

Проектирование и реализация трека «Машинное Обучение» в СПб НИУ ВШЭ

Алексей Александрович Шпильман

Заведующий центром анализа данных и машинного обучения НИУ ВШЭ в Санкт-Петербурге

Руководитель образовательной программы по машинному обучению, НИОЦ JetBrains



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ



Цели программы

- Основная цель – подготовка инженеров (**людей, которые могут делать**) в области анализа данных и машинного обучения, способных выполнять на высоком уровне прикладные задачи с минимальным риском.
- Провели консультацию с компаниями, заинтересованными в выпускниках программы:



Яндекс



Google



Теория оптимизации

Продвинутые техники
машинного обучения

Обработка естественного
языка

Распознавание и
генерация речи

Анализ
изображений

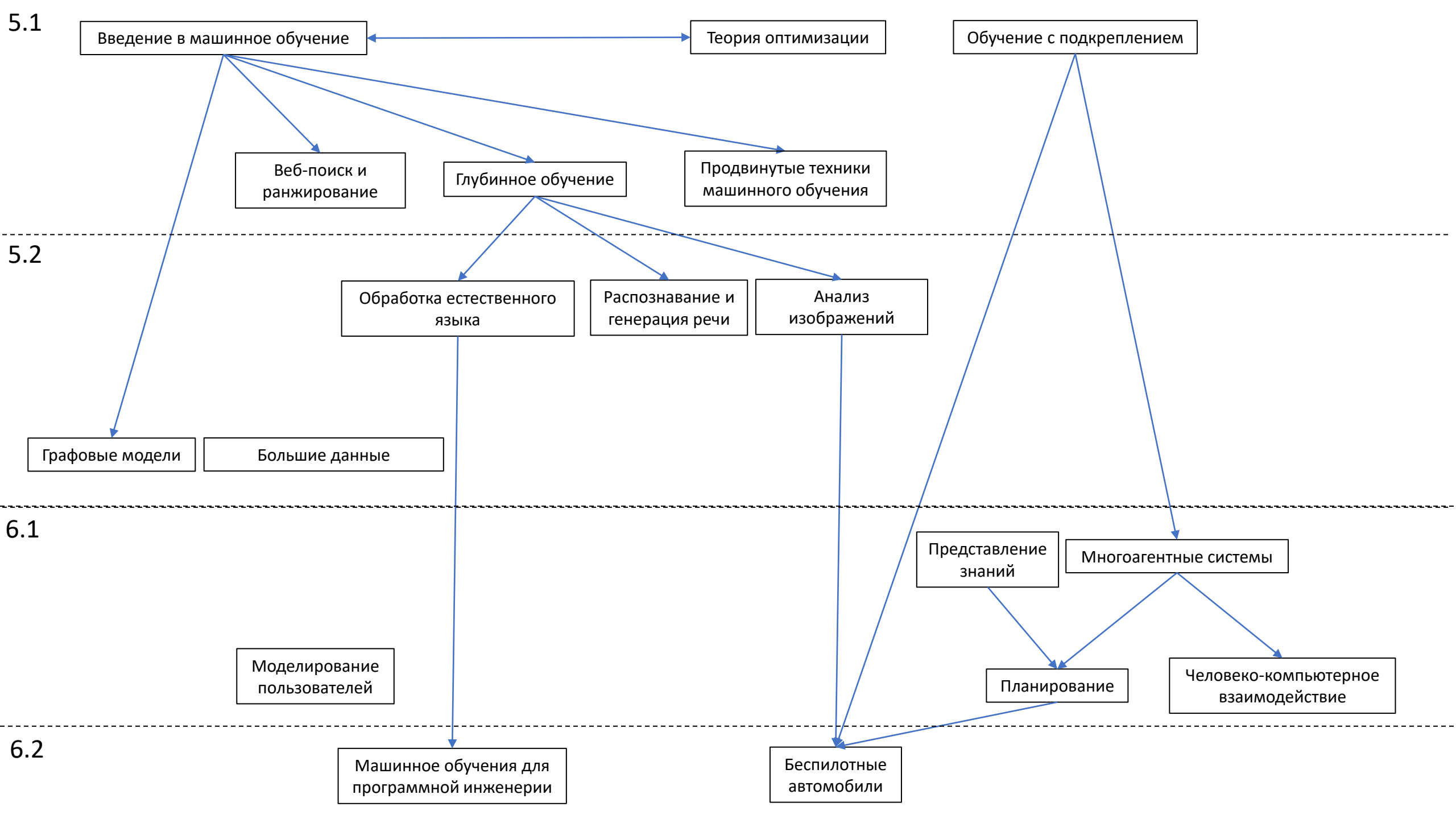
Графовые модели

Большие данные

Человеко-компьютерное
взаимодействие

Машинное обучения для
программной инженерии

Беспилотные
автомобили



Программы западных университетов

Course	Stanford	MIT	Caltech	Cambridge	Edinburgh				
Введение в машинное обучения (Introduction to Machine Learning)	Green	Green	Green	Green	Green		not covered		
Теория оптимизации (Optimization Theory)	Yellow	Yellow	Yellow	Red	Red				
Обучение с подкреплением (Reinforcement Learning)	Green	Green	Red	Green	Yellow		partially covered		
Глубинное обучение (Deep Learning)	Green	Yellow	Red	Yellow	Yellow				
Большие данные (Big Data)	Green	Yellow	Yellow	Red	Red		fully covered		
Продвинутое машинное обучение (Advanced Machine Learning)	Green	Green	Green	Green	Yellow				
Работа с естественным языком (Natural Language Processing)	Yellow	Green	Red	Green	Green				
Обработка изображений и звука (Image and Audio Processing)	Yellow	Green	Green	Green	Green				
Распознавание и генерация речи (Speech processing and generation)	Red	Yellow	Yellow	Yellow	Yellow				
Многоагентные системы (Multi-Agent Systems)	Yellow	Yellow	Red	Red	Yellow				
Веб-поиск и ранжирование (Web-search and Ranking)	Yellow	Red	Yellow	Red	Red				
Графовые модели (Probabilistic Graphical Models)	Green	Yellow	Red	Green	Yellow				
Симуляции и оптимизация моделей (Simulation and Model Optimization)	Red	Red	Red	Red	Red				
Моделирование пользователя (User Modeling)	Yellow	Red	Red	Red	Yellow				
Представление знаний (Knowledge Representation and Reasoning)	Yellow	Green	Yellow	Red	Green				
Вычислительная нейробиология (Computational Neuroscience)	Green	Yellow	Green	Red	Yellow				
Человеко-компьютерное взаимодействие (Human-Computer Interaction)	Green	Green	Yellow	Green	Green				
Планирование (Planning)	Yellow	Yellow	Red	Yellow	Green				
Машинное обучение в программной инженерии (Machine Learning in Software Engineering)	Red	Red	Red	Red	Red				
Беспилотные автомобили (Autonomous Vehicles)	Yellow	Yellow	Red	Red	Green				

