

**МИНОБРНАУКИ РОССИИ**  
**федеральное государственное бюджетное образовательное учреждение**  
**высшего образования**  
**«Нижегородский государственный технический университет**  
**им. Р.Е. Алексеева» (НГТУ)**

---

Институт радиоэлектроники и информационных технологий

(Полное и сокращенное название института, реализующего данное направление)

УТВЕРЖДАЮ:

Директор института:

\_\_\_\_\_ Мякинков А.В.

подпись

ФИО

“10” ИЮНЯ 2021 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Б1.В.ОД.6 Методы Data Mining**

(индекс и наименование дисциплины по учебному плану)

для подготовки бакалавров

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность: Программное обеспечение средств вычислительной техники и автоматизированных систем

Форма обучения: очная

Год начала подготовки 2020, 2021

Выпускающая кафедра ВСТ

Кафедра-разработчик ВСТ

Объем дисциплины 72 / 2  
часов/з.е

Промежуточная аттестация зачет

Разработчик: Суркова А.С., д.т.н., доцент

Нижний Новгород, 2021

Рабочая программа дисциплины: разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования (ФГОС ВО 3++) по направлению подготовки 09.03.01 Информатика и вычислительная техника, утвержденного приказом МИНОБРНАУКИ РОССИИ от 19 сентября 2017 года № 929 на основании учебного плана принятого УМС НГТУ протокол от 10.06.2021 № 6

Рабочая программа одобрена на заседании кафедры ВСТ протокол от 12.05.2021 № 10

Зав. кафедрой д.т.н, доцент, Жевнерчук Д.В. \_\_\_\_\_  
(подпись)

Программа рекомендована к утверждению ученым советом института ИРИТ, Протокол от 10.06.2021 № 1

Рабочая программа зарегистрирована в УМУ, регистрационный № 09.03.01-П-29

Начальник МО \_\_\_\_\_

Заведующая отделом комплектования НТБ

\_\_\_\_\_  
(подпись)

Н.И. Кабанина

## СОДЕРЖАНИЕ

<b>1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....</b>	<b>4</b>
1.1 Цель освоения дисциплины .....	4
1.2 Задачи освоения дисциплины (модуля) .....	4
<b>2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....</b>	<b>4</b>
<b>3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ).....</b>	<b>5</b>
<b>4. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОПВО .....</b>	<b>5</b>
<b>5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....</b>	<b>6</b>
5.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ ПО СЕМЕСТРАМ.....	6
5.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ .....	7
<b>6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ....</b>	<b>ОШИБКА! ЗАКЛАДКА НЕ ОПРЕДЕЛЕНА.</b>
6.1 ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ .....	<b>ОШИБКА! ЗАКЛАДКА НЕ ОПРЕДЕЛЕНА.</b>
6.2 ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ.....	9
<b>7. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ .....</b>	<b>10</b>
<b>8. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....</b>	<b>10</b>
8.1 ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....	11
8.2 ПЕРЕЧЕНЬ СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ .....	11
8.3 ПЕРЕЧЕНЬ СОВРЕМЕННЫХ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....	11
<b>9. ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОВЗ .....</b>	<b>11</b>
<b>10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ, НЕОБХОДИМОЕ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ .....</b>	<b>12</b>
<b>11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.....</b>	<b>13</b>
11.1 ОБЩИЕ МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ, ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ .....	13
11.2 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ЗАНЯТИЙ ЛЕКЦИОННОГО ТИПА .....	14
11.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА ЛАБОРАТОРНЫХ РАБОТАХ .....	14
11.4 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА КУРСОВОЙ РАБОТЕ .....	15
11.5 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО САМОСТОЯТЕЛЬНОЙ РАБОТЕ ОБУЧАЮЩИХСЯ.....	14
<b>12. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....</b>	<b>15</b>
12.1 ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ, НАВЫКОВ И (ИЛИ) ОПЫТА В ХОДЕ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ.....	15

## **1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

### **1.1 Цель освоения дисциплины**

Целью освоения дисциплины «Методы Data Mining» является развитие компетенций в области решения задач интеллектуальной обработки данных, машинного обучения, а также применения различных методов к решению практических профессиональных задач.

### **1.2 Задачи освоения дисциплины (модуля)**

Дисциплина «Методы Data Mining» способствует подготовке студентов к решению следующих профессиональных задач:

- объектно-ориентированное моделирование предметной области, разработка ER-моделей;
- реализация алгоритмов обработки данных;
- применение современных инструментальных средств разработки программного обеспечения;
- разработка моделей программных систем на всех стадиях их жизненного цикла;

## **2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

Учебная дисциплина Б1.В.ОД.6 «Методы Data Mining» включена в перечень обязательных дисциплин вариативной части (формируемой участниками образовательных отношений), определяющий направленность ОП. Дисциплина реализуется в соответствии с требованиями ФГОС, ОП ВО и УП.

