SciPy2019
Scientific Computing with Python
Austin, Texas · July 8-14

Use of Neural Networks to Predict Coffee Rust

Laura Kahn | thedatalass.com/coffee-rust/ @LauraHKahn

Can artificial neural networks predict coffee rust? Coffee rust infestation leads to production losses of over $1 billion annually worldwide. Coffee is the second largest traded commodity worldwide, with about $100 billion in volume traded annually. The livelihood of at least 120 million people worldwide depends on the coffee supply chain. Coffee rust is the deadliest disease that attacks the coffee plant and is caused by Hemileia vastatrix fungus at temperatures between 15-28°C. This research is the first known use of neural networks to predict coffee rust.

Methods

A multi-layer perceptron (MLP) artificial neural network with 5 layers, 6 neurons, and a hyperbolic tangent activation function (converts input to output function) was chosen to compensate for the imprecise dataset's non-linearity and for optimization ease.

Data Acquisition & Exploration

Over 1580 weekly observations of Temperature, Rainfall, Production, and Futures data from Minais Gerais, Brazil (Koppen-Geiger Cwa Humid subtropical climate) from January 1993 - July 2018 were acquired. These co-variates were used to predict rust amount each week in the Minais Gerais, Brazil growing region. The model is generalizable to other coffee growing regions with a Koppen-Geiger climate such as Mexico and Jamaica. Data acquisition challenges included a small set of covariates and data sparsity. These challenges were overcome by using Beretta's k-Nearest Neighbor imputation method with the rkN-imputer and sk-rebate libraries to fill in the 48.8% of missing rust values (773). Missing data was assumed to be missing at random. Matplotlib and Seaborn libraries were used for exploratory data analysis and visualizations.

Results

The Mean Squared Error (MSE) for the MLP ANN model performed 131 times better (54.41) than the Logistic Regression model (6619.777) or Decision Tree Regression model (6179.333).

Conclusion

Artificial neural networks can be applied in the agricultural commodities field to predict with a fair amount of certainty the weekly coffee Rust percentages in a Humid Subtropical coffee-growing region using historical Rain, Temperature, Production and Futures co-variates.