

МИНОБРНАУКИ РОССИИ

**Федеральное государственное бюджетное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"**

Кафедра математического обеспечения вычислительных систем

**Авторы-составители: Постановов Игорь Сергеевич
Чуприна Светлана Игоревна
Анисимов Александр Олегович
Рябинин Константин Валентинович
Юрков Кирилл Александрович**

Рабочая программа дисциплины

**СОВРЕМЕННЫЕ INTERNET-ТЕХНОЛОГИИ РЕШЕНИЯ АКТУАЛЬНЫХ
МАТЕМАТИЧЕСКИХ ЗАДАЧ**

Код УМК 92195

Утверждено
Протокол №5
от «09» июня 2020 г.

Пермь, 2020

1. Наименование дисциплины

Современные Internet-технологии решения актуальных математических задач

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в базовую часть Блока « М.1 » образовательной программы по направлениям подготовки (специальностям):

Направление: **01.04.02** Прикладная математика и информатика
направленность Анализ данных и принятие решений в экономике

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Современные Internet-технологии решения актуальных математических задач** у обучающегося должны быть сформированы следующие компетенции:

01.04.02 Прикладная математика и информатика (направленность : Анализ данных и принятие решений в экономике)

ОПК.1 Способен решать актуальные задачи фундаментальной и прикладной математики

Индикаторы

ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи

ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики

УК.6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

Индикаторы

УК.6.3 Осуществляет выбор направленности профессиональной деятельности в зависимости от собственных интересов, ресурсов и накопленного опыта

4. Объем и содержание дисциплины

Направления подготовки	01.04.02 Прикладная математика и информатика (направленность: Анализ данных и принятие решений в экономике)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	2
Объем дисциплины (з.е.)	3
Объем дисциплины (ак.час.)	108
Контактная работа с преподавателем (ак.час.), в том числе:	36
Проведение лекционных занятий	12
Проведение лабораторных работ, занятий по иностранному языку	24
Самостоятельная работа (ак.час.)	72
Формы текущего контроля	Входное тестирование (1) Защищаемое контрольное мероприятие (3) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Зачет (2 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Современные Internet-технологии решения актуальных математических задач

История развития, основные понятия в области обработки больших и сверхбольших наборов данных (Big Data)

Описание применения BigData в отличие от технологий WEB, CRM, ERP и т.д. Основные поставщики и пользователи решения BigData. Описание семейства решений Apache Hadoop 2.0. Основные особенности распределённой файловой системы HDFS. Основные элементы технологий: Map, Combine, Partitioning, Shuffle&Sort, Reduce. Преимущества и недостатки технологии Map Reduce. Варианты использования: MapReduce, Multiply MapReduce, MapOnly. Недостатки MapReduce при обработке графов. Ключевые особенности технологии Apache Giraph; API, ориентированное на работу с вершинами, парадигма Bulk Synchronous Parallel, суперстепы. Недостатки стандартной модели MRv2. Основные нововведения YARN: HDFS Federation, High Availability, система управления ресурсами кластера.

Технологии обработки больших данных: модель распределенной обработки данных MapReduce и линейка продуктов Apache Hadoop

Недостатки MapReduce. Ключевые особенности Spark: InMemory обработка распределённых наборов данных (RDD). стек технологий Spark: HDFS, Tachion, Spark, Spark Streaming. Graphs, Mllib, Shark. Ключевые концепции Spark: RDD, ленивые (lazy)-вычисления, отказоустойчивость за счёт поддержки родословной RDD-данных. Основные операции с RDD. Основные механизмы Spark SQL и их преимущества. Основные элементы NoSQL: таблица, записи переменной структуры, регионы. Балансировка регионов. Основные команды HBase. Основные элементы: Lily, HBase, Solr. Типизация полей и записей. Версионность типов записей. Варианты записи. Основные операции: GRUD, Scan, MapReduce. Механизмы полной и инкрементальной индексации. Поисковая система Solr: основные принципы индексации документов, типы поисковых запросов, поддержка многоядерности, режимы работы в кластере: classic, cloud.