Дисциплина базируется на дисциплинах программы бакалавриата по направлению «Информатика и вычислительная техника» профиля «Программное обеспечение средств вычислительной техники и автоматизированных систем». Предшествующими курсами, на которых непосредственно базируется дисциплина «Методы Data Mining», являются:

- «Математика»,
- «Дискретные структуры»,
- «Теоретические основы алгоритмизации»,
- «Алгоритмы и структуры данных»,
- «Вычислительная математика».
- «Машинное обучение»
- «Системный анализ и принятие решений»
- «Исследование операций»

Дисциплина «Методы Data Mining» является основополагающей для изучения следующих дисциплин: «Информационно-поисковые системы», «Моделирование систем» также для преддипломной практики и выполнения ВКР.

### 3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)<sup>1</sup>

Таблица 3.1 - Формирование компетенций дисциплинам

Наименование дисциплин, формирующих компетенцию совместно	Семестры, формирования дисциплины							
	Компетенции берутся из Учебного плана по направлению подготовки бакалавра /специалиста/магистра»							
	1	2	3	4	5	6	7	8
<i>ПКС-3. Способен применять системный анализ, методы оптимизации, моделирование при разработке и тестировании программных комплексов</i>								
Дискретные структуры	+							
Теория графов и дискретная математика	+							
Теоретические основы алгоритмизации		+						
Математическая логика и теория алгоритмов		+						
Информационные модели построения АСО и У			+					
Технологическая (проектно-технологическая)				+				
Исследование операций					+			
Системный анализ и принятие решений					+			
Базы знаний					+			
Практика по получению профессиональных умений и опыта профессиональной деятельности						+		
Инструментальные средства разработки систем управления							+	
Организация и проектирование автоматизированных систем							+	
Тестирование программного обеспечения								+
Системы хранения данных								+
Информационно-поисковые системы								+
Моделирование систем								+
Преддипломная практика								+
Выполнение и защита ВКР								+

### 4. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОПВО

Таблица 4.1 - Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине			Оценочные средства	
					Текущего контроля	Промежуточной аттестации
ПКС-3. Способен применять системный анализ, методы оптимизации, моделирование при разработке и тестировании программных комплексов	ИПКС-3.2. Применяет методы оптимизации и моделирования при разработке и тестировании программных комплексов	<b>Знать:</b> - Этапы Data Mining и подходы к построению модели данных	<b>Уметь:</b> - решать задачи классификации, кластеризации, прогнозирования и осуществлять визуализацию результатов. - разрабатывать мо-	<b>Владеть:</b> - современными инструментами Data Mining	Решение индивидуальных заданий по вариантам, выполнение лабораторных работ	Вопросы для устного собеседования – 40 вопросов

			дели Text Mining			
--	--	--	------------------	--	--	--

## 5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 5.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 2 зач. ед. 72 часа, распределение часов по видам работ семестрам представлено в таблице 5.1.

Таблица 5.1 - Распределение трудоёмкости дисциплины по видам работ по семестрам для студентов очного обучения

Вид учебной работы	Трудоёмкость в час	
	Всего час.	В т.ч. по семестрам
		7 сем
<b>Формат изучения дисциплины</b>	с использованием элементов электронного обучения	
<b>Общая трудоёмкость</b> дисциплины по учебному плану	<b>72</b>	<b>72</b>
<b>1. Контактная работа:</b>	<b>38</b>	<b>38</b>
<b>1.1 Аудиторная работа, в том числе:</b>	<b>34</b>	<b>34</b>
занятия лекционного типа (Л)	17	17
занятия семинарского типа (ПЗ-семинары, практ. Занятия и др)		
лабораторные работы (ЛР)	17	17
<b>1.2 Внеаудиторная, в том числе</b>	<b>4</b>	<b>4</b>
курсовая работа (проект) (КР/КП) (консультация, защита)		
текущий контроль, консультации по дисциплине		
контактная работа на промежуточном контроле (КРА)		
<b>2. Самостоятельная работа (СРС)</b>	<b>34</b>	<b>34</b>
реферат/эссе (подготовка)		
расчётно-графическая работа (РГР) (подготовка)		
контрольная работа		
курсовая работа/проект (КР/КП) (подготовка)		
самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиум и т.д.)	<b>28</b>	<b>28</b>
Подготовка к экзамену (контроль)		
Подготовка к зачёту/ зачёту с оценкой (контроль)	<b>6</b>	<b>6</b>

## 5.2 Содержание дисциплины, структурированное по темам

Таблица 5.2 - Содержание дисциплины, структурированное по темам для студентов очного обучения

