INTERNSHIP REPORT

A report submitted in partial fulfillment of the requirements for 3rd Year

BACHELOR OF ENGINEERING
in
ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
by

"Prathamesh Santosh Pagale"

Exam Seat. No.:T190352064

Under Supervision of

Mrs. H.K.Jadhav ,

Industry mentor Mrs. Uzma Sardar at

Edunet Foundation

(Duration: Start date[14/12/2023] to End date[15/01/2024])

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

VIDYA PRATISHTHAN'S INSTITUTE OF ENGINEERING AND TECHNOLOGY,

BARAMATI

ACADEMIC YEAR: 2023-24

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
VIDYA PRATISHTHAN'S INSTITUTE OF ENGINEERING AND TECHNOLOGY
BARAMATI.

CERTIFICATE

This is to certify that the "**Internship report"** submitted by **Prathamesh Santosh Pagale (Exam Seat No.:T190352064)** is work done by her **Online** Mode and submitted during **2023–2024** academic year, in partial fulfillment of the requirements for the **3rd Year** of **BACHELOR OF ENGINEERING in ARTIFICIAL INTELLIGENCE AND DATA SCIENCE,** at Edunet Foundation.

Mr. H.K.Jadhav	**Mrs. H.K.Jadhav**	**Dr. P.M.Paithane**
Internship Mentor	**Dept Internship Coordinator**	**HOD,AI & DS**

Dr. R.S. Bichkar
Principal

CERTIFICATE OF INTERNSHIP

ACKNOWLEDGEMENT

First, I would like to thank **Edunet Foundation**,for giving me the opportunity to do an internship within the organization.

I also would like all the people that worked along with **Edunet Foundation** with their patience and openness they created an enjoyable working environment.

It is indeed with a great sense of pleasure and immense sense of gratitude that I acknowledge the help of these individuals.

I am highly indebted to Principal **Dr. R.S.Bichkar**, for the facilities provided to accomplish this internship.

I would like to thank my Head of the Department **Dr. P.M.Paithane** for his constructive encouragement throughout my internship.

I would like to thank **Mrs. H.K.Jadhav ,**internship coordinator Department of **Artificial Intelligence and Data Science** for their support andadvices to get and complete internship in above said organization.

I am extremely great full to my department staff members and friends who helped me in successful completion of this internship.

Prathamesh Santosh Pagale

(Exam Seat No: T190352064)

ABSTRACT

Edunet foundation provides 4-week , 6-weeks virtual internship program where have internship opportunities in various domains such as Web Development, Data Science, Machine learning, Cloud computing etc. As an intern at Edunet foundation, I had the opportunity to participate in a 4-week Virtual Internship Program in the field of Artificial intelligence and Cloud computing .

Internship started with some lectures on overview of Machine Learning, data analytics, supervised and unsupervised types of Machine Learning, data pre-processing, model selection and evaluation. Doubt sessions were conducted as well. In this internship, I developed a machine learning model for E-commerce customer analysis results shows that time spent on app have more significant impact on sales compared to time spent on website. I used Linear Regression and graphical methods to determine relation between different features.

During this internship, I gained practical work experience and an introduction to creating and improving data-Science and Analytics Projects. Also I get hands on experience on IBM cloud while creating data analysis project . The program involved challenging opportunities and real-world projects, providing me with hands-on experience in the field of machine learning with python using data analytics. vThe guidance and support provided by the award-winning team at Edunet foundation helped me to grow and develop my skills in this field.

Overall, this internship gave me a good opportunity to learn about machine learning firsthand and use my newfound knowledge to solve problems in the real world. Through this effort, we were able to comprehend the fundamentals of machine learning and how they may be used in different scenarios.

Keywords Artificial Intelligence (AI), Cloud Computing, Hands-on Projects, Real-world Scenarios, Machine learning , analytics , Virtual .

INDEX

LIST OF FIGURES

1.INTRODUCTION

The goal of internship is on Data Analytics, Data Science and Cloud computing was an enriching experience that helped me gain practical knowledge and skills in the field of data science. During the internship, I worked on various projects related to data analysis, machine learning, and predictive modeling using python and data analytics techniques.