Технологии Blockchain. Сетевые технологии Интернета вещей

Понятие Blockchain в смысле цепочка блоков. Дерево Меркеля. Понятие Blockchain в смысле распределенной системы. Типы отказоустойчивости. Консенсус. Примеры алгоритмов консенсуса. Трилемма блокчейна. Типы блокчейнов. Смартконтракты и языки их разработки. Практическая значимость технологии блокчейн. Понятие IoT. Ключевые компоненты IoT систем. Существующие технологии и протоколы. Понятие кибер-физических систем. Понятия Ubiquitous Computing, Cloud Computing, Fog Computing и Edge Computing. Вопросы организации интеллектуальных программно-аппаратных решений в Edge Computing.

Использование облачных платформ для решения актуальных математических задач

Облачные вычисления. Облачная платформа AWS

Облачные вычисления. Сравнение существующих облачных платформ. Основные типы предлагаемых услуг. Виртуальные облачные сервера. Автоматическое масштабирование вычислительных ресурсов. Брокеры сообщений. Сервисы аналитики Big Data. Сервисы создания, обучения и развёртывания моделей машинного обучения. Понимание и перевод текста. Поточковая обработка данных. Облачные реляционные и документноориентированные базы данных. Блочное хранилище данных. Архивное хранилище данных. Software as a Service (SaaS). Infrastructure as a service (IaaS). Platform as a Service (PaaS). Основы ценообразования. Serverless.

Технологии автоматизации развёртывания и управления приложениями в среде

виртуализации: Docker

Необходимость контейнеризации и общепринятые практики. Docker. Базовая архитектура. Создание образов. Структура Dockerfile. Допустимые команды Dockerfile. Понятие слоя. Многоэтапная сборка. Управление образами. Настройка сети и видимости контейнеров. Виды сетей. Подпись образов. Хранение данных в томах. Логгирование. Реестры образов. Аутентификация в реестре образов. Загрузка, выгрузка и удаление образов при работе с реестром. Запуск образов. Docker Compose. Проблемы оркестрации. Docker Swarm.

Apache Spark. Распределенная пакетная и потоковая обработка данных. Spark SQL. Apache Kafka. Spark Streaming. Применение облачных сервисов для кластерных вычислений. Контейнеризация.

Архитектура Spark. Основные особенности Spark Streaming: потоковая обработка, обработка данных по временным интервалам - Discretized Stream (Dstream), Sliding Windows. Преимущества Spark Streaming. Основные операции с потоковыми RDD. Сравнение с другими системами потоковой обработки данных. Знание принципов разработки и запуска приложений в стеке технологий Apache Spark при использовании систем очередей сообщений.

Умение разрабатывать приложения потоковой обработки в стеке технологий Apache Spark с использованием системы очередей сообщений Apache Kafka.

Реализация алгоритма k-ближайших соседей с помощью Spark SQL, SQL-скрипт из databricks

Машинное обучение. Предобработка и очистка данных. Платформы для решения задач классификации и кластеризации данных

Общий цикл проекта интеллектуального анализа данных. EDA. Основные типы проблем исходных данных и их источники. Понятие машинного обучения. Основные подходы в машинном обучении. Библиотеки машинного обучения: особенности, преимущества и недостатки. Аппаратная поддержка машинного обучения. Алгоритмы очистки данных. Автоматизированные и автоматические средства по очистке данных. Использование диаграмм потоков данных при конфигурировании правил очистки. Свободно распространяемые и коммерческие платформы для решения задач классификации и кластеризации данных.

Роль и место онтологий в современном Internet и проектах Big Data

Онтологии. Онтологический инжиниринг. Big Data, их ключевые характеристики (от 3 "V" к 42 "V"). Основные технологии Big Data. Роль визуализации в Big Data. Переход от визуализации к визуальной аналитике. Понятие когнитивной графики. Примеры реальных задач аналитики Big Data. Примеры инструментов для эффективного решения таких задач. Принципы построения онтологически управляемых программных решений для работы с Big Data.