Планируемые (контролируемые) результаты освоения: код УК; ОПК; ПК и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы (час)					Вид СРС	Наименование используемых активных и интерактивных образовательных технологий	Реализация в рамках Практической подготовки (трудоемкость в часах)	Наименование разработанного Электронного курса (трудоемкость в часах)	
		Контактная работа				Самостоятельная работа студентов (час)					
		Лекции (час)	Лабораторные работы (час)	Практические занятия (час)	КСР						
Раздел 1. Математические основы Data Mining											
ПКС-3. ИПКС-3.2.	Тема 1.1 Основные понятия и определения Data Mining	1					Подготовка к лекциям [1-4]	Видеоконференция			
	Тема 1.2 Градиентный спуск	0,5				1	Подготовка к лекциям [1-4]	Видеоконференция			
	Тема 1.3 Эвристические алгоритмы. Генетический алгоритм. Алгоритм имитации отжига	0,5				1	Подготовка к лекциям [1-4]	Видеоконференция			
	Лабораторная работа 1: Методы оптимизации в машинном обучении		4			2	Работа над индивидуальным заданием по ЛР	Видеоконференция			
	Тема 1.4 Снижение размерности. Метод главных компонент	2				1	Подготовка к лекциям [1-4]	Видеоконференция			
	Лабораторная работа 2: Методы снижения размерностей		4			2	Работа над индивидуальным заданием по ЛР	Видеоконференция			
	Итого по 1 разделу	4	8		1	7					
Раздел 2. Регрессия и математическая статистика											
ПКС-3. ИПКС-3.2.	Тема 2.1 Линейная и множественная регрессия	2				1	Подготовка к лекциям [1-4]	Видеоконференция			
	Тема 2.2 Робастная и логистическая регрессия	2				1	Подготовка к лекциям [1-4]	Видеоконференция			
	Лабораторная работа 3: Регрессионные модели и анализ данных		4			3	Работа над индивидуальным заданием по ЛР	Видеоконференция			
	Итого по 2 разделу	4	4		1	5					
Раздел 3. Нейронные сети											

Планируемые (контролируемые) результаты освоения: код УК; ОПК; ПК и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы (час)					Вид СРС	Наименование используемых активных и интерактивных образовательных технологий	Реализация в рамках Практической подготовки (трудоемкость в часах)	Наименование разработанного Электронного курса (трудоемкость в часах)
		Контактная работа				Самостоятельная работа студентов (час)				
		Лекции (час)	Лабораторные работы (час)	Практические занятия (час)	КСР					
ПКС-3. ИПКС-3.2.	Тема 3.1 Нейронные сети. Основные понятия и определения.	1				1	Подготовка к лекциям [1-4]	Видеоконференция		
	Тема 3.2. Метод обратного распространения ошибки. Борьба с переобучением в нейронных сетях	1				1	Подготовка к лекциям [1-4]	Видеоконференция		
	Лабораторная работа 4: Нейронные сети		5			5	Работа над индивидуальным заданием по ЛР	Видеоконференция		
	Тема 3.3. Сверточные нейронные сети	2				1	Подготовка к лекциям [1-4]	Видеоконференция		
	Итого по 3 разделу	4	5		1	8				
Раздел 4. Нечеткость										
ПКС-3. ИПКС-3.2.	Тема 4.1 Нечеткие множества и нечеткая логика. Основные понятия и определения.	1				2	Подготовка к лекциям [1-4]	Видеоконференция		
	Тема 4.2. Нечеткие отношения и отображения. Нечеткие правила вывода. Аналогии	1				2	Подготовка к лекциям [1-4]	Видеоконференция		
	Тема 4.3. Алгоритм нечеткой кластеризации FCM	1				2	Подготовка к лекциям [1-4]	Видеоконференция		
	Тема 4.2. Ассоциативные правила	2				2	Подготовка к лекциям [1-4]	Видеоконференция		
	Итого по 4 разделу	5			1	8				
	Подготовка к зачёту/зачёту с оценкой (контроль)					6				
	Итого за семестр	17	17		4	34				



## 6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.

### 6.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

Комплект оценочных средств является неотъемлемой частью ФОС и хранится на кафедре «Вычислительные системы и технологии».

### 6.2 Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Таблица 6.1 - При текущем контроле (контрольные недели) и оценка выполнения лабораторных работ

Шкала оценивания	Экзамен
$40 < R \leq 50$	Отлично
$30 < R \leq 40$	Хорошо
$20 < R \leq 30$	Удовлетворительно
$0 < R \leq 20$	Неудовлетворительно

Таблица 6.2 – Критерии оценивания результата обучения по дисциплине и шкала оценивания

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Критерии оценивания результатов обучения			
		Оценка «неудовлетворительно» / «не зачтено» 0-59% от max рейтинговой оценки контроля	Оценка «удовлетворительно» / «зачтено» 60-74% от max рейтинговой оценки контроля	Оценка «хорошо» / «зачтено» 75-89% от max рейтинговой оценки контроля	Оценка «отлично» / «зачтено» 90-100% от max рейтинговой оценки контроля
ПКС-3. Способен применять системный анализ, методы оптимизации, моделирование при разработке и тестировании программных комплексов	ИПКС-3.2. Применяет методы оптимизации и моделирования при разработке и тестировании программных комплексов	Изложение учебного материала бессистемное, неполное, не знает основные методы интеллектуальной обработки данных	Фрагментарные, поверхностные знания методов интеллектуальной обработки данных. Лабораторные работы выполнены с затруднениями	Знает методы интеллектуальной обработки данных. Все лабораторные работы выполнены.	Имеет глубокие знания методов интеллектуальной обработки данных. Все лабораторные работы выполнены на высоком уровне

Таблица 6.3 - Критерии оценивания

Оценка	Критерии оценивания
Высокий уровень «5» (отлично)	оценку «отлично» заслуживает студент, освоивший знания, умения, компетенции и теоретический материал без пробелов; выполнивший все задания, предусмотренные учебным планом на высоком качественном уровне; практические навыки профессионального применения освоенных знаний сформированы.
Средний уровень «4» (хорошо)	оценку «хорошо» заслуживает студент, практически полностью освоивший знания, умения, компетенции и теоретический материал, учебные задания не оценены максимальным числом баллов, в основном сформировал практические навыки.
Пороговый уровень «3» (удовлетворительно)	оценку «удовлетворительно» заслуживает студент, частично с пробелами освоивший знания, умения, компетенции и теоретический материал, многие учебные задания либо не выполнил, либо они оценены числом баллов близким к минимальному, некоторые практические навыки не сформированы.

