

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION

**Federal State Autonomous Educational Institution of Higher Education
«National Research Lobachevsky State University of Nizhny Novgorod»**

Институт информационных технологий, математики и механики

УТВЕРЖДЕНО

решением президиума Ученого совета ННГУ

протокол № 1 от 16.01.2024 г.

Working programme of the discipline

Big Data Processing

Higher education level

Master degree

Area of study / speciality

02.04.02 - Fundamental Informatics and Information Technology

Focus /specialization of the study programme

Artificial Intelligence and Data Analysis

Mode of study

full-time

Nizhny Novgorod

Year of commencement of studies 2024

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.О.03 Обработка больших данных относится к обязательной части образовательной программы.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

| Формируемые компетенции (код, содержание компетенции) | Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции | | Наименование оценочного средства | |
|--|--|---|------------------------------------|----------------------------------|
| | Индикатор достижения компетенции (код, содержание индикатора) | Результаты обучения по дисциплине | Для текущего контроля успеваемости | Для промежуточной аттестации |
| ОПК-1: Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий | ОПК-1.1: Обладает фундаментальными знаниями в области математических и естественных наук, теории коммуникаций ОПК-1.2: Умеет осуществлять первичный сбор и анализ материала, интерпретировать различные математические объекты ОПК-1.3: Имеет практический опыт работы с решением математических задач и применяет его в профессиональной деятельности | ОПК-1.1: Обладает знаниями в области моделей и методов принятия решений; моделирования данных и моделей обработки данных. ОПК-1.2: Умеет осуществлять сбор, очистку и предварительную обработку данных. ОПК-1.3: Имеет практический опыт разработки программного обеспечения решения математических задач для аналитики данных и принятия решений. | Практическое задание | Экзамен: Практическое задание |
| ОПК-3: Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования | ОПК-3.1: Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей ОПК-3.2: Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные | ОПК-3.1: Знает методы проектирования, разработки и использования систем хранения данных. Знает методы предварительного и проблемного анализа данных. ОПК-3.2: Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, | Практическое задание | Экзамен: Практическое задание |

| | | | | |
|--|--|---|--|--|
| | ресурсы глобальных сетей, образовательного контента, средств тестирования систем ОПК-3.3: Имеет практический опыт применения разработки программного обеспечения и тестирования программных продуктов | образовательного контента, средств тестирования систем. ОПК-3.3: Разработка решения практической задачи и реализация мини-проекта на этой основе | | |
|--|--|---|--|--|

3. Структура и содержание дисциплины

3.1 Трудоемкость дисциплины

| | |
|--|-----------------------------|
| | очная |
| Общая трудоемкость, з.е. | 6 |
| Часов по учебному плану | 216 |
| в том числе | |
| аудиторные занятия (контактная работа): | |
| - занятия лекционного типа | 32 |
| - занятия семинарского типа (практические занятия / лабораторные работы) | 16 |
| - КСР | 2 |
| самостоятельная работа | 130 |
| Промежуточная аттестация | 36 Экзамен |

3.2. Содержание дисциплины

(структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий)

| Наименование разделов и тем дисциплины | Всего (часы) | в том числе | | | |
|--|-----------------|--|--|-------------|---|
| | | Контактная работа (работа во взаимодействии с преподавателем), часы из них | | | Самостоятельная работа обучающегося, часы |
| | | Занятия лекционного типа | Занятия семинарского типа (практические занятия/лабораторные работы), часы | Всего | |
| | о ф о | о ф о | о ф о | о ф о | о ф о |
| Тема 1. Теоретические основы принятия решений. | 10 | 4 | 2 | 6 | 4 |
| Тема 2. Теоретические основы и практика систем хранения данных | 52 | 8 | 4 | 12 | 40 |
| Тема 3. Сбор данных | 34 | 2 | 2 | 4 | 30 |
| Тема 4. Статистический анализ и визуализация | 32 | 8 | 4 | 12 | 20 |

| | | | | | |
|--|-----|----|----|----|-----|
| | | | | | |
| Тема 5. Данные как основа принятия решений | 50 | 10 | 4 | 14 | 36 |
| Аттестация | 36 | | | | |
| КСР | 2 | | | 2 | |
| Итого | 216 | 32 | 16 | 50 | 130 |

Contents of sections and topics of the discipline

Тема 1. Теоретические основы принятия решений.

