

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ
Директор по образовательной
деятельности

_____ С.Т. Князев
«__» _____

РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1150451	Большие данные в экономике и управлении

Екатеринбург

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа 1. Бизнес-информатика	Код ОП 1. 38.03.05/33.01
Направление подготовки 1. Бизнес-информатика	Код направления и уровня подготовки 1. 38.03.05

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Тарасьев Александр Александрович	кандидат экономических наук, без ученого звания	Старший преподаватель	анализа систем и принятия решений

Согласовано:

Управление образовательных программ

Р.Х. Токарева

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Большие данные в экономике и управлении

1.1. Аннотация содержания модуля

Целью модуля Большие данные в экономике и управлении является изучение алгоритмов интеллектуального анализа больших объемов данных (Big Data), при решении экономических задач и принятии управленческих решений. Проект по модулю "Большие данные в экономике и управлении" направлен на применение теоретических знаний в сфере анализа больших данных и методов работы с ними. В процессе выполнения проекта, изученные в рамках дисциплин модуля методы и алгоритмы применяются для решения, как правило, реальных бизнес-задач, связанных с использованием и необходимостью анализа больших объемов данных. По результатам наиболее интересных проектов делаются доклады на студенческих научных конференциях. Целью дисциплины Анализ больших данных является изучение современных алгоритмов и технологий интеллектуального анализа больших объемов данных (Big Data). Рассматриваются парадигма "MapReduce", методы поиска похожих объектов, методы анализа ссылок в социальных сетях и ряд специальных алгоритмов, нацеленных на анализ именно больших объемов данных, такие, например, как метод понижения размерности. На практических занятиях анализируются конкретные бизнес-кейсы, по использованию методов анализа больших данных с помощью имеющихся библиотек, программных продуктов, языка Python. Целью дисциплины Технологии машинного обучения и основы построения рекомендательных систем является изучение методов используемых в технологиях машинного обучения, характерной особенностью которых является не просто проведение анализа данных, а построение модели, позволяющей получить информацию о данных, которые будут предъявлены системе в будущем. Рассматриваются существующие подходы к машинному обучению и его базовые алгоритмы, такие как обучение перцептронов, метод опорных векторов, методы ближайших соседей и регрессия. Рассмотрены основные модели построения рекомендательных систем для решения экономических и управленческих задач.

1.2. Структура и объем модуля

Таблица 1

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Анализ больших данных	3
2	Технологии машинного обучения и основы построения рекомендательных систем	3
3	Проект по модулю "Большие данные в экономике и управлении"	3
ИТОГО по модулю:		9

1.3. Последовательность освоения модуля в образовательной программе

Препреквизиты модуля	<ol style="list-style-type: none">1. Математические методы анализа2. Теория экономических отношений3. Информационные системы и технологии
-----------------------------	---

	<ul style="list-style-type: none"> 4. Архитектура и бизнес-процессы предприятия 5. Архитектуры резервного копирования и восстановления данных 6. Учет и анализ в управлении бизнесом
Постреквизиты и кореквизиты модуля	<ul style="list-style-type: none"> 1. Интеллектуальный анализ данных 2. Программирование в среде 1С 3. Инновационная деятельность в сфере ИТ 4. Информационные технологии финансового менеджмента 5. Информационные технологии на финансовых рынках

1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения (индикаторы)
1	2	3
Анализ больших данных	ПК-1 - Способен выявлять и анализировать целевые сегменты рынка	<p>З-1 - Основные методы анализа конъюнктуры рынков (анализ спроса, предложения, динамики цен, объемов торгов)</p> <p>У-1 - Самостоятельно применять методики анализа конъюнктуры рынков</p> <p>У-2 - Проводить поиск необходимой информации для анализа конъюнктуры рынков в сети интернет</p> <p>У-3 - Использовать язык Python для анализа целевых сегментов рынка</p> <p>П-1 - Устойчивые навыки самостоятельной работы с источниками информации и литературой для поиска информации о целевых сегментах рынка и ее применения для поставленных целей анализа</p> <p>П-2 - Опыт применения методов поиска, сбора и обработки, критического анализа и</p>

