

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Гаранин Максим Александрович
Должность: Ректор
Дата подписания: 21.10.2025 15:09:13
Уникальный программный ключ:
7708e3a47e66a8ee02711b298d7c78bd1e40bf88

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ПРИВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПУТЕЙ СООБЩЕНИЯ»

МОДУЛЬ "СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА"

Практикум по машинному обучению рабочая программа дисциплины (модуля)

Направление подготовки 13.03.02 Электроэнергетика и электротехника

Направленность (профиль) Электрический транспорт

Квалификация **бакалавр**

Форма обучения **очная**

Общая трудоемкость **2 ЗЕТ**

Виды контроля в семестрах:

зачеты 6

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	6 (3.2)		Итого	
Неделя	16			
Вид занятий	УП	РП	УП	РП
Лабораторные	32	32	32	32
Конт. ч. на аттест. в период ЭС	0,15	0,15	0,15	0,15
Итого ауд.	32	32	32	32
Контактная работа	32,15	32,15	32,15	32,15
Сам. работа	31	31	31	31
Часы на контроль	8,85	8,85	8,85	8,85
Итого	72	72	72	72

Программу составил(и):

к.п.н., доцент, Тюжина И.В.

Рабочая программа дисциплины

Практикум по машинному обучению

разработана в соответствии с ФГОС ВО:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 13.03.02 Электроэнергетика и электротехника (приказ Минобрнауки России от 28.02.2018 г. № 144)

составлена на основании учебного плана: 13.03.02-25-2-ЭЭб.plm.plx

Направление подготовки 13.03.02 Электроэнергетика и электротехника Направленность (профиль) Электрический транспорт

Рабочая программа одобрена на заседании кафедры

Цифровые технологии

Зав. кафедрой к.э.н., Ефимова Т.Б.

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)				
1.1	Применение на практике методов машинного обучения при решении задач профессиональной деятельности.			
2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ				
Цикл (раздел) ОП:	Б1.О.23.02			
3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)				
ОПК-1 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности				
ОПК-1.1 Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности				
ПК-2 Способен применять математические методы сбора, систематизации, обобщения и обработки информации для обеспечения требуемого технического состояния подвижного состава электрического транспорта, подстанций, кабельных и воздушных линий электропередачи				
ПК-2.8 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 классы научно-технических задач в области своей профессиональной деятельности, решаемые методами машинного обучения;			
3.1.2	- основные классы интеллектуальных информационных систем, необходимые в процессе решения научно-технических задач в сфере электрического транспорта;			
3.2	Уметь:			
3.2.1	- выполнять агрегирование, фильтрацию и сортировку данных большого объема;			
3.2.2	- использовать методы искусственного интеллекта для решения прикладных задач;			
3.3	Владеть:			
3.3.1	- навыками визуализации данных большого объема;			
3.3.2	- решения научно-технических задач в области своей профессиональной деятельности с использованием технологий анализа данных и машинного обучения.			
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Примечание
	Раздел 1. Практикум по машинному обучению			
1.1	Подготовка и фильтрация данных. Распространенные форматы данных: csv, json. Удаление поврежденных значений. Выбор данных по заданному признаку и группе признаков. /Лаб/	6	2	
1.2	Преобразование данных Формирование нового столбца данных на основе существующих. Формирование вспомогательного столбца данных. Изменение типа данных столбца. /Лаб/	6	4	
1.3	Агрегация данных Группировка данных по признаку. Агрегация данных: подсчет среднего, общего, максимального и минимального значений по определенному признаку. Выбор оптимального метода. /Лаб/	6	4	
1.4	Определение закономерностей в опозданиях железнодорожного транспорта. Графическая обработка данных. Построение графиков зависимости. Построение графиков распределения точек, определение числа кластеров, формируемых наблюдениями. Построение тепловой карты. Различные виды графиков (seaborn, matplotlib, встроенная визуализация pandas). /Лаб/	6	4	
1.5	Определение неисправности железнодорожного полотна. Обучение модели. Подготовка данных. Разделение данных на тренировочный, тестовый и валидационные датасеты. Построение нейронной сети. /Лаб/	6	4	

