



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

УТВЕРЖДАЮ
Директор института

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

основной профессиональной образовательной программы специалитета
по специальности
38.05.01 «ЭКОНОМИЧЕСКАЯ БЕЗОПАСНОСТЬ»

ИНСТИТУТ

отраслевой экономики и управления

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 критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий</p> <p>ОПК-7 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</p> <p>ОПК-6 Способен использовать современные информационные технологии и программные средства при решении профессиональных задач</p>	<p>Анализ данных и искусственный интеллект</p>	<p><u>Знать:</u></p> <ul style="list-style-type: none"> – 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><u>Уметь:</u></p> <ul style="list-style-type: none"> – 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><u>Владеть</u></p> <ul style="list-style-type: none"> – 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

Код и наименование компетенции	Дисциплина	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями
		анализа данных в профессиональной деятельности.

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) в третьем семестре 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 %
Критерий	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
	«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	Может найти необходимую информацию в рамках поставленной задачи	Может найти, интерпретировать и систематизировать необходимую информацию в рамках поставленной	Может найти, систематизировать необходимую информацию, а также выявить новые, дополнительные источники

Система оценок	2	3	4	5
	0-40%	41-60%	61-80 %	81-100 %
Критерий	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»
	«не зачтено»	«зачтено»		
	рамках поставленной задачи		задачи	информации в рамках поставленной задачи
3 Научное осмысление изучаемого явления, процесса, объекта	Не может делать научно корректных выводов из имеющихся у него сведений, в состоянии проанализировать только некоторые из имеющихся у него сведений	В состоянии осуществлять научно корректный анализ предоставленной информации	В состоянии осуществлять систематический и научно корректный анализ предоставленной информации, вовлекает в исследование новые релевантные задаче данные	В состоянии осуществлять 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 критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий

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. _____ (англ.) - технология обнаружения в «сырых» данных ранее неизвестных, нетривиальных, практически полезных и доступных для интерпретации знаний, необходимых для принятия решения в различных сферах человеческой деятельности.

Ответ: Data Mining (интеллектуальный анализ данных)

3. Термин «искусственный интеллект» впервые был введен в _____ году

Ответ: 1956

4. Автором термина «искусственный интеллект» является _____

Ответ: Маккарти (Дж. Маккарти) (Джон Маккарти)

5. В основе имитационного подхода построения систем ИИ лежит модель: _____

Ответ: черный ящик

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

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

Тестовые задания закрытого типа:

7. Методы искусственного интеллекта применяются для:

1. синтаксического анализа текста

2. восприятия и распознавания образов**3. понимания естественного языка**

4. построения базы данных

8. Цель Национальной стратегии развития искусственного интеллекта в России на 2030 год в области разработки программных и технологических решений - разработать решения:

1. аналогичные или способные превосходить человеческие показатели по узкому кругу задач

2. аналогичные или способные превосходить человеческие показатели по широкому кругу задач

3. прогностические возможности которых кардинально отличаются от человеческих

4. аналогичные или способные превосходить показатели суперкомпьютеров по широкому кругу задач

ОПК-7 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

Тестовые задания открытого типа:

9. _____ в данных – это данные, отличающиеся от большинства остальных данных и указывающие на ошибку или исключительные ситуации.

Ответ: аномалии (аномалия)

10. Две или более записи одного набора данных, содержащих идентичные наборы значений всех признаков – это _____

Ответ: дубликаты (дубликат) (копия)

11. Состояние, при котором модель машинного обучения становится слишком хорошо подогнанной под обучающий набор данных, но плохо работает на новых данных – это _____

Ответ: переобучение

12. Метод предобработки числовых признаков в обучающих наборах данных с целью приведения их к некоторой общей шкале измерения без потери информации о различии диапазонов – это _____ данных.

Ответ: нормализация

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

Ответ: классификация

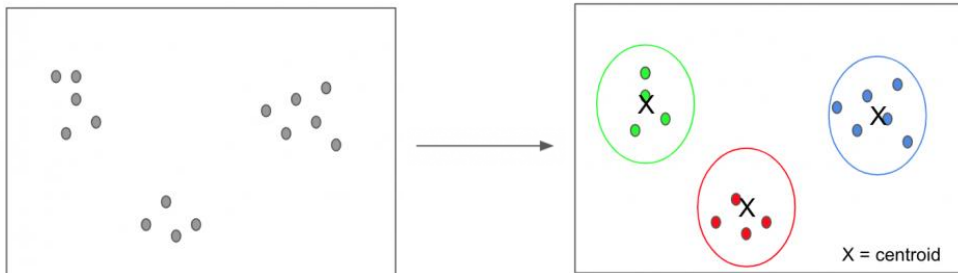
14. Задача машинного обучения, направленная на предсказание значения той или иной непрерывной числовой величины для входных данных – это _____ (тип задачи)

Ответ: регрессия

15. В анализе данных функция, используемая для определения расстояния между многомерными векторами в пространстве признаков – это _____

Ответ: метрика

16. На рисунке



представлена иллюстрация алгоритма: _____

Введите общепринятое название алгоритма (англ.)

