Modern Education Society College of Engineering,pune-411001

DEPARTMENT OF ELECTRONICS & TELECOMMUNICATION

INTERNSHIP PRESENTATION

By Rutuja Borawake
T.E. A (E&TC) R.No-61
Seat no-F190323018

LETS GROW MORE
DATA SCIENCE INTERN
MARCH 2023

ABOUT COMPANY

LetsGrowMore is a ground-based organisation that aims at building the future through nourishing the present. We at LetsGrowMore believe in making our youth especially the students self-aware and exploring the untouched world of technology and tremendous growth-making fields and our belief finally took us where we are standing today. Today we are an officially MSME registered start-up.

LETS GROW MORE VIRTUAL INTERNSHIP PROGRAM

LGMVIP is a 4-week Virtual Internship Program where you are provided internship opportunities for beginners/students who which to excel their career in various domains such as Web Development, Data Science, Campus Ambassadors, Technical Content Writer, etc.
It is a great initiative by AMAN KESARWANI sir(Founder of lets grow more)

BADGE


Lets
GrowMore
VIRTUAL
INTERNSHIP PROGRAM
CONGRATULATIONS
You have been selected!
DATA SCIENCE INTERN

OFFER LETTER

Lets GrowMore

Offer Letter

20 February 2023

Congratulations Rutuja Borawake !!

We would like to congratulate you on being selected for the **"Data Science Intern"** Internship position with **LetsGrowMore**, effective from **"1 March 2023"**. All of us at **LetsGrowMore** are excited that you will be joining our team! We hope you are elevated to start this innovational journey with us.

This Internship is viewed by **LetsGrowMore** as being an educational opportunity for you. As such, your internship will include orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts through hands-on application of the knowledge you learned in class. And, you will find yourself adjoining with numerous opportunities to refine and flaunt your skills.

While performing the internship, you acknowledge your obligation to perform all work allocated to you to the best of your ability and comply with all lawful and reasonable directions and instructions given to you. We look forward to an abiding and fruitful association with you and are sure that you will look back at your engagement with us as a gratifying experience.

Wishing you all the best!

Warm Regards,

Aman Kesarwani

Aman Kesarwani
Founder

MSME

CID- LGMVIPDSWL0011185

Verify Here

/letsgrowmore letsgrowmore.in info@letsgrowmore.in letsgrowmore.in

INTRODUCTION

Internship is a platform where we get an opportunity to explore the fields of our interests at an industrial level and professional environment. During this 4-week internship, I had great experience as DATA SCIENCE intern at LETS GROW MORE I explored many new concepts and technologies. I have also developed important technical as well as non-technical skills. In the presented report, I have put down all the things that I learnt through the course of my of internship

OBJECTIVE

Joining and then completing an internship in Data Science can help you in a lot of ways along your path to becoming a Data Scientist, a data analyst or a data engineer. Internships serve as proof of your accomplishments and your foundational abilities. Through your internship and the projects you have worked on, employers can find out your capabilities and how well you fit inside a Data Science process or a pipeline. Also, without an internship, it is almost impossible to get jobs as a Data Scientist.

Domain-Data science
Duration-4 weeks
Software used-Google Colab

TASK LIST PROVIDED BY LETS GROW MORE

1.Iris Flower classification using Ml
2.Stock Market Prediction
3.Image to pencil sketch
4.Develop Neural Network to read handwriting
5.Prediction using Decision tree
6.Exploratory data analysis on dataset-Terrorism
7.Music Recommendation
8.Next Word Prediction
9.Handwritten equation using cnn
10..ML face recognition to detect mood and suggest songs accordingly

Out of all task i have completed 8 tasks in my internship

TASK-Image to pencil Sketch

LETS GROW MORE VIRTUAL INTERNSHIP PROGRAM

Data Science Intern

Author-Rutuja Borawake

Level-Beginner

Task name-Image to Pencil Sketch with Python

```
import cv2
import matplotlib.pyplot as plt
```

```
from google.colab import files
upload=files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the curren
browser session. Please rerun this cell to enable.

