

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ
Директор института прикладной
математики и компьютерных наук

А. В. Замятин

« 19 » _____ 20 22 г.

Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине
(Оценочные средства по дисциплине)

Введение в интеллектуальный анализ данных

по направлению подготовки / специальности

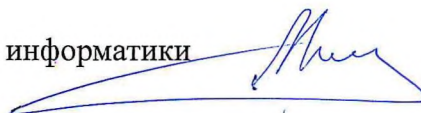
10.05.01 Компьютерная безопасность

Направленность (профиль) подготовки / специализация:

Анализ безопасности компьютерных систем

ОМ составил(и):

д-р техн. наук, профессор,
заведующий кафедрой теоретических основ информатики



А.В. Замятин

Рецензент:

д-р техн. наук, профессор,
заведующий кафедрой прикладной информатики

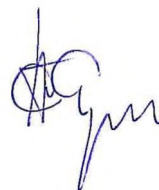


С.П. Сущенко

Оценочные средства одобрены на заседании учебно-методической комиссии
института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 12 мая 2022 г. № 4

Председатель УМК ИПМКН,
д-р техн. наук, профессор



С.П. Сущенко

Оценочные средства (ОС) являются элементом оценивания сформированности компетенций у обучающихся в целом или на определенном этапе ее формирования.

ОС разрабатываются в соответствии с рабочей программой (РП).

1. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Компетенция	Индикатор компетенции	Код и наименование результатов обучения	Критерии оценивания результатов обучения			
			Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
<p>ОПК-1 Способен оценивать роль информации, информационных технологий и информационной безопасности в современном обществе, их значение для обеспечения объективных потребностей личности, общества и государства</p> <p>ОПК-2 Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение, в том числе отечественного происхождения, для</p>	<p>ИОПК-1.1 Учитывает современные тенденции развития информационных технологий в своей профессиональной деятельности</p> <p>ИОПК-2.3 Формулирует предложения по применению программных средств системного и прикладного назначений, в том числе отечественного производства, используемых для решения задач профессиональной деятельности</p> <p>ИОПК-3.2 Осуществляет применение основных понятий, фактов, концепций, принципов математики и информатики для решения задач профессиональной деятельности</p> <p>ИОПК-3.3 Выявляет научную сущность проблем, возникающих в ходе профессиональной деятельности, и применяет соответствующий</p>	<p>ОР-3.1.1: Знать основные методы научно-практического поиска в задачах интеллектуального анализа данных и других областях с использованием информационных технологий.</p> <p>ОР-3.1.2: Знать существующие методы и подходы к интеллектуальному анализу данных различной природы.</p> <p>ОР-3.1.3: Уметь применять существующие методы интеллектуального анализа данных, обоснованно адаптируя и модифицируя их с учетом особенностей задачи предметной области.</p>	<p>Демонстрация высокого уровня знаний; Имеет сформированное представление об основных методах научно-практического поиска в задачах интеллектуального анализа данных и других областях с использованием информационных технологий; умение применять существующие методы интеллектуального анализа данных, обоснованно адаптируя и модифицируя их с учетом особенностей задачи предметной области.</p>	<p>В целом успешные, но содержащие отдельные пробелы знания об основных методах научно-практического поиска в задачах интеллектуального анализа данных и других областях с использованием информационных технологий. умение применять существующие методы интеллектуального анализа данных, обоснованно адаптируя и модифицируя их с учетом особенностей задачи</p>	<p>Фрагментарное, неполное знание без грубых ошибок знания об основных методах научно-практического поиска в задачах интеллектуального анализа данных и других областях с использованием информационных технологий. Фрагментарное, неполное знание без грубых ошибок умение применять существующие</p>	<p>Не имеет представления о современных методах интеллектуального анализа данных.</p>

<p>решения задач профессиональной деятельности</p> <p>ОПК-3 Способен на основании совокупности математических методов разрабатывать, обосновывать и реализовывать процедуры решения задач профессиональной деятельности</p> <p>ОПК-14 Способен проектировать базы данных, администрировать системы управления базами данных в соответствии с требованиями по защите информации</p> <p>ПК-2 Способен разрабатывать требования к программно-аппаратным средствам защиты информации компьютерных систем и сетей</p>	<p>математический аппарат для их формализации, анализа и выработки решения</p> <p>ИОПК-14.1 Понимает модели и структуры данных, физические модели баз данных, принципы организации и методы проектирования баз данных, языки и системы программирования баз данных</p> <p>ИПК-2.2 Разрабатывает математические модели, реализуемые в средствах защиты информации;</p> <p>ИПК-2.3 Проводит исследования с целью нахождения наиболее целесообразных практических решений по обеспечению защиты информации</p>			<p>предметной области.</p>	<p>методы интеллектуального анализа данных.</p>	
--	---	--	--	----------------------------	---	--

