



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное учреждение  
высшего образования  
**«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
(ФГБОУ ВО «ИГУ»)**

Институт математики и информационных технологий  
Кафедра алгебраических и информационных систем



**Рабочая программа дисциплины (модуля)**

**Б1.В.03 Интеллектуальный анализ данных**

Направление подготовки	09.03.03 Прикладная информатика
Направленность (профиль) подготовки информационных систем	Проектирование и разработка информационных систем
Квалификация выпускника	бакалавр
Форма обучения	очная

Иркутск 2025 г.

## 1. Цели и задачи дисциплины

### Цель

Целью курса является изучение современных подходов, моделей, алгоритмов анализа данных и решения задач распознавания, классификации, нахождения зависимостей.

### Задачи:

Задачи дисциплины: изучение сущности и роли анализа информации; ознакомление бакалавров с основными средствами и методами анализа данных при исследовании сложных систем; знакомство с основными технологиями анализа данных; формирование умений и навыков в области поддержки принятия решений в организации с применением современных методов и средств анализа данных.

## 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина (модуль) относится к части программы, формируемой участниками образовательных отношений, и изучается на третьем курсе.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы знания, умения и навыки, сформированные дисциплинами «Математика», «Алгебра», «Программирование».

2.3. Знания, умения и навыки, формируемые в процессе освоения дисциплины, используются при выполнении научно-исследовательских работ, для решения прикладных задач и выполнения выпускной квалификационной работы.

## 3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций (элементов следующих компетенций) в соответствии с ФГОС ВО по соответствующему направлению подготовки.

### Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-3 Способность демонстрации общенаучных базовых знаний математических и естественных наук и информационных технологий; способность применять в профессиональной деятельности современные языки программирования, операционные системы, электронные библиотеки и пакеты программ, сетевые технологии.	ИДК ПК3.1 Способен использовать математические знания в профессиональной деятельности	Знает современные проблемы анализа данных, теории распознавания, классификации, поиска зависимостей; методы и подходы решения практических задач анализа данных и классификации коллективами алгоритмов; программные средства решения основных задач анализа данных и классификации.
	ИДК ПК3.2 Способен использовать теоретические принципы информационных технологий в	Умеет пользоваться своими знаниями для решения фундаментальных, прикладных и технологических задач в различных предметных областях; эффективно

	<p>профессиональной деятельности</p>	<p>использовать информационные технологии и компьютерную технику для достижения необходимых теоретических и прикладных результатов.</p>
<p>ПК-4 Готовность к включению профессиональное сообщество; способность проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности</p>	<p>ИДК ПК4.3 Способен проводить под научным руководством исследования на основе существующих методов в некоторой области профессиональной деятельности</p>	<p>Знает методы и подходы решения практических задач анализа данных и классификации композициями алгоритмов.</p> <p>Умеет пользоваться своими знаниями для решения фундаментальных, прикладных и технологических задач в различных предметных областях; делать правильные выводы из сопоставления результатов теории и эксперимента, выбирать правильно параметры методов, адекватные размерности обучающих выборок.</p> <p>Владеет навыками анализа большого объема частично противоречивых и неполных признаков описаний; навыками грамотной обработки статистических многомерных данных, оформления результатов численных расчетов и их сопоставления с теоретическими оценками; практикой исследования и решения теоретических и прикладных задач; навыками анализа реальных задач из различных предметных областей на уровне отдельных подходов и коллективами алгоритмов.</p>

#### 4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Объем дисциплины составляет 5 зачетных единиц, 180 часов, в том числе 35 часов на контроль, практическая подготовка \_\_\_\_\_.

Форма промежуточной аттестации: 6 семестр - экзамен.

