МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ»
(ТУСУР)

УТВЕРЖДАЮ
Директор департамента образования
_____ П.Е. Троян
«19» 12 2018 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ

Уровень образования: **высшее образование - магистратура**
Направление подготовки / специальность: **11.04.02 Инфокоммуникационные технологии и системы связи**
Направленность (профиль) / специализация: **Системы беспроводной связи и Интернета вещей**
Форма обучения: **очная**
Факультет: **Радиотехнический факультет (РТФ)**
Кафедра: **Кафедра телекоммуникаций и основ радиотехники (ТОР)**
Курс: **2**
Семестр: **3**
Учебный план набора 2019 года

Объем дисциплины и виды учебной деятельности

Виды учебной деятельности	3 семестр	Всего	Единицы
Лекционные занятия	18	18	часов
Практические занятия	18	18	часов
Лабораторные занятия	18	18	часов
Самостоятельная работа	54	54	часов
Общая трудоемкость (включая промежуточную аттестацию)	108	108	часов
	3	3	з.е.

Формы промежуточной аттестация	Семестр
Зачет	3

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Шелупанов А.А.
Должность: Ректор
Дата подписания: 19.12.2018
Уникальный программный ключ:
c53e145e-8b20-45aa-9347-a5e4dbb90e8d

Томск

Согласована на портале № 63293

1. Общие положения

1.1. Цели дисциплины

1. Формирование у студентов профессиональных знаний и практических навыков в области больших данных и искусственного интеллекта.
2. Формирование у студентов теоретических знаний и практических навыков по вопросам анализа данных; поиска управленческих решений; освоение студентами современных математических методов машинного обучения.

1.2. Задачи дисциплины

1. Изучение основных понятий, процесса и технологий «больших данных».
2. Изучение методов статистического анализа данных.
3. Изучение принципов и методов машинного обучения.

2. Место дисциплины в структуре ОПОП

Блок дисциплин: ФТД. Факультативные дисциплины.
Индекс дисциплины: ФТД.03.
Реализуется с применением электронного обучения, дистанционных образовательных технологий.

3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Процесс изучения дисциплины направлен на формирование следующих компетенций в соответствии с ФГОС ВО и основной образовательной программой (таблица 3.1):
Таблица 3.1 – Компетенции и индикаторы их достижения

Компетенция	Индикаторы достижения компетенции	Планируемые результаты обучения по дисциплине
Универсальные компетенции		
-	-	-
Общепрофессиональные компетенции		

Согласована на портале № 63293

ОПК-3. Способен приобретать, обрабатывать и использовать новую информацию в своей предметной области, предлагать новые идеи и подходы к решению задач своей профессиональной деятельности	ОПК-3.1. Знает принципы построения локальных и глобальных компьютерных сетей, основы Интернет-технологий, типовые процедуры применения проблемно-ориентированных прикладных программных средств в дисциплинах профессионального цикла и профессиональной сфере деятельности	Знает математические особенности основных методов и алгоритмов анализа данных. Знает методы обработки данных и их визуального анализа для последующего машинного обучения.
	ОПК-3.2. Умеет использовать современные информационные и компьютерные технологии, средства коммуникаций для эффективного поиска информации из своей предметной области	Умеет создавать функций для последующего машинного обучения. Умеет применить библиотеки машинного обучения. Умеет оценить полученные результаты и переработать алгоритмы обработки данных. Имеет навыки обработки и интерпретации результатов анализа данных с помощью специализированного программного обеспечения.
	ОПК-3.3. Владеет методами научно-технического творчества, способами генерации новых идей и подходов для решения профессиональных задач	Владеет методами разработки оригинальных программных средств для обработки данных.
Профессиональные компетенции		
-	-	-

4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 3 зачетных единиц, 108 академических часов.

Распределение трудоемкости дисциплины по видам учебной деятельности представлено в таблице 4.1.

Таблица 4.1 – Трудоемкость дисциплины по видам учебной деятельности

Виды учебной деятельности	Всего часов	Семестры
		3 семестр
Контактная аудиторная работа обучающихся с преподавателем, всего	54	54
Лекционные занятия	18	18
Практические занятия	18	18
Лабораторные занятия	18	18
Самостоятельная работа обучающихся, в т.ч. контактная внеаудиторная работа обучающихся с преподавателем, всего	54	54
Подготовка к зачету	14	14
Подготовка к тестированию	16	16

Согласована на портале № 63293

Написание отчета по практическому занятию (семинару)	4	4
Подготовка к лабораторной работе, написание отчета	8	8
Написание отчета по лабораторной работе	4	4
Выполнение практического задания	8	8
Общая трудоемкость (в часах)	108	108
Общая трудоемкость (в з.е.)	3	3

5. Структура и содержание дисциплины

5.1. Разделы (темы) дисциплины и виды учебной деятельности

Структура дисциплины по разделам (темам) и видам учебной деятельности приведена в таблице 5.1.

