

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

ФТД.02 Большие данные (Big Data)

рабочая программа дисциплины (модуля)

Закреплена за кафедрой	Информационные технологии и защита информации		
Учебный план	39.03.01 СЛ-2022.plx Направление 39.03.01 Социология		
Направленность (профиль)	Интернет-маркетинг и социальная аналитика		
Квалификация	бакалавр		
Форма обучения	очная		
Объем дисциплины (модуля)	1 ЗЕТ		
Часов по учебному плану	36	Часов контактной работы всего, в том числе:	19
в том числе:		аудиторная работа	18
аудиторные занятия	18	текущие консультации по практическим занятиям	1
самостоятельная работа	18		
Промежуточная аттестация и формы контроля:			
зачет	5		

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	5 (3.1)		Итого	
	18			
Неделя	уп	рп	уп	рп
Вид занятий	уп	рп	уп	рп
Лекции	8	8	8	8
Практические	10	10	10	10
Итого ауд.	18	18	18	18
Контактная работа	18	18	18	18
Сам. работа	18	18	18	18
Итого	36	36	36	36

1. ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1	Цель освоения дисциплины: Изучение теоретических оснований и формирование практических навыков работы с большими данными.
1.2	Задачи дисциплины: изучение теоретических положений и математических моделей анализа больших данных, специальных методов машинного обучения; формирование умений применять математические методы и строить модели анализа больших данных; формирование практических навыков владения инструментарием обработки больших данных.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОП

Цикл (раздел) ОП:	ФТД
2.1 Требования к предварительной подготовке обучающегося:	
Для изучения данной дисциплины требуются знания, умения и навыки, формируемые предшествующими дисциплинами: Методы прикладной статистики для социологов; Математика В результате изучения предшествующих дисциплин обучающийся должен: знать основные элементарные математические факты теории вероятностей и математической статистики; способы использования основных естественнонаучных законов, применения математического аппарата в профессиональной деятельности; математические методы и модели для описания, анализа и решения практических задач; основные положения, методы и технологию математического и статистического анализа, особенностях их применения при решении профессиональных задач; методы обработки, анализа и представления статистических данных. Уметь: использовать основные законы высшей математики при решении практических задач; применять методы математического и статистического анализа при работе с результатами социологических исследований; рассчитывать основные статистические показатели, абсолютные, относительные и средние величины. Владеть: навыками применения математического аппарата (математических методов и моделей) при описании, анализе и решении практических задач; навыком практического использования знаний и методов математики, экономической и социальной статистики; навыком обработки и анализа статистических данных вторичных и самостоятельно проведённых исследований.	
2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
Анализ и визуализация социологических данных Социальное прогнозирование и проектирование Производственная практика (научно-исследовательская работа) Производственная практика (преддипломная практика) Производственная практика (проектно-технологическая практика) Государственная итоговая аттестация	

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

ПК-2.2: Способен обрабатывать, анализировать и интерпретировать данные для подготовки аналитических решений, экспертных заключений и рекомендаций
ПК-2.2.5: Имеет навыки разработки и описания методологии больших данных
ПК-2.2.4: Знает методологию и принципы больших данных, классификацию видов данных и их характеристики, методологию обследования процессов больших данных, базовые алгоритмы обработки больших данных

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	математические объекты и методы, необходимые для работы с большими данными; основные алгоритмы работы с большими данными; основные теоретические и практические задачи в работе с большими данными; современные средства для работы с большими данными
3.2	Уметь:
3.2.1	оценивать качество данных, проводить их анализ, извлекать содержательную информацию; самостоятельно проектировать и реализовывать алгоритмы работы с большими данными; использовать основные информационные системы и среды разработки для самостоятельного решения задач анализа большим данным; строить и оптимизировать модели анализа больших данных
3.3	Владеть:
3.3.1	специальными прикладными программами для создания алгоритмов анализа больших данных; информационными системами и средами, актуальными для решения задач обработки данных

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов (академически)	Компетенции	Литература	Активные формы
	Раздел 1. Общая информация					

1.1	Введение в анализ больших данных /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
1.2	Математические методы для анализа больших данных /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
1.3	Первичное ознакомление с инструментами работы с данными /Пр/	5	2	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
1.4	Решение тестов в BlackBoard /Ср/	5	4	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Л3.2 Э1 Э2 Э3	
	Раздел 2. Основные модели для анализа больших данных					
2.1	Преобразование данных /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
2.2	Практическое преобразование данных /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
2.3	Метод К ближайших соседей (KNN) /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
2.4	Реализация метода KNN, подбор оптимальных гиперпараметров /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
2.5	Линейная и логистическая регрессии /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
2.6	Реализация регрессии, подбор оптимальных гиперпараметров /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
2.7	Деревья принятия решений /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	