Against this backdrop, the four-week intensive internship, facilitated by the Edunet Foundation and sponsored by the All India Council for Technical Education (AICTE), served as a pivotal platform for exploring the practical applications of AI and cloud technologies.

Throughout the internship, participants were immersed in hands-on learning experiences, with a particular focus on leveraging the IBM Cloud Platform and Watson Studio for developing and deploying machine learning models. Guided by experienced mentors and industry experts.

The main objective of the internship was to apply the theoretical concepts of data science to real-world problems and develop solutions that can provide valuable insights to businesses and organizations using machine learning algorithms. Lectures were given by the instructor herself and they were based on basic concepts of machine learning, Cloud Computing

Information related to cloud it's working , interface and real world applications are delivered so well by expert instructor . Due to this I was able to work with IBM cloud and do some projects .

2.LEARNING OBJECTIVES/INTERNSHIP OBJECTIVES

1. **Courses:**
   Access to courseware that includes emerging technologies like data science, cybersecurity, artificial intelligence , cloud computing and Quantum Computing

2. **Access to IBM Cloud and Software:**
   Enhanced access to the IBM Cloud and select cloud-based resources, applications like IBM Watson studio , Watson.ai , etc.

3. **Hands-on Projects:**
   The program includes practical assignments and projects that enable students to apply their knowledge and gain hands-on experience in real-world scenarios which includes machine learning , data science , data analytics , etc.

4. **Badges and Certifications:**
   Upon successful completion of courses on IBM skills build platform students earn digital badges and certifications from IBM, validating their skills and accomplishment

5. **Collaborative work :**
   Forming the small groups and allocate tasks to each respective group as a task which increases collaborative work

3.MOTIVATION/SCOPE OF THE INTERNSHIP

Motivation/Scope of an Internship:

- The program helped me develop a comprehensive understanding of data science processes and methodologies and improve my analytical and problem-solving skills..
- The internship exposed me to different domains and industries where data science is applied, providing valuable insights into potential career paths.
- Working on real-world projects and applying my knowledge and skills to solve challenging problems was incredibly rewarding .
- Opportunities for certification and badges from IBM validate students' skills and accomplishments, enhancing their credibility in the job market.
- The opportunity to gain hands-on experience in the field of data science and its applications was truly unique.
- I am motivated to continue building on the knowledge and skills I have gained during this internship and to pursue a career in this exciting field.

The scope includes:

1. Using data analysis techniques to uncover patterns and trends in data.
2. Hands-on experience working on real-world projects in AI and cloud computing.
3. Gaining practical skills that are valuable in today's job market.
4. Learning how to interpret data and make data-driven decisions
5. Creating visualizations and reports to communicate findings effectively.

4.WEEKLY OVERVIEW OF INTERNSHIP ACTIVITIES

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
1st WEEK	14/12/23	Thursday	Orientation & Registration for Internship
	15/12/23	Friday	Data analysis
	16/12/23	Saturday	IBM academic registration and IBM cloud registration

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
2nd WEEK	20/12/23	Wednesday	Hands on IBM Cloud
	21/12/23	Thusrday	Hands on IBM cloud and Regression analysis
	22/12/23	Friday	Lambda function and EDA

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
3rd WEEK	27/12/23	Wednesday	Text preprocessing in NLP
	28/12/23	Thursday	EDA on unstructured data and project ppt explanation
	29/12/23	Friday	Certification of two courses on IBM skills build
	30/12/23	Saturday	Q.A with Trainer about project
	31/12/23	Sunday	Project PPT & Project Video Explanation

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
4th WEEK	7/01/24	Wednesday	Project submission on Form
	9/01/24	Friday	Project Review
	17/1/24	Saturday	Internship Certificate

5.ABOUT INDUSTRY/ ORGANIZATION

Industry: Technology Education and Training

Organization: Edunet Foundation

Company Background and Activities:

Edunet Foundation is a prominent organization dedicated to providing comprehensive technology education and training programs to individuals seeking to advance their skills and knowledge in the digital domain.

Founded with the vision of bridging the gap between academic learning and industry demands, Edunet Foundation offers a diverse range of courses and workshops covering various emerging technologies, including data science, cybersecurity, artificial intelligence, and cloud computing.

