EVALUATE THE PERFORMANCE OF AMP AND NON-AMP WEBSITES USING MACHINE LEARNING ALGORITHMS

Bhumi patel

X18114865

Supervisor: Manuel Tova-Izquierdo

INTRODUCTION

- Increased the popularity for accessing the web pages through mobile devices.
- Effects of page loading delay on the online web business.
- Google comes with solution called as AMP.
- What is AMP?
  - AMP stands for Accelerated Mobile Pages which are built specifically for high speed access to the content.
- Importance of AMP in web analytics.

LITERATURE REVIEW

Author	Title	Research Work	Motivation
Johanna Dunaway and Paul (2018)	News attention in a mobile era	Examined news websites on the various types of devices, including desktops, tablets, and phones.	Study shows that speed does not only affects the news websites but also affects various domains like e-commerce, hotel, travel and many more categories on the mobile device.
Ren(2018)	A QoE-based governor for web browsing on heterogeneous mobile systems	Paper introduces a novel QoE evaluation model for user experience that takes account of delays for web browsing using different ML algorithms.	This approach does not increase the load of high technology websites on mobile devices. This paper directs me to look for ML algorithms that work well while loading AMP pages.

LITERATURE REVIEW

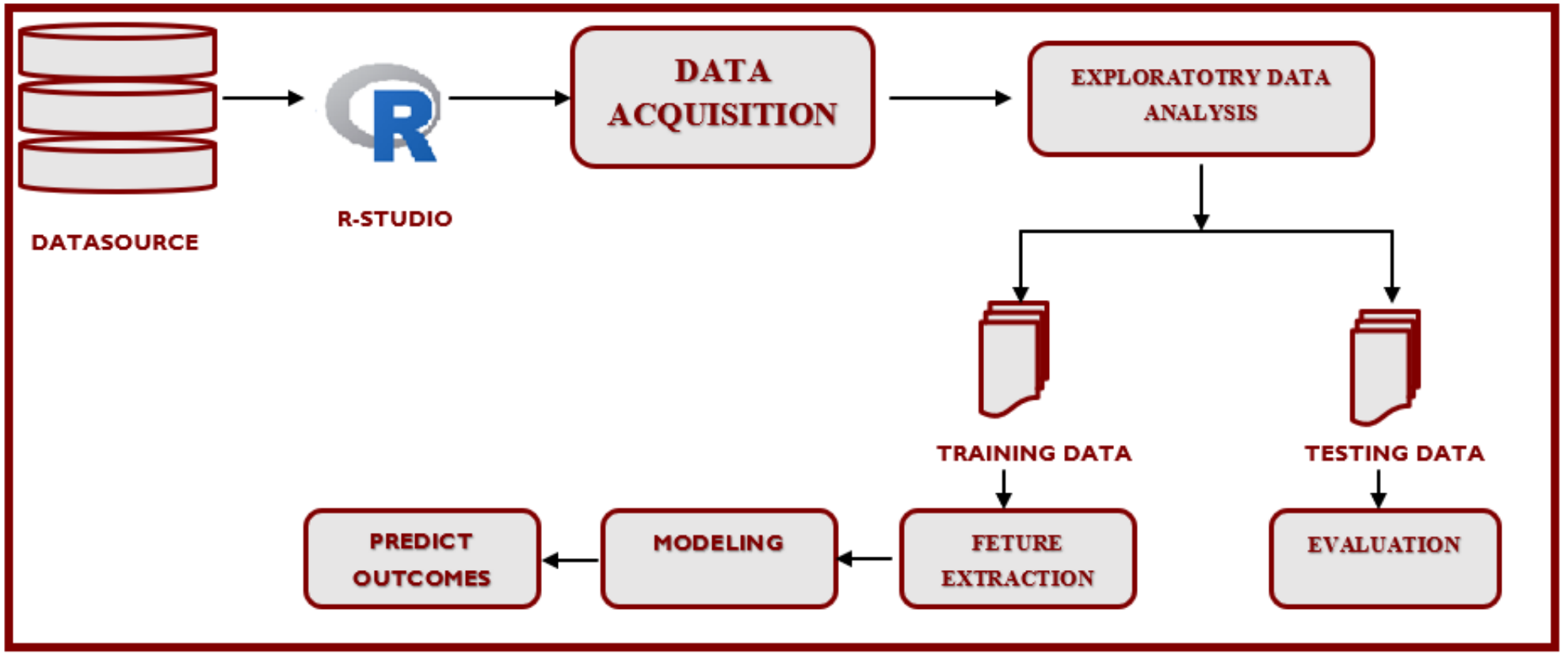
Author	Title	Research Work	Motivation
Byungjin Jun and Bischof (2019)	AMP up your mobile web experience: Characterizing the impact of Google's accelerated mobile project	The test is conducted based on three common attributes of QoE metrics i.e. Page Load Time, Time to First Byte and Speed Index to study the impact mobile web quality experiences using AMP sites.	This result encourages to find the accuracy of performance with additional parameters that affect user experience.
H Jati and Wardani (2018)	Quality analysis of university websites from usability side with multicriteria decision analysis method	GTmetrix and SimilarWeb were the most trusted services ensuring the validity of the collected data.	To gathered the performance factors of non-AMP and AMP-based websites, Gtmetrix and SimilarWeb tools were used in tis research.

