

WEEKLY REPORT

Work done in last week (Attach supporting Documents):

1. Started solving the tasks provided by the company in collaboration with PW Skills.
2. Daily reporting to the manager regarding the progress made in the Data Science Bootcamp.
3. Attended the daily and weekly team meetings.
4. Reporting the team leader on daily basis and discussing the progress.
5. Completed the data pre-processing phase.
6. Submitted the data pre-processing report.
7. Submitted the PW report to the manager.
8. Assisted the CRM Team for 3 days.
9. Submitted the mini-project and pitched the mini-project to the team.

Reason for incomplete work:

Meeting with the manager postponed due to his unavailability.

Plans for next week:

1. To attend the daily and weekly team meetings.
2. To complete the daily practice tasks provided by the team leader.
3. To Report the team leader on daily basis and discussing the progress.
4. Understand and start working on the new project assigned.
5. Attend the team meeting to understand the analysis phase of the project.
6. Assist the CRM team with the customer query
7. Start working on the analysis phase.
8. Complete the project and submit it to the manager.
9. Meeting with the manager for further portfolios and update.

Reporting No: Prof. Dipak Ramoliya
From: 27 / 03 / 2023 To: 31 / 03 / 2023
Project ID: PRJ2022DCS096
Project Title: Data Science & Management-Collab-Team-003,001

Week No:15

Mini-Project Details:

Classify from the data provided, whether the voice is of male or female. Use ML Algorithms specified by the teams, compare the accuracy and give your opinion.

	meanfreq	sd	median	Q25	Q75	IQR	skew	kurt	sp.ent	sfm	...	centroid	meanfun	minfun	maxfun	meandom	mindom	ma
0	0.059781	0.064241	0.032027	0.015071	0.090193	0.075122	12.863462	274.402906	0.893369	0.491918	...	0.059781	0.084279	0.015702	0.275862	0.007812	0.007812	0.00
1	0.066009	0.067310	0.040229	0.019414	0.092666	0.073252	22.423285	634.613855	0.892193	0.513724	...	0.066009	0.107937	0.015826	0.250000	0.009014	0.007812	0.00
2	0.077316	0.083829	0.036718	0.008701	0.131908	0.123207	30.757155	1024.927705	0.846389	0.478905	...	0.077316	0.098706	0.015656	0.271186	0.007990	0.007812	0.00
3	0.151228	0.072111	0.158011	0.096582	0.207955	0.111374	1.232831	4.177296	0.963322	0.727232	...	0.151228	0.088965	0.017798	0.250000	0.201497	0.007812	0.50
4	0.135120	0.079146	0.124656	0.078720	0.206045	0.127325	1.101174	4.333713	0.971955	0.783568	...	0.135120	0.106398	0.016931	0.266667	0.712812	0.007812	5.40

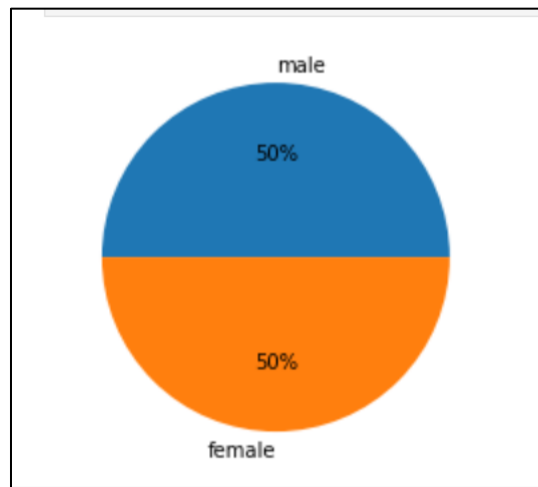
5 rows × 21 columns

Data Set preview

```
meanfreq    0
sd           0
median      0
Q25         0
Q75         0
IQR         0
skew        0
kurt        0
sp.ent      0
sfm         0
mode        0
centroid    0
meanfun     0
minfun      0
maxfun      0
meandom     0
mindom      0
maxdom      0
dfrange     0
modindx     0
label       0
dtype: int64
```

Column Details

Student Id: 19DCS098
Student Name: Parth Niteshkumar Patel



Pre-known details

MOST ACCURATE MODEL FOR THE GIVEN DATA-SET:

- ACCURACY OF DECISION TREE CLASSIFIER : **97%**
- ACCURACY OF RANDOM FOREST CLASSIFIER : **98.4%**
- ACCURACY OF KNN CLASSIFIER : **66.4%**
- ACCURACY OF LOGISTIC REGRESSION : **98.7%**
- ACCURACY OF SVM CLASSIFIER : **97.4%**

• HERE, THE DATASET HAS EQUAL NUMBER OF MALES AND FEMALES AS VALUES IN LABEL COLUMN.

- AS A RESULT, PROBABILITY OF OVERFITTING EVEN IF ACCURACY IS HIGH CAN BE IGNORED.
- ACCORDING TO ME, EXCEPT FOR KNN, ALL MODELS ARE HAVING ACCURACY IN A CLOSE PROXIMITY.
- **HOWEVER, RANDOM FOREST AND LOGISTIC CAN BE USED AS BOTH HAVE ALMOST SAME ACURACY, BUT, LOGISTIC REGRESSION HAS TRAINING SCORE OF 1.0 WHICH SOMETIMES LEAD TO UNREALISTIC RESULTS.**
- AS A RESULT, FOR A SAME SIDE, I WILL GO WITH **RANDOM FOREST CLASSIFIER**

My opinion on the Machine Learning Models

Student Id: 19DCS098

Student Name: Parth Niteshkumar Patel

Reporting No: Prof. Dipak Ramoliya
From: 27 / 03 / 2023 To: 31 / 03 / 2023

Week No:15

Project ID: PRJ2022DCS096

Project Title: Data Science & Management-Collab-Team-003,001

References:

1. <https://learn.pwskills.com/>
2. <https://www.w3schools.com/datascience/>
3. <https://www.geeksforgeeks.org/data-science-tutorial/>
4. <https://www.kaggle.com/>
5. <https://colab.research.google.com/>



Signature of External Guide

Signature of Internal Guide

Student Id: 19DCS098

Student Name: Parth Niteshkumar Patel