEERS

CEO: MR. ASIF AKHTER.

CO- FOUNDER: MRS. FARHEEN FARHATH.

PRINSTON SMART ENGINEERS traces its roots back to 2004 in New Delhi. With Dozens of successful projects under its belt, it is one of the most trusted Engineering, Maintenance & Training Services in Delhi and also extends its services across India in various states such as Karnataka and Gujarat in locations such as Bangalore, Mysore and Jaipur.

In 2016 the company expanded its services in Skill development and Training program for engineering students in various engineering domains. India today produce 1.5 million Engineering graduates a year, it is however agreed by all that 75% of these graduates are unemployed, agreeing to the fact that there is need to provide Skill training to the Engineering students. In 2018 AICTE announced that an Internship is mandatory for engineering students to ensure that technical students get exposure to the Industrial environment, current technology relevant to their subject and opportunities to learn understand and sharpen real time technical and managerial skills. On request from colleges, the company began providing quality internship under the guidance of their expert's team to the students from various universities and colleges.

In 2020 Prinston Smart Engineers collaborated with Wedir-Tech Trading Contracting & Services W.L.L, Doha, Qatar with a MoU for their expertise in technical know-how and skill development.

2.1 OWNERSHIP

Prinston Smart Engineers is wholly owned company by Mr. Asif Akhtar. He is sole investor and there are no shareholders or shares since the company is neither listed on National Stock Exchange of India, nor on Bombay Stock Exchange.

THE SECTOR COMPANY OPERATES IN

PRINSTON SMART ENGINEERS is involved in a wide range of service sectors mentioned below:

1. Operation & Maintenance
2. Mechanical and Electrical Construction
3. Skill Development and training
4. Interior Designing
5. Energy Saving Solutions
6. MEP design & consultancy services

2.2 THE CUSTOMERS OF THE COMPANY

With respect to their Operations, Maintenance and MEP division, customers of the company include clients such as SHIRPA Group, Honeywell, C&S electricals and ITC Hotels. Along with these, the company has worked with several other organisations in providing interior designing and energy saving solutions.

From their Skill Development and training division's perspective, the customers belong to a wide variety of engineering students from over 50+ colleges based in and around Bangalore as well training corporate professionals as well.

The activities of this division mainly involve skilling and training of the students and working professionals in their respective domain of choice with the help of the expertise from the training department.

2.3 ORGANISATION CHART

Prinston Smart Engineers has 4 major divisions in the organisation:

1. Construction department: has 15 employees.
2. Maintenance department: has 15 employees.
3. Designing department: has 15 employees.
4. Training department: has 20 employees.

Altogether, Prinston Smart Engineers has employed over 65 employees of which 20 employees are situated in India and 45 in are based in Qatar.

2.4 LIST OF FUNCTIONS PERFORMED

1. Operation & Maintenance:

This involves Air Conditioning Equipment Maintenance, Refurbishment and renovation of existing system as well as Preventive maintenance.

2. Mechanical and electrical Construction:

The company provides services in both large and small projects pertaining to Mechanical and Electrical Construction. This includes Electrical Wiring, Plumbing, Fabrication, Welding and sprinkler systems.

3. Skill Development and training:

The company offers various courses and programs in reskilling and training engineering graduates. They have tailor made programs which are specifically designed to minimize the gap between the industrial demands and young engineers.

4. Interior designing:

The company consists of professional interior designers that are qualified by education, experience and examination to enhance the function and quality of interior spaces for the purpose of improving the quality of life, increasing productivity and protecting the health safety and welfare of the public.

CHAPTER 3**SCOPE OF THE PROJECT**

In this project I have been implemented all classification algorithms but to get best and approximate result, used Random forest algorithm and certain libraries for data exploration. Random forest algorithm is chosen based on the accuracy got. Random Forest algorithm is popular learning algorithm belongs to supervised learning technique.

It is based on the concept of ensemble learning. It contains the number of decision trees on various subsets of the dataset and takes the average to improve the predictive accuracy of the dataset. In this Greater number of trees leads to higher accuracy.

After training the model, It predict the type of glass depends on the user's input (Input is based on the material used in glass).

