VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI-590014, KARNATAKA

INTERNSHIP REPORT

on

"Mobile Price Range Prediction Using KNN"

Submitted in Partial Fulfillment of the Requirement
For the Award of Degree

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE & ENGINEERING

Submitted By:

SANDESH GC

1SG21CS084

Under the Guidance of:

Prof. Shwetha K
Assistant Professor
Dept of CS&E

Department of Computer Science and Engineering

Accredited by NBA

SAPTHAGIRI COLLEGE OF ENGINEERING

(Affiliated to Visvesvaraya Technological University, Belagavi& Approved by AICTE, New Delhi.)
(Accredited by NAAC with 'A' Grade) (ISO 9001-2015 & 14001-2015 Certified)
#14/5, Chikkasandra, Hesaraghatta Main Road, Bengaluru– **560057**

2023-2024

SAPTHAGIRI COLLEGE OF ENGINEERING

(Affiliated to Visvesvaraya Technological University, Belagavi& Approved by AICTE, New Delhi.)
(Accredited by NAAC with 'A' Grade) (ISO 9001-2015 & 14001-2015 Certified)
#14/5, Chikkasandra, Hesaraghatta Main Road,
Bengaluru– **560 057**

Department of Computer Science and Engineering

Accredited by NBA

Certificate

Certified that the Internship entitled "Mobile Price range Prediction Using KNN" carried out by **SANDESH GC(1SG21CS084),** bonafide student of **Sapthagiri College of Engineering**, in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science and Engineering** of **Visvesvaraya Technological University**, **Belagavi** during the academic year 2023-2024. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the department library. The Internship report has been approved as it satisfies the academic requirements in respect of **Internship (21INT68)** prescribed for the said degree.

Signature of the Guide	**Signature of the HOD**
Prof. Shwetha K	**Dr. Kamalakshi Naganna**
Assistant Professor	**Professor & Head**

Certificate

ACKNOWLEDGEMENT

Any achievement does not depend solely on the individual efforts but on the guidance, encouragement and co-operation of intellectuals, elders and friends. A number of personalities, in their

own capacities have helped us in carrying out this Internship work. We would like to take this opportunity to thank them all.

We would like to express my profound thanks to **Sri. G Dayanand**, Chairman, Sapthagiri College of Engineering Bangalore, for his continuous support in providing amenities to carry out thisInternship.

Special Thanks to **Manoj G D**, Managing Director, Sapthagiri College of Engineering Bangalore, for his valuable suggestion.

I am grateful to thank our Principal **Dr. H Ramakrishna**, Sapthagiri College of Engineering,who patronized throughout our career & for the facilities provided to carry out this work successfully.

It's a great privilege to place on record my deep sense of gratitude to our beloved **HOD Dr. Kamalakshi Naganna** of Computer Science & Engineering, who patronized throughout our career & for the facilities provided to carry out this work successfully.

I am also grateful to thank **Prof. Shwetha K, Assistant Professor, Dept. of CSE** Internship guide for his/her invaluable support and guidance.

I am also grateful to thank Internship Coordinators **Prof. Hemalatha K, Dr. Madhavi Dasari** for their invaluable support and guidance.

I would also like to thank the teaching and non teaching staff members who have helped medirectly or indirectly during the Internship.

Finally, I also thank my family and friends for their co-operation and motivation to completethis internship successfully.

SANDESH GC

1SG21CS084

STUDENT DECLARATION

I, Sandesh GC hereby declare that the presented report of internship titled “Mobile Price Range Prediction Using KNN” is uniquely prepared by me after the completion of one months’ work at KARUNADU TECHNOLGY PRIVATE LIMITED.

I also confirm that the report is only prepared for my academic requirement, not for any other purpose. I further declare that the work reported in this report has not been submittedeither in part or in full for the award of any other degree or diploma in this institute or any other institute or university.

Sandesh GC

1SG21CS084

Abstract

In the rapidly evolving mobile phone market, predicting the price range of new devices is crucial for both manufacturers and consumers. Accurate price range predictions can inform strategic decisions related to marketing, product development, and competitive positioning. This study presents a machine learning approach using the K-Nearest Neighbours (KNN) algorithm to predict the price range of mobile phones based on a comprehensive set of features, including battery power, RAM, internal memory, screen size, camera quality, and processor speed.

