

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования



**Пермский национальный исследовательский
политехнический университет**

УТВЕРЖДАЮ

Проректор по образовательной
деятельности

 А.Б. Петроченков

« 10 » октября 20 23 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина: Информационные технологии и методы обработки больших
объемов данных

(наименование)

Форма обучения: очная

(очная/очно-заочная/заочная)

Уровень высшего образования: бакалавриат

(бакалавриат/специалитет/магистратура)

Общая трудоёмкость: 144 (4)

(часы (ЗЕ))

Направление подготовки: 09.03.01 Информатика и вычислительная техника

(код и наименование направления)

Направленность: Информатика и вычислительная техника (общий профиль,
СУОС)

(наименование образовательной программы)

1. Общие положения

1.1. Цели и задачи дисциплины

Одно из стремительно развивающихся направлений IT-технологий — это большие данные — Big Data. И если само понятие возникло сравнительно давно, то понимание что это такое оформилось недавно. Оказалось, это не объёмы хранимых данных, точнее не только объемы. Это не те данные, с которыми привык работать опытный аналитик. Они скорее всего не будут вписываться в строки и столбцы таблицы электронных таблиц или строки таблицы реляционной базы данных. И они, вероятно, не будут помещаться на жесткий диск обычного компьютера.

1.2. Изучаемые объекты дисциплины

Анализ и развитие технологий обработки больших данных в научных исследованиях
Анализ методов и подходов к технологиям обработки больших данных
Архитектура технологий обработки больших данных
Решение практических задач с помощью технологии обработки больших данных

1.3. Входные требования

Не предусмотрены

2. Планируемые результаты обучения по дисциплине

Компетенция	Индекс индикатора	Планируемые результаты обучения по дисциплине (знать, уметь, владеть)	Индикатор достижения компетенции, с которым соотнесены планируемые результаты обучения	Средства оценки
ПК-2.1	ИД-1ПК-2.1	Знает оценку производительности критических приложений, наиболее влияющих на производительность сетевых устройств и программного обеспечения в целом; планирование требуемой производительности сетевых устройств и программного обеспечения администрируемой сети.	Знает оценку производительности критических приложений, наиболее влияющих на производительность сетевых устройств и программного обеспечения в целом; планирование требуемой производительности сетевых устройств и программного обеспечения администрируемой сети.	Защита лабораторной работы

Компетенция	Индекс индикатора	Планируемые результаты обучения по дисциплине (знать, уметь, владеть)	Индикатор достижения компетенции, с которым соотнесены планируемые результаты обучения	Средства оценки
ПК-2.1	ИД-2ПК-2.1	Умеет оценивать требуемую производительность сетевых устройств и программного обеспечения администрируемой сети; вести специальный документ об оценке готовности системы.	Умеет оценивать требуемую производительность сетевых устройств и программного обеспечения администрируемой сети; вести специальный документ об оценке готовности системы.	Зачет
ПК-2.1	ИД-3ПК-2.1	Владеет навыками использования утилит операционных систем; установки дополнительных программных продуктов и их параметризации.	Владеет навыками использования утилит операционных систем; установки дополнительных программных продуктов и их параметризации.	Дифференцированный зачет

3. Объем и виды учебной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		5	
1. Проведение учебных занятий (включая проведение текущего контроля успеваемости) в форме:	58	58	
1.1. Контактная аудиторная работа, из них:			
- лекции (Л)	18	18	
- лабораторные работы (ЛР)	20	20	
- практические занятия, семинары и (или) другие виды занятий семинарского типа (ПЗ)	18	18	
- контроль самостоятельной работы (КСР)	2	2	
- контрольная работа			
1.2. Самостоятельная работа студентов (СРС)	50	50	
2. Промежуточная аттестация			
Экзамен	36	36	
Дифференцированный зачет			
Зачет			
Курсовой проект (КП)			
Курсовая работа (КР)			
Общая трудоемкость дисциплины	144	144	