The organization is committed to fostering innovation, creativity, and excellence in technology education, with a focus on equipping students with the practical skills and expertise needed to succeed in today's competitive job market.

Edunet Foundation's team comprises experienced educators, industry professionals, and technology experts who are passionate about empowering individuals to realize their full potential and achieve success in their chosen career paths.

Internship Place Details:

Address: IBM India Private Limited, Embassy Golf Links, Off Indira Nagar-Koramangala, Intermediate Ring Rd, Embassy Golf Links Business Park, Domlur, Bengaluru, Karnataka 560071

Contact Information: info@edunetfoundation.org

Website: https://edunetfoundation.org

6.SOFTWARE REQUIREMENTS SPECIFICATIONS

Software Requirement (Minimum): -

Operating system	Windows 10+
Python compiler	IDLE python
Application	Anaconda/Jupyter,Google colab and Watson Studio on IBM Cloud
Programming Language	Python

Hardware Requirement (Minimum): -

Processor	2.0GHz
RAM	8 GB
Hard Disk	512GB

Technologies used –:

- Jupyter Notebook
- Google Collab
- Visual Studio Code
- MS Excel
- Google Docs
- Google Slides
- GitHub
- Linked-In
- Kaggle

7.TECHNOLOGY

Methodology for Customer Churn Prediction in Banking:

1. Data Collection:

- Gathered relevant dataset of E-commerce customer purchasing, purchasing behavior , relationship between mobile applications and desktop sales.

- After gathering the data it's preprocessed to remove it's outliers values and fill the null values for better accuracy .

2. Exploratory Data Analysis (EDA):

- Implemented the graphical representations for more understanding the relation between features

- Pairplot , histplot, plotting straight line , Bar graph , etc. graphical representations are implemented.

3. Feature Engineering:

- Derived new features from existing ones to enhance the predictive power of the model.

4. Data Preprocessing:

- Performed data preprocessing techniques such as scaling numerical features, encoding categorical variables, and splitting the dataset into training and testing sets.

5. Model Selection:

- Evaluated different machine learning algorithms suitable such as Linear Regression, Random Forest and Support Vector Machines.

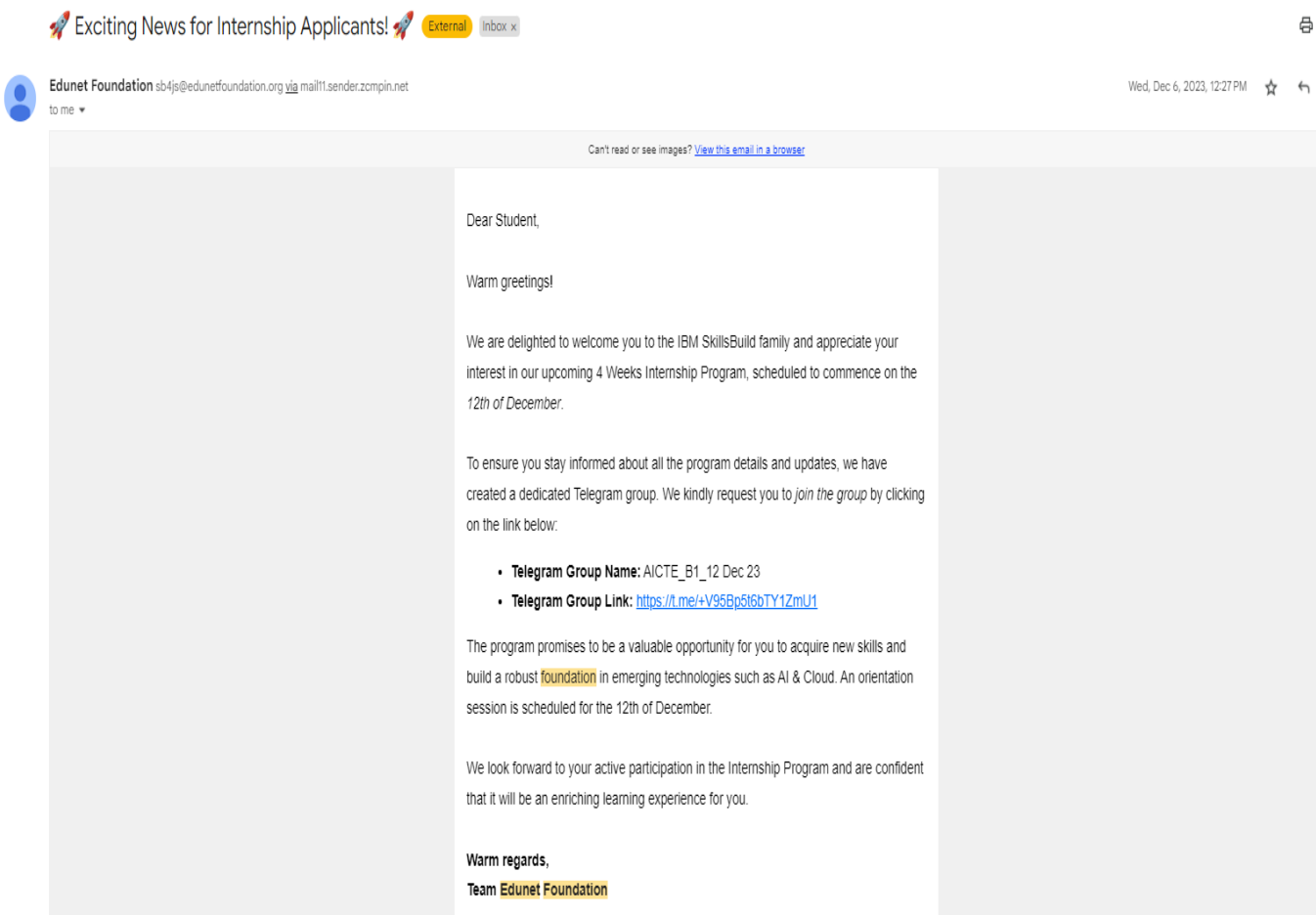
- Selected the most appropriate model based on evaluation metrics such as accuracy.