2. Этапы формирования компетенций и виды оценочных средств

№	Этапы формирования компетенций (разделы дисциплины)	Код и наименование результатов обучения	Вид оценочного средства (тесты, задания, кейсы, вопросы и др.)
1.	Раздел 1. Основные проблемы построения систем	ОР-3.1.1, ОР-3.1.2.	Задания для текущего контроля, вопросы для проведения промежуточной аттестации
2.	Раздел 2. Предварительная обработка данных. Классификация.	ОР-3.1.1, ОР-3.1.2. ОР-3.1.3	Задания для текущего контроля, вопросы для проведения промежуточной аттестации
3.	Раздел 3. Регрессия. Ассоциация, последовательная ассоциация, аномалии и визуализация.	ОР-3.1.1, ОР-3.1.2. ОР-3.1.3	Задания для текущего контроля, вопросы для проведения промежуточной аттестации
4.	Раздел 4. Высокопроизводительная обработка данных. Программные среды для интеллектуального анализа данных.	ОР-3.1.1, ОР-3.1.2. ОР-3.1.3	Задания для текущего контроля, вопросы для проведения промежуточной аттестации

3. Типовые контрольные задания или иные материалы, необходимые для оценки образовательных результатов обучения

3.1. Типовые задания для проведения текущего контроля успеваемости по дисциплине

Реферат (на согласованную тему). К реферату необходимо сделать презентацию.

Примеры тем:

Современные нейронные сети в обработке данных (изображений, видео, технологических сигналов, музыки и т.п.);

Современные алгоритмы классификации (изображений, текстов и т.п.);

Интеллектуальная обработка данных в ... (промышленности, медицине, бизнесе, индустрии развлечений, досуга и др.);

Извлечение знаний из текстов;

Детектирование аномалий;

Разновидности сверточных нейронных сетей;

Интеллектуальные алгоритмы в ранней диагностике заболеваний;

Интеллектуальные алгоритмы в персонализированной медицине;

Интеллектуальные алгоритмы в робототехнике, транспортных системах и т.п.;

Интеллектуальные алгоритмы в банковском деле/страховании/...;

Проект (на согласованную тему). Реализовать небольшой проект по интеллектуальной обработке данных с использованием среды RapidMiner или одного из языков

программирования (например, Python, R), с возможным использованием общедоступных баз данных (или данных из иных источников).

Этапы реализации проекта:

Поиск и подготовка набора данных;

Разработка технического задания;

Пилотная реализация одной модели, выбор метрики и оценка точности (фиксация полученной точности на этом этапе);

Реализация всех пунктов технического задания, настройка параметров моделей, оценка точности (точность, полученная на этом этапе должна быть больше чем на предыдущем):

Подготовка отчета (с описанием предметной области, выбранных алгоритмов и параметров моделей), презентации, публичная защита проекта;

Каждый студент реализует индивидуальный или групповой проект как последовательность лабораторных работ:

Лабораторная работа №1. Индивидуальное задание по теме «Анализ предметной области, формулировка целей и задач исследования. Извлечение и первичное сохранение данных».

Цель работы – научить студентов решать задачи анализа предметной области, ее адаптации для методов анализа данных с учетом принципиальных особенностей предметной области.

Лабораторная работа №2. Индивидуальное задание по теме «Предварительная обработка данных: очистка, интеграция, преобразование».

Цель работы – научить студентов решать задачи предварительной обработки данных, предполагающей трудоемкую процедуру очистки (исключение противоречий, случайных выбросов и помех, пропусков), интеграции (объединение данных из нескольких возможных источников в одном хранилище), преобразования (может включать агрегирование и сжатие данных, дискретизацию атрибутов и сокращение размерности и т.п.).

Лабораторная работа №3. Индивидуальное задание по теме «Содержательный анализ данных методами Data Mining».

