National College of Ireland

Evaluate the performance of AMP and non-AMP websites using machine learning algorithms

MSc Research Project
Data Analytics

Bhumi Patel
Student ID: x18114865

School of Computing
National College of Ireland

Supervisor: Manuel Tova-Izquierdo

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Bhumi Patel
Student ID:	x18114865
Programme:	Data Analytics
Year:	2019
Module:	MSc Research Project
Supervisor:	Manuel Tova-Izquierdo
Submission Due Date:	7/04/2019
Project Title:	Evaluate the performance of AMP and non-AMP websites using machine learning algorithms
Word Count:	7178
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	16th September 2019

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Evaluate the performance of AMP and non-AMP websites using machine learning algorithms

Bhumi Patel
x18114865

Abstract

Due to the enhanced amount of volume and complexity of modern websites, a number of user experiences on mobile devices have been showing a downward trend. To overcome this problem Google AMP project has been conducted to load the content more efficiently which would result in improving user experiences. This research proposes the introduction of Accelerated Mobile Pages(AMP) and to what extent does it can recover the time consumed by a web user on it. To understand mobile Web page performance, this research conducted : (1) an in-depth pairwise evaluation of loading a page of an AMP oriented USA websites URL with low, high and medium technology, (2) and analyzed the outcomes in two different time, Morning and Evening Based on several factors (Monthly Visitor, Page View per visitor, average visit duration and many more) of AMP and an Non-AMP websites URLs using different machine learning algorithms such as Support Vector Machine, Random Forest, Naive Bayes, Decision Tree, and Deep Learning with Keras. The performance of classifiers was found to improve with an increase in the sample of training data. The best performing design was SVM when hyperparameter tuning optimization performed with the specificity of 0.9221. The strategy is effective in improving the efficiency of AMP websites and introducing AMP on multiple websites can enhance the time invested by a web user.

1 Introduction

The web consists of numerous web pages which are accessed by a set of connected linked documents by web browsers. The browser has made available a broad range of services including research, productivity, entertainment, cultural, and many more (Bocchi et al. (2016)). Demand has increased with mobile or tablet devices while accessing the web page[1]. Since October 2016, many websites were loaded on a mobile as compared with the computer. The main reason is that the consumer spends more than three times on smartphone devices as compared with the other systems. The amount of subscribers shows a positive increasing trend thereby passing 7.7 billion people (Byungjin Jun and Bischof (2019)).

While each site should be moved to portable browsing, many of the web sites have been intended for desktop computers (Nejati and Balasubramanian (2016)). According to (Ruamviboonsuk et al. (2017)) the time needed to load a 4G site is around 14 seconds.

[1] `http://gs.statcounter.com/platform-market-share/desktop-mobile-tablet`

Every second is important here. Any delay in a second-page can cost around millions of lost sales to online businesses per year. Therefore enhancing the speed of a particular website is a major factor in company life (Nagy (2013)). Which means, if the website loads the page slowly, it will not only lose spectators (Singh and Verma (2014)) but also obliquely affect traffic which decreases.

To what extent can AMP's implementation on different websites enhance the time a web user spends on it? Facebook Instant Articles (FIA) were presented to increase the loading of the news story. Out of which some browsers implemented proxy-based compression (e.g. Opera Mini, Flywheel) mechanisms to enhance user experience on slow networks (Netravali and Mickens (2018)). Google's Accelerated Mobile Project (AMP) is a new developing technology. Google Research reports that around 53% of the website visitors will leave the respective site if your web page is carrying time to load for more than 3 seconds (O'donoghue (2017)). Due to the significance of efficiency, the research work on AMP still exists.

The purpose of this study is to discover how speed is crucial on the website? It would also be interesting to find how much does it affects the time of the user, using several ML algorithms. Due to much-advanced technology, various parameters influence the loading of the Web page more. Previously, no ML models were used to evaluate the performance of the website. Thus by fulfilling the above questions raised, the novel strategy would be taken for the implementation of various distinctive ML algorithms. Our main aim of this examination is to determine the below problems.

1. How will AMP application be effective when assessing the efficiency of distinct sites based on the features that influence user experience?

2. Which of the ML methods can achieve better results for non-AMP and AMP web pages in terms of accuracy?

To address the above research question, 50 AMP URLs includes 3 categories of U.S. based web pages techniques that are high, low and medium with non-AMP U.S. based URLs using various parameters such as loaded time, page size and many more, as well as analyzing multiple impact factors that may have caused a website to slow down. GTmetrix and SimilarWeb tools are utilized to collect data on two distinct time morning and evening. Once the information is ready, serval ML models like as DT, SVM, Naive Bayes, Keras Deep Learning, and Rf would be conducted.

1.1 Motivation behind this research

Visitors are in the form of traffic which implies that traffic impacts reputation and can create revenue (Ren (2018)). Hence, overcoming site loading time and making it as quick as possible is compulsory. Website quickness will be one of the most vital variables in the online business world. The positive results of the models will not only help to solve the problem of page speed but also it will indirectly increase in revenue and growth rate. The results obtained from the above implementation will indirectly improve the future trend of web analytics.

AMP, however, is a fairly fresh idea, and there has been no comprehensive study on it. The following research document is divided into various sections. In Section 2, Review of Literature to the domain of this subject has been done which explains background understanding and a number of prior web results and AMP studies. Section 3 gives information about methodology using KDD approach. Section 4 and 5 gives the detailed information regarding suggested implementation, Evaluation and Results of modes. Lastly in section 6 conclusion is explained briefly.

2 Related Work

This section introduces the background of web performance with AMP and various techniques for efficiency measurement and enhancement. First, Chapter 2.1 describes the response of the web, how it exists and how it can be measured and improved. Chapter 2.2 not only explains the performance but also several approaches to AMP. Lastly, the final Chapter 2.3 presents tools for site review measurement.

2.1 Evaluation of non AMP websites

Chapter 2.1.1 examines the effectiveness of the website and how does the loading speed impacts the output of the website on different numerous devices. The digital era quality is presented in Chapter 2.1.2.

2.1.1 Websites effectiveness on various devices

(Yu and Kong (2016)) has concluded that mobile sites and layout of websites can have an important effect in terms of obvious user-friendliness, reading the website time and its overall general knowledge. User activity can depend on the structure of portable sites. To comprehend the impact about the web pages and the home page designs on a small screen, researchers have examined 2 sample studies to see how the consumers estimate interface designs that are frequently used on a web page with smaller screens. The suggested study generated two main rules, one for multipage to produce text-based information, and the other for the website with a thumbnail that helps better knowledge processing than the increasing record of views to evaluate and examine the famous mobile news web interface. In this research, the study promotes finding the impact factor of a structured web page. Is speed could be considered as a significant factor while determining the effect on structured websites? Suppose the loading of a page on mobile devices slows down than affects web users?