##### 4.1. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ, С УКАЗАНИЕМ ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ И ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ

№ п/п	Раздел дисциплины/темы	Се мес тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоятельная работа + контроль	
			Лекции	Семинарские (практические занятия)	Контроль обучения		
1	Основные понятия анализа данных.		2				Проверка отчетов по лабораторным работам
2	Линейная регрессия.		6	8	1	15	Проверка отчетов по лабораторным работам
3	Логистическая регрессия.		2	2	1	15	Проверка отчетов по лабораторным работам

№ п/п	Раздел дисциплины/темы	Се мес тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоя тельная работа + контроль	
			Лекции	Семинарские (практические занятия)	Контроль обучения		
4	Регуляризация.		2	4	1	15	Проверка отчетов по лабораторным работам
5	Метрики классификации.		2		1		Проверка отчетов по лабораторным работам
6	Метод k ближайших соседей.		2	2	1		Проверка отчетов по лабораторным работам
7	Метод опорных векторов.		2	2	1		Проверка отчетов по лабораторным работам
8	Деревья решений.		2	2	1		Проверка отчетов по лабораторным работам
9	Композиции деревьев.		4	2	1		Проверка отчетов по

№ п/п	Раздел дисциплины/темы	Се мес тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоя тельная работа + контроль	
			Лекции	Семинарские (практические занятия)	Контроль обучения		
							лабораторным работам
10	Рекомендательные системы.		4	4	1	15	Проверка отчетов по лабораторным работам
11	Нейронные сети.		4	6	1	11	Проверка отчетов по лабораторным работам
12	Подготовка к экзамену.					35	
<b>Итого часов</b>			32	32	10	106	

#### 4.2. План внеаудиторной самостоятельной работы обучающихся по дисциплине

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Затраты времени (час.)		
6	Линейная регрессия.	Выполнение практической работы	01.03	15	Домашняя работа	Материалы курса на платформе ИОС DOMIC
6	Логистическая регрессия.	Выполнение практической работы	01.04	15	Домашняя работа	Материалы курса на платформе ИОС DOMIC
6	Регуляризация.	Выполнение практической работы	01.04	15	Домашняя работа	Материалы курса на платформе ИОС DOMIC
6	Рекомендательные системы.	Выполнение практической работы	01.05	15	Домашняя работа	Материалы курса на платформе ИОС DOMIC
6	Нейронные сети.	Выполнение практической работы	01.06	11	Домашняя работа	Материалы курса на платформе ИОС DOMIC
6	Подготовка к экзамену.		14.06	35		
Общая трудоемкость самостоятельной работы по дисциплине (час)				106		
Из них объем самостоятельной работы с использованием электронного обучения и дистанционных образовательных технологий (час)				106		

#### 4.3. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

##### 1. Основные понятия анализа данных.

Введение. Задачи анализа данных. Основные понятия анализа данных. Регрессия и классификация. Основы Python.

2. Линейная регрессия.  
Линейная регрессия. Гипотеза и функция потерь. Градиентный спуск. Визуализация. Библиотека Matplotlib.
3. Логистическая регрессия.  
Задача классификации. Логистическая регрессия. Гипотеза и функция потерь. Библиотеки Numpy, Pandas.
4. Регуляризация.  
Смещение моделей машинного обучения. Разброс моделей машинного обучения. Компромисс между смещением и разбросом. Недообучение и переобучение. Регуляризация.
5. Метрики классификации.  
Точность, полнота, F-мера. AUC-ROC.
6. Метод k ближайших соседей.  
Метод k ближайших соседей. Основные параметры. Метод k ближайших соседей в scikit-learn.
7. Метод опорных векторов.  
Метод опорных векторов. Основные параметры. Метод опорных векторов в scikit-learn.
8. Деревья решений.  
Деревья решений. Основные параметры. Деревья решений в scikit-learn.
9. Композиции деревьев.  
Композиции деревьев решений. Бэггинг. Случайный лес.
10. Рекомендательные системы.  
Задачи и примеры рекомендательных систем. Рекомендательные системы на основе контента. Ассоциативные правила. Совместная фильтрация. Рекомендательные системы второго поколения. Контекстно-зависимые рекомендательные системы.
11. Нейронные сети.  
Модель нейрона. Персептрон. Линейный нейрон. Функции активации. Метод обратного распространения ошибки. Библиотека TensorFlow. Распознавание образов. Сверточные нейронные сети. Рекуррентные нейронные сети.