CHAPTER 4**REQUIREMENTS****9.1 Hardware Requirements**

1. Processor i3 and above.
2. Ram 4 GB and above.

4.2 Software Requirements

1. Anaconda navigator as an applications wrapper hub.
2. Jupyter Notebook as GUI interface for coding.

4.3 Libraries Required

1. Pandas
2. Numpy
3. Sklearn
4. Matplotlib

CHAPTER 5

IMPLEMENTATION

5.1 Overview

This project includes supervised machine learning model. It was built using Python. The necessary Software needed is listed below. It uses Books.csv dataset.

5.2 Software Used:

5.2.1 Anaconda Navigator:

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda® distribution that allows you to launch applications and easily manage conda packages, environments, and channels without using command-line commands. Navigator can search for packages on Anaconda.org or in a local Anaconda Repository. It is available for Windows, MacOS, and Linux.

Installing Anaconda Navigator

1. Download the Anaconda installer.
2. Double click the installer to launch.
3. Click Next.
4. Read the licensing terms and click “I Agree”.
5. Select an install for “Just Me” unless you’re installing for all users (which requires Windows Administrator privileges) and click Next.
6. Select a destination folder to install Anaconda and click the Next button.
7. Choose whether to add Anaconda to your PATH environment variable. We recommend not adding Anaconda to the PATH environment variable, since this can interfere with other software. Instead, use Anaconda software by opening Anaconda Navigator or the Anaconda Prompt from the Start Menu.
8. Choose whether to register Anaconda as your default Python. Unless you plan on installing and running multiple versions of Anaconda or multiple versions of Python, accept the default and leave this box checked.
9. Click the Install button. If you want to watch the packages Anaconda is installing, click Show Details.
10. Click the Next button.

11. Optional: To install PyCharm for Anaconda, click on the link to <https://www.anaconda.com/pycharm>. Or to install Anaconda without PyCharm, click the Next button.
12. After a successful installation you will see the “Thanks for installing Anaconda” dialog box
13. If you wish to read more about Anaconda.org and how to get started with Anaconda, check the boxes “Anaconda Individual Edition Tutorial” and “Learn more about Anaconda”. Click the Finish button.
14. Verify your installation.

5.3 PYTHON

5.3.1 INTRODUCTION

Python is a popular programming language. It was created by Guido van Rossum, and released in 1991.

- Python can be used on a server to create web applications.
- Python can be used alongside software to create workflows.
- Python can connect to database systems. It can also read and modify files.
- Python can be used to handle big data and perform complex mathematics.
- Python can be used for rapid prototyping, or for production-ready software development.

5.3.2 Why Python?

- Python works on different platforms (Windows, Mac, Linux, Raspberry Pi, etc).
- Python has a simple syntax similar to the English language.
- Python has syntax that allows developers to write programs with fewer lines than some other programming languages.
- Python runs on an interpreter system, meaning that code can be executed as soon as it is written. This means that prototyping can be very quick.

5.3.3 Advantages

- Extensive support libraries (Numpy for numerical calculations, Pandas for data analytics etc)
- Open source and community development
- Easy to learn
- User-friendly data structures
- High level language
- Dynamically Typed language

5.4 MACHINE LEARNING

5.4.1. 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 to learn for themselves.

5.4.2. PHASES OF MACHINE LEARNING

5.4.2.1. Data Acquisition

Data acquisition Machine learning needs two things to work, data (lots of it) and models. When acquiring the data, be sure to have enough features (aspect of data that can help for a prediction, like the surface of the house to predict its price) populated to train correctly your learning model.

5.4.2.2. Data Preparation

Data preparation involves five sub-processes to be followed. They are selection, cleansing, construction, integration, and formatting of data. In other words, all these steps comprise all the activities that must be performed for construction of the final data set.

5.4.2.3. Choosing and Training a Model

When designing a Machine Learning solution for a real-world problem, it's important to remember that the goal is not just to train a model to make accurate predictions on a representative dataset, but to train a model to make accurate predictions on data points seen in the field.

