International Journal of Business Analytics & Intelligence
10 (1) 2022, 04-15
http://publishingindia.com/ijbai/

Customer Attrition Analytics: The Case of a Recruitment Service Provider

Mihir Dash*, Vishnu Raghavan*

Abstract

Customer attrition is the phenomenon wherein a customer leaves a service provider. With the growing competition in the service sector, preventing customer attrition has become critical for sustainability, as it is well established that retaining existing customers is more profitable than acquiring new customers (Jacob, 1994). This gives customer attrition analytics the challenging task of predicting which customers are likely to leave, and of subsequently designing and implementing retention programmes for these customers. Customer analytics has made many strides in marketing, employer desirability, and branding, but has so far made limited strides in the recruitment industry space. The objective of the study is to identify the factors affecting a candidate's decision to accept a job opportunity in an organisation, using predictors such as the industry verticals, the candidate's skillsets, workplace location, gender, compensation offered, and the notice period of the candidate. The model developed is a logistic regression model, to determine whether a candidate selected will accept a job opportunity in an organisation or not. The analysis was performed based on a sample of 443 candidates who were provided job offers in the period 2013-2015 by a recruitment service provider.

Keywords: Customer Attrition Analytics, Factors Affecting Customer Attrition, Logistic Regression Models

Introduction

Customer attrition is the phenomenon wherein a customer leaves a service provider. With the growing competition in the service sector, preventing customer attrition has become critical for sustainability, as it is well established that retaining existing customers is more profitable than acquiring new customers (Jacob, 1994). This gives customer attrition analytics the challenging task of predicting which customers are likely to leave, and of subsequently designing and implementing retention programmes for these customers. Customer analytics has made many strides in marketing, employer desirability, and branding, but has so far made limited strides in the recruitment industry space.

The HR services industry in India has grown exponentially in the last decade, with the need for more specialised providers of services growing rapidly across the world. The services included under its ambit are recruitment and staffing, payroll processing, and several other services (Dash et al., 2009). Of these, a significant portion is occupied by organisations in the recruitment space. The total size of the industry in India is approximately USD 3.5 billion, with major players such as Adecco, Ranstaad, and ADP offering diversified services, and a few organisations specialising in niche skill hires.

Recruitment is generally considered the basis of all HR activity, and as a result, is very important. The nature of services provided in the recruitment space are quite varied. These are classified as follows:

- *Permanent Contingent Hires:* In this case, an organisation will hire a replacement for a position that will turn vacant. Certain large organisations, such as banks and IT product development organisations, are beset with a problem of constant attrition for certain in-demand skillsets, and as a result, hold periodic recruitment drives to fill actual and potential gaps in the organisation.
- *Contract Staffing:* Certain organisations require a huge proportion of employees on contract for proj-

* Alliance University, Bengaluru, Karnataka, India. Email: mihirda@rediffmail.com
** MBA Graduate, Alliance University, Bengaluru, Karnataka, India. Email: raghavan.vishnu@gmail.com

ects that arise from time to time. These employees work for the organisation, but are on the payroll of the recruitment firm. These services require a ready supply of available talent that can be drafted in to plug the gaps.

- *RPO:* Recruitment process involves certain employees working for a client, but on the payrolls of the parent recruitment firm. The role of these employees is to schedule, coordinate, and arrange the entire recruitment process for the client organisation.

There are specialised organisations that provide contract staffing and RPO services. The organisation in question is a contingent manpower placement organisation, with a decade's experience in hiring across verticals and levels. The organisation specialises in 'Niche Skill Hires', which is a rarer set of skills to work on and close. The organisation provides RPO-type services and permanent placements for IT product companies, in addition to manufacturing and banking verticals.

The industry is still largely fragmented in India. Although large players in the HR services, such as ADP, Ranstaad, and Manpower, specialise in contingent hires, the space is also occupied by many small players due to the size of India's workforce and the relatively low cost of operating in this space. In addition, there are established players that handle only CXO level hires (executive search firms), such as RGF Search, Korn Ferry, Spencer Stuart, and so on.

Organisation Overview

The organisation in question, XYZ Pvt. Ltd., is a decade-old player in the recruitment service industry. It has a team of 75 research recruiters and its focus is on niche hires. This is a unique value proposition that focuses predominantly on clients in the IT product development space, BFSI, pharmaceutical, and manufacturing verticals. The candidates placed with the clients are across varying levels of seniority, ranging from junior professionals with 1-2 years of experience, through to senior, CXO-level candidates. The organisation has a dedicated wing for placing CEO/CXO-level candidates, owing to the varied nature of recruiting these professionals. The candidates also differ in terms of their skillsets, ranging from purely technically-qualified candidates, through to techno-functional candidates, and to purely managerial candidates. It is spread across Chennai, Bengaluru, Hyderabad, Coimbatore, Salem, and Puducherry, and is planning to expand to Pune and the Middle East. It services clients not just in India, but also from across the globe, including Malaysia, Singapore, UAE, Qatar, and the USA. The organisation places around 400-500 candidates per annum successfully, and has a success rate of 85% in converting offers into joiners.

