



# US GDP Prediction

Capstone Presentation: Randy Neo

# Table of Contents

01

Problem  
Statement

02

Data

04

Conclusion

03

Modelling

A low-angle, black and white photograph of several tall skyscrapers reaching towards a dark sky. The perspective creates a sense of height and scale. A white rectangular frame is centered on the image, containing the text.

01

**Problem  
Statement**




# Problem Statement

# THE 90/90/90 RULE

# Problem Statement

**90% of new traders lose  
90% of their capital in  
90 days.**





How can we enable investors to  
make better investment  
decisions by having a marco  
view of the economy through  
the prediction of GDP?



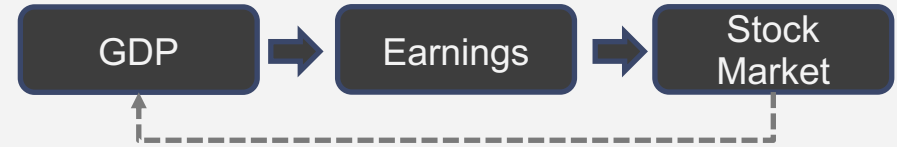
# Mission Statement

## What is GDP?

Gross domestic product (GDP) is the value of final goods and services produced in a country in a certain time frame.

It is the sum of consumer spending, domestic investments, government spending, and the difference in export and imports in a country.

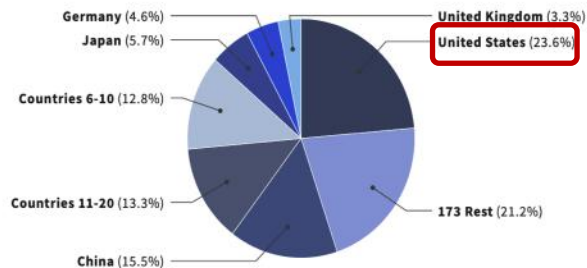
As a broad measure of overall domestic production, it functions as a comprehensive **scorecard** of a given **country's economic health**.



# Why US GDP?

## % Share of the Global Economy

The 173 countries outside the top 20 make up less than a fourth of the total global economy.



Source: IMF

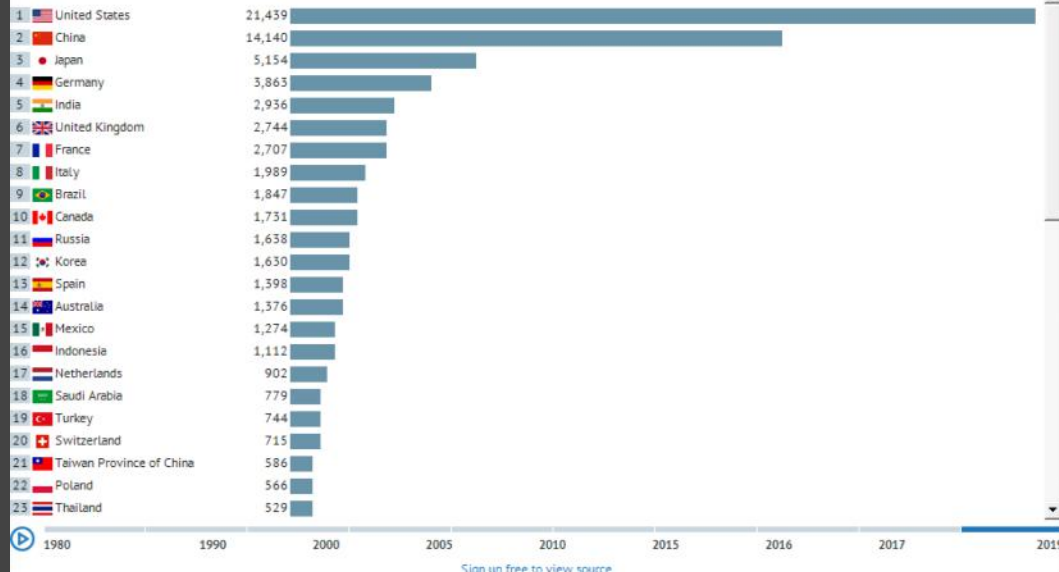
Chart: Investopedia • [Get the data](#) • [Add this chart to your site](#)

Investopedia

The US has historically driven global GDP and it is still the most important economy in the world.

## GDP in Current Prices

Billion USD



Source: <https://knoema.com/nwnfkne/world-gdp-ranking-2019-gdp-by-country-data-and-charts>