Таблица 5.1 – Разделы (темы) дисциплины и виды учебной деятельности

Названия разделов (тем) дисциплины	Лек. зан., ч	Прак. зан., ч	Лаб. раб.	Сам. раб., ч	Всего часов (без экзамена)	Формируемые компетенции
3 семестр						
1 Введение	2	4	-	10	16	ОПК-3
2 Многомерная статистика и её элементы	3	2	-	6	11	ОПК-3
3 Методы статистического анализа данных	7	10	2	14	33	ОПК-3
4 Основы нейронных сетей	3	2	-	10	15	ОПК-3
5 Методы машинного обучения	3	-	16	14	33	ОПК-3
Итого за семестр	18	18	18	54	108	
Итого	18	18	18	54	108	

5.2. Содержание разделов (тем) дисциплины

Содержание разделов (тем) дисциплины (в т.ч. по лекциям) приведено в таблице 5.2.

Таблица 5.2 – Содержание разделов (тем) дисциплины (в т.ч. по лекциям)

Названия разделов (тем) дисциплины	Содержание разделов (тем) дисциплины (в т.ч. по лекциям)	Трудоемкость (лекционные занятия), ч	Формируемые компетенции
3 семестр			
1 Введение	Введение в анализ данных: понятие данных, типы шкал измерений, жизненный цикл данных, этапы решения задачи анализа данных и их взаимосвязи.	2	ОПК-3
	Итого	2	
2 Многомерная статистика и её элементы	Многомерные нормальные массивы данных, условные нормальные распределения, распределения Уишарта; T2-тест Хотеллинга; Коррекция множественного тестирования; Гистограммы; Оценка плотности ядра.	3	ОПК-3
	Итого	3	

Согласована на портале № 63293

3 Методы статистического анализа данных	Основные задачи и методы статистического анализа, корреляционный анализ, дисперсионный анализ, регрессионный анализ, факторный анализ, кластерный анализ, дискриминантный анализ, многомерное шкалирование.	3	ОПК-3
	Регрессия, перекрестная проверка. Классификация, показатели качества. Деревья решений. Объединение наборов данных в ансамбль. Проектирование и выбор функций. Уменьшение Размерности. Кластеризация данных.	4	ОПК-3
	Итого	7	
4 Основы нейронных сетей	Стохастический градиентный спуск, Многослойный персептрон, функции активации (ReLU, tanh), Отсев, обучение и валидация. Ранняя остановка, Сверточные сети; Библиотека Keras. Практические примеры	3	ОПК-3
	Итого	3	
5 Методы машинного обучения	Классификация методов машинного обучения. Решение задач машинного обучения на Python. Методы машинного обучения без учителя: кластеризация, сетевой анализ. Методы машинного обучения с учителем: наивный Байесовский классификатор, комплексные модели, деревья решений, бустинг и бэггинг.	3	ОПК-3
	Итого	3	
Итого за семестр		18	
Итого		18	

5.3. Практические занятия (семинары)

Наименование практических занятий (семинаров) приведено в таблице 5.3.

Таблица 5.3 – Наименование практических занятий (семинаров)

Названия разделов (тем) дисциплины	Наименование практических занятий (семинаров)	Трудоемкость, ч	Формируемые компетенции
3 семестр			
1 Введение	Шкалы измерений	2	ОПК-3
	Анализ главных компонент, обзор нелинейных методов (Isomap, LTSA, tSNE). Практические примеры: DR для аэродинамических профилей и создание новых аэродинамических профилей, генетическая подпись неизвестного происхождения	2	ОПК-3
	Итого	4	

