

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Клочков Юрий Сергеевич  
Должность: и.о. ректора  
Дата подписания: 21.05.2025 15:31:40  
Уникальный программный ключ:  
4e7c4ea90328ec8e65c5d8058549a2538d7400d1

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ  
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное  
образовательное учреждение высшего образования

**«ТЮМЕНСКИЙ ИНДУСТРИАЛЬНЫЙ УНИВЕРСИТЕТ»**

**УТВЕРЖДАЮ**

Заведующий кафедрой ИСТ  
\_\_\_\_\_ Данилов О. Ф.

« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

**РАБОЧАЯ ПРОГРАММА**

дисциплины: **Анализ данных и машинное обучение**

направление подготовки: **09.03.04 Программная инженерия**

направленность (профиль): **Разработка программно-информационных систем**

форма обучения: **очная**

Рабочая программа рассмотрена на заседании кафедры интеллектуальных систем и технологий для направления 09.03.04 Программная инженерия направленность (профиль) «Разработка программно-информационных систем»

### 1. Цели и задачи освоения дисциплины

- формирование компетенций у обучающихся в области машинного или статистического обучения;
- формирование знаний и умений в области методов машинного или статистического обучения;
- получение навыков, позволяющих использовать методы машинного обучения или статистического обучения при решении широкого спектра задач систем искусственного интеллекта.

### 2. Место дисциплины в структуре ОПОП ВО

Дисциплина относится к дисциплинам части учебного плана формируемого участниками образовательных отношений образовательной программы.

Необходимыми условиями для освоения дисциплины являются:

- знание основ обработки больших наборов данных в памяти электронной машины, методов хранения, обработки и передачи электронной информации;
- умение решать стандартные профессиональные задачи с применением методов математического анализа и моделирования;
- владение навыками программирования, отладки и тестирования прототипов программно-технических комплексов задач.

Содержание дисциплины является логическим продолжением содержания дисциплин:

Программирование

Базы данных

Математический анализ

Алгоритмы и структуры данных

и служит основой для освоения дисциплин:

Преддипломная практика и написание ВКР

### 3. Результаты обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих компетенций:

Таблица 3.1

Код и наименование компетенции	Код и наименование индикатора достижения компетенции (ИДК)	Код и наименование результата обучения по дисциплине
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Осуществляет выбор актуальных российских и зарубежных источников, а так же поиск, сбор и обработку информации, необходимой для решения поставленной задачи.	Знать (31) базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы
		Знать (32) методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников

<p>УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</p>	<p>УК-1.1 Осуществляет выбор актуальных российских и зарубежных источников, а так же поиск, сбор и обработку информации, необходимой для решения поставленной задачи.</p>	<p>Уметь (У1) Осуществлять поиск, сбор и обработку информации, необходимой для анализа данных</p>
<p>ПКС 3 – Способность выполнять работы по обеспечению функционирования баз данных и обеспечению их информационной безопасности</p>	<p>ПКС-3.1 Выявляет угрозы безопасности данных и решает задачи администрирования данных.</p>	<p>Знать (З3) Способы защиты данных при проведении анализа и обработки</p> <p>Знать (З4) Способы обновления, преобразования и передачи данных</p> <p>Уметь (У2) Защищать данные при обработке и анализе данных</p> <p>Уметь (У3) Обновлять, преобразовывать и передавать данные</p> <p>Владеть (В2) Навыками защиты данных при обработке и анализе</p> <p>Владеть (В3) Навыками обновления, преобразования и передачи данных</p>

#### 4. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 зачетных единиц, 108 акад. часов.

Таблица 4.1

Курс	Аудиторные занятия/контактная работа, час.			Самостоятельная работа, час.	Контроль, час.	Форма промежуточной аттестации
	Лекции	Практические занятия	Лабораторные занятия			
3	16	32	-	60	-	Зачёт