Минимальный уровень «2» (неудовлетворительно)	оценку «неудовлетворительно» заслуживает студент, не освоивший знания, умения, компетенции и теоретический материал, учебные задания не выполнил, практические навыки не сформированы.
---	--

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 7.1 Учебная литература

7.1.1. Бронфельд Г.Б. Основы искусственного интеллекта: Учеб.пособие / Г.Б. Бронфельд; НГТУ им.Р.Е.Алексеева. - Н.Новгород: [Б.и.], 2014. - 253 с. - ISBN 978-5-502-00111-3

7.1.2. Системы искусственного интеллекта. Практический курс: Учеб.пособие / В.А. Чулюков [и др.]; Под ред.И.Ф.Астаховой. - М.: БИНОМ. Лаб.знаний, 2008. - 293 с. - (Адаптивные и интеллектуальные системы). - ISBN 978-5-94774-731-7

7.1.3. Бажанов Ю.С. Системы искусственного интеллекта и принятия решений. Оптимизация баз знаний нечетких экспертных систем [Электронные текстовые данные]: Учеб.пособие / Ю.С. Бажанов, А.В. Бухнин, Д.А. Кобляков; НГТУ им.Р.Е.Алексеева. - Н.Новгород : [Б.и.], 2012. - 84 с. - ISBN 978-5-502-00067-3.

7.1.4. Сосинская С.С. Представление знаний в информационной системе. Методы искусственного интеллекта и представления знаний: Учеб.пособие / С.С. Сосинская. - Старый Оскол: ООО "ТНТ", 2011. - 216 с. - ISBN 978-5-94178-254-3

### 7.2 Справочно-библиографическая литература

— учебники и учебные пособия

7.2.1 Юре Лесковец, Ананд Раджараман, Джеффри Д. Ульман Анализ больших наборов данных. / Пер. с англ. Слинкин А. А. – М.: ДМК Пресс, 2016.  
[https://www.rulit.me/data/programs/resources/pdf/Analiz-bolshih-naborov-dannyh\\_RuLit\\_Me\\_649408.pdf](https://www.rulit.me/data/programs/resources/pdf/Analiz-bolshih-naborov-dannyh_RuLit_Me_649408.pdf)

7.2.2 Барсегян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009.  
<http://www.kist.ntu.edu.ua/textPhD/AnalizDannyyIProcessov.pdf>

### 7.3 Перечень журналов по профилю дисциплины:

7.3.1 Научно-технический и научно-производственный журнал Информационные технологии Журнал "Информационные технологии" (novtex.ru).

7.3.2 Информационные ресурсы России. Российская ассоциация электронных библиотек. Информационные Ресурсы России — Российская ассоциация электронных библиотек (aselibrary.ru).

7.3.3 Журнал «Информационные технологии и вычислительные системы». Журнал «Информационные технологии и вычислительные системы» - About journal (jitcs.ru)

### 7.4 Методические указания, рекомендации и другие материалы к занятиям

Материалы лекций (слайды), указания по решениям индивидуальных задач, в электронном варианте находятся на кафедре «Вычислительные системы и технологии». Их электронные варианты на электронные адреса групп в начале семестра.

## 8. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Учебный процесс по дисциплине обеспечен необходимым комплектом свободно распространяемого программного обеспечения (состав по дисциплине определен в настоящей РПД и подлежит обновлению при необходимости).

### 8.1 Перечень информационных справочных систем

Таблица 8.1 - Перечень электронных библиотечных систем

№	Наименование ЭБС	Ссылка к ЭБС
1	Лань	<a href="https://e.lanbook.com/">https://e.lanbook.com/</a>
2	Юрайт	<a href="https://biblio-online.ru/">https://biblio-online.ru/</a>

### 8.2 Перечень свободно распространяемого программного обеспечения

Таблица 8.2 – Программное обеспечение, используемое студентами очного обучения

Программное обеспечение, используемое в университете на договорной основе	Программное обеспечение свободного распространения
-	Adobe Acrobat Reader ( <a href="https://acrobat.adobe.com/ru/ru/acrobat/pdf-reader.html">https://acrobat.adobe.com/ru/ru/acrobat/pdf-reader.html</a> )
	Linux ( <a href="https://www.linux.com/">https://www.linux.com/</a> )
	OpenOffice (FreeWare) <a href="https://www.openoffice.org/ru/">https://www.openoffice.org/ru/</a>
	Редактор блок-схем ( <a href="https://app.diagrams.net/">https://app.diagrams.net/</a> )

### 8.3 Перечень современных профессиональных баз данных и информационных справочных систем

В таблице 8.3 указан перечень профессиональных баз данных и информационных справочных систем, к которым обеспечен доступ (удаленный доступ). Данный перечень подлежит обновлению в соответствии с требованиями ФГОС ВО.