4. Содержание дисциплины

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	
				СРС

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
5-й семестр				
Анализ систем обработки больших данных	2	0	2	4
Объем генерируемых данных. Многообразие больших данных Скорость генерирования или обработки данных Изменчивость, нестабильность данных во времени. Достоверность, собранных данных Трудоемкость процесса корреляции и построения взаимосвязей между данными.				
Современные методы визуализации информации в приложениях Big Data	2	4	2	4
сложная столбчатая диаграмма хордовая диаграмма солнечная диаграмма поточковый график матричная диаграмма диаграмма рассеяния диаграмма ohlc				
Задачи параллельной обработки данных в научных исследованиях	2	4	2	6
проблемные точки процесса реализации программы параллельные структуры алгоритмов параллелизм на уровне инструкций параллелизм на уровне задач параллелизм на уровне алгоритмов				
Фреймворк Apache Spark как способ обработки больших объёмов данных	2	2	2	6
горизонтальная масштабируемость основная технология распараллеливания кластеры среды hadoop особенности системы хранения главные конкуренты и вычислительные мощност				
Фреймворк mapreduce как способ обработки больших объёмов данных	2	2	2	6
преимущества и недостатки mapreduce возможности восстановления работы после сбоя узлов примеры работы mapreduce				
Анализ массивов и временных рядов	2	2	2	6
Rolling window estimations Экспоненциальное сглаживание, модель Хольта-Винтерса Кросс-валидация на временных рядах, подбор параметров Линейные модели на временных рядах Извлечение признаков (Feature extraction) Линейная регрессия vs XGBoost				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
Вопросы хранения больших данных	2	2	2	6
вопросы безопасности больших данных резервное копирование облака собственные алгоритмы шифрования типичные проблемы				
Извлечение, сбор и валидация Big Data	2	2	2	6
Определение проблемы; Извлечение данных; Подготовка данных — очистка данных; Подготовка данных — преобразование данных; Предсказательные модели				
Практические примеры применения Big Data	2	2	2	6
Обзор примеров практического применения, успешные и не успешные варианты				
ИТОГО по 5-му семестру	18	20	18	50
ИТОГО по дисциплине	18	20	18	50

Тематика примерных практических занятий

№ п.п.	Наименование темы практического (семинарского) занятия
1	Проектирование и подготовка тестов к сборщику данных
2	Проектирование системы хранения БД. Кластер данных.
3	Постановка и проектирование задачи исследования данных. Машинное обучение

Тематика примерных лабораторных работ

№ п.п.	Наименование темы лабораторной работы
1	Реализация алгоритма извлечения данных. Тестирование.
2	Настройка одной из современных систем хранения данных
3	Поиск закономерностей в данных. Нейросети.

5. Организационно-педагогические условия

5.1. Образовательные технологии, используемые для формирования компетенций

Проведение лекционных занятий по дисциплине основывается на активном методе обучения, при котором учащиеся не пассивные слушатели, а активные участники занятия, отвечающие на вопросы преподавателя. Вопросы преподавателя нацелены на активизацию процессов усвоения материала, а также на развитие логического мышления. Преподаватель заранее намечает список вопросов, стимулирующих ассоциативное мышление и установление связей с ранее освоенным материалом.

Практические занятия проводятся на основе реализации метода обучения действием: определяются проблемные области, формируются группы. При проведении практических занятий преследуются следующие цели: применение знаний отдельных дисциплин и креативных методов для решения проблем и принятия решений; отработка у обучающихся навыков командной работы, межличностных коммуникаций и развитие лидерских качеств; закрепление основ теоретических знаний.

Проведение лабораторных занятий основывается на интерактивном методе обучения, при котором обучающиеся взаимодействуют не только с преподавателем, но и друг с другом. При этом доминирует активность учащихся в процессе обучения. Место преподавателя в интерактивных занятиях сводится к направлению деятельности обучающихся на достижение целей занятия.

При проведении учебных занятий используются интерактивные лекции, групповые дискуссии, ролевые игры, тренинги и анализ ситуаций и имитационных моделей.

5.2. Методические указания для обучающихся по изучению дисциплины

При изучении дисциплины обучающимся целесообразно выполнять следующие рекомендации:

1. Изучение учебной дисциплины должно вестись систематически.
2. После изучения какого-либо раздела по учебнику или конспектным материалам рекомендуется по памяти воспроизвести основные термины, определения, понятия раздела.
3. Особое внимание следует уделить выполнению отчетов по практическим занятиям, лабораторным работам и индивидуальным комплексным заданиям на самостоятельную работу.
4. Вся тематика вопросов, изучаемых самостоятельно, задается на лекциях преподавателем. Им же даются источники (в первую очередь вновь изданные в периодической научной литературе) для более детального понимания вопросов, озвученных на лекции.