#### 5. Структура и содержание дисциплины

##### 5.1. Структура дисциплины.

Структура дисциплины	Аудиторные занятия, час.			СРС, час.	Всего, час.	Код ИДК	Оценочные средства
	Л.	Пр.	Лаб.				
1. Введение в курс							
1.1 Введение в курс	1	2	-	4	7	31, 32, У1, В1	Опрос Защита практического задания
Итого по разделу	1	2	-	4	7		
2. Этапы жизненного цикла решения							
2.1 Этапы жизненного цикла решения	1	3	-	6	10	34, 33, У3, В3, У2, В2	Защита практического задания
Итого по разделу	1	3	-	6	10		
3. Подготовка данных							
3.1 Подготовка данных	2	4	-	7	13	34, У3, В3	Опрос Защита практического задания
Итого по разделу	2	4	-	7	13		
4. Информационное обучение							
4.1 Информационное обучение	2	4	-	7	13	34, У3, В3	Защита практического задания
Итого по разделу	2	4	-	7	13		
5. Обучение на основе сходства							
5.1 Обучение на основе сходства	2	4	-	7	13	34, У3, В3	Опрос защита практического задания
Итого по разделу	2	4	-	7	13		
6. Вероятностное обучение							
6.1 Вероятностное обучение	2	4	-	7	13	34, У3, В3	Защита практического задания
Итого по разделу	2	4	-	7	13		
7. Обучение на основе ошибок							
7.1 Обучение на основе	2	4	-	7	13		Опрос

ошибок			-			34, У3, В3	Защита практического задания
Итого по разделу	2	4	-	7	13		
8. Оценка моделей							
8.1 Оценка моделей	2	4	-	7	13	34, У3, В3	Опрос Защита практического задания
Итого по разделу	2	4	-	7	13		
9. Перспективы							
9.1 Перспективы	2	3	-	8	13	31, 33, 34, 32, У1, В1, У2, У3, В2, В3	Защита практического задания
Итого по разделу	2	3	-	8	13		
Зачет				4	4		Вопросы к зачёту
Итого по дисциплине	16	32	-	60	108		

## 5.2. Содержание дисциплины

### 1. Введение в курс

#### 1.1 Введение в курс

Прогноз. Машинное обучение и основные принципы работы. Жизненный цикл. Основные инструменты. Перспективы развития машинного обучения.

### 2. Этапы жизненного цикла решения.

#### 2.1 Этапы жизненного цикла решения.

Преобразование проблемы в аналитическую модель. Оценка возможности реализации. Подготовка данных. Проектирование и реализация.

### 3. Подготовка данных

#### 3.1 Подготовка данных

Качество данных. Обзор. Определение проблем. Обработка пропущенных значений и выбросов. Визуализация отношений между признаками. Ковариация и корреляция.

Нормализация. Статистическое группирование. Формирование выборки.

### 4. Информационное обучение

#### 4.1 Информационное обучение

Основы. Деревья решений. Модель энтропии Шеннона. Алгоритм ID3.

Обобщения.

Альтернативный выбор признаков и показатели неоднородности. Обработка непрерывных

целевых признаков. Прогнозирование непрерывных целевых признаков.

Усечение деревьев.

Ансамбль моделей.

5. Обучение на основе сходства.

### 5.1 Обучение на основе сходства.

Основы. Пространство признаков. Измерение сходства с помощью расстояния.

Алгоритм ближайшего соседа. Обработка зашумленных данных. Поиск.

Нормализация данных.

Прогнозирование непрерывных целевых признаков. Меры сходства. Отбор признаков.

### 6. Вероятностное обучение.

#### 6.1 Вероятностное обучение.

Теорема Байеса. Байесовское прогнозирование. Условия факторизации. Наивная Байесовская модель. Сглаживание. Функция плотности. Группирование.

Байесовские сети.

### 7. Обучение на основе ошибок

#### 7.1 Обучение на основе ошибок.

Линейная регрессия. Множественная линейная регрессия. Градиентный спуск.

Скорость обучения. Интерпретация моделей. Определение скорости обучения.

Обработка категориальных признаков. Моделирование нелинейных зависимостей.

Многоклассовая логистическая регрессия.

### 8. Оценка моделей

#### 8.1 Оценка моделей

Оценочный эксперимент. Показатели эффективности. Статистика Колмогорова-Смирнова. Оценка моделей после внедрения.

### 9. Перспективы

#### 9.1 Перспективы

Перспективы развития методов машинного обучения. Выбор метода машинного обучения.

## 5.2.2. Содержание дисциплины по видам учебных занятий.

### Лекционные занятия

Номер раздела дисциплины	Объем, час.	Тема лекционного занятия
1. Введение в курс	1	Введение в курс
2. Этапы жизненного цикла решения	1	Этапы жизненного цикла решения
3. Подготовка данных	2	Подготовка данных
4. Информационное обучение	2	Информационное обучение
5. Обучение на основе сходства	2	Обучение на основе сходства
6. Вероятностное обучение	2	Вероятностное обучение
7. Обучение на основе ошибок	2	Обучение на основе ошибок
8. Оценка моделей	2	Оценка моделей
9. Перспективы	2	Перспективы
Итого	16	