The speed site while loading reflects the performance. Mobile pages are loaded more slowly compared with pages which are not mobile oriented (Nejati and Balasubramanian (2016)). In the web page quality standards and scales, each domain of website researchers has established variables which are correlated with fully load times. (Johanna Dunaway and Paul (2018)) came on the were not identical to computer news on the mobile device. Myriad devices have various exploration, referral, traffic speeds for the users. The experiments of the travel website studies indicate that customers hit various web pages per stay, spend less time on a site, and bounce rate on mobile is also higher. (Pan and Wang (2016)). Another research Stringam and Gerdes (2019) illustrates that loading on hotel web pages on a small screen was considerably slower than on desktops. If the website

takes long times to load it may gives the outcomes in perspective of the customer to left the page and indirectly dropping out of the website rate and affect the revenues. Separate research demonstrates that frequency is not only affected travel websites but is also helpful on the mobile device for multiple domains such as e-commerce site, hotel web pages, live news channel websites and much more.

2.1.2 Used machine learning to measure QoE in mobile browsing

Using different kinds of websites on various types of smartphones has become essential in everyone's day to day lives. (Barakovic and Skorin-Kapov (2017)) shows the effect of structure based on distinctive variables such as influence factors (IFs) and different features of web browsing Quality of Experience (QoE). (Barakovic and Skorin-Kapov (2015)) recently published multidimensional QoE on mobile devices as part of browsing, thematic, data and email gateways. The mutual relationship between QOR Ifs and QoE characteristics is deeply correlated. It is then further quantified by using multiple linear regression models. Meanwhile, the effect on mobile web browsing of only chosen QoE variables has been discussed. In this research, various variables are considered which can be treated at different degrees. Thus, this study will help in recognizing the significance of each factor in terms of browsing QoE. It also gives insight into the incentive to execute various ML on various QoE variables to achieve a precise outcome.

Expressing QOE on browsing seems quite complex, as sites are presently much complicated while some of them have numbers of artifacts as well(Weiwang Li and Zhao (2017)). The following compromise is faced by evaluating Web users (WebqoE) quality of experience. In studies (Bocchi et al. (2016), Xiao Sophia Wang and Wetherall (2016)), the WebQoE assessment of today expresses QoE through the completion of the page, i.e. onLoad time. Ren (2018) indicates an adaptive predictive QoE model. It will be helpful for web browsers and a broad spectrum of internet rendering-supported applications. This article presents a novel QoE assessment model for user experience. It takes account energy consumption and web surfing delays that are different from the regular user experience. Many types of machine learning algorithms are used to predict the results among which SVM performs well. The study makes a vital contribution to the imminent model based on ML, Experimental results indicate that this method gives the results 47.3% greater QoE than the conventional Interactive System Governor It also overcomes sophisticated WS (web-aware schedulers) experience, which Interactive only raises 29.1%.WS performs better than the method performed when considered that the average speed of 1.187x for the load time. This strategy does not improve the speed on mobile devices of advance websites. This document will lead to a search for methods which will work well while loading the specific website.

According to the above research conducted, various factors are considered while measuring the performances of websites. The Web service experience relies on several determinants. Which parameters have an important impact on website efficiency and to what extent does it affect user experience? Which factor impacts more when using machine learning algorithms to load non-AMP websites? Many complicated apps on smartphones have now been created and operated. Faster-loaded technology must be found to check the efficiency of sophisticated websites and assess where there is a lack of a non-AMP website to enhance performance speed.

2.2 Accelerated Mobile Pages

When your website takes more than 3 seconds to load,then the Google Research claims that around 53 percent of the respective visitors would be leaving the particular website (O'donoghue (2017)). By restricting the HTML, CSS, and JavaScript with optimized resources management and intelligent caching, AMP resolves the internet performance issues on mobiles, while offering a structure for delivering quick web pages that are capable, media-rich and flashy.

In the field of web performance, comprehensive study was carried out. But AMP is a fairly fresh idea and no extensive study has been done on it. Therefore, there is a study gap that needs to be resolved with respect to AMP.

2.2.1 Efficiency of AMP web pages

The application of the practical life of freshly advanced information technology at this digital era is an essential component of every marketing team in all fields.(Miklosik (2017)) examined accessible mobile optimization techniques and evaluated the options for applying AMP mobile device on specific webpage formats. It has also explored their business marketing implementation, such as the Visegrad countries.

Table 1: Comparison of AMP and non-AMP pages (Miklosik (2017))

	Performance	**Loading time**	**Faster than other**	**Total Page size**	**Total Requests**	**Tested from**
AMP	78	696 ms	94%	806.0	60	Dallas
non-AMP	70	3.60s	46%	3.2MB	205	Dallas

A major index when comparing AMP efficiency with non-AMP website parameters is the speed index parameter (Byungjin Jun and Bischof (2019)). This test is carried on 3 common QoE metrics, Load time, First Byte time and Speed Index. These metrics are used for assessing the effect of AMP on mobile QoE. In this paper, firstly AMP's effect on user QoE is explained in much detail. It is based on a corpus of over 2.100 AMP websites and non-AMP counterparts are based on keyword-based searches which are trendy. Outcomes indicate AMP considerably enhances SI which results in 60% less SI on average than unprefixed AMP websites. This benefit is increased even quicker than non-prefetched AMP pages when AMP web sites are prefabricated with pages up to over 2,000 ms. This test outcome concentrated more on SI rather than PLT because the primary goal of the AMP is to enhance user perception. This outcome promotes the precision of performance with extra user experience-influencing parameters.

Google has freshly published the hosting of more than 2B+ websites which includes higher than 900 K domains. The AMP aims to reduce websites complications and load times. In developing nations, owing to the limited internet connectivity, AMP provides particularly interesting possibilities to improve website user experience in these nations (Phokeer et al. (2019)). Research of local news websites evaluates the efficiency of Google AMP in Africa. Researchers then assessed the benefits of AMP by establishing a scarce and low-performing atmosphere for local content access in developing areas. In Africa,

this research shows that AMP can decrease page load and sizes by the factor of 3 or 8. The primary barrier facing this suggested methodology was : 1) Some web sites were not available and are removed from list 2) sites analyzed news pages articles only 3) Factors considered in the research were quite limited.

Accelerator tool was used to examine the efficiency of e-business websites in further studies. The e-commerce industry involves all kinds of characteristics and company activities using electronic information for enhancing corporate earnings (Agus Wibowo1 and Mufadhol (2018)). It is important to use an E-commerce accelerator tool in E-business. It directly enhances web service efficiency. The test operation is done with the GTmetrix open-source tool, including the required period and efficiency before AMP is used. This is very crucial when accessing the website through a resource-limited smartphone. It will make the website fast, simple and light.