The organisation is structured into various levels, viz., the directors, followed by the operations manager, the assistant managers, and team leads, followed by the recruiters. The assistant managers lead specialised teams that focus on specific industry verticals.

The organisation uses a standard procedure during its recruitment process. The first step is obtaining a job description from the client. The same is entered in the organisation's ERP system and shared with the recruitment team. The next step is to explain the profile to the recruiters. The recruiters then source profiles and speak to the candidates. The profiles are sourced from job portals and social media sites such as LinkedIn, Zoominfo, Skillpages, and so on. The profiles are then screened by a validation team for their suitability and the details of the candidate are noted. The profiles deemed suitable are then subject to a process of validation by a specialised team. The profiles must avoid duplication within the recruitment firm, which is facilitated by means of an ERP system and a set of robust internal procedures. Additionally, the profile must not be processed by a competing recruitment firm. This is ensured by means of the speed with which the recruiters work. Also, the recruiters cross-check with the candidate if they have attended interviews with the client earlier.

Next, the profiles are sent to the client with the relevant details, and feedback is awaited. If the profile is shortlisted, the feedback is sent to the recruitment firm with a request for an interview. The candidates are then contacted and scheduled for an interview on a suitable date and time in agreement with the client. The interview usually consists of a few rounds and the candidate is followed by the recruiter to ensure attendance. The candidates are sent a call letter with detailed instructions and are followed up with, to ensure that a maximum number of candidates attend the interview. This will help increase the chances of offers. Following the selection of the candidate, the

client releases an offer letter which must be accepted by the candidate. Next, the candidate must resign from his/her organisation. The tentative date of relieving is obtained and the same is communicated to the client. Following this, the candidate must actually join the client organisation. The notice period ranges from a few days or weeks to a few months. The candidate may at times hold a series of other offers. Following the release of the offer, the intent of the candidate to resign or to take up the offer is gauged by periodic follow-ups by the recruiter, as well as the respective team leads. The follow-ups continue till the candidate finally joins the organisation. In addition, an independent team meant for process improvement has been assigned to monitor the candidates and their seriousness in taking the offers. The team makes a series of periodic random calls to assess the candidate and provides information in advance to the team leads, who in their turn pass this information to the clients.

Problem Statement

The requirements to be filled by the client organisation usually are mandates that the client is unable to accomplish owing to a host of factors. The factors include paucity of time, shortfall of candidates, and also, exhausting the resources available on job boards. As a consequence, the same is outsourced to a service provider with specialisation in recruitment. The requirements are usually shared on a non-exclusive basis, and as a result, there are certain issues that are addressed by the recruitment firm. These are the speed of generation of profiles, the quality of the profiles, and a certain critical mass of profiles. This will ensure a greater chance of closures for the client from the service provider's end.

After a candidate is given an offer, the recruitment firm must first gain the acceptance of the offer and then seek the candidate to resign from his/her current organisation. These two steps are mandatory and will determine if the candidate was serious about the offer or was merely testing his/her suitability for the job market.

Following the acceptance of the job offer, a candidate usually follows one of three courses of action:

- He/she will cease to seek any new offers. This happens in a small proportion of candidates. They are satisfied with the offer on hand and they confine their search to the client from which they get their first offer, and except in some contingencies, they join the organisation.
- The candidate will continue to seek job offers. This may be due to a host of factors that differentiate the candidates, and as a result, the candidate will be reluctant to take up the offer on hand in search of something better.
- The candidate may get retained by his parent organisation, and as a result, will decline the offer.

There are several possible factors at play for the candidates while selecting a new workplace. The requirements are broadly classed into IT- and non-IT-based requirements, depending on the nature of skillsets. In addition, the candidate's level of experience may influence the decision to shift. The location of the offer is also an important factor. There is generally no bar on candidates based on gender, marital status, or age, although clients sometimes specify gender requirements in advance, owing to the nature of the job and the attendant duties. In addition, the candidates will have notice periods that vary. Some may have negotiable notice periods, others may be serving notice, while still others may have fixed notice periods. Intuitively, in-station candidates are more likely to join an organisation; male candidates are more likely to accept relocation; and candidates with longer notice periods are more likely to attempt a switch. In addition, non-IT candidates have fewer job opportunities as of date vis-à-vis IT candidates; the latter are more likely to decline or avoid a new job offer.

