

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Сибирский государственный индустриальный университет»  
Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ  
Директор института  
информационных технологий и  
автоматизированных систем  
\_\_\_\_\_ Л.Д. Павлова  
подпись  
« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Математические и инструментальные методы анализа данных

09.04.01 «Информатика и вычислительная техника»  
(направленность (профиль): «Информатика и вычислительная техника»)

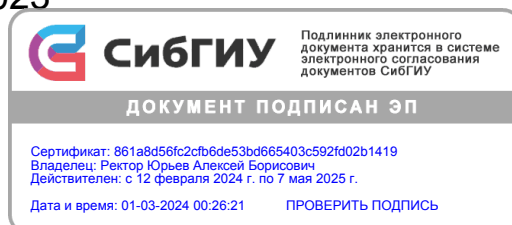
Квалификация выпускника  
Магистр

Форма обучения  
Очная форма

Срок обучения: 2 года

Год начала подготовки 2023

Новокузнецк  
2023



## 1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование целостного представления о методах и алгоритмах сбора, хранения и анализа больших объемов данных.

Задачами учебной дисциплины являются:

- знакомство с основными понятиями в области обработки «больших данных»; знакомство с методами и инструментами получения, хранения и анализа данных; получение практических навыков обработки и анализа данных с использованием специализированного программного обеспечения.

## 2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.01 «Информатика и вычислительная техника».

Учебная дисциплина опирается на базовые знания и компетенции, полученные в процессе получения предыдущего образования.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Анализ и оптимизация бизнес-процессов;
- Машинное обучение для задач управления;
- Разработка и реализация проектов 1;
- Python для анализа данных;
- Разработка и реализация проектов 2;
- Научно-исследовательская работа;
- Выполнение и защита выпускной квалификационной работы.

## 3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

### – Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен разрабатывать и применять алгоритмы интеллектуального	ПК-2.2 Собирает данные из различных источников и осуществляет их подготовку для	– знать: методы сбора и интеграции данных из различных источников, их подготовки,

	анализа больших объемов данных для управления технологическими системами	анализа	первичной обработки и визуализации . – уметь: осуществлять сбор, первичную обработку и визуализацию данных при решении профессиональных задач. – владеть: навыками сбора, подготовки, первичной обработки и визуализации данных с использованием специализированных аналитических платформ.
		ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных	– знать: теоретические и прикладные основы анализа больших объемов данных и построения моделей на основе данных. – уметь: использовать типовые алгоритмы анализа и обработки больших объемов данных. – владеть: навыками построения моделей на основе данных в специализированных аналитических платформах.

#### 4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа включает в себя занятия лекционного типа (лекции), занятия семинарского типа (семинары, практические занятия, практикумы), промежуточную аттестацию обучающихся и иные формы взаимодействия обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации ООП на иных условиях, в том числе при проведении промежуточной аттестации обучающихся. Контактная работа может проводиться с применением электронного обучения, дистанционных образовательных технологий.

## Объем учебной дисциплины

Семестр / курс		<b>ИТОГО</b>	<b>1 семестр</b>
Форма промежуточной аттестации			экзамен, зачет с оценкой по КР
Трудоёмкость	<i>академ. час.</i>	<b>216</b>	216
	<i>зачетных единиц</i>	<b>6</b>	6
Лекции, <i>академ. час.</i>		<b>16</b>	16
в форме практической подготовки		<b>0</b>	0
Лабораторные работы, <i>академ. час.</i>		<b>0</b>	0
в форме практической подготовки		<b>0</b>	0
Практические занятия, <i>академ. час.</i>		<b>32</b>	32
в форме практической подготовки		<b>0</b>	0
Курсовая работа, <i>академ. час.</i>		<b>36</b>	36
в форме практической подготовки		<b>0</b>	0
Консультации, <i>академ. час.</i>		<b>0</b>	0
в форме практической подготовки		<b>0</b>	0
Самостоятельная работа, <i>академ. час.</i>		<b>87</b>	87
в форме практической подготовки		<b>0</b>	0
Контроль, <i>академ. час.</i>		<b>45</b>	45
в форме практической подготовки		<b>0</b>	0

## Содержание учебной дисциплины

Раздел 1 Введение в науку о данных;

Тема 1.1 Основы анализа данных (Принципы анализа данных, классический и информационный подходы. Структурированные данные. Источники данных. Сбор данных. Требования к данным. Подготовка данных к анализу);

Тема 1.2 Концепция Big Date (Большие данные) (Предпосылки формирования тренда больших данных; определение термина «большие данные». Сферы использования больших данных. Модель MapReduce; пути развития передовых аналитических инструментов; подходы к построению моделей для эффективного использования);