We begin by preprocessing a dataset of existing mobile phones, ensuring data quality through normalization and handling missing values. The KNN model is then trained on this dataset to classify new phones into predefined price ranges: low, medium, high, and premium. By selecting the optimal number of neighbors (k) and using cross-validation techniques, we fine-tune the model for better accuracy and robustness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCREENSHOTS

WEEKLY OVERVIEW OF INTERNSHIP ACTIVITIES

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
WEEK 1	06/11/2023	MONDAY	Introduction
	07/11/2023	TUESDAY	Introduction to Python and AI,ML
	08/11/2023	WEDNESDAY	Setting Up Jupyter Notebook
	09/11/2023	THURSDAY	Basics Of Python - l
	10/11/2023	FRIDAY	Basics Of Python - ll

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
WEEK 2	13/11/2023	MONDAY	Introduction to Artificial Intelligence
	14/11/2023	TUESDAY	Real World Examples Of AI
	15/11/2023	WEDNESDAY	Introduction to AI libraries in Python
	16/11/2023	THURSDAY	Projects using basic AI libraries in Python
	17/11/2023	FRIDAY	Gathering Inputs from various domains.

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
WEEK 3	20/11/2023	MONDAY	Introduction to Pandas
	21/11/2023	TUESDAY	Python to Manage the Data
	22/11/2023	WEDNESDAY	Designing of UI using PyQt Designer
	23/11/2023	THURSDAY	Assigning appropriate constraints on the databases
	24/11/2023	FRIDAY	Code Development of the login page.

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
WEEK 4	27/11/2023	MONDAY	Introduction to Machine Learning
	28/11/2023	TUESDAY	Introduction to Various ML Models
	29/11/2023	WEDNESDAY	Linear Regression
	30/11/2023	THURSDAY	Logistic Regression
	01/12/2023	FRIDAY	K-Nearest Neighbors

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION TO TOPIC

1.1.1 Purpose of the project

The primary purpose of this project is to develop a machine learning model using the K-Nearest Neighbors (KNN) algorithm to predict the price range of mobile phones based on various features such as battery power, RAM, internal memory, screen size i.e.(width, height), and processor speed, dual sim, no of cores, four g, Wi-Fi and touch screen. This prediction model aims to assist manufacturers, guide consumers, Enhance market analysis, optimize product development and promote the data driven decisions

1.1.2 Problem Identification

The mobile phone market's rapid innovation and diverse features create challenges for predicting pricing trends. With numerous models and continuous releases, it’s difficult for manufacturers and consumers to understand and forecast prices. Identifying which features most impact price is complex, and consumers struggle to make informed decisions due to overwhelming options and unclear pricing benchmarks

1.1.3 Proposed System

The proposed system utilizes KNN to predict the mobile price range based on attributes such as battery power, RAM, internal memory, screen size i.e.(width, height), and processor speed, dual sim, no of cores, four g, Wi-Fi and touch screen.

1.1.4 Advantages of Proposed System

1. Consumer Guidance: By categorizing mobile phones into predefined price ranges, the system aids consumers in making informed purchasing decisions, enhancing their shopping experience
2. Efficiency: KNN can be computationally efficient, enabling faster analysis than traditional methods.
3. Continuous Improvement: Accuracy can be enhanced by incorporating new data into the training process.

1.2 COMPANY INTRODUCTION

Karunadu Technologies is a Bangalore-based company established in 2018 that stands at the forefront of technological innovation. Their passionate team of engineers and developers are dedicated to creating cutting- edge solutions that address the challenges of today's world. Their areas of expertise span various domains, including artificial intelligence, machine learning, data science, and the Internet of Things (IoT). They are committed to empowering businesses and organizations with the tools they need to thrive in the digital age.

Karunadu Technologies believes in the transformative power of technology and its ability to improve lives across industries. Whether you're a startup looking to disrupt your market or an established company seeking tooptimize operations, Karunadu Technologies can be a valuable partner. They offer a range of services, includingsoftware development, web design, product development, embedded system solutions, and skill development programs. By leveraging their expertise, you can achieve your goals and stay ahead of the curve in the ever- evolving technological landscape.

1.2.1 About Department:

Karunadu Technologies appears to focus on software development, particularly in areas like machine learning,data science, and artificial intelligence. This suggests they likely have a dedicated Software Development department responsible for designing, building, and maintaining these software solutions. Additionally, their services include time tracking and billing software, indicating a potential Business Development or Sales department to manage client relationships and project acquisition.