Clients are dependent on the replacement candidate to ensure continuity of business. In case a candidate is assessed as a potential ditch, it helps the client and the recruitment firm schedule and prepare back-up candidates to ensure business continuity. However, it is not possible to determine which candidate will join, and the likelihood of joining of a candidate. Knowledge of these details will enable us to optimise the follow-up calls, and also to inform the client in advance regarding the intention of the candidate. This will enable the recruitment firm to take suitable steps and mitigate any losses with time.

The candidates, on receiving the offer, are often approached by more than one recruiting firm. The competing firms may approach the same candidate for a

position with a competing organisation. A combination of the above factors, along with an ever-fluctuating job market, makes determining the intention of the candidate extremely difficult. During follow-up by a recruiter post-offer and salary negotiations, candidates may disguise the fact that they are not satisfied with the current offer. A candidate may be unwilling to relocate from their current location. Sometimes the candidate may not be happy with the organisation that they will be joining, owing to negative reviews they have read about the organisation online. Also, a candidate may be lured by a better salary package to join a competing organisation. Additionally, there may be pressure from the family of the candidate to seek a better offer. Candidates also will disguise the fact that they hold extra offers till the last minute. This results in a loss of revenue for the client as well as the recruitment firm, and the value proposition of the recruiter is weakened. While customer analytics has made many strides in business and management, it has so far not addressed the issues discussed above. This paper is an attempt to address these issues.

Literature Review

Recruitment firms the world over are faced with a problem of whether a candidate will accept a job offer from a particular company or client. There are a few factors that need to be taken into account in this context, including:

- *The Notice Period of the Candidate:* Notice periods can range from two weeks to four months. Candidates who serve short notices tend to have other offers already or are currently unemployed, and accept the clients' offer. With longer notice periods, candidates tend to attempt interviews with many organisations. Thus, candidates with longer notice periods are usually more likely to not take up a job offer, compared to those with shorter notice periods.
- *The Location of the Job Opportunity:* Candidates tend to prefer a job opportunity that is located near their hometown, or in some cases, in a cosmopolitan city. Hence, cities such as Pune, Bengaluru, or Mumbai will attract candidates from across India, while areas like NCR or Chennai will attract people from areas that are closer. Additionally, candidates who are within a city will prefer to stay there and will not be willing to relocate.
- *Nature of the Skill Set:* Generally, employees from non-IT backgrounds tend to have fewer opportunities, and hence have fewer options for career advancement. Candidates from IT and related areas tend to have a greater number of opportunities and represent a threat to the organisation. Skillsets in heavy demand, such as Java developers, command better salaries and/or salary hikes, and candidates tend to shift readily.
- *Employer Brand Value Proposition:* Employers with a good reputation tend to attract a greater number of applicants. Usually, such clients are faced with fewer ditches by the candidate, while lesser known clients tend to have a problem in terms of joiners. The reputation of an employer is also important. Reviews posted by employees are available on the Internet and these are a source of feedback to candidates who may join the organisation.
- *Compensation Package on Offer:* Candidates are swayed, to a good extent, on the compensation package they receive. A better employer with a satisfactory compensation structure will pull potential employees to the organisation. Candidates may at times be swayed by better compensation from rival organisations and may not be willing to take up an offer with certain clients.
- *Gender:* Male candidates, as a rule, will usually be willing to relocate more readily, compared to female candidates.
- *Marital Status:* Married candidates tend to prefer a stable work location, owing to the attendant factor of relocation of their spouses and children.

Several statistical techniques are commonly applied for customer attrition analytics, including classification and regression trees (Gray & Fan, 2008), logistic regression (Au et al., 2003), artificial neural networks (Datta et al., 2001), survival analysis (Ma & Li, 1994), and several others (Hadden et al., 2006). There are mixed results concerning the most appropriate technique; however, several studies support the logistic regression model. For instance, Mozer et al. (2000) and Hwang et al. (2004) suggested that logistic regression predicted customer attrition better than decision trees and neural networks.

Methodology

The primary objective of the study is to identify the factors affecting a candidate's decision to accept a job

opportunity in an organisation. The data sources employed were primarily the internal databases of the organisation, XYZ Ltd. The period of the study was 2013-2015, based on data availability.

The independent variables (predictors) considered included the industry verticals (banking, telecom, pharma, IT, non-IT, and so on), the candidate's skill sets (IT, non-IT, and so on), workplace location, the level of the position offered, compensation offered, notice period of the candidate, and the gender of the candidate.