- 1.1. Кратко о теории систем и моделирования.
- 1.2. Многокритериальный анализ и оценка альтернатив.
- 1.3. Решение задач оптимизации.

Глава 2. Теоретические и практические основы систем хранения данных

- 2.1. Реляционные базы данных - общие принципы построения. Язык SQL.
- 2.2. Агрегатные и аналитические функции.
- 2.3. Подзапросы и CTE (Common Table Expressions).
- 2.4. Процедурное расширение SQL.

Тема 3. Сбор данных.

- 3.1. Классификация типов данных.
- 3.2. Современное состояние и подходы сбора и предварительной обработки
- 3.3. Анализ достоверности и очистка данных
- 3.4. Группировка данных.
- 3.5. Выборочный сбор данных.

Тема 4. Статистический анализ и визуализация.

- 4.1. Классификация данных.
- 4.2. Моделирование и сбор данных с учётом времени.

Тема 5. Данные как основа принятия решений.

- 5.1. Принципы моделей проблем принятия решений.
- 5.2. Задачи планирования запасов и производства
- 5.3. Сетевые модели и задачи принятия решений.
- 5.4. Применение программных средств поиска рациональных решений.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведенным в п. 5.

Самостоятельная работа обучающихся включает в себя подготовку к контрольным вопросам и заданиям для текущего контроля и промежуточной аттестации по итогам освоения дисциплины, приведенным в п. 5.

Шапошников Д.Е. Выбор вариантов в проектировании аппаратно-программных комплексов:

Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2020. – 87 с.

5. Assessment tools for ongoing monitoring of learning progress and interim certification in the discipline (module)

5.1 Model assignments required for assessment of learning outcomes during the ongoing monitoring of learning progress with the criteria for their assessment:

5.1.1 Model assignments (assessment tool - Practical task) to assess the development of the competency ОПК-1:

1. Построение модели и решение задачи о назначениях команд исполнителей на работы

Построить модель и решить двухкритериальную задачу о назначениях. Размерность и исходные данные изменить в соответствии с конкретной предметной областью. Построить область эффективных решений (не менее 5 точек).

2. Задача о планировании запасов

Магазин прессует и складывает в поддоны пустые картонные упаковочные коробки для их последующей переработки. За день штабелируется пять поддонов. Стоимость хранения одного поддона на заднем дворе магазина составляет 0,10 долл. в день. Компания, которая перевозит поддоны в перерабатывающий центр, устанавливает оплату в 100 долл. за аренду своего погрузочного оборудования плюс 3 долл. за перевозку каждого поддона. Изобразите графически изменение количества поддонов с течением времени и разработайте оптимальную стратегию доставки поддонов в перерабатывающий центр.

3. Задача о рейтинге банков

Есть данные об активах банков РФ, собранные в файле CSV. Требуется произвести аналитику. 1. Составить интервальный вариационный ряд. 2. Рассчитать медиану. 3. Представить графически полученные результаты. Для выполнения задачи загрузить данные в Excel (или другую электронную таблицу) и сформировать результаты.

5. Разработка программного комплекса анализа рейтинга банков.

Разработать программную систему, которая позволяет загрузить и анализировать таблицу рейтинга банков. Программный комплекс должен работать с реляционной базой данных (СУБД и платформа разработки клиентского приложения – на выбор студентов). Подсчитать рейтинг по областям принадлежности банков (используя аналитические функции).

5.1.2 Model assignments (assessment tool - Practical task) to assess the development of the competency ОПК-3:

Фирма может производить изделие или покупать его у подрядчика. Если фирма сама выпускает изделие, то каждый запуск его в производство обходится в 20 долл. Мощность производства составляет 100 единиц в день. Если изделие закупается, затраты на размещение каждого заказа равны 15 долл. Затраты на содержание изделия на складе, независимо от того, закупается оно или производится на фирме, равны 0,02 долл. в день. Потребление изделия фирмой оценивается в 260 000 единиц в год. Если предположить, что фирма работает без дефицита, определите, что выгоднее — закупать или производить изделия?