2.2.2 Web optimization on mobile device

As there is an important impact on the user experience of page loading, both sector and academia have put a lot of effort into identifying the root causes of web bad effectiveness. KLOTSKI is nothing but a server and a mobile web optimizer push (Butkiewicz et al. (2015)). KLOTSKI consists of two parts, namely a back end and a front end proxy. Klotski pushes the high priority items on the web page when the browser hits the HTML page. Klotski defines high priority objects in the offline stage using a utility function. VROOM is the same as Klotski-like, but instead of receiving server pushes, clients prefix data (Ruamviboonsuk et al. (2017)). A VROOM server utilizes preload link headers to show customers with helpful HTTP answers that items can be preloaded. As a result, VROOM reduces by more than five seconds the average load time for a news site and sports locations.

Mobile page loads are increased in AMP O'donoghue (2017) by requiring that websites be loaded faster in a limited dialect of CSS, HTML, and JavaScript. For instance, AMP fixes all internal $<$ script $>$ material in an async property to proceed with HTML browser. AMP makes a website to have one CSS file using $<$ style $>$ tag with a size of fewer than 50 kB. Prophecy holds random pages using applications such as HTML, CSS, and JavaScript which are arbitrary. However, prophecy can be implemented to AMP websites.

Does the question still arise about which are the most important factors for successful mobile optimization? Accessibility of picture and video, visible text, click ability and interactivity, speed of loading, compatibility, and minimization of the pop-up are some of the factors which motivate the Mobile optimization. VROOM, KLOTSKI, and AMP are recent technologies that enhance mobile optimization how will AMP help increase page efficiency depending on user activity parameters? This research conducted a novel approach which is to monitor the web pages of AMP and non-AMP websites with ML algorithms. In determining the parameter factor, several tools are required to verify speed.

2.3 Measuring the Performance of Websites using tools

Having so much option of the website is essential that you can not use the first one which visited, you go to the next website. The primary reason for the scalability is essential. You will leave the website instantly if you do not know how to navigate your site quickly. Many website owners and visitors can use instruments to conduct website examsPunjab's college websites Kaur et al. (2016) have been evaluated using four automated tools with similar outcomes of several variables, as shown Table 2.3.

Table 2: Measuring parameters using tools(Kaur et al. (2016))

Parameters	**Efficiency**	**Requests**	**Speed**	**Load Time**	**Page Size**
GTMetrix	✓	✓	✓	✓	✓
WebsiteGrader	✓			✓	✓
SimilarWeb	✓	✓	✓		✓
Site Speed checker	✓	✓	✓	✓	✓

The charging page speed represents the efficiency of the website. It affects user experience and satisfaction considerably. Evaluated web load speed with various tools (Bartuskova et al. (2016)). In the in-depth literature research and website performance testing facilities evaluation, including characteristics, interfaces, and additional settings, from different perspectives. The methodology suggested shows that the most trusted services were GTmetrix and SimilarWeb. (H Jati and Wardani (2018)) To guarantee the validity of the information gathered, parameters such as SimilarWeb and GT matrix were considered in this proposition. It creates the usability of the Indonesian University website. Data were acquired for around 15 times over three months of span using the website tools.

SimilarWeb and GT Metrix have been used in these studies from the above research. These two tools would collect performance variables for non-AMP and AMP-based websites. As per the above research, execution of AMP and non-AMP websites can be evaluated using unique elements. Web benefit participation relies on several elements for illustration waiting times, such as the ability to use the service itself and effects on QoE. Many complicated apps have been developed and worked on smartphones nowadays. **It is important to observe faster-loaded technology that will estimate the results of advanced websites and interventions where the non-AMP websites are unable to magnify performance and how useful the AMP can be used to improve website efficiency based on user experience parameters?** This research examine and compare AMP websites with non-AMP websites using multiple machine learning algorithms such as Naive Bayes, DT, SVM, and RF and Deep Learning with Keras. The outcome of this model solves the question of how to find important performance effect variables on AMP and non-AMP locations. These methods are used to predict the precise outcomes of real-world information and provide complete accuracy.

3 Methodology

The above-literary review examined the question of slow load speeds on websites. This chapter highlights the Knowledge Discovery Databases strategy for pre-processing information and proposes models using different measuring equipment and executed how these tools are used to analyzed the efficiency of non-AMP and AMP U.S. based websites. Based on the information and the accuracy of the outcomes, various mythological stages are taken into account. This study selects the methodology of Knowledge Discovery Databases. The following stages of KDD (Fernández et al. (2018)) are described in (Figure 1).

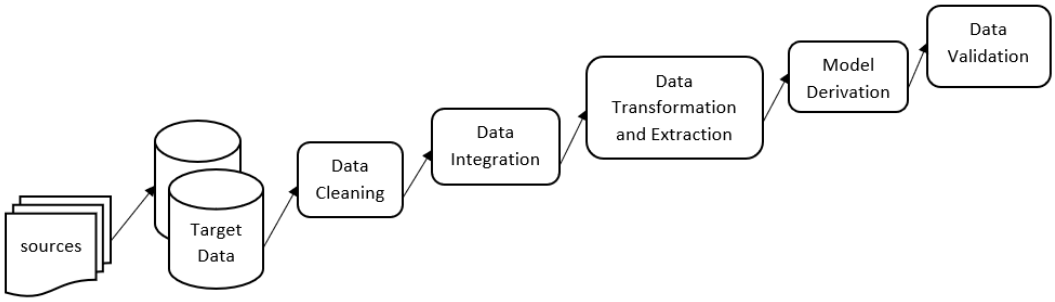

Figure 1: Flow of the process using KDD

3.1 Sources

The first phase of the KDD method collect, analyse and verify the information to organize data. To conclude the above research problems implementation must be divided into two-part AMP and NON-AMP. AMP related U.S. website URL are downloaded into CSV format from BuiltWith using the free trial. BuiltWith[2] is a website profiler that presents online approval of the technology, eCommerce data, and usage analytics. NON-AMP U.S. based websites URL are collected from government-site[3] which is further downloaded into CSV format. This information arrives from a unified Google Analytics that offers a gateway to find the people interacting with websites of government. In this research, AMP and Non-AMP URL of data are collected from GTmetrix and Similarweb tools.

3.2 Target Data

Using a public source of data, AMP and Non-AMP U.S. based websites URL are selected to gather information from performance testing tools. GTmetrix and SimilarWeb are two measurement tools used to fetch the different parameters such as Yslow_Grade, PageSpeed_Grade, Total_Requests, Avg. Visit Duration, Pages Per Visit, Bounce Rate, and many other attributes. AMP-based websites further divided into 3 categories based on technology such as high, low and medium. All the information is downloaded manually from the tools. Three different technology of URL contains the following parameters to perform model. Several websites are chosen to evaluate web performance, including AMP and responsive websites on a various category such as News, Movies, Television, Financial, World Football, Sports, Basketball, and a lot more.

