# MACHINE LEARNING

(Human Activity Prediction Using Smartphone Data)

*Summer Internship Report Submitted in partial fulfillment of the requirement for undergraduate degree of*

## Bachelor of Technology

In

## **Computer Science and Engineering**

By

## **Gopi Bollineni** **221710307029**

*Under the Guidance of*

**XXXX**

Assistant Professor



Department Of Computer Science and Engineering

GITAM School of Technology

GITAM (Deemed to be University) Hyderabad-502329

i

# DECLARATION

I submit this industrial training work entitled **“HUMAN ACTIVITY PREDICTION USING SMARTPHONE DATA**” to GITAM (Deemed To Be University), Hyderabad in partial fulfillment of the requirements for the award of the degree of “**Bachelor of Technology**” in “**Computer Science and Engineering**”. I declare that it was carried out independently by me under the guidance of **Mr. XXXX**, XXXX, GITAM (Deemed To Be University), Hyderabad, India.

The results embodied in this report have not been submitted to any other University or Institute for the award of any degree or diploma.

Place: HYDERABAD Gopi Bollineni

Date: XXXX 221710307029

ii



**GITAM (DEEMED TO BE UNIVERSITY)**

Hyderabad-502329, India

Dated:XXXXXX

**CERTIFICATE**

This is to certify that the Industrial Training Report entitled **“HUMAN ACTIVITY PREDICTION USING SMARTPHONE DATA”** is being submitted by Gopi Bollineni (221710307029) in partial fulfillment of the requirement for the award of **Bachelor of Technology in Computer Science and Engineering** at GITAM (Deemed To Be University), Hyderabad during the academic year 2018-19

It is faithful record work carried out by him at the **Computer Science and Engineering Department**, GITAM University Hyderabad Campus under my guidance and supervision.

### XXXX Dr. S. Phani Kumar

Assistant Professor Professor and HOD

Department of CSE Department of CSE

iii

**ACKNOWLEDGEMENT**

Apart from my effort, the success of this internship largely depends on the encouragement and guidance of many others. I take this opportunity to express my gratitude to the people who have helped me in the successful competition of this internship.

I would like to thank respected **Dr. N. Siva Prasad,** Pro Vice Chancellor, GITAM Hyderabad and **Dr. XXXX,** Principal, GITAM Hyderabad

I would like to thank respected **Dr. S. Phani Kumar,** Head of the Department of Computer Science and Engineering for giving me such a wonderful opportunity to expand my knowledge for my own branch and giving me guidelines to present a internship report. It helped me a lot to realize of what we study for.

I would like to thank the respected faculties **Mr. XXXX** who helped me to make this internship a successful accomplishment.

I would also like to thank my friends who helped me to make my work more organized and well-stacked till the end.

Gopi Bollineni 221710307029

iv

## ABSTRACT

V

**Table of Contents**

1. [MACHINE LEARNING 9](#_bookmark0)
   1. [Introduction: 9](#_bookmark1)
   2. [Importance of Machine Learning: 9](#_bookmark2)
   3. [Uses of Machine Learning: 10](#_bookmark3)
   4. [Types of Machine Learning: 13](#_bookmark4)
      1. [Supervised Learning: 14](#_bookmark5)
      2. [Unsupervised Learning: 15](#_bookmark6)
      3. [Reinforcement Learning: 15](#_bookmark7)
2. [DEEP LEARNING 16](#_bookmark8)
   * 1. Deep Learning and It’s Importance: 16
   1. [Uses of Deep Learning: 17](#_bookmark10)
   2. [Relation between Data Mining, Machine Learning and Deep Learning: 19](#_bookmark11)

[3.PYTHON 20](#_bookmark12)

* 1. [Introduction: 20](#_bookmark13)
  2. [Setup of Python: 21](#_bookmark14)
     1. [Installation(using python IDLE): 21](#_bookmark15)
     2. [Python Installation using Anaconda: 22](#_bookmark16)
  3. [Features: 23](#_bookmark17)
  4. [Variable Types 24](#_bookmark18)
     1. [Python Numbers: 25](#_bookmark19)
     2. [Python Strings: 25](#_bookmark20)
     3. [Python lists: 25](#_bookmark21)
     4. [python tuples: 26](#_bookmark22)
     5. [python Dictionary: 26](#_bookmark23)
  5. [Functions: 26](#_bookmark24)
     1. [Defining a Function: 26](#_bookmark25)
     2. [Calling a Function: 26](#_bookmark26)
  6. [OOPs Concepts: 27](#_bookmark27)
     1. [Class: 27](#_bookmark28)

1. [Mobile Price Classification 28](#_bookmark29)
   1. [Project Requirements: 28](#_bookmark30)
      1. [Packages used: 28](#_bookmark31)
      2. [Versions of the packages: 29](#_bookmark32)
      3. [Algorithms used: 30](#_bookmark33)
   2. [Problem Statement: 30](#_bookmark34)
   3. [Dataset Description: 30](#_bookmark35)
   4. [Objective of the Case Study: 31](#_bookmark36)
2. [DATA PREPROCESSING/FEATURE ENGINEERING AND EDA 32](#_bookmark37)
   1. [Statistical Analysis: 32](#_bookmark38)
   2. [Data Type Conversions: 33](#_bookmark39)
   3. [Detection of Outliers: 34](#_bookmark40)
   4. [Handling Missing Values: 35](#_bookmark41)
   5. [Encoding Categorical Data: 36](#_bookmark42)
   6. [Generating Plots: 36](#_bookmark43)
      1. [Visualize the data between Target and the Features: 36](#_bookmark44)

