When will the Cherry Trees bloom?

Cherry Blossom Prediction Competition by Department of Statistics, George Mason University

By Dr. Zeehasham Rasheed

Adjunct Professor, CS and IST Departments, George Mason University

zrasheed@gmu.edu

For Northern Virginia residents, it is February, and our Metro DC weather is turning warmer. In our Washington, DC Metro Area, the sentinel event of late March and April is the blooming of our cherry blossoms. For a short 5- to 14-day period we marvel at the beauty of spring. Lets see if our Machine Learning model can help us predicting the correct dates.

Cherry blossoms are an icon of spring and are celebrated in many cultures of the temperate region. For its sensitivity to winter and early spring temperatures, the timing of cherry blossoms is an ideal indicator of the impacts of climate change on tree phenology. In this competition work, we applied a Deep Learning model for temperate deciduous trees to predict peak bloom dates using temperature data.

We collected the temperature data from different sources. Please see reference section to access these sources.

Exploratory Data Analysis

At the beginning of our analysis, we found some interesting insights, especially related to correlation between temperature and blossom dates. Below is some analysis snapshots.

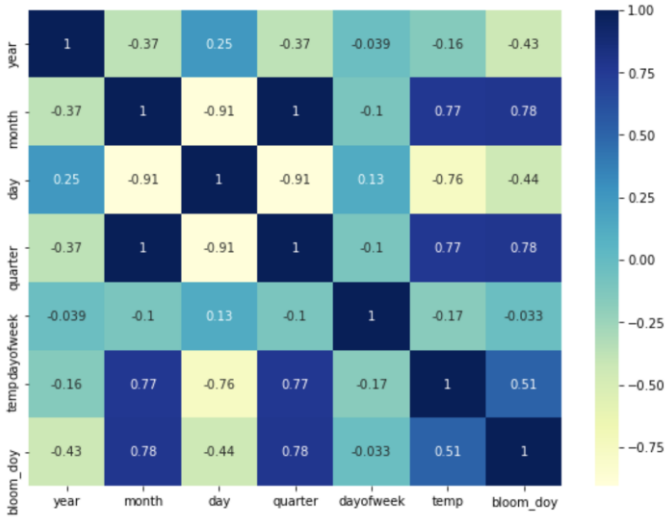

Figure 1: Correlation Coefficients of different variables with bloom doy

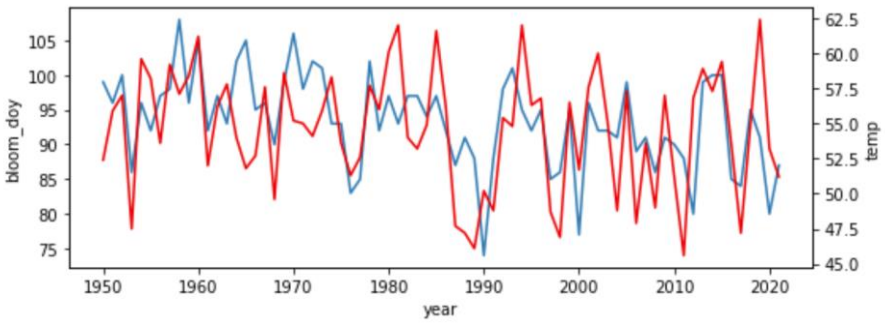

Figure 2: Timeline showing Temperature and Bloom doy relationship

From Figure 1 Correlation plot, we see that Quarter and Month shows strong positive correlation with bloom doy. If we see temperature, it is showing a correlation coefficient of 0.51 which is although not very strong but provides some confidence to predict bloom doy with respect to temperature. (For readers, Correlation Coefficient value close to 1 means strong positive correlation between variables and 0 means no correlation between variables)

From Figure 2 Timeline plot, we can notice that most of the years showing higher temperature with longer bloom doy. So, we finally decided to go with temperature data.

Machine Learning

Now, we are confident about our data, next step is to select the appropriate Machine Learning algorithm. We use Long short-term memory (LSTM) algorithm which is an artificial recurrent neural network (RNN) architecture and used in the field of deep learning. Long Short-Term Memory (LSTM) are capable of learning order dependence in sequence prediction problems. This is a behavior required in complex problem domains like time series, speech recognition and more. Since we assume that cherry blossom date depends on series of temperature in the past, so learning these patters through LSTM is a good direction for our prediction.

During the model training and testing, our evaluation shows reasonable performance with low error rate.

Final Predictions

We plot last 10 years of actual bloom doy along with next 10 year bloom doy in our results. Blue color represents actual bloom doy and red color represents next 1o years prediction. We show results for all the four different locations.