```
image = cv2.imread("Frangipani rust_ what it is and how to treat it.jfif")
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
plt.figure(figsize=(8,6))
plt.imshow(image)
plt.axis('off')
plt.title('Original Image')
plt.show()
```

Original Image

```
gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
plt.figure(figsize=(8,6))
plt.imshow(gray_image, cmap='gray')
plt.axis('off')
plt.title('Gray Image')
plt.show()
```

Gray Image

```
inverted_image = 255 - gray_image
plt.figure(figsize=(8,6))
plt.imshow(inverted_image,cmap='gray')
plt.axis('off')
plt.title('Inverted image')
plt.show()
```

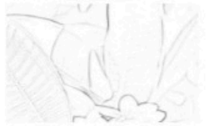
Inverted image

```
blurred_image = cv2.GaussianBlur(inverted_image, (21, 21), 0)
plt.figure(figsize=(8,6))
plt.imshow(blurred_image,cmap='gray')
plt.axis('off')
plt.title('Blurred Inverted image')
plt.show()
```

Blurred Inverted image

```
inverted_blurred = 255 - blurred_image
pencil_sketch = cv2.divide(gray_image, inverted_blurred, scale=256.0)
plt.figure(figsize=(8,6))
plt.imshow(pencil_sketch,cmap='gray')
plt.axis('off')
plt.title('Pencil Sketch')
plt.show()
```

Pencil Sketch



```
image = cv2.imread("Frangipani rust_ what it is and how to treat it.jfif")
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
plt.figure(figsize=(8,6)).add_subplot(1, 2, 1)
plt.imshow(image)
plt.axis('off')
plt.title('Original Image')
pencil_sketch = cv2.divide(gray_image, inverted_blurred, scale=256.0)
plt.figure(figsize=(8,6)).add_subplot(1, 2, 2)
plt.imshow(pencil_sketch,cmap='gray')
plt.axis('off')
plt.title('Pencil Sketch')
plt.show()
```

Original Image

Pencil Sketch

Task -Iris Flower classification using ML

LETS GROW MORE VIRTUAL INTERNSHIP PROGRAM

Data Science Internship

Author-Rutuja borawake

Task No.1

Level-Beginner

Task Name:- Iris Flowers Classification ML Project

This particular ML project is usually referred to as the "Hello World" of Machine Learning. The iris flowers dataset contains numeric attributes, and it is perfect for beginners to learn about supervised ML algorithms, mainly how to load and handle data. Also, since this is a small dataset, it can easily fit in memory without requiring special transformations or scaling capabilities.

Importing Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Dataset

```
from google.colab import files;
upload=files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving iris.data.csv to iris.data.csv

```
df=pd.read_csv("iris.data.csv")
```

```
df.head()#top 5 values
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

```
df.tail() #last 5 values
```

	sepal_length	sepal_width	petal_length	petal_width	species
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

```
df.shape #no. of rows and columns
```

```
(150, 5)
```

```
df.isnull() #returns a Dataframe object where all the values are replaced with a boolean, True for null,otherwise false
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
...	...	...	...	...	...
145	False	False	False	False	False
146	False	False	False	False	False
147	False	False	False	False	False
148	False	False	False	False	False
149	False	False	False	False	False

150 rows × 5 columns

```
df.isnull().sum()#returns no of missing values
```

```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
species         0
dtype: int64
```

```
df.describe() # used to view statistical details
```

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

```
df.columns
```

```
Index(['sepal_length', 'sepal_width', 'petal_length', 'petal_width',
       'species'],
      dtype='object')
```

```
df.nunique()
```

```
sepal_length    35
sepal_width     23
petal_length    43
petal_width     22
species          3
dtype: int64
```

```
df.species.nunique()
```

```
3
```

```
df.species.value_counts()
```

```
Iris-setosa        50
Iris-versicolor    50
Iris-virginica     50
Name: species, dtype: int64
```

```
df.max()
```

