

# Web Traffic Time Series Forecasting

Xin Kang

Institute of Statistics and Big Data

April 8, 2019

# Contents

This project focuses on solving the problem of predicting the future web traffic for approximately 145,000 Wikipedia articles. Detailed data description is covered in the following section. Making future prediction on sequential or temporal observations has emerged in many key real-world problems. By forecasting the future values of multiple web traffic time series, we can answer some questions like how many servers you need in reality and what your total cost for next month is when you need to use external servers. If the performance is satisfactory, similar methods can be applied to other websites to predict their web traffic, and it can help people make smart advertisement decisions and make profit.

# Exploratory Data Analysis

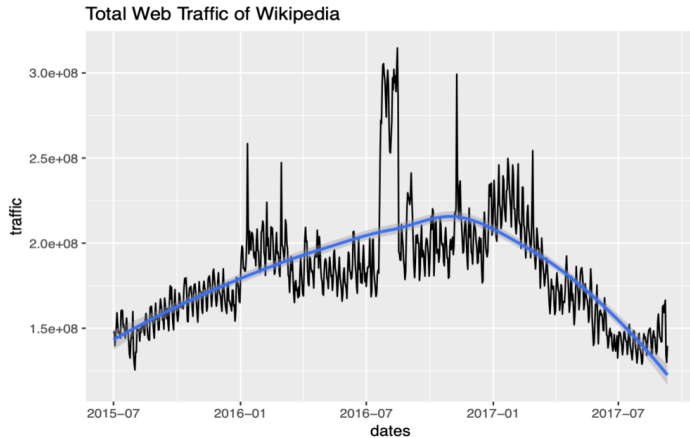


