

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Ярославский государственный университет им. П.Г. Демидова

Кафедра компьютерных сетей

УТВЕРЖДАЮ

Декан факультета ИВТ

 Д.Ю. Чалый

« 24 » мая 2022 г.

Рабочая программа дисциплины
«Интеллектуальный анализ данных»

Направление подготовки
01.03.02 Прикладная математика и информатика

Направленность (профиль)
«Программирование и технологии искусственного интеллекта»

Квалификация выпускника
Бакалавр

Форма обучения
очная

Программа рассмотрена
на заседании кафедры
от 22 марта 2022 г.,
протокол № 7

Программа одобрена НМК
факультета ИВТ
протокол № 6 от
18 апреля 2022 г.

Ярославль

1. 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.

Дисциплина «Интеллектуальный анализ данных» относится к вариативной части ОП бакалавриата.

2. Место дисциплины в структуре ОП бакалавриата

Дисциплина «Интеллектуальный анализ данных» относится к вариативной части (дисциплина по выбору) ОП бакалавриата.

Для освоения данной дисциплиной студенты должны обладать знаниями по математике и информатике, в т.ч.: основы программирования на языках С и/или С++, алгоритмов и структур данных, алгебре, геометрии, дифференциальными уравнениями, уравнениям в частных производных и численным методам; а также проявлять настойчивость, целеустремленность и инициативу в процессе обучения.

3. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОП бакалавриата

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-2 Способен к разработке и применению алгоритмов, моделей данных в профессиональной области | ПК – 2.3 Умеет разрабатывать и реализовывать алгоритмы математических моделей на базе языков и пакетов прикладных программ моделирования | Знать: – области применимости параллельных вычислений; Уметь: – интерпретировать результаты полученных вычислений; Владеть навыками: – обработки полученных вычислений для формирования соответствующих выводов об эффективности применяемых параллельных вычислений. |

4. Объем, структура и содержание дисциплины

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

| № п/п | Темы (разделы) дисциплины, их содержание | Семестр | Виды учебных занятий, включая самостоятельную работу студентов, и их трудоемкость (в академических часах) | | | | | | Формы текущего контроля успеваемости |
|----------|--|---------|---|--------------|--------------|--------------|-----------------------------|---------------------------|--|
| | | | Контактная работа | | | | | | Форма |
| | | | лекции | практические | лабораторные | консультации | аттестационные испытания | самостоятельная работа | |
| 1 | Введение | 7 | 3 | | 6 | | | 2 | |
| 2 | Язык программирования R как универсальный инструмент анализа данных | 7 | 3 | | 6 | | | 2 | |
| 3 | Применение R для классических задач статистического анализа | 7 | 3 | | 6 | | | 2 | |
| 4 | Применение R для задач машинного обучения | 7 | 3 | | 6 | | | 2 | |
| 5 | Применение R для создания рекомендательных систем и систем поддержки принятия решений. | 7 | 3 | | 6 | | | 2 | |
| 6 | Автоматизация рутинных задач обработки данных с помощью R. Формирование отчётов | 7 | 3 | | 6 | | | 3 | |
| | Всего за 7 семестр | | 18 | | 36 | 5 | | 13 | Зачет |
| | Всего | | 18 | | 36 | 5 | | 13 | |

Содержание разделов дисциплины:

Раздел 1. Введение. Программные инструменты интеллектуального анализа данных. Виды программных инструментов. Выбор используемого программного инструмента. Совместимость используемых форматов данных между программными инструментами.

Раздел 2. 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.

Раздел 3. Применение R для классических задач статистического анализа. Применение R для анализа многомерных данных.

Раздел 4. Применение R для задач машинного обучения.

Раздел 5. Применение R для создания рекомендательных систем и систем поддержки принятия решений.

Раздел 6. Автоматизация рутинных задач обработки данных с помощью R. Формирование отчётов.

5. 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.

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

6. 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 RUS OLP NL Acsmc 021-10232, LibreOffice (свободное), издательская система LaTeX;

- компиляторы с высокоуровневых языков программирования;

- вычислительный гибридный кластер ЯрГУ.

- программные библиотеки OpenMP, MPI и NvidiaCUDA.

– для поиска учебной литературы библиотеки ЯрГУ – Автоматизированная библиотечная информационная система "БУКИ-NEXT" (АБИС "Буки-Next").

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

а) основная:

1 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, 2021. — 160 с. — ISBN 978-5-8114-1802-2. — Текст : электронный // Лань : электронно-библиотечная система.

2 Макшанов, А. В. Технологии интеллектуального анализа данных : учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система.

3 Нестеров, С. А., Основы интеллектуального анализа данных. Лабораторный практикум [Электронный ресурс] : учебное пособие / С. А. Нестеров, СПб., Лань, 2020, 40с

б) дополнительная:

в) ресурсы сети «Интернет»

Электронно-библиотечная система «Юрайт»(<https://urait.ru/>).