		<p>синтеза информации о целевых сегментах рынка</p> <p>Д-1 - Демонстрировать внимательность, усердие и целеустремленность в поиске и обработке информации</p> <p>Д-2 - Демонстрировать развитый интеллект и критическое мышление</p>
	<p>ПК-4 - Способен применять технологии интеллектуального анализа данных и статистического анализа применять решения практических задач экономики и управления</p>	<p>З-1 - Основные методы интеллектуального анализа данных</p> <p>У-1 - Проводить статистический и интеллектуальный анализ данных в процессе решения поставленных задач экономики и управления</p> <p>П-1 - Опыт проведения регрессионного и факторного анализа данных</p> <p>П-2 - Опыт проверки гипотез при заданном уровне достоверности</p> <p>П-3 - Опыт решения проведения интеллектуального анализа данных с помощью языка Python (R)</p> <p>Д-1 - Демонстрировать целеустремленность, внимательность и ответственность при обработке информации</p> <p>Д-2 - Демонстрировать творческий подход и способность критического анализа результатов деятельности</p>
	<p>ПК-6 - Способен проводить экономический и финансовый анализ деятельности отдельных экономических субъектов</p>	<p>З-1 - Основные методы экономического и финансового анализа</p> <p>З-2 - Основные методы статистического анализа данных</p> <p>З-3 - Методы интеллектуального анализа данных (в том числе больших данных)</p> <p>У-1 - Умеет проводить экономический и финансовый анализ экономических субъектов</p> <p>У-2 - Умеет применять методы интеллектуального анализа данных для проведения экономического и финансового анализа экономических субъектов</p>

		<p>П-1 - Имеет опыт проведения экономического и финансового анализа экономических субъектов</p>
	<p>ПК-7 - Способен проводить оценку экономической эффективности применения информационных сервисов и систем на уровне отдельной организации</p>	<p>З-1 - Знает методы анализа экономической и финансовой эффективности применения информационных сервисов и систем</p> <p>У-1 - Умеет оценивать финансовую эффективность отдельных информационных сервисов и систем</p> <p>П-1 - Имеет устойчивые навыки проведения оценок финансовой эффективности информационных сервисов и систем</p>
	<p>ПК-22 - Способен проводить исследования в области экономики, управления и финансов с применением информационных технологий, и методов интеллектуального анализа данных</p>	<p>З-1 - Методы проведения исследований в области экономики, управления и финансов</p> <p>З-2 - Информационные технологии, применяемые в области экономики, управления и финансов</p> <p>З-3 - Методы интеллектуального анализа данных</p> <p>У-1 - Применять методы исследования в области экономики, управления и финансов</p> <p>У-2 - Применять информационные технологии для проведения исследований в области экономики, управления и финансов</p> <p>У-3 - Применять методы интеллектуального анализа данных</p> <p>П-1 - Навыками применения информационных технологий в области экономики, управления и финансов</p> <p>П-2 - Навыками интеллектуального анализа данных</p> <p>Д-1 - Личные качества: развитый интеллект, системное мышление, креативность</p>
<p>Проект по модулю "Большие данные в экономике и управлении"</p>	<p>ПК-1 - Способен выявлять и анализировать целевые сегменты рынка</p>	<p>З-1 - Основные методы анализа конъюнктуры рынков (анализ спроса, предложения, динамики цен, объемов торгов)</p> <p>У-1 - Самостоятельно применять методики анализа конъюнктуры рынков</p>

