



Федеральное агентство по рыболовству
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Калининградский государственный технический университет»
(ФГБОУ ВО «КГТУ»)

УТВЕРЖДАЮ
И.о. директора института

Фонд оценочных средств
(приложение к рабочей программе модуля)
«АНАЛИЗ ДАННЫХ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ»

основной профессиональной образовательной программы бакалавриата
по направлению подготовки
36.03.02 ЗООТЕХНИЯ

ИНСТИТУТ

агроинженерии и пищевых систем

РАЗРАБОТЧИК

кафедра прикладной математики и информационных
технологий

1 РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ, ПЕРЕЧЕНЬ ОЦЕНОЧНЫХ СРЕДСТВ И КРИТЕРИИ ОЦЕНИВАНИЯ

1.1 Результаты освоения дисциплины

Таблица 1 – Планируемые результаты обучения по дисциплине, соотнесенные с установленными компетенциями

Код и наименование компетенции	Дисциплина	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями
<p>УК-1: Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</p>	<p>Анализ данных и искусственный интеллект</p>	<p><u>Знать:</u> основные методы анализа и оценки информации, полученной в том числе с помощью цифровых средств; классификацию задач систем искусственного интеллекта; методы и средства искусственного интеллекта для решения профессиональных задач.</p> <p><u>Уметь:</u> критически оценивать надежность источников информации в условиях неопределенности и избытка/недостатка информации для решения поставленных задач, в том числе в цифровой среде; осуществлять поиск данных в открытых источниках и специализированных библиотеках; выявлять и анализировать проблемную ситуацию, выделяя ее структурные составляющие и связи между ними; применять методы анализа данных при решении профессиональных задач; выбирать современные технологии искусственного интеллекта для решения профессиональных задач.</p> <p><u>Владеть</u> аналитическими и системными навыками, способностью к поиску информации; навыками подготовки и очистки данных, предназначенных для обработки системами искусственного интеллекта; современными программными средствами для решения задач анализа данных в профессиональной деятельности.</p>

1.2 К оценочным средствам текущего контроля успеваемости относятся:

- тестовые задания открытого и закрытого типов;
- контрольная работа (для заочной формы обучения).

Промежуточная аттестация в форме дифференцированного зачета (зачета с оценкой) проходит по результатам прохождения всех видов текущего контроля успеваемости. В отдельных случаях (при не прохождении всех видов текущего контроля) зачет может быть проведен в виде тестирования.

1.3 Критерии оценки результатов освоения дисциплины

Универсальная система оценивания результатов обучения включает в себя системы оценок: 1) «отлично», «хорошо», «удовлетворительно», «неудовлетворительно»; 2) «зачтено», «не зачтено»; 3) 100 – балльную/процентную систему и правило перевода оценок в пятибалльную систему (табл. 2).

Таблица 2 – Система оценок и критерии выставления оценки

Система оценок	2	3	4	5
	0-40%	41-60%	61-80 %	81-100 %
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
Критерий	«не зачтено»	«зачтено»		
1 Системность и полнота знаний в отношении изучаемых объектов	Обладает частичными и разрозненными знаниями, которые не может научно-корректно связывать между собой (только некоторые из которых может связывать между собой)	Обладает минимальным набором знаний, необходимым для системного взгляда на изучаемый объект	Обладает набором знаний, достаточным для системного взгляда на изучаемый объект	Обладает полнотой знаний и системным взглядом на изучаемый объект
2 Работа с информацией	Не в состоянии находить необходимую информацию, либо в состоянии находить отдельные фрагменты информации в рамках поставленной задачи	Может найти необходимую информацию в рамках поставленной задачи	Может найти, интерпретировать и систематизировать необходимую информацию в рамках поставленной задачи	Может найти, систематизировать необходимую информацию, а также выявить новые, дополнительные источники информации в рамках поставленной задачи
3 Научное осмысление	Не может делать научно корректных	В состоянии осуществлять	В состоянии осуществлять	В состоянии осуществлять