#### 4.3.1. Перечень семинарских, практических занятий и лабораторных работ

№ п/п	№ раздела и темы	Наименование семинаров, практических и лабораторных работ	Трудоемкость (час.)		Оценочные средства	Формируемые компетенции (индикаторы)*
			Всего часов	Из них практическая подготовка		
1	2	3	4	5	6	7



1	2	Линейная регрессия с одной переменной.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
2	2	Визуализация.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
3	2	Линейная регрессия в scikit-learn.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
4	2	Линейная регрессия с несколькими переменными.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
5	2	Линейная регрессия с несколькими переменными в scikit-learn.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
6	3	Логистическая регрессия.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
7	4	Логистическая регрессия с регуляризацией в scikit-learn.	4	4	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
8	6	Метод k ближайших соседей.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
9	6	Подбор параметров для метода kNN.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
10	7	Метод опорных векторов.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
11	8	Деревья решений.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)

12	8	Визуализация деревьев решений.	1	1	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
13	9	Случайный лес.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
14	10	Поиск ассоциативных правил.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
15	10	Метод коллаборативной фильтрации.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
16	11	Алгоритм обучения персептрона.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
17	11	TensorFlow playground.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
18	11	Полносвязная сеть в TensorFlow.	2	2	Отчет по лабораторной работе	ПК-3 (ИДК ПК3.1, ИДК ПК3.2), ПК-4 (ИДК ПК4.3)
		<b>Всего</b>	32	32		

#### 4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение студентами в рамках самостоятельной работы (СР)

Тема	Задание	Формируемые компетенции
Линейная регрессия.	Решить задачи для самостоятельного выполнения	ПК-3, ПК-4
Логистическая регрессия.	Решить задачи для самостоятельного выполнения	ПК-3, ПК-4
Регуляризация.	Решить задачи для самостоятельного выполнения	ПК-3, ПК-4
Рекомендательные системы.	Решить задачи для самостоятельного выполнения	ПК-3, ПК-4
Нейронные сети.	Решить задачи для самостоятельного выполнения	ПК-3, ПК-4

#### 4.4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Методические рекомендации по организации самостоятельной работы студентов доступны на странице курса в ИОС Домик.

#### **4.5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ РАБОТ (ПРОЕКТОВ)**

Не предусмотрено.

## **5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

### **а) основная литература**

1. Грас, Джоэл. Data Science. Наука о данных с нуля [Текст] : учеб. пособие / Д. Грас. - СПб. : БХВ - Петербург, 2019. - 336 с. ; 23 см. - ISBN 978-5-9775-3758-2 (30 экз.).
2. Цильковский, И. А. Методы анализа знаний и данных [Электронный ресурс] : уч.-методич. пособие / И. А. Цильковский. — Новосибирск : Изд-во НГТУ, 2010. — 68 с. — Режим доступа: ЭБС «Руконт».
3. Чубукова, Ирина Александровна. Data Mining [Текст] : учеб. пособие / И. А. Чубукова. - М. : Интернет-Ун-т информ. технологий : Бином. Лаб. знаний, 2006. - 382 с. (9 экз.)
4. Сапрыкин, О. Н. Интеллектуальный анализ данных : учебное пособие / О. Н. Сапрыкин. — Самара : Самарский университет, 2020. — 80 с. — ISBN 978-5-7883-1563-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/188906> (дата обращения: 05.03.2023). — Режим доступа: для авториз. пользователей.

### **б) дополнительная литература**

1. Люгер, Джордж Ф. Искусственный интеллект [Текст] : стратегии и методы решения сложных проблем / Д.Ф. Люгер. - 4-е [межд.] изд. - М. : Вильямс, 2003. - 863 с. (4 экз.)

### **в) список авторских методических разработок:**

Материалы курса, опубликованные в ИОС «DOMIC».