Ответ: k-means (к-средних)

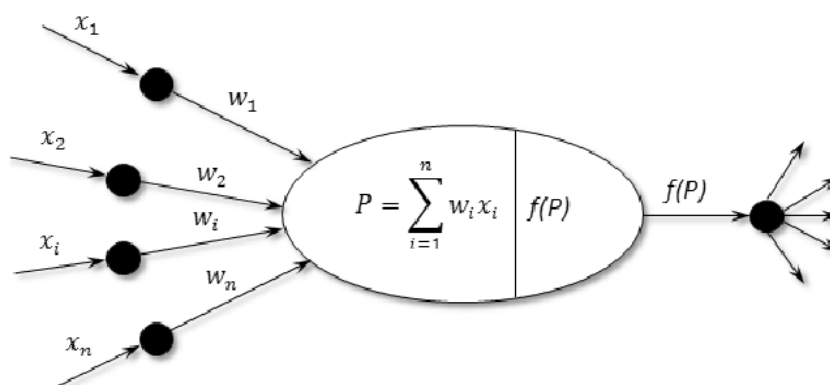
17. Тесноту связи и направление линейной зависимости между двумя переменными определяет коэффициент корреляции _____

Ответ: Пирсона (Пирсон)

18. Одним из первых попробовал смоделировать структуру человеческого мозга: _____ (фамилия ученого)

Ответ: Розенблатт (Ф. Розенблатт) (Фрэнк Розенблатт) (Розенблат)

19. На рисунке



представлена обобщенная математическая модель: _____ (объект)

Ответ: нейрона (нейрон)

20. В нейронной сети один или несколько нейронов, на входы которых подается один и тот же общий сигнал – это _____

Ответ: слой

21. В процессе обучения нейронных сетей для вычисления ошибки обычно используется функция _____

Ответ: потерь

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

Ответ: рекуррентная (RNN) (PHC)

23. _____ - процесс разделения текста на составляющие (слова, предложения, группы слов), является важным обязательным процессом при применении алгоритмов машинного обучения и нейронных сетей для анализа текста.

Ответ: токенизация

Тестовые задания закрытого типа:

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

1. грязные
2. старые
3. странные
4. противоречивые

25. Порядок шагов в модели машинного обучения:

Шаг	
1	трансформация данных
2	обучение модели
3	подготовка данных
4	валидация модели
5	загрузка данных

Ответ: 5, 3, 1, 2, 4

26. Для поиска ассоциативных правил используется алгоритм:

1. РСА

2. Apriori

3. SVM

4. K-ближайших соседей

27. Для бинарной задачи обычно используется результат классификации:

1. числовой

2. категорийный

3. временной ряд

4. мультиклассовый

28. Установление соответствия:

Тип НС		Задача	
1	Перцептроны	а	Анализ временных последовательностей
2	RNN	б	Распознавание изображений
3	CNN	в	Простые задачи классификации

Ответ: 1в, 2а, 3б

ОПК-6 Способен использовать современные информационные технологии и программные средства при решении профессиональных задач

Тестовые задания открытого типа:

29. Сценарий АП Loginom хранится в файле с расширением: _____

Ответ: lgp (*.lgp)

30. В АП Loginom последовательность действий, которые необходимо провести для анализа данных – это _____

Ответ: сценарий

31. В АП Loginom при импорте данных (см. рис)

Обновить все		Обновить данные		Кол-во строк для анализа
Поля	7 Дата	ab Объект	0/1 Потребление эл энергия,кВт час	
Имя	COL1	COL2	COL3	
Метка	Дата	Объект	Потребление эл энергия,кВт час	
Тип данных	7 Дата/Время	ab Строковый	0/1 Логический	
Вид данных	Непрерывный	Дискретный	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)

Тестовые задания закрытого типа:

32. В АП Loginom установление соответствие обозначений типов данных:

Знак (АП Loginom)		Тип данных (АП 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 одной контрольной работы (для заочной формы обучения).

КОНТРОЛЬНАЯ РАБОТА

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 предобработанных 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), использовав модель линейной регрессии может (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 специалитета по специальности 38.05.01 Экономическая безопасность.

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 экономической безопасности (протокол № 11 от 23.05.2024).

Заведующий кафедрой



Т.Е. Степанова

Фонд оценочных средств рассмотрен и одобрен методической комиссией ИНОТЭКУ (протокол № 5 от 20.05.2024).

Фонд оценочных средств актуализирован, рассмотрен и одобрен методической комиссией ИНОТЭКУ (протокол № 8 от 28.08.2024).

Председатель методической комиссии



И. А. Крамаренко