

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА**  
 Федеральное государственное бюджетное образовательное учреждение  
 высшего образования  
 "Уральский государственный университет путей сообщения"  
 (ФГБОУ ВО УрГУПС)

## ФТД.03 Анализ больших данных

### рабочая программа дисциплины (модуля)

Закреплена за кафедрой	<b>Мехатроника</b>		
Учебный план	38.03.01 ЭК-2023.plx 38.03.01 Экономика		
Направленность (профиль)	Экономика труда		
<b>Квалификация</b>	<b>бакалавр</b>		
Форма обучения	<b>очная</b>		
Объем дисциплины (модуля)	<b>2 ЗЕТ</b>		
Часов по учебному плану	72	Часов контактной работы всего, в том числе:	37,8
в том числе:		аудиторная работа	36
аудиторные занятия	36	текущие консультации по лабораторным занятиям	1,8
самостоятельная работа	36		
Промежуточная аттестация и формы контроля:			
зачет	7		

#### Распределение часов дисциплины по семестрам

Семестр (<Курс>. <Семестр на курсе>)	7 (4.1)		Итого	
	18			
Неделя	уп	рп	уп	рп
Вид занятий	уп	рп	уп	рп
Лекции	18	18	18	18
Лабораторные	18	18	18	18
Итого ауд.	36	36	36	36
Контактная работа	36	36	36	36
Сам. работа	36	36	36	36
Итого	72	72	72	72

## 1. ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1	Цели дисциплины: формирование у обучающихся целостного представления о современных проблемах анализа и обработки больших данных, получение опыта разработки и анализа концептуальных и теоретических моделей прикладных задач анализа больших данных.
1.2	Задачи дисциплины: овладение математическими и интеллектуальными методами обработки больших данных, освоение основных принципов исследования информационных моделей при помощи методов больших данных.

## 2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОП

Цикл (раздел) ОП:	ФТД
<b>2.1 Требования к предварительной подготовке обучающегося:</b>	
В результате изучения предыдущих дисциплин у обучающихся должны быть сформированы: Знания: основ высшей математики; способов применения математического аппарата в профессиональной деятельности; математических методов и моделей для описания, анализа и решения практических задач; основных понятий и методов сбора, обработки и анализа информации; основных понятий и инструментов теории статистики и социально-экономической статистики; статистической методологии сбора, обработки и анализа результатов деятельности предприятий и организаций; современных принципов работы и систематизации статистической информации; иметь представление о статистических информационных системах, базах данных и программном обеспечении статистических процессов Умения: анализировать практические ситуации, выделять базовые составляющие задачи, подбирать варианты решения и разрабатывать алгоритмы решения практической задачи; использовать возможности вычислительной техники и программного обеспечения; применять системы управления базами данных для решения профессиональных задач; выполнять поиск необходимой информации; обрабатывать эмпирические и экспериментальные цифровые данные, систематизировать их; интерпретировать результаты статистических расчетов и показателей; использовать статистические методы для выявления основных тенденций изменения социально-экономических явлений Владения: навыками применения математического аппарата (математических методов и моделей) при описании, анализе и решении практических задач; методами анализа и интерпретации результатов финансово-хозяйственной деятельности; инструментарием обработки и анализа статистической информации; использования статистической методологии в аналитических целях; программным обеспечением для работы со статистическими данными и деловой информацией	
<b>2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>	
Государственная итоговая аттестация	

## 3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

ПК-3.3: Способен вырабатывать управленческие решения по эффективному управлению трудовыми ресурсами и затратами на персонал
ПК-3.3.5: Имеет навыки разработки и описания методологии больших данных
ПК-3.3.4: Знает методологию и принципы больших данных, базовые алгоритмы обработки больших данных

В результате освоения дисциплины обучающийся должен

<b>3.1 Знать:</b>	
3.1.1	методы решения задач обработки и анализа больших данных, возможности высокопроизводительных вычислительных систем, технологии распределенных вычислений, методы и модели Data Mining; основные интеллектуальные модели: искусственные нейронные сети, метод решающих деревьев, метод ближайших соседей; системы стандартизации в области больших данных; требования к информационной безопасности в области больших данных.
<b>3.2 Уметь:</b>	
3.2.1	анализировать концептуальные и теоретические модели прикладных задач анализа больших данных; настраивать параметры интеллектуальных моделей и систем для оптимизации их характеристик; использовать алгоритмы анализа и обработки большого объема данных с применением моделей Data Mining
<b>3.3 Владеть:</b>	
3.3.1	навыками применения программных систем, предназначенных для анализа больших данных; терминологией в области больших данных; навыками разработки и описания методологии больших данных; навыками стандартизации процессов в области больших данных

