

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Тульский государственный университет»

Институт прикладной математики и компьютерных наук
Кафедра «Прикладная математика и информатика»

Утверждено на заседании кафедры
«Прикладная математика и информатика»
24 января 2023 г., протокол № 5

И.о. заведующего кафедрой

 Н.В. Ларин

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ (ОЦЕНОЧНЫЕ МАТЕРИАЛЫ) ДЛЯ
ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И
ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО
ДИСЦИПЛИНЕ (МОДУЛЮ)**

«Инструменты прикладной статистики»

**основной профессиональной образовательной программы
высшего образования – программы магистратуры**

по направлению подготовки

01.04.02 Прикладная математика и информатика

с направленностью (профилем)

**Перспективные методы искусственного интеллекта
в сетях передачи и обработки данных**

Форма обучения: очная

Идентификационный номер образовательной программы: 010402-03-23

Тула 2023 год

ЛИСТ СОГЛАСОВАНИЯ
фонда оценочных средств (оценочных материалов)

Разработчик:

Смирнов О.И., доцент каф. ПМиИ, к.ф.-м.н., доцент

(ФИО, должность, ученая степень, ученое звание)



(подпись)

1. Описание фонда оценочных средств (оценочных материалов)

Фонд оценочных средств (оценочные материалы) включает в себя контрольные задания и (или) вопросы, которые могут быть предложены обучающемуся в рамках текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю). Указанные контрольные задания и (или) вопросы позволяют оценить достижение обучающимся планируемых результатов обучения по дисциплине (модулю), установленных в соответствующей рабочей программе дисциплины (модуля), а также сформированность компетенций, установленных в соответствующей общей характеристике основной профессиональной образовательной программы.

Полные наименования компетенций и индикаторов их достижения представлены в общей характеристике основной профессиональной образовательной программы.

2. Оценочные средства (оценочные материалы) для проведения текущего контроля успеваемости обучающихся по дисциплине (модулю)

Перечень контрольных заданий и (или) вопросов для оценки сформированности компетенции ПК-8 (контролируемый индикатор достижения компетенции ПК-8.1)

1. Цели и задачи анализа данных в разных видах деятельности. Методология анализа данных как науки и индустрии. Связь анализа данных с фундаментальной и прикладной математикой.
2. Трехуровневая классификация аналитических задач и технологий. Уровень сбора и хранения информации. Уровень запросов к данным, описания данных и проверки гипотез. Уровень генерации новых гипотез и выявления закономерностей.
3. Понятие об инструментах прикладной статистики и фундаментальных задачах статистического распознавания.
4. Основные модели в анализе данных. Понятие модели данных. Понятие информационной модели.
5. Классификация инструментов статистики и фундаментальных задач интеллектуального анализа данных: по наличию целевых признаков, по типу признаков, по существованию распределения, по модели исходных данных.
6. Теория вероятностей и статистика как формализмы. Понятие эксперимента в теории вероятностей. Основные задачи теории вероятностей. Понятие о параметрических и полупараметрических моделях эксперимента.
7. Понятие случайной величины. Совместные распределения, маргинальные распределения, условная вероятность, теорема произведения, формула Байеса.
8. Виды распределений. Способы задания распределений, функция распределения и ее свойства, плотность распределения.
9. Дополнительные способы задания распределений: квантили, p-value.
10. Числовые характеристики распределений: матожидание, дисперсия, моменты.
11. Основные дискретные и непрерывные распределения.
12. Интерпретация вероятности (классическая, геометрическая, частотная, субъективная)

Перечень контрольных заданий и (или) вопросов для оценки сформированности компетенции ПК-8 (контролируемый индикатор достижения компетенции ПК-8.2)

1. Задача точечного оценивания.
2. Статистики как функции выборки: вариационный ряд, выборочное среднее, выборочная дисперсия, выборочная медиана.

3. Свойства точечных оценок: состоятельность, несмещенность, эффективность, робастность.
4. Принципы статистики: принцип максимального правдоподобия, принцип максимальной апостериорной вероятности, принцип максимальной обоснованности. Сравнение разных принципов на одной и той же задаче.
5. Метод максимального правдоподобия как метод получения точечных оценок.
6. Метод наименьших квадратов, его связь с методом максимального правдоподобия. Регуляризация при настройке линейных моделей регрессии: ridge, lasso, elasticnet.
7. Свойства точечных оценок в западной культуре: accuracy, precision, trueness.
8. Задача разделения смеси распределений. Идентифицируемые распределения. EM-алгоритм.
9. Задача интервального оценивания, доверительный интервал, уровень надежности.
10. Методы построения распределения точечной оценки (параметрический, наивный, бутстреп), построение доверительных интервалов по распределению точечной оценки.
11. Непараметрическое восстановление распределений, метрические методы, ядровое сглаживание.
12. Понятие классов и их традиционные наименования в статистике. Типы задач в проверке гипотез.

3. Оценочные средства (оценочные материалы) для проведения промежуточной аттестации обучающихся по дисциплине (модулю)

Перечень контрольных заданий и (или) вопросов для оценки сформированности компетенции ПК-8 (контролируемый индикатор достижения компетенции ПК-8.1)

1. Фишеровская задача распознавания, нулевая гипотеза, функция правдоподобия, р-значение и его использование, ошибки первого рода и специфичность, уровень значимости.
2. Неймановская задача распознавания, альтернативная гипотеза, отношение правдоподобия, ошибки второго рода и чувствительность, мощность критерия. Минимаксная задача распознавания, равный уровень ошибок.
3. Совместное распределение, априорная вероятность, апостериорная вероятность. Функция потерь, средний риск, байесовская задача распознавания.
4. Матрица ошибок. Основные показатели качества в задачах классификации и восстановления регрессии, доступные показатели качества в разных типах задач. Парадоксы их использования (проблема группирования, проблема редких событий), проблемы теории рационального выбора.
5. Важнейшие функции потерь, соответствующие байесовские стратегии.
6. Средние и эмпирические показатели качества.
7. Эмпирический риск, обобщающая способность стратегии. Явления недообучения и переобучения. Роль обучающей, валидационной и контрольной выборок при обучении по прецедентам. Кросс-валидация.
8. Разложение среднего риска на части, дилемма смещения-дисперсии, теоретическое обоснование ансамблей классификаторов.
9. Проверка параметрических гипотез и проверка непараметрических гипотез.
10. Многоэтапная диагностика.
11. Метрические модели в распознавании. Парзеновские окна. Метод k ближайших соседей.
12. Множественная проверка гипотез.
13. Последовательный анализ.

Перечень контрольных заданий и (или) вопросов для оценки сформированности компетенции ПК-8 (контролируемый индикатор достижения компетенции ОПК-8.2)

1. Дисперсионный анализ как инструмент статистики.
2. Корреляционный анализ как инструмент статистики.
3. Регрессионный анализ как инструмент статистики.
4. Меры качества регрессионных моделей как инструмент статистики.
5. ROC-анализ описаний объектов, индекс Джини (сравнение распределения признака между классами без согласования их размера).
6. ROC-анализ классификаторов, ROC-AUC и его свойства.
7. Отбор и генерация признаков на основе операционных характеристик признаков (информативности).
8. Перестановочные тесты.
9. Обобщенные линейные модели. Логистическая регрессия. Переход от линейных моделей к нелинейным при помощи ядерной функции.
10. Анализ временных рядов.
11. Анализ выживаемости как задача на цензурированных данных. Основы модели Кокса.
12. Теория измерений (метрологии), характеристики средства измерений, выпуклые комбинации предикторов.
13. Теория тестов, валидация шкал.