

Саратовский филиал ПривГУПС

МОДУЛЬ «СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА»

Системы искусственного интеллекта рабочая программа дисциплины (модуля)

Специальность 23.05.04 Эксплуатация железных дорог
Направленность (профиль) Магистральный транспорт
Квалификация **инженер путей сообщения**

Форма обучения **заочная**

Общая трудоемкость **3 ЗЕТ**

Виды контроля на курсах:
зачеты 5

Распределение часов дисциплины по курсам

Курс	5		Итого	
	УП	РП		
Лекции	4	4	4	4
Лабораторные	4	4	4	4
Практические	4	4	4	4
Конт. ч. на аттест. в период ЭС	0,15	0,15	0,15	0,15
Итого ауд.	12	12	12	12
Контактная работа	12,15	12,15	12,15	12,15
Сам. работа	92	92	92	92
Часы на контроль	3,85	3,85	3,85	3,85
Итого	108	108	108	108

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
1.1	овладение студентами основными методами методов машинного обучения, получение навыков программирования алгоритмов в области искусственного интеллекта и анализа полученных результатов.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	
Цикл (раздел) ОП:	Б1.О.32.01

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
ОПК-10 Способен формулировать и решать научно-технические задачи в области своей профессиональной деятельности	
ОПК-10.3 Использует методы искусственного интеллекта (машинного обучения) и анализа больших данных для решения прикладных задач	

В результате освоения дисциплины (модуля) обучающийся должен

3.1	Знать:
3.1.1	основные методы машинного обучения; классификацию задач машинного обучения; метрики качества модели; функции и методы библиотеки Pandas: query, tail, sample, head; методы фильтрации и агрегации данных; признаки переобученности модели; методы библиотек sklearn, seaborn, matplotlib и numpy.
3.2	Уметь:
3.2.1	агрегировать данные средствами Python; выполнять фильтрацию данных средствами Python; настраивать вес нейронна; выбирать тип классификатора в зависимости от поставленной задачи; визуализировать данные средствами библиотек pandas, seaborn, matplotlib;
3.2.2	выполнять многоклассовую классификацию методами библиотеки sklearn;
3.2.3	строить деревья решений и выполнять по ним предсказание.
3.3	Владеть:
3.3.1	навыками построения дерева решений;
3.3.2	навыками реализации алгоритма случайного леса (Random forest);
3.3.3	навыками написания нейронных сетей;
3.3.4	навыками обучения модели.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Примечание
	Раздел 1. Основы анализа данных			
1.1	Основные понятия анализа данных. Искусственный интеллект. Машинное обучение. Глубокое обучение. Обучение с учителем и без учителя. /Ср/	5	2	
1.2	Данные, модель, обработка данных. Методологические принципы анализа данных. Цели, этапы, методы и техники. Фильтрация группировка и агрегация данных. /Лек/	5	2	
1.3	Визуализация данных. Визуализация данных в Python. Визуализация в Pandas. Библиотеки Seaborn и Matplotlib. /Ср/	5	2	
1.4	Алгоритмы и структуры данных. Библиотека Pandas: структуры данных: Series и DataFrame. Работа с csv. /Пр/	5	2	
1.5	Фильтрация данных. Библиотека Pandas. Вывод строк и колонок. Методы query, tail, sample, head. Объединение методов. /Ср/	5	4	
1.6	Агрегация данных в Pandas: count, sum, max, min, mean, median. Методы groupby и aggregate. Объединение методов. /Ср/	5	2	
1.7	Агрегация данных в Pandas: count, sum, max, min, mean, median. Методы groupby и aggregate. Объединение методов. /Лаб/	5	2	
1.8	Визуализация данных в Python. Методы визуализации в Pandas. Библиотеки Seaborn и Matplotlib. Графики, тепловые карты, диаграммы рассеивания. Анализ и интерпретация результатов визуализации. /Ср/	5	2	
	Раздел 2. Методы машинного обучения			
2.1	Решающие деревья: основные понятия. Классификация. Деревья решений. Основные параметры дерева. Дерево решений в задачи регрессии. Метод ближайших соседей. /Лек/	5	2	
2.2	Концепция случайного леса. Случайная выборка тренировочных образцов. Усреднение прогнозов. Проблема переобучения. Метрики качества модели. Тестирование ROC AUC. /Ср/	5	4	

