

## **АННОТАЦИЯ рабочей программы дисциплины «Б1.В.13 Многомерный статистический анализ»**

**Профиль** Современные методы машинного обучения и компьютерного зрения

**Объем трудоемкости:** 3 зачетных единицы.

**Цель дисциплины:** изучение основных методов многомерного статистического анализа данных с точки зрения их практического применения; формирование навыков работы с соответствующими разделами ППП STATISTICA.

**Задачи дисциплины:**

- помочь студентам понять и, освоить методологию многомерного статистического анализа данных;
- привить теоретические и практические знания в области прикладного многомерного анализа данных;
- познакомить студентов и обучить максимально широкому инструментарию многомерного анализа данных в среде ППП STATISTICA;
- выработать в процессе обучения у студентов навыки грамотного использования аппарата вероятно-статистического моделирования посредством применения передовых информационных технологий.

**Место дисциплины в структуре ООП ВО:**

Дисциплина «Многомерный статистический анализ» относится к дисциплинам вариативной части Блока 1 Дисциплины (модули).

Данная дисциплина тесно связана с дисциплинами: «Математический анализ», «Векторная алгебра», «Математические методы и модели исследования операций».

Материал курса предназначен для использования в дисциплинах, связанных с количественным анализом реальных экономических явлений, таких как, например, «Математические модели анализа экономических субъектов», «Теория риска и моделирование рискованных ситуаций» и др.

В совокупности изучение этой дисциплины готовит студентов, как к различным видам практической деятельности, так и к научно-теоретической, исследовательской работе.

Особо следует подчеркнуть, что многомерный статистический анализ лежит в основе методов ИИ. Ниже перечислены наиболее часто используемые разделы многомерного анализа в методах ИИ.

РСА (метод главных компонент) – уменьшение размерности.

Кластерный анализ (k-means, hierarchical clustering) – неконтролируемое обучение (без учителя).

Факторный анализ – сокращение размерности данных и выявление скрытых переменных.

Дискриминантный анализ, общие модели дискриминантного анализа, деревья классификации, логистическая регрессия – в решении задач классификации (контролируемое обучение) методами машинного обучения.

Регрессионный анализ, общие линейные модели в решении задач регрессии методами машинного обучения.

Дискриминантный, кластерный, корреляционный анализ – в выявлении признаков построения моделей машинного обучения с высокими прогностическими свойствами.

**Требования к уровню освоения дисциплины**

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

| Код и наименование индикатора*                                                                                                                              | Результаты обучения по дисциплине                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| MF-1 Способен применять современную теоретическую математику для разработки новых алгоритмов и формулирования перспективных задач искусственного интеллекта |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| MF-1.2 Применяет современный математический аппарат теории вероятностей для исследования методов и моделей машинного обучения.                              | <p>Знать: Основы теории вероятностей, математической статистики и теории информации для анализа данных ИИ</p> <p>Уметь: Оценивать распределения, выявлять зависимости, применять энтропию и взаимную информацию</p> <p>Владеть: Выбором и применением статистико-информационного аппарата для задач ИИ</p>                                                                                                                                                                                                                                                                                                                                                        |
| MF-2 Способен применять байесовский подход для построения вероятностных моделей анализа неопределенности и создания адаптивных систем ИИ                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| MF-2.2 Применяет байесовские методы оценивания и байесовские интервалы для решения задач статистики и построения моделей.                                   | <p>Знать: Определения байесовских точечных оценок (апостериорное среднее, апостериорная мода), байесовских интервалов, а также условия их применения для построения вероятностных моделей в задачах статистики и ИИ</p> <p>Уметь: Вычислять байесовские оценки параметров (апостериорное среднее, медиану, моду) и строить байесовские интервалы, интерпретировать их как вероятностные утверждения о параметре (в отличие от доверительных интервалов в частотной статистике)</p> <p>Владеть: Навыками применения байесовских методов оценивания для построения вероятностных моделей в задачах ИИ (включая биномиальные, нормальные, иерархические модели).</p> |
| BD-1 Способен осуществлять поиск сбор очистку и предварительный анализ данных                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| BD-1.3 Применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных                                                 | <p>Знать: Цели понижения размерности, методы (PCA, t-SNE, UMAP), критерии выбора числа компонент</p> <p>Уметь: Применять PCA, t-SNE, UMAP в Python (sklearn), интерпретировать нагрузки и визуализировать</p> <p>Владеть: Выбором подходящего метода (PCA или t-SNE или UMAP) для визуализации и интерпретации многомерных данных</p>                                                                                                                                                                                                                                                                                                                             |

