**VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI-590018**



**B.L.D.E.A’S V.P. Dr. P.G. HALAKATTI COLLEGE OF ENGINEERING AND TECHNOLOGY, VIJAYAPUR – 586 103**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**INTERNSHIP ON**

**"**MACHINE LEARNING USING PYTHON**"**

**UNDER THE GUIDANCE OF**

**Prof. PRABHU R BEVINMARAD**

**SUBMITTED BY:**

PADMA B SASNUR (2BL17CS045)

****

**VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI-590018**



**B.L.D.E. A’S V.P. DR. P.G. HALAKATTI COLLEGE OF ENGINEERING AND TECHNOLOGY, VIJAYAPUR – 586 103**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

This is to certify that the Internship [17CS84] on work entitled “MACHINE LEARNING USING PYTHON” is a bonafide work carried out by PADMA B SASNUR(2BL17CS045) submitted in partial fulfillment for the VIII Semester of Bachelor of Engineering degree of the Visvesvaraya Technological University, Belgaum during the year 2020-2021. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report. The internship report has been approved as it satisfies the academic requirements in respect of internship work prescribed for internship prescribed in Computer Science and Engineering of VIII semester.

**GUIDE HOD PRINCIPAL**

**PROF.PRABHU R BEVINMARAD DR. PUSHPA PATIL DR. ATUL AYARE**

**EXAMINERS SIGNATURE WITH DATE**

**1.**

**2.**

**ACKNOWLEDGEMENT**

I am glad to have this great opportunity to thank all the people who helped, guided and co-operated to complete internship named **"**Machine Learning Using Python **"** with a great success.

I would like to express deep sense of gratitude to our beloved principal **DR. ATUL AYARE** for providing all facilities in the college.

I would like to thank our Head of Department **DR. PUSHPA. PATIL** for providing facilities and fostering congenial academic environment in the college.

I feel deeply indebted to our esteemed guide **Prof. Prabhu R Bevinmarad** and for motivating and guiding us throughout this internship.

I would take this opportunity to thank all the faculty members and supporting staff for helping us in this endeavour.

# CONTENTS

1. ABOUT THE COMPANY 1
   1. [Company Profile 1](#_TOC_250009)
   2. [Mission](#_TOC_250008)
   3. [Vision](#_TOC_250007)
2. Tools Used 2
   1. [Jupyter Notebook 2](#_TOC_250006)
   2. [Jupyter Notebook Interface 3](#_TOC_250005)
3. TASK PERFORMED 4
   1. Project Introduction 4
      1. [Problem Statement 4](#_TOC_250004)
      2. [Data Set Information 4](#_TOC_250003)
      3. Algorithms Used 5
4. REFLECTION NOTES 6
   1. [Internship Experience 6](#_TOC_250002)
   2. [Activities during Internship 7](#_TOC_250001)
5. SOFTWARE SPECIFICATION 8
6. SOURCE CODE 9
   1. [Project Code 9](#_TOC_250000)
   2. Screenshots 11

**REFERENCES APPENDIX**

**CHAPTER 1 ABOUT THE COMPANY**

### COMPANY PROFILE

The Tech Fortune group was originally incepted in 2013 to focus only on EPC assignment of Infra and Real Estate Projects. Later the Company shifted its focus towards software development, Training ,BPO, Sourcing, Food Business, Health Care and Strategic Advisory Services. The Tech Fortune group Tech Fortune was born in 2013 with an objective to create a landmark initiative by a group of highly qualified technology oriented professionals in the software domain. A Software development Firm head quartered in Vijapur and operating for 5 years in with 3 offices across Karnataka. Since its inception in 2013, Tech fortune group has grown rapidly with the help of its valued Customers, professionals & Business Associates who have been continuously contributing and monitoring the Company’s business activities in the operations of Project Management, Education Consultancy, QMS and Six Sigma implementation and many other domain of expertise.

**Tech Fortune Technologies**, is an emerging technology organization in the fields of business process outsourcing, software development, end-to-end ERP solutions, Artificial Intelligence, Blockchain technology with a focus on providing customized solutions to the various business needs of a diverse global clientele.

### Mission

Being slow and steady, our mission is to gain the confidence of our clients and by dint of our integrity, innovation and dynamism, deliver their requirements on time with full quality thus bridging the gap between demand and delivery.

### Vision

With an unyielding focus on integrity and backed by strong founders and management team , Sourcing wants to make a mark in the field of IT services by applying innovation to simplify complex business processes and add value to clients’ business.

### Jupyter Notebook

**CHAPTER 2 TOOLS USED**

Jupyter Notebook (formerly IPython Notebooks) is a web-based interactive computational environment for creating Jupyter notebooks documents. The "notebook" term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context. A Jupyter Notebook document is a JSON document, following a versioned schema, and containing an ordered list of input/output cells which can contain code, text (using Markdown), mathematics, plots and rich media, usually ending with the ".ipynb" extension.