Source: <https://www.investopedia.com/insights/worlds-top-economies/>



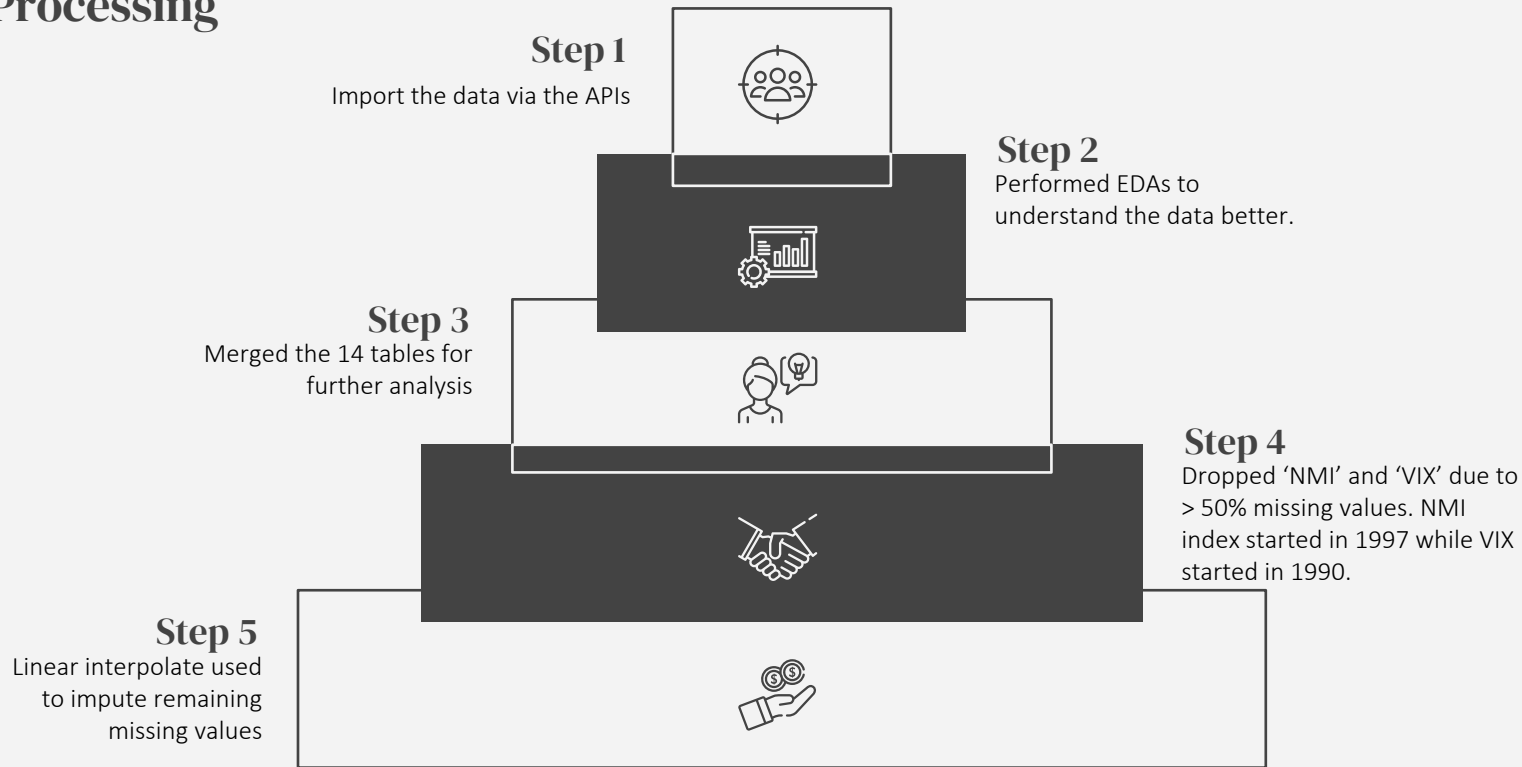


Data

02

Data	Frequency	Source	Description
ISM Manufacturing: PMI Composite Index	Monthly	Quandl	ISM measures manufacturing activity based on a monthly survey of purchasing managers at more than 300 manufacturing firms. Significant as it represents 20% of business contribution to GDP in U.S.
NMI Non-Manufacturing (Services) Index	Monthly	Quandl	An economic index based on surveys of more than 400 non-manufacturing (or services) firms' purchasing and supply executives, within 60 sectors across the nation. Very significant as service sector represents 80% of the U.S. economy.
Consumer Sentiment	Monthly	Quandl	The Michigan Consumer Sentiment Index (MCSI) is a monthly survey of U.S. consumer confidence levels conducted by the University of Michigan. It is based on telephone surveys that gather information on consumer expectations regarding the overall economy.
Building Permits	Monthly	Quandl	Measures the change in the number of new building permits issued by the government. A Key indicator of demand in the housing market.
Unemployment Level	Monthly	Quandl	This statistic displays the seasonally-adjusted unemployment rate in the United States
U.S. Nonfarm Payrolls	Monthly	Quandl	measures the change in the number of people employed during the previous month, excluding the farming industry.
Initial Claims	Weekly	Quandl	measures the number of individuals who filed for unemployment insurance for the first time during the past week.
Continuing_claims	Weekly	Quandl	measures the number of unemployed individuals who qualify for benefits under unemployment insurance.
AAll Investor Sentiment Data	Weekly	Quandl	The AAll Sentiment Survey is a weekly survey of its members which asks if they are "Bullish," "Bearish," or "Neutral" on the stock market over the next six months.
S&P500	Daily	Yahoo Finance	The S&P 500 or Standard & Poor's 500 Index is a market-capitalization-weighted index of the 500 largest U.S. publicly traded companies. The index is widely regarded as the best gauge of large-cap U.S. equities.
VIX	Daily	Yahoo Finance	The CBOE Volatility Index, is a real-time market index representing the market's expectations for volatility over the coming 30 days. Investors use the VIX to measure the level of risk, fear, or stress in the market when making investment decisions.
Net Export of Goods and Services	Quarterly	Quandl	Net exports are a measure of a nation's total trade. A nation that has positive net exports enjoys a trade surplus, while negative net exports mean the nation has a trade deficit.
10-Year Treasury Constant Maturity Minus Federal Funds Rate	Monthly	Quandl	Interest Rate Spread between 10 year Treasury yield and federal funds rate: wider spreads anticipate economic upswings and narrower spreads anticipate downturns
Target - US GDP	Quarterly	Quandl	Gross domestic product (GDP), the featured measure of U.S. output, is the market value of the goods and services produced by labor and property located in the United States.