В данном разделе могут быть приведены ресурсы (ссылки на сайты), на которых можно найти полезную для курса информацию, в т.ч. статистические или справочные данные, учебные материалы, онлайн курсы и т.д.

Таблица 8.3 – Перечень современных профессиональных баз данных и информационных справочных систем

№	Наименование профессиональной базы данных, информационно-справочной системы	Доступ к ресурсу (удаленный доступ с указанием ссылки/доступ из локальной сети университета)
1	2	3
1	База данных стандартов и регламентов РОССТАНДАРТ	<a href="https://www.rst.gov.ru/portal/gost/home/standarts">https://www.rst.gov.ru/portal/gost/home/standarts</a>
2	Перечень профессиональных баз данных и информационных справочных систем	<a href="https://cyberpedia.su/21x47c0.html">https://cyberpedia.su/21x47c0.html</a>
3	Каталог паттернов проектирования	<a href="https://refactoring.guru/ru/design-patterns/catalog">https://refactoring.guru/ru/design-patterns/catalog</a>

## 9. ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОВЗ

В таблице 9.1 указан перечень образовательных ресурсов, имеющих формы, адаптированные к ограничениям их здоровья, а также сведения о наличии специальных технических средств обучения коллективного и индивидуального пользования. При заполнении таблицы может быть использована информация, размещенная в подразделе «Доступная среда» специализированного раздела сайта НГТУ «Сведения об образовательной организации» <https://www.nntu.ru/sveden/accenv/>

Таблица 9.1 - Образовательные ресурсы для инвалидов и лиц с ОВЗ

№	Перечень образовательных ресурсов, приспособленных для использования инвалидами и лицами с ОВЗ	Сведения о наличии специальных технических средств обучения коллективного и индивидуального пользования
1	ЭБС «Консультант студента»	озвучка книг и увеличение шрифта
2	ЭБС «Лань»	специальное мобильное приложение - синтезатор речи, который воспроизводит тексты книг и меню навигации
3	ЭБС «Юрайт»	версия для слабовидящих

## 10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ, НЕОБХОДИМОЕ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Для контактной и самостоятельной работы обучающихся выделены помещения, оснащённые компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду организации:

- зал электронно-информационных ресурсов (ауд. 2210 – 11 компьютеров, ауд. 6119 – 9 компьютеров);
- читальный зал открытого доступа (ауд. 6162 – 2 компьютера);
- ауд. 2303, 2202, оборудованные Wi-Fi.

Перечень материально-технического обеспечения, необходимого для реализации программы бакалавриата и проведения лабораторных работ для студентов очного обучения, включает в себя:

1. Компьютерные классы НГТУ им. Р.Е.Алексеева (6 корпус НГТУ, аудитории 6342, 6339), оснащенные необходимым оборудованием, техническими и электронными средствами обучения и контроля знаний студентов (12 рабочих мест), оборудованных компьютерами:

- процессор: CPU IntelCore i3-2120 3.3 GHz;
- материнская плата: Asus p8h61-M LX2;
- оперативная память: 4 Gb (2\*2Gb) DDR 3;
- жесткий диск: 500 Gb.

с пакетами ПО общего назначения:

- Windows 7;
- Linux;
- Open Office.

Перечень материально-технического обеспечения, необходимого для реализации программы бакалавриата и проведения лабораторных работ для студентов очного обучения, включает в себя компьютерные классы

### 1. Ауд. 5412 кафедры «Вычислительные системы и технологии»,

Компьютеры оснащенные необходимым оборудованием, техническими и электронными средствами обучения и контроля знаний студентов. 6 рабочих мест, включающих моноблоки Lenovo S710 Intel Core i3-3240/4 Gb RAM, в составе локальной вычислительной сети, с подключением к сети Интернет.

Пакеты ПО (лицензионное): Лицензия Windows OEM (входила в поставку моноблоков)

Пакеты ПО (распространяемое по свободной лицензии):

- JDK 8 и выше (<https://adoptopenjdk.net/>);
- Фреймворк Java Spring 5(<https://spring.io/projects/spring-framework>)
- Eclipse (<https://www.eclipse.org/>)
- IntelliJ Idea (<https://www.jetbrains.com/ru-ru/idea/>)
- git (<https://git-scm.com/>)
- Maven (<https://maven.apache.org/>)

### 2. Ауд. 5422 кафедры «Вычислительные системы и технологии»,

Компьютеры оснащенные необходимым оборудованием, техническими и электронными средствами обучения и контроля знаний студентов. 5 рабочих мест, включающих персональные компьютеры Intel Core i5-9400/8 Gb RAM (5 шт.), в составе локальной вычислительной сети, с подключением к сети Интернет.

