

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

Математическое моделирование в медицине

по направлению подготовки

01.03.02 Прикладная математика и информатика

Направленность (профиль) подготовки:
Прикладная математика и инженерия цифровых проектов

Форма обучения
Очная

Квалификация
Бакалавр

Год приема
2024

СОГЛАСОВАНО:
Руководитель ОП
Д.Д. Даммер

Председатель УМК
С.П. Сущенко

Томск – 2024

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности.

ОПК-3 Способен применять и модифицировать математические модели для решения задач в области профессиональной деятельности.

ПК-2 Способен собирать, обрабатывать и анализировать данные для проведения научно-исследовательских работ в зависимости от проблемной и предметной области, создавать математическую модель исследуемого объекта.

ПК-4 Способен проводить анализ проблемной ситуации заинтересованных лиц, разрабатывать концепции, техническое задание на разработку в рамках цифрового проекта, сопровождать мероприятия по вводу в эксплуатацию готовых решений.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-1.1 Демонстрирует навыки работы с учебной литературой по основным естественнонаучным и математическим дисциплинам.

ИОПК-1.2 Демонстрирует навыки выполнения стандартных действий, решения типовых задач с учетом основных понятий и общих закономерностей, формулируемых в рамках базовых математических и естественнонаучных дисциплин.

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.

ИОПК-1.4 Демонстрирует понимание и навыки применения на практике математических моделей и компьютерных технологий для решения практических задач, возникающих в профессиональной деятельности

ИОПК-3.1 Демонстрирует навыки применения современного математического аппарата для построения адекватных математических моделей реальных процессов, объектов и систем в своей предметной области.

ИОПК-3.2 Демонстрирует умение собирать и обрабатывать статистические, экспериментальные, теоретические и т.п. данные для построения математических моделей, расчетов и конкретных практических выводов.

ИОПК-3.3 Демонстрирует способность критически переосмысливать накопленный опыт, модифицировать при необходимости вид и характер разрабатываемой математической модели.

ИОПК-3.4 Демонстрирует понимание и умение применять на практике математические модели и компьютерные технологии для решения различных задач в области профессиональной деятельности.

ИПК-2.1 Осуществляет сбор, обработку, анализ данных в исследуемой предметной области

ИПК-2.2 Способен строить математическую модель исследуемого объекта и/или процесса в зависимости от проблемной и предметной области

ИПК-2.3 Осуществляет проведение экспериментов, внедрение результатов исследований в соответствии с установленными полномочиями

ИПК-2.4 Составляет отчеты (разделы отчетов) по теме или по результатам проведенных исследований

ИПК-4.1 Способен проводить анализ проблемной ситуации заинтересованных лиц, определять задачи, которые могут быть решены математическими методами и/или автоматизацией процессов в зависимости от предметной области

ИПК-4.2 Способен формализовывать, согласовывать и документировать требования к системе и подсистеме в проблемной области

2. Задачи освоения дисциплины

– Освоить аппарат построения интеллектуальных систем в здравоохранении и медицине на базе машинного обучения.

– Научиться применять понятийный аппарат интеллектуальных систем с использованием инструментария библиотек Python, R, публичных облачных сервисов, оценивать эффективность их работы и внедрять в приложения для решения практических задач здравоохранения и медицины.

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 «Прикладная математика и инженерия цифровых проектов».

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

Седьмой семестр, зачет с оценкой

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: «Статистические методы машинного обучения», «Введение в интеллектуальный анализ данных».

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 16 ч.

-лабораторные: 16 ч.

в том числе практическая подготовка: 16 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Разведочный анализ данных. Предварительная обработка в биомедицинских данных.

Анализ признакового пространства. Описательные статистики. Обработка категориальных данных. Обработка пропущенных значений. Поиск аномалий в данных. Преобразование признаков к единой шкале. Балансировка классов.

Тема 2. Решение задач классификации

Постановка задачи классификации. Метрики оценки качества моделей классификации. Алгоритм К-ближайших соседей. Линейные модели. Наивный Байес. Логистическая регрессия. Метод опорных векторов. Ядерный трюк. Деревья решений.

Тема 3. Регрессионные модели в медицине

Постановка задачи регрессии. Метрики оценки качества моделей регрессии. Линейная регрессия. Полиномиальная регрессия. Деревья решений для регрессии. Регрессоры на основе K-ближайших соседей. Метод опорных векторов для регрессии.

Тема 4. Поиск лучших моделей

K-блочная перекрестная проверка. Случайный лес. Бэггинг. Голосование моделей. Бустинг. Стекинг. Поиск гиперпараметров. Автоматизация поиска моделей.

Тема 5. Оптимизация признакового пространства в биомедицинских данных

Снижение размерности. Backward Elimination. Forward Selection. Метод главных компонент. Факторный анализ. Линейный дискриминантный анализ. Применение алгоритма SVM для преобразования признакового пространства. SVD. Усеченное сингулярное разложение. Многомерное масштабирование. Изометрическое картографирование (Isomap). Локальное линейное владение (LLE). t-распределенное стохастическое вложение соседей (t-SNE). Разномерное приближение и проекция многообразия (UMAP).