Программы западных университетов

MIT

Artificial Intelligence	Green
Machine Vision	Green
Computational Cognitive Science	Yellow
Game Theory with Engineering Applications	Yellow
Automatic Speech Recognition	Green
Techniques in Artificial Intelligence	Green
Cognitive Robotics	Yellow
Natural Language and the Computer Representation of Knowledge	Green
Advanced Natural Language Processing	Green
Machine Learning	Green
The Society of Mind	Red
Knowledge-Based Applications Systems	Green
Medical Decision Support	Green
Computational Models of Discourse	Yellow

Составление курсов

- Требование к курсам от индустрии и науки.
- Пререквизиты от других курсов.
- «Полнота» курсов – если студент отчислится в середине, он все равно должен приобрести полезные знания.
- Обязательная практическая составляющая! Все, что рассказано на лекции, должно быть опробовано на практике (возможно в облегченной форме).

Поиск преподавателей

- Главный принцип – человек должен уметь делать то, про что он рассказывает.
- Преподаватели – почти на 100% сотрудники компаний IT-лидеров, имеющие опыт применения методов машинного обучения и анализа данных на практике.
- Мотивация для преподавателей – увеличивать количество полезных кадров для их компании/области. Привлекать студентов на исследовательские или промышленные проекты.

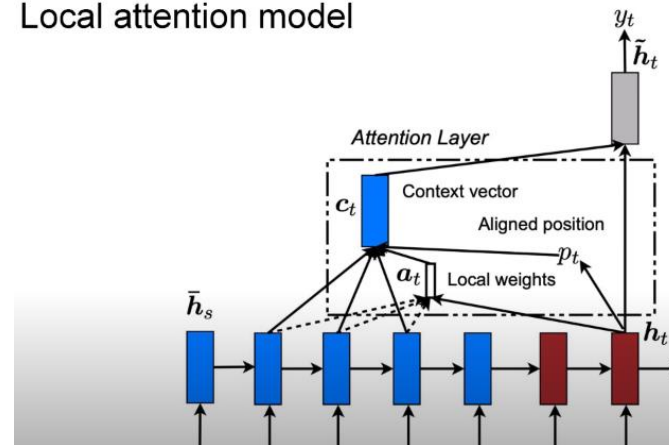
Обновление программы

- Программа постоянно обновляется, добавляются последние разработки, как в рамках курсов, так и полностью новые курсы.
- Пример: программа курса по глубокому обучению обновляется каждый год на 20-30%.
- Появились новые курсы: глубокое обучение с подкреплением, глубокое обучение без учителя. Реализуются бывшими выпускниками, ставшими специалистами в новых областях науки.
- Учитываются замечания и предложения студентов.
- Добавляются курсы от новых индустриальных партнеров.

Семинары для выступления студентов

- 5 семинаров на разные тематики машинного обучения
- Все студенты должны выступить хотя бы на одном из семинаров, сделать доклад по статье или набору статей по современным исследованиям по машинному обучению.
- В процессе студентам оказывается помощь от руководителя как по разбору статьи, так и по подготовке выступлений.
- Видео всех выступлений доступно в открытом доступе.

Local attention model



Научно-исследовательские работы

- Неотъемлимой частью программы является работа над научным или техническим проектом. Работа начинается со второго семестра обязательно, может переходить
- Темы для работ предоставляют компании-партнеры, исследовательские лаборатории. Несколько работ являются личной инициативой студентов.
- Часть работ (~20%) доводится до публикаций на международных конференциях. Стремимся к увеличению до 100%.