6. Model Training and Evaluation:

- Trained the selected model on the training dataset using cross-validation techniques to optimize performance.

- Calculate the Root mean absolute error , f1-score , mean absolute error , etc.

8.SCREENSHOTS



🚀 Exciting News for Internship Applicants! 🚀 External Inbox x

Edunet Foundation sb4js@edunetfoundation.org via mail11.sender.zcmpin.net
to me

Wed, Dec 6, 2023, 12:27 PM

Can't read or see images? View this email in a browser

Dear Student,

Warm greetings!

We are delighted to welcome you to the IBM SkillsBuild family and appreciate your interest in our upcoming 4 Weeks Internship Program, scheduled to commence on the *12th of December*.

To ensure you stay informed about all the program details and updates, we have created a dedicated Telegram group. We kindly request you to *join the group* by clicking on the link below:

- **Telegram Group Name:** AICTE_B1_12 Dec 23
- **Telegram Group Link:** https://t.me/+V95Bp5t6bTY1ZmU1

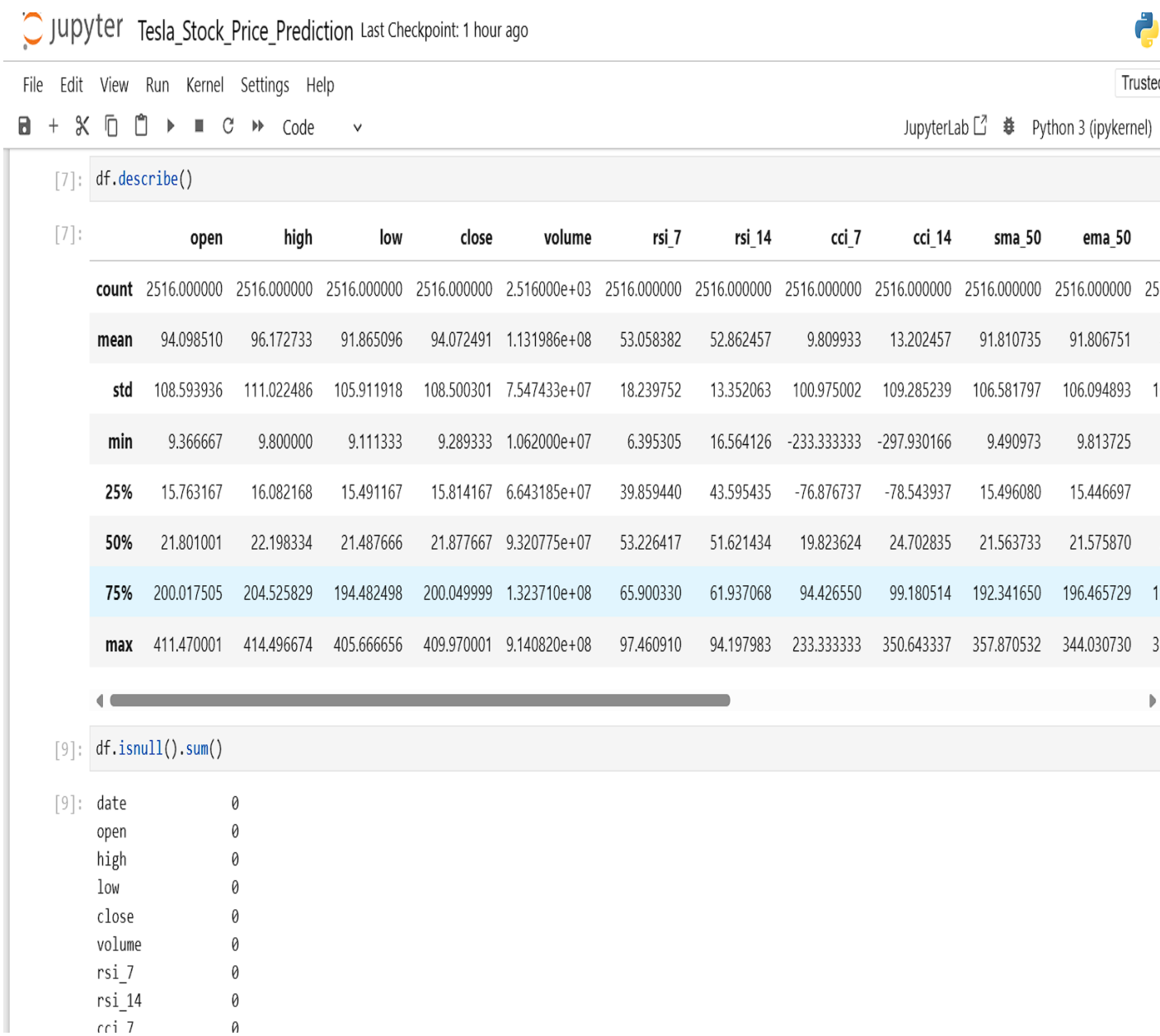
The program promises to be a valuable opportunity for you to acquire new skills and build a robust foundation in emerging technologies such as AI & Cloud. An orientation session is scheduled for the 12th of December.

