IMPACTSURE ASSIGNMENT

Tejas R. Nachan

219045

37

BUSINESS PROBLEM

WE ARE GIVEN A BIG ONLINE RETAIL COMPANY "AMAZING ZONE" WHICH SELLS VARIOUS PRODUCTS FROM
THEIR PORTAL. THE RETAILER GIVES A GUARANTEE THAT THE PRODUCTS PURCHASED WILL BE DELIVERED WITHIN THE
TIME GIVEN. THE LAST MILE OF THE SHIPMENT PROCESS IS DONE BY SEVERAL DELIVERY BOYS WHICH ARE
WORKING IN THIS COMPANY.

THESE DELIVERY BOYS ARE OFTEN OVERLOADED WITH WORK BECAUSE OF THE IMPRESSION THAT THEY HAVE TO
KEEP AS COMPANY WITH THE DELIVERY TIME. DUE TO THE WORKLOAD, THE DELIVERY BOYS ARE NOT ABLE TO
WORK WELL TO THEIR FULL POTENTIAL WHICH OFTEN RESULTS IN THEIR ABSENTEEISM FROM WORK. THIS IN TURN
INCREASES THE WORKLOAD ON ALL THE OTHER DELIVERY BOYS.

WE HAVE TO SUGGEST WAYS IN WHICH WE CAN OPTIMIZE THE DELIVERY PERFORMANCE. WE ALSO HAVE TO
MAKE A MODEL TO HELP DETERMINE THE HOURS FOR AN EMPLOYEE.

To solving this we will use Cross Industry Standard Process for Data Mining (CRISP-DM) is a process model that serves as the base for a data science process. It has six sequential phases:

1)Business understanding – What does the business need

2)Data understanding – What data do we have / need? Is it clean

3)Data preparation – How do we organize the data for modeling

4)Modeling – What modeling techniques should we apply

5)Evaluation – Which model best meets the business objectives

6)Exploratory Data Analysis

MODELLING AND EVALUATION

- MODELLING: After doing preprocessing, EDA and splitting. our data is ready for modelling. we will use Linear regresssion, Stocastic Gradient Descent and Random Forest Regressor, then model which perform better we will do hyperparameter on that model and further increase the score.

- EVALUATION: To measure the score of a model we will use Mean squared error, Root mean sqaured error and R2 which measure how variation in dependent variable is explain by independent variable.

- 1) Linear Regression Model   2) Stochastic Gradient Descent Model   3) Random Forest Model

>MSE: 99.12343560674795
RMSE: 9.956075311424073
R2: 0.0903912367487486

>MSE: 95.13865026438425
RMSE: 9.753904360018312
R2: 0.12695772221106638

>MSE: 148.31909173973077
RMSE: 12.17863258907099
R2: -0.361053970517756

OBSERVATIONS

- In March or February, the majority of employees are absent from work more frequently. This can be because the kids are having their final examinations at this time, and the parents want to be there to help them study.

- Employees that have 0–1 or 2 children have missed more work than other employees. They might be newlyweds who want to spend some time with their spouse, which may be the reason for this. Additionally, it's possible that they have to care for their kids because they are parents.

- The cost of transportation has an impact on the workers' absenteeism as well.

SOLUTIONS

- Company should hire more employees in the company.

- The company should take up the transportation expenses for the employees.

- They can start a stress relief program in the company for the employees to relax.