RESEARCH OBJECTIVE

Enhancing the performance of the websites based on the features that influence user experience

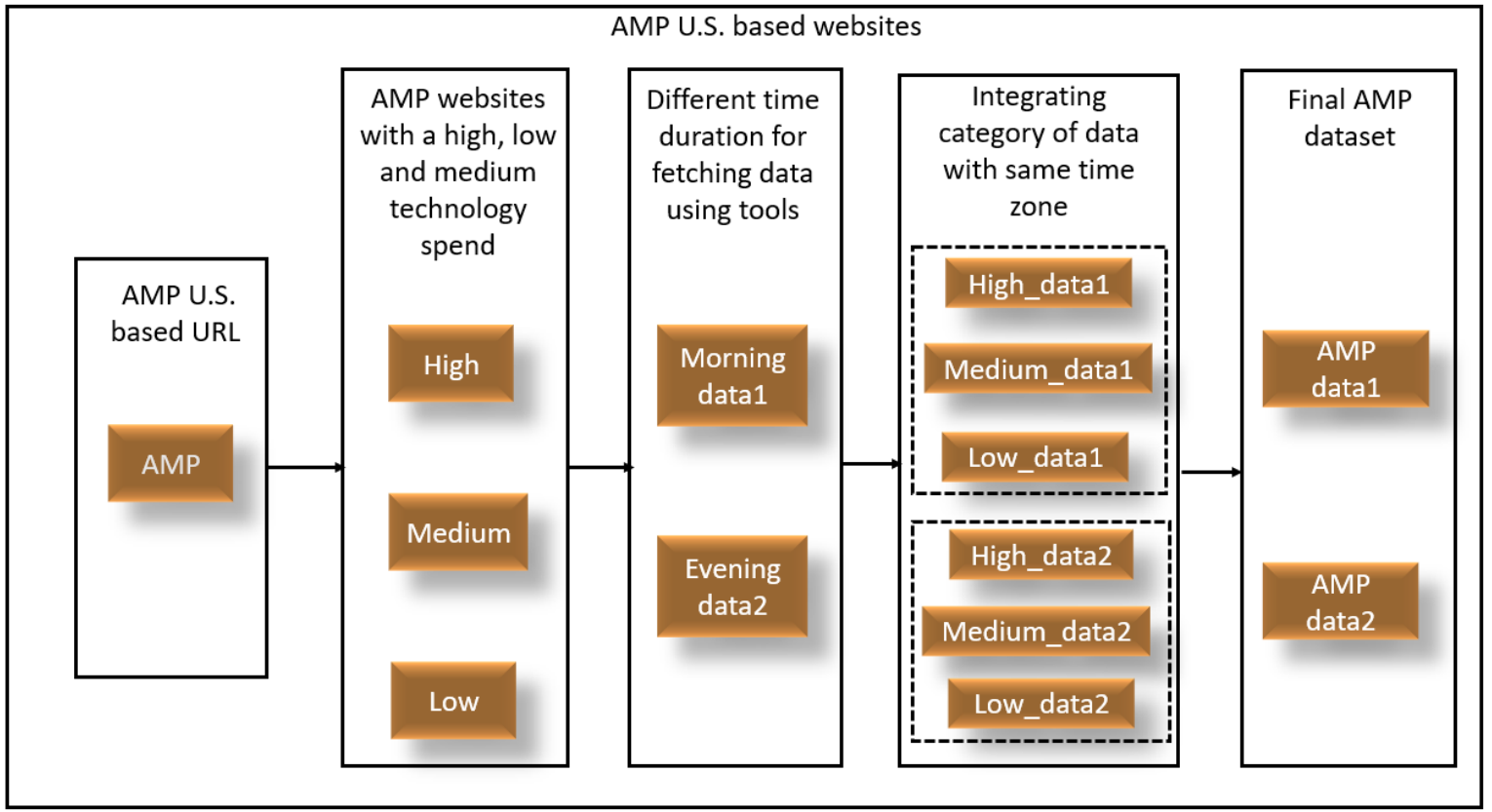
Find the result of non-AMP v/s AMP web pages using ML.