1.6	Исследование данных о дорожных происшествиях и серьёзности аварий. Дерево решений. Построение матрицы ошибок (confusion matrix) Подсчет метрик качества (Precision, Recall, F1 score). /Лаб/	6	4	
1.7	Исследование данных о дорожных происшествиях и серьёзности аварий. Случайный лес Выбор оптимальных параметров дерева решений. Создание модели с указанным числом деревьев и максимальной глубины. Предсказание класса по тренировочной выборке. Отбор топ-5 самых важных переменных для классификации /Лаб/	6	4	
1.8	Определение тональности текста. Нейронные сети. Определение тональности текста с помощью нейронных сетей. Библиотеки tensorflow, Keras, набор данных IMDb. /Лаб/	6	6	
	Раздел 2. Самостоятельная работа			
2.1	Подготовка и фильтрация данных. Регистрация на портале kaggle. Распространенные форматы данных: csv, json. Подготовка к лабораторным работам. /Ср/	6	2	
2.2	Преобразование данных. Методы работы структурами данных pandas. Операции с pandas. series. Повторение материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. /Ср/	6	4	
2.3	Агрегация данных. Группировка. Функции агрегирования в pandas. Повторение материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. /Ср/	6	4	
2.4	Определение закономерностей в опозданиях железнодорожного транспорта. Графическая обработка данных. Методы визуализации pandas. Библиотеки matplotlib и seaborn. Параметры методов. Повторение материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. ЭИОС ПривГУПС /Ср/	6	4	
2.5	Определение неисправности железнодорожного полотна. Обучение модели. Работа с изображениями в python. Построение нейронной сети. Виды слоев. Повторение теоретического материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. /Ср/	6	4	
2.6	Определение закономерностей в опозданиях железнодорожного транспорта. Графическая обработка данных. Методы визуализации pandas. Библиотеки matplotlib и seaborn. Параметры методов. Повторение материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. /Ср/	6	4	
2.7	Исследование данных о дорожных происшествиях и серьёзности аварий. Случайный лес. Основные параметры случайного леса. Повторение теоретического материала, изученного в курсе «Технологии искусственного интеллекта». Подготовка к лабораторным работам. /Ср/	6	4	
2.8	Определение тональности текста. Нейронные сети Методы библиотеки tensorflow. Подготовка к лабораторным работам. /Ср/	6	2	
2.9	Набор данных IMDb. Исследование данных. Библиотека keras. Изучение синтаксиса методов библиотеки. Подготовка к лабораторной работе. /Ср/	6	3	
	Раздел 3. Контактные часы на аттестацию			
3.1	Зачёт /КЭ/	6	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) в рамках контактной работы и самостоятельной работы обучающихся. Для фиксирования результатов текущего контроля может использоваться ЭИОС.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательс тво, год	Эл. адрес
Л1.1	Никольский С. Н.	Автоматизация информационного поведения и искусственный интеллект: Учебное пособие	Москва: МИРЭА, 2020	https://e.lanbook.com/book/163824

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательс тво, год	Эл. адрес
Л2.1	Платонов А. В.	Машинное обучение: учебное пособие для вузов	Москва: Юрайт, 2024	https://urait.ru/bcode/544780

6.2 Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине (модулю)

6.2.1 Перечень лицензионного и свободно распространяемого программного обеспечения

6.2.2 Перечень профессиональных баз данных и информационных справочных систем

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.
7.2	Помещения для хранения и профилактического обслуживания учебного оборудования.
7.3	Лаборатории, оснащенные специальным лабораторным оборудованием: компьютерной техникой с установленным ПО: Python, а также с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ
ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

ПРАКТИКУМ ПО МАШИННОМУ ОБУЧЕНИЮ

(наименование дисциплины(модуля))

Направление подготовки / специальность

13.03.02 Электроэнергетика и электротехника

(код и наименование)

Направленность (профиль)/специализация

Электрический транспорт

(наименование)

Содержание

1. Пояснительная записка.
2. Типовые контрольные задания или иные материалы для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих уровень сформированности компетенций.
3. Методические материалы, определяющие процедуру и критерии оценивания сформированности компетенций при проведении промежуточной аттестации.