The dependent variable (outcome) was the binary variable of whether a candidate had accepted or had declined to accept a job opportunity in an organisation. The model developed is a logistic regression model, to determine whether a candidate selected will accept a job opportunity in an organisation or not.

The analysis was based on a sample of 443 candidates, who were provided job offers through XYZ Ltd. during the period 2013-2015. Around 81.3% of the candidates were men, while 18.7% were women; 54.9% of the candidates were IT-skilled, while 45.1% were non-IT-skilled. There were 31 clients represented in the data, of which the top five were Scope (28.4%), Barclays (12.2%), Aon Hewitt (9.0%), Hospira (9.0%), and UST (6.8%). The positions were offered across 17 locations, of which the top five were Chennai (66.4%), Bengaluru (12.9%), Mumbai (6.1%), Coimbatore (2.7%), and Thiruvananthapuram (2.3%). The clients belonged to nine verticals, of which the top five were BFSI (45.8%), IT (13.3%), Pharma (12.4%), IT Products (9.5%), and Telecom (7.7%). The positions offered were classified into seven hierarchical levels, and the basic salaries were in the range Rs. 6,000 p.m. to Rs. 5,50,000 p.m., with a mean of Rs. 73,352.59, a standard deviation of Rs. 57,683.66, and a median of Rs. 58,310.00. The month of joining varied roughly uniformly across the 12 months, except for lows in January and February. The notice period of the candidates varied between one and three months, with a mean of 1.7020 months, a standard deviation of 0.7455 months, and a median of 1.5 months. Finally, 78.1% of the candidates accepted the positions offered, while 21.9% did not. The frequency tables are available in the Appendix.

Analysis and Findings

Initial exploratory cross-tabulations indicated various significant differences in the percentage of non-joiners between groups. For ease of presentation, some prominent differences are listed point-wise as follows:

- IT skillset (25.8%) vs. non-IT skillset (17.9%); $\chi^2 = 3.986$, $p = 0.046$
- for women, Chemicals (66.7%), IT Products (50.0%), Telecom (42.9%), BFSI (16.7%), and others (0.0%); $\chi^2 = 18.782$, $p = 0.009$
- for women with IT skillset, Chemicals (100.0%), IT Products (50.0%), Telecom (42.9%), BFSI (17.4%), and others (0.0%); $\chi^2 = 15.901$, $p = 0.014$
- for IT vertical, men (34.8%) vs. women (0.0%); $\chi^2 = 6.204$, $p = 0.013$
- for IT vertical with IT skillset, men (31.6%) vs. women (0.0%); $\chi^2 = 5.034$, $p = 0.025$
- for IT skillset, Middle Level Managers (100.0%), Project Level Managers (50.0%), Senior Entry Level (26.3%), Junior Entry Level (25.0%), Project/Team Lead (23.1%), Junior Level Managers (6.7%), and Senior Level Managers (0.0%); $\chi^2 = 15.473$, $p = 0.030$
- for IT skillset in the IT vertical, Project Level Managers (100.0%), Senior Entry Level (29.6%), and others (0.0%); $\chi^2 = 18.232$, $p = 0.006$
- for men in the IT vertical, Project Level Managers (100.0%), Senior Entry Level (41.7%), Junior Level Managers (20.0%), and others (0.0%); $\chi^2 = 16.758$, $p = 0.010$
- for IT vertical, Project Level Managers (60.0%), Senior Entry Level (32.3%), Junior Level Managers (14.3%), and others (0.0%); $\chi^2 = 14.317$, $p = 0.026$
- for Junior Entry Level, IT skillset (25.0%) vs. non-IT skillset (3.3%); $\chi^2 = 5.975$, $p = 0.015$
- for women at Junior Entry Level, Telecom (100.0%), IT Products (75.0%), BFSI (8.3%), and others (0.0%); $\chi^2 = 11.111$, $p = 0.049$
- for women at Senior Entry Level, Chemicals (100.0%), IT Products (40.0%), BFSI (16.7%), and others (0.0%); $\chi^2 = 13.876$, $p = 0.031$
- for women Project Level Managers, IT skillset (20.0%) vs. non-IT skillset (0.0%); $\chi^2 = 7.000$, $p = 0.008$

In view of these observations, the logistic regression analysis was performed for each of the major verticals, as well as for men and women, separately. Table 1 presents the overall logistic regression results.