Согласована на портале № 63293

2 Многомерная статистика и её элементы	Применение разных типов распределений для предсказания событий	2	ОПК-3
	Итого	2	
3 Методы статистического анализа данных	Практический пример: набор данных о спросе на совместное использование велосипедов. Контролируемое обучение, k ближайших соседей с помощью линейной регрессии, концепции переобучения и недообучения (компромисс между смещением и дисперсией).	4	ОПК-3
	Задачи классификации, логистическая регрессия, SVM, функции потерь, точность и отзыв, кривая ROC, Обзор, обработка пропущенных значений, вычисление важности объектов, сложность алгоритмов, визуализация.	2	ОПК-3
	Пакетирование, Бустинг, Случайный лес, Градиентный Бустинг, библиотека XGBoost. Практический пример: Прогнозирование типа лесного покрова	2	ОПК-3
	K-средние, Модель Гауссовой смеси, Иерархическая кластеризация, Спектральная кластеризация. Практический пример: кластеризация текстовых документов	2	ОПК-3
	Итого	10	
4 Основы нейронных сетей	Распознавание ключевых точек лица	2	ОПК-3
	Итого	2	
Итого за семестр		18	
Итого		18	

5.4. Лабораторные занятия

Наименование лабораторных работ приведено в таблице 5.4.

Таблица 5.4 – Наименование лабораторных работ

Названия разделов (тем) дисциплины	Наименование лабораторных работ	Трудоемкость, ч	Формируемые компетенции
3 семестр			
3 Методы статистического анализа данных	Модели математической статистики для предсказания поведения покупателей. Их отличия.	2	ОПК-3
	Итого	2	

Согласована на портале № 63293

5 Методы машинного обучения	Кластерный анализ для сегментирования клиентской базы	4	ОПК-3
	Наивный байесовский классификатор для классификации твитов	4	ОПК-3
	Кластеризация на основе сетевых графов	4	ОПК-3
	Ансамбли моделей (бэггинг и бустинг)	4	ОПК-3
	Итого	16	
Итого за семестр		18	
Итого		18	

5.5. Курсовой проект / курсовая работа

Не предусмотрено учебным планом

5.6. Самостоятельная работа

Виды самостоятельной работы, трудоемкость и формируемые компетенции представлены в таблице 5.6.

Таблица 5.6 – Виды самостоятельной работы, трудоемкость и формируемые компетенции

Названия разделов (тем) дисциплины	Виды самостоятельной работы	Трудоемкость, ч	Формируемые компетенции	Формы контроля
3 семестр				
1 Введение	Подготовка к зачету	2	ОПК-3	Зачёт
	Подготовка к тестированию	4	ОПК-3	Тестирование
	Написание отчета по практическому занятию (семинару)	4	ОПК-3	Отчет по практическому занятию (семинару)
	Итого	10		
2 Многомерная статистика и её элементы	Подготовка к зачету	2	ОПК-3	Зачёт
	Подготовка к тестированию	4	ОПК-3	Тестирование
	Итого	6		
3 Методы статистического анализа данных	Подготовка к зачету	4	ОПК-3	Зачёт
	Подготовка к тестированию	4	ОПК-3	Тестирование
	Подготовка к лабораторной работе, написание отчета	4	ОПК-3	Лабораторная работа
	Написание отчета по лабораторной работе	2	ОПК-3	Отчет по лабораторной работе
	Итого	14		

Согласована на портале № 63293

4 Основы нейронных сетей	Подготовка к зачету	4	ОПК-3	Зачёт
	Выполнение практического задания	4	ОПК-3	Практическое задание
	Подготовка к тестированию	2	ОПК-3	Тестирование
	Итого	10		
5 Методы машинного обучения	Подготовка к зачету	2	ОПК-3	Зачёт
	Выполнение практического задания	4	ОПК-3	Практическое задание
	Подготовка к тестированию	2	ОПК-3	Тестирование
	Подготовка к лабораторной работе, написание отчета	4	ОПК-3	Лабораторная работа
	Написание отчета по лабораторной работе	2	ОПК-3	Отчет по лабораторной работе
	Итого	14		
Итого за семестр		54		
Итого		54		

5.7. Соответствие компетенций, формируемых при изучении дисциплины, и видов учебной деятельности

Соответствие компетенций, формируемых при изучении дисциплины, и видов учебной деятельности представлено в таблице 5.7.

Таблица 5.7 – Соответствие компетенций, формируемых при изучении дисциплины, и видов занятий

Формируемые компетенции	Виды учебной деятельности				Формы контроля
	Лек. зан.	Прак. зан.	Лаб. раб.	Сам. раб.	
ОПК-3	+	+	+	+	Зачёт, Лабораторная работа, Практическое задание, Тестирование, Отчет по лабораторной работе, Отчет по практическому занятию (семинару)

6. Рейтинговая система для оценки успеваемости обучающихся

6.1. Балльные оценки для форм контроля

Балльные оценки для форм контроля представлены в таблице 6.1.

