

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

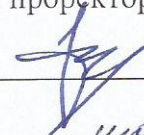
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«УФИМСКИЙ ГОСУДАРСТВЕННЫЙ АВИАЦИОННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Кафедра *высокопроизводительных вычислительных технологий и систем*

УТВЕРЖДАЮ

Первый проректор по науке


_____ Р.Д. Еникеев

« 23 » *июль* 2022 г.

РАБОЧАЯ ПРОГРАММА

УЧЕБНОЙ ДИСЦИПЛИНЫ

«СОВРЕМЕННЫЕ ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ»

Уровень подготовки

высшее образование - подготовка научных и научно-педагогических кадров в аспирантуре

Научная специальность

1.2.2 Математическое моделирование, численные методы и комплексы программ

Квалификация (ученая степень): кандидат наук


Форма обучения



очная


Уфа 2022

Рабочая программа учебной дисциплины «*СОВРЕМЕННЫЕ ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ*»

Рабочая программа дисциплины обсуждена на заседании кафедры ВВТиС 18.04.2022 г., протокол № 12 и рекомендована к реализации в образовательном процессе для подготовки аспирантов по ПА 1.2.2 «Математическое моделирование, численные методы и комплексы программ».

И.о. заведующего кафедрой:  А.А. Гайнетдинова

Составители:  Р.К. Газизов, д.ф.-м.н., профессор, профессор кафедры ВВТиС,
 В.О. Лукашук, к.ф.-м.н., доцент кафедры ВВТиС

Согласовано:  Р.К. Фаттахов, к.т.н., доцент, начальник ОАиД

Содержание

| | |
|---|----|
| 1. Место дисциплины в структуре образовательной программы | 4 |
| 2. Содержание и структура дисциплины (модуля)..... | 4 |
| 3. Учебно-методическое обеспечение самостоятельной работы аспирантов..... | 6 |
| 4. Фонд оценочных средств | 6 |
| 5. Учебно-методическое и информационное обеспечение дисциплины | 10 |
| 6. Адаптация рабочей программы для лиц с ОВЗ | 11 |

1. Место дисциплины в структуре образовательной программы

Дисциплина Современные технологии машинного обучения является частью подготовки научных и научно-педагогических кадров высшей квалификации, образовательного компонента (факультативные дисциплины) программы аспирантуры подготовки научных и научно-педагогических кадров в аспирантуре по научной специальности 1.2.2 Математическое моделирование, численные методы и комплексы программ.

Рабочая программа составлена в соответствии с Федеральными государственными требованиями к структуре программам подготовки научных и научно-педагогических кадров в аспирантуре (адъюнктуре)», утвержденных приказом Министерства науки и высшего образования Российской Федерации (Минобрнауки России) от 20 октября 2021 года № 951; Постановление Правительства Российской Федерации от 30.11.2021 № 2122 "Об утверждении Положения о подготовке научных и научно-педагогических кадров в аспирантуре (адъюнктуре)".

Целью освоения дисциплины является углубление фундаментальных знаний обучающегося, а также его практической подготовки в области математических основ и современных технологий машинного обучения.

Задачи: углубленное изучение математических основ современного машинного обучения, освоение методов и алгоритмов интеллектуального анализа данных, изучение теоретических основ и формирование практических навыков создания и обучения искусственных нейронных сетей, формирование практических навыков использования технологий машинного обучения для решения прикладных задач в различных областях профессиональной деятельности.

2. Содержание и структура дисциплины (модуля)

Общая трудоемкость дисциплины составляет 3 зачетных единиц (108 часов).

Трудоемкость дисциплины по видам работ

| Вид работы | Трудоемкость, час. |
|--|--------------------|
| | 2 курс |
| Лекции (Л) | 14 |
| Практические занятия (ПЗ) | 6 |
| Лабораторные работы (ЛР) | |
| КСР | 1 |
| Курсовая проект работа (КР) | |
| Расчетно - графическая работа (РГР) | |
| Самостоятельная работа (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам, рубежному контролю и т.д.) | 78 |
| Подготовка и сдача экзамена | |
| Подготовка и сдача зачета | 9 |
| Вид итогового контроля (зачет, экзамен) | зачет |

