МИНОБРНАУКИ РОССИИ

Ярославский государственный университет им. П.Г. Демидова Кафедра компьютерных сетей

УТВЕРЖДАЮ

Декан факультета ИВТ



Д.Ю. Чалый

«22» мая 2024 г.

# Рабочая программа дисциплины

«Программные инструменты интеллектуального анализа данных»

# Направление подготовки

02.04.02 Фундаментальная информатика и информационные технологии

# Профиль

«Искусственный интеллект и компьютерные науки»

# Квалификация выпускника

Магистр

# Форма обучения

очная

Программа рассмотрена на заседании кафедры от «22» апреля 2024 г.,

протокол № 8

Программа одобрена НМК факультета ИВТ

протокол № 6 от

«26» апреля 2024 г.

Ярославль

# Цели освоения дисциплины

Цель дисциплины — познакомить с современными инструментами анализа данных. Дисциплина предполагает, что с самими методами анализа данных обучающийся уже знаком (успешно освоил курс «Статистические методы анализа данных» или его аналог). Также предполагается знакомство обучающегося с, как минимум, основами программирования.

# Место дисциплины в структуре ОП магистратуры

Дисциплина «ПРОГРАММНЫЕ ИНСТРУМЕНТЫ ИНТЕЛЛЕКТУАЛЬНОГО

АНАЛИЗА ДАННЫХ» относится к вариативной части ОП магистратуры.. Для освоения данной дисциплиной студенты должны владеть математическим аппаратом линейной алгебры, для программной реализации алгоритмов знать один из языков программирования, проявлять настойчивость, целеустремленность и инициативу в процессе обучения.

# Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОП магистратуры

Процесс изучения дисциплины направлен на формирование следующих элементов компетенций в соответствии с ФГОС ВО, ОП ВО и приобретения следующих знаний, умений, навыков и (или) опыта деятельности:

|  |  |  |
| --- | --- | --- |
| **Формируемая компетенция**  **(код и формулировка)** | **Индикатор достижения компетенции**  **(код и формулировка)** | **Перечень планируемых результатов обучения** |
| **Профессиональные компетенции** | | |
| ПК-4. Способен руководить проектами по созданию комплексных систем  искусственного интеллекта | ПК-4.1 Осуществляет руководство созданием комплексных систем искусственного интеллекта с применением новых  методов и алгоритмов машинного обучения | Знать:  технологическое развитие центров обработки данных, наращивание и поддержание технологических мощностей и компетенций подразделений  Уметь:  Участвовать в создании (модернизации)  общедоступных платформ для хранения наборов данных, соответствующих методологиям описания, сбора и разметки данных; хранения наборов данных (в том числе звуковых, речевых, медицинских, метеорологических,  промышленных данных и данных систем видеонаблюдения) на  общедоступных платформах  для обеспечения потребностей организаций- разработчиков в  области искусственного |

|  |  |  |
| --- | --- | --- |
|  |  | интеллекта  Владеть навыками:  личное участие в проектах в роли архитектора центра обработки данных, технологического эксперта |

# Объем, структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 3 зач.ед., 108 акад.час.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Темы (разделы) дисциплины, их содержание** | **Сем ест р** | **Виды учебных занятий, включая самостоятельную работу студентов,**  **и их трудоемкость**  **(в академических часах)** | | | | | | **Формы текущего контроля успеваемости**  **Форма промежуточной аттестации**  **(по семестрам)** |
|  |  |  | **Контактная работа** | | | | |  |  |
|  |  |  | лек ции | пра кти чес кие | лаб ора тор ны е | кон сул ьта ции | атте стац ион ные исп ыта  ния | самос тояте льная работ а |  |
| 1. | Введение.  Программные инструменты интеллектуального анализа данных. Виды программных инструментов. Выбор используемого программного инструмента.  Совместимость используемых форматов данных между программными инструментами. | 3 | 2 |  |  |  |  | 14 |  |
| 2 | Язык программирования R как универсальный инструмент анализа данных. Простейшие сценарии работы с R. Обработка числовой и  текстовой информации. | 3 | 2 | 2 |  |  |  | 14 |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Зеркало пакетов CRAN. Менеджер пакетов Packrat. Различные способы представления информации. |  |  |  |  |  |  |  |  |
| 3 | Применение R для классических задач статистического анализа. Применение R для анализа многомерных данных. | 3 | 3 | 12 |  | 1 |  | 14 |  |
| 4 | Применение R для задач машинного обучения. | 3 | 2 | 4 |  |  |  | 14 |  |
| 5 | Применение R  для создания  рекомендательных систем и систем поддержки принятия решений. | 3 | 2 | 2 |  | 1 |  | 12 |  |
|  | Автоматизация рутинных задач обработки данных с помощью R.  Формирование отчётов. | 3 | 1 | 4 |  |  |  | 1,7 |  |
|  | Всего за 3 семестр |  | **12** | **24** |  | **2** | **0,3** | **69,7** |  |
|  | **Всего** |  | **12** | **24** |  | **2** | **0,3** | **69,7** | **Экзамен** |