A Jupyter Notebook can be converted to a number of open standard output formats (HTML, presentation slides, LaTeX, PDF, ReStructuredText, Markdown, Python) through "Download As" in the web interface, via the nbconvert library or "jupyternbconvert" command line interface in a shell.

To simplify visualisation of Jupyter notebook documents on the web, the nbconvert library is provided as a service through NbViewer which can take a URL to any publicly available notebook document, convert it to HTML on the fly and display it to the user.

### Jupyter Notebook Interface

Jupyter Notebook provides a browser-based REPL built upon a number of popular open- source libraries:

* + - I Python
    - Tornado (web server)
    - J query
    - Bootstrap (front-end framework)
    - MathJax

Jupyter Notebook can connect to many kernels to allow programming in many languages. By default Jupyter Notebook ships with the IPython kernel. As of the 2.3 release[9][10] (October 2014), there are currently 49 Jupyter-compatible kernels for as many programming languages, including Python, R, Julia and Haske.

The Notebook interface was added to IPython in the 0.12 release December 2011), renamed to Jupyter notebook in 2015 (IPython 4.0 – Jupyter 1.0). Jupyter Notebook is similar to the notebook interface of other programs such as Maple, Mathematica, and SageMath, a computational interface style that originated with Mathematica in the 1980s. According to The Atlantic, Jupyter interest overtook the popularity of the Mathematics notebook interface in early 2018.

Jupyter Notebook is a web application that allows you to create and share documents that contain:

* live code (e.g. Python code)
* visualizations
* explanatory text (written in markdown syntax

Jupyter Notebook is great for the following use cases:

* learn and try out Python
* data processing / transformation
* numeric simulation
* statistical modeling
* Machine learning

## CHAPTER 3 TASKS PERFORMED

### Introduction to Project

Shopping malls play an important role in a retail sector and these malls have been in existence for more than 90 years. They have adapted to new designs and tenant varieties to meet the changes in consumers needs, desires, values, and lifestyles. There are different types of shopping malls, namely, convenience shopping mall, neighbourhood shopping mall, community shopping mall, regional shopping mall and super-regional shopping mall. The shopping mall offers business products, domestic appliances, a variety of services and entertaining equipment.

Shopping malls are characterized as venues that provide a comfortable shopping experience and have turned into social centres and recreational and entertainment facilities for various activities. Shopping mall customers visit shopping malls not only for searching for particular products, but they also view these visits as an entertainment activity that provides fun and pleasure from the shopping experience. Shopping mall visitors tend to engage in various activities during shopping malls visits.

The Mall customer’s dataset contains information about people visiting the mall. The dataset has gender, customer id, age, annual income, and spending score. It collects insights from the data and group customers based on their behaviours.

### Problem Statement

Segment the customers based on the age, gender, interest. Customer segmentation is an important practise of dividing customers base into individual groups that are similar. It is useful in customised marketing.

### Data Set Information:

Data includes data related to Customers , Age , Gender , Annual Income , Spending Score ( 0

– 100 ) . Publicly Available to Use.

**Available at :** kagle.com

1. Number of Instances : 200
2. Number of Attributes: 5
3. Attribute Information - Customer Id - Age - Gender - Annual Income - Spending Score

### Algorithm Used :

K Means Clustering k-means is one of the simplest unsupervised learning algorithms that solve the clustering problems. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters). The main idea is to define k centers, one for each cluster. To start with k-means algorithm, you first have to randomly initialize points called the cluster centroids (K). K-means is an iterative algorithm and it does two steps: 1. Cluster assignment 2. Move centroid step.

#### Cluster assignment

The algorithm goes through each of the data points and depending on which cluster is closer, it assigns the data points to one of the three cluster centroids.

#### Move centroid

Here, K-means moves the centroids to the average of the points in a cluster. In other words, the algorithm calculates the average of all the points in a cluster and moves the centroid to that average location. This process is repeated until there is no change in the clusters (or possibly until some other stopping condition is met). K is chosen randomly or by giving specific initial starting points by the user

## CHAPTER 4 REFLECTION NOTES

### Internship Experience

I had been interned in the course and I have a very nice experience to share. The team was amazing and flexible. They are ready to help us at our convenience. There was lot to know about the subject, sensors in specific. To me from CS background they made it as a very easy task. I want to thank everyone who helped me to make it possible.”

"I thoroughly enjoyed my internship this summer and now have very valuable experience under my belt. I know this will help when looking for jobs and needing references. I was dreading it in the beginning, but now I am so happy it was required. As much as the curriculum changes, I hope that class remains constant."

"We all know that practical experience is the best, and internships give students that hands-on experience they need. I feel that quality internships are essential to develop key skills that you can't get in a classroom. Skills such as multitasking, communicating, learning to deal with diversity, and dealing with deadlines are different when you are working for someone else, not yourself like you do in college. Internships are also a great way to network with people in the industry. My boss and co-workers were great about giving me contacts and referring me to open positions in the industry."

