Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»

Кафедра систем сбора и обработки данных
Кафедра вычислительной техники
Кафедра автоматизированных систем управления

“УТВЕРЖДАЮ”
ДЕКАН АВТФ
к.т.н. Рева И. Л.

ДОКУМЕНТ ПОДПИСАН ПРОСТОЙ НЕКВАЛИФИЦИРОВАННОЙ ЭЛЕКТРОННОЙ ПОДПИСЬЮ
31.08.2026
Владелец: к.т.н. Рева И. Л.
Срок действия: не ограничен
Адрес хранения электронного документа:
https://www.nstu.ru/university/info/sveden/education

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Анализ больших данных

Образовательная программа: 09.03.02 Информационные системы и технологии, профиль: Информационные системы в промышленности и бизнесе

Курс: 4, семестр : 7

Факультет автоматики и вычислительной техники,

		Семестр
№	**Вид деятельности**	**7**
1	Всего зачетных единиц (кредитов)	4
2	Всего часов	144
3	Всего занятий в контактной форме, час.	40
4	Лекции, час.	18
5	Практические занятия, час.	0
6	Лабораторные занятия, час.	18
7	из них в активной и интерактивной форме, час.	9
8	из них практической подготовки, час.	8
9	Аттестация, час.	2
10	Консультации, час.	2
11	Самостоятельная работа, час.	104
12	Виды самостоятельной работы (курсовой проект, курсовая работа, РГЗ(Р)/реферат, подготовка к контрольной работе)	
13	Вид аттестации	3

Новосибирск 2026

Рабочая программа составлена на основании федерального государственного образовательного стандарта высшего образования по направлению (специальности): 09.03.02 Информационные системы и технологии

ФГОС введен в действие приказом №926 от 19.09.2017 г. , дата утверждения: 12.10.2017 г.

Место дисциплины в структуре учебного плана: Блок 1, формируемая участниками образовательных отношений

Рабочая программа разработана на основе компетентностной модели выпускника по направлению (специальности): 09.03.02 Информационные системы и технологии

Рабочая программа обсуждена на заседании кафедры

ССОД, 31.08.2026
ВТ, 31.08.2026
АСУ, 31.08.2026

Утверждена на совете факультета автоматики и вычислительной техники, протокол № 8 от 31.08.2026

Заведующий кафедрой:

доцент, к.т.н. Бакаев М. А.
доцент, к.т.н. Якименко А. А.
доцент, к.т.н. Томилов И. Н.

Ответственный за образовательную программу:

1. Внешние требования

Таблица 1.1

Группа	Компетенции
	Индикаторы достижения компетенций
Профессиональные компетенции	ПК-1/НА Способность проводить исследования на всех этапах жизненного цикла программных средств
	ПК-1/НА. 1 Иметь навыки завершения проекта в соответствии с полученным заданием
Универсальные компетенции	УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений
	УК-2. 2 Умеет определять круг задач в рамках избранных видов профессиональной деятельности, планировать собственную деятельность исходя из имеющихся ресурсов; соотносить главное и второстепенное, решать поставленные задачи в рамках избранных видов профессиональной деятельности.

2. Требования НГТУ к планируемым результатам обучения, соотнесенным с индикаторами достижения компетенций

Таблица 2.1

Индикаторы достижения компетенций	Формы организации занятий
Результаты обучения	
ПК-1/НА. 1 Иметь навыки завершения проекта в соответствии с полученным заданием	
Уметь разрабатывать алгоритмы и методы для обработки больших данных.	Лекции; Лабораторные работы; Самостоятельная работа
УК-2. 2 Умеет определять круг задач в рамках избранных видов профессиональной деятельности, планировать собственную деятельность исходя из имеющихся ресурсов; соотносить главное и второстепенное, решать поставленные задачи в рамках избранных видов профессиональной деятельности.	
Знать принципы обработки больших данных в распределенных вычислительных системах.	Лекции; Лабораторные работы; Самостоятельная работа

3. Содержание и структура дисциплины

Таблица 3.1

Темы лекций	Часы	Из них в форме практ. подг., час.	Активные формы, час.	Индикаторы достижения компетенций	Учебная деятельность
Семестр: 7					
Дидактическая единица: Большие данные. Основы систем больших данных.					
1. Понятие «большие данные». Характеристики больших данных. Принципы работы с большими данными. Большие данные в научных исследованиях. Экосистема больших данных.	2	0	0	ПК-1/НА.1, УК-2.2	Лекция.

