

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 (a set of experiments), 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 facing challenges handling and processing the data due to the amount and complexity of data. In order to manipulate these data, data science is very helpful and also it is equipped with variety of techniques from statistics. As botanic epidemiologists, we explore the causes of plant disease epidemics. To investigate the causes, we analyze the data by combining weather or climatic data (e.g., temperature, relative humidity) and relevant variables that we think can be used to explain why disease occurs. This process creates more data, and consequently may lead to a new extensive result. Because either the data or the tools to analyze the data didn't exist before, data science potentially gives botanic epidemiologists the opportunities to discover previously hidden insights into why diseases occur. In this seminar, Sith will present how scientists apply the ideas and processes from data science, give some examples of the result applied data science. Finally, Sith will show the possible ways that data science can involve in botanical epidemiological studies.