		<p>У-2 - Проводить поиск необходимой информации для анализа конъюнктуры рынков в сети интернет</p> <p>У-3 - Использовать язык Python для анализа целевых сегментов рынка</p> <p>П-1 - Устойчивые навыки самостоятельной работы с источниками информации и литературой для поиска информации о целевых сегментах рынка и ее применения для поставленных целей анализа</p> <p>П-2 - Опыт применения методов поиска, сбора и обработки, критического анализа и синтеза информации о целевых сегментах рынка</p> <p>Д-1 - Демонстрировать внимательность, усердие и целеустремленность в поиске и обработке информации</p> <p>Д-2 - Демонстрировать развитый интеллект и критическое мышление</p>
	<p>ПК-4 - Способен применять технологии интеллектуального анализа данных и статистического анализа применять решения практических задач экономики и управления</p>	<p>З-1 - Основные методы интеллектуального анализа данных</p> <p>У-1 - Проводить статистический и интеллектуальный анализ данных в процессе решения поставленных задач экономики и управления</p> <p>П-1 - Опыт проведения регрессионного и факторного анализа данных</p> <p>П-2 - Опыт проверки гипотез при заданном уровне достоверности</p> <p>П-3 - Опыт решения проведения интеллектуального анализа данных с помощью языка Python (R)</p> <p>Д-1 - Демонстрировать целеустремленность, внимательность и ответственность при обработке информации</p> <p>Д-2 - Демонстрировать творческий подход и способность критического анализа результатов деятельности</p>
	<p>ПК-7 - Способен проводить оценку экономической эффективности применения</p>	<p>З-1 - Знает методы анализа экономической и финансовой эффективности применения информационных сервисов и систем</p>

<p>информационных сервисов и систем на уровне отдельной организации</p>	<p>У-1 - Умеет оценивать финансовую эффективность отдельных информационных сервисов и систем</p> <p>П-1 - Имеет устойчивые навыки проведения оценок финансовой эффективности информационных сервисов и систем</p>
<p>ПК-13 - Способен организовать управление требованиями к программному обеспечению (ПО), продукту, средству, программно-аппаратному комплексу, автоматизированной информационной системе или автоматизированной системе управления</p>	<p>З-1 - Знает технологии управления требованиями к продуктам информационной сферы</p> <p>У-1 - Умеет организовать управление требованиями к программному обеспечению (ПО), продукту, средству, программно-аппаратному комплексу, автоматизированной информационной системе или автоматизированной системе управления</p> <p>П-1 - Владеет организаторскими способностями</p> <p>П-2 - Опыт организации управления требованиями к продуктам информационной сферы</p> <p>Д-1 - Личные качества: системное мышление, коммуникабельность, настойчивость в достижении цели</p>
<p>ПК-19 - Способен консультировать и проектировать в консалтинговых проектах по разработке и внедрению информационных систем и сервисов</p>	<p>З-1 - Технологии консультирования</p> <p>З-2 - Методы проектирования консалтинговых проектов</p> <p>З-3 - Методы внедрения информационных сервисов и систем</p> <p>У-1 - Проектировать консалтинговые проекты</p> <p>У-2 - Внедрять информационные сервисы и системы</p> <p>У-3 - Консультировать по реализации консалтинговых проектов по разработке и внедрению информационных систем и сервисов</p> <p>П-1 - Навыками проектирования консалтинговых проектов</p> <p>П-2 - Навыками внедрения информационных сервисов и систем</p>

		<p>П-3 - Навыками консультирования</p> <p>Д-1 - Личные качества: коммуникабельность, внимательность, креативность</p>
	<p>ПК-22 - Способен проводить исследования в области экономики, управления и финансов с применением информационных технологий, и методов интеллектуального анализа данных</p>	<p>З-1 - Методы проведения исследований в области экономики, управления и финансов</p> <p>З-2 - Информационные технологии, применяемые в области экономики, управления и финансов</p> <p>З-3 - Методы интеллектуального анализа данных</p> <p>У-1 - Применять методы исследования в области экономики, управления и финансов</p> <p>У-2 - Применять информационные технологии для проведения исследований в области экономики, управления и финансов</p> <p>У-3 - Применять методы интеллектуального анализа данных</p> <p>П-1 - Навыками применения информационных технологий в области экономики, управления и финансов</p> <p>П-2 - Навыками интеллектуального анализа данных</p> <p>Д-1 - Личные качества: развитый интеллект, системное мышление, креативность</p>
<p>Технологии машинного обучения и основы построения рекомендательных систем</p>	<p>ПК-1 - Способен выявлять и анализировать целевые сегменты рынка</p>	<p>З-1 - Основные методы анализа конъюнктуры рынков (анализ спроса, предложения, динамики цен, объемов торгов)</p> <p>У-3 - Использовать язык Python для анализа целевых сегментов рынка</p> <p>П-1 - Устойчивые навыки самостоятельной работы с источниками информации и литературой для поиска информации о целевых сегментах рынка и ее применения для поставленных целей анализа</p> <p>П-2 - Опыт применения методов поиска, сбора и обработки, критического анализа и синтеза информации о целевых сегментах рынка</p>