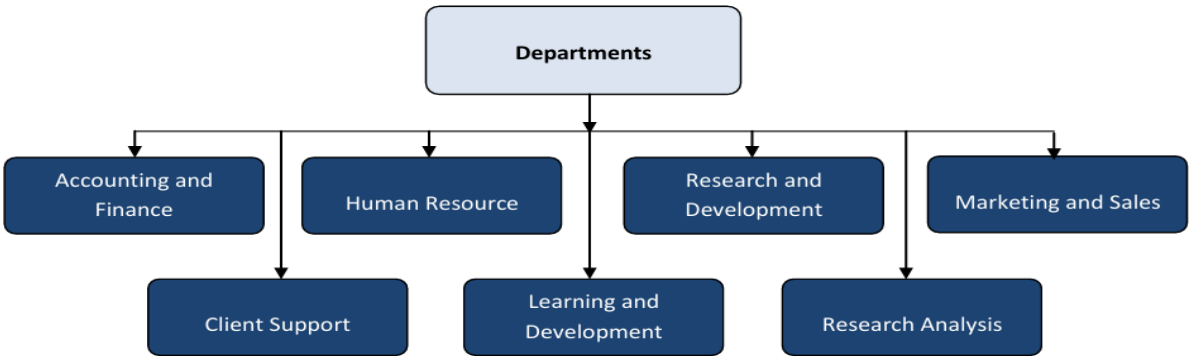

Figure 1.1: Departments of Karunadu Technologies

CHAPTER 2

TRAINING PROGRAM

Throughout this internship program, I undertook a variety of tasks to successfully build the KNN-based breast cancer prediction system. Initially, I delved into the domain of breast cancer, familiarizing myself with various cancer types, diagnosis methods, and the importance of early detection. This knowledge foundation was crucialfor understanding the problem statement and the potential impact of the proposed system.

Next, I focused on acquiring the technical skills necessary for building the machine learning model. This involved learning the KNN algorithm, its working principles, and its suitability for classification tasks. I explored different distance metrics used in KNN to determine the similarity between data points. Furthermore,I practiced data preprocessing techniques to clean and prepare the breast cancer dataset for model training. Thisincluded handling missing values, identifying and addressing outliers, and potentially feature scaling to ensureall features contribute equally during the classification process.

The KNN model for breast cancer prediction requires a well-defined training program to ensure optimal performance. Here are some key steps involved:

i. Data Preprocessing: The raw data containing patient information and cancer classifications might requirecleaning and preprocessing steps. This could involve handling missing values, identifying and correctingoutliers, and potentially scaling features to a common range for improved distance calculations.

ii. Data Splitting: The preprocessed data is then divided into two sets: a training set and a testing set. The training set (typically the larger portion) is used to train the KNN model. The testing set is used to evaluate the model's performance on unseen data.

iii. K-Value Selection: A crucial step is determining the optimal value of K, the number of nearest neighborsto consider for prediction. Techniques like cross-validation can be employed. Here, the training data is further divided into smaller folds.

iv. Model Training: With the preprocessed data split and the K value chosen, the KNN algorithm is trainedon the training set. This involves calculating distances between all data points in the training set.

v. Model Evaluation: Once trained, the KNN model's performance is evaluated on the testing set. Commonmetrics for classification tasks include accuracy, precision, recall, and F1-score.

vi. Model Refinement: Based on the evaluation results, the training program might involve further refinement. This could include adjusting hyperparameters like K value, exploring different distance metrics, or potentially feature selection techniques to improve model performance.

CHAPTER 3

LEARNING EXPERIENCE

Knowledge acquired:

- Machine Learning Fundamentals: Gained a deeper understanding of core machine learning concepts like supervised learning, classification algorithms, and KNN's working principles.
- Breast Cancer Diagnosis: Learned about different breast cancer types, diagnostic techniques, and the importance of early detection.
- Data Preprocessing and Analysis: Explored techniques for data cleaning, feature engineering,and exploratory data analysis to prepare the data for modeling.
- KNN Model Development: Understood the process of KNN model training, hyperparameter tuning, and evaluation metrics like accuracy and sensitivity.

Skills learned:

- Data Manipulation: Developed skills in using programming languages (Python likely) to handle, manipulate, and visualize breast cancer data.
- Machine Learning Libraries: Learned to utilize machine learning libraries like scikit-learn to implement and train the KNN model.
- Model Evaluation: Gained practical experience in evaluating the performance of the KNN model using appropriate metrics and interpreting the results.
- Technical Report Writing: Enhanced skills in writing technical reports by structuring information, presenting findings clearly, and effectively communicating technical concepts.

Observed attitudes and gained values:

- Importance of Interdisciplinary Work: Witnessed the value of collaboration between data science and healthcare professionals for developing impactful solutions.
- Ethical Considerations in AI: Understood the ethical considerations surrounding AI inhealthcare, particularly issues like bias and fairness in model development.
- Importance of Continuous Learning: Recognized the value of continuous learning in the rapidly evolving field of machine learning and healthcare technology.
- Patient-Centric Approach: Appreciated the importance of keeping the patient's well-being at the forefront when developing machine learning solutions for healthcare applications.

The most challenging tasks performed:

- Data Cleaning and Preprocessing: Dealing with missing values, outliers, and inconsistencies in thebreast cancer dataset required careful attention and exploration.
- Hyperparameter Tuning: Finding the optimal hyperparameters for the KNN model involved experimentation and evaluating various configurations.
- Interpreting Model Results: Understanding the limitations and interpreting the rationale behind theKNN model's predictions was a challenging but valuable exercise.
- Communicating Complex Concepts: Translating technical machine learning concepts into anunderstandable format for a broader audience required clear and concise communication skills.