PROCESS FLOW

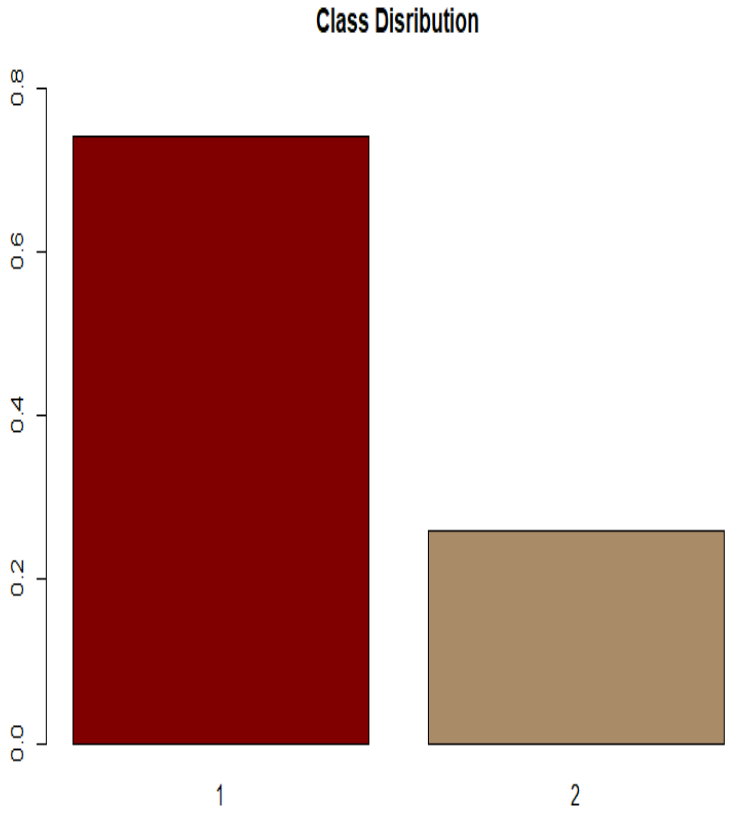


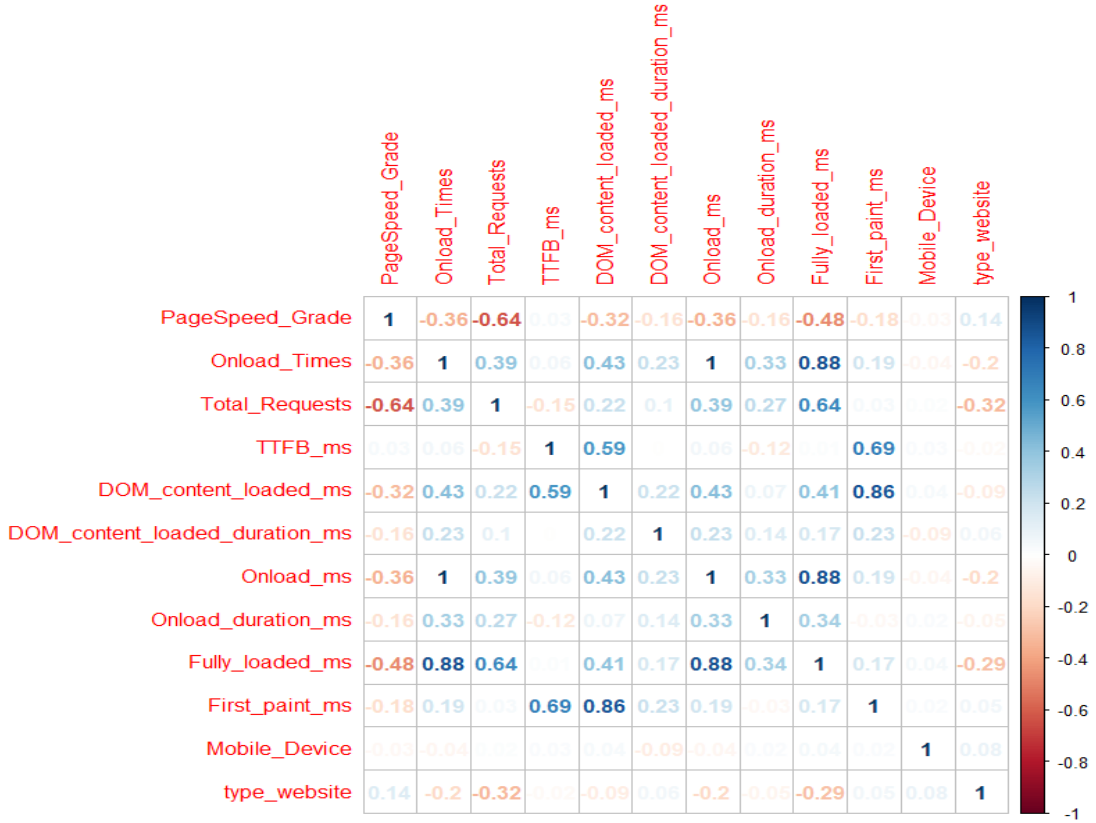
DATA PRE-PROCESSING

- Target Data
  - Tools for data acquisition
  - Data Collection
- Data Cleaning
- Data Integration