Зачёт

Проверка

1. Освоения современных облачных платформ для решения задач аналитики данных
2. Знания технологий Apache Hadoop 2.0: HDFS, MapReduce, HBase, Solr
3. Навыков решения аналитических задач на базе Spark, Spark SQL, Spark Streaming
4. Знания ключевых понятий в области Интернета вещей
5. Знания основ онтологического инжиниринга и принципов построения онтологически управляемых программных решений

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторские занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Воронова Л. И. Big Data. Методы и средства анализа: Учебное пособие/Воронова Л. И..- Москва:Московский технический университет связи и информатики,2016.-33.
<http://www.iprbookshop.ru/61463.html>
2. Гергель В. П. Высокопроизводительные вычисления для многопроцессорных многоядерных систем:учебник для студентов высших учебных заведений, обучающихся по направлениям ВПО 010400 "Прикладная математика и информатика" и 010300 "Фундаментальная информатика и информационные технологии"/В.П. Гергель.-Москва:Изд-во Московского университета,2010, ISBN 978-5-211-05937-5.- 5394.-Библиогр.: с. 534-539

Дополнительная:

1. Valentina Janev. Knowledge Graphs and Big Data Processing / Valentina Janev, Damien Graux, Hajira Jabeen, Emanuel Sallinger // — 2020. — 209 p. — ISBN978-3-030-53199-7. [Электронный ресурс].
<https://link.springer.com/book/10.1007/978-3-030-53199-7>
2. Суперкомпьютерные технологии в науке, образовании и промышленности:альманах/Моск. гос. ун-т им. М. В. Ломоносова, Суперкомпьютерный консорциум ун-тов России, Рос. акад. наук.- Москва:Издательство Московского университета,2012, ISBN 978-5-211-06528-4.-232.
3. Высокопроизводительные вычисления на графических процессорах:материалы II Всероссийской научно-практической конференции с международным участием с элементами научной школы для молодежи 2-6 июня 2014 года/М-во образования и науки РФ, Перм. гос. нац. исслед. ун-т.- Пермь:Пермский государственный национальный исследовательский университет,2014, ISBN 978-5-7944-2341-9.-65.

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<http://spark.apache.org/docs/latest/programming-guide.html> Spark Programming Guide

<http://spark.apache.org/docs/latest/streaming-programming-guide.html> Spark Streaming Programming Guide

<http://spark.apache.org/docs/latest/graphx-programming-guide.html> GraphX Programming Guide

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Современные Internet-технологии решения актуальных математических задач** предполагает использование следующего программного обеспечения и информационных справочных систем:

доступ в режиме on-line в Электронную библиотечную систему (ЭБС);

доступ в электронную информационно-образовательную среду университета.

Необходимое лицензионное и (или) свободно распространяемое программное обеспечение:

Пакет JetBrains: DataGrip, PyCharm Community, IntelliJ Idea Ultimate

При освоении материала и выполнения заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (student.psu.ru).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).

система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.

система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для лекционных занятий требуется аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения лабораторных занятий - меловая и (или) маркерная доска, компьютерный класс (аппаратное и программное обеспечение определено в Паспортах компьютерных классов)

Для групповых (индивидуальных) консультаций - аудитория, оснащенная меловой (и) или маркерной доской.

Для проведения текущего контроля - аудитория, оснащенная меловой (и) или маркерной доской.

Самостоятельная работа студентов: аудитория, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», с обеспеченным доступом в электронную информационно-образовательную среду университета, помещения Научной библиотеки ПГНИУ.

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с доступом к локальной и глобальной компьютерным сетям.

2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с

доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Современные Internet-технологии решения актуальных математических задач**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ОПК.1