Тема 1.3 Консолидация, визуализация и подготовка данных (Организация хранилища данных. Хранилища данных и анализ. Визуализаторы общего назначения. Многомерная модель данных. OLAP-анализ. Качество данных. Сбор и интеграция данных. Виды источников данных. Методы интеграции данных. Преобразование и обогащение данных. Методы трансформации данных.);

Раздел 2 Интеллектуальный анализ данных и машинное обучение;

Тема 2.1 Методы и инструменты интеллектуального анализа данных (Задачи интеллектуального анализа данных (ИАД) и их классификация. Модели и методы ИАД. Цели, задачи и методы машинного обучения.);

Тема 2.2 Кластерный анализ (Введение в кластеризацию. Алгоритмы k-means, g-means. Иерархический кластерный анализ. Сети и карты Кохонена);

Тема 2.3 Классификация и регрессия (Метод опорных векторов. Метод k-ближайших соседей. Байесовский классификатор. Основы логистической регрессии. Правила классификации. Деревья решений, случайный лес. Простая линейная регрессия. Искусственные нейронные сети).

### 5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	Введение в науку о данных		
Тема 1.1.	Основы анализа данных	2	
Тема 1.2.	Концепция Big Data (Большие данные)	2	
Тема 1.3.	Консолидация, визуализация и подготовка данных	2	
Раздел 2.	Интеллектуальный анализ данных и машинное обучение		
Тема 2.1.	Методы и инструменты интеллектуального анализа данных	2	
Тема 2.2.	Кластерный анализ	2	
Тема 2.3.	Классификация и регрессия	6	
<b>Итого:</b>		<b>16</b>	<b>0</b>

### 6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	Темы практических занятий (семинаров)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Тема 1.3.	Консолидация, визуализация и подготовка данных	10	
Тема 2.2.	Кластерный анализ	8	
Тема 2.3.	Классификация и регрессия	14	
<b>Итого:</b>		<b>32</b>	<b>0</b>

### 7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	Темы лабораторных работ	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
	<i>Отсутствуют</i>		
<b>Итого:</b>		<b>0</b>	<b>0</b>

### 8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	Темы курсовых работ (проектов)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1; Раздел 2.	<p>1) Построение описательной модели и анализ ассоциаций по данным о клиентах организации предоставляющей услуги (мобильный оператор, интернет магазин, розничная торговая сеть, банк, государственные органы и т.п.).</p> <p>2) Построение описательной модели и сегментация клиентов организации предоставляющей услуги (мобильный оператор, интернет магазин, розничная торговая сеть, банк, страховая компания и т.п.).</p> <p>3) Построение модели оценки качества конечного продукта на основе физико-химических свойств, протекания технологического процесса и т.п.</p> <p>4) Построение предсказательной модели поведения клиента организации предоставляющей услуги (банк, издательство, мобильный оператор и т.д.).</p> <p>5) Построение скоринговой модели для оценки риска невозврата</p>	36	

	кредита заемщиком. 6) Построение диагностической модели по данным предметной области (медицина, техническая диагностика и т.п.). 7) Построение рекомендательной системы для посетителей интернет ресурса.		
<b>Итого:</b>		<b>36</b>	<b>0</b>

## 9 Виды самостоятельной работы

№ раздела / темы дисциплины	Виды самостоятельной работы	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	19	
Раздел 2.	1. Изучение лекционного материала; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	68	
<i>Курсовая работа</i>	<i>Выполнение курсовой работы</i>	36	0
<i>Контроль</i>	<i>Подготовка к экзамену</i>	45	
<b>Итого:</b>		<b>168</b>	<b>0</b>

## 10 Учебно-методическое и информационное обеспечение учебной дисциплины

### а) литература:

1 Гасанов, Э. Э. Интеллектуальные системы. Теория хранения и поиска информации : учебник для вузов / Э. Э. Гасанов, В. Б. Кудрявцев. – 2-е изд., испр. и доп. – Москва : Юрайт, 2023. – 271 с. – ISBN 978-5-534-08684-3. – URL: <https://urait.ru/bcode/513151> (дата обращения: 10.04.2023);

2 Новиков, Ф. А. Символический искусственный интеллект: математические основы представления знаний : учебное пособие для вузов / Ф. А. Новиков. – Москва : Юрайт, 2023. – 278 с. – ISBN 978-5-534-

00734-3. – URL: <https://urait.ru/bcode/512382> (дата обращения: 10.04.2023);

3 Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. – Москва : Юрайт, 2022. – 121 с. – ISBN 978-5-534-09837-2. – URL: <https://urait.ru/bcode/492609> (дата обращения: 10.04.2023);