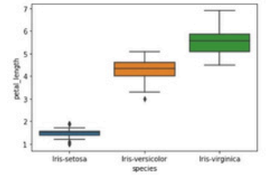
```
sepal_length             7.9
sepal_width              4.4
petal_length             6.9
petal_width              2.5
species          Iris-virginica
dtype: object
```

```
df.min()
```

```
sepal_length             4.3
sepal_width              2.0
petal_length             1.0
petal_width              0.1
species          Iris-setosa
dtype: object
```

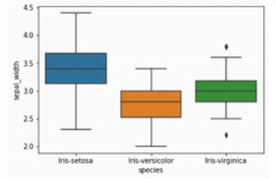
Visualization

```
#boxplot
sns.boxplot(x="species",y='petal_length',data=df)
plt.show()
```




```
sns.boxplot(x="species",y="sepal_width",data=df)
```

```
<AxesSubplot:xlabel='species', ylabel='sepal_width'>
```



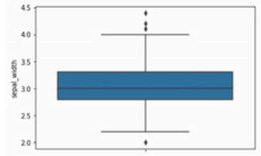

```
sns.boxplot(x="species",y="petal_width",data=df)
```

TASK -Iris Flower Classification

```
<AxesSubplot:xlabel='species', ylabel='petal_width'>
```

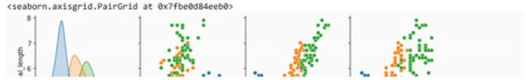
```
sns.boxplot(y="sepal_length",data=df);
plt.show()
```

```
sns.boxplot(y="sepal_width",data=df);
plt.show()
```



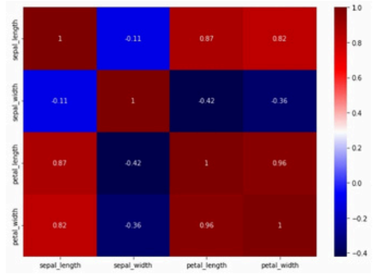
```
sns.pairplot(df,hue="species") #Pairplot is a pairwise relationship in dataset
```

```
<seaborn.axisgrid.PairGrid at 0x7fbe0d84eeb0>
```



Data Preprocessing

```
#heatmap is used to show data in graphical format
plt.figure(figsize=(10,7))
sns.heatmap(df.corr(),annot=True,cmap="seismic")
plt.show()
```



Label Encoder

```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
```

```
df['species']=le.fit_transform(df['species'])
df.head()
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

```
X=df.drop(columns=['species'])
y=df['species']
X[:5]
```

```
sepal_length  sepal_width  petal_length  petal_width
y[:5]

0    0
1    0
2    0
3    0
4    0
Name: species, dtype: int64
```

Splitting the dataset into train and test **

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X, y, test_size=0.3, random_state=1)
```

Selecting the models and metrics

```
from sklearn.linear_model import LogisticRegression
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.naive_bayes import GaussianNB
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score

lr=LogisticRegression()
knn=KNeighborsClassifier()
svm=SVC()
nb=GaussianNB()
dt=DecisionTreeClassifier()
rf=RandomForestClassifier()
```

Training and evaluating the models

```
models=[lr,knn,svm,nb,dt,rf]
scores=[]

for model in models:
  model.fit(X_train,y_train)
  y_pred=model.predict(X_test)
  scores.append(accuracy_score(y_test,y_pred))
  print("Accuracy of"+ type(model).__name__+"is",accuracy_score(y_test,y_pred))
```