1. Пояснительная записка

Цель промежуточной аттестации – оценивание промежуточных и окончательных результатов обучения по дисциплине, обеспечивающих достижение планируемых результатов освоения образовательной программы.

Формы промежуточной аттестации: Зачёт, 6 семестр

Перечень компетенций, формируемых в процессе освоения дисциплины

Код и наименование компетенции	Код индикатора достижения компетенции
ОПК-1: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-1.1: Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности
ПК-2: Способен применять математические методы сбора, систематизации, обобщения и обработки информации для обеспечения требуемого технического состояния подвижного состава электрического транспорта, подстанций, кабельных и воздушных линий электропередачи	ПК-2.8: Использует методы искусственного интеллекта (машинного обучения) и анализа больших данных для решения прикладных задач

Результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине	Оценочные материалы
ОПК-1.1: Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности	Обучающийся знает : - основные классы научно-технических задач в области своей профессиональной деятельности, решаемые методами машинного обучения;	Вопросы (№1 - №10)
	Обучающийся умеет : - выполнять агрегирование, фильтрацию и сортировку данных большого объёма;	Задания (№1 - №15)
	Обучающийся владеет : - навыками визуализации данных большого объёма;	Задания (№ 31- №32)
ПК-2.8: Использует методы искусственного интеллекта (машинного обучения) и анализа больших данных для решения прикладных задач	Обучающийся знает : - основные классы интеллектуальных информационных систем, необходимые в процессе решения научно-технических задач в сфере электрического транспорта;	Вопросы (№11 - №15)
	Обучающийся умеет : - использовать методы искусственного интеллекта для решения прикладных задач;	Задания (№1 - №30)
	Обучающийся владеет : - навыками решения научно-технических задач в области своей профессиональной деятельности с использованием технологий анализа данных и машинного обучения.	Задания (№ 34- №35)

Промежуточная аттестация (зачет) проводится в одной из следующих форм:

- 1) собеседование;
- 2) выполнение заданий в ЭИОС университета.

2. Типовые¹ контрольные задания или иные материалы для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих уровень сформированности компетенций

2.1. Типовые вопросы (тестовые задания) для оценки знаниевого образовательного результата

Проверяемый образовательный результат:

Код и наименование индикатора достижения компетенции	Образовательный результат
ОПК-1.1: Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности	Обучающийся знает: - основные классы научно-технических задач в области своей профессиональной деятельности, решаемые методами машинного обучения;
1. Какой метод машинного обучения лучше всего подходит для прогнозирования потребления электроэнергии на зарядных станциях для электрических автомобилей? А) Линейная регрессия В) Случайный лес С) Нейронные сети D) Градиентный бустинг 2. Какой алгоритм машинного обучения наиболее эффективен для оптимизации работы электрической сети с учетом колебаний в потреблении энергии? А) Кластеризация данных В) Генетические алгоритмы С) Рекуррентные нейронные сети D) Метод опорных векторов 3. Какой метод машинного обучения можно использовать для диагностики и прогнозирования неисправностей в электрических транспортных средствах? А) Деревья решений В) Кластерный анализ С) Сверточные нейронные сети D) Логистическая регрессия 4. Какой алгоритм машинного обучения можно применить для оптимизации зарядки электрических транспортных средств с учетом времени суток и стоимости электроэнергии? А) Генетический алгоритм В) Рекурсивное дерево решений С) Градиентный спуск D) Случайный лес 5. Какой метод машинного обучения наиболее эффективен для предсказания энергопотребления в зданиях с учетом погодных условий и количества людей? А) Рекуррентные нейронные сети В) Метод k-ближайших соседей С) Градиентный бустинг D) Линейная регрессия 6. Какой алгоритм машинного обучения можно использовать для оптимизации работы системы управления энергопотреблением в больших промышленных предприятиях? А) Сверточные нейронные сети В) Генетический алгоритм С) Деревья решений D) Метод опорных векторов 7. Какой метод машинного обучения наиболее подходит для предсказания возможных аварийных ситуаций в электрической сети на основе исторических данных? А) Логистическая регрессия В) Случайный лес С) Рекуррентные нейронные сети	