Содержание разделов и формы текущего контроля

| № | Наименование и содержание раздела | Количество часов | | | | | Литература, рекомендуемая аспирантам* | |
|---|--|-------------------|----|----|-----|-----|---|----------------------------|
| | | Аудиторная работа | | | | СРС | | Всего |
| | | Л | ПЗ | ЛР | КСР | | | |
| 1 | Математические основы машинного обучения | 4 | | | | 14 | 18 | 5.1.1, 5.1.2, 5.1.3 |
| 2 | Классификация и кластеризация | 2 | 2 | | | 16 | 20 | 5.1.1, 5.1.2, 5.1.3, 5.4.1 |
| 3 | Искусственные нейронные сети | 4 | 2 | | 1 | 24 | 31 | 5.1.2, 5.1.3, 5.1.4, 5.4.1 |
| 4 | Прикладные задачи машинного обучения | 4 | 2 | | | 24 | 30 | 5.1.2, 5.1.3, 5.4.1 |

_____.

3. Учебно-методическое обеспечение самостоятельной работы аспирантов

Тема 1 Математические основы машинного обучения

Вопросы для самостоятельного изучения (подготовке к обсуждению):

1. Метод стохастического градиента.
2. Сингулярное разложение и регуляризация.
3. Нелинейная регрессия.
4. Градиентный бустинг.
5. Метод случайных подпространств.
6. Методы восстановления плотности распределения.

Тема 2 Классификация и кластеризация

Вопросы для самостоятельного изучения:

1. Решающие деревья.
2. Генеративные и дискриминативные модели классификации.
3. Алгоритм k-средних и EM-алгоритм для разделения гауссовской смеси.
4. Агломеративная кластеризация, Алгоритм Ланса-Вильямса и его частные случаи.
5. Модели внимания на графах. Задача классификации вершин графа.

Тема 3 Искусственные нейронные сети

Вопросы для самостоятельного изучения:

1. Рекуррентные нейронные сети (RNN), особенности их обучения.
2. Нейронные сети долгой кратковременной памяти.
3. ResNet: остаточная нейронная сеть.
4. Обучение нейронных сетей без учителя с использованием привилегированной информации.
5. Метод SARSA. Метод Q-обучения.

Тема 4 Прикладные задачи машинного обучения

Вопросы для самостоятельного изучения:

1. Векторные представления графов.
2. Задача тематического моделирования коллекции текстовых документов. Метод максимума правдоподобия.
3. Задачи коллаборативной фильтрации, транзакционные данные.
4. Задача прогнозирования временных рядов. Примеры приложений.

4. Фонд оценочных средств

Оценка уровня освоения дисциплины осуществляется в виде текущего и промежуточного контроля успеваемости аспирантов университета, и на основе критериев оценки уровня освоения дисциплины.

Активность обучающегося оценивается на занятиях и на основе выполненных работ и заданий, предусмотренных ФОС дисциплины.

Оценивание проводится преподавателем независимо от наличия или отсутствия

обучающегося (по уважительной или неуважительной причине) на занятии. Оценка носит комплексный характер и учитывает достижения обучающегося по основным компонентам образовательного процесса за текущий период.

| № п/п | Контролируемые разделы (темы) дисциплины | Наименование оценочного средства* |
|-------|--|-----------------------------------|
| 1 | Математические основы машинного обучения | Круглый стол |
| 2 | Классификация и кластеризация | Круглый стол, комплексное задание |
| 3 | Искусственные нейронные сети | Круглый стол |
| 4 | Прикладные задачи машинного обучения | Круглый стол, комплексное задание |

Критерии оценки:

- оценка «зачтено» выставляется аспиранту, если была проявлена дискуссионная активность в рамках круглого стола, задания решены корректно;

- оценка «не зачтено» - отсутствовала дискуссионная активность в рамках круглого стола, задания решены некорректно.

Типовые оценочные материалы

Раздел (тема) дисциплины Математические основы машинного обучения

Перечень дискуссионных тем для круглого стола (дискуссии, полемики, диспута, дебатов)

1. Эвристики.
2. Логистическая регрессия. Принцип максимума правдоподобия и логарифмическая функция потерь.
3. Задача квадратичного программирования и двойственная задача.
4. Метод главных компонент и декоррелирующее преобразование Карунена-Лоэва, его связь с сингулярным разложением.
5. Робастная регрессия, функции потерь с горизонтальными асимптотами.

