

Ausschreibung Masterarbeit:

Training eines DNA Methylierungs basierten Krebs Klassifikators.

Junior research group transcriptome Bioinformatics

Viele Krebserkrankungen zeichnen sich durch eine Veränderung der DNA Methylierung in Bereichen des Genoms aus, die in normalen Zellen der Kontrolle einer bestimmten Art von Chromatin, genannt bivalentes Chromatin, unterliegen. Diese Stellen unterliegen Hypermethylierung in Krebs, sie sind also im Krebs stärker methyliert als in gesunden Zellen. Wir haben im Zuge einer Veröffentlichung in Scientific Reports einen Cancer Klassifikator vorgeschlagen, der sehr gute Ergebnisse in der Klassifizierung von Krebs ermöglicht. Normalerweise wird diese Art von Classifiern an Methylierungsdaten trainiert, das heisst es werden diejenigen Datenpunkte als Klassifikatoren ausgewählt, die die größte Methylierungsdifferenz in einem Trainingsset zeigen. Wir haben uns für unsere Veröffentlichung bewusst dagegen entschieden, das zu tun. Wir gehen allerdings davon aus, dass eine Verfeinerung des von uns vorgeschlagenen Klassifikators sowie eine Spezialisierung für gewisse Krebsarten möglich ist.

Wir suchen eine Masterstudentin oder einen Masterstudenten, der oder die Kenntnisse in LINUX, bash, einer Skriptsprache (Perl oder python) sowie idealerweise in R besitzt und keine Berührungsängste mit Machine Learning hat. Im Zuge der Masterarbeit sollen verschieden Ansätze des Trainings erprobt werden, um einen qualitativ besseren Klassifikator zu erhalten.

Bei Interesse mailen Sie bitte Ihre Bewerbung an:
berni@bioinf.uni-leipzig.de

Master thesis:

Training of a DNA-methylation based cancer classifier.

Junior research group transcriptome Bioinformatics

Many cancer types show a distinct DNA-methylation in distinct regions of the genome. These regions are under control of bivalent chromatin in normal cells. In cancer, these regions are hypermethylated, i.e. they show higher methylation than in normal cells. We recently published a cancer classifier that showed promising results in classifying a diverse set of different cancers. Most methylation classifiers are trained on methylation data, using the data points showing the biggest differences between two groups. For our publication in scientific reports, we decided against this approach. However, we think that our classifier can be improved by training methods, and be adapted to work even better for single cancer types. We are looking for a student to perform these training steps in course of her or his master thesis in bioinformatics. We expect knowledge in LINUX, bash, a scripting language (Perl or python) and possibly R as well as

willingness to learn about machine learning if necessary. In course of the thesis different approaches to get to a better classifier should be implemented. If you are interested in this work please contact:
berni@bioinf.uni-leipzig.de