CHAPTER 4

STRENGTH, WEAKNESS, OPPORTUNITIES, THREATS (SWOT) ANALYSIS

STRENGTH:

- Accuracy: KNN can achieve good accuracy in classification tasks, especially when dealingwith well-defined datasets.
- Interpretability: Compared to complex machine learning models, KNN offers some level of interpretability, allowing us to understand the factors influencing its predictions.
- Efficiency: KNN is a relatively simple algorithm with efficient computational requirements, making it suitable for real-time applications.
- Scalability: The system can be potentially scaled to handle larger datasets as more data becomes available.
- Low Training Cost: KNN requires minimal data pre-processing and can be trained withmoderate computational resources.

WEAKNESS:

- Sensitivity to Data Quality: KNN performance heavily relies on the quality and relevance of the training data. Issues like noise and irrelevant features can negatively impact accuracy.
- Curse of Dimensionality: As the number of features in the data increases, KNN's effectivenesscan diminish. Feature selection techniques may be needed to address this.
- Instance-based Learning: KNN stores the entire training dataset for classification, which can be memory intensive for large datasets.
- Parameter Tuning: Selecting the optimal value for the K parameter in KNN can be challenging and can significantly impact performance.
- Limited Class Imbalance Handling: KNN may struggle with imbalanced datasets where certain cancer types are less frequent.

OPPORTUNITIES:

- Market Insights: Gain valuable insights into market trends, consumer preferences, and competitive positioning within the mobile phone industry, enabling informed business decisions.
- Product Development: Facilitate better alignment of product features with pricing strategies based on accurate price range predictions, potentially increasing market competitiveness.
- Consumer Engagement: Enhance consumer engagement by providing transparent pricing information, aiding in informed purchasing decisions and improving customer satisfaction.
- Data Analytics skills: Strengthen expertise in data analytics and machine learning, which are increasingly in demand across various industries for predictive modeling and decision support.
- Research and Development: Further research can explore advancements in KNN algorithms andensemble methods to improve the system's accuracy and robustness.

THREATS:

- Data Privacy and Security: The KNN system should be used as a complementary tool,not a replacement for physician expertise and judgment.
- Model accuracy: Challenges in maintaining model accuracy over time due to evolving market dynamics, new product introductions, and changes in consumer preferences
- Competitive Pressures: Intense competition within the mobile phone industry may lead to rapid shifts in pricing strategies and market positioning, impacting the reliability of predictive models.
- Technological Advancements: Constant advancements in machine learning and data analytics may require continuous learning and adaptation to stay competitive and relevant.
- Ethical Considerations: Ethical implications related to the use of predictive analytics in influencing consumer behavior and market dynamics, necessitating responsible and transparent practices.

CHAPTER 5

IMPLEMENTATION

5.1 PROBLEM STATEMENT

The problem statement for predicting mobile price ranges using KNN involves developing a machine learning model that can accurately classify mobile phones into predefined price categories based on their features such as battery power, RAM, internal memory, camera quality, and processor speed. The goal is to provide stakeholders, including manufacturers, consumers, and market analysts, with a tool that can predict price ranges effectively, aiding in strategic decision-making, product positioning, and consumer guidance in the competitive mobile phone market. The project aims to leverage KNN's ability to identify similarities between mobile phone features and historical pricing data to automate and enhance the accuracy of price range predictions, thereby addressing challenges related to market complexity, feature variability, and consumer decision-making.

5.1.1 Dataset

The dataset used in this project consists of 20 variables: ‘battery_power,’blue’,’clock_speed’,’dual_sim’,’fc’,’int_memory’,’m_dep’,’mobile_wt’,’n_cores’,’pc’,’px_height’,’px_width’,’ram’,’sc_h’,’sc_w’,’talktime’,’three_g’,’touch_screen’,’wifi’