Пакеты ПО (распространяемое по свободной лицензии):

- Linux Ubuntu 20.04 (<https://releases.ubuntu.com/20.04/>)
- JDK 8 и выше (<https://adoptopenjdk.net/>);
- Фреймворк Java Spring 5(<https://spring.io/projects/spring-framework>)
- IntelliJ Idea (<https://www.jetbrains.com/ru-ru/idea/>)
- git (<https://git-scm.com/>)
- Maven (<https://maven.apache.org/>)

Также, для самостоятельной работы обучающихся выделены помещения, оснащённые компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду организации:

- аудитория 6543;
- аудитория 6545 (Проектор Acer – 1шт; ПК на базе IntelCoreDuo 2.93 ГГц, 2 Гб ОЗУ, 320 Гб HDD, монитор Samsung 19` – 11 шт. ПК подключены к сети «Интернет» и обеспечивают доступ в электронную информационно-образовательную среду университета).

## **11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

### **11.1 Общие методические рекомендации для обучающихся по освоению дисциплины, образовательные технологии**

Дисциплина реализуется посредством проведения контактной работы с обучающимися (включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

При преподавании дисциплины «Методы Data Mining», используются современные образовательные технологии, позволяющие повысить активность студентов при освоении материала курса и предоставить им возможность эффективно реализовать часы самостоятельной работы.

Весь лекционный материал курса может сопровождаться компьютерными презентациями, в которых наглядно преподносятся материал различных разделов курса и что дает возможность обсудить материал со студентами во время чтения лекций, активировать их деятельность при освоении материала. Электронные материалы лекций в период дистанционного обучения отправляются по электронной почте на адреса групп и могут быть получены до чтения лекций и проработаны студентами в ходе самостоятельной работы.

На лекциях, лабораторных занятиях реализуются интерактивные технологии, приветствуются вопросы и обсуждения, используется личностно-ориентированный подход, технология работы в малых группах, что позволяет студентам проявить себя, получить навыки самостоятельного изучения материала, выровнять уровень знаний в группе.

Все вопросы, возникшие при самостоятельной работе над домашним заданием, подробно разбираются на практических занятиях и лекциях. Проводятся индивидуальные и групповые консультации с использованием современных информационных технологий: электронная почта, мессенджеры, Zoom, Discord.

Иницируется активность студентов, поощряется задание любых вопросов по материалу, практикуется индивидуальный ответ на вопросы студента, рекомендуются методы успешного самостоятельного усвоения материала в зависимости от уровня его базовой подготовки.

Для оценки знаний, умений, навыков и уровня сформированности компетенции применяется балльно-рейтинговая система контроля и оценки успеваемости студентов в процессе текущего контроля.

Промежуточная аттестация проводится в форме зачета с учетом текущей успеваемости.

**Результат обучения считается сформированным на повышенном уровне**, если теоретическое содержание курса освоено полностью. При устных собеседованиях студент исчерпывающе, последовательно, четко и логически излагает учебный материал; свободно справляется с заданиями, вопросами, использует в ответе дополнительный материал. Все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты, проявляет самостоятельность при выполнении заданий.

**Результат обучения считается сформированным на пороговом уровне**, если теоретическое содержание курса освоено полностью. При устных собеседованиях студент последовательно, четко и логически излагает учебный материал; справляется с заданиями, вопросами, требующих применения знаний; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий.

**Результат обучения считается несформированным**, если студент при выполнении заданий не демонстрирует знаний учебного материала, допускает ошибки, неуверенно, с большими затруднениями выполняет задания, не демонстрирует необходимых умений, качество выполненных заданий не соответствует установленным требованиям, качество их выполнения оценено числом баллов ниже трех по оценочной системе, что соответствует допороговому уровню.

### **11.2 Методические указания для занятий лекционного типа**

Лекционный курс предполагает систематизированное изложение основных вопросов тематического плана. В ходе лекционных занятий раскрываются базовые вопросы в рамках каждой темы дисциплины (Таблица 5.2). Обозначаются ключевые аспекты тем, а также делаются акценты на наиболее сложные и важные положения изучаемого материала. Материалы лекций являются опорной основой для подготовки обучающихся к лабораторным работам и выполнения заданий самостоятельной работы, а также к мероприятиям текущего контроля успеваемости и промежуточной аттестации по дисциплине.

### **11.3 Методические указания по освоению дисциплины на занятиях семинарского типа**

Практические занятия представляют собой детализацию лекционного теоретического материала, проводятся в целях закрепления курса и охватывают все основные разделы. Основной формой проведения практических занятий является решение задач и разбор примеров и ситуаций в аудиторных условиях.

Практические занятия обучающихся обеспечивают:

- проверку и уточнение знаний, полученных на лекциях;
- получение умений и навыков составления докладов и сообщений, обсуждения вопросов по учебному материалу дисциплины;
- подведение итогов занятий по рейтинговой системе, согласно технологической карте дисциплины.

При оценивании практических занятий учитывается следующее:

- качество решения индивидуальных задач
- качество оформления решения;
- качество устных ответов на дополнительные вопросы.