Таблица 6.1 – Балльные оценки

Формы контроля	Максимальный балл на 1-ую КТ с начала семестра	Максимальный балл за период между 1КТ и 2КТ	Максимальный балл за период между 2КТ и на конец семестра	Всего за семестр
3 семестр				
Зачёт	5	10	15	30
Лабораторная работа	0	5	15	20
Практическое задание	0	0	10	10
Тестирование	0	5	10	15

Согласована на портале № 63293

Отчет по лабораторной работе	0	5	10	15
Отчет по практическому занятию (семинару)	5	5	0	10
Итого максимум за период	10	30	60	100
Нарастающим итогом	10	40	100	100

6.2. Пересчет баллов в оценки за текущий контроль

Пересчет баллов в оценки за текущий контроль представлен в таблице 6.2.

Таблица 6.2 – Пересчет баллов в оценки за текущий контроль

Баллы на дату текущего контроля	Оценка
≥ 90% от максимальной суммы баллов на дату ТК	5
От 70% до 89% от максимальной суммы баллов на дату ТК	4
От 60% до 69% от максимальной суммы баллов на дату ТК	3
< 60% от максимальной суммы баллов на дату ТК	2

6.3. Пересчет суммы баллов в традиционную и международную оценку

Пересчет суммы баллов в традиционную и международную оценку представлен в таблице 6.3.

Таблица 6.3 – Пересчет суммы баллов в традиционную и международную оценку

Оценка	Итоговая сумма баллов, учитывает успешно сданный экзамен	Оценка (ECTS)
5 (отлично) (зачтено)	90 – 100	A (отлично)
4 (хорошо) (зачтено)	85 – 89	B (очень хорошо)
	75 – 84	C (хорошо)
	70 – 74	D (удовлетворительно)
3 (удовлетворительно) (зачтено)	65 – 69	
	60 – 64	E (посредственно)
2 (неудовлетворительно) (не зачтено)	Ниже 60 баллов	F (неудовлетворительно)

7. Учебно-методическое и информационное обеспечение дисциплины

7.1. Основная литература

1. Юре, Л. Анализ больших наборов данных / Л. Юре, Р. Ананд, Д. У. Джеффри ; перевод с английского А. А. Слинкин. — Москва : ДМК Пресс, 2016. — 498 с. — ISBN 978-5-97060-190-7. — Текст : электронный // Лань : электронно-библиотечная система. [Электронный ресурс]: — Режим доступа: https://e.lanbook.com/book/93571.

2. Прикладная математическая статистика: Учебное пособие / А. А. Мицель - 2019. 113 с. [Электронный ресурс]: — Режим доступа: https://edu.tusur.ru/publications/9151.

7.2. Дополнительная литература

1. Прикладная математическая статистика: Учебное пособие / А. А. Мицель - 2016. 113 с. [Электронный ресурс]: — Режим доступа: https://edu.tusur.ru/publications/6252.

2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2022. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. [Электронный ресурс]: — Режим доступа: https://urait.ru/bcode/469306.

Согласована на портале № 63293

3. Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст : электронный // Лань : электронно-библиотечная система. [Электронный ресурс]: — Режим доступа: https://e.lanbook.com/book/131721.

7.3. Учебно-методические пособия

7.3.1. Обязательные учебно-методические пособия

1. Вычислительные методы: Учебное пособие / А. А. Мицель - 2013. 198 с. [Электронный ресурс]: — Режим доступа: https://edu.tusur.ru/publications/4863.
2. Прикладная математическая статистика: Лабораторный практикум / А. А. Мицель - 2019. 72 с. [Электронный ресурс]: — Режим доступа: https://edu.tusur.ru/publications/9152.
3. Прикладная математическая статистика: Методические указания по выполнению самостоятельной работы / А. А. Мицель - 2016. 10 с. [Электронный ресурс]: — Режим доступа: https://edu.tusur.ru/publications/6255.

7.3.2. Учебно-методические пособия для лиц с ограниченными возможностями здоровья и инвалидов

Учебно-методические материалы для самостоятельной и аудиторной работы обучающихся из числа лиц с ограниченными возможностями здоровья и инвалидов предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации.

Для лиц с нарушениями зрения:

– в форме электронного документа;

– в печатной форме увеличенным шрифтом.

Для лиц с нарушениями слуха:

– в форме электронного документа;

– в печатной форме.

Для лиц с нарушениями опорно-двигательного аппарата:

– в форме электронного документа;

– в печатной форме.

7.4. Современные профессиональные базы данных и информационные справочные системы

1. При изучении дисциплины рекомендуется обращаться к современным базам данных, информационно-справочным и поисковым системам, к которым у ТУСУРа открыт доступ: https://lib.tusur.ru/ru/resursy/bazy-dannyh.