Электронно-библиотечная система «Лань»(<https://e.lanbook.com/>).

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

- специальные помещения:
 - компьютерный класс, оборудованный для проведения лекций и практических 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 практических занятий (семинаров)– списочному составу группы обучающихся.

- фонд библиотеки,
- компьютерная техника с доступом к вычислительному гибриднему кластеру ЯрГУ.

Автор(ы) :

Старший преподаватель

кафедры компьютерных сетей, к.ф.-м.н.

_____ С.В. Алешин

**Приложение №1 к рабочей программе дисциплины
«Интеллектуальный анализ данных»
Фонд оценочных средств
для проведения текущей и промежуточной аттестации студентов
по дисциплине**

1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

1.1. Контрольные задания и иные материалы, используемые в процессе текущей аттестации

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.

2. Перечень компетенций, этапы их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкалы оценивания

2.1. Шкала оценивания сформированности компетенций и ее описание

Оценивание уровня сформированности компетенций в процессе освоения дисциплины осуществляется по следующей трехуровневой шкале:

Пороговый уровень - предполагает отражение тех ожидаемых результатов, которые определяют минимальный набор знаний и (или) умений и (или) навыков, полученных студентом в результате освоения дисциплины. Пороговый уровень является обязательным уровнем для студента к моменту завершения им освоения данной дисциплины.

Продвинутый уровень - предполагает способность студента использовать знания, умения, навыки и (или) опыт деятельности, полученные при освоении дисциплины, для решения профессиональных задач. Продвинутый уровень превосходит пороговый уровень по нескольким существенным признакам.

Высокий уровень - предполагает способность студента использовать потенциал интегрированных знаний, умений, навыков и (или) опыта деятельности, полученных при освоении дисциплины, для творческого решения профессиональных задач и самостоятельного поиска новых подходов в их решении путем комбинирования и использования известных способов решения применительно к конкретным условиям. Высокий уровень превосходит пороговый уровень по всем существенным признакам.

2.2. Перечень компетенций, этапы их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования

| Код компетенции | Форма контроля | Этапы формирования (№ темы (раздела)) | Показатели оценивания | Шкала и критерии оценивания компетенций на различных этапах их формирования | | |
|----------------------------------|--|---------------------------------------|---|---|--|--|
| | | | | Пороговый уровень | Продвинутый уровень | Высокий уровень |
| Общепрофессиональные компетенции | | | | | | |
| ОПК-2 | Самостоятельная работа по разделам 2-6. Зачет. | 1-6 | Знать: – архитектуру параллельных вычислительных систем; Уметь: – разбивать программу на независимые процессы ; Владеть навыками: – разработки программ с использованием технологий параллельного программирования OpenMP, MPI и CUDA. | Знание основы архитектуры параллельных систем. Умение выделять в программе независимые процессы Владеть основами работы с использованием технологий параллельного программирования OpenMP, MPI и CUDA | Знать основные параллельные алгоритмы работы для решения практических задач. Владеть продвинутыми средствами работы для технологий параллельного программирования OpenMP, MPI и CUDA. | Знать продвинутые параллельные алгоритмы работы для решения практических задач. Уметь практически решать задачи численного характера с использованием адекватных средств параллельных вычислений. Владеть продвинутыми средствами работы для совместного использования технологий параллельного программирования OpenMP, MPI и CUDA. |
| Профессиональные компетенции | | | | | | |
| ПК-1 | Самостоятельная | 1-5 | Знать: | Базовые знания и | Продвинутые знания и | Продвинутые знания и |

| | | | | | | |
|--|--------------------------------------|--|--|--|--|--|
| | работа по разделам 2-5. Зачет. | | <p>– области применимости параллельных вычислений; Уметь:</p> <p>– интерпретировать полученные результаты вычислений; Владеть навыками:</p> <p>– обработки полученных вычислений для формирования соответствующих выводов об эффективности применяемых параллельных вычислений</p> | <p>понимание применимости параллельных вычислений</p> <p>Базовое умение интерпретировать результаты полученных вычислений.</p> <p>Владение базовыми навыками по обработке вычислений</p> | <p>понимание применимости параллельных вычислений</p> <p>Продвинутое умение интерпретировать результаты полученных вычислений.</p> <p>Владение продвинутыми навыками по обработке вычислений</p> | <p>понимание применимости параллельных вычислений в различных сферах науки и производства</p> <p>Продвинутое умение интерпретировать результаты полученных вычислений.</p> <p>Владение продвинутыми навыками по обработке вычислений</p> |
|--|--------------------------------------|--|--|--|--|--|

3. Методические рекомендации преподавателю по процедуре оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Целью процедуры оценивания является определение степени овладения студентом ожидаемыми результатами обучения (знаниями, умениями, навыками и (или) опытом деятельности).