"I learned that customer service is hard. You can’t please everyone. There are people in this world that are simply difficult. I have learned that stressing over little things will not get me anywhere. I have learned to work well as a team and that without my counterparts the work would not get done. Another aspect that I learned throughout my internship is to never be afraid to ask lots of questions. By asking questions I got answers."

"I learned more than I ever expected. I felt in the beginning that I would really have a head start being that I am a senior and have taken all my classes already but I realized that I don’t think any amount of classes will ever prepare you for all the world has to offer. Don’t get me wrong, I feel that Auburn did a wonderful job teaching me but I feel that it is most valuable if combined with first hand work . I feel that during my time interning my most challenging part of my work was having to share my tasks and coordinate my work with others. I have always been kind of a perfectionist and liked to have things done my way so working with others and sharing duties was a slight change I must say. It took me a couple of weeks to

really feel comfortable allowing someone else to share my work but eventually I learned it was OK and that two heads were better than one."

### ACTIVITIES DURING INTERNSHIP:

* Introduction to ML
* Domain Training
* Training on Python
* Assignment

**Week 1 Activities**

* Training on statistics.
* More use of python libraries.
* Assignment on stats.

**Week 2 Activities**

* Training on ML.
* Algorithms of ML.
* Models in ML
* Assignment on ML
* Activities on Model building

**Week 3 Activities**

Internship Report Preparation

Project Based Activities

Assessment Test

Internship Project Work Presentation

●

●

●

●

**Week 4 Activities**

**CHAPTER 5 SOFTWARE SPECIFICATION**

#### Software used:

* + - Python Version: Python IDE 3.7
    - IDE :Anaconda
    - Packages used: sklearn, pandas, numpy

#### Hardware used:

* RAM: 2GB
* Hard disk: 500MB
* Processor: I3

### Project Code

**CHAPTER 6 SOURCE CODE**

import numpy as np import pandas as pd

import matplotlib.pyplot as plt import seaborn as sns

import shap data=pd.read\_csv("Mall\_Customers.csv")

from sklearn.preprocessing import LabelEncoder le=LabelEncoder() data['Gender']=le.fit\_transform(data['Gender']) sns.countplot('Gender',data=data,palette='winter') size=data['Gender'].value\_counts()

print('Female :',size[0]/(size[0]+size[1])\*100)

print('Male :',size[1]/(size[0]+size[1])\*100) plt.title("Gender distirbution")

plt.figure(1 , figsize = (15 ,6))

n = 0 color=['red','green','blue'] count=0

for x in ['Age' , 'Annual Income (k$)' , 'Spending Score (1-100)']: n += 1

plt.subplot(1 , 3 , n)

plt.subplots\_adjust(hspace =0.5 , wspace = 0.5) sns.distplot(data[x] , color=color[count]) plt.title('Distplot of {}'.format(x))

count+=1 plt.show()

sns.pairplot(data) plt.plot()

plt.rcParams['figure.figsize'] = (18, 8) corr=data.corr()

sns.heatmap(corr)

plt.title("Data correleation", fontsize=14) plt.plot()

plt.rcParams['figure.figsize'] = (18, 6)

sns.violinplot(data['Gender'], data['Spending Score (1-100)'], palette = 'pastel') plt.title('Gender vs Spending Score', fontsize = 14)

plt.show() plt.rcParams['figure.figsize'] = (18, 6)

sns.violinplot(data['Age'], data['Spending Score (1-100)'], palette = 'pastel') plt.title('Age vs Spending Score', fontsize = 14)

plt.show() plt.rcParams['figure.figsize'] = (18, 6)

sns.violinplot(data['Annual Income (k$)'], data['Spending Score (1-100)'], palette = 'pastel') plt.title('Gender vs Spending Score', fontsize = 14)

plt.show()

### Screenshot

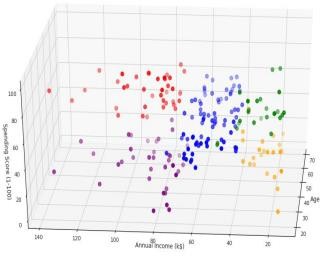


Fig 6.2.1 Mall Customers Segmentation

## REFERENCES

* + 1. [https://www.valuecoders.com/blog/technology-and-apps/how-ai-and-ml- have-](https://www.valuecoders.com/blog/technology-and-apps/how-ai-and-ml-have-)revamped-mobile-app-development/
    2. https://theappsolutions.com/blog/development/machine-learning-in-mobile-app/
    3. https://en.wikipedia.org/wiki/Python\_(programming\_language)
    4. https:/[/www.w3schools.com/](http://www.w3schools.com/python/python_intro.asp)p[ython/python\_intro.asp](http://www.w3schools.com/python/python_intro.asp) [5]https:/[/www.ana](http://www.analyticsvidhya.com/blog/2014/06/introduction-random-forest-)l[yticsvidhya.com/blog/2014/06/introduction-random-forest-](http://www.analyticsvidhya.com/blog/2014/06/introduction-random-forest-) simplified/