We look forward to your active participation in the Internship Program and are confident that it will be an enriching learning experience for you.

Warm regards,
Team Edunet Foundation

Fig.8.1.Offer letter(Mail)

Project Screenshots:



	open	high	low	close	volume	rsi_7	rsi_14	cci_7	cci_14	sma_50	ema_50
count	2516.000000	2516.000000	2516.000000	2516.000000	2.516000e+03	2516.000000	2516.000000	2516.000000	2516.000000	2516.000000	2516.000000
mean	94.098510	96.172733	91.865096	94.072491	1.131986e+08	53.058382	52.862457	9.809933	13.202457	91.810735	91.806751
std	108.593936	111.022486	105.911918	108.500301	7.547433e+07	18.239752	13.352063	100.975002	109.285239	106.581797	106.094893
min	9.366667	9.800000	9.111333	9.289333	1.062000e+07	6.395305	16.564126	-233.333333	-297.930166	9.490973	9.813725
25%	15.763167	16.082168	15.491167	15.814167	6.643185e+07	39.859440	43.595435	-76.876737	-78.543937	15.496080	15.446697
50%	21.801001	22.198334	21.487666	21.877667	9.320775e+07	53.226417	51.621434	19.823624	24.702835	21.563733	21.575870
75%	200.017505	204.525829	194.482498	200.049999	1.323710e+08	65.900330	61.937068	94.426550	99.180514	192.341650	196.465729
max	411.470001	414.496674	405.666656	409.970001	9.140820e+08	97.460910	94.197983	233.333333	350.643337	357.870532	344.030730

