# ABSTRACT

The report presents the three tasks completed during summer internship at INTERNSHALA which are listed below:

1. Understand the problem objective & business implication

2. Understanding the data & build the model

3. Evaluation of the model

All these tasks have been completed successfully and results were according to expectations. All the tasks were need very systematic approach, starting from the behavior of the data to the application of the algorithm and till evaluation of the model. The most challenging task was the domain knowledge, to understand the behavior of the data. Once the data has been prepared, we applied statistical algorithm for model building. It is one of the major area and really need very fundamental and conceptual knowledge of Advanced Statistics.

Surbhi

**DECLARATION**

I hereby declare that that the work reported in the report entitled **“Machine Learning”** submitted to the Department of Electronics & Communication Engineering, **ABESIT GHAZIABAD** is an authentic record of my work. I have not submitted this work elsewhere for any other degree. I am fully responsible for the contents of my report.

SURBHI

Department of Electronics and communication Engineering

ABES Institute of Technology

December 2020

**ACKONOWLEDGEMENT**

# Simply put, I could not have done this work without the lots of help I received cheerfully from INTERNSHALA TEAM. The work culture in Internshala Training Course really motivates. Everybody is such a friendly and cheerful companion here that work stress is never comes in way. I would specially like to thank Mr. Sarvesh Agrawal for proving the nice ideas to work upon. Not only did they advised about my project but listening to their discussions in Discussion Forum have evoked a good interest in Image analysis. I am also highly indebted to all Internshala team, who seemed to have solutions to all my problems.

***SURBHI***

**CERTIFICATE**

**TABLE OF CONTENTS**

**Page No.**

**ABSTRACT ii**

**DECLARATION iii**

**ACKNOWLEDGEMENT iv**

**CERTIFICATE(S) v**

**TABLE OF CONTENTS vi**

**LIST OF FIGURES viii**

**LIST OF TABLES ix**

**CHAPTER-1: AN INTRODUCTION TO MACHINE LEARNING 10**

**1.1: MACHINE LEARNING 11**

**1.2: DATA IN MACHINE LEARNING 11**

**CHAPTER-2: BEST PYTHON LIBRARIES FOR MACHINE LEARNING 13**

**2.1: NUMPY 13**

**2.2: SCIPY 13**

**2.3: SCIKIT-LEARN 14**

**2.4: THEANO 14**

**2.5: TENSOR FLOW 14**

**2.6: KERAS 14**

**2.7: PYTORCH 14**

**2.8: PANDAS 15**

**2.9: MATPLOTLIB 15**

**CHAPTER-3: CLASSIFICATION OF MACHINE LEARNING 16**

**3.1: SUPERVISED & UNSUPERVISED LEARNING 16**

**3.2: HYBRID MACHINE LEARNING 16**

**3.3: REINFORCEMENT LEARNING 17**

**CHAPTER-4: REGRESSION ANALYSIS IN MACHINE LEARNING 19**

**4.1: TERMINOLOGIES 19**

**4.2: WHY DO WE USE REGRESSION ANALYSIS? 20**

**4.3: TYPES OF REGRESSION 20**

**4.4: LINEAR REGRESSION 21**

**4.5: LOGISTIC REGRESSION 22**

**4.6: POLYNOMIAL REGRESSION 24**

**4.7: SUPPORT VECTOR REGRESSION 24**

**CHAPTER-5: DECISION TREE REGRESSION 26**

**5.1: SOME ASSUMPTION OF DECISION TREE REGRESSION 26**

**CHAPTER-6: RANDOM FOREST REGRESSION 28**

**6.1: ADVANTAGES OF RANDOM FOREST 29**

**CHAPTER-7: LINEAR REGRESSION IN MACHINE LEARNING 30**

**7.1: TYPES OF LINEAR REGRESSION 31**

**7.2: LINEAR REGRESSION LINE 31**

**7.3: FINDING THE BEST FIT LINE 32**

**7.4: COST FUNCTION 32**

**7.5: GRADIENT DESCENT 33**

**7.6: MODEL PERFORMANCE 33**

**7.7: ASSUMPTION OF LINEAR REGRESSION 34**

**CHAPTER-8: CLASSIFICATION ALGORITHM IN ML 35**

**8.1: WHAT IS CLASSFICATION ALGORITHM? 35**

**8.2: LEARNERS IN CLASSIFICATION PROBEMS 36**

**8.3: TYPES OF CLASSIFICATION ALGORITHMS 37**

**8.4: EVALUATING A CLASSIFICATION MODEL 37**

**8.5: USE CASES OF CLASSIFICATION ALGORITHM 38**

**CHAPTER-9: LOGISTIC REGRESSION IN ML 39**

**9.1: LOGISTIC (SIGMOID) FUNCTION 40**

**9.2: ASSUMPTION FOR LOGISTIC REGRESSION 40**

**9.3: LOGISTIC REGRESSION EQUATION 40**

**CHAPTER-10: K-NEAREST NEIGHBOUR (K-NN) ALGORITHM 41**

**10.1: WHY WE NEED K-NN ALGORITHM? 42**

**10.2: HOW DOES K-NN WORK? 42**

**10.3: HOW TO SELECT THE VALUE OF K IN K-NN 44**

**10.4: ADVANTEGES AND DISADVANTAGES OF K-NN 44**

**PROJECT 45**

**REFERENCES 53**

**LIST OF FIGURES**

**Figure Caption Page No. Number**

3.1 The HML (Hybrid Machine Learning) work flow. 17

4.1 **Types of Regression.** 21

4.2 **Linear Relationship between two variables. 22**

4.3 **Sigmoid function curve. 23**

4.4 **Polynomial Regression data points. 24**

**4.5 Support vector data points. 25**

**5.1 Decision tree. 26**

**5.2 Decision tree example. 27**

**6.1 Random Forest Prediction 28**

**7.1 Linear Regression relationship between variables. 30**

7.2 **Positive linear Relationship. 31**

**7.3 Negative Linear Relationship. 32**

8.1 **Classification Algorithm graph of example. 36**

9.1 **Logistic Linear Function. 40**

**10.1 K-NN Classifier. 41**

**10.2 Need of K-NN. 42**

**10.3 How to put new data points in K-NN. 43**

**10.4** Euclidean distance between the data points. 43

10.5 Calculating Nearest Neighbors **using** Euclidean distance. 44

**LIST OF TABLES**

**Figure Caption Page No. Number**

3.1 Difference between Reinforcement learning and Supervised learning 17

8.1 Prediction of Confusion Matrix. 38