## 4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов (академических)	Компетенции	Литература	Активные формы
	Раздел 1. Технологии анализа больших данных					

1.1	Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных. /Лек/	7	2	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	
1.2	Начало работы. Понятие сценария и узла обработки. Консолидация данных. Трансформация данных. Визуализация данных. Стандартизация процессов в области больших данных при проектировании ИС, в т. ч. требования безопасности. /Лаб/	7	2	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	Работа в группе со специализированным программным обеспечением
1.3	Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных. /Лек/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	
1.4	Ассоциативные правила. Поиск ассоциативных правил. /Лаб/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Л3.2 Э1	Работа в группе со специализированным программным обеспечением
1.5	Самостоятельное изучение материалов раздела с помощью литературы и ресурсов сети Интернет. Оформление отчетов по практическим занятиям. /Ср/	7	20	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Л3.2 Э1 Э2 Э3 Э4	
<b>Раздел 2. Методы и технологии Data Mining</b>						
2.1	Технологии KDD и Data Mining. Подготовка данных к анализу. Методика извлечения знаний. Data Mining. Мультидисциплинарный характер Data Mining. Причины распространения KDD и Data Mining. Актуальность технологий Data Mining как средств обработки больших объемов информации. /Лек/	7	2	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	
2.2	Кластеризация. Алгоритм кластеризации k-means. /Лаб/	7	2	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	Работа в группе со специализированным программным обеспечением
2.3	Программное обеспечение в области анализа данных. Аналитические платформы: классификация и особенности применения. Языки визуального моделирования. /Лек/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	
2.4	Прогнозирование с помощью линейной регрессии. /Лаб/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	Работа в группе со специализированным программным обеспечением
2.5	Системы стандартизации в области больших данных. Требования к информационной безопасности в области больших данных. /Лек/	7	6	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	

2.6	Классификация больших данных с помощью нейросети. /Лаб/	7	2	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Э1	Работа в группе со специализированным программным обеспечением
2.7	Классификация больших данных с помощью деревьев решений. /Лаб/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Л3.2 Э1	Работа в группе со специализированным программным обеспечением
2.8	Самостоятельное изучение материалов раздела с помощью литературы и ресурсов сети Интернет. Оформление отчетов по практическим занятиям. /Ср/	7	12	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Л3.2 Э1 Э2 Э3 Э4	
2.9	Подготовка к промежуточной аттестации /Ср/	7	4	ПК-3.3.4 ПК-3.3.5	Л1.1 Л1.2 Л1.3 Л1.4Л2.1 Л2.2Л3.1 Л3.2 Э1 Э2 Э3 Э4	

### 5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

Фонд оценочных материалов по дисциплине (модулю), состоящий из ФОМ для текущего контроля и проведения промежуточной аттестации обучающихся, разрабатывается по каждой дисциплине. Оценочные материалы размещаются на странице данного курса в системе электронной поддержки обучения Blackboard Learn (сайт bb.usurt.ru), доступной через личный кабинет обучающегося.

Фонд оценочных материалов для проведения промежуточной аттестации обучающихся по дисциплине (модулю), включая порядок проведения промежуточной аттестации, систему оценивания результатов промежуточной аттестации и критерии выставления оценок, примеры типовых заданий или иных материалов, необходимых для оценки знаний, умений, навыков, используемых для промежуточной аттестации по дисциплине, приведен в приложении 1 к рабочей программе дисциплины.

### 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

#### 6.1 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

##### 6.1.1. Основная учебная литература

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л1.1	Форман Д., Соколова А.	Много цифр. Анализ больших данных при помощи Excel: Учебное пособие	Москва: ООО "Альпина Паблишер", 2016	<a href="http://znanium.com">http://znanium.com</a>
Л1.2	Парамонов И. Ю., Смагин В. А., Косых Н. Е., Хомоненко А. Д.	Методы и модели исследования сложных систем и обработки больших данных: монография	Санкт-Петербург: Лань, 2020	<a href="http://e.lanbook.com">http://e.lanbook.com</a>
Л1.3	Волк В. К.	Базы данных. Проектирование, программирование, управление и администрирование: учебник	Санкт-Петербург: Лань, 2022	<a href="http://e.lanbook.com">http://e.lanbook.com</a>
Л1.4	Тарасян В. С.	Анализ больших данных: курс лекций для обучающихся по направлению подготовки 38.03.01 «Экономика» всех форм обучения	Екатеринбург: УрГУПС, 2022	<a href="http://biblioserver.usurt.ru">http://biblioserver.usurt.ru</a>