Результаты обучения достигаются в рамках осуществления всех видов контактной и самостоятельной работы обучающихся в соответствии с утвержденным учебным планом.

Индикаторы достижения компетенций считаются сформированными при достижении соответствующих им результатов обучения.

Разделы дисциплины, изучаемые в 4 семестре:

| №<br>п/<br>п                             | Наименование раздела, темы                                                                    | Количество часов |                   |           |                      |
|------------------------------------------|-----------------------------------------------------------------------------------------------|------------------|-------------------|-----------|----------------------|
|                                          |                                                                                               | Всего            | Аудиторная работа |           | Внеаудиторная работа |
|                                          |                                                                                               |                  | Л                 | ЛР        | СРС                  |
| 1.                                       | Элементарные понятия статистики. Измерительные шкалы                                          | 4                | 2                 | 2         | –                    |
| 2.                                       | Работа с данными. Основные операции над случаями и переменными                                | 7                | 2                 | 2         | 3                    |
| 3.                                       | Основные статистики, корреляционный анализ, графика                                           | 6                | 2                 | 2         | 2                    |
| 4.                                       | Подгонка закона распределения                                                                 | 6                | 2                 | 2         | 2                    |
| 5.                                       | Сравнение средних величин критерием Стьюдента                                                 | 6                | 2                 | 2         | 2                    |
| 6.                                       | Группировка и однофакторная ANOVA                                                             | 7                | 2                 | 2         | 3                    |
| 7.                                       | Дисперсионный анализ                                                                          | 6                | 2                 | 2         | 2                    |
| 8.                                       | Непараметрическая статистика. Сравнение средних величин методами непараметрической статистики | 7                | 2                 | 2         | 3                    |
| 9.                                       | Таблицы частот, сопряженности, флагов и заголовков                                            | 8                | 2                 | 2         | 4                    |
| 10.                                      | Линейное и нелинейное моделирование взаимосвязей                                              | 4,8              | 2                 | 2         | <b>0,8</b>           |
| 11.                                      | Кластерный анализ                                                                             | 8                | 2                 | 2         | 4                    |
| 12.                                      | Дискриминантный анализ                                                                        | 8                | 2                 | 2         | 4                    |
| 13.                                      | Деревья классификации                                                                         | 7                | 2                 | 2         | 3                    |
| 14.                                      | Факторный анализ                                                                              | 9                | 2                 | 2         | 5                    |
| 15.                                      | Канонический анализ. Многомерное шкалирование                                                 | 6                | 2                 | 2         | 2                    |
| 16.                                      | Обзор пройденного материала.                                                                  | 4                | 2                 | 2         | –                    |
| Контроль самостоятельной работы (КСР)    |                                                                                               | <b>4</b>         |                   |           |                      |
| Промежуточная аттестация (ИКР)           |                                                                                               | <b>0,2</b>       |                   |           |                      |
| Подготовка к текущему контролю           |                                                                                               |                  |                   |           |                      |
| <b>Общая трудоемкость по дисциплине:</b> |                                                                                               | <b>108</b>       | <b>32</b>         | <b>32</b> | <b>39,8</b>          |

Примечание: Л – лекции, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

**Курсовые работы:** не предусмотрены

**Форма проведения аттестации по дисциплине:** зачет

Автор:

профессор кафедры анализа данных и искусственного интеллекта д-р техн. наук, доц., Халафян А.А.