

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

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ:
Директор



А. В. Замятин

« 16 » мая 20 22 г.

Рабочая программа дисциплины

Введение в биоинформатику

по направлению подготовки

01.03.02 Прикладная математика и информатика

Направленность (профиль) подготовки :

Интеллектуальный анализ больших данных

Форма обучения

Очная

Квалификация

Бакалавр

Год приема

2022

Код дисциплины в учебном плане: Б1.В.ДВ.01.02.01

СОГЛАСОВАНО:

Руководитель ОП

А.В. Замятин

Председатель УМК

С.П. Сущенко

Томск – 2022

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

– ПК-4 – способность 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:

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.

ИПК-4.3 Способен проводить оценку соответствия требованиям существующих систем и их аналогов, осуществлять постановку задачи на разработку требований к системам и контролировать их качество.

2. Задачи освоения дисциплины

Цели дисциплины: дать студентам знания в области методов биоинформатики.

Изучение дисциплины предполагает выполнение следующих задач:

- знакомство с основными понятиями и методами биоинформатики, машинного обучения;
- знакомство с технологиями выделения информативных признаков в задачах биоинформатики на основе методов машинного обучения;
- знакомство с подходами к разработке предиктивных моделей в задачах биоинформатики на основе методов машинного обучения.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль «Специализация» («Биоинформатика и биомедицина»).

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

Второй семестр, зачет

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Информационная безопасность и работа с персональными данными», «Математические методы и модели для компьютерных наук»

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 2 з.е., 72 часа, из которых:

-лекции: 8 ч.

-практические занятия: 16 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Основы методов машинного обучения в задачах биоинформатики

Краткое содержание темы. Общие сведения и тенденции развития методов машинного обучения в задачах биоинформатики.

Тема 2. Методы выделения информативных признаков в задачах биоинформатики

Краткое содержание темы. Обзор и основные сведения о методах выделения информативных признаков в задачах биоинформатики, включая одномерные и многомерные методы. Достоинства и недостатки конкретных методов, практики применения.

Тема 3. Экстракция признаков, обнаружение выбросов, удаление шумов.

Краткое содержание темы. Обзор и основные сведения о методах экстракции, обнаружения выбросов и удаления шумов в задачах биоинформатики. Достоинства и недостатки конкретных методов, практики применения.

Тема 4. Методы кластеризации и создания предиктивных моделей в задачах биоинформатики

Краткое содержание темы. Обзор и основные сведения о методах кластеризации и создания предиктивных моделей в задачах биоинформатики. Достоинства и недостатки конкретных методов, практики применения. Методы валидации и оценки точности полученных моделей.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, устных опросов по материалам лекций, устных опросов по материалам практических занятий, устных опросов по материалам СРС, и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Для итоговой аттестации (зачет) студент должен успешно выполнить все практические задания в рамках курса, а также успешно пройти устный опрос.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=22112>

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

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

– Y.V. Kistenev; A.V. Borisov; D.A. Vrazhnov. Medical Applications of Laser Molecular Imaging and Machine Learning. Volume: PM333. SPIE. 2021. ISBN: 9781510645349

– Курс лекций К. Воронцова по машинному обучению (<http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>)

б) дополнительная литература:

– Bellman, R. E., Dynamic Programming, s.l.: Princeton University Press (2003).

– Shapiro, L., G. C. Stockman, Computer vision, Prentice Hall, (2002).

– Рашка С. Python и машинное обучение. – Litres, 2021.

в) ресурсы сети Интернет:

- Открытые онлайн-курсы
- Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных - <http://www.machinelearning.ru>.

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

- ПО Microsoft Visual Studio Code версии 1.61 и выше, язык Python версии 3.8 или выше, с набором библиотек sklearn, matplotlib, scipy, pandas актуальных версий
- публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

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>
- Электронная база данных научных статей Google Scholar – <https://scholar.google.com/>

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.

15. Информация о разработчиках

Вражнов Денис Александрович, младший научный сотрудник, Научное управление, лаборатория лазерного молекулярного имиджинга и машинного обучения