HUMAN RESOURCE ANALYTICS

Who are likely to join the company ?

INNOMATİCS
RESEARCH LABS

INNOMATICS
RESEARCH LABS

Problem Statement

The dataset hr_data.csv contains sample of candidates that we part of a recruitment process of particular client of ScaleneWorks. ScaleneWorks supports several information technology (IT) companies in India with their talent acquisition. One of the challenge they face is about **30% of the candidates who accept the jobs offers, do not join the company**. This lead to huge loss of revenue and time as the companies initiate the recruitment process again to fill the workforce demand. ScaleneWorks want to find out if a model can be build to predict the likelihood of a candidate joining the company. If the likelihood is high, then the company can go ahead and offer the jobs to the candidates.

INNOMATICS
RESEARCH LABS

DataSet Description

Candidate - Reference number; it is a unique to identify the candidate

DOJextended - Binary variable identifying candidate asked for date of joining extension (yes/no)

Duration to accept the offer- Number of days taken by the candidate to accept the offer (Scale variable)

Notice Period - Notice period to be served in the parting company before candidate can join this company

Offered band - Band offered to the candidate based on experience, performance in the interview rounds (C0/C1/C2/C3/C4/C5/C6)

INNOMATICS
RESEARCH LABS

Dataset Description cond...

Percentage hike expected - Percentage hike expected by the candidate (Scale variable)

Percentage hike offered - Percentage hike offered by the company (Scale variable)

Joining bonus - Binary variable indicating if joining bonus was given or not (yes/no)

Gender - Gender of the candidate (male/female)

Candidate source - Source form which resume of the candidates was obtained (Employee referral/Agency/Direct)

Dataset Description cond...

REX (in yrs) - Relevant years of experience of the candidate for the position offered

LOB - Line of business for which offer was rolled out

Date of Birth - Date of birth of the candidate

Joining location - Company location for which the offer was rolled out for the candidate to join

Candidate relocation status - Binary Variable indicating whether the candidate has to relocate from one city to another city for joining (yes/no)

HR Status - Final joining status of the candidate (joined/ no joined)

INNOMATICS RESEARCH LABS

Task - 1

1. Build a logistic Regression model to predict the probability of the candidate joining the company.
   - *Not Joined - 1 (positive case)*
   - *Joined - 0 (negative case)*
2. Find the significant features form the above model and build another logistic regression model with only the significant features
3. Use the following cost to find optimal cut-off probability to determine if a candidate will join or not
   - *Cost of predicting "Not Joining" as "Joining" (FPs) cases is 3 times more than predicting "Joining" as "Not Joining" (FNs)*
4. Build a confusion matrix based on the cut-off probability found in question 3 and report the precision and recall of the model for joining cases

INNOMATICS
RESEARCH LABS

Task - 2

HR Wants to understand the **key parameters** affecting the joining of candidates. So, build a *decision tree* with optimal parameters and provide some rules to HR for building strategies to ensure candidate offer job most likely will the company in future

INNOMATICS
RESEARCH LABS

Submission

---

1. Create seperate jupyter notebooks and document properly with markdown.
2. Prepare only 5-10 slides powerpoint presentation on HR Analytics.
3. Upload jupyter files and ppt(pdf version) in Github
4. Post an article about **HR Analytics** in Linkedin with this dataset with *Innomatics Research Labs Logo*.
5. Submits it by 21-Oct-2019, 6:00 PM.

TOP 15 CANDIDATES WILL SELECTED FOR MOCK INTERVIEW BASED ON THE PERFORMANCE WITH THIS DATASET

INNOMATICS
RESEARCH LABS

INNOVATION IS OUR TRADITION