Capstone Project: CO2 Emissions Forecasting

DSI-TOR-9
Nov 20 2020
Ayshnoor Dewji

Problem Definition

Climate change is a significant problem

- Which countries are contributing the most currently to CO2 emissions?
- What are possible forecasts of CO2 given current trends (on a per country basis)?

Goal of the Web App:

To provide reporting on the current landscape & to predict possible future emissions (country level)

Ultimately this web app is to inspire and to trigger actions (in order to develop solutions to climate change)

Data Science Process

Future Steps

Data sourced from Gapminder.org on a per Country basis

GDP: University of Groningen - Faculty of Economics and Business
CO2 Emissions: Carbon Dioxide Information Analysis Center
Population: Gapminder based on Maddison & UN

Data as far back as 1800; but data for modelling is 1949 to 2018

Forecasts generated on a per country basis

Three different forecasting models were used to generate predictions

Nearly 600 predictive models ran in total.

Three Different Forecasting Models (Analyze)

Variable Autoregression or VAR
(Multivariate)

- GDP
- Population
- CO2 Emissions

FB Prophet Model
(Univariate)

- CO2 Emissions

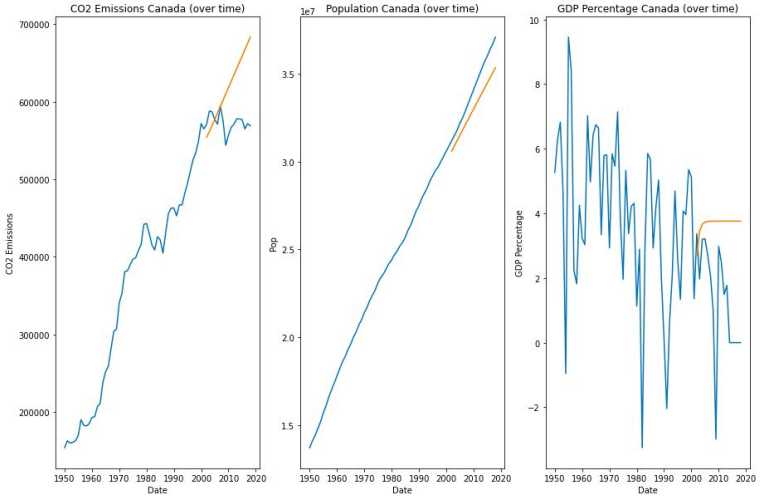
Seasonal Autoregressive integrated moving average or SARIMA
(Univariate)

- CO2 Emissions

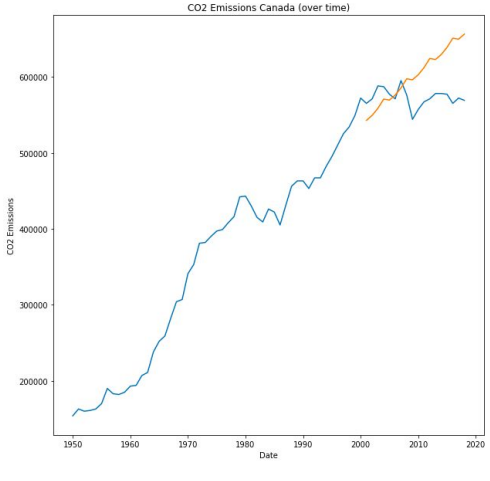
Models ran for each individual country
Forecasts in the app are go forward forecasts, meaning they use the whole data set to project future values.

Forecasting is difficult and the idea to run many types of models was to get a range of possible future outcomes.

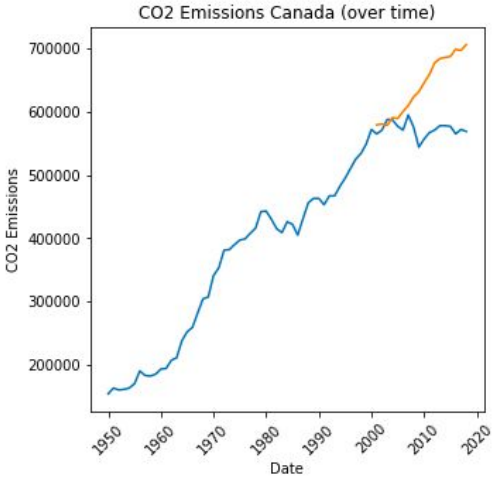
Model Train & Test Split - Canada Comparisons



VAR



FBProphet



SARIMA

Variation of forecast accuracy across the different types of models.
Train/Test and MSE/R2 Scores between the models show terrible results!
Forecasting is difficult and there are additional factors that are likely influencing the results.

Demonstration Time!

Let’s look at Australia and ask the app for a 30 year projection

Future Work/Next Steps

1) Deployment on Heroku and Refinement of models (additional hyperparameter tuning)
2) Aggregation of forecast to create a regional or global projection
3) Recommendation on how to reduce CO2 emissions in various regions (based on possible projections)