

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1) NAME: RAHUL KUMAR GUPTA

Email : rahulgupta41298@gmail.com

CONTRIBUTION:

- 1) DATA PREPERATION**
- 2) DATA CLEANING**
- 3) EDA**
- 4)ML ALGORITHM**
- 5) REPRORT MAKING(TECHNIQUAL DOCUMENTAION)**

2) NAME: AAFIAN BIN NISHANT

Email : affanbn.ind@gmail.com

CONTRIBUTION:

- 1) RESEARCH PAPER ANALYSIS**
- 2) DATA CLEANING**
- 3) EDA**
- 4)ML ALGORITHM**
- 5) SUMMARY PREPERATION**

3) NAME: AMISHA KHARE

Email : amishakhare2000@gmail.com

CONTRIBUTION:

- 1) RESEARCH PAPER ANALYSIS**
- 2) DATA CLEANING**
- 3) EDA**
- 4)REPORT MAKING**
- 5) SUMMARY PREPERATION**

4) NAME: RAHUL THUMAR

Email : thumarrahul@gmail.com

CONTRIBUTION:

- 1) RESEARCH PAPER ANALYSIS**
- 2) DATA CLEANING**
- 3) EDA**
- 4)REPORT MAKING**
- 5) SUMMARY PREPERATION**

Please paste the GitHub Repo link.

Github Link:

<https://github.com/RAHULGUPTA41298/Email-Campaign-Effectiveness>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Email Campaign Effectiveness

In the data set of Email Campaign Effectiveness, there are total 68353 row and 12 columns are present there, in today world email marketing is one of the main marketing are here we will see in many mail sent how many of them read, ignored and acknowledge with the help of EDA, ML models. In EDA, we analyzed the customer location feature we can find that irrespective of the location, the percentage ratio of emails being ignored, read and acknowledged are kind of similar. It does not exclusively influence our target variable. It would be better to not consider location as a factor in people ignoring, reading or acknowledging our emails. In the Email Campaign Type feature, it seems like in campaign type 1 very few emails were sent but has a very high likelihood of getting read. Most emails were sent under email campaign type 2 and most ignored. Seems like campaign 3 was a success as even when less number of emails were sent under campaign 3, more emails were read and acknowledged. Time email sent category cannot be considered as a relevant factor in classifying the emails. Both feature importance showed this particular thing. If we consider Time email sent category 2 as middle of the of course they are going to be read and acknowledged more than morning and night. Analyzing total past communications, we can see that the more the number of previous emails, the more it leads to read and acknowledged emails. This is just about making connection with your customers. The more the words in an email, the more it has a tendency it has to get ignored. Too lengthy emails are getting ignored. More images were there in ignored emails. There are outliers in almost every continuous variable except Word Count and upon analyzing, it was found that outliers make up for more than 5% of the minority data and will influence the results either way, so it was better not to get rid of them. Upon this in-depth exploratory data analysis and feature importance of various good models, we came to conclusion and mentioned some factors that are leading to a higher number of ignored emails and accordingly here are some recommend. Email Campaign Type 1 and 3 are doing better than 2. So, focusing on improving 2, can do the trick. The word count should be reasonable. The content should be crisp and to the point with a few marketing gimmicks. The number of images and links should be kept in check. Total past communications had a positive influence, hence having a healthy relationship with customers is a big yes.