# Содержание разделов дисциплины:

|  |
| --- |
| Введение. Программные инструменты интеллектуального анализа данных. Виды программных инструментов. Выбор используемого программного инструмента. Совместимость используемых форматов данных между программными инструментами. |
| Язык программирования R как универсальный инструмент анализа данных. Простейшие сценарии работы с R. Обработка числовой и текстовой информации. Зеркало пакетов CRAN. Менеджер пакетов Packrat. Различные способы представления информации. |
| Применение R для классических задач статистического анализа. Применение R для анализа многомерных данных. |
| Применение R для задач машинного обучения. |
| Применение R для создания рекомендательных систем и систем поддержки принятия решений. |

|  |
| --- |
| Автоматизация рутинных задач обработки данных с помощью R. Формирование отчётов. |

1. **Образовательные технологии, используемые при осуществлении образовательного процесса по дисциплине**

В процессе обучения используются следующие образовательные технологии:

Вводная лекция – дает первое целостное представление о дисциплине и ориентирует студента в системе изучения данной дисциплины. Студенты знакомятся с назначением и задачами курса, его ролью и местом в системе учебных дисциплин и в системе подготовки в целом. Дается краткий обзор курса, история развития науки и практики, достижения в этой сфере, имена известных ученых, излагаются перспективные направления исследований. На этой лекции высказываются методические и организационные особенности работы в рамках данной дисциплины, а также дается анализ рекомендуемой учебно-методической литературы.

Академическая лекция (или лекция общего курса) – последовательное изложение материала, осуществляемое преимущественно в виде монолога преподавателя. Требования к академической лекции: современный научный уровень и насыщенная информативность, убедительная аргументация, доступная и понятная речь, четкая структура и логика, наличие ярких примеров, научных доказательств, обоснований, фактов.

Лабораторное занятие – занятие, посвященное освоению и углублению конкретных умений и навыков и закреплению полученных на лекции знаний.

# Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень лицензионного программного обеспечения и информационных справочных систем (при необходимости)

В процессе осуществления образовательного процесса используются:

* для формирования текстов материалов для промежуточной и текущей аттестации, для разработки документов, презентаций, для работы с электронными таблицами - программы OfficeStd 2013 RUSOLPNLAcdmc 021-10232, LibreOffice (свободное), издательская система LaTeX;

- компиляторы с высокоуровневых языков программирования;

* для поиска учебной литературы библиотеки ЯрГУ– Автоматизированная библиотечная информационная система "БУКИ-NEXT" (АБИС "Буки-Next").
* PacketTracer 6.3, CiscoSDM, CiscoNetworkAssistant, CiscoConfigurationProfessional.

1. **Перечень основной и дополнительной учебной литературы, ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

Основная литература:

Буховец, А. Г. Алгоритмы вычислительной статистики в системе R : учебное пособие / А. Г. Буховец, П. В. Москалев. — 2-е изд., перераб. и доп. — Санкт-Петербург : Лань, 2022. — 160 с. — ISBN 978-5-8114-1802-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/212195>

Макшанов, А. В. Технологии интеллектуального анализа данных : учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2022. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/206711>

# Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

* + специальные помещения:

-учебные аудитории для проведения занятий лекционного типа и лабораторных занятий;

* + - учебные аудитории для проведения групповых и индивидуальных консультаций,
    - учебные аудитории для проведения текущего контроля и промежуточной аттестации;

-помещения для самостоятельной работы;

-помещения для хранения и профилактического обслуживания технических средств обучения.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Число посадочных мест в лекционной аудитории больше либо равно списочному составу потока, а в аудитории для лабораторных занятий – списочному составу группы обучающихся.

* + фонд библиотеки.

Компьютерные классы, оборудованные ПЭВМ класса не ниже IntelСоrе2Duo , 4gbRAM, 60GHDDc установленным программным обеспечением: Windows7/8/10, Linux, PacketTracer 6.3, CiscoSDM, CiscoNetworkAssistant, CiscoConfigurationProfessional. Из расчета одна ПЭВМ на одного человека.

# Приложение №1 к рабочей программе дисциплины

**«Программные инструменты интеллектуального**

# анализа данных»

**Фонд оценочных средств**

**для проведения текущей и промежуточной аттестации студентов по дисциплине**

Типовые задания для практических занятий

* Анализ одного из популярных наборов данных (на выбор студента) с помощью набора выбираемых студентом (с помощью знаний, полученных ранее на курсе «Статистические методы анализа данных») статистических методов, представление результатов анализа (в т.ч. графическое).
* Выбор наиболее подходящего пакета из зеркала пакетов CRAN для решения одной из задач многомерного анализа на основе изучения документации.
* Построение дерева решений с помощью R, его интерпретация.
* Анализ одного из популярных наборов текстовых данных средствами R, формирование рекомендаций на их основе.
* Формирование отчёта средствами R в формате .pdf по результатам автоматического анализа данных

Итоговое задание по курсу

Выбрать доступный набор данных (числовой, текстовый или смешанный), проанализировать его с помощью R, результат анализа оформить в виде рекомендательной системы или автоматически формирующегося отчёта. При этом система формирования отчёта должна отслеживать собственную актуальность и не предоставлять недостоверную информацию.