8. Материально-техническое и программное обеспечение дисциплины

8.1. Материально-техническое и программное обеспечение для лекционных занятий

Для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используется учебная аудитория с достаточным количеством посадочных мест для учебной группы, оборудованная доской и стандартной учебной мебелью. Имеются мультимедийное оборудование и учебно-наглядные пособия, обеспечивающие тематические иллюстрации по лекционным разделам дисциплины.

8.2. Материально-техническое и программное обеспечение для практических занятий

Учебная аудитория "Цифровая связь": учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий практического типа, учебная аудитория для проведения занятий лабораторного типа, учебная аудитория для проведения занятий семинарского типа, помещение для курсового проектирования (выполнения курсовых работ), помещение для проведения групповых и индивидуальных консультаций, помещение для проведения текущего контроля и промежуточной аттестации, помещение для самостоятельной работы; 634034, Томская область, г. Томск, Вершинина улица, д. 47, 309 ауд.

Описание имеющегося оборудования:

- Магнитно-маркерная доска;
- Комплект специализированной учебной мебели;

Согласована на портале № 63293

- Рабочее место преподавателя.

Программное обеспечение:
- Algorithm Builder;
- Google Chrome;
- Microsoft Office 2010 и ниже;
- Microsoft Windows 8.1;

8.3. Материально-техническое и программное обеспечение для лабораторных работ

Учебная аудитория "Цифровая связь": учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий практического типа, учебная аудитория для проведения занятий лабораторного типа, учебная аудитория для проведения занятий семинарского типа, помещение для курсового проектирования (выполнения курсовых работ), помещение для проведения групповых и индивидуальных консультаций, помещение для проведения текущего контроля и промежуточной аттестации, помещение для самостоятельной работы; 634034, Томская область, г. Томск, Вершинина улица, д. 47, 309 ауд.

Описание имеющегося оборудования:
- Магнитно-маркерная доска;
- Комплект специализированной учебной мебели;
- Рабочее место преподавателя.

Программное обеспечение:
- Google Chrome;
- Microsoft Office 2010 и ниже;
- Microsoft Windows 8.1;
- Oracle VirtualBox;

8.4. Материально-техническое и программное обеспечение для самостоятельной работы

Для самостоятельной работы используются учебные аудитории (компьютерные классы), расположенные по адресам:
- 634050, Томская область, г. Томск, Ленина проспект, д. 40, 233 ауд.;
- 634045, Томская область, г. Томск, ул. Красноармейская, д. 146, 201 ауд.;
- 634034, Томская область, г. Томск, Вершинина улица, д. 47, 126 ауд.;
- 634034, Томская область, г. Томск, Вершинина улица, д. 74, 207 ауд.

Описание имеющегося оборудования:
- учебная мебель;
- компьютеры;
- компьютеры подключены к сети «Интернет» и обеспечивают доступ в электронную информационно-образовательную среду ТУСУРа.

Перечень программного обеспечения:
- Microsoft Windows;
- OpenOffice;
- Kaspersky Endpoint Security 10 для Windows;
- 7-Zip;
- Google Chrome.

8.5. Материально-техническое обеспечение дисциплины для лиц с ограниченными возможностями здоровья и инвалидов

Освоение дисциплины лицами с ограниченными возможностями здоровья и инвалидами осуществляется с использованием средств обучения общего и специального назначения.

При занятиях с обучающимися с **нарушениями слуха** предусмотрено использование звукоусиливающей аппаратуры, мультимедийных средств и других технических средств приема/передачи учебной информации в доступных формах, мобильной системы преподавания для обучающихся с инвалидностью, портативной индукционной системы. Учебная аудитория, в которой занимаются обучающиеся с нарушением слуха, оборудована компьютерной техникой, аудиотехникой, видеотехникой, электронной доской, мультимедийной системой.

Согласована на портале № 63293

При занятиях с обучающимися с **нарушениями зрения** предусмотрено использование в лекционных и учебных аудиториях возможности просмотра удаленных объектов (например, текста на доске или слайда на экране) при помощи видеоувеличителей для комфортного просмотра.

При занятиях с обучающимися с **нарушениями опорно-двигательного аппарата** используются альтернативные устройства ввода информации и другие технические средства приема/передачи учебной информации в доступных формах, мобильной системы обучения для людей с инвалидностью.