[5.6.1. Visualize the data between all the Features: 38](#_bookmark45)

1. [FEATURE SELECTION 41](#_bookmark46)
   1. [Select relevant features for the analysis: 41](#_bookmark47)
   2. [Drop irrelevant features: 42](#_bookmark48)
   3. [Train-Test-Split: 42](#_bookmark49)
   4. [Feature Scaling: 44](#_bookmark50)
2. [MODEL BUILDING AND EVALUATION 45](#_bookmark51)
   1. [Brief about the algorithms used: 45](#_bookmark52)
   2. [Train the Models: 48](#_bookmark53)
   3. [Make Predictions: 49](#_bookmark54)
   4. [Validate the Models: 50](#_bookmark55)
   5. [Parameter Tuning: 52](#_bookmark56)

[i. Grid search: 53](#_bookmark57)

[iii.Bayesian optimization: 54](#_bookmark58)

* 1. [Predictions from raw data: 55](#_bookmark59)

1. [Conclusion 58](#_bookmark60)
2. [REFERENCES 60](#_bookmark61)

**List of Figures**

FIG 1.2 USAGE OF MACHINE LEARNING IN DIFFERENT FIELDS **9**FIG 1.3 USES OF MACHINE

LEARNING **13**FIG 1.4 TYPES OF ML **13**FIG 2.1 DEEP NEURAL NETWORK **16**FIG 2.2 THE DEEP LEARNING PROCESS **17**FIG 2.3 RELATION BETWEEN DM,ML,DL **18**FIG 2.3 PROCESS IN MACHINE LEARNING AND DEEP LEARNING **18**FIGURE 3.2.1 : PYTHON DOWNLOAD **20**FIG

3.2.1.1 PYTHON INSTALLATION **21**FIG 3.2.1.2 IDLE **21**FIG 3.2.2.1 AFTER INSTALLATION

**22**FIG 3.2.2.2 JUPYTER NOTEBOOK **22**FIG 3.6.1 CLASS DEFINING **26**FIG 3.6.1.1

EXAMPLE OF CLASS **26**FIG 4.1.1 PACKAGES **28**FIG 4.1.2 VERSIONS**28**FIG 5.1 LOADING DATA SET **30**FIG 5.1.1 STATISTICAL DATA **31**FIG 5.2 DATATYPES**32**FIG 5.3 DETECTION OF OUTLIER USING BOXPLOT **33**FIG 5.5 NUMERICAL IN THE DATASET **34**FIG 5.6.1 CORRELATION **34**FIG 5.6.1.1 CO-RELATION GRAPH **35**FIG 5.6.1.2 RAM VS PRICE **35**FIG 5.6.1.3 BLUETOOTH,WIFI **36**FIG 5.6.2.1 3G SUPPORT PHONES **36**FIG 5.6.2.2 DUAL SIMS **37**FIG

5.6.2.3 BATTERY POWER **37**FIG 5.6.2.4 4G SUPPORT **38**FIG 6.1 TOP FEATURES IN DATASET

**39**FIG 6.2 CORRELATION BETWEEN ATTRIBUTES **40**FIG 6.4.2 SCALED X TRAIN **42**FIG

6.4.3 SCALED X TEST **42**FIG 7.1.1 LOGISTIC REGRESSION FOR TRAIN DATA **43**FIG 7.1.2

IMPORTING DECISION PACKAGES **44**FIG 7.1.3 RANDOM FORESTS **45**FIG 7.2.1 TRAINING DATA **46**FIG 7.3.2 PREDICTING TEST IN LOGISTIC REGRESSION **46**FIG 7.3.3 PREDICTING TEST IN DECISION TREE **47**FIG 7.3.4 PREDICTING TEST IN RANDOM FOREST **47**FIG

7.4.1 TEST AND TRAIN ACCURACY IN LOGISTIC REGRESSION **48**FIG 7.4.2 TEST VS TRAIN IN LOGISTIC REGRESSION **48**FIG 7.4.3 TEST AND TRAIN ACCURACY IN DECISION TREE **48**FIG 7.4.4 TEST VS TRAIN IN DECISION TREE **49**FIG 7.4.5 TEST ACCURACY IN RANDOM FOREST **49**FIG

7.4.6 TEST VS TRAIN IN RANDOM FOREST **49**FIG 7.5.1 GRID SEARCH CV **51**FIG 7.5.2

ACCURACY OF TEST USING GRID SEARCH CV I.E ACCURACY IS INCREASED IN DECISION TREE **52**FIG

7.6.1 BEST ACCURACY MODEL AMONG ALGORITHMS **52**FIG 7.6.2 LOADING TEST DATASET **53**FIG

7.6.3 PREDICTING TEST DATASET USING RANDOM FOREST **53**FIG 7.6.4 ADDED NEW COLUMN AS PRICE RANGE IN TEST DATA SET **53**FIG 7.6.5 RANGE OF TARGET IN TRAIN DATASET **53**FIG

* + 1. OUTPUT OF PREDICTED TEST DATASET **54**FIG 7.6.7 EXAMPLE OF PREDICTED TEST DATASET

**54**