fig.8.2(a).Data Preprocessing

[11]:
```
#Visualize the stocks of Tesla  performing from year 2014 to 2023.
plt.figure(figsize=(20, 10))
plt.plot(df.index, df['open'], label='Open')
plt.plot(df.index, df['close'], label='Close')
plt.plot(df.index, df['high'], label='High')
plt.plot(df.index, df['low'], label='Low')
plt.legend()
```

[11]:
```
<matplotlib.legend.Legend at 0x237aeae6c50>
```

Fig.8.2(b).Data visualization

Spliting Dataset for Training and Testing

```
[19]: from sklearn.model_selection import train_test_split
      X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=2)

[23]: print(X.shape,X_train.shape,X_test.shape,)

      (2516, 18) (2012, 18) (504, 18)

[21]: print(X_train,X_test)

      2016-07-22  52.708013  53.575524  -60.607939   13.971244   14.453640
      2020-02-20  75.528429  74.352347   81.929886  116.049156   37.138760
      2016-03-21  90.383736  78.298705  128.127301  157.042120   12.736227
      2022-02-03  39.319470  40.869472    9.713902  -50.735089  339.571332
      2015-01-20  20.339633  28.844051  -65.854370 -109.880208   14.970760
      ...               ...        ...         ...         ...         ...

[31]: print(Y.shape,Y_train.shape,Y_test.shape,)

      (2516,) (2012,) (504,)

[24]: print(Y_train,Y_test)

      date
      2016-07-22     15.334000
      2020-02-20     60.066666
      2016-03-21     15.616000
      2022-02-03    307.773346
      2015-01-20     13.104667
                       ...
      2016-09-06     13.447333
      2023-07-28    267.429993
```

Fig.8.2(c).Training and Testing of Data

Linear Regression

```
[31]: from sklearn.linear_model import LinearRegression
      lr=LinearRegression()

[32]: lr.fit(X_train,Y_train)
      lr_predict=lr.predict(X_test)

[33]: #For regression tasks, mean absolute error (MAE) and mean squared error (MSE) are common evaluation metrics.
      #You're already using mean_absolute_error, but you might also consider mean_squared_error for your regression evaluation.

[34]: from sklearn.metrics import mean_absolute_error
      lr_mean_abs_error=mean_absolute_error(Y_test,lr_predict)
      lr_mean_abs_error

[34]: 3.0322865822549976
```

Decision Tree Regressor

```
[25]: from sklearn.tree import DecisionTreeRegressor
      dr=DecisionTreeRegressor()

[26]: dr.fit(X_train,Y_train)
      dr_predict=dr.predict(X_test)

[29]: from sklearn.metrics import mean_absolute_error
      dr_mean_abs_error=mean_absolute_error(Y_test,dr_predict)
      dr_mean_abs_error

[29]: 4.463828085317462
```

Fig.8.2(d).Algorithm

K-Nearest Neighbors (KNN) regressor

```
[36]: from sklearn.neighbors import KNeighborsRegressor
      kr=KNeighborsRegressor()

[37]: kr.fit(X_train,Y_train)
      kr_predict=kr.predict(X_train)

[38]: from sklearn.metrics import mean_absolute_error
      kr_mean_abs_error=mean_absolute_error(Y_test,dr_predict)
      kr_mean_abs_error

[38]: 4.463828085317462
```

Random Forest Regressor

```
[40]: from sklearn.ensemble import RandomForestRegressor
      rfr = RandomForestRegressor()

[41]: rfr.fit(X_train, Y_train)
      rfr_prediction = rfr.predict(X_test)

[42]: rfr_mean_abs_error = mean_absolute_error(Y_test, rfr_prediction)
      rfr_mean_abs_error

[42]: 3.0709980662103136
```

