ABSTRACT

Retaining the most valuable customers is a major problem companies face in this information age. Especially, the field of telecommunication faces complex challenges due to a number of vibrant competitive service providers. Therefore, it has become very difficult for them to retain existing customers. Since the cost of acquiring new customers is much higher than the cost of retaining the existing customers, it is the time for the telecom industries to take necessary steps to retain the customers to stabilize their market value. CRM uses data mining (one of the elements of CRM) techniques to interact with customers. This study investigates the use of a technique, supervised learning, for the management and analysis of customer-related data warehouse and information. Data mining technologies extract hidden information and knowledge from large data stored in databases or data warehouses, which supports the corporate in decision making process. Several data mining techniques have been proposed in the literature for predicting the happy and stressed customer using heterogeneous customer records. Probably, the stressed customers are in the urge of moving out to competitive service providers. This project analysis the telecom customer data available in open data set and predict the customer stress by applying supervised machine-learning algorithms mainly using Deep Neural Network , K Nearest Neighbour , Support Vector Machine and Random Forest.

TABLE OF CONTENTS

CHAPTER No.	CHAPTER NAME	PAGE No.
	ABSTRACT	**V**
	LIST OF FIGURES	**VIII**
	LIST OF ABBREVATIONS	**IX**
1	**INTRODUCTION**	**01**
	1.1 OBJECTIVE	**01**
	1.2 EXISTING SYSTEM	**01**
	1.2.1 SVM ALGORITHM	**02**
	1.2.2 K NEAREST NEIGHBOUR	**02**
	1.3 PROPOSED SYSTEM	**03**
2	**LITERATURE SURVEY**	**05**
	2.1 CUSTOMER CHURN	**05**
	2.2 DATA MINING TECHNIQUES	**06**
	2.2.1 ARTIFICIAL NEURAL NETWORK	**06**
	2.2.2 SELF ORGANIZING MAPS	**07**
	2.2.3 HYBRID APPROCHES	**08**
	2.3 RELATED WORK	**09**
3	**METHODOLOGY**	**10**
	3.1 HARDWARE REQUIREMENTS	**10**
	3.2 SOFTWARE REQUIREMENTS	**10**
	3.3 SYSTEM FEATURES	**11**
	3.3.1 PYTHON	**11**
	3.3.2 SPYDER	**12**

3.3.3 KERAS **12**
3.3.4 ANACONDA **13**
3.4 ALGORITHMS **13**
3.4.1 DEEP NEURAL NETWORK **13**
3.5 DATAFLOW DIAGRAM **14**
3.5.1 DFD AT THE INITIAL LEVEL **14**
3.5.2 DFD LEVEL 1 FOR ENTIRE PROJECT **15**
3.5.3 DFD LEVEL1 FOR DATA PREPROCESSING **15**
3.5.4 DFD LEVEL 2 FOR FEATURE EXTRACTION **16**
3.5.5 DFD LEVEL 2 FOR SPLITTING DATA **17**
3.5.6 DFD LEVEL 2 FOR CLASSIFYING DATA **18**
3.5.7 DFD LEVEL 2 FOR PREDICTION **19**
3.6 UNIFIED LANGUAGE **20**
3.6.1 USE CASE DIAGRAM FOR LOAD FORECASTING **20**
3.6.2SEQUENCE DIAGRAM FOR LOAD FORECASTING **21**
3.7 ARCHITECTURE DIAGRAM **23**
3.8 MODULES **23**
3.9 MODULER DESCRIPTION **25**
3.9.1 DATA EXTRACTION **25**
3.9.2 PREPROCESSING **25**
3.9.3 LABEL ENCODING **25**
3.9.4 ONE HOT CODING **25**
3.9.5 STANDARD SCALE **25**
3.9.6 DATA VISUVALIZATION AND DESCRIPTION **26**
3.9.7 DATA SPLITTING INTO TRAINING AND TEST SETS **26**
3.9.7.1 OVER FITTING **26**
3.9.7.2 UNDER FITTING **26**
3.9.7.3 TRAIN /TEST SPLIT **27**
3.9.8 DEEP NEURAL NETWORK **27**