### **11.4 Методические указания по самостоятельной работе обучающихся**

Самостоятельная работа обеспечивает подготовку обучающегося к аудиторным занятиям и мероприятиям текущего контроля и промежуточной аттестации по изучаемой дисциплине. Результаты этой подготовки проявляются в активности обучающегося на занятиях и в качестве выполненных практических заданий и других форм текущего контроля.

При выполнении заданий для самостоятельной работы рекомендуется проработка материалов лекций по каждой пройденной теме, а также изучение рекомендуемой литературы, представ-

ленной в Разделе 7.

В процессе самостоятельной работы при изучении дисциплины студенты могут работать на компьютере в специализированных аудиториях для самостоятельной работы, указанных в Разделе 10. В аудиториях имеется доступ через информационно-телекоммуникационную сеть «Интернет» к электронной информационно-образовательной среде университета (ЭИОС) и электронной библиотечной системе (ЭБС), где в электронном виде располагаются учебные и учебно-методические материалы, которые могут быть использованы для самостоятельной работы при изучении дисциплины.

### 11.5 Методические указания по освоению дисциплины на курсовой работе

Курсовая работа не предусмотрена учебным планом

## 12. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

### 12.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе текущего контроля успеваемости

Для текущего контроля знаний студентов по дисциплине проводится комплексная оценка знаний, включающая выполнение и защита лабораторных работ для студентов всех форм обучения. Зачет для студентов очной формы обучения в 7 семестре.

Типовые задания для лабораторных работ приведены в учебно-методических пособиях по проведению лабораторных работ.

Курсовая работа не предусмотрена учебным планом

Типовые вопросы для промежуточной аттестации в форме зачета для студентов всех форм обучения:

#### 12.1.1. Типовые задания для текущего контроля и промежуточной аттестации знаний обучающихся

С помощью метода градиентного спуска найти минимум функции. Построить график и отметить все пройденные точки в алгоритме.

Вариант	$F(x)$	$X_0$	step	eps
1	$f(x) = x^2 - 3x - 15$	-2	0.2,0.3,0.4,0.45	0.005
2	$f(x) = 2x^2 - 2x$	1	0.3,0.2,0.1,0.05	0.005
3	$f(x) = 2x - \sin(5x)$	7	0.5,0.6,0.7	0.05

Применить метод главных компонент к данным (задан dataset размерностью 3x10). Определить количество главных компонент для описания 70% данных.

#### Вариант №1.

№	Возраст, года	Рост, см	Зрение
1	3	92	-1
2	15	130	- 0,25
3	80	175	+5
4	54	180	- 0,89
5	2	80	-1,3
6	6	120	-5,4
7	20	164	+4
8	35	170	-3,2
9	49	182	-1,3
10	5	110	+3

## Вариант №2.

№	Возраст, года	Размер обуви	Вес, кг
1	3	26	17
2	15	37	50
3	80	42	75
4	54	38	82
5	2	18	12
6	6	30	25
7	20	40	55
8	35	41	68
9	49	45	72
10	5	28	19

Исследовать взаимосвязь в данных (Второй столбец – зависимая переменная).

1. Построить регрессионную модель  $y(\text{второй столб.}) = b_0 + b_1 * x(\text{первый столб.})$ . Визуализировать полученную модель.
2. Рассчитать коэффициент детерминации.
3. Сформулировать нулевую гипотезу, произвести анализ с помощью t-критерия. ( $H_0$  – никакой взаимосвязи в данных нет).
4. Произвести анализ остатков: построить scatter-plot и qq-plot остатков.
5. Предсказать значение по регрессионной модели.

<b>Вариант 1.</b> Исследовать взаимосвязь бедности и уровня образования. Предсказать значение бедности, основываясь на уровне образования = 95. «Уровень образования» - количество людей, имеющих высшее обр.		<b>Вариант 2.</b> Исследовать взаимосвязь процента людей, проживающих в столице и уровень образования. Предсказать процент людей, проживающих в столице, основываясь на уровне образования = 95.		<b>Вариант 3.</b> Исследовать взаимосвязь процента домохозяек в штате и уровня образования. Предсказать процент домохозяек, основываясь на уровне образования = 95.	
Уровень образ.(%)	Бедность (%)	Уровень образ.(%)	Столица (%)	Уровень образ.(%)	Домохоз. (%)
79,9	14,6	79,9	55,4	79,9	14,2
90,6	8,3	90,6	65,6	90,6	10,8
83,8	13,3	83,8	88,2	83,8	11,1
80,9	18	80,9	52,5	80,9	12,1
81,1	12,8	81,1	94,4	81,1	12,6
88,7	9,4	88,7	84,5	88,7	9,6
87,5	7,8	87,5	87,7	87,5	12,1
88,7	8,1	88,7	80,1	88,7	13,1
86	16,8	86	100	86	18,9
84,7	12,1	84,7	89,3	84,7	12
85,1	12,1	85,1	71,6	85,1	14,5
88,5	10,6	88,5	91,5	88,5	12,6

Построить отношение. Произвести заданные расчеты.