6. Перечень учебно-методического и информационного обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Печатная учебно-методическая литература

№ п/п	Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)	Количество экземпляров в библиотеке
1. Основная литература		
1	Мэтлофф Н. Искусство программирования на R. Погружение в большие данные : пер. с англ. / Н. Мэтлофф. - Санкт-Петербург [и др.]: Питер, 2019.	2
2	Плас Дж. В. Python для сложных задач: наука о данных и машинное обучение : пер. с англ. / Дж. В. Плас. - Санкт-Петербург [и др.]: Питер, 2018.	1

3	Прохоренок Н. А. Python 3 и PyQt. Разработка приложений / Н. А. Прохоренок. - Санкт-Петербург: БХВ-Петербург, 2014.	5
4	Сенько А. Работа с BigData в облаках. Обработка и хранение данных с примерами из Microsoft Azure / А. Сенько. - Санкт-Петербург [и др.]: Питер, 2019.	2
5	Бн А. Теоретический минимум по Big Data. Все что нужно знать о больших данных : пер. с англ. / А. Бн, К. Су. - Санкт-Петербург [и др.]: Питер, 2019.	2
2. Дополнительная литература		
2.1. Учебные и научные издания		
1	2nd International Conference of Young Scientists Solutions of Applied Problems in Control, Data Processing and Data Analysis, Koethen, Germany, 4 April 2011 / Perm State Technical University ; Anhalt University of Applied Sciences ; Ed. U. Fissgus. - Perm: PSTU publ., 2011.	2
2	Proceedings of the International Symposium on Numerical Reference Data for Science and Technology, Warsaw, August 29-30, 1969 / Polskiej Akademii Nauk ; Zaklad Narodowy im. Ossolinskiх ; Ed. by J. Malczynski. - Warszawa: Akademii Nauk, 1973.	1
3	О'Нил К. Data Science. Инсайдерская информация для новичков. Включая язык R / К. О'Нил, Р. Шатт. - Санкт-Петербург [и др.]: Питер, 2019.	1
4	Ручкин В. Н. Системы искусственного интеллекта. Нейросети и нейрокомпьютеры : учебник для вузов / В. Н. Ручкин, Б. В. Костров, А. Г. Свирина. - Москва: КУРС, 2019.	2
5	Сузи Р. А. Язык программирования PYTHON : учебное пособие / Р. А. Сузи. - Москва: ИНТУИТ, БИНОМ. Лаб. знаний, 2006.	3
6	Усков А. А. Интеллектуальные технологии управления. Искусственные нейронные сети и нечёткая логика / А. А. Усков, А. В. Кузьмин. - Москва: Горячая линия-Телеком, 2004.	1
7	Шамис А.Л. Пути моделирования мышления. Активные синергические нейронные сети. Мышление и творчество. Формальные модели поведения и распознавания с пониманием / А.Л. Шамис. - М.: КомКнига, 2006.	1
2.2. Периодические издания		
	Не используется	
2.3. Нормативно-технические издания		
	Не используется	
3. Методические указания для студентов по освоению дисциплины		
	Не используется	
4. Учебно-методическое обеспечение самостоятельной работы студента		
	Не используется	

6.2. Электронная учебно-методическая литература

Вид литературы	Наименование разработки	Ссылка на информационный ресурс	Доступность (сеть Интернет / локальная сеть; авторизованный / свободный доступ)
Дополнительная литература	Big Data и ECM: рассмотрим практические примеры	https://ecm-journal.ru/post/Big-Data-i-ECM-rassmotrim-prakticheskie-primery.aspx	сеть Интернет; свободный доступ
Дополнительная литература	Обзор кейсов интересных внедрений Big Data в компаниях финансового сектора	https://habr.com/ru/post/353008/	сеть Интернет; свободный доступ

6.3. Лицензионное и свободно распространяемое программное обеспечение, используемое при осуществлении образовательного процесса по дисциплине

Вид ПО	Наименование ПО
Операционные системы	Debian (GNU GPL)
Прикладное программное обеспечение общего назначения	Scilab лиц.GNU GPL v2
Среды разработки, тестирования и отладки	PIP (The Python Package Installer) Free
Среды разработки, тестирования и отладки	PostgreSQL (PostgreSQL License)
Среды разработки, тестирования и отладки	Язык R