¹ Приводятся типовые вопросы и задания. Оценочные средства, предназначенные для проведения аттестационного мероприятия, хранятся на кафедре в достаточном для проведения оценочных процедур количестве вариантов. Оценочные средства подлежат актуализации с учетом развития науки, образования, культуры, экономики, техники, технологий и социальной сферы. Ответственность за нераспространение содержания оценочных средств среди обучающихся университета несут заведующий кафедрой и преподаватель – разработчик оценочных средств.

D) Градиентный бустинг 8. Какой алгоритм машинного обучения можно использовать для оптимизации распределения электроэнергии на зарядных станциях в городе с учетом пикового потребления? A) Кластерный анализ B) Генетический алгоритм C) Сверточные нейронные сети D) Деревья решений 9. Какой метод машинного обучения наиболее эффективен для прогнозирования изменений в потреблении электроэнергии в зависимости от дня недели и времени суток? A) Рекуррентные нейронные сети B) Линейная регрессия C) Градиентный бустинг D) Метод k-ближайших соседей 10. Какой алгоритм машинного обучения можно использовать для оптимизации работы системы управления зарядными станциями на основе данных о прогнозах погоды и количестве заряжаемых транспортных средств? A) Генетический алгоритм B) Случайный лес C) Рекурсивное дерево решений D) Сверточные нейронные сети	
ПК-2.8: Использует методы искусственного интеллекта (машинного обучения) и анализа больших данных для решения прикладных задач	Обучающийся знает: - основные классы интеллектуальных информационных систем, необходимые в процессе решения научно-технических задач в области в сфере электрического транспорта;
11. Это включает в себя экосистема «умных дорог»? Типовой ответ: детекторы транспортного потока; адаптивные (умные) светофоры; средства автоматической фиксации нарушений ПДД; электронные средства безостановочной оплаты проезда; паркоматы; подключенные информационные табло; системы автоматизированного управления освещением; 12. Какие элементы включает в себя Интеллектуальная транспортная система? Типовой ответ: транспортные средства, участники системы, а также дорожно-транспортное регулирование. 13. Каково основное назначение интеллектуальной транспортной системы? Типовой ответ: обеспечение безопасности дорожного движения и оптимизация транспортных потоков. 14. Назовите основные препятствия для внедрения интеллектуальных транспортных систем на междугородних трассах? Типовой ответ: из-за больших расстояний строительство инфраструктуры (электропитание, канал передачи данных, опоры и технические помещения) нецелесообразно с экономической точки зрения, кроме того, оборудование нужно обслуживать и обеспечивать его сохранность. 15. В чем отличие между нейронными сетями и генетическими алгоритмами в контексте интеллектуальных информационных систем? Типовой ответ: Нейронные сети и генетические алгоритмы являются различными методами машинного обучения. Нейронные сети моделируют работу человеческого мозга и используются для обработки данных, распознавания образов, прогнозирования и других задач. Генетические алгоритмы основаны на принципах естественного отбора и эволюции, используя генетические операторы для поиска оптимальных решений задач.	

2.2. Типовые задания для оценки навыкового образовательного результата

Проверяемый образовательный результат:

Код и наименование индикатора достижения компетенции	Образовательный результат
ОПК-1.1: Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности	Обучающийся умеет: - выполнять агрегирование, фильтрацию и сортировку данных большого объема;
ПК-2.8: Использует методы	- использовать методы искусственного интеллекта для решения прикладных задач;

Приложение 2

Задание №1

Для ответа на этот вопрос вам понадобится набор данных imdb из библиотеки keras from keras.datasets import imdb

Определите, какое слово стоит на 1000 месте (имеет индекс 1000) в частотном словаре датасета IMDb?