Fig.8.2(e).Algorithm

Comparing The Mean Absolute Error Of Models

[73]:
```
Classifiers=['LinearRegression','Decision Tree Regressor','K-Nearest Neighbors regressor','Random Forest Regressor']
mean_abs_errors=[lr_mean_abs_error,dr_mean_abs_error,kr_mean_abs_error,rfr_mean_abs_error]
import matplotlib.pyplot as plt
plt.bar(Classifiers,mean_abs_errors,color=['cyan','red','green','magenta'])
plt.xticks(rotation=45)
plt.xlabel('Classifiers')
plt.ylabel('Mean_Absolute_Error')
plt.title('Comparing Mean Absolute Error Of Different Models ')
plt.show()
```

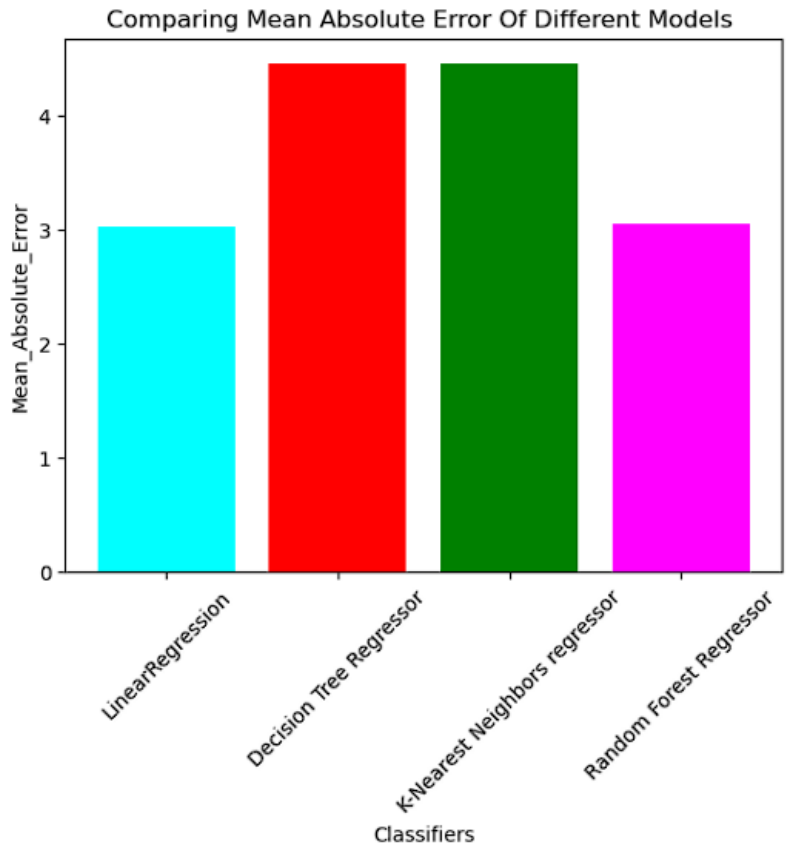


Fig.8.2(f).Result

9.RESULTS/ANALYSIS/INFERENCES

Results :

Based on the evaluation results, the Linear Regression algorithm achieved the best performance in predicting Tesla stock prices, with the minimum mean absolute error. The MAE values for each model:-

Linear Regression: 3.03
Decision Tree Regressor: 4.46
K-Nearest Neighbors Regressor: 4.46
Random Forest Regressor: 3.07

Analysis and Inference:

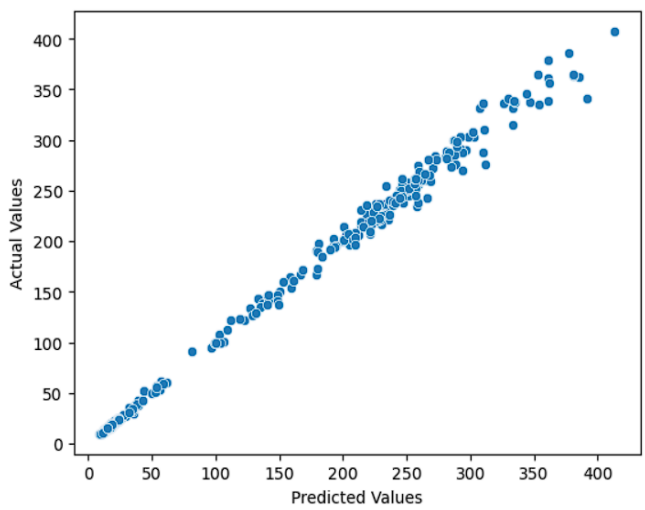
Among the four regression models analyzed for predicting Tesla stock prices, Linear Regression stood out with the lowest mean absolute error (MAE). Its specialty lies in its simplicity and ease of interpretation. Linear Regression is effective at capturing straightforward relationships between predictor variables and stock prices, making it valuable for understanding the factors influencing stock price movements.

