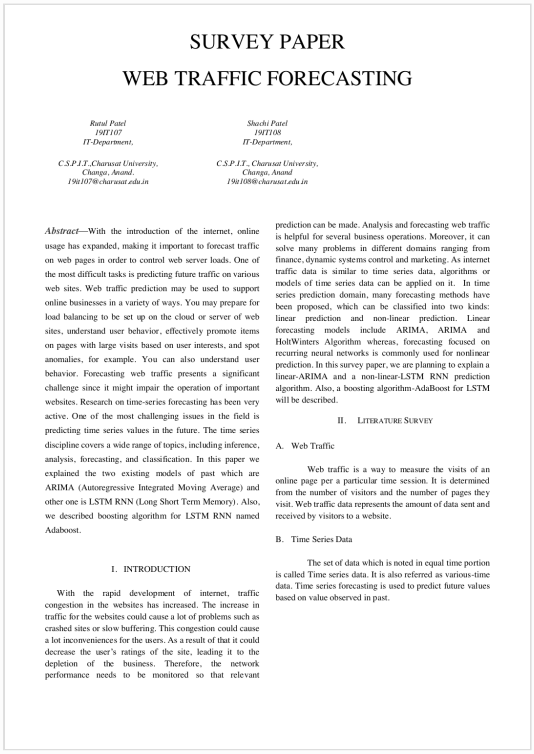
Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	19it107 PATEL RUTUL MANUBHAI
Assignment title:	7-IT450-SGP-2022-Technical-Paper
Submission title:	19it107_19it108_SurveyPaper_WebTrafficForcasting
File name:	19IT107_19IT108_WEB_TRAFFIC_FORECASTING.pdf
File size:	508.95K
Page count:	4
Word count:	2,019
Character count:	10,839
Submission date:	05-Nov-2022 01:37PM (UTC+0530)
Submission ID:	1945223229



SURVEY PAPER

WEB TRAFFIC FORECASTING

Rutul Patel
19IT107
IT-Department,
C.S.P.I.T.,Charusat University, Changa, Anand.
19it107@charusat.edu.in

Shachi Patel
19IT108
IT-Department,
C.S.P.I.T., Charusat University, Changa, Anand
19it108@charusat.edu.in

***Abstract*—With the introduction of the internet, online usage has expanded, making it important to forecast traffic on web pages in order to control web server loads. One of the most difficult tasks is predicting future traffic on various web sites. Web traffic prediction may be used to support online businesses in a variety of ways. You may prepare for load balancing to be set up on the cloud or server of web sites, understand user behavior, effectively promote items on pages with large visits based on user interests, and spot anomalies, for example. You can also understand user behavior. Forecasting web traffic presents a significant challenge since it might impair the operation of important websites. Research on time-series forecasting has been very active. One of the most challenging issues in the field is predicting time series values in the future. The time series discipline covers a wide range of topics, including inference, analysis, forecasting, and classification. In this paper we explained the two existing models of past which are ARIMA (Autoregressive Integrated Moving Average) and other one is LSTM RNN (Long Short Term Memory). Also, we described boosting algorithm for LSTM RNN named Adaboost.**

1. INTRODUCTION

With the rapid development of internet, traffic congestion in the websites has increased. The increase in traffic for the websites could cause a lot of problems such as crashed sites or slow buffering. This congestion could cause a lot inconveniences for the users. As a result of that it could decrease the user's ratings of the site, leading it to the depletion of the business. Therefore, the network performance needs to be monitored so that relevant prediction can be made. Analysis and forecasting web traffic is helpful for several business operations. Moreover, it can solve many problems in different domains ranging from finance, dynamic systems control and marketing. As internet traffic data is similar to time series data, algorithms or models of time series data can be applied on it. In time series prediction domain, many forecasting methods have been proposed, which can be classified into two kinds: linear prediction and non-linear prediction. Linear forecasting models include ARIMA, ARIMA and HoltWinters Algorithm whereas, forecasting focused on recurring neural networks is commonly used for nonlinear prediction. In this survey paper, we are planning to explain a linear-ARIMA and a non-linear-LSTM RNN prediction algorithm. Also, a boosting algorithm-AdaBoost for LSTM will be described.

II. LITERATURE SURVEY

A. Web Traffic

Web traffic is a way to measure the visits of an online page per a particular time session. It is determined from the number of visitors and the number of pages they visit. Web traffic data represents the amount of data sent and received by visitors to a website.

B. Time Series Data

The set of data which is noted in equal time portion is called Time series data. It is also referred as various-time data. Time series forecasting is used to predict future values based on value observed in past.