

We got information from an education company called X Education, which sells online courses to business professionals. When these people enter the site, they can browse courses, fill out a form or watch some videos. When those people filled out the form to provide their email address or phone number,

were classified as administrators. Although X Education has recruited many leaders, the transition has not been smooth.

We searched the scientific literature on these documents to confirm probability.

We use python libraries such as pandas and matplotlib for this. First, we cleaned files by removing unnecessary lines to identify duplicates and redundant values.

We then analysed the data using matplotlib and the seaborn library. We performed univariate analysis of categorical system and bivariate analysis of number. From these analyses, we found that newsletter or email recipients had more than

exchange exchanges. From the heatmap, we determined a correlation between total hours spent on site and change; however, page views of visits and total visits were not attributed to conversions.

Our next step is the machine learning model. We split the data into training and test data, making up 70% of the training data. Then we did RFE (Recursive Feature Elimination) and used logistic regression. The final model was obtained after resampling and modifying the process 14 times.

Features of the final model were obtained and the obtained accuracy was 83%.

From these results we draw some conclusions, such as our logistic regression model, and compared with the model used by PCA (Principal Component Analysis). To improve overall conversion, X Education should increase

user engagement on websites as it facilitates conversion, increase sending of SMS notifications as it helps increase conversions, increase "total visits" as it helps increase conversion rates