##### 6.1.2. Дополнительная учебная литература

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л2.1	Лемешко Б. Ю., Постовалов С. Н., Лемешко С. Б., Чимитова Е. В.	Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход	Москва: ООО "Научно-издательский центр ИНФРА-М", 2015	<a href="http://znanium.com">http://znanium.com</a>

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л2.2	Полищук Ю. В., Боровский А. С.	Базы данных и их безопасность: Учебное пособие	Москва: ООО "Научно-издательский центр ИНФРА-М", 2020	<a href="http://znanium.com">http://znanium.com</a>

### 6.1.3. Методические разработки

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л3.1	Тарасян В. С.	Анализ больших данных: методические указания по организации самостоятельной работы обучающихся по направлению подготовки 38.03.01 «Экономика» всех форм обучения	Екатеринбург: УрГУПС, 2022	<a href="http://biblioserver.usurt.ru">http://biblioserver.usurt.ru</a>
Л3.2	Тарасян В. С.	Анализ больших данных: методические рекомендации по выполнению лабораторных работ для обучающихся по направлению подготовки 38.03.01 «Экономика» всех форм обучения	Екатеринбург: УрГУПС, 2022	<a href="http://biblioserver.usurt.ru">http://biblioserver.usurt.ru</a>

### 6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Э1	Система электронной поддержки обучения Blackboard (bb.usurt.ru)
Э2	Онлайн-курсы и материалы для разработки ( <a href="http://www.coursera.org/browse/computer-science">www.coursera.org/browse/computer-science</a> )
Э3	Учебные материалы для самостоятельного изучения программирования ( <a href="http://www.codecademy.com">www.codecademy.com</a> )
Э4	Курсы и ресурсы для программирования и разработки ( <a href="http://www.edx.org/course/subject/computer-science">www.edx.org/course/subject/computer-science</a> )

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

#### 6.3.1 Перечень программного обеспечения

6.3.1.1	Неисключительные права на ПО Windows
6.3.1.2	Неисключительные права на ПО Office
6.3.1.3	Платформа управления базами данных: SQL Server
6.3.1.4	Система электронной поддержки обучения Blackboard Learn
6.3.1.5	Matlab
6.3.1.6	Statistics and Machine Learning

#### 6.3.2 Перечень информационных справочных систем и профессиональных баз данных

6.3.2.1	ЭБС издательства ЛАНЬ по адресу <a href="http://www.e.lanbook.com/">www.e.lanbook.com/</a>
6.3.2.2	ЭБС издательства Инфра-М (znanium.com) по адресу <a href="http://www.znanium.com/">www.znanium.com/</a>
6.3.2.3	ЭБ УМЦ ЖДТ по адресу <a href="https://umczt.ru/books/">https://umczt.ru/books/</a>
6.3.2.4	Электронная библиотека диссертаций Российской государственной библиотеки по адресу <a href="https://dvs.rsl.ru/">https://dvs.rsl.ru/</a>
6.3.2.5	ЭБС eLibrary, содержит электронные версии российских научно-технических журналов ( <a href="http://www.elibrary.ru">www.elibrary.ru</a> )
6.3.2.6	Консультант-Плюс - кроссплатформенная справочная правовая система.
6.3.2.7	АСПИ ЖТ - автоматизированная система правовой информации на железнодорожном транспорте.

## 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Назначение	Оснащение
Учебная аудитория для проведения занятий лекционного типа	Специализированная мебель Демонстрационное оборудование - Комплект мультимедийного оборудования Учебно-наглядные пособия - презентационные материалы
Компьютерный класс - Учебная аудитория для проведения текущего контроля и промежуточной аттестации	Специализированная мебель Компьютерная техника с установленным лицензионным ПО, включая ПО АСТ-Тест, с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета
Компьютерный класс - Учебная аудитория для проведения практических (занятий семинарского типа) и лабораторных занятий, групповых и индивидуальных	Специализированная мебель Компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета Технические средства обучения - Комплект мультимедийного оборудования