2. Распределенные файловые системы. Распределенные фреймворки. Бенчмаркинг. Серверное программирование. Планирование. Системы развертывания.	2	0	0	ПК-1/НА.1, УК-2.2	Лекция.
3. Интеграция данных. Машинное обучение. Базы данных NoSQL и новые SQL базы данных.	2	0	0	ПК-1/НА.1, УК-2.2	Лекция.
Дидактическая единица: Архитектура системы обработки больших данных.					
4. Уровни в системах обработки больших данных. Прием данных (Data Ingestion). Сбор данных (Data Staging). Анализ данных (Analysis Layer). Представление результатов (Consumption Layer).	3	0	0	ПК-1/НА.1, УК-2.2	Лекция.
Дидактическая единица: Параллельные алгоритмы для работы с данными.					
5. Операторы Map и Reduce. Оператор Map (предварительная обработка). Оператор Reduce (свертка). Лямбда-архитектура. Каппа-архитектура.	3	0	0	ПК-1/НА.1, УК-2.2	Лекция.
Дидактическая единица: Программные платформы и системы для больших данных .					
6. Системы управления потоками данных. Системы хранения больших данных. Платформы больших данных. Обработка данных в реальном времени	2	0	0	ПК-1/НА.1, УК-2.2	Лекция.
7. Системы управления большими данными. Аналитические платформы. Состав и возможности программного комплекса Apache Hadoop.	2	0	1	ПК-1/НА.1, УК-2.2	Лекция.
8. Языки поисковых запросов для Hadoop. Принципы работы Apache Spark. Прочие компоненты экосистемы Hadoop.	2	0	1	ПК-1/НА.1, УК-2.2	Лекция.

Темы лабораторных работ	Часы	Из них в форме практ. подг., час.	Активные формы, час.	Индикаторы достижения компетенций	Учебная деятельность
Семестр: 7					
Дидактическая единица: Большие данные. Основы систем больших данных.					

1. Поиск и определение Big Data.	2	0	1	ПК-1/НА.1	Лабораторная работа.
Дидактическая единица: Архитектура системы обработки больших данных.					
2. Хранение больших данных.	4	0	1	ПК-1/НА.1	Лабораторная работа.
Дидактическая единица: Параллельные алгоритмы для работы с данными.					
3. Аналитические платформы: классификация и особенности применения	2	0	1	ПК-1/НА.1, УК-2.2	Лабораторная работа.
4. Прогнозирование с помощью линейной регрессии.	2	0	1	ПК-1/НА.1, УК-2.2	Лабораторная работа.
5. Алгоритм класстеризации больших данных.	2	0	1	ПК-1/НА.1, УК-2.2	Лабораторная работа.
Дидактическая единица: Программные платформы и системы для больших данных .					
6. Поиск ассоциативных правил.	2	0	0	ПК-1/НА.1, УК-2.2	Лабораторная работа.
7. Классификация с помощью нейронной сети.	2	0	1	ПК-1/НА.1, УК-2.2	Лабораторная работа.
8. Классификация с помощью деревьев решений.	2	0	1	ПК-1/НА.1, УК-2.2	Лабораторная работа.

Темы для самостоятельного изучения	Часы	Из них в форме практ. подг., час.	Активные формы, час.	Индикаторы достижения компетенций	Учебная деятельность
Семестр: 7					
Дидактическая единица: Большие данные. Основы систем больших данных.					
1. История появления.Технические сложности работы с большими данными. Общая схема анализа. Формы представления данных, типы и виды данных.	18	3	0	ПК-1/НА.1, УК-2.2	Самостоятельная работа.
Дидактическая единица: Параллельные алгоритмы для работы с данными.					
2. Программное обеспечение в области больших данных. Языки визуального моделирования.	18	5	0	ПК-1/НА.1, УК-2.2	Самостоятельная работа.

3.1 Практическая подготовка

Таблица 3.2

№	Темы занятий	Формы организации занятий	Содержание практической подготовки (виды работ)
1	История появления.Технические сложности работы с большими данными. Общая схема анализа. Формы представления данных, типы и виды данных.	Самост.	Выполняет следующие виды работ: этапы моделирования. Процесс построения моделей.
2	Программное обеспечение в области больших данных. Языки визуального моделирования.	Самост.	Выполняет следующие виды работ: реляционные базы данных. Параллельные базы данных.