3.9.9 PREDICTING ADMISSIONS BASED ON THE TEST **29**
4 RESULT AND DISCUSSION 31
5 CONCLUSION AND FUTURE WORK 34
5.1 CONCLUSION **34**
5.2 FUTURE WORK **34**
REFERENCES 35
APPENDICES 36
A.SOURCE CODE **36**
B.SCREENSHOTS **42**

LIST OF FIGURES

FIGURE No.	FIGURE NAME	PAGE No.
2.1	MULTILAYER NEURAL NETWORK	**07**
2.2	A 4*4 SELF KOHONENS SELF ORGANIZING MAP	**08**
3.1	DATA FLOW DIAGRAM LEVEL 0 FOR ENTIRE PROJECT	**14**
3.2	DATA FLOW DIAGRAM LEVEL 1 FOR ENTIRE PROJECT	**15**
3.3	DATA FLOW DIAGRAM LEVEL1 FOR PREPROCESSING DATA	**16**
3.4	DATA FLOW DIAGRAM LEVEL 2 FOR FEATURE EXTRACTION	**17**
3.5	DATA FLOW DIAGRAM LEVEL 2 FOR SPLITTING DATA	**18**
3.6	DATA FLOW DIAGRAM LEVEL 2 FOR CLASSIFYING DATA	**18**
3.7	DATA FLOW DIAGRAM LEVEL 2 FOR ACCURACY WEATHER PREDICTION	**19**
3.8	USE CASE DIAGRAM OF LOAD FORECASTING	**20**
3.9	SEQUENCE DIAGRAM OF MODULE	**21**
3.10	ARCHITECTURE DIAGRAM	**23**
3.11	DEEP NEURAL NETWORK	**27**
4.1	SPLITTING OF DATA	**31**
4.2	EPOCH COUNT FOR PREDICTION	**32**
4.3	ACCURACY PREDICTION FOR SUPPORT VECTOR MACHINE	**32**
4.4	ACCURACY PREDICTION FOR DEEP NEURAL NETWOR	**33**

LIST OF ABBREVATIONS

CRM	CUSTOMER RESOURCE MANAGEMENT
DNN	DEEP NEURAL NETWORK
DFD	DATA FLOW DIAGRAM
IDE	INTEGRATED DEVELOPMENT ENVIRONMENT
KNN	K NEAREST NEIGHBOUR
RDF	RANDOM FOREST
SGD	STOCHASTIC GRADIENT DESCENT
SVM	SUPER VECTOR METHOD
UML	UNIFIED MODELING LANGUAGE

CHAPTER 1

INTRODUCTION

1.1 OBJECTIVE

Our Aim Customer Stress-ML should meet the basic requirements, i.e., predict the churning of the customer, find the accuracy of the above algorithms *and* find the efficient one. Efficient Churn model predicts the Customer churning and helps in relieving the customer from the stress and help them in retaining in the respective telecom service.

1.2 EXISTING SYSTEM

Stressed customer, tend to cease doing business with a company in a given time period which has become a significant problem for many firms. These include publishing industry, investment services, insurance, electric utilities, health care providers, credit card providers, banking, Internet service providers, telephone service providers, online services, and cable services operators.

There are numerous predictive modeling techniques for predicting customer behavior and their satisfactory level. These vary in terms of statistical technique (e.g., neural nets versus logistic regression), variable selection method (e.g., theory versus stepwise selection), and number of variables included in the model.

We have highlighted below two algorithm and its predicting techniques.

- SVM algorithm
- KNN algorithm

1.2.1 SVM ALGORITHM

SVM algorithm developed by Vapnik, is based on statistical learning theory. In some classification cases, we try to find an optimal hyper-plane that separates two classes. When the two classes of points in the training set can be separated by a linear hyper-plane, it is natural to use the hyper-plane that separates the two groups of points in the training set by the largest margin. In order to find an optimal hyper-plane, we need to minimize the norm of the vector w, which defines the separating hyper-plane. This is equivalent to maximizing the margin between two classes.

LIMITATIONS

Customer stress prediction is a problem of classification between “happy” and “stressed” customer. But when the number of the negative examples is too small, the generalization performance of SVM classifier must be weak, and the error rates is proved unsatisfactory.