# Data Pre Processing



EDAs

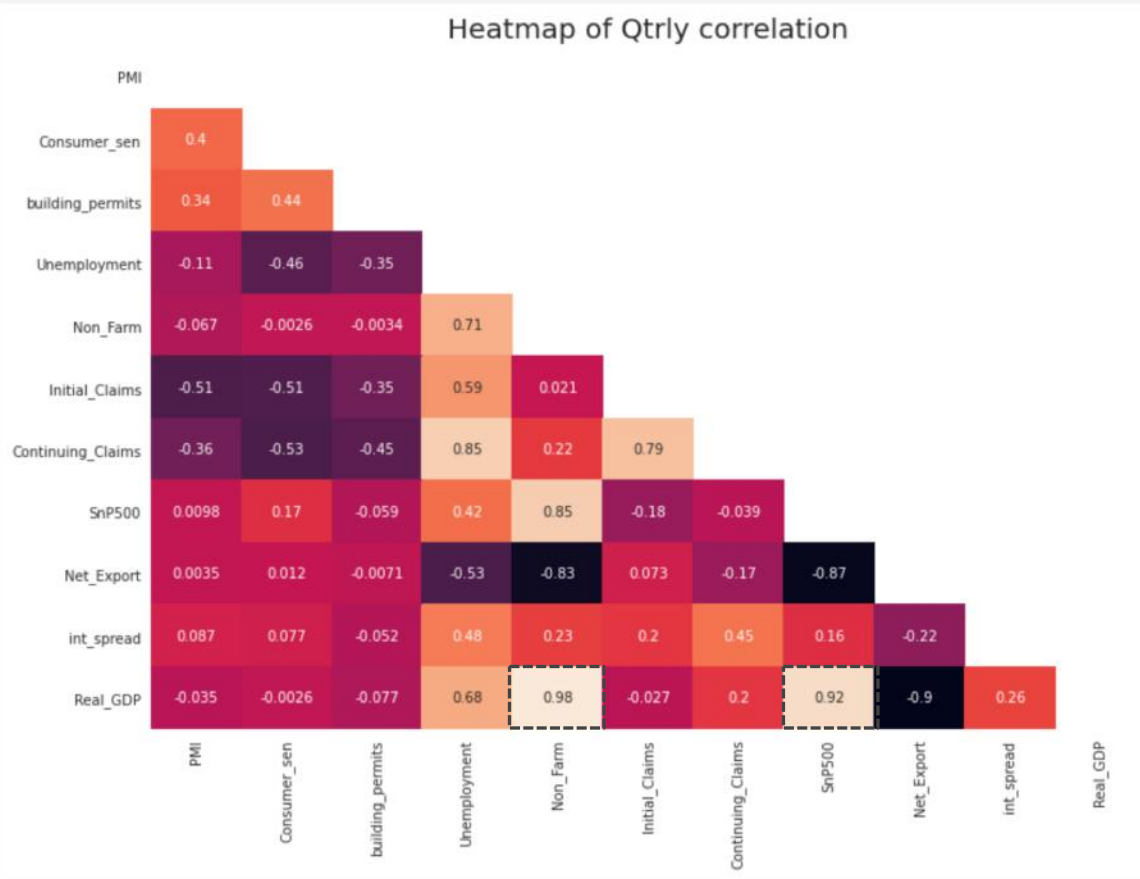
### Overview of Monthly Indexes



### Daily S&P 500

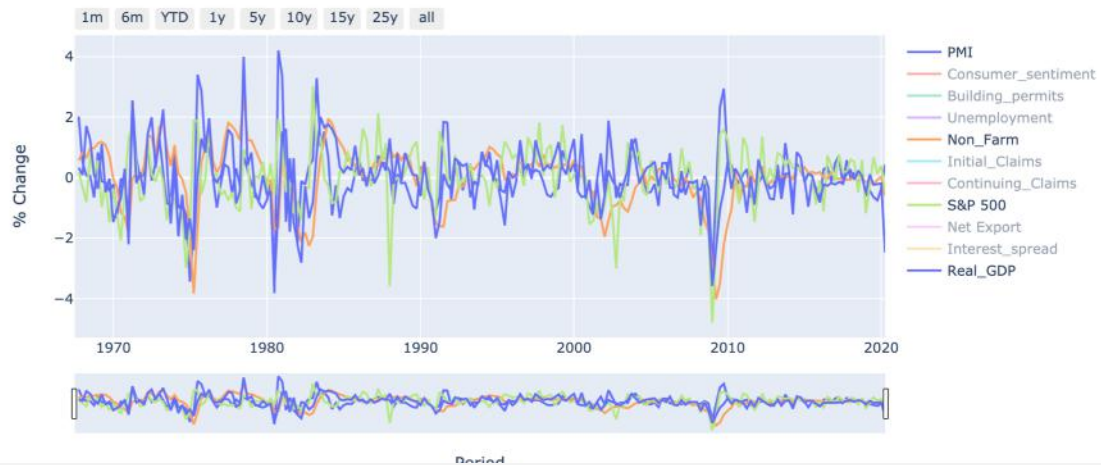


# EDAs



EDAs

Correlation of Quarterly Indexes





03

Modelling

## Classification

Logistic Regression

True Negatives: 5  
False Positives: 18  
False Negatives: 0  
True Positives: 136

	pred GDP down	pred GDP up
actual GDP down	5	18
actual GDP up	0	136

Accuracy : 0.8867924528301887  
Sensitivity : 1.0  
Specificity : 0.21739130434782608  
Precision : 0.8831168831168831

BASE  
MODEL

## Regression

Linear Regression

$R^2$ : 0.99613  
RMSE: 276.174