*Нечетные варианты – минимаксный подход и импликация KD. Четные варианты – вероятностный подход и импликация G.*



Вариант	Дано $A \rightarrow B$	Расчет
1	$\left(\frac{0,8}{a_1} + \frac{0,4}{a_2} + \frac{0,3}{a_3}\right) \rightarrow \left(\frac{0,6}{b_1} + \frac{0,9}{b_2} + \frac{0,7}{b_3}\right)$	$\mu_{A \rightarrow B} \left(\frac{0,6}{a_1} + \frac{0,3}{a_2} + \frac{0,5}{a_3}\right)$
2	$\left(\frac{0,3}{a_1} + \frac{0,1}{a_2} + \frac{0,9}{a_3}\right) \rightarrow \left(\frac{0,8}{b_1} + \frac{0,8}{b_2} + \frac{0,4}{b_3}\right)$	$\mu_{A \rightarrow B} \left(\frac{0,3}{a_1} + \frac{0,3}{a_2} + \frac{0,6}{a_3}\right)$
3	$\left(\frac{0,8}{a_1} + \frac{0,8}{a_2} + \frac{0,1}{a_3}\right) \rightarrow \left(\frac{0,8}{b_1} + \frac{0,5}{b_2} + \frac{0,5}{b_3}\right)$	$\mu_{A \rightarrow B} \left(\frac{0,9}{a_1} + \frac{0,6}{a_2} + \frac{0,2}{a_3}\right)$

Произвести кластеризацию данных fcm-алгоритмом. Евклидова норма,  $q = 2$ . Произвести вычисления одной итерации, т.е. получить матрицы  $U_0$  и  $U_1$ . Построить график с данными, а также центрами кластеров (как изначальными, так и новыми вычисленными). Посчитать критерий останова.

Вариант	Данные	Кластеры(0)
1	$(-7;7), (-8;5), (3;3), (10;2), (6;7)$	$(-2;0), (4;10)$
2	$(0;3), (-1;2), (2;1), (7;7), (9;5)$	$(3;3), (6;2)$
3	$(10;10), (5;5), (7;5), (-1;3), (-2;5)$	$(4;8), (0;2)$

### 12.1.2. Перечень вопросов, выносимых на промежуточную аттестацию (зачет):

1. Что такое data mining
2. Градиентный спуск
3. Генетический алгоритм
4. Алгоритм имитации отжига
5. Постановка задачи снижения размерности признакового пространства
6. Необходимость и возможность снижения размерности признакового пространства
7. Z – преобразование
8. Основные числовые характеристики главных компонент
9. Этапы и реализации метода главных компонент.
10. Связь метода главных компонент с собственными числами и собственными векторами
11. Оценка важности признаков
12. Определение линейной регрессии
13. Метод наименьших квадратов
14. Линейный коэффициент корреляции
15. Проверка гипотезы о значимости коэффициентов регрессии
16. Множественная линейная регрессия
17. Коэффициент детерминации
18. Регуляризация
19. Робастные регрессионные модели
20. Определение логистической регрессии
21. Преобразование уравнения линейной регрессии
22. Однослойный перцептрон
23. Наивное обучение и интерпретация
24. Многослойный перцептрон
25. Метод обратного распространения ошибки
26. Борьба с переобучением в нейронных сетях
27. Карты признаков
28. Операция свертки
29. Операция пулинга
30. Обратное распространение ошибки в CNN
31. Архитектуры CNN
32. Нечеткие множества. Основные определения
33. Лингвистическая переменная

34. Нечеткая логика
35. Квантификаторы и их смысл
36. Нечеткие отношения и отображения
37. Нечеткие правила вывода
38. Алгоритмы нечеткой кластеризации
39. FCM-алгоритм
40. Алгоритмы поиска ассоциативных правил

Полный комплект оценочных средств является неотъемлемой частью ФОС и хранится на кафедре «Вычислительные системы и технологии».

УТВЕРЖДАЮ:  
Директор института ИРИТ

“\_\_\_” \_\_\_\_\_ 20\_\_ г.

**Лист актуализации рабочей программы дисциплины**  
**« Б1.В.ОД.6 Методы Data Mining »**  
индекс по учебному плану, наименование

для подготовки **бакалавров**/ специалистов/ магистров

Направление: {шифр – название} 09.03.01 Информатика и вычислительная техника

Направленность: Программное обеспечение средств вычислительной техники и автоматизированных систем

Форма обучения очная

Год начала подготовки: 2020, 2021

Курс 4

Семестр 7

В рабочую программу не вносятся изменения. Программа актуализирована для 2021 г. начала подготовки.

Разработчик (и): Суркова А.С., д.т.н., доцент  
(ФИО, ученая степень, ученое звание)

«\_\_» \_\_\_\_\_ 20\_\_ г.

Рабочая программа пересмотрена и одобрена на заседании кафедры ВСТ  
\_\_\_\_\_ протокол № \_\_\_\_\_ от «\_\_» \_\_\_\_\_ 20\_\_ г.

Заведующий кафедрой \_\_\_\_\_

**Лист актуализации принят на хранение:**

Заведующий выпускающей кафедрой ВСТ \_\_\_\_\_ «\_\_» \_\_\_\_\_ 20\_\_ г.

Методический отдел УМУ: \_\_\_\_\_ «\_\_» \_\_\_\_\_ 20\_\_ г.

\_\_\_\_\_