[2]https://pro.builtwith.com/
[3]https://analytics.usa.gov/

3.2.1 Tools for data acquisition

Once the website URL is prepared, it is requisite to forward estimates of performance to their specific sites. GTmetrix[4] is the most perfect for evaluating the performance of specific sites. SimilarWeb[5] tools are used as a free trial. GTmetrix data are freely available and can be downloaded into CSV format. GTmetrix can check speed test without registration. It gives an appraisal of an HTML page loading speed with all objects including images, RSS, and many more. SimilarWeb data is generated from Apr 2019 to Jun 2019 and downloaded in PDF format. It contains attributes like Total Visits, Mobile Device, Desktop Device, Monthly visits, Avg. Visit Duration, Pages Per Visit and Bounce Rate are affect the user experience of the sites whereas GTmetrix contains the attributes such as PageSpeed Score, YSlow, Timing and Waterfall.

3.2.2 Data Collection

The information obtained from data source and measurement tools is available from Apr 2019 to Jun 2019. AMP oriented U.S. based data sets are distributed into three different categories based on technology spend i.e high, low and medium. Each categories has a 50 different rows with 28 columns as discussed above. AMP is most popular for websites speeds. So there are lots of AMP oriented websites URL lists such as Amazon, Facebook, YouTube, Yahoo, and many more publicly available which are already faster and no need to experiment. Main motive of this research is to check at what extent AMP is useful to increases the speed. To fulfill this requirement Non-popular AMP and Non-AMP URL are considered which is difficult to get. Because of this difficulty only 150 distinct U.S. based AMP URL were considered on which GTmetrix and SimilarWeb tools were used to obtain performance parameter of websites at two different time duration which are morning and evening as seen in Figure 2.

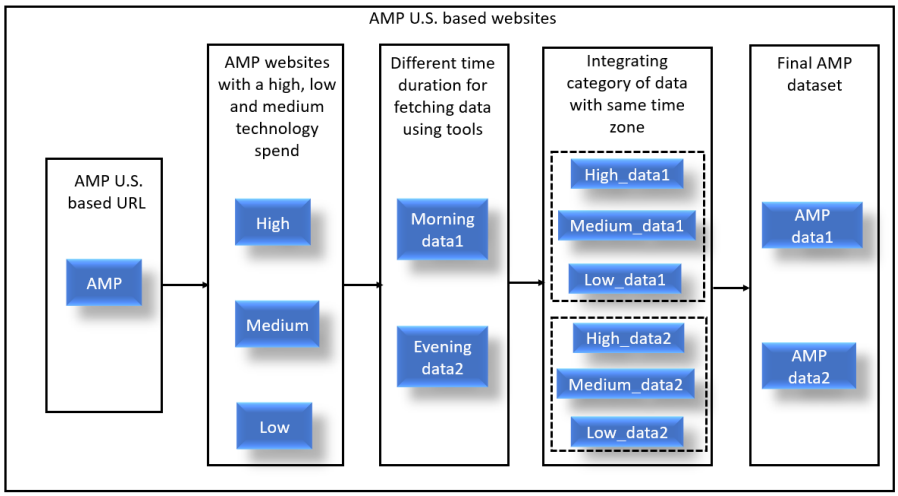

Figure 2: Data Collection

Non-AMP data are also loaded into these tools at two different time duration. Once the data was generated further step to consider is data pre-possessing.

[4]https://gtmetrix.com
[5]https://www.similarweb.com/

3.3 Data Cleaning

It is important to remove unnecessary data after targeting information from a separate data source. In this phase, noise and inconsistent information are deleted. The U.S. based websites performance data is obtained from BuiltWith and USA gov analytics site which carry an excessive amount of information. Real-world information is often incorrect, incompatible, missing and can have a lot of errors. Such instances may result in bad performance and may have an impact on accuracy. In this step, the problem is fixed by removing the missing value from the information set. The data set cleaning process is accomplished by using the R programming language. R library such as dplyr, tidyr, pdftools, etc. are used to improve the outcomes. Pdf format data has been fetched using function pdf_text and convert into standardizing CSV format. Special characters have been removed using gsub function. This information purification and standardization decreased the danger of duplication and made further predictive analysis easy.

3.4 Data Integration

Once the necessary information has been gathered from the above stage, the next step of the KDD process is the integration of data in which various sources of information are combined. GTmetrix and SimilarWeb tools are used to calculate attributes and check the impact factor of attributes on sites. Using R, SimilarWeb data has been fetched and combined with GTmetrix parameters. As per the Figure 2 four CSV file would be generated(AMP data1, AMP data2, Non-AMP data1, Non-AMP data2). Two different time zone of data are downloaded (for example AMP data1 and Non-AMP data1) and integrated to perform classification of the model. Once the U.S. based websites URL are combined, further, a phase conduct data transformation and extraction.

3.5 Data Transformation and Extraction

Through summary or aggregation operations, data is then processed or consolidated into mining formats. Data pre-possession is a method of data mining that involves transforming raw information into an understandable format. Data from GTmetrix and SimilarWeb tools can be downloaded in CSV and PDF structure which need to be organized and turned into a well-defined format. Using R, a semi-structured form of PDF data was fetched using PDF_tool, and save it in CSV format. All the fetched information is transformed into label form. Each tool has a range of parameters and To find outcomes compared to others, the most significant parameters must be analyzed to extract the results of the comparison.

3.6 Model Derivation

Data collection and handling strategy are one of the main components of this project. Compared to unsupervised technology, supervised ML algorithms are used for the labelled data. A fresh QOE strategy has been introduced (Ren (2018)) and a series of supervised methods have been assessed to define website efficiency. Classification model is used because this study is based upon the comparisons of AMP or Non-AMP websites URL. Datasets are split into a two part, 70% training and 30% testing to perform ML algorithms. DT, RF, Deep learning with Keras, SVM, and Naive Bayes are evaluated a predictive model.

3.7 Data Validation

Various methods are used inn this study to analyze the impact factor of AMP and non-AMP websites on users. The next stage is model validation that evaluates the outcome of the model after pre- processing and model derivation. Each model has a number of characteristics, which help to accomplish efficient results. The efficiency of the model would be evaluated using F1-Score, precision, recall, confusion table, error rate, accuracy, specificity and sensitivity. In order to verify the accuracy, misclassification costs are used.

4 Implementation

For evaluation of U.S. based websites efficiency with attributes, the implementation would be necessary. Implementation of this research has been followed as per Figure 3. which demonstrate the basic structure of this research evaluation. The first step of data is data collection which is in the form of CSV and PDF format. Then in R studio, data would be loaded and integrated. A further stage was data pre-processed to perform classification model. EDA (Exploratory Data Analysis) was used to explore the attributes of data for deep understanding. The different classification model is used and compared to identify the best results.