# Time-Series Modelling Techniques used in this Capstone

## Classification



### LSTM- Neural Network

A type of recurrent neural network capable of remembering the past information and while predicting the future values

## Regression



### FB Prophet

Prophet is an open source software released by Facebook's Core Data Science team

## Regression



### Sarimax

SARIMAX model is a linear regression model that uses a SARIMA-type process

## Regression



### Auto-Arima

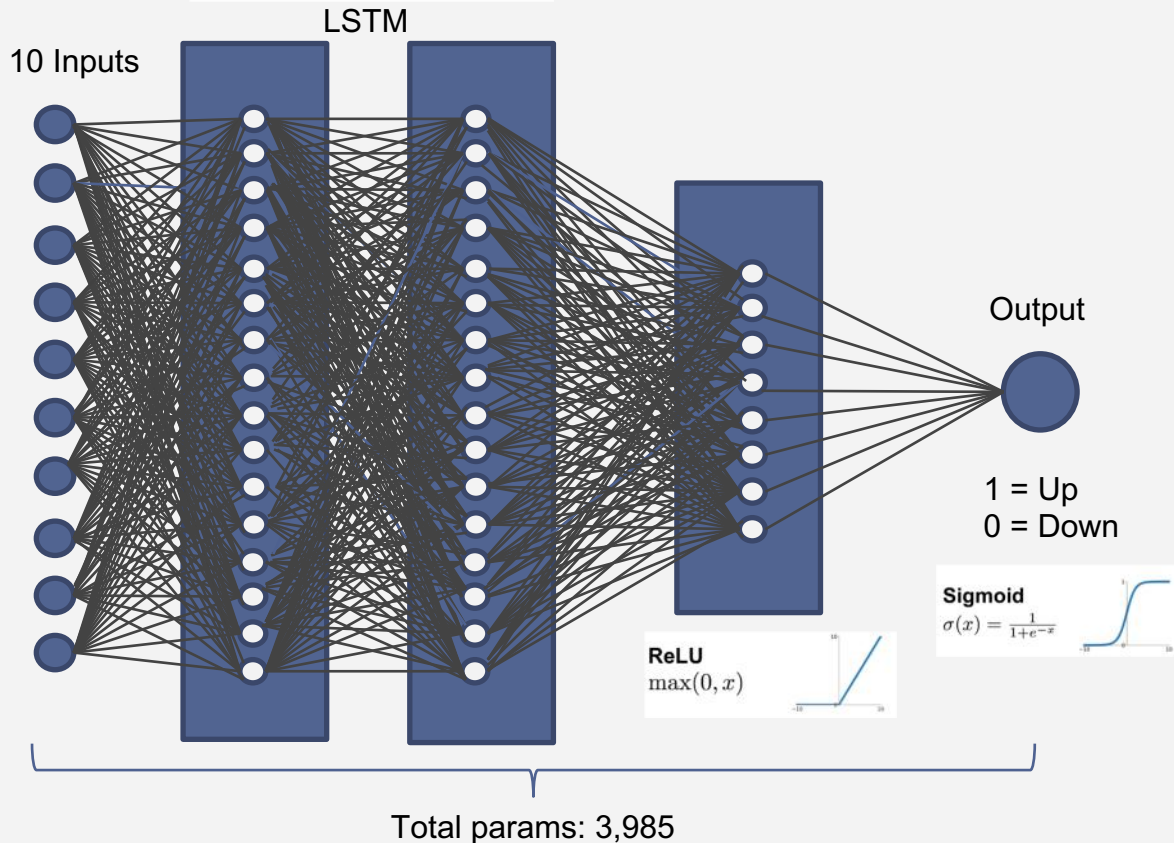
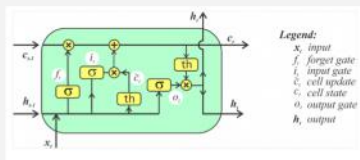
Returns best ARIMA model according to either AIC, AICc or BIC value