### Практические занятия

Номер раздела дисциплины	Объем, час.	Тема практического занятия
1. Введение в курс	2	Введение в курс

2. Этапы жизненного цикла решения	3	Этапы жизненного цикла решения
3. Подготовка данных	4	Подготовка данных
4. Информационное обучение	4	Информационное обучение
5. Обучение на основе сходства	4	Обучение на основе сходства
6. Вероятностное обучение	4	Вероятностное обучение
7. Обучение на основе ошибок	4	Обучение на основе ошибок
8. Оценка моделей	4	Оценка моделей
9. Перспективы	3	Перспективы
Итого	32	

### **Самостоятельная работа студента**

Номер раздела дисциплины	Объем, час.	Тема	Вид СРС
1. Введение в курс	4	Введение в курс	
2. Этапы жизненного цикла решения	6	Этапы жизненного цикла решения	
3. Подготовка данных	7	Подготовка данных	
4. Информационное обучение	7	Информационное обучение	
5. Обучение на основе сходства	7	Обучение на основе сходства	
6. Вероятностное обучение	7	Вероятностное обучение	
7. Обучение на основе ошибок	7	Обучение на основе ошибок	
8. Оценка моделей	7	Оценка моделей	
9. Перспективы	4	Перспективы	
Итого	60		

5.2.3. Преподавание дисциплины ведется с применением следующих видов образовательных технологий:

- визуализация учебного материала в PowerPoint в диалоговом режиме (лекционные занятия);
- работа на компьютерах (практические занятия).

### **6. Тематика курсовых работ/проектов**

не предусмотрено

### **7. Контрольные работы**

не предусмотрено

### **8. Оценка результатов освоения дисциплины**

8.1. Критерии оценивания степени полноты и качества освоения компетенций в соответствии с планируемыми результатами обучения приведены в Приложении 1.

8.2. Рейтинговая система оценивания степени полноты и качества освоения компетенций обучающихся представлена ниже.

№ п/п	Виды мероприятий в рамках текущего контроля	Количество баллов
1 текущая аттестация		
1	Опрос на лекции, работа на лекции	15
2	Защита практических заданий	15
Итого:		30
2 текущая аттестация		
1	Опрос на лекции, работа на лекции	15
2	Защита практических заданий	15
Итого:		30
3 текущая аттестация		
1	Опрос на лекции, работа на лекции	25
2	Защита практических заданий	15
Итого:		40
ВСЕГО:		100

### 9. Учебно-методическое и информационное обеспечение дисциплины

9.1. Перечень рекомендуемой литературы представлен в Приложении 2.

- 9.2. Современные профессиональные базы данных и информационные справочные системы
- Электронный каталог/Электронная библиотека ТИУ <http://webirbis.tsogu.ru/>;
  - Цифровой образовательный ресурс – библиотечная система IPR SMART — <https://www.iprbookshop.ru/>;
  - Электронно-библиотечная система «Консультант студента» [www.studentlibrary.ru](http://www.studentlibrary.ru/);
  - Электронно-библиотечная система «ЛАНЬ» [https://e.lanbook.com](https://e.lanbook.com/);
  - Образовательная платформа ЮРАЙТ [www.urait.ru](http://www.urait.ru/);
  - Научная электронная библиотека ELIBRARY.RU [http://www.elibrary.ru](http://www.elibrary.ru/);
- Библиотеки нефтяных вузов России:
- Электронная нефтегазовая библиотека РГУ нефти и газа им. Губкина <http://elib.gubkin.ru/>;
  - Электронная библиотека Уфимского государственного нефтяного технического университета <http://bibl.rusoil.net/>;
  - Библиотечно-информационный комплекс Ухтинского государственного технического университета УГТУ <http://lib.ugtu.net/books>
  - Электронная справочная система нормативно-технической документации «Технорматив»;
  - ЭКБСОН – информационная система доступа к электронным каталогам библиотек сферы образования и науки.
- Лицензионное и свободно распространяемое программное обеспечение, в т.ч. отечественного производства:
1. Microsoft Windows – операционная система.
  2. Microsoft Office Professional Plus – набор офисных приложений.
  3. Deductor Academic (Бесплатная некоммерческая версия) – инструмент для анализа данных.
  4. GPSS Studio Student (Бесплатная студенческая версия) – программное обеспечение для моделирования систем.
  5. Orange (Свободно-распространяемое ПО) – инструмент для анализа данных и визуализации.