Figure: 1

# Exploratory Data Analysis

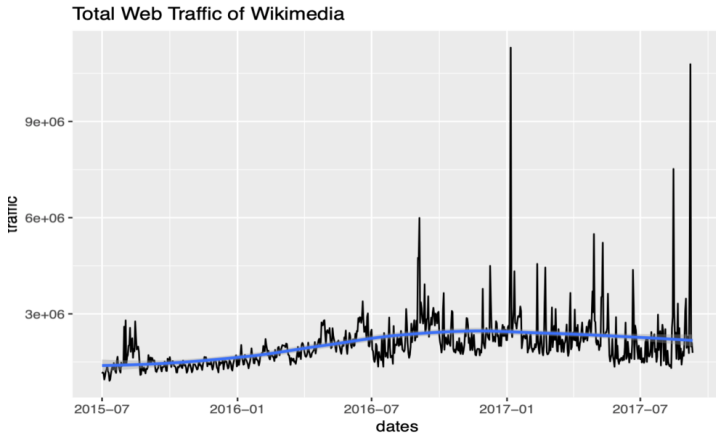


Figure: 2

# Exploratory Data Analysis

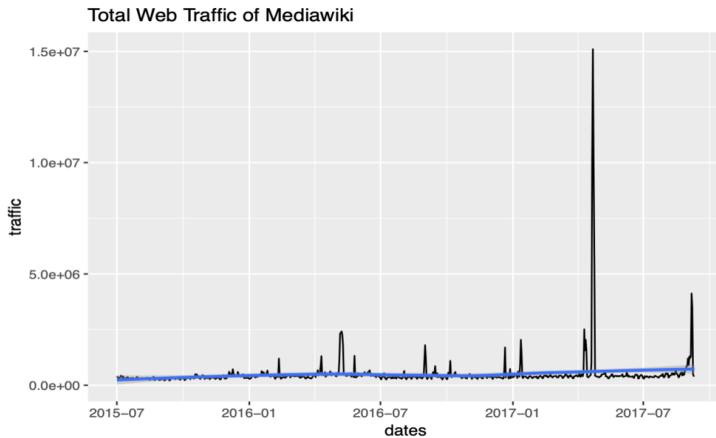


Figure: 3

# Exploratory Data Analysis

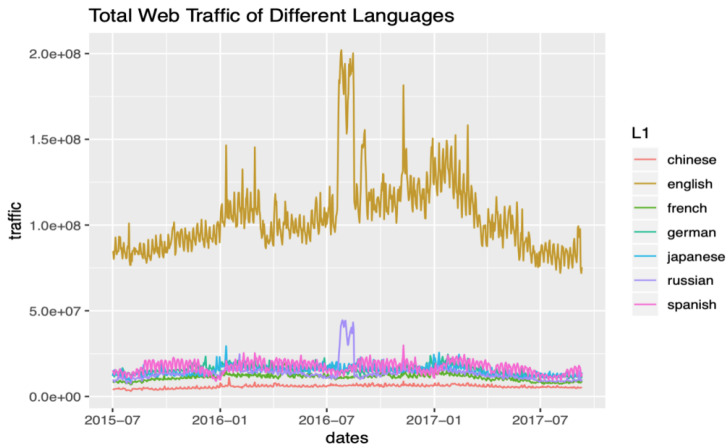


Figure: 4

# Exploratory Data Analysis

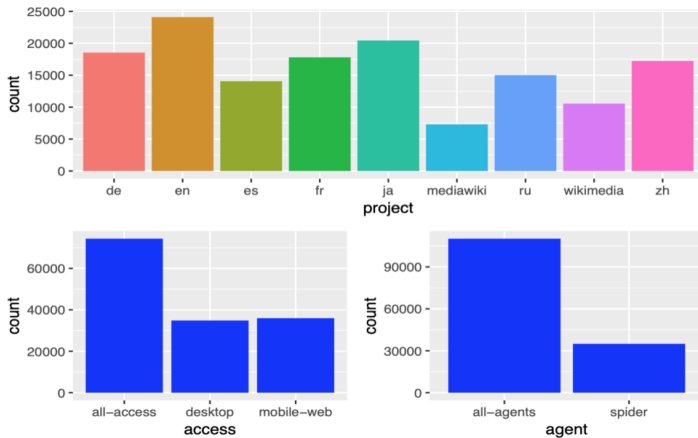


Figure: 5



# Exploratory Data Analysis

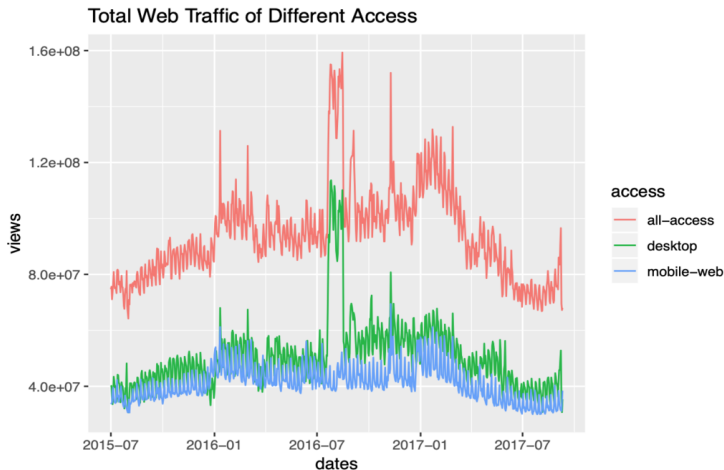