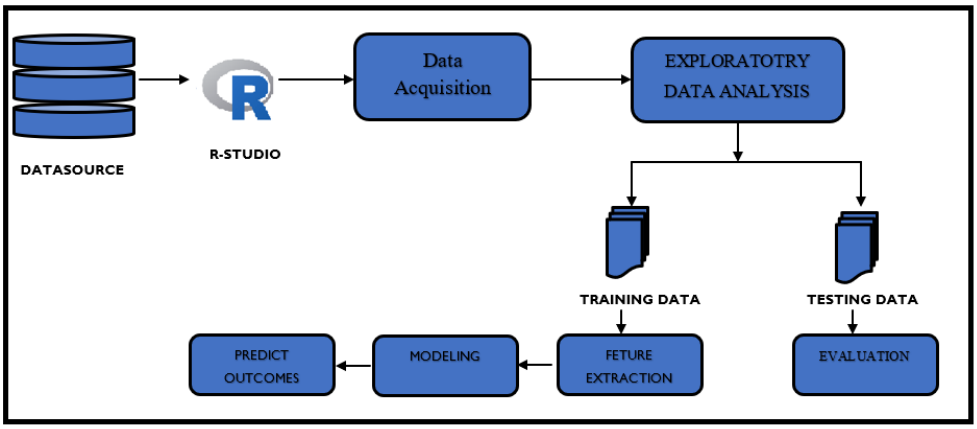

Figure 3: Flow Diagram

4.1 Data Pre-processing

Data Acquisition was done in R studio using different libraries.

1. **Data Acquisition**
   Data was downloaded into PDF format which is difficult to transform into structure data. R studio has been apply to convert into CSV form using "pdftools", "stringr", and "dplyr" libraries. Different function such as "pdf_text" and "str_extract_all" are used. Manually one by one URL are downloaded from GTmetrix and Similar-Web tools. At the end "rbind.data.frame" was used to merge all variables. Once the data has been fetched as shown in above Figure 2, two data (AMP and Non-AMP) sets are integrated as per the time zone Morning and evening using "rbind" function.

2. **Exploratory Data Analysis**

Once the final data set was generated next step is to understand data for evaluating classification model. The EDA was performed to better comprehend information, such as class imbalance, included characteristics, correlation among characteristics, missing values, and outliers. To comprehending the class imbalance, two case study was considered such as What is the distribution of AMP and Non-AAMP websites? In below Figure 4 two-thirds of the data belong to 1 categories where as 1 represent websites which was AMP oriented and one-thirds belong to 2 categories were 2 represent websites which was non-AMP.

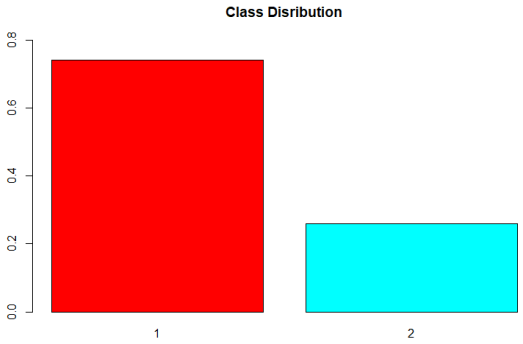

Figure 4: Distribution of websites

This shows a big difference in the annulling class imbalance of the data set in each of the classifications. To handle this class imbalance, over sampling or under sampling would be used in the model to get the better performance. A further step of EDA is correlation which indicates if the correlation exceeds 0.7 suggests a powerful correlation, below 0.3 shows a poor correlation and between 0.3 to 0.7 highlights moderate correlation. In our data sets shows correlation between the variables which are almost moderate.

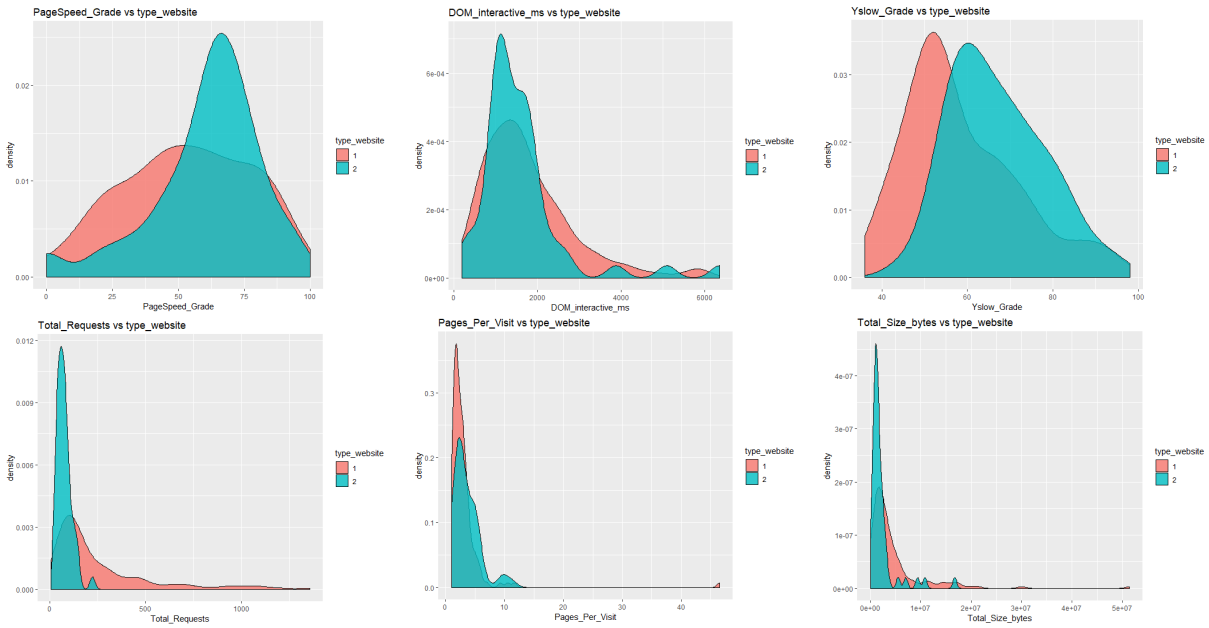

Figure 5: Variable dependencies

In data exploration, its vital to the analyzed outline's. In this research, outliers are neglected because all the data are downloaded manually and out of 28, only important attributes are selected. Additional investigation is done based on the dependencies. This proposition contains 29 columns, one of which had strong links or one of which was not important. Its necessary to check the efficiency of independent variables on target variable as shown in Figure 5.

3. **Feature Selection**
   In R studio, Boruta package was installed, using this package all the attributes are compared with target attributes. Boruta performed 99 iterations in 3.825091 secs using Boruta function. Boruta run "getSelectedAttributes" to checks which attributes are confirmed. U.S. based dataset contains 29 variables out of which 10 variables are strongly confirmed as shown in table. Table 3 shows 11 variables are further used for modeling and validation.

