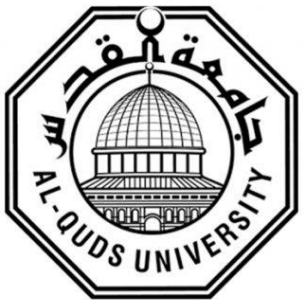
Deanship of Graduate Studies

Al-Quds University



Performance Evaluation of Websites Using Machine Learning

Mohammad Rebhi Rateb Ghattas

M.Sc. Thesis

Jerusalem - Palestine

1440 - 2019

Performance Evaluation of Websites Using Machine Learning

Prepared By:

Mohammad Rebhi Rateb Ghattas

B.Sc. Information Technology from Palestine Polytechnic University - Palestine.

Supervisor: Dr. Badie Sartawi

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science / Department of Computer Science Faculty of Science &Technology / Deanship of Graduate Studies / Al-Quds University

Jerusalem - Palestine

1440 - 2019

Al-Quds University
Deanship of Graduate Studies
Computer Science Department

Thesis Approval

Performance Evaluation of Websites Using Machine Learning

Prepared By: Mohammad Rebhi Rateb Ghattas

Registration No: 21610053

Supervisor: Dr. Badie Sartawi

Master Thesis submitted and accepted, Date 23 June 2019

The names and signatures of the examining committee members are as follows:

1- Head of Committee: Dr. Badie Sartawi Signature .....

2- Internal Examiner: Dr. Nidal Kafri Signature .....

3- External Examiner: Dr. Jamil Itmazi Signature .....

Jerusalem – Palestine

1440 / 2019

Dedication

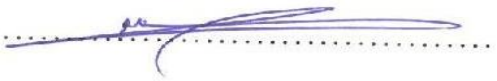
I dedicate my thesis to my family. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears, and who never left my side and are very special. I also dedicate this dissertation to my beloved wife, and I will always appreciate her patience, encouragement and all she has done for me. I also dedicate this dissertation to my brothers, sister, my children and friends who have supported me throughout the process.

Thank you all

Mohammad Rebhi Rateb Ghattas

Declaration

I certify that this thesis submitted for the degree of Master is the result of my own research, except where otherwise acknowledgment and that this thesis or any part of the same has not been submitted for a higher degree to any other university or institution.

Signed: 

Mohammad Rebhi Rateb Ghattas

Date: 23 June 2019

Acknowledgment

All praise is to Allah.....

I would like to express my sincere gratitude to my supervisor, Dr. Badie Sartawi. He has offered me great freedom on choosing my favorite research topic and developing my research interests, and continuously provided me help and encouragement with extensive knowledge. His ideas have been a source of inspiration for this work. I also thank the examining committee, all my colleagues and relatives for their support. Finally, I also thank all the lecturers.

Many thanks to my mother and my wife for them praying and providing support and to my father thank you very much.

Abstract

The Web is playing a major role in various application domains such as business, education, engineering, and entertainment. As a result, there are increasing interests in designing and developing an effective website to deliver a high degree of performance. Therefore, automated support for web designers is becoming more important to evaluate websites performance. Hence, many of the previous studies tried to evaluate websites performance by developing a static model and it's unless used for more domain.

The aims of this thesis are: (i) to explore the best metrics that most affect website performance; (ii) propose a dynamic model for performance evaluation of websites by using machine learning that called is PEML ; and (iii) to help webmaster and decision-makers to know what improvements are needed to enhance the performance and the final relative weights of metrics in the level of the hierarchy.

This research proposes a dynamic model to performance evaluation of websites using machine learning method by applied two regression methods experiments namely, multiple linear regression and support vector machine regression on the same dataset that collected, to take the best performance of regression methods to generate weight for every metric and then developing a new dynamic model to evaluate websites performance.

Keywords

website performance, regression, machine learning, web metrics, support vector machine, multiple linear regression, evaluation, RapidMiner.

Table of Content

List of Tables

List of Figures

List of Appendices

List of Abbreviations

Abbreviation	Abbreviations Full Name
LR	Linear Regression
SVM	Support Vector Machine
LWM	Linear Weightage Model
AHP	Analytical Hierarchy Process
FAHP	Fuzzy Analytical Hierarchy Process
RP	Rapid miner

Chapter 1

Introduction

This chapter introduces the thesis. It describes the problem statement, purpose, research questions, limitations, contributions, methodology and organization of the thesis.

1.1 Research Overview

Lately, we have got become witness to an important alteration of our lives to a worldwide with the incipience of the web era. The web is an increasingly more vital asset in many sides of life: government, education, commerce and more [3]. Hence, Websites are a key element in obtaining the right information about the institutions. However, when it comes to a huge number of synchronous users these websites performance decreases considerably.

Utilizing the web devices many institutions become been able to raise their being customer-focused and their attributes of services and products. The analysis of the web site is currently thought to be an essential facet of attracting customers' attention[3]. In this study, it is logical to explore metrics into measure the performance of websites, whether to study the communication efficiency that they represent or in order to build useful appraisal metrics.

As result of the above requirements, it is important to provide a method to evaluate the performance quality of websites which include various technological and logical factors. Each definition of performance quality from literature leads to lists of criteria about what constitutes a good quality website and how to measure the performance [8]. Therefore, it is

important to build a model into evaluation websites performance, thus ensuring the development of modern websites and keeping abreast of modern technology.

This study employed machine learning to build a mathematical model approach to evaluating the performance quality of websites. In this thesis, we suggest an method based on appropriate metrics for evaluating websites performance.

This study proposed to build an understandable and applicable dynamic model for evaluating websites performance by using previous studies as a case study. By establishing a practical model, it is expected that organizations can better understand whether a given website can meet the expectations of its users, they serve in order to grow their satisfaction level.

1.2 Problem Statement

The website is becoming more important each day for conducting business, sharing information, and communication. Each passing day, the number of organizations, companies, and individuals propagation their websites is increasing.

Hence, the task of evaluating and improving the websites can be intimidating, considering the number of websites available, and the frequency of updates. As a result, automated support for web designers is becoming more important to evaluate websites performance. It is necessary to provide an easy method to performance evaluation of websites, which include several technological and logical factors, as a contribution to addressing this need.

Therefore, the problems in this study are : How to determine the best metrics that affect websites performance, what are the weights of every metric of website performance, how can arrangement for metrics that more affect websites performance in the level of the hierarchy, and how to evaluate the performance of the proposed approach.