# Приложение №2 к рабочей программе дисциплины

**«**Программные инструменты интеллектуального анализа данных**»**

# Методические указания для студентов по освоению дисциплины

Основной формой изложения учебного материала по дисциплине «Программные инструменты интеллектуального анализа данных» являются практические занятия, на которых происходит закрепление материала путем применения его к конкретным задачам и отработка практических навыков.

Для успешного освоения дисциплины очень важно решение достаточно большого количества задач, как в аудитории, так и самостоятельно в качестве домашних заданий. Примеры решения задач разбираются на лекциях и лабораторных занятиях, при необходимости по наиболее трудным темам проводятся дополнительные консультации. Для решения всех задач необходимо знать и понимать лекционный материал. Поэтому в процессе изучения дисциплины рекомендуется регулярное повторение пройденного лекционного материала. Материал, законспектированный на лекциях, необходимо дома еще раз прорабатывать и при необходимости дополнять информацией, полученной на консультациях, лабораторных занятиях или из учебной литературы.

Освоить вопросы, излагаемые в процессе изучения дисциплины «Программные инструменты интеллектуального анализа данных» самостоятельно студенту крайне сложно. Это связано со сложностью изучаемого материала и большим объемом курса. Поэтому посещение всех аудиторных занятий является совершенно необходимым. Без упорных и регулярных занятий в течение семестра сдать зачет по итогам изучения дисциплины студенту практически невозможно.

# Учебно-методическое обеспечение самостоятельной работы студентов по дисциплине

Для самостоятельной работы особенно рекомендуется использовать учебную литературу.

Также для подбора учебной литературы рекомендуется использовать широкий спектр интернет-ресурсов:

1. Электронно-библиотечная система «Университетская библиотека online» ([www.biblioclub.ru](http://www.biblioclub.ru/) ) - электронная библиотека, обеспечивающая доступ к наиболее востребованным материалам-первоисточникам, учебной, научной и художественной литературе ведущих издательств (\*регистрация в электронной библиотеке – только в сети университета. После регистрации работа с системой возможна с любой точки доступа в Internet.).
2. Информационная система "Единое окно доступа к образовательным ресурсам" (<http://window.edu.ru/library>).

Целью создания информационной системы "Единое окно доступа к образовательным ресурсам" (ИС "Единое окно ") является обеспечение свободного доступа к интегральному каталогу образовательных интернет-ресурсов и к электронной библиотеке учебно-методических материалов для общего и профессионального образования.

Информационная система "Единое окно доступа к образовательным ресурсам" создана по заказу Федерального агентства по образованию в 2005-2008 гг. Головной разработчик проекта - Федеральное государственное автономное учреждение Государственный научно-исследовательский институт информационных технологий и телекоммуникаций (ФГАУ ГНИИ ИТТ "Информика") [www.informika.ru](http://www.informika.ru/).

Для самостоятельного подбора литературы в библиотеке ЯрГУ рекомендуется использовать:

1. Личный кабинет (<http://lib.uniyar.ac.ru/opac/bk_login.php>)дает возможность получения on-line доступа к списку выданной в автоматизированном режиме литературы, просмотра и копирования электронных версий изданий сотрудников университета (учеб. и

метод. пособия, тексты лекций и т.д.) Для работы в «Личном кабинете» необходимо зайти на сайт Научной библиотеки ЯрГУ с любой точки, имеющей доступ в Internet, в пункт меню

«Электронный каталог»; пройти процедуру авторизации, выбрав вкладку «Авторизация», и заполнить представленные поля информации.

1. Электронная библиотека учебных материалов ЯрГУ (<http://www.lib.uniyar.ac.ru/opac/bk_cat_find.php>) содержит более 2500 полных

текстов учебных и учебно-методических материалов по основным изучаемым дисциплинам, изданных в университете. Доступ в сети университета, либо по логину/паролю.

1. Электронная картотека [«Книгообеспеченность»](http://10.1.0.4/buki/bk_bookreq_find.php) (<http://www.lib.uniyar.ac.ru/opac/bk_bookreq_find.php>) раскрывает учебный фонд

научной библиотеки ЯрГУ, предоставляет оперативную информацию о состоянии книгообеспеченности дисциплин основной и дополнительной литературой, а также цикла дисциплин и специальностей. Электронная картотека [«Книгообеспеченность»](http://10.1.0.4/buki/bk_bookreq_find.php) доступна в сети университета и через Личный кабинет.