

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

Институт прикладной математики и компьютерных наук

УТВЕРЖДЕНО:
Директор
А. В. Замятин

Рабочая программа дисциплины

Статистические методы машинного обучения

по направлению подготовки

09.03.03 Прикладная информатика

Направленность (профиль) подготовки:
Искусственный интеллект и большие данные

Форма обучения
Очная

Квалификация
Бакалавр

Год приема
2024

СОГЛАСОВАНО:
Руководитель ОП
С.П. Сущенко

Председатель УМК
С.П. Сущенко

Томск – 2024

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-10 Способен решать задачи в профессиональной деятельности на основе информационной и библиографической культуры, цифровых технологий и систем искусственного интеллекта.

ПК-3 Способен осуществлять научно-исследовательские и опытно-конструкторские разработки как при исследовании самостоятельных тем, так и разработки по тематике организации.

ПК-6 Способен разрабатывать и применять методы машинного обучения для решения задач.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-10.1 Выбирает, применяет и адаптирует методы исследования для решения задач профессиональной деятельности с использованием систем искусственного интеллекта

ИПК-3.1 Осуществляет проведение работ по обработке и анализу научно-технической информации и результатов исследований

ИПК-6.1 Проводит анализ требований и определяет необходимые классы задач машинного обучения

ИПК-6.2 Принимает участие в оценке и выборе используемых методов машинного обучения

2. Задачи освоения дисциплины

– Освоить аппарат основных методов машинного обучения, применяемых при анализе данных в экономике, медицине, социологии и других областях

– Научиться решать задачи статистического анализа данных с помощью моделей машинного обучения.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к Блоку 1 «Дисциплины (модули)».

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль Модуль «Введение в искусственный интеллект».

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Шестой семестр, экзамен

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: математический анализ, линейная алгебра, методы оптимизаций, теория вероятностей и математическая статистика.

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 4 з.е., 144 часов, из которых:

-лекции: 16 ч.

-практические занятия: 32 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Введение в статистический анализ и первичная статистическая обработка

Задачи и этапы статистического анализа. Типы и структуры данных.

Предварительная обработка данных.

Тема 2. Критерии сравнения групп.

Параметрические критерии. Непараметрические критерии.

Тема 3. Корреляционный анализ

Количественная корреляция. Ранговая корреляция. Корреляционный анализ количественных данных.

Тема 4. Регрессионный анализ

Парная регрессия. Множественная регрессия.

Тема 5. Дисперсионный анализ

Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ

Тема 6. Задача классификации.

Методы классификации. Логистическая регрессия.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине осуществляется на основании проверки практических заданий, выполняемых студентами на компьютерах в течение семестра и фиксируется в форме контрольной точки не менее одного раза в семестр.

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Экзамен в шестом семестре проводится в форме теста из 20 вопросов. Продолжительность экзамена 45 минут.

Оценочные материалы для проведения промежуточной аттестации размещены на сайте ТГУ в разделе «Информация об образовательной программе» - <https://www.tsu.ru/sveden/education/eduop/>.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «LMS IDO»

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

в) План практических занятий по дисциплине.

г) Методические указания по проведению лабораторных работ.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

– Джеймс Г., Уиттон Д., Хастис Е., Тибширани Р. Введение в статистическое обучение с примерами на языке R М.: ДМК Пресс, 2016. – 450 с.

– Кабанова Т. В. Применение пакета R для решения задач прикладной статистики : учебное пособие : [для студентов и аспирантов университетов]. Томск : Издательский дом Том. гос. ун-та, 2019. – 124 с.

- Марголис Н. Ю., Кабанова Т. В. Прикладная статистика: учебно-методическое пособие. Ч. 1. Том. гос. ун-т, 2007. – 46 с.
- Марголис Н. Ю., Кабанова Т. В. Прикладная статистика: учебно-методическое пособие. Ч. 2. Том. гос. ун-т, 2007. – 58 с.

б) дополнительная литература:

- М. Кендалл, А. Стьюарт Статистические выводы и связи. Наука.: Физматлит, 1973. – 432 с.
- С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин Прикладная статистика. Классификация и снижение размерности. Финансы и статистика, 1989. – 608 с.
- Айвазян С. А, Мхитарян В. С. Прикладная статистика. Основы эконометрики: Учебник для экономических специальностей вузов: В 2 т. . Т. 1. ЮНИТИ-ДАНА, 2001. – 270 с.
- Айвазян С. А. Прикладная статистика. Основы эконометрики: Учебник для экономических специальностей вузов: В 2 т. . Т. 2, ЮНИТИ-ДАНА, 2001. – 432 с.

в) ресурсы сети Интернет:

- <http://statsoft.ru/#tab-STATISTICA-link>
- <https://www.r-project.org/>
- <http://www-01.ibm.com/software/ru/analytics/spss/index.html>
- <http://itmu.vsuet.ru/Posobija/MathCAD/gl13/index.htm#anc1323>
- <http://www.exponenta.ru/>
- Общероссийская Сеть КонсультантПлюс Справочная правовая система. <http://www.consultant.ru>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

- MS Windows,
- MS Office,
- Mathcad,
- Statistica,
- R, R Studio.

б) информационные справочные системы:

- Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>
- Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>
- ЭБС Лань – <http://e.lanbook.com/>
- ЭБС Консультант студента – <http://www.studentlibrary.ru/>
- Образовательная платформа Юпайт – <https://urait.ru/>
- ЭБС ZNANIUM.com – <https://znanium.com/>
- ЭБС IPRbooks – <http://www.iprbookshop.ru/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения практических занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

15. Информация о разработчиках

Кабанова Татьяна Валерьевна, канд. физ.-мат. наук, доцент кафедры теории вероятностей и математической статистики ИПМКН ТГУ.