	<p>Д-1 - Демонстрировать внимательность, усердие и целеустремленность в поиске и обработке информации</p> <p>Д-2 - Демонстрировать развитый интеллект и критическое мышление</p>
<p>ПК-4 - Способен применять технологии интеллектуального анализа данных и статистического анализа применять решения практических задач экономики и управления</p>	<p>З-1 - Основные методы интеллектуального анализа данных</p> <p>У-1 - Проводить статистический и интеллектуальный анализ данных в процессе решения поставленных задач экономики и управления</p> <p>П-1 - Опыт проведения регрессионного и факторного анализа данных</p> <p>П-2 - Опыт проверки гипотез при заданном уровне достоверности</p> <p>П-3 - Опыт решения проведения интеллектуального анализа данных с помощью языка Python (R)</p> <p>Д-1 - Демонстрировать целеустремленность, внимательность и ответственность при обработке информации</p> <p>Д-2 - Демонстрировать творческий подход и способность критического анализа результатов деятельности</p>
<p>ПК-7 - Способен проводить оценку экономической эффективности применения информационных сервисов и систем на уровне отдельной организации</p>	<p>З-1 - Знает методы анализа экономической и финансовой эффективности применения информационных сервисов и систем</p> <p>У-1 - Умеет оценивать финансовую эффективность отдельных информационных сервисов и систем</p> <p>П-1 - Имеет устойчивые навыки проведения оценок финансовой эффективности информационных сервисов и систем</p>
<p>ПК-13 - Способен организовать управление требованиями к программному обеспечению (ПО), продукту, средству, программно-аппаратному комплексу, автоматизированной информационной</p>	<p>З-1 - Знает технологии управления требованиями к продуктам информационной сферы</p> <p>У-1 - Умеет организовать управление требованиями к программному обеспечению (ПО), продукту, средству, программно-аппаратному комплексу, автоматизированной информационной</p>

	<p>системе или автоматизированной системе управления</p>	<p>системе или автоматизированной системе управления</p> <p>П-1 - Владеет организаторскими способностями</p> <p>П-2 - Опыт организации управления требованиями к продуктам информационной сферы</p> <p>Д-1 - Личные качества: системное мышление, коммуникабельность, настойчивость в достижении цели</p>
	<p>ПК-19 - Способен консультировать и проектировать в консалтинговых проектах по разработке и внедрению информационных систем и сервисов</p>	<p>З-1 - Технологии консультирования</p> <p>З-2 - Методы проектирования консалтинговых проектов</p> <p>З-3 - Методы внедрения информационных сервисов и систем</p> <p>У-1 - Проектировать консалтинговые проекты</p> <p>У-2 - Внедрять информационные сервисы и системы</p> <p>У-3 - Консультировать по реализации консалтинговых проектов по разработке и внедрению информационных систем и сервисов</p> <p>П-1 - Навыками проектирования консалтинговых проектов</p> <p>П-2 - Навыками внедрения информационных сервисов и систем</p> <p>П-3 - Навыками консультирования</p> <p>Д-1 - Личные качества: коммуникабельность, внимательность, креативность</p>
	<p>ПК-22 - Способен проводить исследования в области экономики, управления и финансов с применением информационных технологий, и методов интеллектуального анализа данных</p>	<p>З-1 - Методы проведения исследований в области экономики, управления и финансов</p> <p>З-2 - Информационные технологии, применяемые в области экономики, управления и финансов</p> <p>З-3 - Методы интеллектуального анализа данных</p> <p>У-1 - Применять методы исследования в области экономики, управления и финансов</p>