Figure: 6

# Exploratory Data Analysis

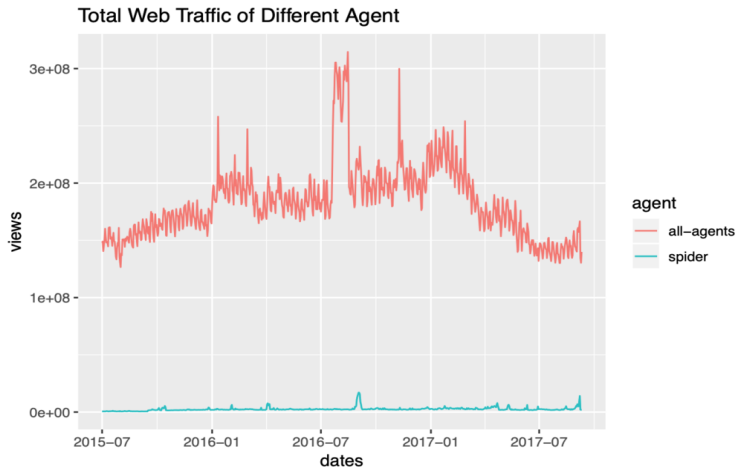


Figure: 7

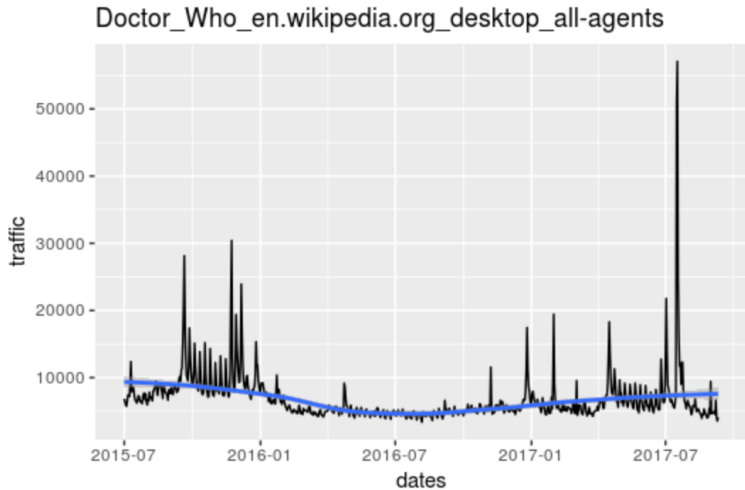


Figure: 8

# Exploratory Data Analysis

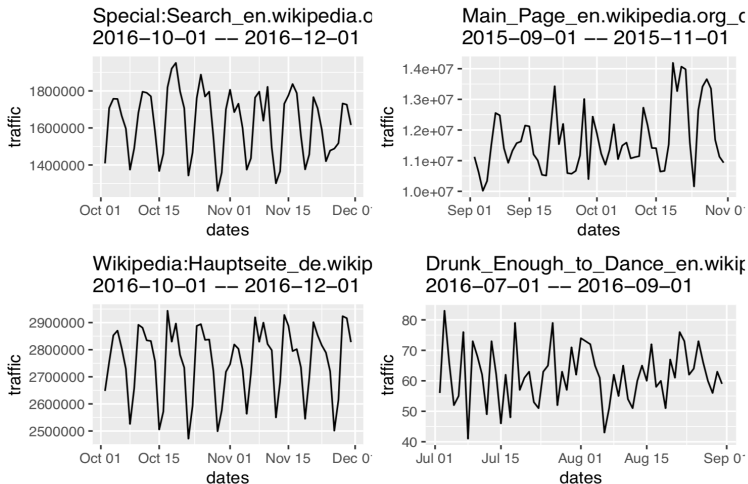


Figure: 9

# Model Fitting – ARIMA

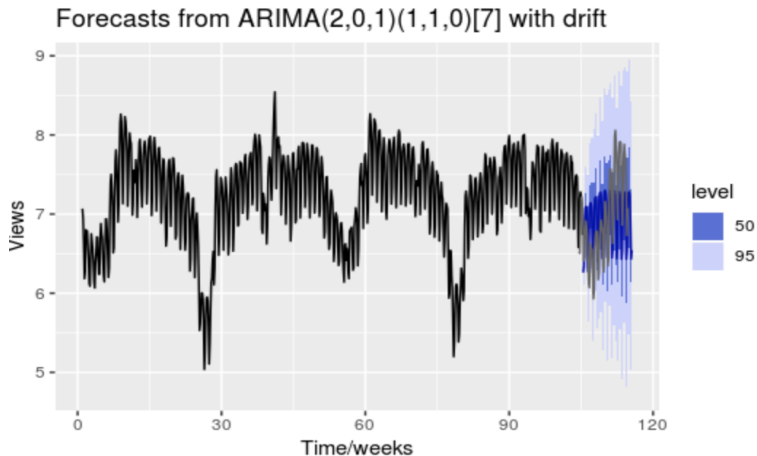


Figure: 10

# Model Fitting – ARIMA

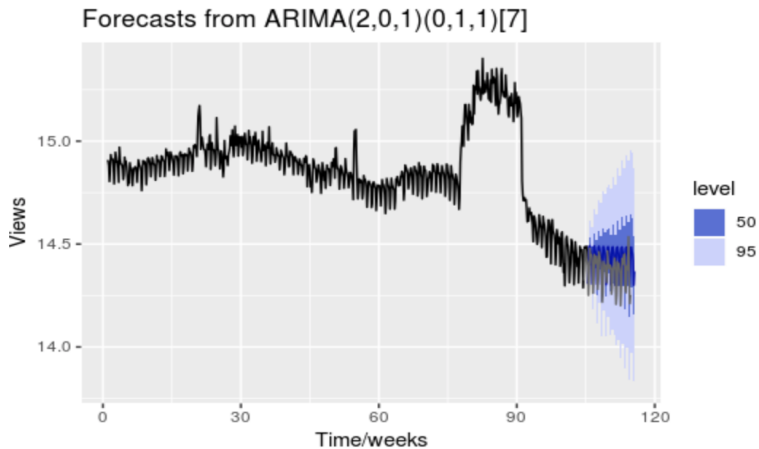


Figure: 11

# Model Fitting – ARIMA

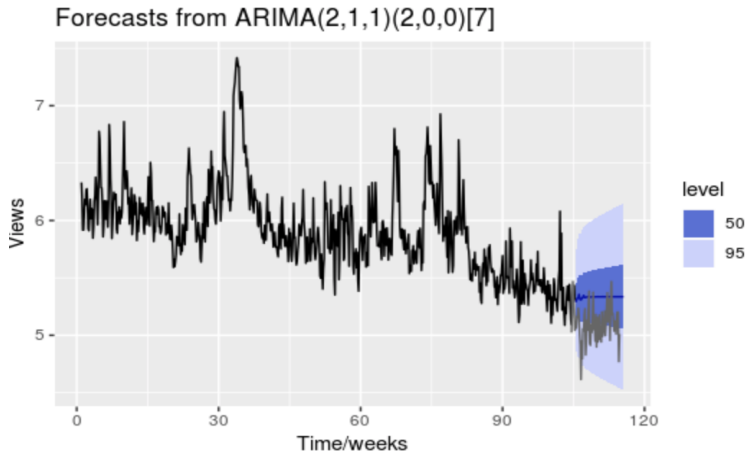