battery_po	blue	clock_spee	dual_sim	fc	four_g	int_memor	m_dep	mobile_wt	n_cores	pc	px_height	px_width	ram	sc_h	sc_w	talk_time	three_g	touch_scr	wifi	price_range
842	0	2.2	0	1	0	7	0.6	188	2	2	20	756	2549	9	7	19	0	0	1	1
1021	1	0.5	1	0	1	53	0.7	136	3	6	905	1988	2631	17	3	7	1	1	0	2
563	1	0.5	1	2	1	41	0.9	145	5	6	1263	1716	2603	11	2	9	1	1	0	2
615	1	2.5	0	0	0	10	0.8	131	6	9	1216	1786	2769	16	8	11	1	0	0	2
1821	1	1.2	0	13	1	44	0.6	141	2	14	1208	1212	1411	8	2	15	1	1	0	1
1859	0	0.5	1	3	0	22	0.7	164	1	7	1004	1654	1067	17	1	10	1	0	0	1
1821	0	1.7	0	4	1	10	0.8	139	8	10	381	1018	3220	13	8	18	1	0	1	3
1954	0	0.5	1	0	0	24	0.8	187	4	0	512	1149	700	16	3	5	1	1	1	0
1445	1	0.5	0	0	0	53	0.7	174	7	14	386	836	1099	17	1	20	1	0	0	0
509	1	0.6	1	2	1	9	0.1	93	5	15	1137	1224	513	19	10	12	1	0	0	0
769	1	2.9	1	0	0	9	0.1	182	5	1	248	874	3946	5	2	7	0	0	0	3
1520	1	2.2	0	5	1	33	0.5	177	8	18	151	1005	3826	14	9	13	1	1	1	3
1815	0	2.8	0	2	0	33	0.6	159	4	17	607	748	1482	18	0	2	1	0	0	1
803	1	2.1	0	7	0	17	1	198	4	11	344	1440	2680	7	1	4	1	0	1	2
1866	0	0.5	0	13	1	52	0.7	185	1	17	356	563	373	14	9	3	1	0	1	0
775	0	1	0	3	0	46	0.7	159	2	16	862	1864	568	17	15	11	1	1	1	0
838	0	0.5	0	1	1	13	0.1	196	8	4	984	1850	3554	10	9	19	1	0	1	3
595	0	0.9	1	7	1	23	0.1	121	3	17	441	810	3752	10	2	18	1	1	0	3
1131	1	0.5	1	11	0	49	0.6	101	5	18	658	878	1835	19	13	16	1	1	0	1
682	1	0.5	0	4	0	19	1	121	4	11	902	1064	2337	11	1	18	0	1	1	1
772	0	1.1	1	12	0	39	0.8	81	7	14	1314	1854	2819	17	15	3	1	1	0	3
1709	1	2.1	0	1	0	13	1	156	2	2	974	1385	3283	17	1	15	1	0	0	3
1949	0	2.6	1	4	0	47	0.3	199	4	7	407	822	1433	11	5	20	0	0	1	1
1602	1	2.8	1	4	1	38	0.7	114	3	20	466	788	1037	8	7	20	1	0	0	0
503	0	1.2	1	5	1	8	0.4	111	3	13	201	1245	2583	11	0	12	1	0	0	1

train

Figure 5.1 Overview of Dataset

5.1.2 Variable Information

The dataset used for training and prediction consists of 20 features extracted from mobile ranges These features aim to capture various characteristics that might be able to predict the price range Here's a breakdown ofthe variables:

- Battery_power: Represents the battery power in mAh.
- Blue: Has Bluetooth or not (1 = Yes, 0 = No).
- clock_speed: Speed at which microprocessor executes instructions.
- dual_sim: Has dual SIM support or not (1 = Yes, 0 = No).
- fc: Front Camera megapixels.
- four_g: Has 4G support or not (1 = Yes, 0 = No).
- int_memory: Internal Memory in Gigabytes.
- m_dep: Mobile Depth in cm.
- mobile_wt: Weight of mobile phone.
- n_cores: Number of cores of processor.
- pc: Primary Camera megapixels.
- px_height: Pixel Resolution Height.
- px_width: Pixel Resolution Width.
- ram: Random Access Memory in Megabytes.
- sc_h: Screen Height of mobile in cm.
- sc_w: Screen Width of mobile in cm.
- talk_time: Longest time that a single battery charge will last when you are talking.
- three_g: Has 3G support or not (1 = Yes, 0 = No).
- touch_screen: Has touch screen or not (1 = Yes, 0 = No).
- wifi: Has WiFi support or not (1 = Yes, 0 = No).
- price_range: This is the target variable with values ranging from 0 (lowest) to 3 (highest).

5.2Algorithm

5.2.1. K-Nearest Neighbors:

- K-Nearest Neighbors (KNN) is a simple yet powerful machine learning algorithm for classification and regression.
- It predicts the class of a new data point based on the majority class of its k nearest neighbors in the training data. KNN is versatile, handling various data types and requiring minimal assumptions about the underlying data distribution.

STEPS:

The K-Nearest Neighbors (KNN) algorithm follows a straightforward approach for classification:

1. Define K: This is a crucial step where you determine the number of nearest neighbors to consider for prediction. A higher K value reduces the impact of noise but might lead to overfitting.
2. Calculate Distance: For a new data point (patient record), KNN calculates the distance to alldata points in the training set. Commonly used distance metrics include:
    - Euclidean Distance: This measures the straight-line distance between two points in n-dimensional space (where n is the number of features). It's calculated as:
    - Manhattan Distance: This calculates the total distance traveled along each axis to reach another point.
3. Identify Nearest Neighbors: After calculating distances, KNN selects the k data points from the training set closest to the new data point.
4. Prediction:
    - Classification: For classification tasks like breast cancer type prediction, KNN determines the most frequent class label (malignant or benign) among the k nearest neighbors. The new data point is assigned the majority class label.
    - Regression: For regression tasks, KNN predicts the value for the new data point by averaging the values of its k nearest neighbors.