		<p>У-2 - Применять информационные технологии для проведения исследований в области экономики, управления и финансов</p> <p>У-3 - Применять методы интеллектуального анализа данных</p> <p>П-1 - Навыками применения информационных технологий в области экономики, управления и финансов</p> <p>П-2 - Навыками интеллектуального анализа данных</p> <p>Д-1 - Личные качества: развитый интеллект, системное мышление, креативность</p>
--	--	--

1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной, очно-заочной и заочной формах.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Анализ больших данных

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Тарасьев Александр Александрович	кандидат экономических наук, без ученого звания	Старший преподавателе ль	анализа систем и принятия решений

Рекомендовано учебно-методическим советом института Институт экономики и управления

Протокол № 13 от 11.06.2021 г.

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Тарасьев Александр Александрович, Старший преподаватель, анализа систем и принятия решений

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Базовый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
001	Введение в анализ больших данных	Введение в анализ данных, кейсы из предметной области. Основы синтаксиса Python. Объекты Python. Типы данных. Манипуляции с типами. Бинарные операторы и операторы сравнения. Основные алгоритмические конструкции. Регулярные выражения. Добавление комментариев.
002	Поиск и обработка данных	Поиск данных. Открытые данные, используемые в государственном управлении. Базы данных. Портал открытых данных. Подбор данных для конкретных задач.
003	Основы статистического анализа данных	Генеральная совокупность, выборка. Репрезентативность выборки. Распределение. Типы переменных (количественные, порядковые, категориальные). Числовые характеристики: МЦТ (медиана, среднее арифметическое, мода) и меры разброса (стандартное отклонение, дисперсия), квантили, перцентили. Соотношение МЦТ и типов переменных. Анализ взаимосвязи переменных, корреляция. Свойства корреляции. Проверка статистической гипотезы, p-value, сравнение средних (критерий Стьюдента). Ложные корреляции. Отличие взаимосвязи от причинно-следственной связи.
004	Машинное обучение	Задачи машинного обучения: обучение с учителем и без учителя, выбор метода для конкретной задачи. Основная терминология. Регрессия. Классификация и кластеризация: основной смысл, кейсы.

005	Визуализация данных	Визуализация результатов анализа данных. График плотности распределения. Диаграммы рассеяния. Гистограммы. Столбиковые диаграммы. Линейные диаграммы. Коробчатые диаграммы. Круговые диаграммы. Необычные типы диаграмм. Удачные и неудачные варианты визуализации данных. Подбор графического представления для различных результатов исследований.
-----	---------------------	--

1.3. Направление, виды воспитательной деятельности и используемые технологии

Таблица 1.2

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения
Профессиональное воспитание	проектная деятельность профориентационная деятельность предпринимательская деятельность целенаправленная работа с информацией для использования в практических целях	Технология формирования уверенности и готовности к самостоятельной успешной профессиональной деятельности Технология проектного образования Технология самостоятельной работы	ПК-4 - Способен применять технологии интеллектуального анализа данных и статистического анализа применять решения практических задач экономики и управления	З-1 - Основные методы интеллектуального анализа данных У-1 - Проводить статистический и интеллектуальный анализ данных в процессе решения поставленных задач экономики и управления П-3 - Опыт решения проведения интеллектуального анализа данных с помощью языка Python (R)

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Анализ больших данных

Электронные ресурсы (издания)

