

Введение в вычислительные социальные науки

Иван Смирнов

кандидат наук, заведующий лабораторией вычислительных социальных наук

Социальные науки стремительно компьютеризируются. Данные из социальных сетей и другие цифровые следы позволяют проводить исследования на гораздо больших выборках и с гораздо большей детализацией, чем это возможно с помощью традиционных методов. Новыми становятся не только данные, но и методы их анализа. Так, например, методы машинного обучения все чаще используются в социальных исследованиях. В связи с этим появилось и активно развивается новое направление исследований, получившее название вычислительные социальные науки (computational social science).

Цель курса – познакомить слушателей с основными исследованиями в этой области за последние годы и помочь им понять, как они могут использовать новые методы в собственных исследованиях.

Так как лучше всего учиться на примере ведущих специалистов, то курс построен вокруг обсуждения статей в области computational social science, опубликованных в лучших зарубежных журналах. Помимо этого курс включает обсуждение этики исследований, а также типичных ошибок и ограничений изучаемых методов.

Практические занятия проходят в формате перевернутого класса. Слушателям предоставляются материалы для самостоятельного изучения, а на очных занятиях обсуждаются возникшие проблемы и вопросы. На практических занятиях используется язык программирования Python, однако предварительных знаний они требуют. Тем, кого интересует прежде всего получение технических навыков, этот курс, скорее всего, не подойдет, так он направлен на достижение другой цели.

Итоговая оценка за курс выставляется по результатам подготовки и устной защиты проекта исследования (само исследование при этом выполнять не нужно). Выполнение итогового проекта возможно в группах.

Примерный тематический план

1. Введение в вычислительные социальные науки.
2. Новые источники данных в социальных исследованиях (большие, быстрые, глубокие и др. данные). Возможности их применения.
3. Методы машинного обучения и искусственный интеллект в социальных исследованиях.
4. Сетевой анализ.
5. Компьютерное моделирование социальных процессов.
6. Типичные ошибки количественных исследований и ограничения новых методов.
7. Этика исследований.

Примерный план практических занятий

1. Введение в Python
2. Использование веб-данных (на примере livejournal, ЦИАН и др.)
3. Использование API Вконтакте
4. Работа с сетевыми данными
5. Работа с текстовыми данными
6. Работа с геоданными

Итоговая оценка

Итоговая оценка складывается из оценки за выполнение практического задания (0-2 балла, где 0 – не выполнено, 2 – выполнено, 1 – выполнено не полностью или с ошибками) и оценки за итоговый проект.

Итоговый проект представляет собой проект исследования (именно проект, то есть реализовывать его нужно) объемом около 2 страниц. Проект должен включать в себя постановку вопроса или проблемы, то есть должно быть понятно, почему именно важно изучать то, что предлагается изучать. Проект также должен включать в себя обзор литературы, то есть должно быть понятно, что именно уже делалось в этом направлении, и описание методологии, то есть должно быть понятно, как и какие данные предлагается собирать, как их предполагается анализировать, и как полученные результаты могут быть интерпретированы. Проект оценивается по четырем критериям, за каждый из которых выставляет от 0 до 2 баллов:

Постановка вопроса/проблемы

Обзор литературы

Описание методологии

Соответствие предложенных методов исходному исследовательскому вопросу.

Баллы за практическое задание суммируются с баллами за проект, полученное число от 0 до 10 и является оценкой за курс.