# 4 layer LSTM RNN

Classification problem  
(1 for GDP is up and 0 for GDP is down)

Lookback length=2, Batch size=8

Layer (type)	Output Shape	Param #
lstm_18 (LSTM)	(None, 2, 16)	1728
dropout_7 (Dropout)	(None, 2, 16)	0
lstm_19 (LSTM)	(None, 16)	2112
dense_18 (Dense)	(None, 8)	136
dense_19 (Dense)	(None, 1)	9
Total params: 3,985		
Trainable params: 3,985		
Non-trainable params: 0		



# 4 layer LSTM RNN

Classification problem  
(1 for GDP is up and 0 for GDP is down)

Lookback length=2, Batch size=8

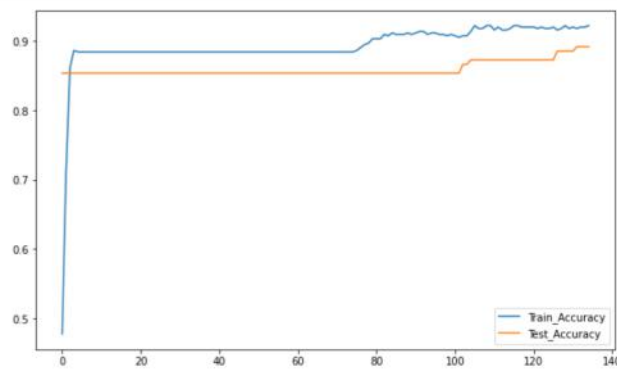
Layer (type)	Output Shape	Param #
lstm_18 (LSTM)	(None, 2, 16)	1728
dropout_7 (Dropout)	(None, 2, 16)	0
lstm_19 (LSTM)	(None, 16)	2112
dense_18 (Dense)	(None, 8)	136
dense_19 (Dense)	(None, 1)	9
Total params: 3,985		
Trainable params: 3,985		
Non-trainable params: 0		

Split train/test  
data (75/25)