Система оценок	2	3	4	5
	0-40%	41-60%	61-80 %	81-100 %
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
Критерий	«не зачтено»	«зачтено»		
изучаемого явления, процесса, объекта	выводов из имеющихся у него сведений, в состоянии проанализировать только некоторые из имеющихся у него сведений	научно корректный анализ предоставленной информации	систематический и научно корректный анализ предоставленной информации, вовлекает в исследование новые релевантные задаче данные	систематический и научно-корректный анализ предоставленной информации, вовлекает в исследование новые релевантные поставленной задаче данные, предлагает новые ракурсы поставленной задачи
4 Освоение стандартных алгоритмов решения профессиональных задач	В состоянии решать только фрагменты поставленной задачи в соответствии с заданным алгоритмом, не освоил предложенный алгоритм, допускает ошибки	В состоянии решать поставленные задачи в соответствии с заданным алгоритмом	В состоянии решать поставленные задачи в соответствии с заданным алгоритмом, понимает основы предложенного алгоритма	Не только владеет алгоритмом и понимает его основы, но и предлагает новые решения в рамках поставленной задачи

1.4 Оценивание тестовых заданий закрытого типа осуществляется по системе зачтено/ не зачтено («зачтено» – 41-100% правильных ответов; «не зачтено» – менее 40 % правильных ответов) или пятибалльной системе (оценка «неудовлетворительно» - менее 40 % правильных ответов; оценка «удовлетворительно» - от 41 до 60 % правильных ответов; оценка «хорошо» - от 61 до 80% правильных ответов; оценка «отлично» - от 81 до 100 % правильных ответов).

Тестовые задания открытого типа оцениваются по системе «зачтено/ не зачтено». Оценивается верность ответа по существу вопроса, при этом не учитывается порядок слов в словосочетании, верность окончаний, падежи.

2 ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

УК-1: Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач

Тестовые задания открытого типа:

1. Процесс исследования, очистки, преобразования и моделирования данных с целью выявления полезной информации, предсказания и поддержки принятия решений – это _____ данных.

Ответ: анализ

2. Научное направление, связанное с попытками формализовать мышление человека – это _____

Ответ: искусственный интеллект (ИИ) (AI)

3. Метод для определения способности машины (компьютера) проявлять интеллектуальное поведение, неотличимое от человеческого в 1950 году предложил _____

Ответ: Тьюринг (А. Тьюринг) (Алан Тьюринг)

4. Выявление и удаление ошибок в данных в целях улучшения их качества – это _____ данных (*процесс*)

Ответ: очистка

5. Нерелевантная или ложная информация в данных, которая может быть уменьшена через фильтрацию, сглаживание или использование алгоритмов очистки данных – это _____

Ответ: шум

6. Проверка правильности работы (предсказательной способности) аналитической модели, построенной на основе машинного обучения, а также удостоверение, что она соответствует требованиям решаемой задачи – это _____

Ответ: валидация

7. Метод главных компонент (РСА) уменьшает _____ данных

Ответ: размерность

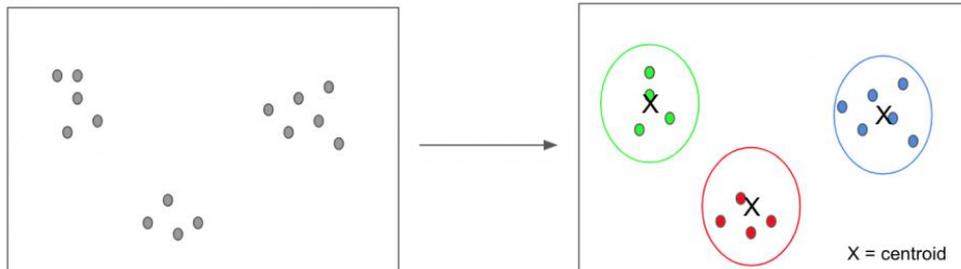
8. Деревья решений, логистическая регрессия, метод k-ближайших соседей и нейронные сети используются для решения задачи: _____ (*тип*)

Ответ: классификация

9. В задачах анализа покупательской корзины, маркетинговых исследованиях и управления инвентарем обычно используют поиск _____ правил

Ответ: ассоциативных

10. На рисунке



представлена иллюстрация алгоритма: _____

Введите общепринятое название алгоритма (англ.)