While Linear Regression excels in simplicity, Decision Tree Regressor, K-Nearest Neighbors (KNN) Regressor, and Random Forest Regressor offer different strengths. Decision trees are good at capturing non-linear relationships, KNN is flexible and suitable for complex data patterns, and Random Forests are robust and effective for handling noisy data.

Linear Regrerssion Model Relationship Between Label and Predicted Value

[44]:
```
sns.scatterplot(x=lr_predict,y=Y_test)
plt.xlabel("Predicted Values")
plt.ylabel('Actual Values')
```

[44]:
```
Text(0, 0.5, 'Actual Values')
```



10.CONCLUSION AND FUTURE SCOPE

CONCLUSION:

In conclusion, the analysis of various regression models for predicting Tesla stock prices indicates that Linear Regression performed the best, displaying the lowest error in its predictions. This straightforward model is easy to interpret and understand, making it a suitable choice for stock price prediction. Nonetheless, other models such as Decision Tree, K-Nearest Neighbors, and Random Forest possess their own strengths and may excel in capturing more intricate patterns within the data. Looking ahead, further exploration could focus on expanding the range of features utilized, integrating predictions from multiple models, and maintaining continuous updates to enhance prediction accuracy and enable more informed investment decisions in the future.

FUTURE SCOPE:

1)Feature Expansion: Explore additional factors like economic indicators, industry news sentiment, and competitor performance to gain deeper insights into Tesla stock price movements.

2)Ensemble Modeling: Combine predictions from multiple regression models using techniques like stacking or blending to improve overall predictive accuracy.

3)Time-Series Analysis: Utilize advanced time-series techniques such as ARIMA or seasonal decomposition to better capture patterns and dependencies in Tesla stock prices over time.

4)Sentiment Analysis Integration: Incorporate sentiment analysis of news articles, social media posts, and analyst reports to gauge market sentiment and its impact on Tesla stock prices.

5)Continuous Monitoring and Updating: Implement a system to continuously monitor and update predictive models with new data to adapt to changing market conditions, ensuring ongoing relevance and effectiveness.

11. ATTENDANCE RECORD

Name & Address of Organization: -
Vidya Pratishtha's Kamalanayan Bajaj
Institute of Engineering and Technology
Baramati - 413102

Name of Student	Mr. Prathamesh Santosh Pagale
Roll. No	2337004(**T190352064**)
Name of Course	TE - Artificial Intelligence & Data Science
Date of Commencement of Training.:	14/12/23
Date of Completion of Training:	15/01/24

Initials of the student:

Month& Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Jan-2023																															

Note:

1. Attendance Sheet should remain affixed in Daily Training Diary. Do not remove or tear it off.

2. Students should sign/initial in the attendance column. Do not mark 'P'

3. Holidays should be marked in red ink in the attendance column. Absent should bemarked as 'A' in Red Ink**.**

Signature of Company internship supervisor /Dept. Mentor with company stamp/ seal

(Name_____) Contact No.

12.BIBLOGRAPHY

List of references:-

1) https://www.sciencedirect.com/science/article/abs/pii/S0950061819324420
2)https://www.javatpoint.com/machine-learning-algorithms
3) https://www.geeksforgeeks.org/machine-learning/
4) https://youtu.be/JxgmHe2NyeY
5) https://docs.jupyter.org/en/latest/
6) https://colab.research.google.com/
7) https://youtu.be/RLYoEyIHL6A
8) https://youtu.be/FvPJmBoBH1A
10) https://www.kaggle.com/
11) https://github.com/