

**FINAL YEAR MAJOR PROJECT REPORT (PROJECT STAGE – II)**

# **CUSTOMER CHURN ANALYSIS**

**TRACK – 2 (RESEARCH & DEVELOPMENT)**

Submitted in Partial Fulfillment for the Award of Degree of Bachelor of Technology in  
Computer Science and Engineering from Rajasthan Technical University, Kota



**COORDINATOR:**

**Ms. Deepa Modi**  
(Dept. of Computer Science & Engineering)

**SUBMITTED BY:**

**Ayush Bijawat (12ESKCS028)**

**MENTOR:**

**Mr. Basant Agarwal**  
(Dept. of Computer Science & Engineering)

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**SWAMI KESHWANAND INSTITUTE OF TECHNOLOGY,  
MANAGEMENT & GRAMOTHAN  
Ramnagar (Jagatpura), Jaipur – 302017**



**SWAMI KESHWANAND INSTITUTE OF TECHNOLOGY,  
MANAGEMENT & GRAMOTHAN  
Ramnagar (Jagatpura), Jaipur – 302017**

**CERTIFICATE**

This is to certify that Final Year Major Project Report (Project Stage – II) entitled  
“CUSTOMER CHURN ANALYSIS” has been duly submitted by

- Ayush Bijawat (12ESKCS028)

for partial fulfillment of the Degree of Bachelor of Technology of Rajasthan Technical University. It has been found satisfactory and hence approved for submission as Major Project during academic session 2015-2016.

Date: 16 May 2016

**COORDINATOR:**

**Ms. Deepa Modi**

(Dept. of Computer Science  
& Engineering)

**MENTOR:**

**Mr. Basant Agarwal**

(Dept. of Computer Science &  
Engineering)

**HEAD OF  
DEPARTMENT:  
Prof. Dr. C.M.Choudhary**

(Dept. of Computer Science  
& Engineering)



**SWAMI KESHWANAND INSTITUTE OF  
TECHNOLOGY, MANAGEMENT & GRAMOTHAN  
Ramnagar (Jagatpura), Jaipur – 302017**

**ABSTRACT**

The primary goal of churn analysis is to identify those customers that are most likely to discontinue using your service or product. In this dynamic financial industry, companies are progressively providing products and services with similar features. Amidst this ever growing competition, the cost of acquiring a new customer typically exceeds the cost of retaining a current customer. Existing customers are a valuable asset. Furthermore, given the nature of the financial services industry, where customers generally tend to stay with a company for a longer term, churning could lead to substantial revenue loss.

Customer value analysis is critical for a good marketing and a customer relationship management strategy. Customer retention rate has a strong impact on the customer lifetime value, and understanding the true value of a possible customer churn will help the company in its customer relationship management. Conventional statistical methods are very successful in predicting a customer churn. The goal of this study is to apply logistic regression techniques to predict a customer churn and analyze the churning and no-churning customers by using data from a personal retail banking company

**Domain Descriptors:**

7 CS 7      Web Development Lab  
5 CS 4      Database Management Systems  
4 CS 3      Software Engineering

**Keywords:**

**Generic Keywords**

Databases, Middleware, Programming

**Project Type Keywords**

Analysis ,Design, Implementation, Testing



**SWAMI KESHWANAND INSTITUTE OF  
TECHNOLOGY, MANAGEMENT & GRAMOTHAN  
Ramnagar (Jagatpura), Jaipur – 302017**

**DECLARATION**

We hereby declare that the report of the project entitled **CUSTOMER CHURN ANALYSIS** is a record of an original work done by us at **Swami Keshvanand Institute of Technology, Management & Gramathan, Jaipur** under the mentorship of **Mr. Basant Agarwal** (Dept. of Computer Science & Technology) and coordination of **Ms. Deepa Modi** (Dept. of Computer Science & Technology). This project report has been submitted as the proof of original work for the partial fulfillment of the requirement for the award of the degree of **Bachelor of Technology (B.Tech)** in the **Department of Computer Science & Technology**. It has not been submitted anywhere else, under any other program to the best of our knowledge and belief.

**Team Members:**

**Signatures:**

(12ESKCS028) Ayush Bijawat

## **ACKNOWLEDGMENT**

A project of such a vast coverage cannot be realized without help from numerous sources and people in the organization. We take this opportunity to express our gratitude to all those who have been helping us in making this project successful.

We are highly indebted to our faculty mentor **Mr. Basant Agarwal**. He has been a guide, motivator & source of inspiration for us to carry out the necessary proceedings for the project to be completed successfully. We also thank our project coordinator **Ms. Deepa Modi** for her co-operation, encouragement, valuable suggestions and critical remarks that galvanized our efforts in the right direction.

We would also like to convey our sincere thanks to **Prof. C.M. Choudhary**, HOD, Department of Computer Science & Engineering, for facilitating, motivating and supporting us during each phase of development of the project. Also, we pay our sincere gratitude to all the **Faculty Members** of Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur and all our Colleagues for their co-operation and support.

Last but not least we would like to thank all those who have directly or indirectly helped and cooperated in accomplishing this project.

### **Team Members:**

(12ESKCS028) Ayush Bijawat

## **CONTENTS**

<b>Cover Page</b>	<b>i</b>
<b>Certificate</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Declaration</b>	<b>iv</b>
<b>Acknowledgment</b>	<b>v</b>
<b>Contents</b>	<b>vi</b>
<b>List of Figures</b>	<b>viii</b>
<b>Chapter 1</b>	
<b>INTRODUCTION.....</b>	<b>2</b>
1.1 Churn.....	2
1.2 BigData .....	3
1.3 BigData Problem.....	3
1.4 BigData-Where is it?.....	4
1.5 Issues.....	4
<b>Chapter 2</b>	
<b>SOFTWARE REQUIREMENT</b>	
<b>SPECIFICATION.....</b>	<b>7</b>
2.1 Introduction.....	7
2.2 Requirement Analysis.....	8
2.3 Functional Requirement.....	9
2.4 Hardware & Software Req.....	10
<b>Chapter 3</b>	
<b>RELATED WORK.....</b>	<b>11</b>
3.1 Symbolic Technique.....	11
3.2 Machine learning Technique.....	11
<b>Chapter 4</b>	
<b>PROPOSED WORK.....</b>	<b>13</b>
4.1 Project Plan.....	13
4.2 Meet Hadoop.....	13

4.3 Other Frameworks.....	16
4.4 Redhat.....	19
4.5 Scala.....	20
4.6 Cloudera Manager.....	25
4.7 Apache Spark.....	34
<b>Chapter 6</b>	
<b>EXPERIMENTAL SETUP AND IMPLEMENTATION.....</b>	<b>35</b>
5.1 Spark Streaming.....	35
5.2 Spark Mllib.....	35
5.3 Mllib Algorithm for churn analysis.....	36
5.4 Customer Churn Analysis in Telecommunication Sector.....	44
<b>Chapter 7</b>	
<b>RESULT ANALYSIS.....</b>	<b>56</b>
<b>Chapter 8</b>	
<b>CONCLUSION AND FUTURE WORK.....</b>	<b>57</b>
<b>REFERENCES.....</b>	<b>58</b>

## LIST OF FIGURES

Figure 1 Types of data.....	3
Figure 2 Hadoop Architecture.....	15
Figure 3 Cloudera Manager .....	25
Figure 4 Dataset Used. ....	47
Figure 5 Decision tree code .....	48
Figure 6 Naïve Bayes Code.....	49
Figure 7 Execution.....	51
Figure 8 Accuracy .....	52
Figure 9 Trained Model.....	52
Figure 10 Result.....	53