

LEARNING

Total Time Spend
in Learning (video
watched)

28 hours

Total Time in
Completion of
skills

8 weeks

Extra Skill Score
Through
Community Help

39

5

No. of Question
Asked

49

No. of Question
Answered

SKILLS	SCORE	OUT OF
BASIC STATISTICS	50	50
STATISTICS - ADVANCE	46	50
ANALYTICS BASICS	50	50
ADVANCE ANALYTICS	48	50
MACHINE LEARNING	50	50
ADVANCE CONCEPTS	51	51
TABLEAU	48	50
MONGODB	50	50
TOTAL	393	401

JOB READINESS ON TECHNICAL SKILLS

The Candidate has build up the work portfolio working on industry standards, projects which has been professionally reviewed and evaluate by Sr. industry professionals.

1

Churn reduction

86 / 100

Mentor Details

Ravi Shekhar
Data Scientist

<https://www.linkedin.com/in/ravi-shekhar-8b41a95a/>

Problem Statement

Churn (loss of customers to competition) is a problem for companies because it is more expensive to acquire a new customer than to keep your existing one from leaving. This problem statement is targeted at enabling churn reduction using analytics concepts.

Area of Strength

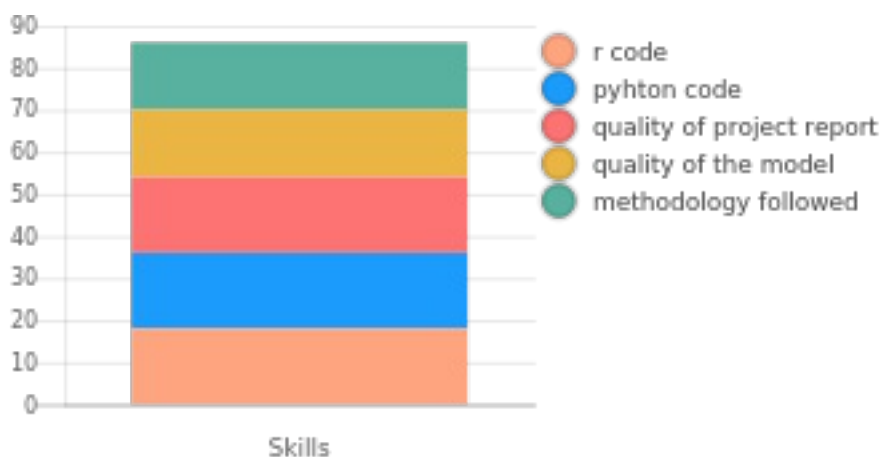
Good understanding of end to end modelling approach. Steps right from data analysis using descriptive analysis involving bivariate and uni-variate analysis done properly. Missing/outlier/multi-colinearity check done and best part is that rather than just logistic model, various other ML models have been implemented and best is then selected.

Area of Weakness

While doing multi-colinearity right now pairwise correlation is done wherein problem is multivariate analysis My suggestion is to try and explore VIF since it handles multivariate and based on VIF values select best variables Do try variable reduction techniques which can get rid of noisy variables even in the beginning. Outlier treatment was missing. You can also create Gain/Lift chart along with the other stats for model validation like concordance, ks value.

Comments

Good work and specially like your end to end PDF report with proper steps and detailed output shown. As next step, you can optimize your models by selecting best hyper-parameters and then improve model performance. Also you need to check some more stats during model validation



Mentor Details

Muquayyar Ahmed

Data Scientist

<https://www.linkedin.com/in/muquayyar-ahmed-04578238>

Problem Statement

Employee Absenteeism

Area of Strength

Data pre processing, Model development, Error metrics, python

Area of Weakness

R code

Comments

Analysis looks fine to me. Below are my detailed feedback 1. Found the detailed view of data and analysis 2. Never include code in between the report. You can add appendix at the bottom 3. Model parameters are as per expectation 4. Applied all relevant pre processing techniques 5. Coding standards need to improve. There are lot of development codes which need not required and also comments are missing. 6. Good in ML algorithms and techniques. Justified the work done in this project 7. I suggest to work on GBM, XgBoost, SVM