Table 3: Confirmed attributes

Column Name	**Description**
Yslow_Grade	YSlow is a tool that analyses your page and tells you why it's slow.
PageSpeed_Grade	Google PageSpeed Insights is a frustrating handy tool that analyzes your site's front-end performance and offers optimization suggestions
DOM_interactive_ms	Point at which the browser has finished loading and parsing HTML
DOM_content_loaded_ms	Point at which the DOM (Document Object Model) is ready
Total_Requests	Number of requests
Bounce_Rate	The percentage of visitors who navigate away from the site
Avg_Visit_Duration	Its shows the average amount of time visitors spend on a website
Fully_loaded	Time to fully loaded websites
Pages_Per_Visit	Number of visitor visit per pages
Onload_Times	its occurs when the processing of the page is complete

4.2 Classification Model

To recognize which data is suitable, several classification algorithms are introduced in this research. Using R studio library, every classifier model is executed. Data are split into 70 % trained and 30 % tested which was further used to construct models. The building of the various models was carried out using the train function. To evaluate

the comparative study of the websites performance based on the parameters that affect the user experience, its necessary to implement the model with best suits attributes. Sampling base approach and hyperparameter tuning were used to remove class imbalance.

Firstly, RF classification is evaluated in two different way to improve accuracy. One is basic RF with hyperparameter tuning and the other one is using the under-sampling approach in two different time zone. Tuning is hyperparameter optimization and it helps to select the best model. RF used tuneRF function with different parameters and get OOB error 0 % when mtry is 10. Random undersampling was performed for better specificity using ROSE library in R studio. ovun.sample function used to split the data randomly with N 84. The model has been predicted on test data with position 1 because its necessary to improve the value of 1 which represent AMP websites in confusion matrix. Secondly, SVM was also performed in two scenarios one is basic SVM and other used tune function. SVM is built using tune function in R to identify the best model and predict the performance of algorithms. Model use tune function with the epsilon which makes segment and starts with o and goes up to 1 with an increment of 0.1 and cost make use of large value which is able to capture optimal cost value. At the end, 8 different cost value and 11 different epsilon i.e 88 different combinations were used.

Thirdly, basic naive Bayes algorithms are executed to improve the result but it does not fit with this data sets. Next algorithm was DT which was predicted using rpart and plot the tree. Our researchshows the websites speed with this DT algorithm, and check that attributes of websites affect the user experience or not. Lastly, Deep learning with kears has been evaluated in R studio using python installation. The TensorFlow backend engine is used by default on the Keras R interface. To prepare train and test data for model reshape and rescale are used in x data. For y data which is an integer vector which was prepared for training using the Keras to_categorical function. For evaluation of models confusion matrix, recall, precision, specificity, sensitivity, F1-Measure, misclassification error, and gamma value are compared based on two different data set and evaluated the result in the form of the case study.

5 Evaluation and Results

Our main objective is to use different machine learning algorithms to find the factors of the AMP and non-AMP websites and their performance on the user of the website. Based on the ideas obtained from the analysis of the literature the impact of all 4 models, RF, DT, SVM and Naive Bayes were evaluated based on accuracy, precision, recall, F Measure, misclassification rate,sensitivity and specificity. Deep learning with Keras and statistical approach chi-square was also executed to verify the results.

5.1 Case Study 1 : Classification Model

Using the values in the confusion matrix, these performance metrics have been calculated. Figure 6, demonstrates the confusion matrix with a binary classification such as 1/2, true/false, and, AMP/non-AMP. Since numerous measurements are used to estimate accuracy, precision, recall, F Measure, misclassification rate, sensitivity and specificity of classification models are predominant. **Accuracy**: is the total number of correct

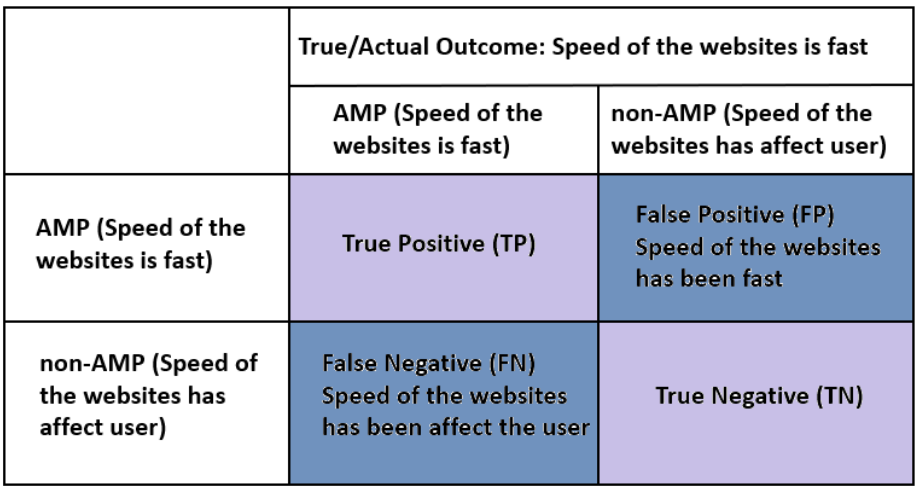

Figure 6: Binary classification of confusion matrix

forecasts for all the predicted cases. However, To measure the performance of real life data sets accuracy is not always reliable. **Precision**: Its ratio of correctly predicted AMP websites to the sum of total websites predicted as fast. **Recall**: It shows the ratio of correctly predicted AMP websites to the total number of speeds of the websites has been affected the user. **F Measures**: It provide the average score between precision and recall score. **Specificity**: Its ratio of correctly predicted total number of non-AMP websites is slow and affected web user to the total number of actual websites that marked as AMP. **Sensitivity** : Its ratio of correctly predicted total number of AMP websites is fast total number of actual websites that marked as non-AMP.

5.1.1 Random forest(RF)

This research was mainly analyzing AMP and non-AMP websites URLs. RF is remarkably good in the scenarios where morning time model precisely classifies 43 occurrences while 4 of the occurrences classifies incorrectly whereas evening time model perfectly recognizes 39 AMP websites out of 47 who loading speed was faster and 8 are classified inaccurately to have speed faster while they didn't have. Firstly, RF performed using hyper-parameter tuning and get the following result as shown in Table 4.

Table 4: Random Forest Outcomes

Random Forest						
Time	**Accuracy**	**Recall**	**Precision**	**F-measures**	**Sensitivity**	**Specificity**
Morning	87.93	93.47	91.48	92.47	91.49	72.73
Evening	75.86	86.66	82.97	84.78	82.98	45.45
Random Forest (Random Under-Sampler)						
Morning	78.7	95.23	78.63	87.79	72.33	87.50
Evening	76.74	90.32	73.68	81.15	73.68	62.10

As per the above Table 4 Evening time get only 45.45 specificity which are less. In our researchactual prediction was based on AMP i.e website speed was faster and doesn't affected web user. To solve this issue Random under sampling technique was performed. Secondly, RF performed using under sampling as shown in Table 4, Evening time period specificity contains 62.50% which is higher than above model.