Тема 6. Кластеризация многомерных данных

Задача кластеризации. Графы и гиперграфы. Жёсткая и мягкая кластеризация. K-средних. Метод локтя. Анализ значений Силуэта и Силуэтных графиков. Статистический подход к кластеризации. EM алгоритм. Определение числа кластеров. Кластеризация, основанная на плотности (DBSCAN). OPTICS. Кластеризация категориальных данных. ROCK. CACTUS. Иерархическая кластеризация. Дивизивные и агломеративные методы. Спектральная кластеризация. Корреляционная и консенсусная кластеризация. Внутренние и внешние метрики оценки качества кластеризации. Относительные методы оценки качества кластеризации.

Тема 7. Анализ временных последовательностей и биомедицинских сигналов

Временные последовательности. Преобразование временных серий. ARMA модели для стационарных временных последовательностей. ARMA модели для нестационарных временных последовательностей. Определение трендов. Прогноз с помощью ARIMA. Модели ненаблюдаемых компонентов, извлечение сигналов и фильтры. Сезонность и экспоненциальное сглаживание. Волатильность и обобщенные авторегрессионные условные гетероскедастические процессы. Нелинейные случайные процессы. Передаточные функции и авторегрессионное распределенное моделирование задержек. Векторные авторегрессии и причинность по Грейнджеру. Векторные авторегрессии с интегрированными переменными, модели коррекции векторных ошибок и общие тенденции. Композиционный и счетный временной ряд.

Тема 8. Объяснимый искусственный интеллект

Глобальная и локальная интерпретируемость. Графики частичной зависимости. Индивидуальное условное ожидание. График накопленных локальных эффектов (ALE). Важность признаков для моделей, основанных на деревьях. Аддитивные объяснения Шепли (SHAP). Локальные интерпретируемые объяснения независимые от модели (LIME). Профиль Ceteris-paribus. Осцилляции Ceteris-paribus. Локальные диагностические диаграммы. Использование якорей. Понимание о семантической схожести. Контрфактуальные объяснения. Контрастные объяснения. Количественное тестирование с векторами активации концепций (TCAV).

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, тестов по лекционному материалу, выполнения домашних заданий 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

Зачет с оценкой в седьмом семестре проводится в письменной форме по билетам. Экзаменационный билет состоит из трех частей. Продолжительность зачета с оценкой 1 час.

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 «Moodle» - <https://moodle.tsu.ru/course/view.php?id=00000>

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

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

1) Sudip Mandal (Editor). Artificial Intelligence (AI) in Cell and Genetic Engineering. – Humana Press. Springer Protocols. Methods in Molecular Biology 2952. New York, 2025. ISBN 978-1-0716-4989-2

2) Dukka B.KC (Editor). Large Language Models (LLMs) in Protein Bioinformatics. – Humana Press Springer Protocols. Methods in Molecular Biology 2941. New York, 2025. ISBN 978-1-0716-4622-9

3) Muzafar A. Macha, Assif Assad, Muzafar Rassol Bhat (Editors) Artificial Intelligence in Human Health and Diseases. Springer, Singapore, 2025. ISBN 978-981-96-8175-4

4) Potter C.Chang. Data Analysis of Medical Studies. Reading and Reporting. Springer, Los Angeles, 2025. ISBN 978-3-031-49983-8

5) Saiyed Umed, Ranjeet Kumar Rout, Monika Khandelwal, Smitarani Pati (Editors) Computational Techniques for Biological Sequence Analysis. CRC Press. Taylor & Francis Group, Boca Raton, 2025. ISBN 978-1-032-63026-7

6) Charu C. Aggarwal. Neural Networks and Deep Learning. Springer, Cham, 2023. ISBN 978-3-031-29641-3

7) Mehdi Ghayoumi. Mathematical Foundations for Deep Learning. CRC Press. Taylor & Francis Group, Boca Raton, 2026. ISBN 978-1-032-69073

8) Hui Jiang. Machine Learning Fundamentals. A Concise Introduction. Cambridge University Press, New York, 2021. ISBN 978-1-108-83-704-0

9) Ajay Kumar, Deepak Dembla, Seema Tinker, Surbhi Bhatia Khan (Editors). Handbook of Deep Learning Models for Healthcare Data Processing. Disease Prediction, Analysis, and Applications. CRC Press, Taylor & Francis Group, Boca Raton, 2026. ISBN 978-1-032-73939-7

10) Pin-Yu Chen, Sijia Liu. Introduction to Foundation Models. Springer, Cham, 2025. ISBN 978-3-031-76769-2

11) Dimitrious P. Panagoulas, George A. Tsihrintzis, Maria Virvou (Editors). Artificial Intelligence-Empowered Bio-medical Applications. Challenges, Solutions and Development Guidelines. Springer, Cham, 2025. ISBN 978-3-031-90173-7