5.3 FORMULAS

$$Manhattan\ Distance = \sum_{i=1}^{d} |x_{1i} - x_{2i}|$$

$$Euclidean\ Distance = \left(\sum_{i=1}^{d} (x_{1i} - x_{2i})^2 \right)^{\frac{1}{2}}$$

$$Minkowski\ Distance = \left(\sum_{i=1}^{d} |x_{1i} - x_{2i}|^p \right)^{\frac{1}{p}}$$

Figure 5.2.1 : Distance Formulas

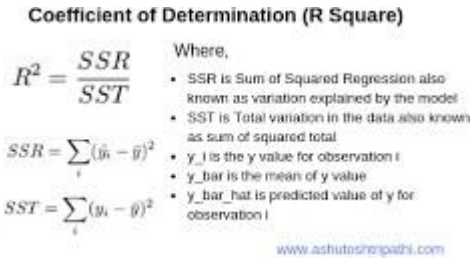

Figure 5.2.2 : Residual Sum Of Squares

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

Where:

$\hat{y}_i$ = Predicted value for the i^{th} data point

y_i = Actual value for the i^{th} data point

n = number of observations

Figure 5.2.3 : Mean Square Error

5.4 CODE :

```
import pandas as pd

path = "C:\\programs\\internship final\\train.csv"

data=pd.read_csv(path)
print(data)
print(data.info())
print(data.shape)
print(data.isnull().sum())
```

```
inputs=data.drop(['price_range'],axis=1)
outputs=data.drop(['battery_power','blue','clock_speed','dual_sim','fc','four_g','int_memory','m_dep','mobile_wt','n_cores','pc',
print(inputs)
print(outputs)
```

```
from sklearn.metrics import accuracy_score
```

```
y_pred = model.predict(x_test)
print(y_pred)

acc=accuracy_score(y_test,y_pred)
print(acc*100)
```

```
from sklearn.model_selection import train_test_split
```

```
x_train,x_test,y_train,y_test = train_test_split(inputs,outputs,test_size=0.2,random_state=32)
```

```
import sklearn
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC

model= DecisionTreeClassifier(max_depth=5)
# model = KNeighborsClassifier(n_neighbors = 45)
model = SVC(kernel ='rbf')
model.fit(x_train,y_train)
```