Ответ: k-means (к-средних)

11. Метод в статистике и машинном обучении, используемый для моделирования и анализа отношений между зависимой и одной или несколькими независимыми переменными – это _____ (*вид*) анализ

Ответ: регрессионный

12. Тесноту связи и направление линейной зависимости между двумя переменными определяет коэффициент корреляции _____

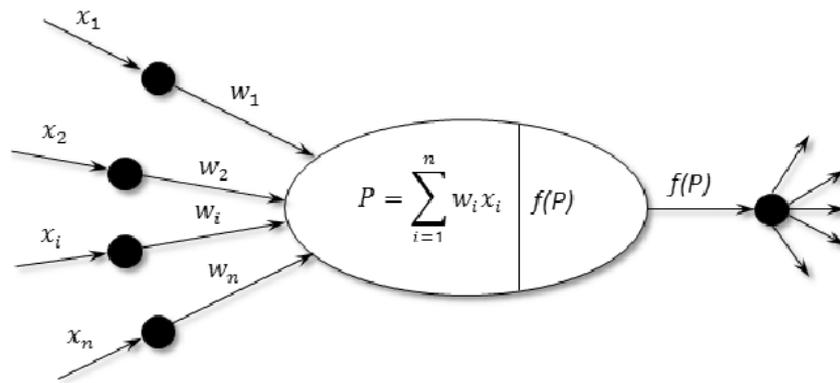
Ответ: Пирсона (Пирсон)

13. Количество слоев в перцептроне Розенблатта: ____

Введите число

Ответ: 1

14. На рисунке



представлена обобщенная математическая модель: _____ (объект)

Ответ: нейрона (нейрон)

15. В нейронной сети коэффициент, определяющий важность входных данных для выхода нейрона – это _____

Ответ: вес

16. Тип нейронной сети, идеально подходящий для работы с последовательными данными, такими как текст или временные ряды: _____

Ответ: рекуррентная (RNN) (РНС)

17. Тип нейронной сети, чаще всего используемый для анализа визуальных данных, например, для распознавания изображений: _____

Ответ: сверточная (CNN) (СНС)

18. Наиболее востребованы в технологии Datas Science современные языки программирования: _____

Ответ: Python, R (Питон, R)

19. Сценарий API Loginom хранится в файле с расширением: _____

Ответ: lgp (*.lgp)

20. В API Loginom при импорте данных (см. рис)

Обновить все		Обновить данные		Кол-во строк для анализа
Поля	7	ab	0/1	25
Имя	COL1	COL2	COL3	
Метка	Дата	Объект	Потребление эл энергия,кВт час	
Тип данных	7 Дата/Время	ab Строковый	0/1 Логический	
Вид данных	0 Непрерывный	0 Дискретный	0 Дискретный	
Использовать	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
1	01.05.2010, 0:00	Объект 2		
2	01.06.2010, 0:00	Объект 2		
3	01.07.2010, 0:00	Объект 2	true	
4	01.08.2010, 0:00	Объект 2		
5	01.09.2010, 0:00	Объект 2	true	
6	01.10.2010, 0:00	Объект 2	true	
7	01.11.2010, 0:00	Объект 2		
8	01.12.2010, 0:00	Объект 2	true	
9	01.01.2011, 0:00	Объект 2	true	
25	01.02.2011, 0:00	Объект 2		

неверно распознанные данные находятся в поле: ###

Введите имя поля

Ответ: COL3 (col3)

21. Методы искусственного интеллекта применяются для восприятия и распознавания _____ и понимания естественного _____.

Ответ: образов, языка

22. Цель Национальной стратегии развития искусственного интеллекта в России на 2030 год в области разработки программных и технологических решений - разработать решения аналогичные или способные превосходить _____ по широкому кругу задач.

Ответ: человеческие показатели

23. Когда одна транзакция изменяет какой-либо объект базы данных, а другая транзакция в этот момент читает данные из того же объекта, возникают _____ данные.

Ответ: грязные

Тестовые задания закрытого типа:

24. Установление соответствия:

Машинное обучение		Действия	
1	с учителем	а	Действовать с целью добиться максимального результата. Создавать автономных агентов, которые умеют учиться.
2	без учителя	б	Обучить необходимому поведению через размеченные данные. Разбираться с новыми данными на основе предыдущих данных
3	с подкреплением	в	Делать выводы без размеченных данных. Открывать неизвестные или скрытые закономерности.