- Washington DC, USA
- Kyoto, Japan
- Liestal Switzerland
- Vancouver BC, Canada

Washington DC Predictions

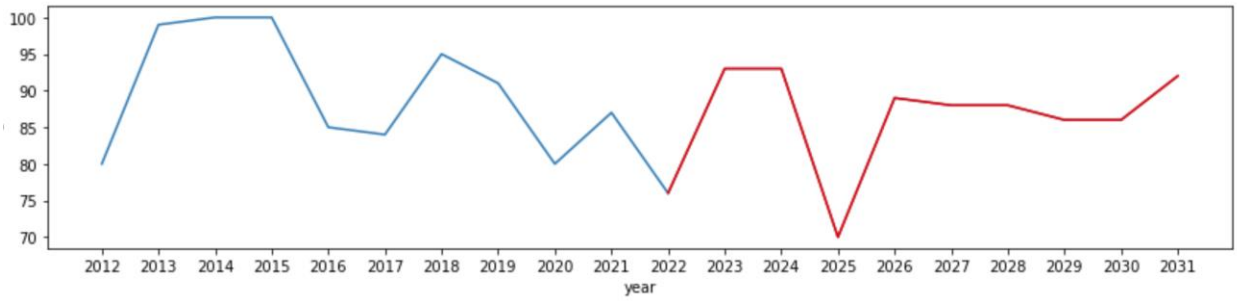

Figure 3: Washington DC Bloom DOY prediction trend

Kyoto Predictions

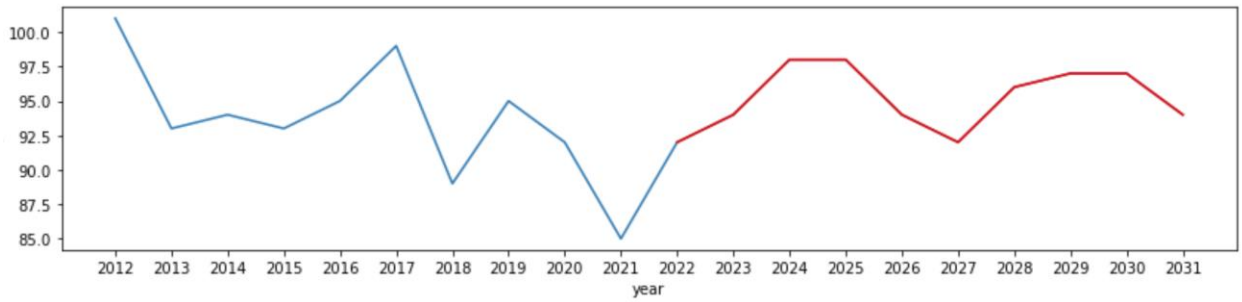

Figure 4: Kyoto Bloom DOY prediction trend

Liestal Predictions

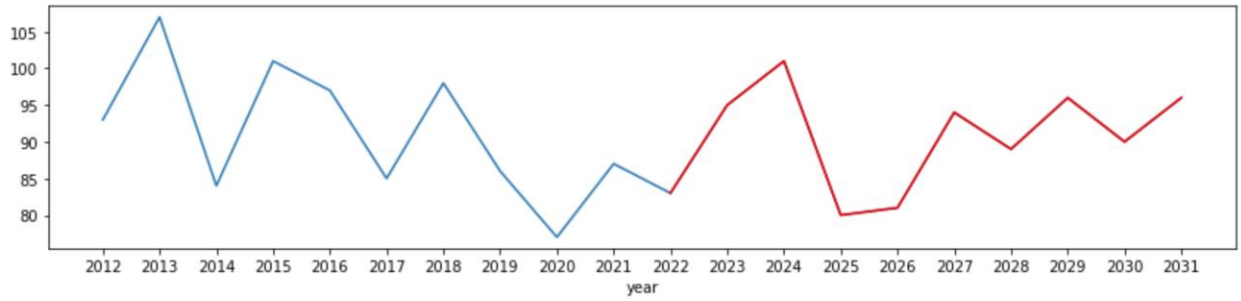

Figure 5: Liestal Bloom DOY prediction trend

Vancouver Predictions

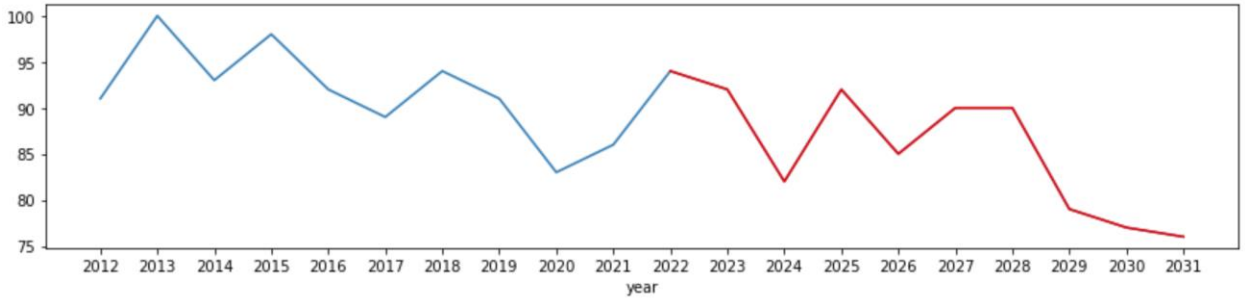

Figure 6: Vancouver Bloom DOY prediction trend

Conclusion

With our model, we are able to produce the following predictions for the next 10 years.

year	kyoto	liestal	washingtondc	vancouver
2022	92	83	76	94
2023	94	95	93	92
2024	98	101	93	82
2025	98	80	70	92
2026	94	81	89	85
2027	92	94	88	90
2028	96	89	88	90
2029	97	96	86	79
2030	97	90	86	77
2031	94	96	92	76

This is a very interesting competition. We anticipate that this type of research and simple but robust model based climate data points would provide valuable insights for developing adaptation strategies to climate change in conservation planning, restoration and other related disciplines and industries.

References

Temperature Data Sources

- Washington DC https://www.weather.gov/wrh/Climate?wfo=lwx
- Kyoto https://www.data.jma.go.jp/obd/stats/etrn/view/monthly_s3_en.php?block_no=47759&view=1
- Liestal Switzerland https://climateknowledgeportal.worldbank.org/country/switzerland/climate-data-historical
- Liestal Switzerland 2 https://climateknowledgeportal.worldbank.org/download-data
- Vancouver BC https://vancouver.weatherstats.ca/download.html