1. Переславцева, О. Н.; Комплексные решения для создания инфраструктуры предприятия на основе суперкомпьютерных систем : учебно-методическое пособие.; Тамбовский государственный университет имени Г.Р. Державина, Тамбов; 2020; <http://www.iprbookshop.ru/109753.html> (Электронное издание)
2. Дубровский, С. А.; Методы обработки и анализа экспериментальных данных : учебное пособие.;

Липецкий государственный технический университет, ЭБС АСВ, Липецк; 2015; <http://www.iprbookshop.ru/55640.html> (Электронное издание)

3. Пальмов, С. В.; Интеллектуальный анализ данных : учебное пособие.; Поволжский государственный университет телекоммуникаций и информатики, Самара; 2017; <http://www.iprbookshop.ru/75376.html> (Электронное издание)

4. Секлетова, Н. Н.; Системный анализ и принятие решений : учебное пособие.; Поволжский государственный университет телекоммуникаций и информатики, Самара; 2017; <http://www.iprbookshop.ru/75407.html> (Электронное издание)

5. Маккинли, Слинкина, А.; Python и анализ данных; Профобразование, Саратов; 2019; <http://www.iprbookshop.ru/88752.html> (Электронное издание)

6. Чубукова, И. А.; Data Mining : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва, Саратов; 2020; <http://www.iprbookshop.ru/89404.html> (Электронное издание)

7. Бекетнова, Ю. М.; Модели и методы решения аналитических задач финансового мониторинга : монография.; Прометей, Москва; 2018; <http://www.iprbookshop.ru/94464.html> (Электронное издание)

8. Полубояров, В. В.; Использование MS SQL Server Analysis Services 2008 для построения хранилищ данных : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва; 2021; <http://www.iprbookshop.ru/102014.html> (Электронное издание)

9. Синева, И. С.; Анализ данных в среде R. Ч. 1 : учебное пособие.; Московский технический университет связи и информатики, Москва; 2018; <http://www.iprbookshop.ru/92422.html> (Электронное издание)

Печатные издания

1. Докучаев, В. А., Докучаев, В. А.; Архитектура центров обработки данных : [монография].; Горячая линия - Телеком, Москва; 2020 (1 экз.)

2. Петрунин, Ю. Ю.; Информационные технологии анализа данных. Data Analysis : учеб. пособие по дисциплине "Информатика" для студентов вузов, обучающихся по упр. и экон. специальностям и направлениям.; КДУ, Москва; 2008 (11 экз.)

3. Петрунин, Ю. Ю.; Информационные технологии анализа данных. Data Analysis : учеб. пособие по дисциплине "Информатика" для студентов вузов, обучающихся по упр. и экон. специальностям и направлениям.; КДУ, Москва; 2008 (11 экз.)

4. Чубукова, И. А.; Data Mining : учеб. пособие.; Интернет-Университет информационных технологий, Москва; 2006 (1 экз.)

Профессиональные базы данных, информационно-справочные системы

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Анализ больших данных

Сведения об оснащении дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES Adobe Acrobat Professional 2017 Multiple Platforms
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Персональные компьютеры по количеству обучающихся Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES Adobe Acrobat Professional 2017 Multiple Platforms
3	Самостоятельная работа студентов	Персональные компьютеры по количеству обучающихся Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES Adobe Acrobat Professional 2017 Multiple Platforms

4	Текущий контроль и промежуточная аттестация	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Периферийное устройство</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p>	<p>Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>Adobe Acrobat Professional 2017 Multiple Platforms</p>
5	Консультации	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p>	<p>Office 365 EDUA5 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>Adobe Acrobat Professional 2017 Multiple Platforms</p>

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Технологии машинного обучения и основы
построения рекомендательных систем

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Тарасьев Александр Александрович	кандидат экономических наук, без ученого звания	Старший преподавате ль	анализа систем и принятия решений

Рекомендовано учебно-методическим советом института Институт экономики и управления

Протокол № 13 от 11.06.2021 г.