2.8	Реализация ДПР, подбор оптимальных гиперпараметров /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
2.9	Решение тестов в BlackBoard /Ср/	5	4	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Л3.2 Э1 Э2 Э3	
Раздел 3. Специальные задачи машинного обучения						
3.1	Обработка текстов. Распознавание естественного языка(NLP) /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
3.2	Практическое применение библиотек nltk, fasttext, word2vec /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
3.3	Анализ тональности постов социальных сетей /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
3.4	Временные ряды. /Лек/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1 Э1 Э2 Э3	
3.5	Практическая работа с временными рядами /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
3.6	Анализ поведенческой динамики в социальных сетях /Пр/	5	1	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Э1 Э2 Э3	Работа в малых группах по решению практико-ориентированных задач с использованием прикладного программного обеспечения
3.7	Решение тестов в BlackBoard /Ср/	5	4	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Л3.2 Э1 Э2 Э3	
3.8	Подготовка к промежуточной аттестации /Ср/	5	6	ПК-2.2.4 ПК-2.2.5	Л1.1 Л1.2Л2.1Л3.1 Л3.2 Э1 Э2 Э3	

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	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	http://e.lanbook.com
Л1.2	Чернышов Ю. Ю.	Большие данные (Big Data): курс лекций для обучающихся по направлению подготовки 39.03.01 «Социология» всех форм обучения	Екатеринбург: УрГУПС, 2021	http://biblioserver.usurt.ru

6.1.2. Дополнительная учебная литература

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л2.1	Форман Д., Соколова А.	Много цифр. Анализ больших данных при помощи Excel: Учебное пособие	Москва: ООО "Альпина Паблишер", 2016	http://znanium.com

6.1.3. Методические разработки

	Авторы, составители	Заглавие	Издательство, год	Web-ссылка
Л3.1	Чернышов Ю. Ю.	Большие данные (Big Data): методические указания к практическим занятиям для обучающихся по направлению подготовки 39.03.01 «Социология» всех форм обучения	Екатеринбург: УрГУПС, 2021	http://biblioserver.usurt.ru
Л3.2	Чернышов Ю. Ю.	Большие данные (Big Data): методические рекомендации по организации самостоятельной работы обучающихся по направлению подготовки 39.03.01 «Социология» всех форм обучения	Екатеринбург: УрГУПС, 2021	http://biblioserver.usurt.ru

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

Э1	bb.usurt.ru . Система электронной поддержки обучения BlackBoard Learn.
Э2	Открытый курс "Машинное обучение" mlcourse.ai
Э3	Открытая библиотека статей habr.com

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	Система электронной поддержки обучения Blackboard Learn
6.3.1.4	Statistics and Machine Learning
6.3.1.5	IDLE Python

6.3.2 Перечень информационных справочных систем и профессиональных баз данных

6.3.2.1	Автоматизированная система правовой информации на железнодорожном транспорте АСПИ ЖТ (профессиональная БД)
6.3.2.2	Интерактивный справочник по математике, физике, химии (ИСС открытого доступа, https://www.fxyz.ru)
6.3.2.3	MathTree - каталог математических интернет-ресурсов (ИСС открытого доступа, http://www.mathtree.ru)
6.3.2.4	Образовательный математический сайт Exponenta.ru (БД и ИСС открытого доступа по решению математических и прикладных задач в среде математических пакетов Mathcad, Matlab, Maple, Mathematica, Statistica, http://www.old.exponenta.ru)

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
Учебная аудитория для проведения занятий лекционного типа, семинарского типа (практических занятий, лабораторных занятий), курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также для самостоятельной работы. Кабинет «Информатика, технологии и методы программирования».	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 Технические средства обучения - Комплект мультимедийного оборудования
Компьютерный класс - Учебная аудитория для проведения практических (занятий семинарского типа) и лабораторных занятий, групповых и индивидуальных консультаций	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
Учебная аудитория для проведения текущего контроля и промежуточной аттестации	Специализированная мебель
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
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
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

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), 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).

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) размещены тестовые материалы. Число тренировочных попыток ограничено. Тестовые материалы сформированы в логической последовательности в соответствии с изученными темами. 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). Для корректной работы в системе обучающиеся в разделе "Личные сведения" должны ввести актуальный адрес своей электронной почты. 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), 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)) в рамках созданного курса, что позволяет реализовывать асинхронное и синхронное взаимодействие участников образовательных отношений.