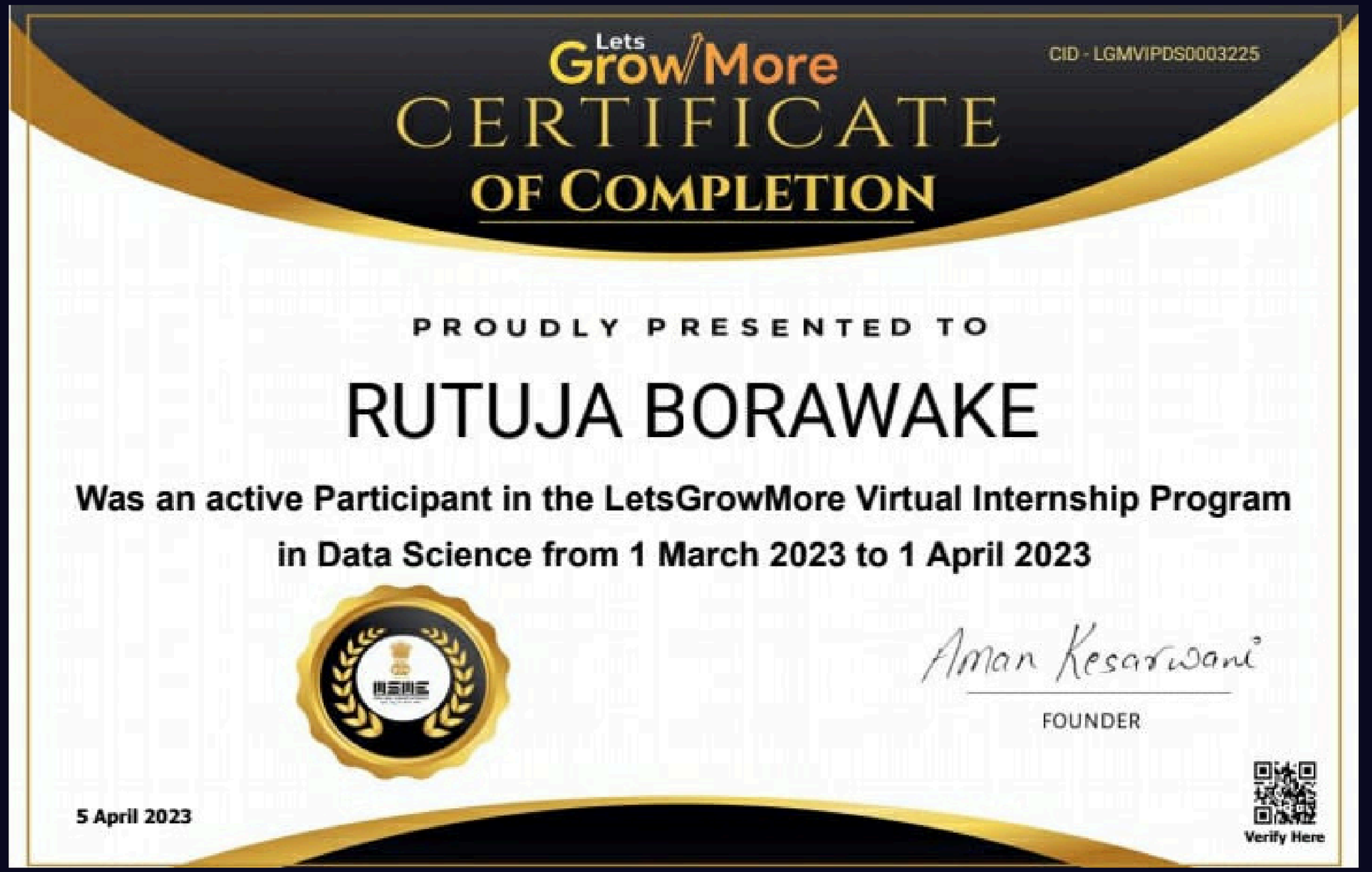
```
/usr/local/lib/python3.8/dist-packages/sklearn/linear_model/_logistic.py:458: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
Accuracy ofLogisticRegressionis 0.9777777777777777
Accuracy ofKNeighborsClassifieris 0.9777777777777777
Accuracy ofSVCis 0.9777777777777777
Accuracy ofGaussianNBis 0.9333333333333333
Accuracy ofDecisionTreeClassifieris 0.9555555555555556
Accuracy ofRandomForestClassifieris 0.9555555555555556
```

Conclusion

In conclusion, I am happy with my data science internship so far and I learned a lot of new things. . I was also exposed to some new technology at my workplace that I want to implement in my workflow very soon such as Docker. Work is very different from university and therefore, an internship as a data scientist is a great learning experience. I encourage everyone to do it!
This was my first internship and had best experience and exposure to many data science concepts

COMPLETION CERTIFICATE



THANK YOU