EXPLORATORY DATA ANALYSIS





MACHINE LEARNING MODELS

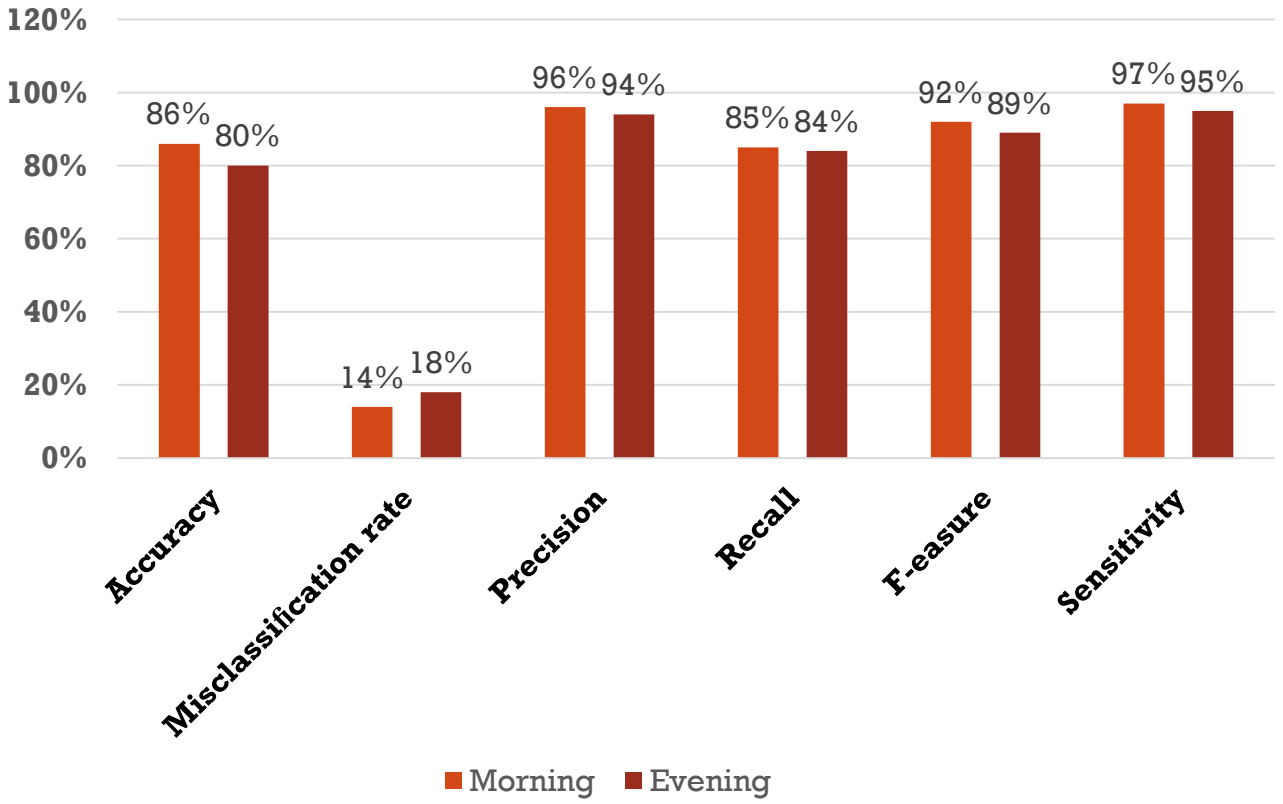
- Naive Bayes
- Decision Tree (DT)
- Random Forest (RF)
- Support Vector Machine (SVM)
- Deep Learning with Keras

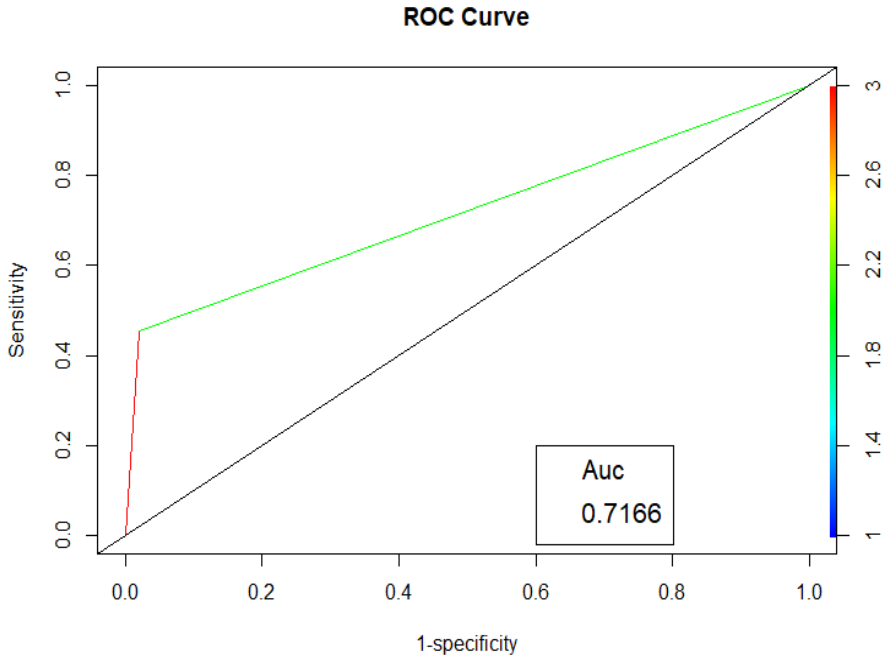
RESULTS

Algorithms	Accuracy	
	Morning	Evening
Random Forest (RF)	88%	76%
RF (Random Under-Sampler)	79%	77%
Support Vector Machine (SVM)	84%	71%
SVM (Hyper-parameter tuning)	**87%**	**81%**
Decision Tree (DT)	74%	72%
Naive Bayes	67%	66%
Deep learning with Keras	80%	75%

