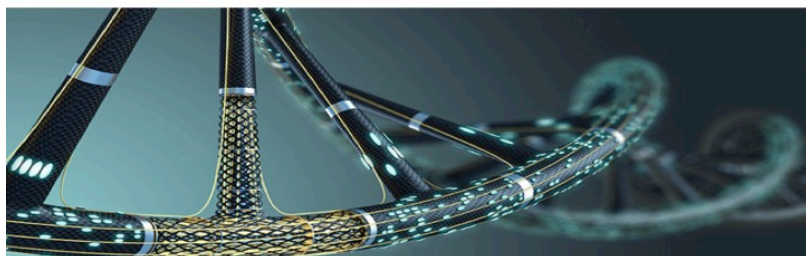


24
октября
вторник

Коллоквиум
факультета
компьютерных наук
НИУ ВШЭ



Мария Попцова

НИУ ВШЭ

Методы машинного
обучения и большие
данные
биоинформатики

Проект расшифровки первого генома человека занял 13 лет, потребовал около 1,5 миллиарда долларов и работы огромного числа институтов и университетов мира. Революция в технологиях секвенирования, произошедшая в начале 21 века, позволила сократить затраты до 2 дней и 1000 долларов. Технологии секвенирования следующего поколения (Next Generation Sequencing, или NGS) производят данные геномики, эпигеномики, транскриптомики, протеомики, метаболомики и других “омик” молекулярной биологии. Как только стало возможно секвенировать буквально “все”, были запущены международные консорциумные проекты, такие как проект 1000 геномов человека, Хар Мар – исследование разнообразия человека на 450 геномов трех рас, ENCODE – энциклопедия ДНК-элементов, The Roadmap Epigenomics (маркировка эпигенетических факторов, формирующих ткани) и проекты по секвенированию всех типов раковых опухолей (The Cancer Genome Atlas, TCGA и International Cancer Genome Consortium, ICGC). Биоинформатика на наших глазах стала областью, быстро генерирующей большие данные, нуждающиеся в обработке и интерпретации. В лекции я расскажу о том, что это за данные с точки зрения аналитика данных и как методы машинного обучения успешно применяются для решения задач аннотации и поиска новых связей между функциональными элементами генома. Предварительное знание биоинформатики не предполагается.

24 октября, 18:10 – 19:30
Кочновский проезд, 3, ауд. 205

Заказать пропуск на проход в здание
можно на computerscience@hse.ru