2.3	Нейронные сети. Искусственные нейроны. Перцептрон. Обучение нейрона. /Ср/	5	4	
2.4	Градиентный спуск. Дробление шага при градиентном спуске. Стохастический градиентный спуск. Метод наискорейшего спуска. /Ср/	5	4	
2.5	Нейронная сеть. Однослойная модель. Сверточные нейронные сети. Свёрточный слой, слой подвыборки, полносвязный слой. Целевая функция. /Ср/	5	4	
2.6	Решающие деревья: обучение модели. Библиотеки sklearn, pandas, numpy. Выбор параметров модели. Кросс-валидация. Предварительная обработка данных. Тренировочный набор данных. Тестовый набор данных. /Лпр/	5	2	
2.7	Библиотеки sklearn, pandas, numpy. Выбор параметров модели. Кросс-валидация. Предварительная обработка данных. Тренировочный набор данных. Тестовый набор данных. /Лаб/	5	2	
2.8	Концепция случайного леса. Случайная выборка тренировочных образцов. Усреднение прогнозов. Проблема переобучения. Типы ошибок: true negative, false positive. Метрики качества: Precision, Recall, F1 score. Тестирование ROC AUC. /Ср/	5	4	
2.9	Распознавание рукописных цифр. Алгоритмы случайного леса. Оптимальные параметры дерева и леса. /Ср/	5	4	
2.10	Искусственный нейрон. Вход, функция активации (сигмоида, гиперболический тангенс, SoftMax). Подбор весов. /Ср/	5	4	
2.11	Определение тональности текста с помощью нейронных сетей. Библиотеки tensorflow, Keras, набор данных IMDB. /Ср/	5	4	
2.12	Библиотека Keras. Полносвязный и сглаживающий слои, слой Dropout. Решение проблем переобучения. /Ср/	5	4	
2.13	Классификация изображений по нескольким классам (более двух). Набор данных Fashion Mnist. /Ср/	5	4	
2.14	Внешние источники данных для анализа. Kuggle. Подготовка данных. Работа с изображениями разного формата. /Ср/	5	2	
2.15	Распознавание образов. Классификация изображений кошек и собак. Библиотека keras, ImageDataGenerator. Достижение заданной точности модели на валидационной выборке. /Ср/	5	4	
	Раздел 3. Самостоятельная работа			
3.1	Алгоритм обратного распространения ошибки. Функции активации. Оценка работы сети. /Ср/	5	4	
3.2	Компьютерное зрение. Библиотека OpenCV. Сегментация изображений. Детектирование объектов. /Ср/	5	4	
3.3	Компьютерное зрение. Библиотека OpenCV. Отслеживание движущихся объектов во времени. Распознавание лиц. /Ср/	5	4	
3.4	Обработка естественного языка. Основные понятия. Токенизация. Лемматизация. /Ср/	5	4	
3.5	Обработка естественного языка. Парсинг зависимостей. Распознавание именованных сущностей /Ср/	5	4	
3.6	Написание собственных классов для описания нейронной сети. Нейрон. Слой. Сеть. Обучение модели. Выполнение предсказания. /Ср/	5	4	
3.7	Минимизация значения функции потерь. Обучение: стохастический градиентный спуск. /Ср/	5	4	
3.8	Подготовка к лекциям /Ср/	5	2	
3.9	Подготовка к лабораторным занятиям /Ср/	5	4	
	Раздел 4. Контактные часы на аттестацию			
4.1	Зачёт /КЭ/	5	0,15	
5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ				

Оценочные материалы для проведения промежуточной аттестации обучающихся приведены в приложении к рабочей программе дисциплины.				
Формы и виды текущего контроля по дисциплине (модулю), виды заданий, критерии их оценивания, распределение баллов по видам текущего контроля разрабатываются преподавателем дисциплины с учетом ее специфики и доводятся до сведения обучающихся на первом учебном занятии.				
Текущий контроль успеваемости осуществляется преподавателем дисциплины (модуля), как правило, путем проверки письменных работ, предусмотренных рабочими программами дисциплин в рамках контактной работы и самостоятельной работы обучающихся.				
6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
6.1. Рекомендуемая литература				
6.1.1. Основная литература				
	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л1.1	Никольский С. Н.	Автоматизация информационного поведения и искусственный интеллект: Учебное пособие	Москва: МИРЭА, 2020	https://e.lanbook.com/bo
6.1.2. Дополнительная литература				
	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л2.1	Сидоркина И.Г.	Системы искусственного интеллекта.	Москва: КноРус, 2020	http://www.book.ru/boo
6.2 Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине (модулю)				
6.2.1 Перечень лицензионного и свободно распространяемого программного обеспечения				
6.2.1.1	Ubuntu			
6.2.2 Перечень профессиональных баз данных и информационных справочных систем				
6.2.2.1	Информационная справочная система Техэксперт https://tech.company-dis.ru			
6.2.2.2	Информационная справочная система «Гарант» http://www.garant.ru			
6.2.2.3	База данных Государственных стандартов http://gostexpert.ru/			
6.2.2.4	База данных «Железнодорожные перевозки» https://cargo-report.info/			
6.2.2.5	База данных для конкурсов среди специалистов по обработке данных и машинному обучению kaggle.com/			
7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
7.1	Учебные аудитории для проведения занятий лекционного типа, укомплектованные специализированной мебелью и техническими средствами обучения: мультимедийное оборудование для предоставления учебной информации большой аудитории и/или звукоусиливающее оборудование (стационарное или переносное).			
7.2	Учебные аудитории для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения: мультимедийное оборудование и/или звукоусиливающее оборудование (стационарное или переносное)			
7.3	Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.			
7.4	Помещения для хранения и профилактического обслуживания учебного оборудования.			