### **г) базы данных, информационно-справочные и поисковые системы**

Google Colaboratory <https://colab.research.google.com/>

## **6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **6.1. УЧЕБНО-ЛАБОРАТОРНОЕ ОБОРУДОВАНИЕ:**

Для проведения лекционных занятий необходима аудитория с презентационным оборудованием, для проведения лабораторных занятий необходима аудитория на 25-30 рабочих мест (в зависимости от численности учебной группы), оборудованная доской, презентационной техникой, компьютерами.

### **6.2. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:**

1. Браузер.
2. Python 3.
3. Anaconda.

### **6.3. ТЕХНИЧЕСКИЕ И ЭЛЕКТРОННЫЕ СРЕДСТВА:**

ИОС DOMIC, презентационное оборудование, персональный компьютер с возможностью демонстрации презентаций в формате pdf.

## **7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ**

При реализации данного курса используются следующие образовательные технологии: технологии традиционного обучения, игровые технологии, технологии проблемного обучения, технологии обучения в сотрудничестве, технологии контекстного обучения, интерактивные технологии, технологии дистанционного обучения, активные педагогические технологии.

## 8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

### 8.1. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ВХОДНОГО КОНТРОЛЯ

Входной контроль обеспечивается выполнением учебного плана по предшествующим дисциплинам.

Входной контроль не предусмотрен в виде отдельных мероприятий.

### 8.2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ

Оценочные материалы представлены в виде лабораторных работ. Все задания доступны на странице курса в ИОС DOMIC. Имеется 20 лабораторных работ, каждая из которых содержит набор последовательно выполняемых заданий. Оценка за выполнение лабораторных работ выставляется пропорционально числу выполненных заданий

В качестве примера далее приводится фрагмент из заданий лабораторной работы по теме “Случайный лес”

1. Прочитайте набор данных `Wine_Quality_Data.csv`. Измените категорию признака "color" с качественной на количественную. (Вместо "red", "white" требуется записать 0, 1). Это можно сделать с использованием класса `LabelEncoder` или любым другим способом. Выведите сумму значений в этом столбце `print(wine['color'].sum())`  
Правильный ответ: 4898
2. Далее мы будем строить модель предсказания качества вина по всем остальным признакам. Поэтому требуется выделить признак "quality" в отдельную переменную `y`, а из всех остальных создать таблицу `x`. Для проверки выведите список имен всех столбцов `print(x.columns)`
3. Разделите набор данных на обучающую и тестовую выборку при помощи метода `x_train, x_test, y_train, y_test = train_test_split(x, y, random_state=42, test_size=0.5)`  
Выведите количество записей в каждой выборке.  
Правильный ответ: 3248 3249
4. Следующие шаги потребуются повторить в цикле 6 раз с изменением параметра `n_estimators`. Значения параметра следует взять из списка `[1, 5, 20, 50, 100, 200]`.
5. Постройте модель классификации методом случайного леса с заданным значением параметра `n_estimators`. Найдите предсказанные значения для обучающей и тестовой выборки. Посчитайте значение метрики `accuracy_score`. Выведите результат.
6. Посчитайте количество различных значений в разности `y_test - y_test_pred`. Здесь `y_test_prev` — предсказанные значения в тестовой выборке. Выведите количество значений для которых разность больше единицы по модулю.
7. Выведите столбчатую диаграмму с возможными разностями и их количеством, как показано в примере

### 8.3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПРОМЕЖУТОЧНОГО КОНТРОЛЯ

Итоговая сумма первичных баллов складывается из баллов, полученных за выполнение лабораторных работ (20 работ с максимальной оценкой в 3 балла) и баллов, полученных на экзамене (максимум 30 баллов).

Для достижения соответствия с бально-рейтинговой системой ИГУ первичные баллы пересчитываются в аттестационные баллы по формуле  $3 * \max(x, 20) + 0.5 * \max(x - 20, 20)$ , где  $x$  –


первичные баллы за курс. Далее оценка выставляется в соответствии с бально-рейтинговой системой ИГУ (80 баллов – отлично, 70 баллов – хорошо, 60 баллов – удовлетворительно).