9. Оценочные материалы и методические рекомендации по организации изучения дисциплины

9.1. Содержание оценочных материалов для текущего контроля и промежуточной аттестации

Для оценки степени сформированности и уровня освоения закрепленных за дисциплиной компетенций используются оценочные материалы, представленные в таблице 9.1.

Таблица 9.1 – Формы контроля и оценочные материалы

Названия разделов (тем) дисциплины	Формируемые компетенции	Формы контроля	Оценочные материалы (ОМ)
1 Введение	ОПК-3	Зачёт	Перечень вопросов для зачета
		Тестирование	Примерный перечень тестовых заданий
		Отчет по практическому занятию (семинару)	Темы практических занятий
2 Многомерная статистика и её элементы	ОПК-3	Зачёт	Перечень вопросов для зачета
		Тестирование	Примерный перечень тестовых заданий
3 Методы статистического анализа данных	ОПК-3	Зачёт	Перечень вопросов для зачета
		Лабораторная работа	Темы лабораторных работ
		Тестирование	Примерный перечень тестовых заданий
		Отчет по лабораторной работе	Темы лабораторных работ
4 Основы нейронных сетей	ОПК-3	Зачёт	Перечень вопросов для зачета
		Практическое задание	Темы практических заданий
		Тестирование	Примерный перечень тестовых заданий

Согласована на портале № 63293

5 Методы машинного обучения	ОПК-3	Зачёт	Перечень вопросов для зачета
		Лабораторная работа	Темы лабораторных работ
		Практическое задание	Темы практических заданий
		Тестирование	Примерный перечень тестовых заданий
		Отчет по лабораторной работе	Темы лабораторных работ

Шкала оценки сформированности отдельных планируемых результатов обучения по дисциплине приведена в таблице 9.2.

Таблица 9.2 – Шкала оценки сформированности планируемых результатов обучения по дисциплине

Оценка	Баллы за ОМ	Формулировка требований к степени сформированности планируемых результатов обучения		
		знать	уметь	владеть
2 (неудовлетворительно)	< 60% от максимальной суммы баллов	отсутствие знаний или фрагментарные знания	отсутствие умений или частично освоенное умение	отсутствие навыков или фрагментарные применение навыков
3 (удовлетворительно)	от 60% до 69% от максимальной суммы баллов	общие, но не структурированные знания	в целом успешно, но не систематически осуществляемое умение	в целом успешное, но не систематическое применение навыков
4 (хорошо)	от 70% до 89% от максимальной суммы баллов	сформированные, но содержащие отдельные проблемы знания	в целом успешное, но содержащие отдельные пробелы умение	в целом успешное, но содержащие отдельные пробелы применение навыков
5 (отлично)	≥ 90% от максимальной суммы баллов	сформированные систематические знания	сформированное умение	успешное и систематическое применение навыков

Шкала комплексной оценки сформированности компетенций приведена в таблице 9.3.

Таблица 9.3 – Шкала комплексной оценки сформированности компетенций

Оценка	Формулировка требований к степени компетенции
2 (неудовлетворительно)	Не имеет необходимых представлений о проверяемом материале или Знать на уровне **ориентирования**, представлений. Обучающийся знает основные признаки или термины изучаемого элемента содержания, их отнесенность к определенной науке, отрасли или объектам, узнает в текстах, изображениях или схемах и знает, к каким источникам нужно обращаться для более детального его усвоения.

Согласована на портале № 63293

3 (удовлетворительно)	Знать и уметь на **репродуктивном** уровне. Обучающихся знает изученный элемент содержания репродуктивно: произвольно воспроизводит свои знания устно, письменно или в демонстрируемых действиях.
4 (хорошо)	Знать, уметь, владеть на **аналитическом** уровне. Зная на репродуктивном уровне, указывать на особенности и взаимосвязи изученных объектов, на их достоинства, ограничения, историю и перспективы развития и особенности для разных объектов усвоения.
5 (отлично)	Знать, уметь, владеть на **системном** уровне. Обучающийся знает изученный элемент содержания системно, произвольно и доказательно воспроизводит свои знания устно, письменно или в демонстрируемых действиях, учитывая и указывая связи и зависимости между этим элементом и другими элементами содержания дисциплины, его значимость в содержании дисциплины.