CHAPTER 6

RESULTS

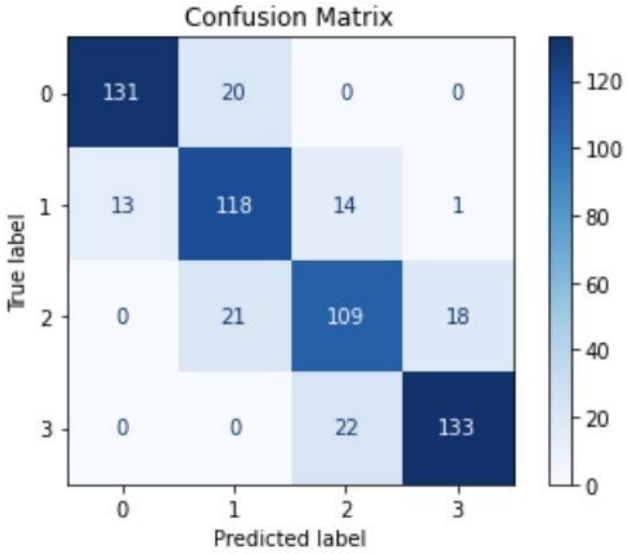

Figure 6.1 : Confusion Matrix

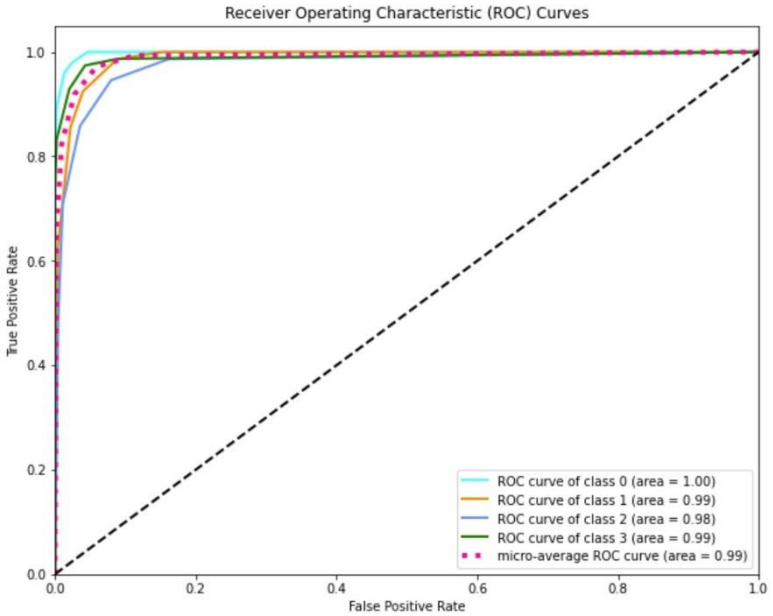

Figure 6.2 : ROC Curve -KNN

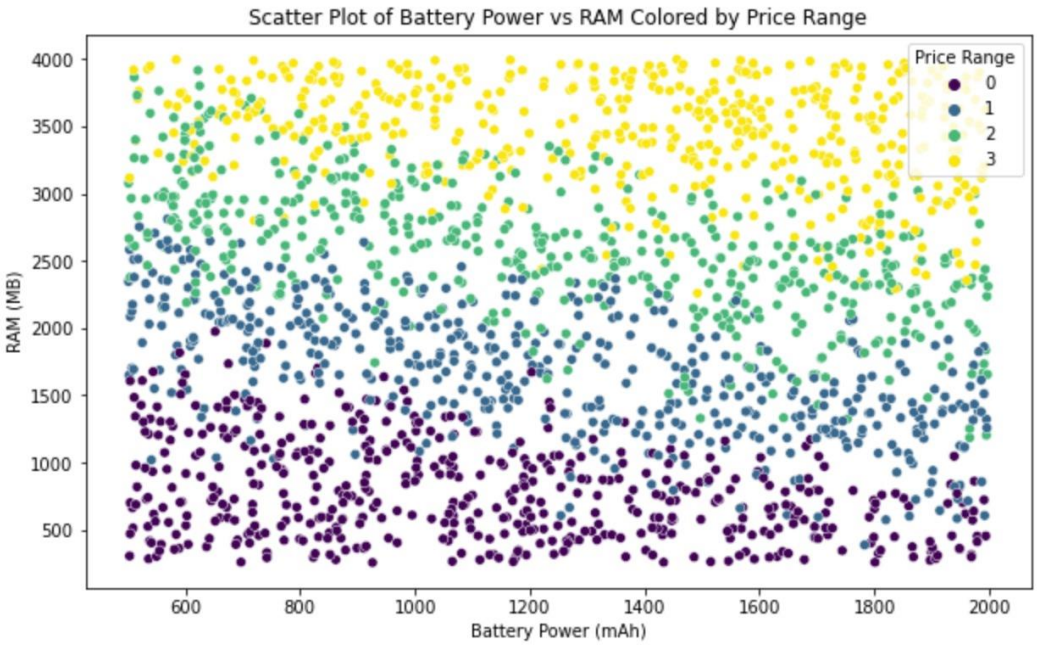

Figure 6.3 : Scatter Plot with Decision Boundary - KNN

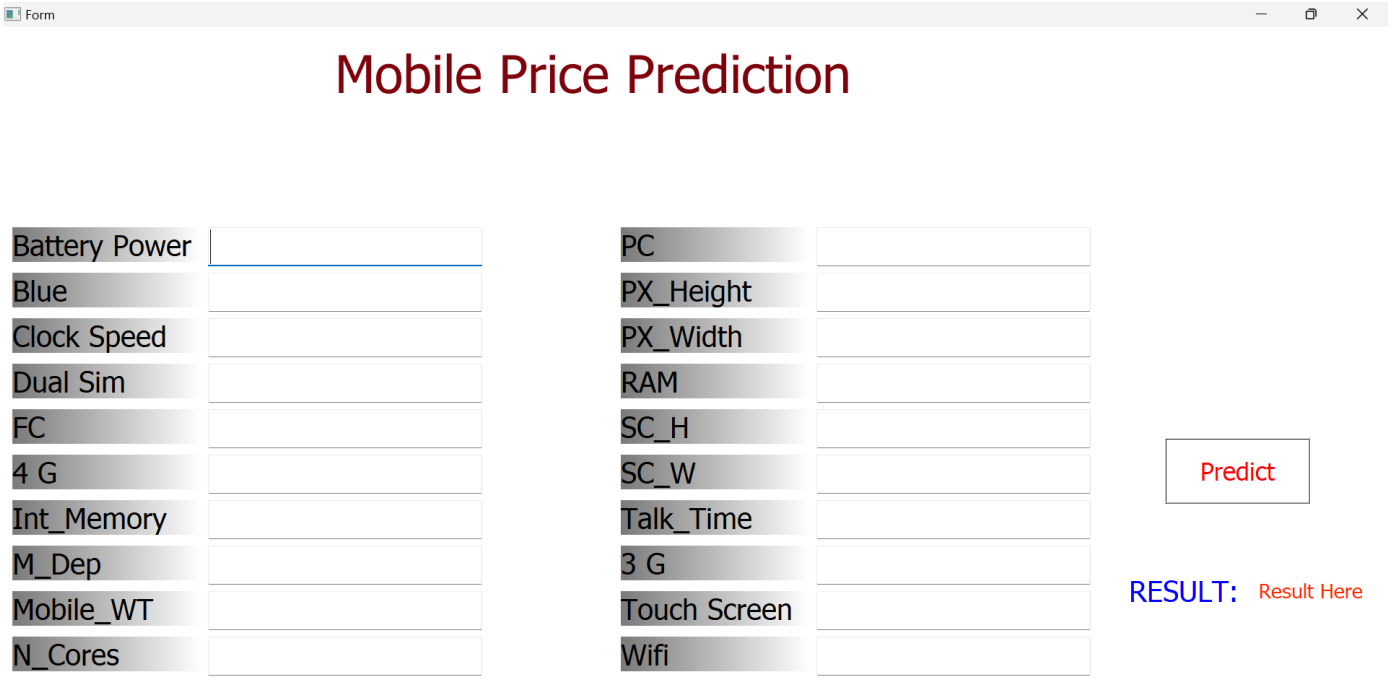

Figure 6.4 : Output

CHAPTER 7

CONCLUSION AND FUTURE ENHANCEMENTS

Conclusion

In conclusion, developing a K-Nearest Neighbors (KNN) based model for predicting mobile phone price ranges addresses the complexities and challenges inherent in the rapidly evolving mobile phone market. By leveraging various features such as battery power, RAM, internal memory, camera quality, and processor speed, the model provides accurate and reliable price range predictions. This tool not only aids manufacturers in strategic product positioning and development but also helps consumers make informed purchasing decisions by categorizing phones into well-defined price segments.

Future Enhancements:

- **Integration of Additional Features**: Incorporate more features such as user reviews, brand reputation, and market trends to improve the model's accuracy and comprehensiveness in predicting mobile phone price ranges.
- **Advanced Algorithms**: Explore and compare advanced machine learning algorithms like Random Forest, Gradient Boosting, and Neural Networks to potentially enhance predictive performance and robustness.
- **Dynamic Model Updating**: Implement a dynamic updating mechanism where the model is periodically retrained with new data to keep up with the latest market trends and newly released mobile phone models.
- **Geographical Customization**: Tailor the model to account for regional variations in pricing and consumer preferences, providing more localized and relevant predictions for different markets.
- **User Interface Development**: Develop an intuitive user interface that allows non-technical users, such as consumers and business analysts, to easily input mobile phone features and obtain price range predictions.