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Тарасьев Александр Александрович, Старший преподаватель, анализа систем и принятия решений

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Базовый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
001	Введение в машинное обучение	Обучение с учителем (supervised learning): регрессия и классификация; обучение без учителя (unsupervised learning): кластеризация, снижение размерности; semi-supervised learning, рекомендательные системы, обработка текстов: тематическое моделирование, построение аннотаций, извлечение ответов на вопросы, машинный перевод; обработка изображений: порождение, преобразование; обучение представлений; обучение с подкреплением.
002	Статистическое обучение	Машинное обучение как математическое моделирование. Статистические модели. Теоретико-вероятностная постановка задачи обучения с учителем. Минимизация ожидаемой ошибки. Пример: задача регрессии, минимизация квадрата отклонения. Регрессионная функция: условное матожидание. Линейная регрессия и метод k ближайших соседей. Переобучение и недообучение. Разложение ошибки на шум, смещение и разброс. Проклятие размерности
003	Линейная регрессия	Линейная регрессия, метод наименьших квадратов и максимизация правдоподобия. Теорема Гаусса—Маркова. Явный вид решения в методе наименьших квадратов. Ковариационная матрица для коэффициентов. Практические соображения: что делать с категориальными данными? Вычислительные соображения: точное решение vs градиентный спуск.

004	Задачи классификации	Общая постановка задачи классификации. 0-1 ошибка. Байесовский классификатор. Линейные методы для классификации. Логистическая регрессия, максимизация правдоподобия, кросс-энтропия.
005	Выбор и оценка моделей	Кросс-валидация: тонкости (отбор переменных, переобучение на валидационное множество). Оценки ожидаемой ошибки для линейной регрессии: AIC и другие. L1 и L2 регуляризация
006	Решающие деревья и случайные леса	Ограничения линейных методов (пример: XOR). Решающие деревья. CART. Ансамбли. Бутстреп. Бэггинг. Случайный лес.
007	Градиентный бустинг над деревьями	Аддитивные модели. Градиентный бустинг. Популярные алгоритмы градиентного бустинга: XGBoost, lightgbm, catboost.
008	Особенности практической работы с данными	Feature engineering. Кодирование категориальных данных. Пропущенные значения. Обработка текстов: bag of words, tf-idf, векторные эмбединги
009	Нейронные сети	Нейронные сети: общая архитектура. Реализация XOR с помощью трёх персептронов. Теорема об универсальной аппроксимации. Многослойные сети. Обратное распространение ошибки. Стохастический градиентный спуск. Проблемы: затухающие и взрывающиеся градиенты, невыпуклость функции потерь.
010	Современные нейросетевые архитектуры	Нейронные сети в обработке изображений. Фильтры. Сверточные слои. Нейронные сети и обучение представлений. Обработка последовательностей. Рекуррентные нейронные сети.
011	Задачи кластеризации	Кластеризация. K-means. EM-алгоритм. Другие методы кластеризации. Иерархическая кластеризация.
012	Методы снижения размерности	Снижение размерности. SVD-разложение. Метод главных компонент. t-SNE, UMAP.
013	Матричные разложения	Матричные разложения и их применение. Рекомендательные системы. Тематическое моделирование.

1.3. Направление, виды воспитательной деятельности и используемые технологии

Таблица 1.2

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения
Профессиональное воспитание	проектная деятельность целенаправленная работа с информацией для использования в практических	Технология проектного образования Технология самостоятельной работы	ПК-22 - Способен проводить исследования в области экономики, управления и финансов с применением информационных технологий, и	3-1 - Методы проведения исследований в области экономики, управления и финансов 3-2 - Информационные

	целях		методов интеллектуального анализа данных	технологии, применяемые в области экономики, управления и финансов 3-3 - Методы интеллектуального анализа данных
--	-------	--	--	---

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Технологии машинного обучения и основы построения рекомендательных систем

Электронные ресурсы (издания)