6.4. Современные профессиональные базы данных и информационные справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Наименование	Ссылка на информационный ресурс
База данных Web of Science	http://www.webofscience.com/
База данных научной электронной библиотеки (eLIBRARY.RU)	https://elibrary.ru/
Научная библиотека Пермского национального исследовательского политехнического университета	http://lib.pstu.ru/
Электронно-библиотечная система Лань	https://e.lanbook.com/
Электронно-библиотечная система IPRbooks	http://www.iprbookshop.ru/
Информационные ресурсы Сети КонсультантПлюс	http://www.consultant.ru/

7. Материально-техническое обеспечение образовательного процесса по дисциплине

Вид занятий	Наименование необходимого основного оборудования и технических средств обучения	Количество единиц
-------------	---	-------------------

Вид занятий	Наименование необходимого основного оборудования и технических средств обучения	Количество единиц
Лабораторная работа	Компьютер	15
Лекция	Проектор	1
Практическое занятие	Компьютер	15

8. Фонд оценочных средств дисциплины

Описан в отдельном документе

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего
образования
**«Пермский национальный исследовательский политехнический
университет»**

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

**для проведения промежуточной аттестации обучающихся по дисциплине
«Информационные технологии и методы обработки больших объемов
данных»**

Приложение к рабочей программе дисциплины

Направление подготовки:	09.03.01 Информатика и вычислительная техника
Направленность (профиль) образовательной программы:	Информатика и вычислительная техника (общий профиль, СУОС)
Квалификация выпускника:	бакалавр
Выпускающая кафедра:	Информационных технологий и автоматизированных систем
Форма обучения:	очная

Курс: 3

Семестр: 5

Трудоёмкость:

Кредитов по рабочему учебному плану: 4

Часов по рабочему учебному плану: 144

Форма промежуточной аттестации:

Экзамен

Пермь 2022 г.

Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине является частью (приложением) к рабочей программе дисциплины. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине разработан в соответствии с общей частью фонда оценочных средств для проведения промежуточной аттестации основной образовательной программы, которая устанавливает систему оценивания результатов промежуточной аттестации и критерии выставления оценок. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине устанавливает формы и процедуры текущего контроля успеваемости и промежуточной аттестации, обучающихся по дисциплине.

1. Перечень контролируемых результатов обучения по дисциплине, объекты оценивания и виды контроля

Согласно РПД освоение учебного материала дисциплины запланировано в течение одного семестра (3-го семестра учебного плана) и разбито на 4 учебных модуля. В каждом модуле предусмотрены аудиторские лекционные и лабораторные занятия, а также самостоятельная работа студентов. В рамках освоения учебного материала дисциплины формируется компоненты компетенций *знать, уметь, владеть*, указанные в РПД, которые выступают в качестве контролируемых результатов обучения по дисциплине (таблица 1.1).

Контроль уровня усвоенных знаний, усвоенных умений и приобретенных владений осуществляется в рамках текущего, рубежного и промежуточного контроля при изучении теоретического материала, сдаче отчетов по лабораторным работам. Виды контроля сведены в таблицу 1.1.

Таблица 1.1. Перечень контролируемых результатов обучения по дисциплине

Контролируемые результаты обучения по дисциплине (ЗУВы)	Вид контроля		
	Текущий	Промежуточный /рубежный	Итоговый
	ТО	ОЛР	Зачет
Усвоенные знания			
З.1 Знает оценку производительности критических приложений, наиболее влияющих на производительность сетевых устройств и программного обеспечения в целом; планирование требуемой производительности сетевых устройств и программного обеспечения администрируемой сети.	ТО1	ОЛР1-ОЛР7	по результатам текущего и рубежного контроля
Освоенные умения			
У.1 Умеет оценивать требуемую производительность сетевых устройств и программного обеспечения администрируемой сети; вести специальный документ об оценке готовности системы.		ОЛР1- ОЛР7	по результатам текущего и рубежного контроля
Приобретенные владения			

В.1 Владеет навыками использования утилит операционных систем; установки дополнительных программных продуктов и их параметризации.		ОЛР8	по результатам текущего и рубежного контроля
--	--	------	--

ТО – коллоквиум (теоретический опрос); ОЛР – отчет по лабораторной работе.