Запишите ответ:

1)

Ответ:

Задание №2

Для ответа на этот вопрос вам понадобится датасет imdb из библиотеки keras from keras.datasets import imdb

Определите, какое слово стоит на 3355 месте (имеет индекс 3355) в частотном словаре датасета IMDb?

Запишите ответ:

1)

Ответ:

Задание №3

Для ответа на этот вопрос вам понадобится датасет imdb из библиотеки keras from keras.datasets import imdb

Определите на каком месте в списке самых часто встречающихся слов в наборе данных IMDb стоит слово "joker" (впишите полученный индекс)?

Запишите ответ:

1)

Ответ:

Задание №4

Для ответа на этот вопрос вам понадобится датасет imdb из библиотеки keras
from keras.datasets import imdb

Определите на каком месте в списке самых часто встречающихся слов в наборе данных IMDb стоит слово "study" (впишите полученный индекс)?

Запишите ответ:

1)

Ответ:

Задание №5

В какой стране значение соотношения между количеством вакцинированных и общей численностью населения на текущую дату в стране максимально? Поле Daily vaccinations per million.

Источник данных - country_vaccinations.csv

Запишите ответ:

1)

Ответ:

Задание №6

В какой строке (укажите индекс) находится максимальное значение по полю daily_vaccinations_per_million для стран Италия, Испания и Португалия.

Ответ - одно число, индекс строки.

Источник данных - country_vaccinations.csv

Запишите число:

1)

Ответ:

Задание №7

В какой строке указаны максимальные показатели по daily_vaccinations_per_million для Аргентины, России и Уэlsa.

Укажите номера строк через пробел. Например: 45 678 2223

Источник данных - country_vaccinations.csv

Запишите ответ:

1)

Ответ:

Задание №8

Укажите название библиотеки позволяющей строить диаграммы рассеивания, графики и гистограммы в Python

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №9

Имеется набор данных о работе датчиков в грузовых вагонах поезда. Необходимо по датчикам автоматически разделять вагоны на полностью заполненные и пустые.

Напишите команду, создающую дерево решений dt для этого набора.

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №10

Какое значение имеет параметр `daily_vaccinations` для строки с индексом 3567?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №11

Какое значение имеет параметр `daily_vaccinations` строки датасета `country_vaccinations` с индексом 3367?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №12

Какое значение имеет параметр `daily_vaccinations` строки с индексом 25?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)

Ответ:

Задание №13

Какое значение имеет параметр `people_fully_vaccinated_per_hundred` для строки датасета с индексом 4036?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)

Ответ:

Задание №14

Какое значение имеет параметр `people_vaccinated` для последней строки датасета `country_vaccinations`?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)

Ответ:

Задание №15

Какое значение имеет параметр `total_vaccinations_per_hundred` для строки датасета с индексом 4036?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)

Ответ:

Задание №16

Работа с файлом `country_vaccinations.csv`

Какое число указано в строке с индексом "3000" в столбце `daily_vaccinations_per_million`?

Источник данных - `country_vaccinations.csv`

Запишите ответ:

1)

Ответ:

Задание №17

При обучении дерева решений на обучающей выборке, получена 100% точность классификатора, однако на тестовых классификатор показал точность около 50%. Что может быть причиной такого результата?

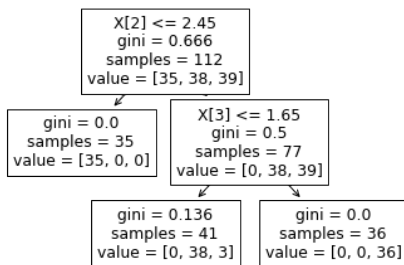
Запишите ответ:

1)

Ответ:

Задание №18

Перед вами дерево решений. Укажите самый высокий коэффициент загрязнения в листе (не во внутреннем узле)



Запишите ответ:

1)

Ответ:

Задание №19

Отберите данные за 5 февраля 2021 года. В какой стране количество привитых за день было максимально?