Assessment criteria (assessment tool — Practical task)

| Grade | Assessment criteria |
|-------|---|
| pass | Задание выполнено полностью |
| fail | Задание не выполнено полностью или частично |

5.2. Description of scales for assessing learning outcomes in the discipline during interim certification

Шкала оценивания сформированности компетенций

| Уровень сформированности компетенций (индикатор достижения компетенций) | плохо | неудовлетворительно | удовлетворительно | хорошо | очень хорошо | отлично | превосходно |
|---|---|--|--|---|---|--|--|
| | не зачтено | | зачтено | | | | |
| <u>Знания</u> | Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа | Уровень знаний ниже минимальных требований. Имели место грубые ошибки | Минимально допустимый уровень знаний. Допущено много негрубых ошибок | Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок | Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок | Уровень знаний в объеме, соответствующем программе подготовки. Ошибок нет. | Уровень знаний в объеме, превышающем программу подготовки. |
| <u>Умения</u> | Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа | При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки | Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме | Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания в полном объеме, но некоторые с недочетами | Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания в полном объеме, но некоторые с недочетами. | Продemonстрированы все основные умения. Решены все основные задачи с отдельным и несущественными недочетами, выполнены все задания в полном объеме | Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов |
| <u>Навыки</u> | Отсутствие базовых навыков. Невозможность оценить наличие навыков вследствие отказа обучающегося от | При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки | Имеется минимальный набор навыков для решения стандартных задач с некоторым | Продemonстрированы базовые навыки при решении стандартных задач с некоторым и | Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов | Продemonстрированы навыки при решении нестандартных задач без ошибок и | Продemonстрирован творческий подход к решению нестандартных задач |

| | | | | | | | |
|--|--------|--|-----------------|------------|--|-----------|--|
| | ответа | | и недочетами | недочетами | | недочетов | |
|--|--------|--|-----------------|------------|--|-----------|--|

Scale of assessment for interim certification

| Grade | | Assessment criteria |
|-------|----------------|---|
| pass | outstanding | All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "outstanding", the knowledge and skills for the relevant competencies have been demonstrated at a level higher than the one set out in the programme. |
| | excellent | All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "excellent", |
| | very good | All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "very good", |
| | good | All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "good", |
| | satisfactory | All the competencies (parts of competencies) to be developed within the discipline have been developed at a level no lower than "satisfactory", with at least one competency developed at the "satisfactory" level. |
| fail | unsatisfactory | At least one competency has been developed at the "unsatisfactory" level. |
| | poor | At least one competency has been developed at the "poor" level. |

5.3 Model control assignments or other materials required to assess learning outcomes during the interim certification with the criteria for their assessment:

5.3.1 Model assignments (assessment tool - Practical task) to assess the development of the competency ОПК-1

1. Построение модели и решение задачи о назначениях команд исполнителей на работы

Построить модель и решить двухкритериальную задачу о назначениях. Размерность и исходные данные изменить в соответствии с конкретной предметной областью. Построить область эффективных решений (не менее 5 точек).

2. Задача о планировании запасов

Магазин прессует и складывает в поддоны пустые картонные упаковочные коробки для их последующей переработки. За день штабелируется пять поддонов. Стоимость хранения одного поддона на заднем дворе магазина составляет 0,10 долл. в день. Компания, которая перевозит поддоны в перерабатывающий центр, устанавливает оплату в 100 долл. за аренду своего погрузочного оборудования плюс 3 долл. за перевозку каждого поддона. Изобразите графически изменение количества поддонов с течением времени и разработайте оптимальную стратегию доставки поддонов в перерабатывающий центр.

3. Задача о рейтинге банков

Есть данные об активах банков РФ, собранные в файле CSV. Требуется произвести аналитику. 1. Составить интервальный вариационный ряд. 2. Рассчитать медиану. 3. Представить графически полученные результаты. Для выполнения задачи загрузить данные в Excel (или другую электронную таблицу) и сформировать результаты.

5. Разработка программного комплекса анализа рейтинга банков.

Разработать программную систему, которая позволяет загрузить и анализировать таблицу рейтинга банков. Программный комплекс должен работать с реляционной базой данных (СУБД и платформа разработки клиентского приложения – на выбор студентов). Подсчитать рейтинг по областям принадлежности банков (используя аналитические функции).

5.3.2 Model assignments (assessment tool - Practical task) to assess the development of the competency ОПК-3

1. Построение модели и решение задачи о назначениях команд исполнителей на работы

Построить модель и решить двухкритериальную задачу о назначениях. Размерность и исходные данные изменить в соответствии с конкретной предметной областью. Построить область эффективных решений (не менее 5 точек).