EVALUATION

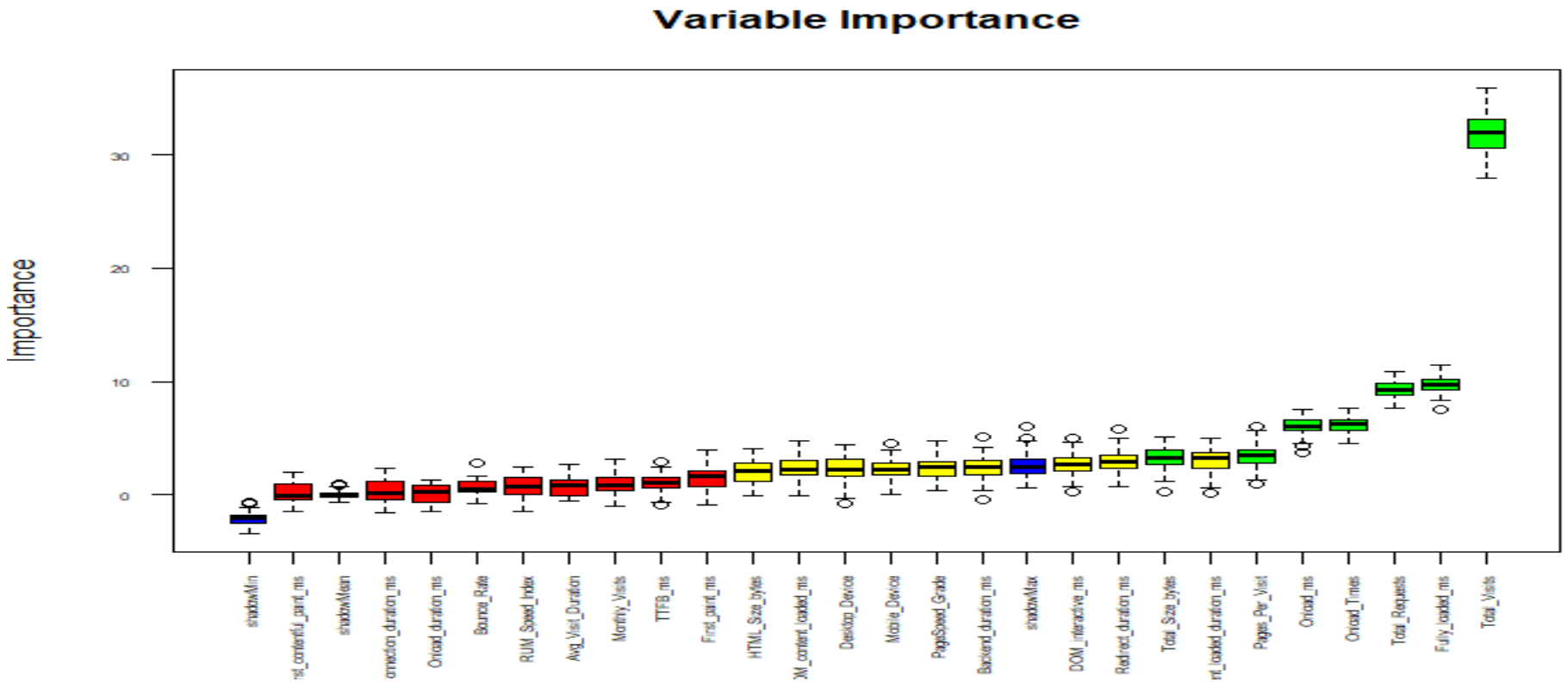
Results FOR SVM (Hyper-parameter tuning)





EVALUATION

Feature selection using Boruta Packages



LIMITATION AND FUTURE WORK

- Large dataset should be considered for enhancing model performance.
- Limited for only mobile web pages.

APPLICATION

- Enhancing speed of websites
- Increase in number of web users
- Improving future trend of web analytics

REFERENCES

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

THANK YOU!