Figure: 12

# Model Fitting – ARIMA

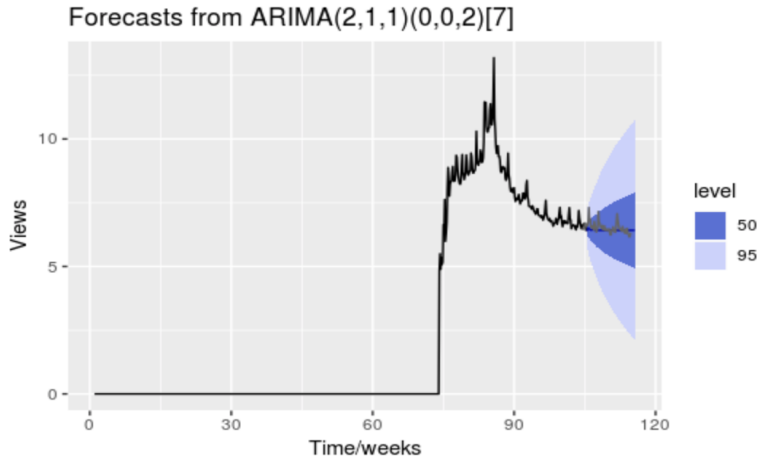


Figure: 13



# Model Fitting – LSTM

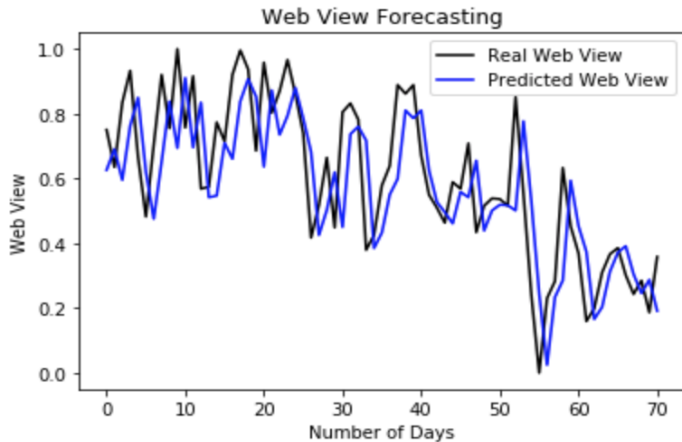


Figure: 14

# Model Fitting – LSTM

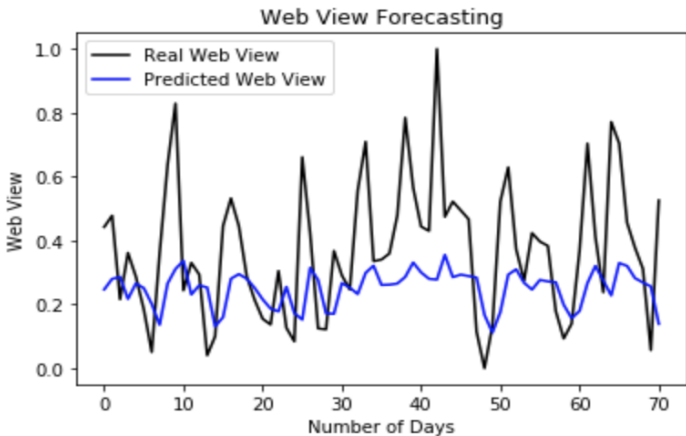


Figure: 15

# Model Fitting – LSTM

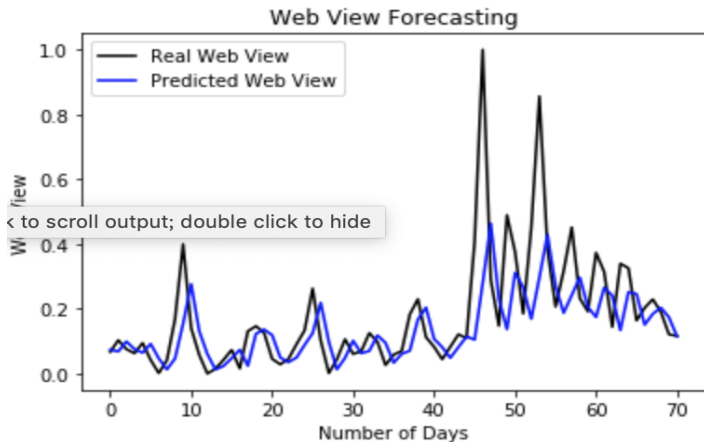


Figure: 16

# Model Fitting – LSTM

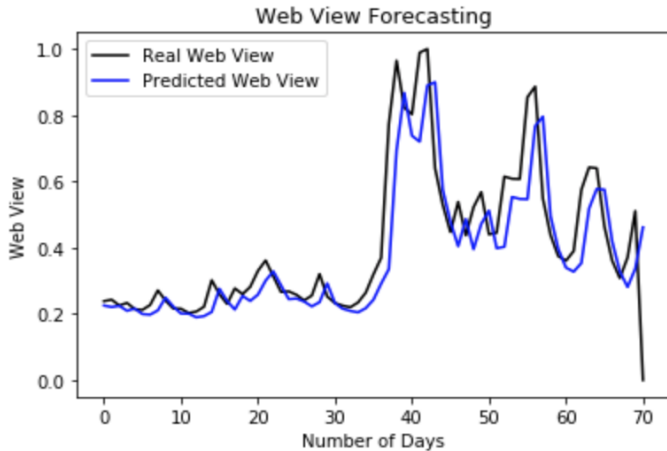


Figure: 17

**Thank you!**