Ответ: 1б, 2в, 3а

25. Для разработки системы классификации электронных письма на 'спам' и 'не спам' лучше всего подходит метод машинного обучения:

1. обучение с учителем
2. обучение без учителя
3. подкрепляющее обучение
4. активное обучение

26. Порядок шагов при создании модели регрессионного анализа:

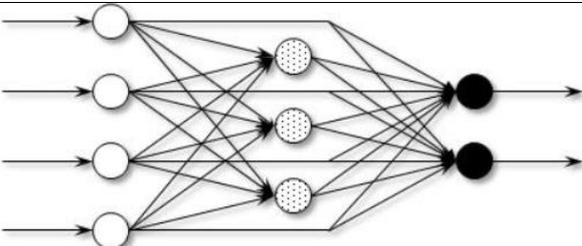
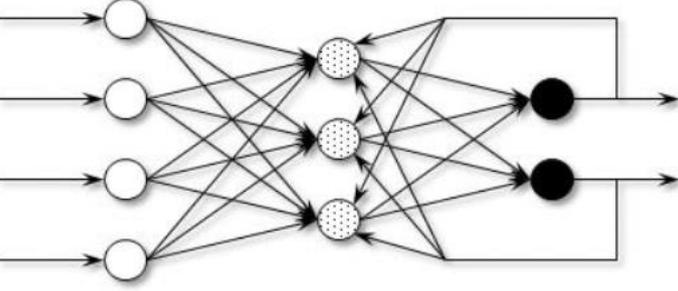
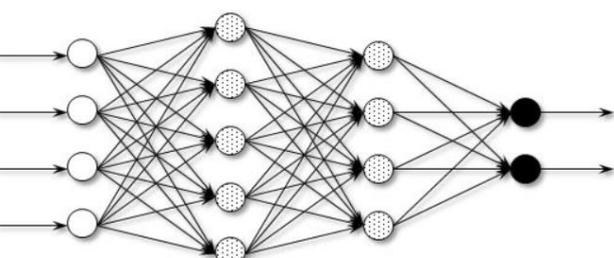
Шаг	
1	определение модели
2	расчет коэффициентов
3	оценка допущений модели
4	интерпретация результатов

Ответ: 1, 3, 2, 4

27. Алгоритм обучения перцептрона относится к типу обучения:

1. с учителем
2. без учителя
3. с подкреплением
4. глубоким

28. Установление соответствия:

Искусственная нейронная сеть		Тип межнейронных связей	
1		а	прямые
2		б	перекрестные
3		в	рекуррентные

Ответ: 1б, 2в, 3а

29. К области обработки естественного языка относится:

1. метод опорных векторов
2. генетический алгоритм
- 3. рекуррентная нейронная сеть**
4. случайный лес

30.

В API Loginom установление соответствие обозначений типов данных:

Знак (API Loginom)		Тип данных (API Loginom)	
1		а	Логический
2		б	Дата/Время
3		в	Вещественный
4		г	Целый
5		д	Строковый
6		е	Переменный

Ответ: 1г, 2а, 3б, 4в, 5е, 6д

3 ТИПОВЫЕ ЗАДАНИЯ НА КОНТРОЛЬНУЮ РАБОТУ, КУРСОВУЮ РАБОТУ/ КУРСОВОЙ ПРОЕКТ, РАСЧЕТНО-ГРАФИЧЕСКУЮ РАБОТУ

Учебным планом предусмотрено выполнение контрольной работы для студентов заочной формы обучения. Типовые задания для выполнения представлены ниже.

Задание 1. Подготовить реферат по заданной теме (по вариантам):

- Понятие анализа данных. Задачи анализа данных. Методы анализа данных. Принципы анализа данных.
- Инструменты анализа данных. Области применения анализа данных.
- Очистка данных. Преобразование данных. Интеграция данных. Разделение данных. Определение типов данных.
- Обработка отсутствующих значений. Обработка выбросов. Масштабирование данных. Нормализация данных.
- Кодирование категориальных признаков. Слияние данных. Удаление дубликатов. Индексирование данных.