4 Горбаченко, В. И. Интеллектуальные системы: нечеткие системы и сети : учебное пособие для вузов / В. И. Горбаченко, Б. С. Ахметов, О. Ю. Кузнецова. – 2-е изд., испр. и доп. – Москва : Юрайт, 2023. – 105 с. – ISBN 978-5-534-08359-0. – URL: <https://urait.ru/bcode/514580> (дата обращения: 10.04.2023);

5 Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. – Москва : Юрайт, 2023. – 397 с. – ISBN 978-5-534-02126-4. – URL: <https://urait.ru/bcode/511651> (дата обращения: 10.04.2023);

6 Загорюлько, Ю. А. Искусственный интеллект. Инженерия знаний : учебное пособие для вузов / Ю. А. Загорюлько, Г. Б. Загорюлько. – Москва : Юрайт, 2022. – 93 с. – ISBN 978-5-534-07198-6. – URL: <https://urait.ru/bcode/494205> (дата обращения: 10.04.2023);

7 Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для вузов / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. – Москва : Юрайт, 2023. – 243 с. – ISBN 978-5-534-01042-8. – URL: <https://urait.ru/bcode/511999> (дата обращения: 10.04.2023);

8 Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. – 2-е изд., испр. – Москва : ИНТУИТ. Бинوم. Лаборатория знаний, 2008. – 383 с. – ISBN 978-5-94774-819-2. – URL: <https://biblioclub.ru/index.php?page=book&id=233055> (дата обращения: 10.04.2023).

**б) ресурсы информационно-телекоммуникационной сети «Интернет»:**

1 Консультант студента : электронно-библиотечная система / ООО «КОНСУЛЬТАНТ СТУДЕНТА». – Москва, [200 – ]. – URL: <http://www.studentlibrary.ru>. – Режим доступа: для авторизир. пользователей;

2 ЛАНЬ : электронно-библиотечная система : [коллекция «Инженерно-технические науки»] / ООО «Издательство ЛАНЬ». – Санкт-Петербург, [200 – ]. – URL: <http://e.lanbook.com>. – Режим доступа: для авторизир. пользователей;

3 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 – ]. – URL: <http://elibrary.ru>. – Режим доступа: по подписке;

4 Национальная электронная библиотека (НЭБ) : информационная система / ФГБУ «РГБ». – Москва, [2015 – ]. – URL: <http://rusneb.ru>. – Режим доступа: по подписке;



5 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство ЮРАЙТ». – Москва, [200 – ]. – URL: <https://urait.ru>. – Режим доступа: для авторизир. пользователей;

6 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 – ]. – URL: <https://biblioclub.ru>. – Режим доступа: для авторизир. пользователей. – URL: <http://www.biblioclub.ru>;

7 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 – ]. – URL: <http://library.sibsiu.ru/LibrELibraryFullText.asp>. – Режим доступа: для авторизир. пользователей. – URL: <https://library.sibsiu.ru/LibrELibraryFullText.asp>;

8 Электронные периодические издания ИВИС : универсальная база данных / ООО «ИВИС». – Москва, [200 – ]. – URL: <http://eivis.ru>. – Режим доступа: по подписке;

9 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 – ]. – URL: <http://libr.sibsiu.ru>. – URL: <https://libr.sibsiu.ru>.

**в) лицензионное и свободно распространяемое программное обеспечение:**

- Adobe Acrobat Reader;
- Deductor Academic;
- Loginom Community;
- Microsoft Office.

**г) базы данных и информационно-справочные системы:**

1 ГАРАНТ : справочно-правовая система / ООО «Правовой центр «Гарант». – Кемерово, [200 – ]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

2 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 – ]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 – ]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.

## **11 Материально-техническое обеспечение учебной дисциплины**

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;

- учебную аудиторию для проведения занятий семинарского типа (практических занятий);
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ;
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.04.01 «Информатика и вычислительная техника».

Составитель(и):

доцент Кораблина Татьяна Валентиновна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

## Приложение

### Аннотация

**рабочей программы дисциплины «Математические и инструментальные методы анализа данных»**

**по направлению подготовки (специальности)**

**09.04.01 «Информатика и вычислительная техника»  
(направленность (профиль): «Информатика и вычислительная техника»)**

**форма обучения – Очная форма**

### **1 Цели и задачи освоения учебной дисциплины**

Целями учебной дисциплины являются:

- формирование целостного представления о методах и алгоритмах сбора, хранения и анализа больших объемов данных.

Задачами учебной дисциплины являются:

- знакомство с основными понятиями в области обработки «больших данных»; знакомство с методами и инструментами получения, хранения и анализа данных; получение практических навыков обработки и анализа данных с использованием специализированного программного обеспечения.