1. Дубровский, С. А.; Методы обработки и анализа экспериментальных данных : учебное пособие.; Липецкий государственный технический университет, ЭБС АСВ, Липецк; 2015; <http://www.iprbookshop.ru/55640.html> (Электронное издание)
2. Пальмов, С. В.; Интеллектуальный анализ данных : учебное пособие.; Поволжский государственный университет телекоммуникаций и информатики, Самара; 2017; <http://www.iprbookshop.ru/75376.html> (Электронное издание)
3. Секлетова, Н. Н.; Системный анализ и принятие решений : учебное пособие.; Поволжский государственный университет телекоммуникаций и информатики, Самара; 2017; <http://www.iprbookshop.ru/75407.html> (Электронное издание)
4. Маккинли, Слинкина, А.; Python и анализ данных; Профобразование, Саратов; 2019; <http://www.iprbookshop.ru/88752.html> (Электронное издание)
5. Чубукова, И. А.; Data Mining : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва, Саратов; 2020; <http://www.iprbookshop.ru/89404.html> (Электронное издание)
6. Бекетнова, Ю. М.; Модели и методы решения аналитических задач финансового мониторинга : монография.; Прометей, Москва; 2018; <http://www.iprbookshop.ru/94464.html> (Электронное издание)
7. Полубояров, В. В.; Использование MS SQL Server Analysis Services 2008 для построения хранилищ данных : учебное пособие.; Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, Москва; 2021; <http://www.iprbookshop.ru/102014.html> (Электронное издание)
8. Синева, И. С.; Анализ данных в среде R. Ч. 1 : учебное пособие.; Московский технический университет связи и информатики, Москва; 2018; <http://www.iprbookshop.ru/92422.html> (Электронное издание)

Печатные издания

1. Лимановская, О. В., Обабков, И. Н.; Основы машинного обучения : учебное пособие для студентов вуза, обучающихся по направлениям подготовки: 09.03.04 - Программная инженерия; 09.03.03 - Прикладная информатика; 02.04.02 - Фундаментальная информатика и информационные технологии; 09.04.02 - Информационные системы и технологии.; Издательство Уральского университета, Екатеринбург; 2020 (15 экз.)

2. Глухих, И. Н.; Интеллектуальные информационные системы : учеб. пособие для студентов учреждений высш. проф. образования.; Академия, Москва; 2010 (1 экз.)
3. Петрунин, Ю. Ю.; Информационные технологии анализа данных. Data Analysis : учеб. пособие по дисциплине "Информатика" для студентов вузов, обучающихся по упр. и экон. специальностям и направлениям.; КДУ, Москва; 2008 (11 экз.)
4. Лимановская, О. В., Обабков, И. Н.; Основы машинного обучения : учебное пособие для студентов вуза, обучающихся по направлениям подготовки: 09.03.04 - Программная инженерия; 09.03.03 - Прикладная информатика; 02.04.02 - Фундаментальная информатика и информационные технологии; 09.04.02 - Информационные системы и технологии.; Издательство Уральского университета, Екатеринбург; 2020 (15 экз.)

Профессиональные базы данных, информационно-справочные системы

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Технологии машинного обучения и основы построения рекомендательных систем

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Оборудование, соответствующее требованиям организации учебного процесса в соответствии с	Adobe Acrobat Professional 2017 Multiple Platforms Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES

		санитарными правилами и нормами Подключение к сети Интернет	
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Персональные компьютеры по количеству обучающихся Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	Adobe Acrobat Professional 2017 Multiple Platforms Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
3	Консультации	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами Подключение к сети Интернет	Adobe Acrobat Professional 2017 Multiple Platforms Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES
4	Текущий контроль и промежуточная аттестация	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство	Adobe Acrobat Professional 2017 Multiple Platforms Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES

		<p>Персональные компьютеры по количеству обучающихся</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p>	
5	Самостоятельная работа студентов	<p>Персональные компьютеры по количеству обучающихся</p> <p>Оборудование, соответствующее требованиям организации учебного процесса в соответствии с санитарными правилами и нормами</p> <p>Подключение к сети Интернет</p>	<p>Adobe Acrobat Professional 2017 Multiple Platforms</p> <p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p>