5.1.2 Support Vector Machine(SVM)

The second model is evaluated in 2 scenarios, one was basic SVM and other was performed hyper-parameter tuning with SVM to get optimal output. Hyperparameter tuning was executed to identify the best kernel because SVM contains polynomial, linear and sigmoid. In both cases, morning time has higher accuracy then evening as shown in Table 5. Gamma value of tuning model is 0.0833 and cost is 4 in the morning time. Accuracy of morning time while executing the basic SVM model is 84.48% and with tuning its 86.20%.

Table 5: Support Vector Machine Outcomes

Support Vector Machine					
Time	**Accuracy**	**Misclassification rate**	**Precision**	**Recall**	**F-Measure**
Morning	84.48	15.51	76.68	83.92	91.26
Evening	70.68	29.31	74.48	87.5	80.45
Support Vector Machine (Hyper-parameter tuning)					
Morning	86.20	13.79	95.81	85.45	92.21
Evening	81.03	18.24	93.61	84.61	88.88

5.1.3 Decision Tree(DT)

This model breaks the tree-shaped information. The tree's root node is considered the best predictor. The model perfectly recognizes 36 AMP websites while 11 are classified incorrectly. Accuracy of Morning has been higher compared to Evening as shown in Table 6. All the value of the performance matrix is higher at Morning. Model get almost 64% of Specificity which is quite good.

Table 6: Decision Tree Outcomes

Decision Tree						
Time	**Accuracy**	**Precision**	**Recall**	**F-Measure**	**Sensitivity**	**Specificity**
Morning	74.14	89.58	91.48	95.52	76.60	63.64
Evening	72.41	92.70	89.89	90.52	78.72	45.45

5.1.4 Naive Bayes

A Naive Bayes classification model used Bayes theorem for classification task which is a probabilistic machine learning model. As per the Table 7, this model doesn't fit for data. All the value was less compared to the other classification model. Accuracy of the model is 67.34% in the morning and 65.78% at Evening. Specificity and Recall percentage of both the time is good to compare to other performance matrices.

Table 7: Naive Bayes Outcomes

Naive Bayes						
Time	**Accuracy**	**Precision**	**Recall**	**F-Measure**	**Sensitivity**	**Specificity**
Morning	67.34	56.15	95.16	68.20	59.38	84.35
Evening	65.78	53.25	85.71	67.92	56.25	82.35

5.2 Case Study 3: Deep learning with Keras

Keras core data structure is a model for layers organization. Sequential model is a simplest model type. First layer of the input data which specifies the shape represent as input_shape and a length 9 numeric vector. The last layer output a numeric vector length 3. Model performance are evaluated on the test data and get 0.62742 loss and 0.80 accuracy. As per the above Morning data set, same model was evaluated in evniing data set. Model performance are evaluated on the test data and get 0.60742 loss and 0.75 accuracy.

5.3 Data Discussion

As discussed above section 5, separate models were performed to verify the accuracy of the consequence. In this study website efficiency was identified based on multiple parameters. The most impacted parameters are found on websites using machine learning algorithms such as RF and DT. (Ren (2018)) conducted experimental study and get a result in term of value that SVM perform well compare to all the model. As seen from Figure 7 in terms of accuracy, SVM with a value of 86.20 % in the morning and 81.03 % in the evening are outperforms than other models. It is followed by RF with a value of 78.70 %. Naive Bayes are the poorest performer based on the accuracy with the value 67.34 % in Morning and 65.78 % in Evening. DT and RF have mild results while DL (Deep Learning) with Keras perform well in comparison with classification models such as DT, RF, and SVM. Therefore, this study concludes that SVM and DL have superior predictive efficiency over other models. As per the model, more individuals scrolling the web page are predicting successful outcomes in the morning period. Web page speed is therefore important on mobile devices for web users. Every second-page delay can possibly cost internet businesses millions of lost revenues each year and velocity enhancement tends to play an significant part in company life (Nagy (2013)). AMP solves the issue of velocity affecting the web user and making it easy to enhance the online business in the future.

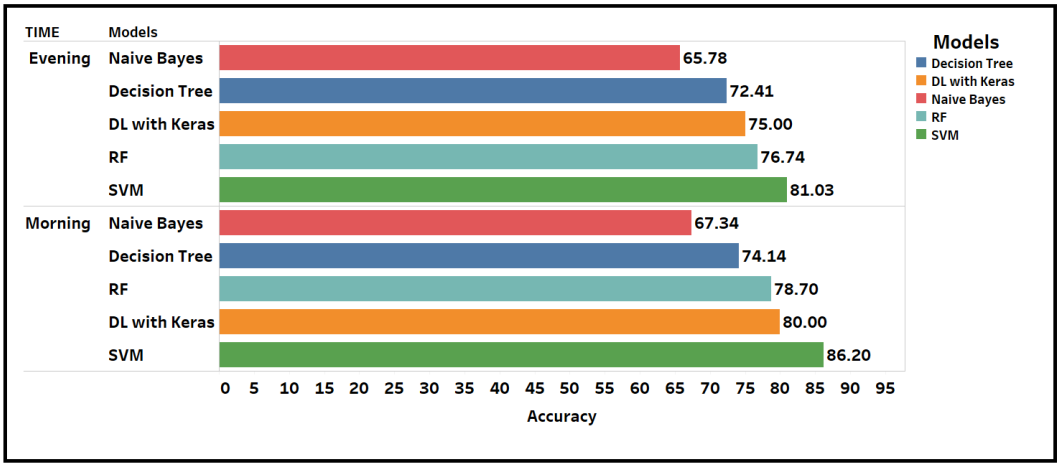

Figure 7: Performance Comparison of Models Graphically

6 Conclusion and Future Work

This research evaluated the performance impact of AMP and non-AMP websites based on the numerous parameters such as Yslow_Grade, PageSpeed_Grade, Avg_Visit_Duration, Total_Requests, Onload_Times, and many more. The novelty of this research is to test various ML classifier such as SVM, DT, RF, NaÃŕve Bayes, and DL with to Keras to predict better accuracy of AMP and non-AMP websites. To identify the parameters that affect user experience while enhancing the performance of AMP sites feature selection algorithm was performed and get 11 confirmed attributes. DT and RF also performed to verify the feature importance. SVM and DL with Keras have a higher value which means perfectly fit into the dataset than another model. Further, performance matric is evaluated using confusion matrix by considering under a sampling-based approach and hyperparameter tuning.

The specificity of models using these two techniques has higher and shows AMP orient URL are faster. As the model is performed on two different time zone to identify user experience on web performance, it verifies that Morning time has been higher accuracy than Evening. Several web user scrolls the web page daily, so the speed is a major concern while loading the page. Amp implementation is useful while enhancing the performance of websites. Positive findings will fix the page speed issue that indirectly improves revenue and improves conversion rate. This outcome increases the future trend of web analytics indirectly.