### **2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)**

Учебная дисциплина относится к учебным дисциплинам части, формируемой участниками образовательных отношений **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.04.01 «Информатика и вычислительная техника».

Учебная дисциплина опирается на базовые знания и компетенции, полученные в процессе получения предыдущего образования.

Учебная дисциплина дополняет знания, умения и навыки, получаемые по одновременно изучаемым и последующим дисциплинам:

- Анализ и оптимизация бизнес-процессов;
- Машинное обучение для задач управления;
- Разработка и реализация проектов 1;
- Python для анализа данных;
- Разработка и реализация проектов 2;
- Научно-исследовательская работа;
- Выполнение и защита выпускной квалификационной работы.

### 3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

#### – Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен разрабатывать и применять алгоритмы интеллектуального анализа больших объемов данных для управления технологическими системами	ПК-2.2 Собирает данные из различных источников и осуществляет их подготовку для анализа	– знать: методы сбора и интеграции данных из различных источников, их подготовки, первичной обработки и визуализации . – уметь: осуществлять сбор, первичную обработку и визуализацию данных при решении профессиональных задач. – владеть: навыками сбора, подготовки, первичной обработки и визуализации данных с использованием специализированных аналитических платформ.
		ПК-2.3 Выбирает и использует методы и инструментальные средства для проведения анализа больших объемов данных, строит модели на основе данных	– знать: теоретические и прикладные основы анализа больших объемов данных и построения моделей на основе данных. – уметь: использовать типовые алгоритмы анализа и обработки больших объемов данных. – владеть: навыками построения моделей на основе данных в специализированных аналитических платформах.

#### 4 Объем учебной дисциплины

Семестр / курс		<b>ИТОГО</b>	<b>1 семестр</b>
Форма промежуточной аттестации			экзамен, зачет с оценкой по КР
Трудоёмкость	<i>академ. час.</i>	<b>216</b>	216
	<i>зачетных единиц</i>	<b>6</b>	6
Лекции, <i>академ. час.</i>		<b>16</b>	16
в форме практической подготовки		<b>0</b>	0
Лабораторные работы, <i>академ. час.</i>		<b>0</b>	0
в форме практической подготовки		<b>0</b>	0
Практические занятия, <i>академ. час.</i>		<b>32</b>	32
в форме практической подготовки		<b>0</b>	0
Курсовая работа, <i>академ. час.</i>		<b>36</b>	36
в форме практической подготовки		<b>0</b>	0
Консультации, <i>академ. час.</i>		<b>0</b>	0
в форме практической подготовки		<b>0</b>	0
Самостоятельная работа, <i>академ. час.</i>		<b>87</b>	87
в форме практической подготовки		<b>0</b>	0
Контроль, <i>академ. час.</i>		<b>45</b>	45
в форме практической подготовки		<b>0</b>	0

## **5 Краткое содержание учебной дисциплины**

В структуре учебной дисциплины выделяются следующие основные разделы (темы):

Раздел 1 Введение в науку о данных;

Тема 1.1 Основы анализа данных (Принципы анализа данных, классический и информационный подходы. Структурированные данные. Источники данных. Сбор данных. Требования к данным. Подготовка данных к анализу);

Тема 1.2 Концепция Big Date (Большие данные) (Предпосылки формирования тренда больших данных; определение термина «большие данные». Сферы использования больших данных. Модель MapReduce; пути развития передовых аналитических инструментов; подходы к построению моделей для эффективного использования);

Тема 1.3 Консолидация, визуализация и подготовка данных (Организация хранилища данных. Хранилища данных и анализ. Визуализаторы общего назначения. Многомерная модель данных. OLAP-анализ. Качество данных. Сбор и интеграция данных. Виды источников данных. Методы интеграции данных. Преобразование и обогащение данных. Методы трансформации данных.);

Раздел 2 Интеллектуальный анализ данных и машинное обучение;

Тема 2.1 Методы и инструменты интеллектуального анализа данных (Задачи интеллектуального анализа данных (ИАД) и их классификация. Модели и методы ИАД. Цели, задачи и методы машинного обучения.);

Тема 2.2 Кластерный анализ (Введение в кластеризацию. Алгоритмы k-means, g-means. Иерархический кластерный анализ. Сети и карты Кохонена);

Тема 2.3 Классификация и регрессия (Метод опорных векторов. Метод k-ближайших соседей. Байесовский классификатор. Основы логистической регрессии. Правила классификации. Деревья решений, случайный лес. Простая линейная регрессия. Искусственные нейронные сети).

**6 Составитель(и):**

доцент Кораблина Татьяна Валентиновна (кафедра прикладных информационных технологий и программирования).