Table 1: Overall Logistic Regression Results

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Men	−0.2696	0.4231	0.7637			
IT Skillset	−0.2375	0.4796	0.7886	−0.5385	0.0277	0.5836
Level		0.4082				
Junior Entry Level	0.4724	0.4635	1.6038			
Senior Entry Level	−0.0897	0.8742	0.9142			
Project/Team Lead	−0.1165	0.8489	0.8900			
Junior Level Managers	0.8230	0.2692	2.2774			
Project Level Managers	−0.6602	0.3037	0.5167			
Middle Level Managers	−0.2553	0.7133	0.7746			
Notice Period		0.0146			0.0082	
1.00 Month	1.2266	0.0013	3.4095	1.1077	0.0012	3.0274
1.50 Months	0.4762	0.2390	1.6099	0.3717	0.3270	1.4502
2.00 Months	0.3484	0.3726	1.4168	0.2647	0.4754	1.3030
2.50 Months	0.2223	0.6471	1.2489	0.0878	0.8426	1.0917
Basic Salary	0.0000	0.9964	1.0000			
Vertical		0.4423				
BFSI	0.3073	0.5163	1.3597			
Chemicals	−0.9605	0.2077	0.3827			
Electronics	0.4158	0.6590	1.5156			
Infrastructure	0.5213	0.6620	1.6841			
IT	−0.3213	0.5352	0.7252			
IT Products	−0.1377	0.7958	0.8714			
Manufacturing	0.6379	0.4880	1.8924			
Pharma	0.7607	0.1890	2.1398			
[Constant]	0.8270	0.2984	2.2864	1.0218	0.0004	2.7782
Model Fit X^2	35.623	0.0330		18.952	0.0020	
Nagelkerke R^2	12.2%			6.6%		
% Correctly Classified	79.4%			77.3%		

The overall logistic regression results suggest that the only significant variables affecting the likelihood that a candidate joins are the candidate's skillsets and the notice period. In particular, candidates with IT skillsets are less likely to join than candidates with non-IT skillsets; and the lesser the notice period, the more likely the candidate is to join. The explanatory power of the overall logistic regression results is quite low, indicating scope for improvement.

Table 2 presents the logistic regression results for the BFSI vertical.

Table 2: Logistic Regression Results for BFSI Vertical

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Men	−0.3417	0.4761	0.7105			
IT Skillset	0.2547	0.5938	1.2900			
Level		0.4680				
Junior Entry Level	0.7349	0.4603	2.0854			
Senior Entry Level	0.0504	0.9435	1.0517			
Project/Team Lead	−0.8199	0.2853	0.4405			

	Coeff.	P-Value	Exp(B)	Coeff.	P-Value	Exp(B)
Junior Level Managers	0.6540	0.5049	1.9232			
Project Level Managers	0.8228	0.4181	2.2769			
Middle Level Managers	0.0124	0.9878	1.0124			
Notice Period		0.0225			0.0177	
1.00 Month	1.6171	0.0119	5.0383	1.6337	0.0063	5.1228
1.50 Months	0.2039	0.7463	1.2261	0.2744	0.6392	1.3158
2.00 Months	0.1646	0.7815	1.1789	0.1646	0.7687	1.1789
2.50 Months	−0.5640	0.4511	0.5689	−0.3542	0.5952	0.7018
Basic Salary	0.0000	0.8435	1.0000			
[Constant]	0.9918	0.2427	2.6961	0.8650	0.0401	2.3750
Model Fit X^2	22.645	0.066		14.674	0.005	
Nagelkerke R^2	17.6%			11.6%		
% Correctly Classified	79.8%			80.3%		

The logistic regression results for the BFSI vertical suggest that the only significant variable affecting the likelihood that a candidate joins is the notice period. In particular, the lesser the notice period, the more likely the candidate is to join. The explanatory power of the logistic regression results for the BFSI vertical is relatively low, indicating scope for improvement.

Table 3 presents the logistic regression results for the IT vertical.

Table 3: Logistic Regression Results for IT Vertical

	Coeff.	P-Value	Exp(B)	Coeff.	P-Value	Exp(B)
Men	−20.9601	0.9985	0.0000	−20.8344	0.9986	0.0000
IT Skillset	0.5452	0.6578	1.7250			
Level		0.8664			0.8774	
Junior Entry Level	1.6128	1.0000	5.0168	−0.2843	1.0000	0.7526
Senior Entry Level	−19.0146	0.9994	0.0000	−20.8664	0.9994	0.0000
Project/Team Lead	2.0703	0.9999	7.9273	−19.8166	0.9994	0.0000
Junior Level Managers	−18.1976	0.9994	0.0000	0.0000	1.0000	1.0000
Project Level Managers	−21.1692	0.9993	0.0000	−22.3015	0.9994	0.0000
Notice Period		0.9071				
1.00 Month	21.3840	0.9996	1.94E+09			
1.50 Months	21.8670	0.9996	3.14E+09			
2.00 Months	22.8788	0.9995	8.63E+09			
2.50 Months	41.4084	0.9992	9.63E+17			
Basic Salary	0.0000	0.5604	1.0000			
[Constant]	18.8920	0.9997	1.60E+08	42.0373	0.9989	1.81E+18
Model Fit X^2	30.432	0.004		25.568	0.001	
Nagelkerke R^2	59.5%			52.0%		
% Correctly Classified	80.7%			78.9%		

The logistic regression results for the IT vertical suggest that the only significant variable affecting the likelihood that a candidate joins is the level. In particular, candidates are less likely to join at lower hierarchical levels. The explanatory power of the logistic regression results for the IT vertical is moderate.

