**A Dissertation**

**On**

# ANALYSIS OF WEBSITES USING WEB METRICS

**Submitted in partial fulfillment of the requirement**
**For the award of degree of**
**MASTER OF TECHNOLOGY**
**Software Engineering**

**Submitted By:**

**RAVI SHARMA**

**2K11/SWE/13**

**Under the Guidance of**
**Dr. Kapil Sharma**
**Associate Professor**
**Computer Engineering Department**
**Delhi Technological University**


**DEPARTMENT OF COMPUTER ENGINEERING**
**DELHI TECHNOLOGICAL UNIVERSITY**
**(*Formerly Delhi College of Engineering)***
**Bawana Road, Delhi-110042 JULY, 2013**

# CERTIFICATE

This is to certify that **RAVI SHARMA** has carried out the work presented in this thesis report entitled, **"Analysis of Websites Using Web Metrics"**,under my supervision. The report embodies result of work and studies carried out by him and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else.

**Dr. Kapil Sharma**
Associate Professor
Computer Engineering Department
Delhi Technology University, Delhi

**DATE:-…….**

# ACKNOWLEDGMENT

First and foremost, praises and thanks to the God, the Almighty, for his showers of blessings throughout my M.Tech to get it completed successfully.

I want to express my sincere thanks to my HOD, Dr. Daya Gupta, Department of Computer Engineering, Delhi Technological University, Delhi for providing well equipped infrastructure support. I would like to express my deep and sincere gratitude to my project mentor Dr. Kapil Sharma, Associate Professor, Computer Engineering Department, Delhi Technological University, Delhi for giving me the opportunity to do research and providing invaluable guidance and constant encouragement throughout the project.

**RAVI SHARMA**

2K11/SWE/13

M.Tech (Software Engineering)

Department of Computer Engineering

Delhi Technology University, Delhi

## Table of Contents

# List of figures

# List of Tables

## ABSTRACT

World Wide Web is a source of enormous information and has massive influence on our lives. A large number of websites nowadays are designed without sufficient resources and professional skills. Therefore evaluating the quality of a website is a very important issue, further for success of any web; quality is one of the most important attribute. Various guidelines, tools and methodologies have been described by many authors to maintain the quality of a websites but their implementation is not much clear. Web metrics are used to measure various attributes of websites quantitatively and can be used to evaluate the quality of a website. So it is important to assess the website to enhance the quality of websites and web development process.

In this research work twenty web page metrics were computed using an automated tool WEB METRICS CALCULATOR developed in ASP.NET language. Data from websites of various categories from pixel awards of year 2010, 2011 and 2012 have been collected, to categorize the websites into good or bad. Logistic regression and 10 machine learning techniques (Bayes net, Naïve bayes, Multilayer perceptron, Adaboost, Decision table, Nnge, Part, Bf-tree, J-48 and Random forest) were employed. Out of all these techniques results shows that area under ROC curve is greatest for Random forest within range .842-.891 for all year data set, so performance of Random forest model is greater as compared to all other models.