Итоговой оценкой достижения результатов обучения по дисциплине является промежуточная аттестация в виде зачета, проводимая с учетом результатов текущего и рубежного контроля.

2. Виды контроля, типовые контрольные задания и шкалы оценивания результатов обучения

Текущий контроль успеваемости имеет целью обеспечение максимальной эффективности учебного процесса, управление процессом формирования заданных компетенций обучаемых, повышение мотивации к учебе и предусматривает оценивание хода освоения дисциплины. В соответствии с Положением о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по образовательным программам высшего образования – программам бакалавриата, специалитета и магистратуры в ПНИПУ предусмотрены следующие виды и периодичность текущего контроля успеваемости обучающихся:

- входной контроль, проверка исходного уровня подготовленности обучаемого и его соответствия предъявляемым требованиям для изучения данной дисциплины;
- текущий контроль усвоения материала (уровня освоения компонента «знать» заданных компетенций) на каждом групповом занятии и контроль посещаемости лекционных занятий;
- промежуточный и рубежный контроль освоения обучаемыми отдельных компонентов «знать», «уметь» заданных компетенций путем компьютерного или бланчного тестирования, контрольных опросов, контрольных работ (индивидуальных домашних заданий), защиты отчетов по лабораторным работам, рефератов, эссе и т.д.

Рубежный контроль по дисциплине проводится на следующей неделе после прохождения модуля дисциплины, а промежуточный – во время каждого контрольного мероприятия внутри модулей дисциплины;

- контроль остаточных знаний.

2.1. Текущий контроль усвоения материала

Текущий контроль усвоения материала в форме собеседования или выборочного теоретического опроса студентов проводится по каждой теме. Результаты по 4-балльной шкале оценивания заносятся в книжку преподавателя и учитываются в виде интегральной оценки при проведении промежуточной аттестации.

2.2. Рубежный (промежуточный) контроль

Рубежный (промежуточный) контроль для комплексного оценивания усвоенных знаний, освоенных умений и приобретенных владений (таблица 1.1) проводится в форме защиты лабораторных работ.

2.2.1. Защита лабораторных работ

Всего запланировано 8 лабораторных работ. Типовые темы лабораторных работ приведены в РПД.

Защита лабораторной работы проводится индивидуально каждым студентом или группой студентов. Типовые шкала и критерии оценки приведены в общей части ФОС образовательной программы.

2.3. Промежуточная аттестация (итоговый контроль)

Допуск к промежуточной аттестации осуществляется по результатам текущего и рубежного контроля. Условиями допуска являются успешная сдача всех лабораторных работ и положительная интегральная оценка по результатам текущего и рубежного контроля.

Промежуточная аттестация, согласно РПД, проводится в виде зачета по результатам текущего и рубежного контроля.

3. . Шкалы оценивания результатов обучения на экзамене

Оценка результатов обучения по дисциплине в форме уровня сформированности компонентов *знать, уметь, владеть* заявленных компетенций проводится по 4-х балльной шкале оценивания.

Типовые шкала и критерии оценки результатов обучения при сдаче зачета для компонентов *знать, уметь и владеть* приведены в общей части ФОС образовательной программы.

4. Критерии оценивания уровня сформированности компонентов и компетенций

4.1. Оценка уровня сформированности компонентов компетенций

При оценке уровня сформированности компетенций в рамках выборочного контроля при экзамене считается, что *полученная оценка за компонент проверяемой в билете компетенции обобщается на соответствующий компонент всех компетенций, формируемых в рамках данной учебной дисциплины.*

Типовые критерии и шкалы оценивания уровня сформированности компонентов компетенций приведены в общей части ФОС образовательной программы.

4.2. Оценка уровня сформированности компетенций

Общая оценка уровня сформированности всех компетенций проводится путем агрегирования оценок, полученных студентом за каждый компонент формируемых компетенций, с учетом результатов текущего и рубежного контроля в виде интегральной оценки по 4-х балльной шкале. Все результаты контроля заносятся в оценочный лист и заполняются преподавателем по итогам промежуточной аттестации.

Форма оценочного листа и требования к его заполнению приведены в общей части ФОС образовательной программы.

При формировании итоговой оценки промежуточной аттестации в виде дифференцированного зачета используются типовые критерии, приведенные в общей части ФОС образовательной программы.