Источник данных - country_vaccinations.csv

Запишите ответ:

1)

Ответ:

Задание №20

Отсортируйте данные о вакцинации в Испании по полю daily_vaccinations. Используйте сортировку по убыванию. Укажите значение поля daily_vaccinations в третьей строке.

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №21

Посчитайте в скольких наблюдениях в качестве вакцины использовался "Sputnik V" и "Pfizer".

Укажите оба числа через пробел. Сначала меньшее, потом большее.

Чтобы найти страны, в которых применялась вакцина, можно воспользоваться методом `Series.str.contains`

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №22

Для ответа на этот вопрос вам понадобится датасет imdb из библиотеки keras

```
from keras.datasets import imdb
```

Определите какое из слов lucky или drama чаще встречается в датасете IMDb? УВ качестве ответа укажите абсолютную (по модулю) разницу между позициями этих слов в частотном словаре?

Запишите число:

1)	Ответ:	
----	--------	--

Задание №23

Сгруппируйте данные по дате наблюдения. Посчитайте сумму по полю daily_vaccinations. Укажите второе по величине значение.

Ответ - сумма, а не дата.

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №24

Сгруппируйте данные по странам, для каждой страны вычислите сумму по полю "daily_vaccinations", отсортируйте полученные данные по убыванию. Укажите значение в третьей строке.

Ответ - одно число.

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №25

Сгруппируйте данные по странам. Укажите страну, которая находится на втором месте по максимальному числу полностью вакцинированных (people_fully_vaccinated).

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №26

Сколько людей было вакцинировано в России 10го февраля? Поле daily_vaccinations.

Источник данных - country_vaccinations.csv

Запишите ответ:

1)	Ответ:	
----	--------	--

Задание №27

Создайте случайные лес rf состоящий из15 деревьев. В ответе укажите строчку кода, отвечающую за его создание.

Запишите число:

1)	Ответ:	
----	--------	--

Задание №28

Имеется матрица ошибок:

Прогнозируемый	1	2	3	All
Фактический				
1	189	894	3028	4111
2	912	7195	30044	38151
3	4022	36797	223878	264697
All	5123	44886	256950	306959

Посчитайте показатель recall для класса с меткой 3.

Запишите число:

1)

Ответ:

Задание №29

Имеется матрица ошибок:

Прогнозируемый	1	2	3	All
Фактический				
1	189	894	3028	4111
2	912	7195	30044	38151
3	4022	36797	223878	264697
All	5123	44886	256950	306959

Посчитайте показатель precision для класса с меткой 2.

Запишите число:

1)

Ответ:

Задание №30

Загрузите встроенный набор данных digits.

```
from sklearn.datasets import load_digits
```

```
digs = load_digits()
```

Разделите данные на тестовый и тренировочный наборы в пропорции 75% и 25%.

Создайте дерево решений.

Выявите лучшие параметры дерева решений на следующих данных: максимальная глубина дерева от 2 до 10, минимальное количество выборок, необходимых создания листа (внешнего узла) от 4 до 6. В ответ запишите полученные показатели через пробел, сначала максимальную глубину, затем минимальное количество выборок, необходимых для создания узла.

Запишите ответ:

1)	Ответ:	
Ответы:		
#1 (1 б.)	Ответ = 'secret'	
#2 (1 б.)	Ответ = 'loser'	
#3 (1 б.)	Ответ = 6196	
#4 (1 б.)	Ответ = 2075	
#5 (1 б.)	Ответ = Falkland Islands	
#6 (1 б.)	Ответ = 2034	
#7 (1 б.)	Ответ = 143 3154 4024	
#8 (1 б.)	Ответ = Python	
#9 (1 б.)	Ответ = dt = DecissionTreeClassifier()	
#10 (1 б.)	Ответ = 70598.0	
#11 (1 б.)	Ответ = 2383.0	
#12 (1 б.)	Ответ = 24.0	
#13 (1 б.)	Ответ = 1.58	
#14 (1 б.)	Ответ = 4041.0	
#15 (1 б.)	Ответ = 29.16	
#16 (1 б.)	Ответ = 2165.0	
#17 (1 б.)	Ответ = Переобучение	
#18 (1 б.)	Ответ = 0,136	
#19 (1 б.)	Ответ = China	
#20 (1 б.)	Ответ = 74783.0	
#21 (1 б.)	Ответ = 391 3126	
#22 (1 б.)	Ответ = 1589	
#23 (1 б.)	Ответ = 6085565.0	
#24 (1 б.)	Ответ = 17868335.0	

#25 (1 б.)	Ответ = Israel
#26 (1 б.)	Ответ = 103571.0
#27 (1 б.)	Ответ = rf = RandomForestClassifier(n_estimators = 15)
#28 (1 б.)	Ответ = 0.845790
#29 (1 б.)	Ответ = 0.160295
#30 (1 б.)	Ответ = 9 4

ПК-1.1: Применяет основные методы представления информации и алгоритмы обработки данных в профессиональной деятельности	Обучающийся владеет: - навыками визуализации данных большого объёма;
ПК-2.8: Использует методы искусственного интеллекта (машинного обучения) и анализа больших данных для решения прикладных задач	- навыками решения научно-технических задач в области своей профессиональной деятельности с использованием технологий анализа данных и машинного обучения.
Задание 31 Скачайте набор данных о поездах и транспортной сети Франции. https://www.kaggle.com/datasets/gatandubuc/public-transport-traffic-data-in-france. Прочтите данные из файла Regularities_by_liaisons_Trains_France.csv в датафрейм gb. Получите распределение числовых величин датафрейма gb используя базовые методы визуализации библиотеки pandas. Определите графически в каком месяце было больше всего отмененных поездов? Связаны ли пики данных с общим количеством запланированных поездов? Исследуйте графически корреляцию между числовыми признаками в наборе данных gb. Постройте тепловую карту. Задание 32 Скачайте данные из набора https://www.kaggle.com/code/ambaniverma/uk-traffic-accidents/data Используя средства визуализации, проведите исследование, есть ли параметры, которые наиболее сильно влияют на дорожные происшествия? В наборе не наблюдается сильных положительных корреляций, с погодой, местностью, возрастом водителя и т.д. Определите количество инцидентов в зависимости от зоны ограничения скорости. Задание 33 Зарегистрируйтесь в системе организации конкурсов по исследованию данных Kaggle. Скачайте набор данных о поездах и транспортной сети Франции. https://www.kaggle.com/datasets/gatandubuc/public-transport-traffic-data-in-france В разделе Details изучите информацию о файле Regularities_by_liaisons_Trains_France.csv. Ответьте на вопрос, в каком случае поезд считается прибывшим вовремя? Отсортируйте годы ведения статистики по общему количеству поездов, опоздавших более, чем на 15 минут. Выведите количество отмененных поездов по каждой станции отправления. Задание 34 Зарегистрируйтесь в системе организации конкурсов по исследованию данных Kaggle. Скачайте набор данных о поездах и транспортной сети Франции. https://www.kaggle.com/datasets/gatandubuc/public-transport-traffic-data-in-france В разделе Details изучите информацию о файле Regularities_by_liaisons_Trains_France.csv. Проанализировав полученные данные предложите способы оптимизации работы транспортной сети.	

2.3. Перечень вопросов для подготовки обучающихся к промежуточной аттестации

1. Что такое машинное обучение и какие основные методы машинного обучения используются в сфере электрического транспорта, электроэнергетики и электротехники?
2. Какие преимущества предоставляет применение машинного обучения в электротранспорте и электроэнергетике?
3. Какие задачи можно решать с помощью машинного обучения в области электрического транспорта?
4. Какие алгоритмы машинного обучения используются для прогнозирования потребления электроэнергии?
5. Какие методы машинного обучения применяются для оптимизации работы электрических сетей?
6. Какие технологии и инструменты используются для обработки и анализа данных в сфере электрического транспорта с использованием машинного обучения?
7. Какие проблемы могут возникнуть при применении методов машинного обучения в электротехнике и как их можно решить?
8. Какие примеры успешного применения машинного обучения в электроэнергетике можно назвать?