2. Задача о планировании запасов

Магазин прессует и складывает в поддоны пустые картонные упаковочные коробки для их последующей переработки. За день штабелируется пять поддонов. Стоимость хранения одного поддона на заднем дворе магазина составляет 0,10 долл. в день. Компания, которая перевозит поддоны в перерабатывающий центр, устанавливает оплату в 100 долл. за аренду своего погрузочного оборудования плюс 3 долл. за перевозку каждого поддона. Изобразите графически изменение количества поддонов с течением времени и разработайте оптимальную стратегию доставки поддонов в перерабатывающий центр.

3. Задача о рейтинге банков

Есть данные об активах банков РФ, собранные в файле CSV. Требуется произвести аналитику.

1. Составить интервальный вариационный ряд.
2. Рассчитать медиану.
3. Представить графически полученные результаты.

Для выполнения задачи загрузить данные в Excel (или другую электронную таблицу) и сформировать результаты.

4. Дополнительная задача о рейтинге банков (на оценку "Превосходно" при условии правильности остальных задач).

Загрузить таблицу рейтинга банков в реляционную базу данных (СУБД – на выбор студентов). Подсчитать рейтинг по областям принадлежности банков (используя аналитические функции).

Assessment criteria (assessment tool — Practical task)

| Grade | Assessment criteria |
|----------------|--|
| outstanding | Уровень знаний в объеме, превышающем программу подготовки. Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов. Продемонстрирован творческий подход к решению нестандартных задач. |
| excellent | Уровень знаний в объеме, соответствующем программе подготовки, без ошибок. Продемонстрированы все основные умения, решены все основные задачи с отдельными несущественным недочетами, выполнены все задания в полном объеме. Продемонстрированы навыки при решении нестандартных задач без ошибок и недочетов. |
| very good | Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько несущественных ошибок. Продемонстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами. Продемонстрированы базовые навыки при решении стандартных задач без ошибок и недочетов. |
| good | Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок. Продемонстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами. Продемонстрированы базовые навыки при решении стандартных задач с некоторыми недочетами |
| satisfactory | Минимально допустимый уровень знаний. Допущено много негрубых ошибки. Продемонстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме. Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами. |
| unsatisfactory | Уровень знаний ниже минимальных требований. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки. При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки. |
| poor | Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа. Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа. Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа. |

6. Учебно-методическое и информационное обеспечение дисциплины (модуля)

Основная литература:

1. Таха Хэмди А. Введение в исследование операций : в 2 кн. Кн. 1 / пер. с англ. В. Я. Алтаева [и др.]. - М. : Мир, 1985. - 479 с. : ил. - 2.30., 3 экз.

2. Саати Томас Л. Принятие решений : метод анализа иерархий / пер. с англ. Р. Г. Вачнадзе. - М. : Радио и связь, 1993. - 314, [1] с. : ил. - ISBN 5-256-00443-3 : 300.00., 2 экз.
3. Ногин Владимир Дмитриевич. Принятие решений в многокритериальной среде : количеств. подход. - М. : Физматлит, 2002. - 176 с. : 20 ил. - ISBN 5-9221-0274-5 : 18.00., 1 экз.

Дополнительная литература:

1. Крылов В. В. Большие данные и их приложения в электроэнергетике : от бизнес-аналитики до виртуальных электростанций. - М. : Нобель Пресс, 2014. - 168 с. : ил. - ISBN 978-5-519-02625-3 : 100.00., 4 экз.

Программное обеспечение и Интернет-ресурсы (в соответствии с содержанием дисциплины):

СУБД PostgreSQL

Пакет линейной оптимизации LP_Solv

7. Материально-техническое обеспечение дисциплины (модуля)

Учебные аудитории для проведения учебных занятий, предусмотренных образовательной программой, оснащены мультимедийным оборудованием (проектор, экран), техническими средствами обучения, компьютерами.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Программа составлена в соответствии с требованиями ОС ННГУ по направлению подготовки/специальности 02.04.02 - Fundamental Informatics and Information Technology.

Author(s): Шапошников Дмитрий Евгеньевич, кандидат физико-математических наук.

Заведующий кафедрой: Баркалов Константин Александрович, доктор технических наук.

Программа одобрена на заседании методической комиссии от 13.12.2023, протокол № 3.