

Программа учебной дисциплины «Научный семинар»

Утверждена
Академическим советом ООП
Протокол № 2.03-09/2706-01 от «27» июня 2018г.

Автор	Масютин Алексей Александрович
Число кредитов	5
Контактная работа (час.)	40
Самостоятельная работа (час.)	150
Курс	2
Формат изучения дисциплины	full time

I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Целями освоения дисциплины «Научный семинар» являются овладение студентами основными концепциями и навыками для решения продвинутых задач прикладного анализа данных.

В результате освоения дисциплины студент должен:

знать:

- продвинутые разделы машинного обучения, такие как обучение с подкреплением и компьютерное зрение;
- подходы к обучению агентских систем;
- приемы работы с изображением, видео в целях сегментации и детектирования объектов;
- современные методы и архитектуры моделей для работы с видео и изображениями.

уметь:

- обучать агентские системы;
- строить модели сегментации произвольных изображений;

владеть:

- навыками использования инструментов для построения моделей сегментации изображений;
- навыками обработки неструктурированных данных;
- навыками постановки и решения продвинутых задач машинного обучения.

Изучение дисциплины «Научный семинар» базируется на следующих дисциплинах:

- Машинное обучение;
- Современные методы анализа данных: Глубинное обучение.
- Анализ текстов. Генеративные модели

Для освоения учебной дисциплины студенты должны владеть следующими знаниями и компетенциями:

- программирование на Python;
- построение моделей машинного обучения;
- валидация предиктивных моделей.

Основные положения дисциплины должны быть использованы в дальнейшем при подготовке выпускной квалификационной работы (магистерской диссертации).

II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Тема 1. Введение в обучение с подкреплением

Понятие агента и окружающей среды. Состояния системы. Обратная связь, понятие положительного и отрицательного вознаграждения. Типы систем подкрепления.

Тема 2. Многорукые бандиты.

Метрика потерь по сравнению с оптимальной стратегией. Алгоритм линейного вознаграждения–бездействия. Правило инкрементального обновления.

Тема 3. Марковский процесс принятия решений.

Value-based подходы. Задача принятия решений для полностью наблюдаемой среды с марковской моделью перехода. Модель конечного и бесконечного горизонта. Модель среднего вознаграждения. Метод Монте-Карло.

Тема 4. Приближенное обучение с подкреплением.

Model-based методы, Model-free методы.

Тема 5. Policy Gradients I,II.

Методы, основанные на градиенте. Уравнения Беллмана. Улучшение стратегий. Архитектура Дуна. Набор инструментов Gym OpenAI.

Тема 6. Компьютерное зрение. Дополнительные главы.

Сегментация изображений. Практическое упражнение: сегментация изображений МРТ для замера частей сердца с помощью таких инструментов, как TensorBoard и TensorFlow Python API.

Тема 7. Захват картинки и видео.

Сочетание компьютерного зрения и обработки естественного языка для целей описания сцены. Обучение модели, генерирующей описание изображения из необработанных данных о пикселях, комбинируя выходные данные нескольких сетей (CNNs и RNNs) путем конкатенации и/или усреднения.

Тема 8. Использование GPU ресурсов

Библиотеки с открытым исходным кодом для ускорения подготовки данных и машинного обучения на GPU. Работа с продуктом RAPIDS. Обзор поддерживаемых API. Установка и запуск.

III. ОЦЕНИВАНИЕ

Оценка за «Научный семинар» складывается из двух заданий по продвинутому прикладному анализу данных. Первое представляет домашнее задание по теме обучения с подкреплением, второе – подготовка лабораторного проекта по разработке модели сегмента-

ции и генерации описания сцен с помощью библиотек с открытым кодом TensorFlow, а также RAPIDS.

Итоговая оценка складывается как среднее арифметическое из двух оценок.

$$O_{\text{итог}} = 0,5 * O_{\text{дз}} + 0,5 * O_{\text{лабораторная}}$$

Домашнее задание и лабораторная выполняется индивидуально.

IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

Оценочные средства для текущего контроля студента:

Домашнее задание по разделу Обучение с подкреплением.

Оценочные средства для промежуточной аттестации:

Защита проекта по построению и валидации моделей компьютерного зрения. Темы: сегментация изображений, генерация описания сцены.

V. РЕСУРСЫ

5.1 Основная литература

1. Саттон Р.С., Барто Э.Г. Обучение с подкреплением, Москва, «Бином. Лаборатория знаний», 2017
2. Solem J.E. Programming Computer Vision with Python: Tools and Algorithms for Analyzing Images, O'Reilly, 2012.

5.2 Дополнительная литература

1. Kapur S., Computer Vision with Python 3. Packt Publishing Ltd, 2017.

5.3 Программное обеспечение

№ п/п	Наименование	Условия доступа
1.	Облачные ресурсы Microsoft Azure	<i>Внутренняя подписка Сбербанка Azure Cloud</i>
2.	Python 2x, 3x	<i>Из внутренней сети университета</i>

5.4 Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

№ п/п	Наименование	Условия доступа
	<i>Профессиональные базы данных, информационно-справочные системы</i>	

1.	База статей Arxiv	URL: https://arxiv.org/
<i>Интернет-ресурсы (электронные образовательные ресурсы)</i>		
1.	Открытые лекции MIT	URL: https://ocw.mit.edu/courses/#electrical-engineering-and-computer-science

5.5 Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.
- в процессе занятий используется подписка Microsoft Azure, предоставляемая Сбербанком.

Учебные аудитории для лабораторных и самостоятельных занятий по дисциплине оснащены компьютерами, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.