Biogen

Position Summary

The nonclinical and clinical biomarker group in Biostatistics Department provides statistical support in activities including nonclinical, preclinical and clinical biomarker projects.
The intern's research will focus on evaluating and/or developing innovative and multivariate machine learning methods that may provide improvements in classification accuracy compared to the common machine learning approaches for small-size overlapping data with class imbalance issue. The work encompasses literature review and simulation.

This is a full-time internship role from June – August 2018.

The Biostatistics Intern position provides an exciting opportunity to gain hands-on experience and exposure to the biotechnology/pharmaceutical industry. Working closely with Biometrics team in Biogen, the intern could expect to learn the following:

- To develop and compare machine learning methods with subsampling approaches under various scenarios to mimic the clinical settings
- To compare the performance of those learning methods using different performance metrics
- To identify any potential methods in differentiating patients from different treatments cohorts
- To identify any potential disease subclasses and/or patient subgroups

Key Responsibilities:
- Perform literature review on current and innovative methodologies.
- Conduct simulation studies to evaluate the selected methods.
- Present and write up statistical research projects

The position is based in Cambridge.

The ideal candidate will have statistical simulation study experience under UNIX/Linux environment.

Qualifications

To participate in the Biogen Summer Intern Program, students must meet the following eligibility criteria:
- Legal authorization to work in the U.S.
- Enrollment in a full-time Ph.D program, returning to the academic program following Biogen internship assignment
- At least 18 years of age prior to the scheduled start date
- Completed at least one year of graduate courses
- Passed the qualification exam (preferred)
- Deep knowledge in statistical theory and methods required

- Proficient in R programming; Familiarity of UNIX/Linux preferred
- Experience in statistical simulation preferred
- Some knowledge in machine learning (supervised, semi-supervised and unsupervised learning) preferred
- Basic knowledge of biomarkers, clinical development and neurodegenerative diseases preferred

Education
- Ph.D student completed at least one year of graduate program
- Desired majors include: Statistics (preferred), Biostatistics (preferred), Data Science, Computer Science