

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

Прикладные аспекты машинного обучения - I

по направлению подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки :

Big Data and Data Science

Форма обучения

Очная

Квалификация

Магистр

Год приема

2022

Код дисциплины в учебном плане: Б1.О.07

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:

- УК-1 – способность осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий;
- ПК-6 – способность управлять получением, хранением, передачей, обработкой больших данных.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИУК-1.1 Выявляет проблемную ситуацию, на основе системного подхода осуществляет её многофакторный анализ и диагностику.

ИУК-1.2 Осуществляет поиск, отбор и систематизацию информации для определения альтернативных вариантов стратегических решений в проблемной ситуации.

ИУК-1.3 Предлагает и обосновывает стратегию действий с учетом ограничений, рисков и возможных последствий.

ИПК-6.1 Осуществляет мониторинг и оценку производительности обработки больших данных.

ИПК-6.2 Использует методы и инструменты получения, хранения, передачи, обработки больших данных.

ИПК-6.3 Разрабатывает предложения по повышению производительности обработки больших данных.

2. Задачи освоения дисциплины

Цель дисциплины – обучить студентов осуществлять работы по исследованию больших данных с применением технологий искусственного интеллекта и больших данных и разрабатывать интеллектуальные системы с использованием инструментария библиотек Python, R, публичных облачных сервисов, оценивать эффективность их работы и внедрять в приложения;

Задачи дисциплины:

1. Научить студентов выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий искусственного интеллекта и больших данных;

2. Научить студентов принципам планирования и организации аналитических работ с использованием технологий искусственного интеллекта и больших данных;

3. Научить студентов подготавливать данные для проведения аналитических работ по исследованию больших данных методами искусственного интеллекта;

4. Научить студентов проводить аналитическое исследование и разрабатывать приложения с применением технологий искусственного интеллекта и больших данных в соответствии с требованиями заказчика.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к обязательной части образовательной программы.

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

Второй семестр, экзамен.

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Интеллектуальные системы», «Нейронные сети».

6. Язык реализации

Английский.

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 5 з.е., 180 часов, из которых:

-лекции: 20 ч.

-лабораторные: 40 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Разведочный анализ данных

Предварительный анализ данных. Визуализация. Проведение разведочного анализа данных.

Тема 2. Классификационные и регрессионные модели

Обучение классификаторов. Обучение регрессоров. Оценка качества обучения. Ознакомление с программными инструментами проектирования систем машинного обучения. Построение простых моделей классификации и регрессии.

Тема 3. Ансамбли моделей

Подбор гиперпараметров моделей. Ансамбли моделей. Построение ансамблей моделей машинного обучения для решения задач классификации.

Тема 4. Работа с признаковым пространством

Снижение размерности признакового пространства. Конструирование признаков. Кластеризация данных. Оценка информативности признаков и снижение размерности признакового пространства для анализа объектов из баз данных.

Тема 5. Основы нейросетевых вычислений

Архитектуры нейронных сетей для разнородных задач анализа данных. Применение нейронных сетей для решения задач классификации и регрессии. Проектирование нейросетевых моделей классификации и регрессии.

Тема 6. Обработка изображений сверточными нейронными сетями

Сверточные нейронные сети. Визуализация параметров и активности нейронной сети. Проектирование нейросетевых моделей классификации изображений.

Тема 7. Автокодировщики

Передача обучения. Основы работы с автокодировщиками. Повышение эффективности работы глубоких нейронных сетей. Проектирование нейросетевых моделей сегментации изображений.

Тема 8. Анализ сигналов и временных рядов

Векторное представление текстовых данных. Рекуррентные нейронные сети. Проектирование нейросетевых моделей анализа временных рядов.

Тема 9. Использование генеративно- конкурирующих моделей

Основы генеративно-конкурирующих моделей. Реализация генеративной нейросетевой модели.

Тема 10. Практические аспекты использования обучения с подкреплением
Обучение с подкреплением. Построение систем искусственного интеллекта с помощью обучения с подкреплением.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, тестов по лекционному материалу и фиксируется в форме контрольной точки не менее одного раза в семестр.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Теоретический материал по дисциплине дается в виде лекций с применением стандартных средств демонстрации мультимедиа в формате .pdf. На практических занятиях студенты решают задачи по построению систем искусственного интеллекта с помощью средств библиотек Python и R. Текущий контроль по практическим работам осуществляется в виде обсуждения алгоритма и результатов его работы.

При сдаче каждой практической работы проверяется достижение компетенции ИПК-3.1, закрепленной за дисциплиной.

Итоговая оценка выставляется как среднеарифметическое по результатам контрольных и практических работ с округлением до ближайшего целого.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=221290>

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

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

– Джоэл Грас. Data Science: Наука о данных с нуля. 2-е издание. СПб: БХВ-Петербург, 2021.

– Себастьян Рашка, Вахид Мирджалили. Python и машинное обучение. М.: Диалектика, 2020.

– Ameet V. Joshi. Machine Learning and Artificial Intelligence. Springer Nature Switzerland AG, 2020.

– Denis Rothman. Artificial Intelligence by Example. Second Edition. Packt Publishing, 2020.

– Stuart Russel, Peter Norvig. Artificial Intelligence. A Modern Approach. 4th Edition. Hoboken: Pearson, 2021.

– Stuart Russel, Peter Norvig. Artificial Intelligence. A Modern Approach. 4th Edition. Hoboken: Pearson, 2021.

– Эндрю Гласснер. Глубокое обучение без математики. Том 1. Основы. М.: ДМК Пресс, 2020.

– Эндрю Гласснер. Глубокое обучение без математики. Том 2. Основы. М.: ДМК Пресс, 2020.

– Ян Гудфеллоу, Йошуа Бенджио, Аарон Курвилль. Глубокое обучение. Второе цветное издание, исправленное. М.: ДМК Пресс, 2018.

– Roman Shirkin. Artificial Intelligence. The Complete Beginners' Guide to Artificial Intelligence. ISBN: 9798609154415. Amazon KDP Printing and Publishing, 2020.

– Франсуа Шолле. Глубокое обучение на Python. СПб: Питер, 2018.

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

- Anaconda (Python, R), библиотеки для глубокого обучения Tensorflow, Keras, Pytorch, сервисы OpenAI.
- публично доступные облачные технологии (Google Colab.).

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>

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

Аксёнов Сергей Владимирович, канд. техн. наук, доцент кафедры теоретических основ информатики.