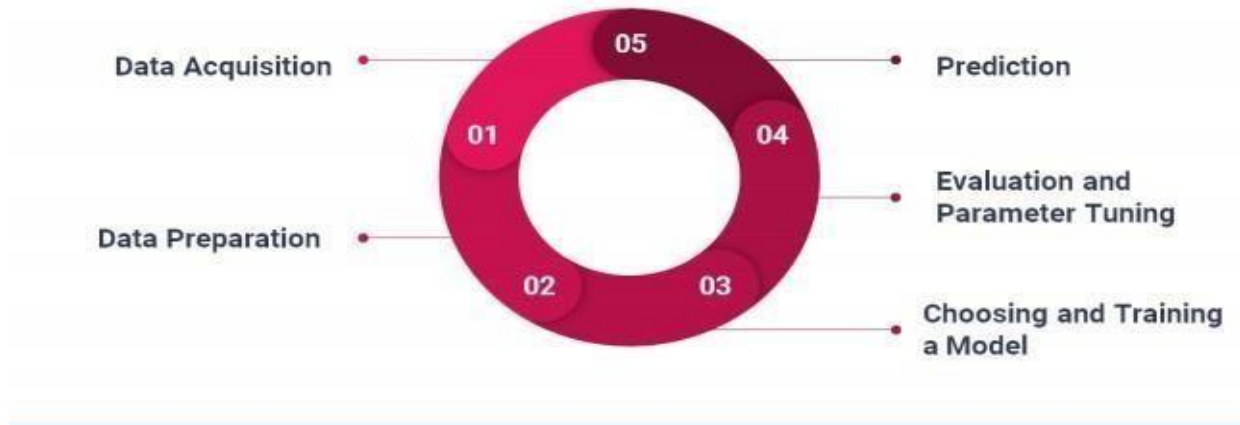
5.4.2.4. Evaluation and Parameter Tuning

In machine learning, the specific model you are using is the function and requires parameters in order to make a prediction on new data. Whether a model has a fixed or variable number of parameters determines whether it may be referred to as “parametric” or “nonparametric“. The weights in an artificial neural network.

5.4.2.5 Prediction

“Prediction” refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome. Prediction is at the heart of almost every scientific discipline, and the study of generalization (that is, prediction) from data is the central topic of machine learning and statistics, and more generally, data mining. Machine learning and statistical methods are used throughout the scientific world for their use in handling the "information overload" that characterizes our current digital age.

