ACKNOWLEDGEMENTS

We sincerely thank Mr. Jim Mathew Philip Assistant Professor (Sr. Gr.), CSE Dept., for guiding through our project for the successful completion. We also thank our beloved principal, Dr. M. Paulraj Ph.D, for providing us with the resources required for the successful completion.

We also thank our IDP Course Instructor, Mr. Devendra Kumar, Assistant Professor, CSE Dept., for properly conducting this course. Lastly, we would also like to thank our Reviewer Mr. Rajesh, Assistant Professor, CSE Dept., for reviewing our project sessions.

APPROVAL AND DECLARATION

This project report titled **Web Server Log Analysis System** was prepared and submitted by **Vijay Krishnaa S (1702128), Praveen K (1702082), Santhosh Kumar P (1702106)** and has been found satisfactory in terms of scope, quality and presentation as partial fulfilment of the requirement for the **Bachelor of Engineering (Computer Science and Engineering)** in Sri Ramakrishna Institute of Technology, Coimbatore (SRIT).

Checked and Approved by

_____

Mr. Jim Mathew Philip
Assisstant Professor

Department of Computer Science and Engineering
Sri Ramakrishna Institute of Technology, Coimbatore – 10
October 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS OR NOMENCLATURE

1.	IP	Internet Protocol
2.	UA	User-Agent
3.	GUI	Graphical User Interface
4.	CSV	Comma Separated Values
5.	HTML	Hyper Text Mark-up Language
6.	DoS	Denial of Service
7.	SEO	Search Engine Optimization

WEB SERVER LOG ANALYSIS SYSTEM

ABSTRACT

This project analyzes the log files of the web server and generate visuals for better understanding of the information in the logs. Log files from the web server are fed as input to the system and it reads the entire log file line by line and stores the data in memory and performs analysis on the data and then the final part of the program is to generate the visuals based on the analysis of the data. The data that is visualized: IP address and the total number of requests from them, request methods like GET, POST and the count of each request methods, the User-Agents used the clients, finally the files that are requested by the clients and frequency of their access. This project can analyze logs from the Apache Web server. From the output provided by the project, we can perform a wide range of tasks. If we find a specific User-Agent requests are increasing, we may develop the features for that specific User-Agent. If we see some alarming rates of requests from a specific IP address, it may be an attempt of Denial-of-Service which may need further investigation, it is also useful for Incident response.