Способен решать актуальные задачи фундаментальной и прикладной математики

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи</p>	<p>Знает: основные понятия в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения, распределенной пакетной и потоковой обработки данных; понятие облачных вычислений и основные модели обслуживания облачных сервисов; модель распределенной обработки данных MapReduce и линейку продуктов Apache Hadoop; основные составляющие технологии Blockchain; основные понятия и стек технологий Интернета вещей (IoT); что такое контейнеризация, классификацию и примеры контейнеров; роль и место онтологий в современном Internet и проектах Big Data.</p>	<p align="center">Неудовлетворител</p> <p>Отсутствие знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения, распределенной пакетной и потоковой обработки данных; понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT); не имеет представление о том, что такое контейнеризация, не знает классификацию и примеры контейнеров; роль и место онтологий в современном Internet и проектах Big Data.</p> <p align="center">Удовлетворительн</p> <p>Сформированы базовые знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения, распределенной пакетной и потоковой обработки данных; понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT); имеет представление о том, что такое контейнеризация, знает классификацию и примеры контейнеров; роль и место онтологий в современном Internet и проектах Big Data. При этом студент допускает ошибки в определении основных понятий</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Удовлетворительн</p> <p>курса или не знает менее половины понятий из пяти основных разделов курса (технологии Big Data, машинное обучение, облачные вычисления, контейнеризация, IoT, Blockchain).</p> <p style="text-align: center;">Хорошо</p> <p>Сформированы знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения, распределенной пакетной и потоковой обработки данных; понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT); имеет представление о том, что такое контейнеризация, знает классификацию и примеры контейнеров; роль и место онтологий в современном Internet и проектах Big Data. При этом студент допускает неточности в определении основных понятий курса.</p> <p style="text-align: center;">Отлично</p> <p>Сформированы систематические знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения, распределенной пакетной и потоковой обработки данных; понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT); имеет представление о том, что такое контейнеризация, знает классификацию и примеры контейнеров; роль и место онтологий в современном Internet и проектах Big Data.</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики</p>	<p>Знает: понятие облачных вычислений и основные модели обслуживания облачных сервисов; модель распределенной обработки данных MapReduce и линейку продуктов Apache Hadoop.</p> <p>Умеет использовать облачные платформы для решения актуальных математических задач: выполнять предобработку, очистку и анализ данных в среде открытых аналитических и облачных платформ; использовать Spark SQL для анализа данных; развёртывать и управлять приложениями в среде виртуализации Docker; строить онтологии в среде визуальных редакторов онтологий для управляемых онтологиями систем.</p> <p>Владеет навыками анализа данных в среде доступных облачных платформ</p>	<p>Неудовлетворител</p> <p>Не знает понятие облачных вычислений и основные модели обслуживания облачных сервисов; модель распределенной обработки данных MapReduce и линейку продуктов Apache Hadoop.</p> <p>Не умеет использовать облачные платформы для решения актуальных математических задач: выполнять предобработку, очистку и анализ данных в среде открытых аналитических и облачных платформ; использовать Spark SQL для анализа данных; развёртывать и управлять приложениями в среде виртуализации Docker; строить онтологии в среде визуальных редакторов онтологий для управляемых онтологиями систем.</p> <p>Не владеет навыками анализа данных в среде доступных облачных платформ.</p> <p>Удовлетворительн</p> <p>Сформированы базовые знания понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop. При этом студент допускает ошибки в определении основных понятий курса.</p> <p>В целом умеет использовать облачные платформы для решения актуальных математических задач: выполнять предобработку, очистку и анализ данных в среде открытых аналитических и облачных платформ; использовать Spark SQL для анализа данных; развёртывать и управлять приложениями в среде виртуализации Docker; строить онтологии в среде визуальных редакторов онтологий для управляемых онтологиями систем.</p> <p>Владеет начальными навыками анализа данных в среде доступных облачных платформ.</p> <p>Студент выполнил минимальный объем заданий по аналитике данных; в отчетах отсутствует анализ результатов.</p> <p>Хорошо</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо</p> <p>Сформированы знания понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop. При этом студент допускает неточности в определении основных понятий курса.</p> <p>Умеет на достаточном уровне использовать облачные платформы для решения актуальных математических задач: выполнять предобработку, очистку и анализ данных в среде открытых аналитических и облачных платформ; использовать Spark SQL для анализа данных; развёртывать и управлять приложениями в среде виртуализации Docker; строить онтологии в среде визуальных редакторов онтологий для управляемых онтологиями систем.</p> <p>Владеет базовыми навыками анализа данных в среде доступных облачных платформ. При выполнении практических заданий в среде открытых аналитических платформ студент допускает упрощение заданий и/или некачественную очистку данных, что ухудшает качество анализа; в отчетах допускает поверхностный анализ результатов.</p> <p style="text-align: center;">Отлично</p> <p>Сформированы систематические знания понятия облачных вычислений и основных моделей обслуживания облачных сервисов; модели распределенной обработки данных MapReduce и линейки продуктов Apache Hadoop.</p> <p>В совершенстве умеет использовать облачные платформы для решения актуальных математических задач: выполнять предобработку, очистку и анализ данных в среде открытых аналитических и облачных платформ; использовать Spark SQL для анализа данных; развёртывать и управлять приложениями в среде виртуализации Docker; строить онтологии в</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p align="center">Отлично</p> <p>среде визуальных редакторов онтологий для управляемых онтологиями систем. Уверенно владеет навыками анализа данных в среде доступных облачных платформ; выполнил и составил отчеты по всем практическим заданиям.</p>