Table 4 presents the logistic regression results for the Pharma vertical.

Table 4: Logistic Regression Results for Pharma Vertical

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Men	−17.9120	0.9990	0.0000			
IT Skillset	2.5031	0.2257	12.2198			
Level		0.8000				
Junior Entry Level	10.3550	0.9994	3.14E+04			
Senior Entry Level	−9.7970	0.1218	0.0001			
Project/Team Lead	14.6203	0.9991	2.24E+06			
Junior Level Managers	−10.9619	0.1173	0.0000			
Project Level Managers	−13.0461	0.0874	0.0000			
Middle Level Managers	7.1471	0.9998	1270.4238			
Notice Period		0.9234				
1.00 Month	25.5640	0.9979	1.27E+11			
1.50 Months	2.3287	0.5291	10.2650			
2.00 Months	1.2656	0.4168	3.5452			
2.50 Months	25.9374	0.9983	1.84E+11			
Basic Salary	−0.0001	0.0267	0.9999	−0.0000	0.0146	1.0000
[Constant]	35.1282	0.9980	1.80E+15	3.3800	0.0002	29.3719
Model Fit X^2	31.304	0.003		6.630	0.010	
Nagelkerke R^2	75.1%			19.9%		
% Correctly Classified	90.4%			88.5%		

The logistic regression results for the Pharma vertical suggest that the only significant variable affecting the likelihood that a candidate joins is the salary. In particular, candidates are less likely to join at higher salary levels. The explanatory power of the logistic regression results for the Pharma vertical is low.

Table 5 presents the logistic regression results for the IT Products vertical.

Table 5: Logistic Regression Results for IT Products Vertical

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Men	0.9238	0.3313	2.5189			
Level		0.9827				
Junior Entry Level	−18.7421	0.9996	0.0000			
Senior Entry Level	−18.5389	0.9996	0.0000			
Project/Team Lead	−17.6939	0.9996	0.0000			
Junior Level Managers	1.6308	1.0000	5.1079			
Middle Level Managers	−19.9158	0.9996	0.0000			
Notice Period		0.6975				
1.00 Month	−0.2593	0.8250	0.7716			
1.50 Months	−0.2173	0.8876	0.8047			
2.00 Months	−0.4993	0.6771	0.6069			
2.50 Months	−1.7610	0.1940	0.1719			
Basic salary	0.0000	0.4013	1.0000			
[Constant]	17.7266	0.9996	5.00E+07	0.8023	0.0162	2.2308
Model Fit X^2	9.908	0.624		-	-	
Nagelkerke R^2	29.6%			-		
% Correctly Classified	73.8%			69.0%		

The logistic regression results for the IT Products vertical suggest that none of the variables considered affect the likelihood that a candidate joins.

Table 6 presents the logistic regression results for the Telecom vertical.

Table 6: Logistic Regression Results for Telecom Vertical

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Men	1.3428	0.4449	3.8297			
IT Skillset	−22.6703	0.9995	0.0000			
Level		0.9603				
Junior Entry Level	−1.0559	0.4329	0.3479			
Project/Team Lead	−0.3848	0.7862	0.6806			
Junior Level Managers	19.9889	0.9994	4.80E+08			
Project Level Managers	−40.4920	0.9987	0.0000			
Notice Period		0.9923				
1.00 Month	0.6774	0.6155	1.9688			
1.50 Months	0.2863	0.8385	1.3314			
2.00 Months	21.1121	0.9989	1.48E+09			
2.50 Months	62.5131	0.9990	1.41E+27			
Basic Salary	0.0000	0.4289	1.0000			
[Constant]	23.0563	0.9995	1.03E+10	0.8755	0.0200	2.4000
Model Fit X^2	15.020	0.182		-	-	
Nagelkerke R^2	50.8%			-		
% Correctly Classified	82.4%			70.6%		

The logistic regression results for the Telecom vertical suggest that none of the variables considered affect the likelihood that a candidate joins.