12) Kishor Kumar Reddy C, Anindya Nag (Editors). Federated Learning for Neural Disorders in Healthcare 6.0. CRC Press, Taylor & Francis Group, Boca Raton, 2025. ISBN 978-1-032-96887-2

13) Nima Rezaei (Editor). Artificial Intelligence and Bioinformatics in Cancer: An Interdisciplinary Approach. Springer, Cham, 2025. ISBN 978-3-031-92205-3

14) Thomas P. Trappenberg. Fundamentals of Computational Neuroscience. Oxford University Press, Oxford, 2023. ISBN 978-0-19-286936-4

15) D.P. Acharjiya, Kun Ma (Editors). Computational Intelligence in Healthcare Informatics. Springer, Studies in Computational Intelligence 1132, Singapore, 2024, ISBN 978-981-99-88-52-5

б) дополнительная литература:

1) Bioinformatics with Python Cookbook. Packt Publishing. Birmingham, 2025. ISBN 978-1-83664-275-6

2) Dr. Parul Verma, Dr. Shahnaz Fatima. Python for Bioinformatics. Using machine learning for drug discovery, cluster analysis, and phylogenetics. BPB Publications, 2025. ISBN 978-93-65897-760

3) Alice Paul. Mastering Health Data Science Using R. CRC Press. Taylor & Francis Group. Boca Raton, 2025. ISBN 978-1-032-72936-7

4) Michael Korvink. Practical Healthcare Statistics with Examples in Python and R. A Guide for the Uninitiated. CRC Press. Taylor & Francis Group. Boca Raton, 2025. ISBN 978-1-041-00423-3

5) Ken Youens-Clark. Reproducible Bioinformatics with Python. O'Reilly Media, Sebastopol, 2022. ISBN 978-1-098-10088-9

6) Antonio Di Cecco, Leonida Gianfagna. Explainable AI with Python. Springer, Cham, 2023. ISBN 978-3-031-92228-2

7) Charles Ravarani, Natasha Latysheva. Deep Learning for Biology. Harness AI to Solve Real-World Biology Problems. O'Reilly Media, Santa Rosa, 2025.

8) Tanupriya Choudhuty, Avita Katal (Editors). Machine Learning for Disease Detection, Prediction, and Diagnosis. Springer, Singapore, 2025. ISBN 978-981-96-4240-3

9) Daniel Sorensen. Statistical Learning in Genetics. An Introduction Using R. Springer, Cham, 2025. ISBN 978-3-031-86273-1

10) Wan Azani Mustafa, Hiam Alquran (Editors). Advances in Medical Image Processing, Segmentation and Classification. MDPI AG, Basel, 2025. ISBN 978-3-7258-4123-3

11) Roebbe Wuenschiers. Computational Biology. A Practical Introduction to Bio Data Juggling with Worked Examples. Springer, Cham, 2025. ISBN 978-3-031-70313-3

12) Dr. Patanjali Kashyap (Editor). Machine Learning for Decision Makers. Cognitive Computing Fundamentals for Better Decision Making. Apress, New York, 2024. ISBN 978-1-4842-9800-8

13) Rangaraj M. Rangayyan, Sridhar Krishnan. Biomedical Signal Analysis. IEEE Press, Piscataway, 2024. ISBN 978-1-119-82585-2

14) Aston Zhang, Zachary C., Mu Li, Alexander J. Smola. Dive into Deep Learning, 2023. <https://d2l.ai/index.html>

15) Stephen S.-T. Yau, Xin Zhao, Kun Tian, Hongyu Yu. Mathematical Principles in Bioinformatics. Springer, Interdisciplinary Applied Mathematics 58, Cham, ISBN 978-3-031-48294-6

в) ресурсы сети Интернет:

– The AI community building the future. The platform where the machine learning community collaborates on models, datasets, and applications. – <https://huggingface.co/>

– OpenAI. – <https://openai.com/>

- Tensorflow. An end-to-end platform for machine learning. – <https://www.tensorflow.org/>
- PyTorch documentation. – <https://pytorch.org/>
- IBM. What is deep learning? – <https://www.ibm.com/topics/deep-learning>
- Общероссийская Сеть КонсультантПлюс Справочная правовая система. <http://www.consultant.ru>

13. Перечень информационных технологий

- а) лицензионное и свободно распространяемое программное обеспечение:
 - Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office OneNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);
 - публично доступные облачные технологии (Google Docs, Google Colab, Яндекс диск).
 - Пакет Anaconda
 - Средства языков программирования и анализа данных R и Python
 - Библиотеки для машинного и глубокого обучения: Scikit-learn, NumPy, Matplotlib.pyplot, Seaborn, PyTorch, Keras/TensorFlow, OpenAI Gym.
- 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 Юрайт – <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 занятий семинарского типа, 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.
- Аудитории для проведения занятий лекционного и семинарского типа индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации в смешенном формате («Актру»).

15. Информация о разработчиках

Аксёнов Сергей Владимирович, кандидат технических наук, доцент кафедры теоретических основ информатики, Институт прикладной математики и компьютерных наук, Национальный исследовательский Томский государственный университет