УК.6

Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>УК.6.3 Осуществляет выбор направленности профессиональной деятельности в зависимости от собственных интересов, ресурсов и накопленного опыта</p>	<p>Знать основные понятия в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения; платформы для решения актуальных математических задач, в том числе задач классификации и кластеризации; основные составляющие технологии Blockchain; основные понятия и стек технологий Интернета вещей (IoT). Уметь оценивать накопленный опыт и доступные ресурсы при выборе направленности дальнейшего профессионального развития. Владеть навыком выбирать инструмент при решении профессиональных аналитических задач в зависимости от доступных ресурсов и накопленного опыта.</p>	<p align="center">Неудовлетворител</p> <p>Отсутствие знаний основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения; платформ для решения актуальных математических задач, в том числе задач классификации и кластеризации; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT). Отсутствие умения оценивать накопленный опыт и доступные ресурсы при выборе направленности дальнейшего профессионального развития. Отсутствие навыка выбирать инструмент при решении профессиональных аналитических задач.</p> <p align="center">Удовлетворительн</p> <p>Общие, но не структурированные знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения; платформ для решения актуальных математических задач, в том числе задач классификации и кластеризации; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT). Частично сформированное умение оценивать накопленный опыт и доступные ресурсы при выборе направленности дальнейшего профессионального развития. Фрагментарное применение навыка</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Удовлетворительн</p> <p>выбирать инструмент при решении профессиональных аналитических задач в зависимости от доступных ресурсов и накопленного опыта.</p> <p style="text-align: center;">Хорошо</p> <p>Сформированные, но содержащие отдельные пробелы знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения; платформ для решения актуальных математических задач, в том числе задач классификации и кластеризации; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT).</p> <p>В целом успешные, но содержащие отдельные пробелы умения оценивать накопленный опыт и доступные ресурсы при выборе направленности дальнейшего профессионального развития.</p> <p>В целом успешное, но содержащее отдельные пробелы применение навыка выбирать инструмент при решении профессиональных аналитических задач в зависимости от доступных ресурсов и накопленного опыта.</p> <p style="text-align: center;">Отлично</p> <p>Сформированные систематические знания основных понятий в области обработки больших и сверхбольших наборов данных (Big Data), машинного обучения; платформ для решения актуальных математических задач, в том числе задач классификации и кластеризации; основных составляющих технологии Blockchain; основных понятий и стека технологий Интернета вещей (IoT).</p> <p>Сформированное умение оценивать накопленный опыт и доступные ресурсы при выборе направленности дальнейшего профессионального развития.</p> <p>Успешное и систематическое применение навыка выбирать инструмент при решении профессиональных аналитических задач в зависимости от доступных ресурсов и</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		Отлично накопленного опыта.

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Зачет

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 50 до 60

«неудовлетворительно» / «незачтено» менее 50 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
Входной контроль	История развития, основные понятия в области обработки больших и сверхбольших наборов данных (Big Data) Входное тестирование	Тест с вопросами закрытого и открытого типов, направленный на определение уровня подготовки студента в области современных информационных технологий.
ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи	Технологии Blockchain. Сетевые технологии Интернета вещей Защищаемое контрольное мероприятие	Знает: основные составляющие технологии Blockchain; основные понятия и стек технологий Интернета вещей (IoT); роль и место онтологий в современном Internet и проектах Big Data. Умеет строить онтологии в среде визуальных редакторов онтологий для управляемых онтологиями систем.
ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи	Apache Spark. Распределенная пакетная и потоковая обработка данных. Spark SQL. Apache Kafka. Spark Streaming. Применение облачных сервисов для кластерных вычислений. Контейнеризация. Защищаемое контрольное мероприятие	Знание принципов разработки и запуска приложений в стеке технологий Apache Spark. Умение разрабатывать приложения пакетной обработки в стеке технологий Apache Spark: реализация алгоритма k-ближайших соседей с помощью Spark SQL, SQL-скрипт из databricks