Критерии оценки:

- оценка «зачтено» выставляется аспиранту, если была проявлена дискуссионная активность в рамках круглого стола;
- оценка «не зачтено» - отсутствовала дискуссионная активность в рамках круглого стола.

Раздел (тема) дисциплины Классификация и кластеризация

Перечень дискуссионных тем для круглого стола (дискуссии, полемики, диспута, дебатов)

1. Метрические методы классификации.
2. Критерии качества классификации.
3. Логические методы классификации.
4. Байесовская теория классификации.
5. Постановка задачи кластеризации. Примеры прикладных задач. Типы кластерных структур.

Комплексное задание:

Дана база данных рукописных цифр, изображения которых центрированы и нормированы по размеру. В качестве примера может быть использована база данных MNIST из 60000 изображений обучающей выборки и 10000 изображений тестовой выборки, доступная по адресу <http://yann.lecun.com/exdb/mnist/>. Необходимо решить задачу кластеризации, т.е. разделить изображения на группы одинаковых цифр. Для этого выполнить следующие действия.

- 1) Реализовать программно алгоритмы k средних, плотностной алгоритм пространственной кластеризации с присутствием шума (DBSCAN) и любой метод иерархической кластеризации.
- 2) С использованием программных реализаций провести кластеризацию заданного набора данных.
- 3) Для алгоритма иерархической кластеризации построить график зависимости межкластерного расстояния от числа кластеров и определить оптимальный порог разделения.
- 4) Выполнить анализ результатов работы алгоритмов с использованием заданной метрики и сделать вывод об их эффективности.

Критерии оценки:

- оценка «зачтено» выставляется аспиранту, если была проявлена дискуссионная активность в рамках круглого стола и задания выполнены без ошибок;
- оценка «не зачтено» - отсутствовала дискуссионная активность в рамках круглого стола и/или задания выполнены с ошибками.

Раздел (тема) дисциплины Искусственные нейронные сети**Перечень дискуссионных тем для круглого стола (дискуссии, полемики, диспута, дебатов)**

1. Нейронные сети: градиентные методы оптимизации.
2. Сверточные нейронные сети.
3. Самостоятельное обучение нейронных сетей.
4. Проблема переобучения нейронных сетей.

Критерии оценки:

- оценка «зачтено» выставляется аспиранту, если была проявлена дискуссионная активность в рамках круглого стола;
- оценка «не зачтено» - отсутствовала дискуссионная активность в рамках круглого стола.

Раздел (тема) дисциплины Прикладные задачи машинного обучения**Перечень дискуссионных тем для круглого стола (дискуссии, полемики, диспута, дебатов)**

1. Задачи обработки и преобразования последовательностей.
2. Обучение ранжированию.
3. Поиск ассоциативных правил.
4. Адаптивные методы прогнозирования.
5. Онлайнное обучение.
6. Активное обучение.

Комплексное задание:

Задан временной ряд с постоянным шагом. Необходимо решить задачу его прогнозирования. Для этого выполнить следующие действия.

- 1) Провести проверку стационарности ряда с использованием теста Дики-Фуллера.
- 2) С использованием метода Фостера-Стюарта и критерия серий подтвердить гипотезу наличия тренда у заданного временного ряда.
- 3) Обучить модель ARIMA с выбранным по критерию AIC параметрами и сделать предсказание на следующие 10 отсчетов вперед.
- 4) Разделить временной ряд на тренировочную и тестовую части.
- 5) Для тестовой части вычислить значения метрик MAPE, SMAPE и MAE.
- 6) Сделать вывод о качестве построенной модели ARIMA.

Критерии оценки:

- оценка «зачтено» выставляется аспиранту, если была проявлена дискуссионная активность в рамках круглого стола и задания выполнены без ошибок;

- оценка «не зачтено» - отсутствовала дискуссионная активность в рамках круглого стола и/или задания выполнены с ошибками.

5. Учебно-методическое и информационное обеспечение дисциплины

5.1 Основная литература

1. Вьюгин, В. В. Математические основы машинного обучения и прогнозирования : учебное пособие [Электронный ресурс] / В. В. Вьюгин. – Москва : МЦНМО, 2014. – 304 с. – URL: <https://e.lanbook.com/book/56397>.