The work can be further enhanced and expanded by taking a larger number of websites URL. This would enable to get more accurate results using different ML algorithms. Non-AMP oriented website also contains the same attributes like bounce rate, average visit duration, and many more are a major parameter which needs to improve at Evening if you want to increase the revenue. It's also interesting to identify the outcomes using regression models if only AMP URLs were considered. Positive results help to increase user experience.

Acknowledgement

I would like to thank my supervisor, Prof. Manuel Tova-Izquierdo, for his valuable time in guiding and promoting me and exchanging understanding with me throughout the studies. His guidelines are the primary motive for the efficient completion and execution of this study. I would also like to thank my family for their assistance and love my whole life. Last but not least, all my friends who have always motivated me to work hard and inspire me.

References

Agus Wibowo1, G. A. and Mufadhol, M. (2018). Accelerated mobile pages from javascript as accelerator tool for web serviceon e-commerce in the e-business, *International Journal of Electrical and Computer Engineering (IJECE)* **8**(4): 2399 –2405.
URL: *http://doi.org/10.11591/ijece.v8i4.pp2399-2405*

Barakovic, S. and Skorin-Kapov, L. (2015). Multidimensional modelling of quality of experience for mobile web browsing, *Computers in Human Behavior* **50**: 314 – 332.
URL: *http://www.sciencedirect.com/science/article/pii/S0747563215002721*

Barakovic, S. and Skorin-Kapov, L. (2017). Modelling the relationship between design/performance factors and perceptual features contributing to quality of experience for mobile web browsing, *Computers in Human Behavior* **74**: 311 – 329.
URL: *http://www.sciencedirect.com/science/article/pii/S0747563217302881*

Bartuskova, A., Krejcar, O., Sabbah, T. and Selamat, A. (2016). Website speed testing analysis using speedtesting model, *Jurnal Teknologi* **78**(12-3): 121âĂŞ134.
URL: *https://jurnalteknologi.utm.my/index.php/jurnalteknologi/article/view/10028/6035*

Bocchi, E., De Cicco, L. and Rossi, D. (2016). Measuring the quality of experience of web users, *ACM SIGCOMM Computer Communication Review* **46**(4): 8–13.
URL: *http://doi.acm.org/10.1145/3027947.3027949*

Butkiewicz, M., Wang, D., Wu, Z., Madhyastha, H. V. and Sekar, V. (2015). Klotski: Reprioritizing web content to improve user experience on mobile devices, *In 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)* pp. 439–453.

Byungjin Jun, FabiÃąn E. Bustamante, S. Y. W. and Bischof, Z. S. (2019). Amp up your mobile web experience: Characterizing the impact of googleâĂŹs accelerated mobile project, *The 25th Annual International Conference on Mobile Computing and Networking (MobiCom âĂŹ19)* pp. 1–14.

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B. and Herrera, F. (2018). *Introduction to KDD and Data Science*, Springer International Publishing, Cham, pp. 1–17.
URL: *https://doi.org/10.1007/978-3-319-98074-4_1*

H Jati, N. and Wardani, R. (2018). Quality analysis of university websites from usability side with multicriteria decision analysis method, *Journal of Physics: Conference Series*

1140(1): 012038.
URL: *https://iopscience.iop.org/article/10.1088/1742-6596/1140/1/012038/meta*

Johanna Dunaway, Kathleen Searles, M. S. and Paul, N. (2018). News attention in a mobile era, *Journal of Computer-Mediated Communication* **23**(2): 107—124.

Kaur, S., Kaur, K. and Kaur, P. (2016). An empirical performance evaluation of universities website, *International Journal of Computer Applications* **146**(15): 10–16.
URL: *https://www.iaescore.com/journals/index.php/IJECE/article/view/9186*

Miklosik, Andrej & ÄŇervenka, P. . H. I. (2017). Impact of accelerated mobile pages format on corporate web sites, *MARKETING IDENTITY* pp. 204–214.

Nagy, Z. (2013). Improved speed on intelligent web sites, *Recent Advances in Computer Science* **1**(14): 215–220.

Nejati, J. and Balasubramanian, A. (2016). An in-depth study of mobile browser performance, *Proceedings of the 25th International Conference on World Wide Web* pp. 1305–1315.
URL: *https://doi.org/10.1145/2872427.2883014*

Netravali, R. and Mickens, J. (2018). Remote-control caching: Proxy-based url rewriting to decrease mobile browsing bandwidth, *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications* pp. 63–68.
URL: *http://doi.acm.org/10.1145/3177102.3177118*

O'donoghue, R. (2017). *AMP: Building Accelerated Mobile Pages: Create lightning-fast mobile pages by leveraging AMP technology*, Packt Publishing Ltd.

Pan, B. and Wang, D. (2016). Mobile internet access patterns for travel: Comparison of desktops, tablets, and phones, *TRAVEL AND TOURISM RESEARCH ASSOCIATION: ADVANCING TOURISM RESEARCH GLOBALLY* .

Phokeer, A., Chavula, J., Johnson, D., Densmore, M., Tyson, G., Sathiaseelan, A. and Feamster, N. (2019). On the potential of google amp to promote local content in developing regions, *In Proceedings IEEE COMSNETS* .

Ren, J. (2018). A qoe-based governor for web browsing on heterogeneous mobile systems, *2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* pp. 650–655.

Ruamviboonsuk, V., Netravali, R., Uluyol, M. and Madhyastha, H. V. (2017). Vroom: Accelerating the mobile web with server-aided dependency resolution, *Proceedings of the Conference of the ACM Special Interest Group on Data Communication* pp. 390–403.
URL: *http://doi.acm.org/10.1145/3098822.3098851*

Singh, M. K. and Verma, M. (2014). Performance evaluation of websites emulating referenced resources, *International Journal of Computer Applications* pp. 0975–8887.

Stringam, B. and Gerdes, J. (2019). Service gap in hotel website load performance, *International Hospitality Review* .
URL: *https://doi.org/10.1108/IHR-09-2018-0012*

Weiwang Li, Zhiwei Zhao, G. M. H. D. Q. N. and Zhao, Z. (2017). Reordering webpage objects for optimizing quality-of-experience, *IEEE Access, Access, IEEE* **5**: 6626—6635.
URL: *https://ieeexplore.ieee.org/abstract/document/7888987/authors#authors*

Xiao Sophia Wang, A. K. and Wetherall, D. (2016). Speeding up web page loads with shandian, *In 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)* pp. 109–122.

Yu, N. and Kong, J. (2016). User experience with web browsing on small screens: Experimental investigations of mobile-page interface design and homepage design for news websites, *Information Sciences* **330**: 427 – 443. SI Visual Info Communication.
URL: *http://www.sciencedirect.com/science/article/pii/S0020025515004363*