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
<p>ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики</p> <p>ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи</p>	<p>Машинное обучение. Предобработка и очистка данных. Платформы для решения задач классификации и кластеризации данных</p> <p>Защищаемое контрольное мероприятие</p>	<p>Знание основных этапов создания модели машинного обучения. Умение использовать платформы для предобработки и очистки данных, решения задач классификации и кластеризации данных.</p>
<p>ОПК.1.1 Осуществляет выбор конкретного метода фундаментальной и/или прикладной математики для решения конкретной задачи</p> <p>ОПК.1.2 Применяет информационные технологии для решения задачи фундаментальной и/или прикладной математики</p> <p>УК.6.3 Осуществляет выбор направленности профессиональной деятельности в зависимости от собственных интересов, ресурсов и накопленного опыта</p>	<p>Зачёт</p> <p>Итоговое контрольное мероприятие</p>	<p>Тест, направленный на определение знаний и навыков в области Big Data, Интернета вещей, методов MapReduce, технологий Blockchain, платформ обработки Big Data, машинного обучения, предобработки данных, очистки данных, классификации и кластеризации данных, онтологического инжиниринга.</p>

Спецификация мероприятий текущего контроля

История развития, основные понятия в области обработки больших и сверхбольших наборов данных (Big Data)

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **0**

Проходной балл: **0**

Показатели оценивания	Баллы
Знание основных методов в области современных информационных технологий.	5.9
Владение основными терминами и понятиями в области современных информационных технологий.	4.1

Технологии Blockchain. Сетевые технологии Интернета вещей

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **14**
 Проходной балл: **7**

Показатели оценивания	Баллы
Создана корректная программа визуализации данных, управляемая онтологией задачи	7
Представлен отчёт, в котором со скриншотами и ссылками представлены основные этапы работы с выбранным блокчейном на выбранном тестнете	4
Создана корректная онтология задачи, управляющая программой визуализации данных	3

Apache Spark. Распределенная пакетная и потоковая обработка данных. Spark SQL. Apache Kafka. Spark Streaming. Применение облачных сервисов для кластерных вычислений. Контейнеризация.

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **39**
 Проходной балл: **19**

Показатели оценивания	Баллы
Приложение решает соответствующую варианту задачу	25
Приложение решает соответствующую варианту задачу, аналогичную задаче из предыдущего текущего контроля	8
Приложение считывает входные данные из HDFS	3
Приложение записывает выходные данные в HDFS	3

Машинное обучение. Предобработка и очистка данных. Платформы для решения задач классификации и кластеризации данных

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**
 Проходной балл: **15**

Показатели оценивания	Баллы
В соответствии с вариантом решена задача классификации или кластеризации с использованием современных технологии.	15
Для решения задачи корректно использованы технологии предобработки и очистки данных.	10
Отчёт содержит подробное описание процесса выполнения задачи, сопровождаемое скриншотами и комментариями.	5

Зачёт

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**
 Условия проведения мероприятия: **в часы аудиторной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **17**
 Проходной балл: **8.5**

Показатели оценивания	Баллы
Знания и навыки в области классификации и кластеризации данных	2
Знания и навыки в области Интернета вещей	2
Знания и навыки в области методов MapReduce	2
Знания и навыки в области технологий Blockchain	2
Знания и навыки в области онтологического инжиниринга	2
Знания и навыки в области Big Data	2
Знания и навыки в области применения машинного обучения	1.5
Знания и навыки в области применения платформ обработки Big Data	1.5
Знания и навыки в области применения предобработки данных	1
Знания и навыки в области применения очистки данных	1