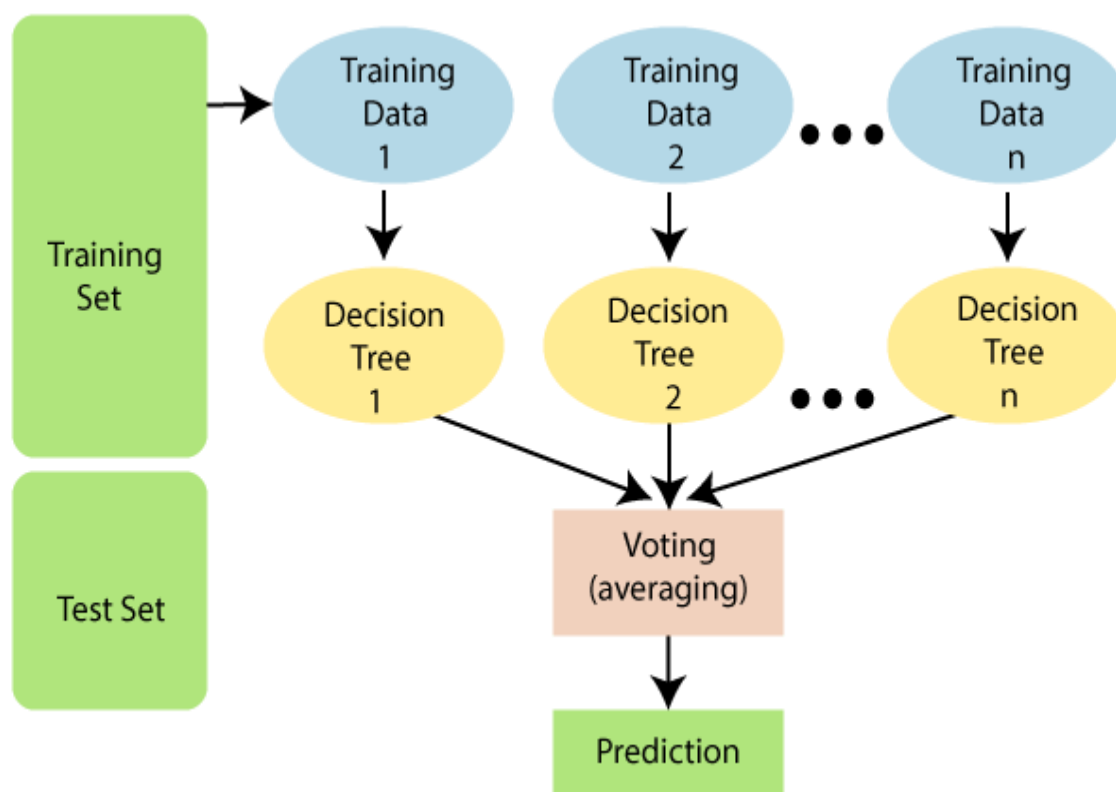
Phases of Machine Learning



5.4.2. Phases Of Machine Learning

5.4.3 Algorithm- Random Forest

- First, start with the selection of random samples from a given dataset.
- Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.
- In this step, voting will be performed for every predicted result.
- At last, select the most voted prediction result as the final prediction result.



5.4.3 Working of Random Forest Algorithm

5.5. Task Performed Weekly Work Plan:

- Week 1:

1. Introduction to installation of Anaconda.
2. Introduction to the Google Colab and Jupyter Notebook.
3. Understanding the importance and features of Google colab, Jupyter Notebook.
4. Basics of Python and its data structures.
5. Brief discussion on the Phases of Machine Learning

- Week 2

1. Installation of Libraries used in Anaconda.
2. Introduction to Pandas and Numpy.
3. Importing Datasets and understanding the features in various Datasets.

4. Usage of Matplotlib library.
5. Correlation.
6. Handling Missing values, Label encoder and adding synthetic columns.

- Week 3

1. Understanding of Splitting test and train sets.
2. Scaling Data – Normalisation and Standardisation.
3. Introduction to Regression.
4. Implementation and deep understanding of various algorithms such as Simple Linear Regression, Multiple Regression, Polynomial Regression, Decision trees, Random Forest, Support Vector Regression.
5. Brief about Kernel
6. Comparison of all the models.

- Week 4

1. Introduction to Classification.
2. Implementation and deep understanding of various algorithms such as Logistic Regression, K-Nearest Neighbours, Decision Trees, Random Forest, Support Vector Machine.
3. Understanding Confusion matrix and accuracy score.
4. Comparison of all the models.
5. Projects were assigned.
6. Working on Glass Classification.

CHAPTER 6

SNAPSHOTS

6.1 Loading Dataset

```
dataset = pd.read_csv('/content/drive/MyDrive/Projects/GlassDetails.csv')
dataset
```

	RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
0	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0.00	0.0	1
1	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0.00	0.0	1
2	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0.00	0.0	1
3	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0.00	0.0	1
4	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0.00	0.0	1
...
209	1.51623	14.14	0.00	2.88	72.61	0.08	9.18	1.06	0.0	7
210	1.51685	14.92	0.00	1.99	73.06	0.00	8.40	1.59	0.0	7
211	1.52065	14.36	0.00	2.02	73.42	0.00	8.44	1.64	0.0	7
212	1.51651	14.38	0.00	1.94	73.61	0.00	8.48	1.57	0.0	7
213	1.51711	14.23	0.00	2.08	73.36	0.00	8.62	1.67	0.0	7

6.1 Loading Dataset

6.2 Overview of the Dataset

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 214 entries, 0 to 213
Data columns (total 10 columns):
#   Column  Non-Null Count  Dtype
---  -
0    RI      214 non-null    float64
1    Na      214 non-null    float64
2    Mg      214 non-null    float64
3    Al      214 non-null    float64
4    Si      214 non-null    float64
5    K       214 non-null    float64
6    Ca      214 non-null    float64
7    Ba      214 non-null    float64
8    Fe      214 non-null    float64
9    Type    214 non-null    int64
```

6.3 Data preprocessing

6.3.1

```
X = dataset.iloc[:, :-1].values
y = dataset.iloc[:, -1].values
```

6.3.1 Separating features and target

6.3.2

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.25)
```

6.3.2 Splitting train and test data

6.3.3

```
sc = StandardScaler()
```

```
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

