syngenta

Using Collaborative filtering algorithms to optimize trial network- Syngenta

Syngenta is a global Swiss agribusiness that produces agrochemicals and seeds. As a leading global organization, we help farmers feed a fast-growing world population by bringing plant potential to life. At Syngenta, you will work with great people and teams with exciting opportunities for personal development.

From research, production to market, Syngenta produces, treats seeds before selling them to farmers. To optimize this process, Seed Production Research (SPR) is a crucial team which assists the seed selection and gives advices both internal and external customers the right "recipe" under different climate conditions.

SPR Engineers and technicians collect data from in field trials every year in different location to evaluate the potential, stability and risk of our hybrids/inbreeds. The challenge we face today is information we collect in the field are not always in the format of rating or score. Actually, we generate rating/score based on comments noted down during field scouting. In this process, valuable information are not fully used.

INTERN WILL LEARN AND BE EXPOSED TO:

- Applying statistics knowledge on real world data (seed industry)
- Team work with our local/global statistician team
- Cutting edge decision making and statistical software/Tool
- Participation in various projects
- Direct interaction with customers

MISSIONS:

- Collaborative filtering (CF) is a technique used by recommender systems. Social networks (Facebook), E-commerce and streaming services (Netflix) use CF to generate short-list of items that a given user will likely appreciate.

- Stage 1: replace missing values

  We would like to apply the same method on our crop performance trial data to replace the missing values and to get more reliable results. Due to certain limits, e.g. limited quantity of trialing seeds, weather damage etc., yield results of certain varieties are not available in certain locations/ certain years.

- Stage 2: optimize trial network by CF algorithms

  Certain trials might not be necessary if CF algorithms can successfully replace a part of the trialing data with high accuracy.

User 1	★★★	★★★★★	★★
User 2	★★★★	?	★
User 3	★★	★★	★★★

Variety A	4	?	3
Variety B	3	1	4
Variety C	2	2	1

- This topic is a part of the project meta-analysis on multi-year multi-loc trial data. Successful candidate will need to work closely with other statisticians in the team on project implementation, e.g., data mining/text mining on risk calculation, R shiny interface prototyping etc.

PREFFERRED QUALIFICATIONS:

- Bac+4/5 or equivalent level in statistics, applied mathematics or related subjects
- Research-oriented and detail-oriented
- Have a good knowledge in R/Python
- Fluent in French and English
- Flexibility and adaptability
- Team player
- Interested in seed/crop protection industry
- Demonstrated skills in selecting the right statistical tools given a data analysis problem.
- Demonstrated effective written and verbal communication skills.

STARTING DATE: April
DURATION: 4 months

Please send your CV and motivation letter to
Yaojie SHI – Yaojie.Shi@syngenta.com
Clara JACOB – Clara.Jacob@syngenta.com