2. Воронина, В. В. Теория и практика машинного обучения : учебное пособие [Электронный ресурс] / В. В. Воронина. – Ульяновск : УлГТУ, 2017. – 290 с. – URL: <https://e.lanbook.com/book/165053>

3. Воронцов, К. В. Машинное обучение: курс лекций [Электронный ресурс] / К. В. Воронцов. – Москва: Школа анализа данных (Яндекс), 2019. – URL <https://bit.ly/1bCmE3Z>.

4. Ростовцев, В. С. Искусственные нейронные сети : учебник для вузов [Электронный ресурс] / В. С. Ростовцев. – Санкт-Петербург : Лань, 2021. – 216 с. – URL: <https://e.lanbook.com/book/160142>.

5.2 Дополнительная литература

1. Баймуратов, И. Р. Методы автоматизации машинного обучения : учебное пособие [Электронный ресурс] / И. Р. Баймуратов. – Санкт-Петербург : НИУ ИТМО, 2020. – 40 с. – URL: <https://e.lanbook.com/book/190871>.

2. Данилов, В. В. Нейронные сети : учебное пособие [Электронный ресурс] / В. В. Данилов. – Донецк : ДонНУ, 2020. – 158 с. – URL: <https://e.lanbook.com/book/179953>.

3. Лимановская, О. В. Основы машинного обучения: учебное пособие [Электронный ресурс] / О. В. Лимановская, Т. И. Алферьева. – Москва : ФЛИНТА, 2022. – 88 с. – URL: <https://e.lanbook.com/book/231677>.

5.3. Интернет-ресурсы (электронные учебно-методические издания, лицензионное программное обеспечение)

На сайте библиотеки <http://library.ugatu.ac.ru/> в разделе «Информационные ресурсы», подраздел «Доступ к БД» размещены ссылки на интернет-ресурсы.

1. Anaconda. Individual Edition [Электронный ресурс]. - Режим доступа: URL: <https://www.anaconda.com/> (дата обращения: 22.06.2021).

2. Pandas - Python Data Analysis Library [Электронный ресурс]. - Режим доступа: URL: <https://pandas.pydata.org/> (дата обращения: 22.06.2021).

3. Rashi Desai Топ-10 библиотек Python для Data Science [Электронный ресурс]. - Режим доступа: - Режим доступа: URL: <https://towardsdatascience.com/top-10-python-libraries-for-data-science-cd82294ec266> (дата обращения: 22.06.2021).

4. Scikit-learn. Машинное обучение на Python [Электронный ресурс]. - Режим доступа: URL: <https://scikit-learn.org> (дата обращения: 22.06.2021).

5. The MNIST database of handwritten digits [Электронный ресурс]. - Режим доступа: URL: <http://yann.lecun.com/exdb/mnist/> (дата обращения: 22.06.2021).

5.4 Методические указания к практическим занятиям

1. Гладилин, П. Е. Технологии машинного обучения : учебно-методическое пособие [Электронный ресурс] / П. Е. Гладилин, К. О. Боченина. – Санкт-Петербург : НИУ ИТМО, 2020. – 75 с. – URL: <https://e.lanbook.com/book/190885>.

6. Адаптация рабочей программы для лиц с ОВЗ

При инклюзивном обучении лиц с ОВЗ предоставляется возможность использовать следующие материально-технические средства:

- для аспирантов с ОВЗ по зрению предусматривается применение средств преобразования визуальной информации в аудио и тактильные сигналы, таких как, брайлевская компьютерная техника, электронные лупы, видеоувеличители, программы невидимого доступа к информации, программы-синтезаторов речи;

- для аспирантов с ОВЗ по слуху предусматривается применение сурдотехнических средств, таких как, системы беспроводной передачи звука, техники для усиления звука, видеотехника, мультимедийная техника и другие средства передачи информации в доступных формах;

для аспирантов с нарушениями опорно-двигательной функции предусматривается применение специальной компьютерной техники с соответствующим программным обеспечением, в том числе, специальные возможности операционных систем, таких, как экранная клавиатура и альтернативные устройства ввода информации.