

# Bioinformatics Analyst-Machine Learning and Predictive Modeling

## Job Description:

I am always seeking to develop a data-science team to solve the many computationally complex problems we encounter in bioinformatics and health data science. We work towards the development of new therapeutics and immunotherapies that leverage the patient's own immune-system to eradicate disease. Our work and findings help to improve the treatment and care for cancer patients around the world.

The data we deal with is truly huge: individual datasets can be hundreds of terabytes, and can include numerical, textual and imaging data. The team we are building is incredibly diverse, with biologists, statisticians, software engineers, bioinformaticians, mathematicians and data scientists. Together, we are all working as a team toward the common goal of developing new therapies for cancer patients. Use your technical skills to tackle one of the hardest problems of our time as you continue to grow into data-science.

We are searching for a machine learning engineer passionate about using cutting edge prediction methods to drive the next round of discovery in immunotherapeutics. We are equipped with state of the art facilities on which to develop your data-science and high performance computing skillset. The project has access to Cyfronet Prometheus (~55, 000 cores) and CI TASK Tryton (~38, 000 cores) clusters, which are consistently represented among the top 500 super computers in the world. As the work is international in nature, we are constantly integrating with sister-

groups around the world. There will be ample training opportunities for developers who join the team and we offer an internationally competitive salary competitive with your experience.

## How to apply:

You are encouraged to contact Dr. Javier Alfaro (Javier.Alfaro@proteogenomics.ca) to ask about potential opportunities. I will put you in touch with the appropriate contact.

## Responsibilities:

The machine learning engineer is responsible for applying state-of-the-art machine learning methods towards bioinformatics projects leveraging on existing large-scale genomics datasets. You will further be involved in algorithm and computational methods development for biomarker discovery analysis for different types of cancers.

Responsibilities may also include:

- Applied machine learning
- Provide consulting on biological data analysis for collaborations conducting cutting-edge molecular biology research
- Advise and help researchers with their experimental design, data analysis and interpretation of results
- Provide training (2-5 days courses) in the area of machine learning and data science in biology
- Mentorship of junior staff and co-op students
- Collaborate and interact with other scientists across partner networks in an international, interdisciplinary, and highly collaborative work environment

Preferred Qualifications:

- Bachelor's degree (MSc preferred) or further training in Machine Learning or Bioinformatics
- Experience with version control systems (SVN, Git, Mercurial, etc)

- Experience with some of the following: R, Python, C, C++, Java
- Knowledge of UNIX/Linux environments

#### Beneficial skills include:

- Experience in analyzing single cell 'omics, imaging, -omics, proteomics, or microfluidics data sets
- Experience in teaching
- Experience in R as a package developer or maintainer
- Knowledge of Python or other scripting languages
- Knowledge of scientific authoring frameworks (LaTeX, Rmarkdown) and R-Shiny
- Experience with version control, method benchmarking, or data management technologies
- experience with Amazon Web Services, Google Cloud Engine or other clouds

- Our software engineers, statisticians and computer-scientists are encouraged to further develop their skills in data-science through the attendance of workshops, conferences etc.
- As we are an international network, our infrastructure is remote-work compatible. A benefit enjoyed by excellent communicators and team players.

#### Benefits

- You will interact internationally with researchers having state-of-the-art facilities for mass spectrometry, virology, protein biochemistry, vaccine technology and computational biology.
- Employees will have the opportunity to develop skills in machine learning and high performance computing. The center has access to Cyfronet Prometheus (~55, 000 cores) and CI TASK Tryton (~38, 000 cores) clusters, which are consistently represented among the top 500 super computers in the world.