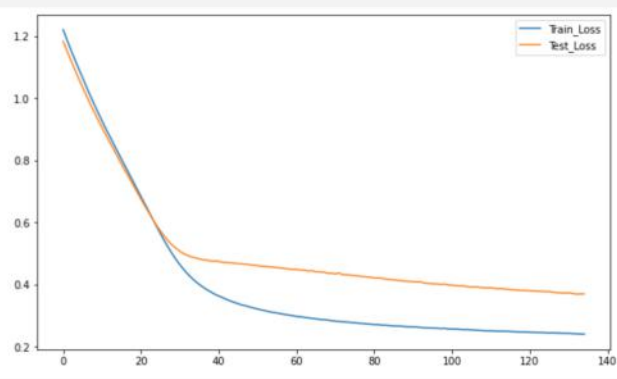
Scale data,  
time series  
generator

Instantiate  
model , added  
regularizers

Compile, fit  
and predict



Train accuracy: 0.9179  
Test accuracy: 0.8917



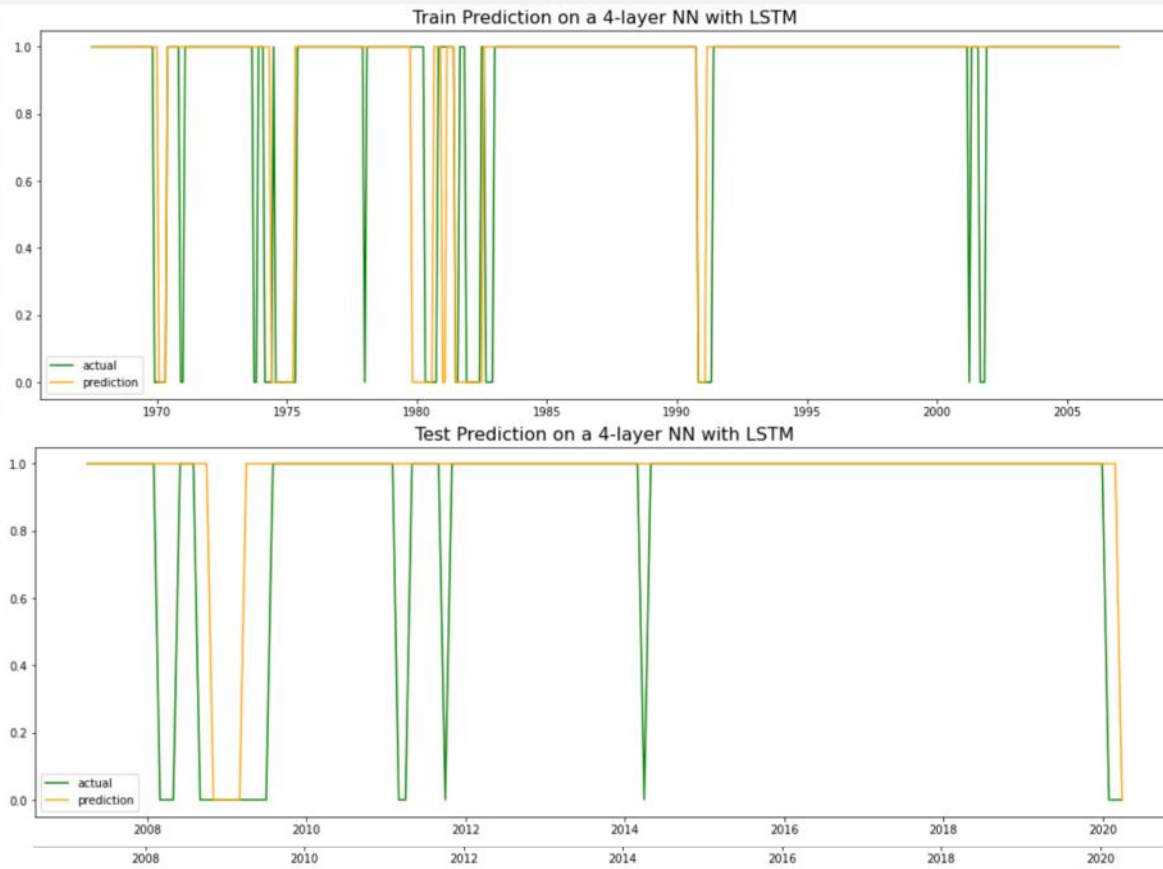
Train loss: 0.2397  
Test loss: 0.3693

# 4 layer LSTM RNN

True Negatives: 6  
False Positives: 15  
False Negatives: 0  
True Positives: 136

Accuracy : 0.9044585987261147  
Sensitivity : 1.0  
Specificity : 0.3  
Precision : 0.9006622516556292

	pred GDP down	pred GDP up
actual GDP down	6	15
actual GDP up	0	136



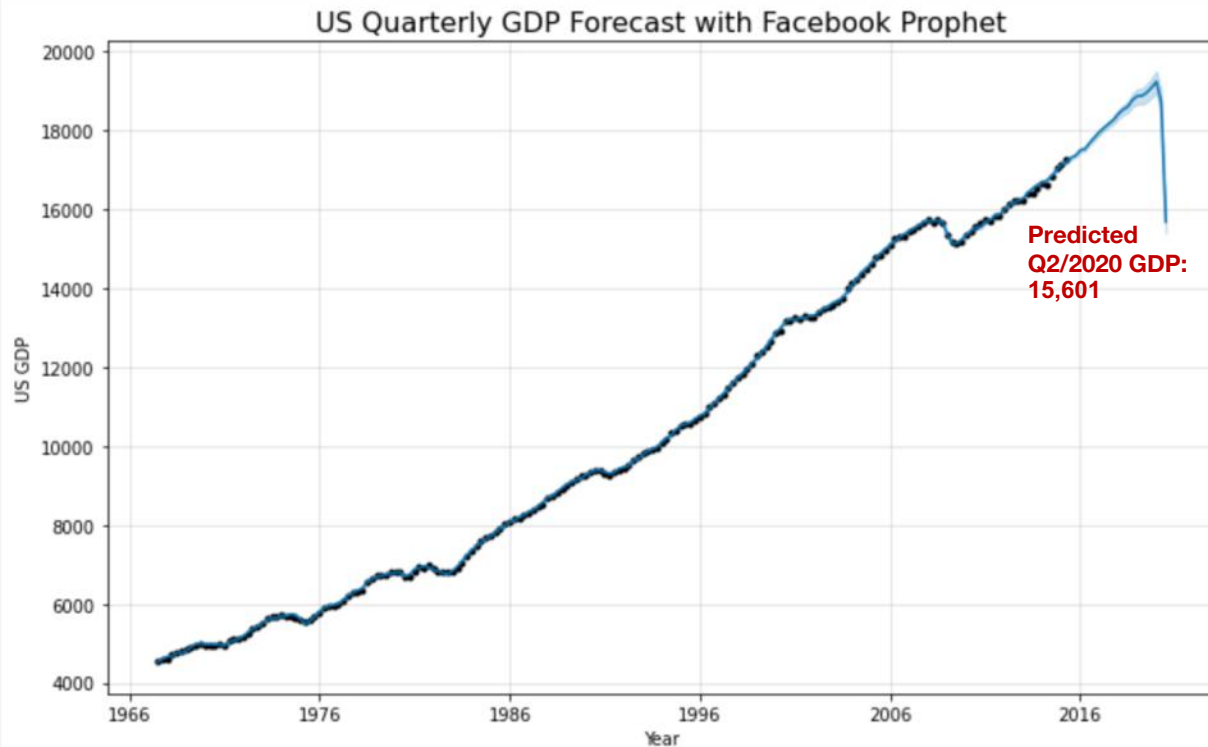
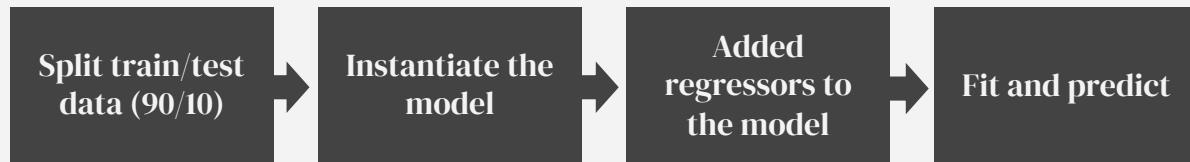
Prediction: Q2 2020 (0) GDP down

# FB Prophet

- $r^2$ : 0.9998
- mean\_squared\_error: 2118.38
- root\_mean\_squared\_error: 46.025
- mean\_absolute\_error: 33.997

- FB Prophet predicted Q2 2020 to drop to 15,601 (~15% drop from previous quarter).