консультаций	
Читальный зал Информационно-библиотечного центра ИБК УрГУПС - Аудитория для самостоятельной работы	Специализированная мебель Компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета
Компьютерный класс - Учебная аудитория для самостоятельной работы студентов	Специализированная мебель Компьютерная техника с установленным лицензионным ПО, предусмотренным пунктом 6.3.1 РПД, с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета
Центр тестирования - Учебная аудитория для проведения текущего контроля и промежуточной аттестации	Специализированная мебель Моноблоки с установленным лицензионным ПО, включая ПО АСТ-Тест, с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета
Учебная аудитория для проведения текущего контроля и промежуточной аттестации	Специализированная мебель

#### **8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ) И ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

Эффективное освоение дисциплины предполагает регулярное посещение всех видов аудиторных занятий, выполнение плана самостоятельной работы в полном объеме и прохождение аттестации в соответствии с календарным учебным графиком.

Обучающемуся рекомендуется ознакомиться со списком основной и дополнительной литературы и взять в библиотеке издания (необходимо иметь при себе персонифицированную электронную карту и уметь пользоваться электронным каталогом «ИРБИС»).

Доступ к информационным ресурсам библиотеки и информационно-справочным системам сети «Интернет» организован в читальных залах библиотеки, в компьютерных классах, в помещениях для самостоятельной работы обучающихся со стационарных ПЭВМ, либо с личного ПЭВМ (ноутбука, планшетного компьютера или иного мобильного устройства) посредством беспроводного доступа при активации индивидуальной учетной записи.

Пользование информационными ресурсами расширяет возможности освоения теоретического курса, выполнения самостоятельной работы и позволяет получить информацию для реализации творческих образовательных технологий.

Комплект учебно-методических материалов по всем видам учебной деятельности, предусмотренным рабочей программой дисциплины (модуля), размещен на странице данного курса в системе электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru)), доступной через личный кабинет обучающегося. Перечень учебно-методических материалов (учебно-методического обеспечения) для самостоятельной работы обучающихся по дисциплине указан по темам дисциплины в разделе 4 РПД "Структура и содержание дисциплины (модуля)", материалы размещены на странице данного курса в системе электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru)).

Методические материалы, разработанные для обеспечения образовательного процесса представлены в электронном каталоге УрГУПС.

Для закрепления теоретического материала в системе электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru)) размещены тестовые материалы. Число тренировочных попыток ограничено. Тестовые материалы сформированы в логической последовательности в соответствии с изученными темами. Самостоятельная работа, связанная с выполнением отчетов по лабораторным занятиям, организована таким образом, чтобы обучающиеся имели возможность получать обратную связь о результатах их выполнения по мере готовности до начала промежуточной аттестации. Для этого отчеты по лабораторным занятиям направляются в адрес преподавателя, который проверяет их и возвращает обучающемуся с комментариями. Совместная деятельность преподавателя и обучающихся по проверке выполнения мероприятий текущего контроля, предусмотренных рабочей программой дисциплины (модуля) организована в системе электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru)). Для корректной работы в системе обучающиеся в разделе "Личные сведения" должны ввести актуальный адрес своей электронной почты.

Требования к объему и содержанию отчетов по лабораторным занятиям, а также качеству их выполнения идентичны для обучающихся всех форм обучения.

Формы самостоятельной работы обучающихся по данной дисциплине разнообразны. Они включают в себя:

- изучение теоретического материала (учебной, научной, методической литературы, материалов периодических изданий);
- подготовку к занятиям, предусмотренным РПД, мероприятиям текущего контроля, промежуточной аттестации и т.д.

Выполнять самостоятельную работу и отчитываться по ее результатам обучающийся должен в соответствии с календарным планом изучения дисциплины, видами и сроками отчетности.

При выполнении самостоятельной работы обучающемуся рекомендуется руководствоваться учебно-методическими материалами, размещенными на странице данного курса в системе электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru)), а также учебно-методическими материалами, которые указаны для самостоятельной работы по темам дисциплины в разделе 4 РПД "Структура и содержание дисциплины (модуля)".

При применении дистанционных образовательных технологий и электронного обучения освоение дисциплины (модуля) осуществляется в электронно-информационной образовательной среде (образовательная платформа электронной поддержки обучения Blackboard Learn (сайт [bb.usurt.ru](http://bb.usurt.ru))) в рамках созданного курса, что позволяет реализовывать