9.1.1. Примерный перечень тестовых заданий

1. Какие виды шкал относятся к качественным?
   А) наименований;
   Б) порядковая;
   В) интервалов;
   Г) отношений;
   Д) разностей;
   Е) абсолютная.
2. Основные характеристики больших данных (4V):
   А) объем, скорость создания и обработки, разнообразие источников и форм хранения, ценность;
   Б) объем, скорость создания и обработки, разнообразие источников и форм хранения, стоимость;
   В) структурированность, скорость создания и обработки, объем, ценность;
   Г) неструктурированность, скорость создания и обработки, объем, ценность
3. К какому типу информации относятся XML-документы и XSD-схемы?
   А) структурированная;
   Б) полуструктурированная;
   В) неструктурированная;
   Г) условно структурированная.
4. В чем состоит назначение задач-распределителей в технологии MapReduce?
   А) обрабатывают данные по одному ключу за раз и каким-то образом комбинируют значения, ассоциированные с этим ключом;
   Б) собирают значения, ассоциированные с каждым ключом, в список;
   В) преобразуют порцию в последовательность пар ключ-значение;
   Г) осуществляют сортировку по ключу.
5. Статистический метод анализа данных, предназначенный для исследования взаимозависимости выборок – это …
   А) кластерный анализ;
   Б) регрессионный анализ;
   В) корреляционный анализ;
   Г) факторный анализ.
6. Эти методы кластеризации последовательно объединяют, начиная с отдельных объектов, каждый раз во всё более крупные группы (кластеры) до тех пор, пока все объекты не попадают в один кластер.
   А) Иерархические дивизимные методы кластеризаци;
   Б) Иерархические агломеративные методы кластеризации;
   В) Неиерархические дивизимные методы кластеризации;
   Г) Неиерархические агломеративные методы кластеризации.
7. Если a – среднее расстояние от данного объекта до объектов из того же кластера, b –

Согласована на портале № 63293

среднее расстояние от данного объекта до объектов из ближайшего кластера (отличного от того, в котором лежит сам объект), то силуэтом данного объекта называется величина:
А) s=min(a,b).
Б) s=(b−a)/min(a,b).
В) s= max(a,b).
Г) s=(b−a)/max(a,b).

8. С какой целью в наивном Байесовском классификаторе применяется логарифмирование?
А) для решения проблемы арифметического переполнения;
Б) для устранения проблемы появления неизвестных слов;
В) для формирования вероятностного пространства;
Г) для реализации bag of words model
9. Технология классификации, где все элементарные классификаторы обучаются и работают параллельно (независимо друг от друга).
А) Бустинг;
Б) Бэггинг;
В) bag of words model;
Г) наивный Байесовский классификатор.
10. Укажите методы статистического анализа данных:
А) регрессивный;
Б) дисперсионный;
В) статистические тесты;
Г) диаграмма разброса

9.1.2. Перечень вопросов для зачета

1. Типы шкал измерений
2. Жизненный цикл данных
3. Этапы решения задачи анализа данных и их взаимосвязи.
4. Предпосылки формирования тренда больших данных
5. Определение больших данных.
6. Источники больших данных
7. Основные характеристики больших данных (4V)
8. Проблема качества и структурированности исходных данных
9. Функции и задачи больших данных, составляющие аналитики больших данных
10. Процесс и принципы аналитики Big Data,

9.1.3. Темы практических занятий

1. Шкалы измерений

9.1.4. Темы лабораторных работ

1. Модели математической статистики для предсказания поведения покупателей. Их отличия.
2. Кластерный анализ для сегментирования клиентской базы
3. Наивный байесовский классификатор для классификации твитов
4. Кластеризация на основе сетевых графов
5. Ансамбли моделей (бэггинг и бустинг)

9.1.5. Темы практических заданий

1. Формирование вероятностного пространства в NBC
2. Комплексные модели (ансамбли моделей)
3. Деревья решений
4. Бустинг: понятие и суть метода на практическом примере
5. Бэггинг: понятие и суть метода на практическом примере
6. Алгоритм «Случайный лес»: сущность, достоинства и недостатки

9.2. Методические рекомендации

Согласована на портале № 63293

Учебный материал излагается в форме, предполагающей самостоятельное мышление студентов, самообразование. При этом самостоятельная работа студентов играет решающую роль в ходе всего учебного процесса.

Начать изучение дисциплины необходимо со знакомства с рабочей программой, списком учебно-методического и программного обеспечения. Самостоятельная работа студента включает работу с учебными материалами, выполнение контрольных мероприятий, предусмотренных учебным планом.

В процессе изучения дисциплины для лучшего освоения материала необходимо регулярно обращаться к рекомендуемой литературе и источникам, указанным в учебных материалах; пользоваться через кабинет студента на сайте Университета образовательными ресурсами электронно-библиотечной системы, а также общедоступными интернет-порталами, содержащими научно-популярные и специализированные материалы, посвященные различным аспектам учебной дисциплины.