Table 7 presents the logistic regression results for men.

Table 7: Logistic Regression Results for Men

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
IT Skillset	−0.0476	0.8987	0.9535			
Level		0.5114				
Junior Entry Level	0.7293	0.3064	2.0736			
Senior Entry Level	−0.0524	0.9326	0.9489			
Project/Team Lead	−0.1058	0.8739	0.8996			
Junior Level Managers	1.1189	0.1782	3.0615			
Project Level Managers	−0.4830	0.4927	0.6169			
Middle Level Managers	0.1784	0.8085	1.1953			
Notice Period		0.0095			0.0138	
1.00 Month	1.4324	0.0007	4.1885	1.1069	0.0024	3.0250
1.50 Months	0.4996	0.2568	1.6480	0.3124	0.4433	1.3667
2.00 Months	0.4875	0.2589	1.6282	0.3365	0.4082	1.4000
2.50 Months	0.2000	0.7044	1.2214	−0.0953	0.8375	0.9091
Basic Salary	0.0000	0.6796	1.0000			
Vertical		0.3110				
BFSI	0.3183	0.5486	1.3748			

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
Chemicals	−0.8561	0.3021	0.4248			
Electronics	0.3623	0.7087	1.4366			
Infrastructure	0.4718	0.6994	1.6029			
IT	−0.6935	0.2209	0.4998			
IT Products	0.0113	0.9850	1.0114			
Manufacturing	0.5765	0.5468	1.7797			
Pharma	0.7125	0.2524	2.0391			
[Constant]	0.3852	0.6320	1.4699	0.6931	0.0114	2.0000
Model Fit X^2	34.037	0.049		13.411	0.009	
Nagelkerke R^2	14.2%			5.7%		
% Correctly Classified	77.4%			76.5%		

The logistic regression results for men candidates suggest that the only significant variable affecting the likelihood that a candidate joins is the notice period. In particular, the lesser the notice period, the more likely the men candidates are to join. The explanatory power of the logistic regression results for men is relatively low, indicating scope for improvement.

Table 8 presents the logistic regression results for women.

Table 8: Logistic Regression Results for Women

	Coeff.	*P-Value*	*Exp(B)*	*Coeff.*	*P-Value*	*Exp(B)*
IT Skillset	−0.7136	0.4677	0.4899			
Level		0.9840				
Junior Entry Level	−19.0334	0.9992	0.0000			
Senior Entry Level	−19.3713	0.9991	0.0000			
Project/Team Lead	−19.9730	0.9991	0.0000			
Junior Level Managers	−19.1141	0.9991	0.0000			
Project Level Managers	−19.8140	0.9991	0.0000			
Middle Level Managers	−20.4831	0.9991	0.0000			
Notice Period		0.9999				
1.00 Month	−0.0449	0.9706	0.9561			
1.50 Months	−0.1981	0.8938	0.8203			
2.00 Months	−0.0751	0.9512	0.9276			
2.50 Months	−0.0165	0.9917	0.9836			
Basic Salary	0.0000	0.6235	1.0000			
Vertical		0.9346			0.3733	
BFSI	0.6144	0.6177	1.8485	1.2993	0.1896	3.6667
Chemicals	−21.8651	0.9996	0.0000	−21.6084	0.9996	0.0000
IT	20.4491	0.9987	7.60E+08	20.7974	0.9987	1.08E+09
IT Products	−0.7502	0.5749	0.4723	−0.4055	0.7255	0.6667
Pharma	20.0157	0.9991	4.93E+08	20.7974	0.9991	1.08E+09
[Constant]	20.4821	0.9991	7.86E+08	0.4055	0.6569	1.5000
Model Fit X^2	18.341	0.368		15.155	0.010	
Nagelkerke R^2	32.9%			27.7%		
% Correctly Classified	82.7%			82.7%		

The logistic regression results for women suggest that the only significant variable affecting the likelihood that a candidate joins is the vertical. In particular, women candidates are more likely to join in the IT and Pharma verticals, and less likely to join the Chemicals vertical. The explanatory power of the logistic regression results for the women is moderate.

Discussion

The model provides a tool for recruitment firms to prioritise job offers to candidates. The model can be used to optimise candidate-job matching, benefiting all parties involved; in particular, the losses to the recruiters and clients can be minimised. Depending on the candidate's preference for a job opportunity, their likelihood of joining can be notified to the client in advance. Based on the client's assessment, a suitable backup can be generated to ensure business continuity. Further, follow-ups to the candidate can be made on the basis of the priority accorded to the candidate.