3.2 Самостоятельная работа обучающегося

Таблица 3.3

№	Виды самостоятельной работы	Индикаторы достижения компетенций	Часы на выполнение	Часы на консультации
Семестр: 7				
1	Подготовка к занятиям	ПК-1/НА.1,УК-2.2	28	0
, подробная информация приведена в приложении №1 : Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 c. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/81324.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 c. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/89404.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей			
2	Подготовка к аттестации	ПК-1/НА.1,УК-2.2	40	2
Подготовка к сдаче зачета.: Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 c. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/89404.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей			
3	Самостоятельное изучение теоретического материала	ПК-1/НА.1,УК-2.2	36	0
Студент изучает темы, приведенные в таблице 3.1 : Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 c. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/81324.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей			

3.3 Технология обучения

Для организации и контроля самостоятельной работы обучающихся, а также проведения консультаций применяются информационно-коммуникационные технологии (табл. 3.4).

Таблица 3.4

Деятельность	Информационно-коммуникационные технологии
Информирование	e-mail; Портал НГТУ; Среда электронного обучения НГТУ
Консультирование	e-mail; Портал НГТУ; Среда электронного обучения НГТУ
Размещение учебных материалов	Среда электронного обучения НГТУ; ЭБС

Таблица 3.5

Активные и интерактивные формы проведения занятий

№	Наименование активных форм
1	Дискуссия
Краткое описание применения: Обсуждение материала.	

4. Правила аттестации обучающихся по дисциплине

Для аттестации обучающихся по дисциплине используется балльно-рейтинговая система (БРС), позволяющая выставлять оценки по традиционной шкале и 15-уровневой ECTS. Краткая информация о БРС приведена в табл. 4.1.

Таблица 4.1

Оцениваемые виды деятельности обучающихся	Мин. балл	Максимальный балл
Семестр: 7		
Самостоятельное изучение теоретического материала:	10	20
Лекция:	5	10
Лабораторная:	20	40
Контролирующие материалы приводятся в "Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/81324.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей"		
Зачет:	15	30
Контролирующие материалы приводятся в "Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/89404.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей"		

В таблице 4.2 представлено соответствие форм контроля заявляемым требованиям к результатам освоения дисциплины.

Таблица 4.2

Коды компетенций	Индикаторы достижения компетенций	Формы контроля	
		Защита Л/Р	Зачет
ПК-1/НА	ПК-1/НА 1. Иметь навыки завершения проекта в соответствии с полученным заданием	+	+
УК-2	УК-2 2. Умеет определять круг задач в рамках избранных видов профессиональной деятельности, планировать собственную деятельность исходя из имеющихся ресурсов; соотносить главное и второстепенное, решать поставленные задачи в рамках избранных видов профессиональной деятельности.	+	+

Фонд оценочных средств по дисциплине представлен в приложении № 1 к рабочей программе.

5. Литература

Основная литература

1. Макшанов, А. В. Большие данные. Big Data : учебник для вузов / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — Санкт-Петербург : Лань, 2021. — 188 с. — ISBN 978-5-8114-6810-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/165835 (дата обращения: 10.03.2021). — Режим доступа: для авториз. пользователей.

Дополнительная литература

1. Виктор, Майер-Шенбергер Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / Майер-Шенбергер Виктор, Кукьер Кеннет ; перевод И. Гайдюк. — Москва : Манн, Иванов и Фербер, 2014. — 220 с. — ISBN 978-5-91657-936-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/39150.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей

Интернет-ресурсы

1. Волкова, В. М. Методические указания к выполнению лабораторных работ по предмету «Методы анализа знаний и данных»: учебно-методическое пособие / В. М. Волкова. – Новосибирск, [2011]. – Текст : электронный // Электронно-библиотечная система НГТУ : [сайт]. – Новосибирск, 2011– . – URL: https://elibrary.nstu.ru/source?id=14481 (дата обращения: 09.03.2021). – Режим доступа: для авториз. пользователей.

6. Методическое и программное обеспечение, информационные технологии

6.1 Методическое обеспечение

1. Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/81324.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей

2. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: http://www.iprbookshop.ru/89404.html (дата обращения: 10.03.2021). — Режим доступа: для авторизир. пользователей

6.2 Специализированное программное обеспечение

1 СУБД Microsoft Microsoft SQL Server

2 Создание отчётов для лабораторных работ. Microsoft Microsoft Office

6.3 Информационные технологии

При осуществлении образовательного процесса по дисциплине применяются элементы дистанционных образовательных технологий, а также синхронного и асинхронного взаимодействия в электронной информационно-образовательной среды НГТУ.

7. Материально-техническое обеспечение

Презентационное оборудование

№	Наименование	Назначение
1	Презентационное оборудование (мультимедиа-проектор, экран, компьютер для управления)	Проведение лекций

2	Компьютерный класс №5	Проведение лабораторных работ