9. Какие показатели эффективности используются для оценки результатов применения методов машинного обучения в электрическом транспорте?
10. Какие основные принципы лежат в основе работы нейронных сетей и как они применяются в электротехнике?
11. Какие методы машинного обучения используются для оптимизации зарядки электромобилей?
12. Какие проблемы безопасности возникают при использовании алгоритмов машинного обучения в сфере электрического транспорта и как их можно предотвратить?
13. Какие тренды и перспективы развития машинного обучения в области электроэнергетики можно выделить?
14. Какие методы машинного обучения используются для диагностики и предотвращения аварийных ситуаций в электротехнике?
15. Какие навыки и знания необходимы специалистам в области электрического транспорта, электроэнергетики и электротехники для успешного применения методов машинного обучения?

3. Методические материалы, определяющие процедуру и критерии оценивания сформированности компетенций при проведении промежуточной аттестации

Критерии формирования оценок по ответам на вопросы, выполнению тестовых заданий

- оценка **«отлично»** выставляется обучающемуся, если количество правильных ответов на вопросы составляет 100 – 90% от общего объёма заданных вопросов;
- оценка **«хорошо»** выставляется обучающемуся, если количество правильных ответов на вопросы – 89 – 76% от общего объёма заданных вопросов;
- оценка **«удовлетворительно»** выставляется обучающемуся, если количество правильных ответов на тестовые вопросы – 75–60 % от общего объёма заданных вопросов;
- оценка **«неудовлетворительно»** выставляется обучающемуся, если количество правильных ответов – менее 60% от общего объёма заданных вопросов.

Критерии формирования оценок по результатам выполнения заданий

«Отлично» – ставится за работу, выполненную полностью без ошибок и недочетов.

«Хорошо» – ставится за работу, выполненную полностью, но при наличии в ней не более одной негрубой ошибки и одного недочета, не более трех недочетов.

«Удовлетворительно» – ставится за работу, если обучающийся правильно выполнил не менее 2/3 всей работы или допустил не более одной грубой ошибки и двух недочетов, не более одной грубой и одной негрубой ошибки, не более трех негрубых ошибок, одной негрубой ошибки и двух недочетов.

«Неудовлетворительно» – ставится за работу, если число ошибок и недочетов превысило норму для оценки «удовлетворительно» или правильно выполнено менее 2/3 всей работы.

Виды ошибок:

- *грубые ошибки: незнание основных понятий, правил, норм; незнание приемов решения задач; ошибки, показывающие неправильное понимание условия предложенного задания.*
- *негрубые ошибки: неточности формулировок, определений; нерациональный выбор хода решения.*
- *недочеты: нерациональные приемы выполнения задания; отдельные погрешности в формулировке выводов; небрежное выполнение задания.*

Критерии формирования оценок по зачету

«Зачтено» – студент демонстрирует знание основных разделов программы изучаемого курса: его базовых понятий и фундаментальных проблем; приобрел необходимые умения и навыки, освоил вопросы практического применения полученных знаний, не допустил фактических ошибок при ответе, достаточно последовательно и логично излагает теоретический материал, допуская лишь незначительные нарушения последовательности изложения и некоторые неточности.

«Не зачтено» – выставляется в том случае, когда студент демонстрирует фрагментарные знания основных разделов программы изучаемого курса: его базовых понятий и фундаментальных проблем. У экзаменуемого слабо выражена способность к самостоятельному аналитическому мышлению, имеются затруднения в изложении материала, отсутствуют необходимые умения и навыки, допущены грубые ошибки и незнание терминологии, отказ отвечать на дополнительные вопросы, знание которых необходимо для получения положительной оценки.