- **Real-Time Prediction**: Enable real-time price range predictions by integrating the model with mobile phone databases and online marketplaces, offering immediate insights for both manufacturers and consumers.

- **Explainability and Transparency**: Enhance the model's interpretability by implementing methods to explain the influence of different features on the price prediction, fostering trust and transparency among users.
- **Scalability**: Ensure the system can handle large-scale data inputs and provide predictions for a vast number of mobile phone models without compromising performance.
- **Cross-Domain Applications**: Extend the application of the predictive model to other consumer electronics such as tablets, laptops, and smartwatches, leveraging similar features and market dynamics.
- **Collaboration with Industry Stakeholders**: Collaborate with mobile phone manufacturers, retailers, and market analysts to gather high-quality data, validate model predictions, and refine the system based on real-world feedback and requirements.

BIBLIOGRAPHY

[1] Manoj Kumar, Urmila Pilania (2023): Predicting Mobile Phone Prices with Machine Learning.

[2] Ningyuan Hu (2022): Classification of Mobile Phone Price Dataset Using Machine Learning Algorithms.

[3] Madhavi Pujari, Chetan Awati, Sonam Kharade (2019). Efficient Classification with an Improved Nearest Neighbor Algorithm

[4] Data Science and Machine Learning (2018).

https://people.smp.uq.edu.au/DirkKroese/DSML/DSML.pdf

[5] Hands on Machine Learning with Sckiti Learn and Tensor Flow

https://github.com/yanshengjia/ml-road/blob/master/resources/Hands.pdf

[6] Scikit-learn documentation. (n.d.). K-Nearest Neighbors. Retrieved from https://scikit-learn.org/stable/modules/neighbors.html

[7]Machine Learning Mastery. (2019). A Gentle Introduction to k-Nearest Neighbors Algorithm. Retrieved from https://machinelearningmastery.com/k-nearest-neighbors-for-machine-learning/

[8] Towards Data Science. (2020). K-Nearest Neighbors (KNN) Algorithm for Machine Learning. Retrieved from https://towardsdatascience.com/k-nearest-neighbors-knn-algorithm-for-machine-learning-e883219c8f26