What data science can offer a botanic epidemiologist

Sith Jaisong
Plant Disease Management Group, CESD, IRRI
Los Baños, Philippines
s.jaisong@irri.org

Whereas a traditional data analyst may look only at data from a single source (often one experiment), data science enables us explore and examine data from multiple disparate sources. Data science is the science adopting techniques and theories from broad areas of mathematics, statistics and computer science for extracting information from large and complex data. Scientists are now facing challenges handling and processing data due to the amount and complexity. In order to manipulate these data, data science can be helpful. As botanic epidemiologists, we explore the causes of plant disease epidemics. To investigate the causes, we often analyze disparate data, combining weather or climatic data (e.g., temperature or relative humidity) and relevant variables that we think can be used to explain why disease occurs. As a result of the sheer amount of data available to us today and because either the data or the tools to analyze the data didn't exist before, data science potentially gives botanic epidemiologists new opportunities to discover previously hidden insights into why diseases occur. In this seminar, I will present how scientists apply ideas and processes from data science and give some examples of the results from applied data science. Finally, I will show the possible ways that data science can benefit botanical epidemiologists in their own research.