## 10. Материально-техническое обеспечение дисциплины

Помещения для проведения всех видов работы, предусмотренных учебным планом, укомплектованы необходимым оборудованием и техническими средствами обучения.

Таблица 10.1

№ п/п	Наименование учебных предметов, курсов, дисциплин (модулей), практики, иных видов учебной деятельности, предусмотренных учебным планом образовательной программы	Наименование помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий	Адрес (местоположение) помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом (в случае реализации образовательной программы в сетевой форме дополнительно указывается наименование организации, с которой заключен договор)
1	2	3	4
1	Анализ данных и машинное обучение	Учебная аудитория для проведения занятий лекционного типа; групповых и индивидуальных консультаций; текущего контроля и промежуточной аттестации. Основное оборудование: столы – 52 шт., стулья – 52 шт, доска аудиторная – 1 шт., моноблок – 1 шт., проектор – 1 шт., проекционный экран – 1 шт., акустическая система (колонки) -2 шт., микрофон - 1 шт.	625001, Тюменская область, г. Тюмень, ул. Луначарского, д.4
		Учебная аудитория для проведения занятий семинарского типа (практические занятия); групповых и индивидуальных консультаций; текущего контроля и промежуточной аттестации Основное оборудование: столы – 25 шт., стулья – 57 шт, доска аудиторная – 1 шт., моноблок – 16 шт., проектор – 1 шт., проекционный экран – 1 шт., акустическая система (колонки) -2 шт., микрофон - 1 шт.	625001, Тюменская область, г. Тюмень, ул. Луначарского, д.4
		Помещение для самостоятельной работы обучающихся с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Учебная мебель: столы – 10 шт., стулья – 15 шт., доска аудиторная – 1 шт., моноблок – 5 шт.,	625001, Тюменская область, г. Тюмень, ул. Луначарского, д.2, корп.1
		Помещение для самостоятельной работы обучающихся с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Учебная мебель: столы – 15 шт., стулья – 25 шт., доска аудиторная – 1 шт., моноблок – 5 шт., проектор - 1 шт., экран - 1 шт.,	625001, Тюменская область, г. Тюмень, ул. Луначарского, д.2, корп.1

## **11. Методические указания по организации СРС**

Самостоятельная работа является одной из важнейших форм изучения любой дисциплины. Она позволяет систематизировать и углубить теоретические знания, закрепить умения и навыки, способствует развитию умений пользоваться научной и учебно-методической литературой. Познавательная деятельность в процессе самостоятельной работы требует от обучающегося высокого уровня активности и самоорганизованности.

Самостоятельная работа включает в себя работу с конспектом лекций, изучение и конспектирование рекомендуемой литературы, подготовка мультимедиа-сообщений/докладов, подготовка реферата, тестирование, решение задач и упражнений по образцу, решение вариативных задач, выполнение чертежей, схем, расчетов (графических работ), решение ситуационных (профессиональных) задач, подготовка к деловым играм, проектирование и моделирование разных видов и компонентов профессиональной деятельности, научно-исследовательскую работу и др.

Контроль результатов внеаудиторной самостоятельной работы обучающихся осуществляется в пределах времени, отведенного на обязательные учебные занятия по дисциплине и внеаудиторную самостоятельную работу обучающихся по дисциплине, может проходить в письменной, устной или смешанной форме.

Самостоятельная работа обучающегося без преподавателя включает в себя подготовку к различным видам контрольных испытаний, подготовку и написание самостоятельных видов работ.

## Планируемые результаты обучения для формирования компетенции и критерии их оценивания

Дисциплина **Анализ данных и машинное обучение**

Код, направление подготовки **09.03.04 Программная инженерия**

Направленность (профиль) **Разработка программно-информационных систем**

Код компетенции	Код и наименование результата обучения по дисциплине (модулю)	Критерии оценивания результатов обучения			
		1-2	3	4	5
УК-1	Знать (31) базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы	Не знает базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы	Частично знает базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы и допускает ряд ошибок	Знает базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы, но допускает неточности	Знает базы данных рецензируемой научной литературы, необходимых для поиска актуальной российской и зарубежной литературы
	Знать (32) методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников	Не знает методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников	Частично знает методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников и допускает ряд ошибок	Знает методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников, но допускает неточности	Знает методы поиска, сбора и обработки информации, полученной из актуальных российских и зарубежных источников
	Уметь (У1) Осуществлять поиск, сбор и обработку информации, необходимой для анализа данных	Не умеет осуществлять поиск, сбор и обработку информации, необходимой для анализа данных	Частично умеет осуществлять поиск, сбор и обработку информации, необходимой для анализа данных и допускает ошибки	Умеет осуществлять поиск, сбор и обработку информации, необходимой для анализа данных, но допускает неточности	Умеет осуществлять поиск, сбор и обработку информации, необходимой для анализа данных