Процедура оценивания степени овладения студентом ожидаемыми результатами обучения осуществляется с помощью методических материалов, представленных в разделе «Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций»

3.1 Критерии оценивания степени овладения знаниями, умениями, навыками и (или) опытом деятельности, определяющие уровни сформированности компетенций

Пороговый уровень (общие характеристики):

- владение основным объемом знаний по программе дисциплины;
- знание основной терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы без существенных ошибок;
- владение инструментарием дисциплины, умение его использовать в решении стандартных (типовых) задач;
- способность самостоятельно применять типовые решения в рамках рабочей программы дисциплины;
- усвоение основной литературы, рекомендованной рабочей программой дисциплины;
- знание базовых теорий, концепций и направлений по изучаемой дисциплине;
- самостоятельная работа на практических и лабораторных занятиях, периодическое участие в групповых обсуждениях, достаточный уровень культуры исполнения заданий.

Продвинутый уровень (общие характеристики):

- достаточно полные и систематизированные знания в объёме программы дисциплины;
- использование основной терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать выводы;
- владение инструментарием дисциплины, умение его использовать в решении учебных и профессиональных задач;
- способность самостоятельно решать сложные задачи (проблемы) в рамках рабочей программы дисциплины;
- усвоение основной и дополнительной литературы, рекомендованной рабочей программой дисциплины;
- умение ориентироваться в базовых теориях, концепциях и направлениях по изучаемой дисциплине и давать им сравнительную оценку;
- самостоятельная работа на практических и лабораторных занятиях, участие в групповых обсуждениях, высокий уровень культуры исполнения заданий.

Высокий уровень (общие характеристики):

- систематизированные, глубокие и полные знания по всем разделам дисциплины;

- точное использование терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;
- безупречное владение инструментарием дисциплины, умение его использовать в постановке и решении научных и профессиональных задач;
- способность самостоятельно и творчески решать сложные задачи (проблемы) в рамках рабочей программы дисциплины;
- полное и глубокое усвоение основной и дополнительной литературы, рекомендованной рабочей программой дисциплины;
- умение ориентироваться в основных теориях, концепциях и направлениях по изучаемой дисциплине и давать им критическую оценку;
- активная самостоятельная работа на практических и лабораторных занятиях, творческое участие в групповых обсуждениях, высокий уровень культуры исполнения заданий.

3.2 Описание процедуры выставления оценки

В зависимости от уровня сформированности каждой компетенции по окончании освоения дисциплины студенту выставляется оценка. Вид оценки («зачтено», «незачтено») определяется рабочей программой дисциплины в соответствии с учебным планом.

Оценка «зачет» выставляется студенту, у которого каждая компетенция (полностью или частично формируемая данной дисциплиной) сформирована не ниже, чем на пороговом уровне.

Оценка «незачтено» выставляется студенту, у которого хотя бы одна компетенция (полностью или частично формируемая данной дисциплиной) сформирована ниже, чем на пороговом уровне.

Приложение №2 к рабочей программе дисциплины «Параллельное программирование»

Методические указания для студентов по освоению дисциплины

Основной формой изложения учебного материала по дисциплине «Параллельное программирование» являются лекции и практические занятия. Это связано с тем, что в основе читаемого курса лежит особый математический аппарат, с помощью которого решаются довольно сложные и громоздкие задачи. По большинству тем предусмотрены практические занятия, на которых происходит закрепление лекционного материала путем применения его к конкретным задачам и отработка практических навыков программирования.

Большое внимание должно быть уделено выполнению домашней работы. В качестве заданий для самостоятельной работы дома студентам предлагаются задачи, аналогичные разобранным на лекциях и практических занятиях или немного более сложные, которые являются результатом объединения нескольких базовых задач.

Для проверки и контроля усвоения теоретического материала, приобретенных практических навыков работы в течение обучения проводятся мероприятия текущей аттестации в виде самостоятельных работ.

В конце изучения дисциплины студенты сдают зачет. Зачет выставляется по итогам выполнения самостоятельных заданий и краткого собеседования по его результатам.

Освоить вопросы, излагаемые в процессе изучения дисциплины «Параллельное программирование» самостоятельно студенту достаточно сложно. Это связано со сложностью изучаемого материала, требованиям к техническому оснащению курса и большим объемом материала курса. Поэтому посещение всех аудиторных занятий является совершенно необходимым. Без упорных и регулярных занятий в течение семестра, в том числе самостоятельных, сдать зачет по итогам изучения дисциплины студенту достаточно сложно.

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) - 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.

Для самостоятельного подбора литературы в библиотеке ЯрГУ рекомендуется использовать:

1. Личный кабинет (http://lib.uniylar.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.

2. 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.

3. Электронная картотека «Книгообеспеченность»

(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» доступна в сети университета и через Личный кабинет.