

МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ
Директор института прикладной
математики и компьютерных наук
А.В. Замятин
« 15 » _____ 2023 г.



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

Статистические методы машинного обучения - I

по направлению подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки:

Big Data and Data Science

ОС составил(и):

канд. физ.-мат. наук, доцент
доцент кафедры теории вероятностей
и математической статистики

Т.В. Кабанова

Рецензент:

канд. физ.-мат. наук, доцент
доцент кафедры системного анализа
и математического моделирования

Ж.Н. Зенкова

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

Протокол от 08.06.2023 г. № 2

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

С.П. Сущенко

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

ОС разрабатывается в соответствии с рабочей программой (РП) дисциплины.

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

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения, характеризующие этапы формирования компетенций)	Критерии оценивания результатов обучения			
			Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
ОПК-1 – способен решать актуальные задачи фундаментальной и прикладной математики	ИОПК-1.3 Демонстрирует навыки использования основных понятий, фактов, концепций, принципов математики, информатики и естественных наук для решения практических задач, связанных с прикладной математикой и информатикой.	ОР-1.1.1. обучающийся сможет: - выбирать адекватные метод для решения поставленной задачи; - реализовывать выбранный метод в программе анализа данных; - делать выводы и интерпретировать полученные результаты.	Демонстрация высокого уровня знаний математических основ и базовых понятий, которые необходимы для понимания статистических методов анализа данных.	В целом успешные, но содержащие отдельные пробелы знания о математических основах и базовых понятиях, которые необходимы для понимания статистических методов анализа данных.	Фрагментарное, неполное знание без грубых ошибок математически базовых понятий, которые необходимы для понимания статистических методов анализа данных.	Не знает математические основы и базовые понятия, которые необходимы для понимания статистических методов анализа данных.

ОПК-2 – способен совершенствовать и реализовывать новые математические методы решения прикладных задач.	ИОПК-2.1 Использует результаты прикладной математики для освоения, адаптации новых методов решения задач в области своих профессиональных интересов.	ОР-2.1.1. обучающийся сможет: адаптировать модели для описания процессов реальной предметной области	Демонстрация высокого уровня знаний математических основ и базовых понятий, которые необходимы для понимания статистических методов анализа данных.	В целом успешные, но содержащие отдельные пробелы знания о математических основах и базовых понятиях, которые необходимы для понимания статистических методов анализа данных.	Фрагментарное, неполное знание без грубых ошибок математически базовых понятий, которые необходимы для понимания статистических методов анализа данных.	Не знает математические основы и базовые понятия, которые необходимы для понимания статистических методов анализа данных.
	ИОПК-2.2 Реализует и совершенствует новые методы, решения прикладных задач в области профессиональной деятельности.	ОР-2.2.1. обучающийся сможет: реализует и интерпретирует построенные модели для описания процессов реальной предметной области				
	ИОПК-2.3. Проводит качественный и количественный анализ полученного решения с целью построения оптимального варианта.	ОР-2.3.1. Обучающийся сможет проводить качественный и количественный анализ построенных моделей и полученных на их основании прогнозов и выбирать наиболее оптимальный в соответствии с выбранной метрикой				

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.	Введение в статистический анализ.	ОР-1.1.1, ОР-2.1.1, ОР-2.2.1, ОР-2.3.1.	Лабораторные работы; ответы на вопросы экзамена или теста
2.	Критерии сравнения групп.	ОР-1.1.1, ОР-2.1.1, ОР-2.2.1, ОР-2.3.1.	Лабораторные работы; ответы на вопросы экзамена или теста
3	Корреляционный анализ	ОР-1.1.1, ОР-2.1.1, ОР-2.2.1, ОР-2.3.1.	Лабораторные работы; ответы на вопросы экзамена или теста
4	Парная регрессия	ОР-1.1.1, ОР-2.1.1, ОР-2.2.1, ОР-2.3.1.	Лабораторные работы; ответы на вопросы экзамена или теста

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

3.1. Типовые задания для проведения текущего контроля успеваемости по дисциплине.

Текущий контроль осуществляется на основании посещаемости и проверки лабораторных работ.

Примеры заданий для лабораторных работ

Лабораторная работа. Предварительная обработка данных

Задание.

1. Импортировать заданный набор данных.
2. Проверить на наличие пропусков и выбросов.
3. Для количественных показателей построить гистограммы.
4. Найти оценки числовых характеристик.
5. Проверить гипотезу о нормальности.
6. Построить диаграммы размаха по группам на основании разбиения количественных показателей по уровням категориальных признаков.

Лабораторная работа. Анализ связи признаков

Выполняется в R.

Задание.

Импортировать таблицу с данными в R.

1. Построить графики для визуализации данных и их взаимосвязей.
2. Проверить связи факторов друг с другом и их влияние на зависимую целевую переменную, выбирая соответствующий критерий, в зависимости от типов данных.
3. Проверить гипотезы о значимости связи.

Лабораторная работа. Парная регрессия. Генерация.

Выполняется в R.

Задание.

1. Определить объем выборки n (от 50 до 150).
2. Сгенерировать вектор значений предсказывающей переменной.
3. Задать вектор шума, удовлетворяющий условиям Гаусса-Маркова.
4. Задать параметры регрессии.
5. Сформировать вектор значений зависимой переменной по линейной модели регрессии.
6. Построить диаграмму рассеяния и при необходимости скорректировать параметры.
7. Построить МНК-оценки параметров, проверить их значимость, сравнить с исходными значениями
8. Найти СКО остатков.
9. Проверить общую адекватность модели.

Лабораторная работа. Парная регрессия для реальных данных. Линейные и нелинейные модели.

Выполняется в R.

Задание.

Импортировать таблицу с данными в R.

1. Построить графики для визуализации данных и их взаимосвязей.
2. Проверить связь фактора с зависимой целевой переменной.
3. Построить и провести анализ парной модели регрессии целевой переменной от фактора.

4. Построить линейную, степенную, экспоненциальную, логарифмическую и обратную зависимости.
5. Оценить качество каждой модели.
6. Выбрать наиболее адекватную.

3.2. Типовые задания для проведения промежуточной аттестации по дисциплине.

1. Типы данных и способы их представления.
2. Параметрические критерии сравнения групп.
3. Непараметрические критерии сравнения групп.
4. Корреляционный анализ количественных данных.
5. Ранговая корреляция.
6. Корреляционный анализ категоризованных данных.
7. Парная регрессии. Модель. МНК-оценки параметров.
8. Числовые характеристики оценок параметров парной регрессии.
9. Теорема Гаусса-Маркова для случая парной регрессии.
10. Проверка качества уравнения парной регрессии.

4. Методические материалы, определяющие процедуры оценивания образовательных результатов обучения

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

Для текущей аттестации необходимо иметь посещаемость не менее 75% от всех проведенных на момент аттестации занятий и сдать все данный на момент аттестации лабораторные работы. Лабораторные работы оцениваются на зачтено/незачтено.

4.2. Методические материалы для проведения промежуточной аттестации по дисциплине.

Результаты экзамена определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Для теста из 15 вопросов. За каждый вопрос в зависимости от его сложности можно получить от 1 до 3 баллов. Максимально 30.

отлично	От 26 до 30 баллов
хорошо	От 21 до 25 баллов
удовлетворительно	От 16 до 20 баллов
неудовлетворительно	От 0 до 15 баллов