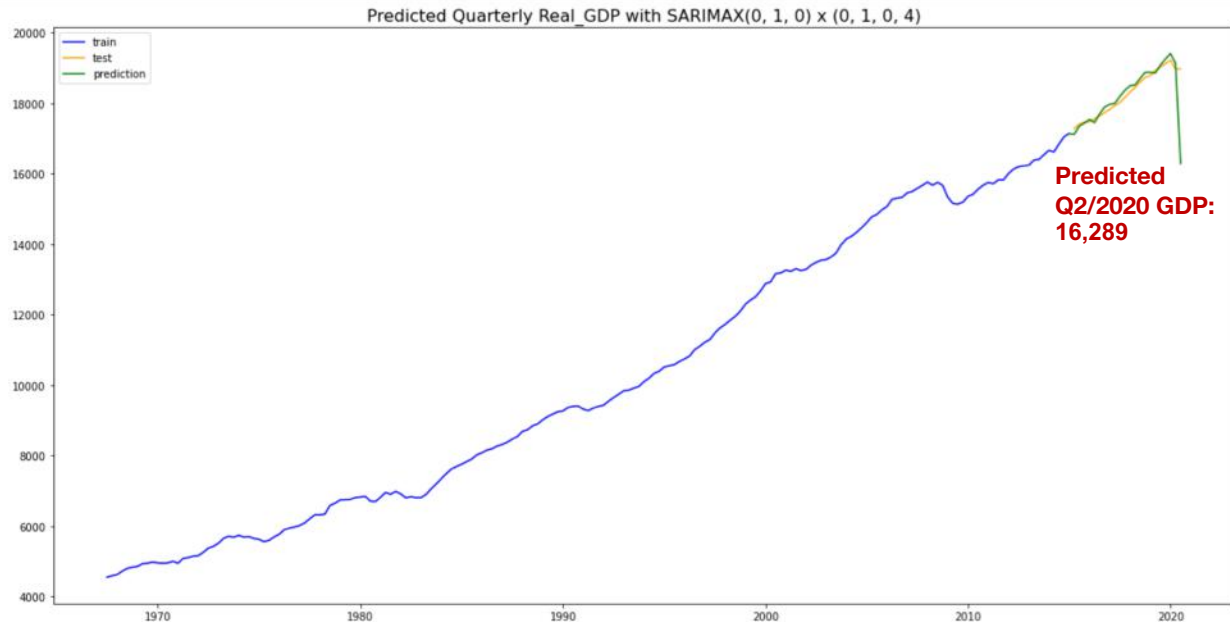
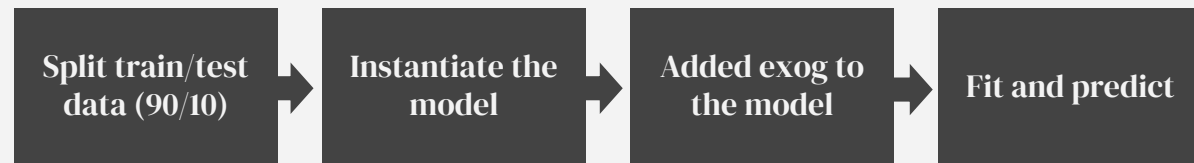
- It is comparable to the analysis by Deloitte Analysis published on 15 Jun 2020.



# Sarimax

- $r^2$ : 0.9597
- mean\_squared\_error: 15,920.6826
- root\_mean\_squared\_error: 126.177
- mean\_absolute\_error: 115,696

- The SARIMAX model predicted a -14.86% drop in Q2 2020 quarterly US GDP
- Predicted Q2/2020 GDP: 16,289.79



# Auto-Arima

- $r^2$ : - 0.630
- mean\_squared\_error: 644,286.6
- root\_mean\_squared\_error: 126.177
- mean\_absolute\_error: 546.263