The results of the study indicate that there are different factors affecting the likelihood of joining; it may vary with different sub-groups. In particular, the factors should be analysed in detail, to clearly identify the factors affecting the decreasing popularity of our teams. More detailed analysis is required to minimise the differences between groups for the Indian respondents.

There are several limitations inherent in the study. The sample size is relatively small, and in particular, some verticals were under-represented. In the same vein, women were quite under-represented in almost all of the groups. A difficulty in some of the results is the low explanatory power, which suggests that other variables should be considered to improve the same. A more serious difficulty is that some of the results indicate multicollinearity between some of the independent variables, suggesting that some variables may have to be excluded from consideration in order to yield more robust results. Also, the data is collected during the period 2013-15, and so may not reflect more up-to-date recruitment trends.

There are several ways through which the results of the study can be updated and/or extended. More variables can be introduced into the model, for example, the candidate's level of experience could be another factor – more experienced candidates would be less likely to switch jobs unless it resulted in better pay or a better profile. Other similar hypotheses may be examined in detail for future studies.

References

Au, W., Chan, C. C., & Yao, X. (2003). A novel evolutionary data mining algorithm with applications to churn prediction. *IEEE Transactions on Evolutionary Computation, 7,* 532-545.

Dash, M., Singh, A., Mishra, N., & Gupta, G. (2009). A study of human resource outsourcing in Indian IT companies. SSRN Working Papers.

Datta, P., Masand, B., Mani, D. R., & Li, B. (2001). Automated cellular modelling and prediction on a large scale. *Issues on the Application of Data Mining,* 485-502.

Gray, J. B., & Fan, G. (2008). Classification tree analysis using TARGET. *Computational Statistics & Data Analysis, 52,* 1362-1372.

Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2006). Churn prediction using complaints data. *Transactions on Engineering, Computing and Technology, Enformatika, 13,* 158-163.

Hwang, H., Jung, T., & Suh, E. (2004). An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunications industry. *Expert Systems with Applications, 26,* 181-188.

Jacob, R. (1994). Why some customers are more equal than others. *Fortune,* 149-154.

Ma, G., & Li, S. (1994). *Applications of the survival analysis techniques in modelling customer retention.* Workbook for the 5th Advanced Research Techniques Forum, American Marketing Association.

Mozer, M. C., Wolniewicz, R., Grimes, D. B., Johnson, E., & Kaushansky, H. (2000). Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. *IEEE Transactions on Neural Networks, 11,* 690-696.

Appendix

Table 9: Frequency Table of Verticals, Clients, and Glassdoor Ratings

Vertical	*Client*	*Glassdoor Rating*	*Frequency*
BFSI	ANZ	3.5	23
	Barclays	3.4	54
	Scope	3.4	126
Chemicals	VWR	3.3	10
Electronics	Flextronics	3.3	6
	Logitech	4.0	3
Infrastructure	Aparna Constructions	2.9	1
	Dar Engineering	3.0	1
	IWL India	2.3	2
	MYK Laticrete	3.7	4
	RGM	4.8	3
IT	Adaptive Mobiles	4.0	3
	Cross Domain	3.2	5
	Durr India	3.7	4
	IDOCZ	1.0	2
	New Age Management	4.0	5
	Software One	3.3	1
	Touchbase	3.5	3
	UST	3.0	30
	Valuelabs	3.0	3
	Williams Lea	3.3	3
IT Products	Aon Hewitt	3.7	40
	Lennox	3.5	2
Manufacturing	Hirotec	2.9	1
	Sharada Ceramics	2.3	1
	TVS Sundaram-Brake Linings Limited	2.3	5
	TVS Sundaram industries	3.5	13
Pharma	Astrazeneca	3.5	15
	Hospira	3.0	40
Telecom	Ericsson	3.6	19
	Prodapt	3.3	15
	Total		443

Table 10: Frequency Table for Locations

	Frequency
Chennai	294
Bengaluru	57
Mumbai	27
Coimbatore	12
Thiruvananthapuram	10
Delhi NCR	9
Hyderabad	8
Visakhapatnam	6
Cochin	5
Dubai	3
Kolkata	3
Pune	3
Ahmedabad	1
Jamshedpur	1
Mysuru	1
Raipur	1
Vijayawada	1
Total	442

Table 11: Frequency Table for Levels

	Frequency
Junior Entry Level	66
Senior Entry Level	180
Project/Team Lead	72
Project Level Manager	32
Junior Level Manager	38
Middle Level Manager	22
Senior Level Manager	28
Total	438

Table 12: Frequency Table for Month of Offer

	Frequency
Jan	6
Feb	14
Mar	40
Apr	52
May	47
Jun	54
Jul	41
Aug	48
Sep	35
Oct	35
Nov	38
Dec	32
Total	442