УК-1	Владеть (В1) Навыками поиска, сбора и обработки информации, необходимой для анализа данных	Не владеет навыками поиска, сбора и обработки информации, необходимой для анализа данных	Частично владеет навыками поиска, сбора и обработки информации, необходимой для анализа данных и допускает ошибки	Владеет навыками поиска, сбора и обработки информации, необходимой для анализа данных, но допускает ряд неточностей	Владеет навыками поиска, сбора и обработки информации, необходимой для анализа данных
ПКС-3	Знать (33) Способы защиты данных при проведении анализа и обработки	Не знает способы защиты данных при проведении анализа и обработки	Частично знает способы защиты данных при проведении анализа и обработки и допускает ошибки	Знает способы защиты данных при проведении анализа и обработки, но допускает неточности	Знает способы защиты данных при проведении анализа и обработки
	Знать (34) Способы обновления, преобразования и передачи данных	Не знает способы обновления, преобразования и передачи данных	Частично знает способы обновления, преобразования и передачи данных и допускает ошибки	Знает способы обновления, преобразования и передачи данных, но допускает неточности	Знает способы обновления, преобразования и передачи данных
	Уметь (У2) Защищать данные при обработке и анализе данных	Не умеет защищать данные при обработке и анализе данных	Частично умеет защищать данные при обработке и анализе данных и допускает ошибки	Умеет защищать данные при обработке и анализе данных, но допускает неточности	Умеет защищать данные при обработке и анализе данных
	Уметь (У3) Обновлять, преобразовывать и передавать данные	Не умеет обновлять, преобразовывать и передавать данные	Частично умеет обновлять, преобразовывать и передавать данные и допускает ошибки	Умеет обновлять, преобразовывать и передавать данные, но допускает неточности	Умеет обновлять, преобразовывать и передавать данные
	Владеть (В2) Навыками защиты данных при обработке и анализе	Не владеет навыками защиты данных при обработке и анализе	Частично владеет навыками защиты данных при обработке и анализе и допускает ошибки	Владеет навыками защиты данных при обработке и анализе, но допускает неточности	Владеет навыками защиты данных при обработке и анализе

ПКС-3	Владеть (В3) Навыками обновления, преобразования и передачи данных	Не владеет навыками обновления, преобразования и передачи данных	Частично владеет навыками обновления, преобразования и передачи данных и допускает ошибки	Владеет навыками обновления, преобразования и передачи данных, но допускает неточности	Владеет навыками обновления, преобразования и передачи данных
-------	--	--	---	--	---

**КАРТА**  
**обеспеченности дисциплины учебной и учебно-методической**  
**литературой**

Дисциплина **Анализ данных и машинное обучение**

Код, направление подготовки **09.03.04 Программная инженерия**

Направленность (профиль) **Разработка программно-информационных систем**

№ п/п	Название учебного, учебно-методического издания, автор, издательство, вид издания, год издания	Количество экземпляров в БИК	Контингент обучающихся, использующих указанную	Обеспеченность обучающихся литературой, %	Наличие электронного варианта в ЭБС (+/-)
1	Воронова, Л. И. Machine Learning: регрессионные методы интеллектуального анализа данных : учебное пособие / Л. И. Воронова, В. И. Воронов. — Москва : Московский технический университет связи и информатики, 2018. — 82 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <a href="https://www.iprbookshop.ru/81325.html">https://www.iprbookshop.ru/81325.html</a>	ЭР*	30	100	+
2	Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук ; перевод с английского А. Б. Огурцова. — Москва : ДМК Пресс, 2018. — 250 с. — ISBN 978-5-97060-508-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/97353">https://e.lanbook.com/book/97353</a>	ЭР*	30	100	+
3	Коэльо, Л. П. Построение систем машинного обучения на языке Python / Л. П. Коэльо, В. Ричарт ; перевод с английского А. А. Слинкин. — 2-е изд. — Москва : ДМК Пресс, 2016. — 302 с. — ISBN 978-5-97060-330-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/82818">https://e.lanbook.com/book/82818</a>	ЭР*	30	100	+

ЭР\* – электронный ресурс для авторизованных пользователей доступен через Электронный каталог/Электронную библиотеку ТИУ <http://webirbis.tsogu.ru/>