- The Auto-Arima model predicted a - 15.91% drop in Q2 2020 quarterly US GDP

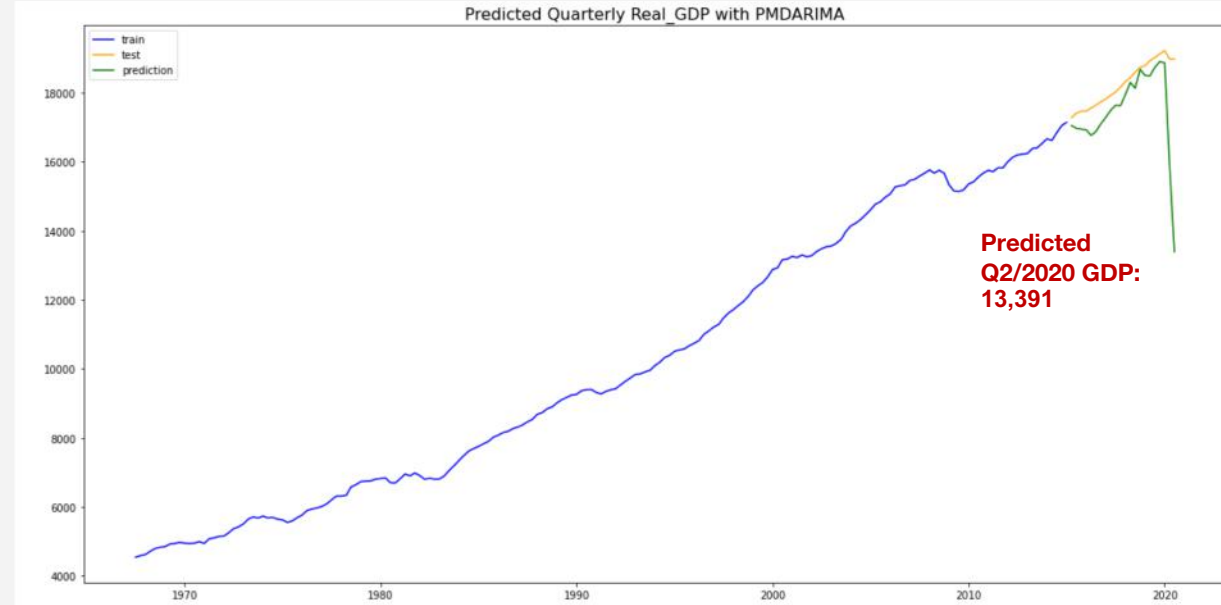
- Predicted Q2/2020 GDP: 13,391

Split train/test  
data (90/10)

Instantiate the  
model

Added exog to  
the model

Fit and predict



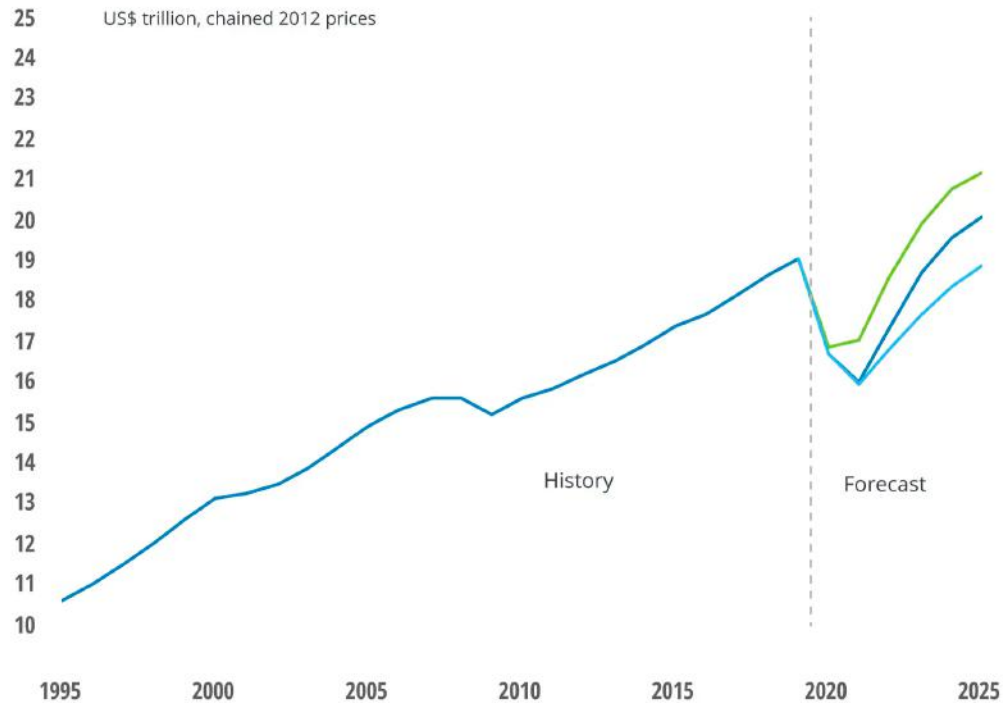
# Deloitte Analysis

<https://www2.deloitte.com/us/en/insights/economy/us-economic-forecast/united-states-outlook-analysis.html>

FIGURE 1

## Real GDP, US\$ trillion

— Baseline — Fast bounce back — No end in sight



Source: Deloitte analysis.





Conclusion

04

# Summary



## Time Series modelling

Forecasting on time series problems can be performed on LSTM-NN, FB prophet, Sarimax models and they provide high accuracy/ good RMSE scores.

## GDP Prediction

We are able to predict the US Quaterly GDP accurately as shown in the results. Though it started as a classification problem, we are able to perform regression thus enabling investors to have more anchor points to support the macro view of the economy, thus making better investment decisions.

## Generalization

The model can be deployed to analyse the GDP in other countries as they all uses similar features.



Thanks

Does anyone have any questions?