При самостоятельном изучении тем следуйте рекомендациям:

– чтение или просмотр материала осуществляйте со скоростью, достаточной для индивидуального понимания и освоения материала, выделяя основные идеи; на основании изученного составить тезисы. Освоив материал, попытаться соотнести теорию с примерами из практики;

– если в тексте встречаются незнакомые или малознакомые термины, следует выяснить их значение для понимания дальнейшего материала;

– осмысливайте прочитанное и изученное, отвечайте на предложенные вопросы.

Студенты могут получать индивидуальные консультации, в т.ч. с использованием средств телекоммуникации.

По дисциплине могут проводиться дополнительные занятия, в т.ч. в форме вебинаров. Расписание вебинаров и записи вебинаров публикуются в электронном курсе / электронном журнале по дисциплине.

9.3. Требования к оценочным материалам для лиц с ограниченными возможностями здоровья и инвалидов

Для лиц с ограниченными возможностями здоровья и инвалидов предусмотрены дополнительные оценочные материалы, перечень которых указан в таблице 9.2.

Таблица 9.2 – Дополнительные материалы оценивания для лиц с ограниченными возможностями здоровья и инвалидов

Категории обучающихся	Виды дополнительных оценочных материалов	Формы контроля и оценки результатов обучения
С нарушениями слуха	Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы	Преимущественно письменная проверка
С нарушениями зрения	Собеседование по вопросам к зачету, опрос по терминам	Преимущественно устная проверка (индивидуально)
С нарушениями опорно-двигательного аппарата	Решение дистанционных тестов, контрольные работы, письменные самостоятельные работы, вопросы к зачету	Преимущественно дистанционными методами
С ограничениями по общемедицинским показаниям	Тесты, письменные самостоятельные работы, вопросы к зачету, контрольные работы, устные ответы	Преимущественно проверка методами, определяющимися исходя из состояния обучающегося на момент проверки

9.4. Методические рекомендации по оценочным материалам для лиц с ограниченными возможностями здоровья и инвалидов

Для лиц с ограниченными возможностями здоровья и инвалидов предусматривается доступная форма предоставления заданий оценочных средств, а именно:

Согласована на портале № 63293

– в печатной форме;
– в печатной форме с увеличенным шрифтом;
– в форме электронного документа;
– методом чтения ассистентом задания вслух;
– предоставление задания с использованием сурдоперевода.

Лицам с ограниченными возможностями здоровья и инвалидам увеличивается время на подготовку ответов на контрольные вопросы. Для таких обучающихся предусматривается доступная форма предоставления ответов на задания, а именно:

– письменно на бумаге;
– набор ответов на компьютере;
– набор ответов с использованием услуг ассистента;
– представление ответов устно.

Процедура оценивания результатов обучения лиц с ограниченными возможностями здоровья и инвалидов по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:
– в форме электронного документа;
– в печатной форме увеличенным шрифтом.

Для лиц с нарушениями слуха:
– в форме электронного документа;
– в печатной форме.

Для лиц с нарушениями опорно-двигательного аппарата:
– в форме электронного документа;
– в печатной форме.

При необходимости для лиц с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения может проводиться в несколько этапов.

Согласована на портале № 63293

ЛИСТ СОГЛАСОВАНИЯ

Рассмотрена и одобрена на заседании кафедры ТОР
протокол № 21 от «15» 11 2018 г.

СОГЛАСОВАНО:

Должность	Инициалы, фамилия	Подпись
Заведующий выпускающей каф. ТОР	С.И. Богомолов	Согласовано, 645961f5-19ed-4d47-a699-64d057f3100c
Заведующий обеспечивающей каф. ТОР	С.И. Богомолов	Согласовано, 645961f5-19ed-4d47-a699-64d057f3100c
Начальник учебного управления	Е.В. Саврук	Согласовано, fa63922b-1fce-4a6a-845d-9ce7670b004c

ЭКСПЕРТЫ:

Заведующий кафедрой, каф. ТОР	Е.В. Рогожников	Согласовано, 89e0aaec-be8a-4f7b-bd1a-f43585db8135
Доцент, каф. ТОР	Д.А. Покаместов	Согласовано, 7d7b7be3-ee63-4218-8302-48c017e45ea9

РАЗРАБОТАНО:

Доцент, каф. ТОР	К.Ю. Попова	Разработано, a993ba93-27bf-4cae-844c-b4909c1fac71

Согласована на портале № 63293