



Bayer at a glance

Innovation Site at Purdue:
Bayer at Convergence
101 Foundry Dr. Suite 4500
West Lafayette, IN 47906.

*Health for all,
hunger for none*

3

divisions:
Pharmaceuticals,
Crop Science,
Consumer Health

87

countries with
legal entities

€43.5bn

net sales
in FY2019

~104 K

employees worldwide
in 2019*

~€5.3bn

investment in R&D in 2019



Tackling two of the most
pressing challenges of our time:
Health & Nutrition



Diverse, international focus with
cross-border and cross-division
people development

* As of December 31, 2019; employees in full-time equivalents



Project: Yield Prediction through ML



Description:

With the recent success of using **advanced machine learning** methods, such as **deep learning** for example, across various domains and problems, several new open-source tools and technologies with various implementations are now available. One of the use cases that we are exploring within Plant Breeding at Bayer Crop Science team is to use **genotypic data and environmental data with deep architectures** or other newer methods understand the plant response (yield, primarily) as function of the genotypic and environmental variables and **build models that can capture the interactions** between them.

The goals of the candidates working in the project will be to **comprehensively evaluate various machine learning methods to predict yield**. We will provide a data set containing the genomic data and the phenotypic data obtained through large scale field experiments across hundreds of locations. When combined with publicly available weather data at these locations, we have a rich data set with several layers of information. Being able to predict the plant response has a significant impact on the future of plant breeding and agriculture in general. We can meet the diverse needs of growing populations by creating new varieties of seed faster and with fewer resources, like land and water, leading to a global food security and sustainability.



Project: Yield Prediction through ML



Keywords:

Plant breeding; genomics; deep learning; environmental data; GxE interactions.



Tools/skills that will be used:

R, Python, Machine Learning methods, Genomic Prediction techniques.



Preferences for student profile:

Introductory Machine Learning and familiarity with genomic data is a plus, but not required. All students welcome regardless of background and citizenship status.