- Визуализация данных. Оценка качества данных. Сохранение преобработанных данных.
- Линейная регрессия. Модель линейной регрессии. Коэффициенты линейной регрессии.
- Метод наименьших квадратов.
- Регрессионный анализ. Критерий оптимальности. Предсказание.
- Множественная линейная регрессия.
- Регуляризация.
- Лассо-регрессия.
- Ридж-регрессия.
- Кросс-валидация. Статистическая значимость. Коэффициент детерминации. Остатки. Анализ выбросов.
- Интерпретация результатов. Применение линейной регрессии в реальных задачах.
- Классификация. Класс. Объект. Признак. Алгоритм.
- Кластеризация. Разбиение. Дерево решений.
- Нейронные сети.
- Метод опорных векторов. Метрики качества. Переобучение. Регуляризация.
- Иерархическая кластеризация.
- Алгоритм k-средних.
- Алгоритм DBSCAN.
- Агломеративная кластеризация.
- Спектральная кластеризация.
- Метрики качества кластеризации. Определение оптимального количества кластеров.
- Проблемы кластеризации и их решения.
- Применение кластеризации в машинном обучении.
- Применение кластеризации на практике в различных областях.
- Ассоциативные правила.
- Правила ассоциации первого порядка. Правила ассоциации высоких порядков. Методы извлечения правил ассоциации.
- Алгоритм Apriori. Алгоритм FP-growth. Параллельные алгоритмы извлечения правил ассоциации.
- Построение дерева принятия решений на основе правил ассоциации.
- Анализ последовательностей событий на основе правил ассоциации.
- Оценка качества правил ассоциации.

- Искусственные нейронные сети. Математические модели.
- Распознавание образов. Классификация данных.
- Многослойный перцептрон.
- Сверточные нейронные сети.
- Компьютерное зрение.
- Обработка естественного языка.
- Архитектура нейронных сетей.
- Распознавание речи.
- Обработка изображений.

Задание 2. Использование возможностей платформы Loginom Community и MS Excel для обработки данных.

Создать сценарий Loginom, включающий в себя:

- импорт данных (формат *.csv, *.xlsx) на платформу Loginom Community;
- обработку данных в соответствии с заданием (данные о продажах товаров по группам и категориям):

а) исключить из набора записи с отсутствующими данными;

б) определить количество продаж товаров по категориям, сумму выручки в каждой категории за определенный период (по вариантам);

в) рассчитать прибыль (N% - по вариантам) от продаж определенной (по вариантам) группы товаров за определенный период (по вариантам)

- визуализацию результатов обработки соответствующих узлов сценария;

- экспорт результатов обработки в файл MS Excel.

Задание 3. Использование возможностей платформы Loginom Community и MS Excel для проведения регрессионного анализа исследовательских данных, планирования и обработки результатов факторного эксперимента.

Обработать и оценить результаты исследований по изучению влияния различных факторов на стоимость жилья.

Создать сценарий Loginom, включающий в себя:

- импорт данных (формат *.xlsx) на платформу *Loginom Community*;
- обработку данных в соответствии с заданием (исходный набор данных - по вариантам), используя модель линейной регрессии (компонент *Loginom* Линейная регрессия (парная)).

Условия отбора данных:

1. отсутствие записей с пустыми полями;
2. год постройки дома не ранее 1940 г. и не позже 2018 г;

3. средний класс жилья (от 2 до 7);
 4. независимые переменные: Площадь и Класс, зависимая переменная Цена, \$.
- визуализацию результатов обработки соответствующих узлов сценария;
 - экспорт результатов обработки в файл MS Excel;
 - сделать вывод.

4 СВЕДЕНИЯ О ФОНДЕ ОЦЕНОЧНЫХ СРЕДСТВ И ЕГО СОГЛАСОВАНИИ

Фонд оценочных средств для аттестации по дисциплине «Анализ данных и искусственный интеллект» представляет собой компонент основной профессиональной образовательной программы бакалавриата по направлению подготовки 36.03.02 Зоотехния.

Преподаватель-разработчик – Руденко А.И., к.ф.-м.н., Романов М.А.

Фонд оценочных средств рассмотрен и одобрен и.о. заведующего кафедрой прикладной математики и информационных технологий.

И.о. заведующего кафедрой



А.И. Руденко

Фонд оценочных средств рассмотрен и одобрен заведующим кафедрой производства и экспертизы качества сельскохозяйственной продукции

Заведующий кафедрой



А.С. Баркова

Фонд оценочных средств рассмотрен и одобрен методической комиссией института агроинженерии и пищевых систем (протокол № 07 от 27 августа 2024 г).

Председатель методической комиссии _____



М.Н. Альшевская