```
X_train
```

```
array([[ 0.06013163,  0.26826047,  0.80899853, ..., -0.41168736,
        -0.34472572,  2.63136848],
       [ 0.28951658,  0.75338818,  0.59890432, ..., -0.42616269,
        -0.34472572, -0.6119167 ],
       [-2.01416371, -0.52007207,  0.51921342, ..., -0.38997437,
        -0.34472572,  2.53001582],
       ...,
       [-0.4608999 , -0.90817424,  0.09178038, ..., -0.25245876,
        -0.34472572, -0.6119167 ],
       [-0.2020226 , -1.04158437,  0.55543656, ..., -0.25245876,
        -0.34472572, -0.6119167 ],
       [ 1.59045694,  0.02569661,  0.42503326, ..., -0.04256651,
        -0.34472572, -0.6119167 ]])
```

6.3.3 Standardizing data

6.4 Outcome

```
print("Enter values RI: refractive index Na: Sodium Mg: Magnesium Al: Aluminum Si: Silicon K: Potassium Ca: Calcium Ba: Barium Fe: Iron")
print("\n For Example [1.52101, 13.64, 4.49, 1.1, 71.78, 0.06, 8.75, 0,0]\n ")
RI = input('RI:')
Na = input('Na:')
Mg = input('Mg:')
Al = input('Al:')
Si = input('Si:')
K = input('K:')
Ca = input('Ca:')
Ba = input('Ba:')
Fe = input('Fe:')
```

Enter values RI: refractive index Na: Sodium Mg: Magnesium Al: Aluminum Si: Silicon K: Potassium Ca: Calcium Ba: Barium Fe: Iron

For Example [1.52101, 13.64, 4.49, 1.1, 71.78, 0.06, 8.75, 0,0]

RI:1.5
Na:13.64
Mg:4.49
Al:1.1
Si:71.7
K:0
Ca:8.7
Ba:0
Fe:0

```
rpred=rcla.predict(sc.transform([[RI,Na,Mg,Al,Si,K,Ca,Ba,Fe]]))
values(rpred[0])
```

'building windows float processed'

CHAPTER 7**REFLECTION NOTES****7.1 Learning from the internship program:**

During the internship period, I had hands-on experience with many real-time projects. By working with numerous datasets, I got deeper understanding on how to understand, analyze and handle data. This internship has provided me a great experience, lessons and the tools that I will need to get a gig in the future. Moreover, it has given me the opportunity to grow and learn fully on how these concepts can be used to solve real-time problem statements.

Work Experience:

Whereas Datasets were used to understand and compare all the algorithms during the training phase. On completing the training, I was in the position to work on this interesting project individually.

Responsibility and keeping commitments:

I had to stay focused and committed to attend all the training sessions. In addition, there was a deadline within which the project along with report and presentation had to be submitted. By this, I had to work on the project sincerely with the time frame in mind.

7.2 Technical Outcome

- Working with Anaconda Navigator
- Working on different kinds of Dataset
- Various regression and classification algorithm
- Data Analysis
- Hands on experience with projects

7.3 Non-Technical Outcome

- Verbal and written Communication

Communication in a formal way while interacting with the trainer. Writing skills improved through making the report and presentation.

- Problem Solving skill

Introduced to real-life work problems and worked on regular tasks. Problem solving skill has improved by working on the challenging project as well. Handling and solving all errors was experienced.

- Work Ethics

Attending training sessions regularly and completing the projects and other tasks within the stipulated time without making excuses. Punctuality and Time Management was also enhanced.

CONCLUSION

Machine learning is quickly growing field in computer science. It has applications in nearly every other field of study and is already being implemented commercially because machine learning can solve problems too difficult or time consuming for humans to solve.

As a result, we have studied the future of Machine Learning. Also, study algorithms of machine learning. Along with we have studied its application which will help you to deal with real life. Common methods and popular approaches used in the field, suitable machine learning programming languages, and also covered some things to keep in mind in terms of unconscious biases being replicated in algorithms.

This has set a direction and a route which will definitely boost my career. I worked on this challenging project independently. Moreover, it has made me take quite a few decisions alone such as deciding the dataset and the approach that this project requires.

RESOURCES

- Training handouts
 - <https://www.kaggle.com>