Цель работы – научить студентов обоснованно применять базовые методы интеллектуального анализа данных, учитывая особенности как теоретического построения применяемых методов, так и выбранной предметной области.

Лабораторная работа №4. Индивидуальное задание по теме «Визуализация и интерпретация полученных результатов».

Цель работы – научить студентов выполнять визуализацию и интерпретация полученных результатов в виде, пригодном для принятия управленческих решений.

Примеры тем для самостоятельного изучения:

- Нейросетевые методы анализа данных, сверточные сети (convolution neural networks), глубинное обучение (deep learning).
- Методы интеллектуального анализа медиа (social media data mining).
- Методы машинного обучения в задачах финансовой аналитики.
- Методы машинного обучения в задачах ранней медицинской диагностики.

- Комбинирование моделей в анализе данных, бустинг.
- Метод анализа независимых компонент (independent component analysis).
- Методы визуализации данных высокой размерности.

3.2. Типовые задания для проведения промежуточной аттестации по дисциплине

Вопросы к зачету с оценкой:

1. Основные понятия, терминология;
2. Data Mining / Data Science;
3. Big Data (основные понятия и свойства);
4. Дедукция и индукция;
5. Интеллектуальный анализ данных в бизнесе примеры применения;
6. Интеллектуальный анализ данных в решении сложных прикладных задач;
7. Интеллектуальный анализ данных в ранней диагностике опасных заболеваний;
8. Интеллектуальный анализ данных в индустриальной предиктивной аналитике;
9. Основные задачи и классификация методов анализа данных;
10. Принципиальные основы машинного обучения;
11. Предварительная обработка данных;
12. Оптимизация признаков пространства;
13. Постановка задачи классификации;
14. Контролируемая непараметрическая классификация;
15. Контролируемая непараметрическая нейросетевая классификация;
16. Классификация по методу машины опорных векторов;
17. Деревья решений;
18. Неконтролируемая классификация (кластеризация);
19. Регрессия (понятие регрессии, основные этапы регрессионного анализа, методы восстановления регрессии);
20. Ассоциация;
21. Последовательная ассоциация (алгоритмы семейства «Априори», алгоритм GSP);
22. Многоуровневое машинное обучение (бутстрэппинг, бэггинг, стекинг, бустинг);
23. Обнаружение аномалий;
24. Визуализация в Data Mining;
25. Функции активации;
26. Основные типы искусственных нейронных сетей;
27. Сверточные нейронные сети;
28. Среды и фреймворки глубинного обучения;
29. Основные задачи обработки текста;
30. Этапы предварительной обработки текста;
31. Метрики качества классификации;
32. Гипотеза A/B, Каппа-индекс согласия, ROC-кривая;
33. Метрика качества прогноза временного ряда;
34. Метрики качества кластеризации;
35. Принципы высокопроизводительных вычислений;
36. Особенности построения вычислительного кластера;
37. Среды и инструменты высокопроизводительных вычислений;
38. Инструменты data mining.

4. Методические материалы, определяющие процедуры оценивания образовательных результатов обучения

4.1. Методические материалы для оценки текущего контроля успеваемости по дисциплине.

Рейтинговая система для оценки текущей успеваемости обучающихся

Таблица – Балльные оценки для элементов контроля

Элементы учебной деятельности	Максимальный балл с начала семестра	Оцениваемая компетенция
Реферат по теме с презентацией	20	ОПК-1; ОПК-2; ОПК-3
Реализация проекта	40	ОПК-1; ОПК-2; ОПК-3; ОПК-14; ПК-2
Опрос на занятиях	10	ОПК-1; ОПК-2; ОПК-3
Зачет с оценкой	30	ОПК-1; ОПК-2; ОПК-3; ОПК-14; ПК-2

4.2. Методические материалы для проведения промежуточной аттестации по дисциплине.

Сумма баллов, набранная студентом в течение семестра, переводится в оценку промежуточной аттестации успеваемости студента по приведенной ниже шкале.

Пересчет баллов в оценки для промежуточной аттестации

Баллы на дату контрольной точки	Оценка
$\geq 90\%$ от максимальной суммы баллов	5
От 70% до 89% от максимальной суммы баллов	4
От 60% до 69% от максимальной суммы баллов	3
$< 60\%$ от максимальной суммы баллов	2