1.2.2 K NEAREST NEIGHBOUR

KNN algorithm is one of the simplest classification algorithm. Even with such simplicity, it can give highly competitive results. KNN algorithm can also be used for regression problems. KNN algorithm which when implemented helps to predict wither the customer will churn or not. But the accuracy level is not so high.

LIMITATIONS

The major disadvantage in KNN is determining the K value . Distance based classifier makes a very strong assumption on the shape of your data distribution, i.e. any two features are independent given the output class.

Another problem happens due to data scarcity. For any possible value of a feature, you need to estimate a likelihood value by a frequents approach. This can result in probabilities going towards 0 or 1, which in turn leads to numerical instabilities and worse results. In this case, you need to smooth in some way your probabilities or to impose some prior on your data.

1.3 PROPOSED SYSTEM

In this paper, we review the existing works on customer stress prediction in three different perspectives: Data sets, methods, and metrics. Firstly, we present the details about the availability of public data sets and what kinds of customer details are available for predicting customer stress, which leads to churn. Secondly, we compare and contrast the various predictive modeling methods that have been used in the literature for predicting the churners using different categories of customer records, and then quantitatively compare their performances. Finally, we summarize what kinds of performance metrics have been used to evaluate the existing churn prediction methods. Analyzing all these three perspectives is very crucial for developing a more efficient churn prediction system for telecom industries.

In a business environment, the term, customer churn simply refers to the customers leaving one business service to another, which is the process of customers switching from one service provider to another anonymously. This is because of customer stress and un-satisfaction with the business. From a machine learning perspective, customer stress (probable churn) prediction is a supervised (i.e. labeled) problem defined as follows: Given a predefined forecast horizon, the goal is to predict the future churners over that horizon, given the data associated with each subscriber in the network.

Churn Prediction is a phenomenon which is used to identify the possible churners in advance before they leave the network. This helps the CRM department to prevent subscribers who are likely to churn in future by taking the required retention

policies to attract the likely churners and to retain them. Thereby, the potential loss of the company could be avoided.

The input for this problem includes the data on past calls for each mobile subscriber, together with all personal and business information that is maintained by the service provider. In addition, for the training phase, labels are provided in the form of a list of churners. After the model is trained with highest accuracy, the model must be able to predict the list of churners from the real data set which does not include any churn label. In the perspective of knowledge discovery process, this problem is categorized as predictive mining or predictive modeling.

We are analyzing the Churn data with DNN – Deep Neural Network and Random forest . As both the algorithms are a deep investigator of data , the accuracy of the prediction can be increased.

2.2 DATA MINING TECHNIQUES

In order to establish effective and accurate customer-churn prediction model, many data mining methods have been recently considered (e.g. Coussement & Van den Poel, 2008; Hung, Yen, & Wang, 2006). The two primary goals of data mining in practice tend to be description and prediction. Description focuses on finding human-interpreted patterns describing the data, and prediction involves using some variables or fifields in the database to predict unknown or future values of other variables of interest (Fayyad, Piatetsky, & Smyth, 1996; Fayyad & Uthurusamy, 1996).The goals of description and prediction can be achieved using a variety of particular data mining methods include classification, clustering, regression, and so on (Berry & Linoff, 2003).

2.2.1 ARTIFICIAL NEURAL NETWORK

Classification is one of the commonly used data mining methods, and called as supervised learning techniques. It calculates the value of some variables, and classifies according to results.The algorithms of classification include decision trees, artificial neural networks, and so on (Han & Kamber, 2001; Tou & Gonzalez,1974), in which artificial neural networks are the most widely considered technique in many business problems.

Artificial neural networks (ANN) attempt to simulate biological neural systems which learn by changing the strength of the synaptic connection between neurons upon repeated stimulation's by the same impulse (Li & Tan, 2006). Neural networks can be distinguished into single-layer perception and multilayer perception (MLP). The multilayer perception consists of multiple layers of simple, two taste, sigmoid processing nodes or neurons that interact by using weighted connections. In addition, the neural network contains one or more several intermediary layers between the input and output layers. Such intermediary layers are called hidden layers and nodes embedded in these layers are called hidden nodes shown in Fig. 1. Based on prior research results (e.g. Cybenco ,1998; Hung et al.,