### **Примерные вопросы к экзамену:**

1. Задача регрессии. Парная и множественная регрессия.
2. Графическое представление задачи парной регрессии.
3. Метод наименьших квадратов как задача оптимизации. Функция потерь.
4. Графическое представление метода наименьших квадратов.
5. Вывод формулы линейной множественной регрессии.
6. Реализация метода `predict` в задаче регрессии.
7. Метрики качества. Коэффициент детерминации.
8. Задача бинарной классификации.
9. Сигмоида и ее свойства.
10. Использование сигмоиды в задаче бинарной классификации.
11. Графическое представление линейной бинарной классификации с двумя признаками.
12. Метод максимального правдоподобия как задача оптимизации.
13. Вывод формулы логарифмической функции потерь из метода максимального правдоподобия.
14. Реализация методов `predict` и `predict_proba` в задаче бинарной классификации.
15. Матрица оценки качества бинарной классификации. Ошибки первого и второго рода.
16. Метрики качества: `accuracy`, `sensitivity`, `specificity`.
17. ROC - кривая. Построение ROC - кривой.
18. Метрики качества. ROC-AUC
19. Масштабирование признаков. Виды масштабирования. Стандартизация. Формулы для стандартизации признаков.
20. Прием перехода к пространствам большей размерности. Расширение пространства признаков многочленами.
21. Проблемы машинного обучения: плохая обусловленность, мультиколлинеарность и переобучение.
22. Регуляризация. Виды регуляризации. Добавление регуляризации в задачи оптимизации.
23. Графическое представление регуляризации L1 и L2.
24. Цель регуляризации. Влияние регуляризационного параметра  $C$  на параметры модели.
25. Метод градиентного спуска с фиксированным шагом.
26. Вывод формулы вычисления градиента для функции потерь.
27. Вывод формулы вычисления градиента для логарифмической функции потерь.
28. Метод наискорейшего градиентного спуска. Решение задачи одномерной оптимизации методом золотого сечения.
29. Метод сопряженных градиентов.
30. Метод k-ближайших соседей. Общее описание
31. Виды метрик в алгоритме k-ближайших соседей.
32. Выбор параметра  $k$  в алгоритме k-ближайших соседей и его влияние на переобучение модели.
33. Графическое представление метода опорных векторов в случае линейной разделимости.
34. Метод опорных векторов, как задача оптимизации. Случай линейной разделимости.
35. Метод опорных векторов, как задача оптимизации. Случай линейной неразделимости.
36. Двойственные коэффициенты в методе опорных векторов, их смысл для метода.
37. Формула для реализации метода `predict` в методе опорных векторов.

38. Определение ядерной функции в методе опорных векторов. Привести пример нетривиальной функции, которая будет ядерной по определению.
39. Формула радиальной базисной функции.
40. Формула полиномиальной ядерной функции.
41. Двоичное решающее дерево. Общее описание.
42. Формула метрики gini impurity.
43. Формула метрики энтропии.
44. Формула вычисления суммарной gini impurity для дерева
45. Алгоритм построения двоичного решающего дерева. Критерий для разбиения узла
46. Методы борьбы с переобучением для двоичного решающего дерева: ограничение глубины, ограничение по количеству образцов, ограничение по уменьшению impurity.
47. Случайный лес. Общее описание.
48. Почему случайный лес, как правило, позволяет добиться более высокого качества обучения, чем случайное дерево?
49. Случайный лес. Алгоритм для метода predict.
50. Какие из методов классификации требуют стандартизации признаков и почему?

**Разработчики:**

 \_\_\_\_\_ доцент \_\_\_\_\_ Казимиров А.С.  
 (подпись) (занимаемая должность) (Ф.И.О.)

 \_\_\_\_\_ доцент \_\_\_\_\_ Кириченко Константин Дмитриевич  
 (подпись) (занимаемая должность) (Ф.И.О.)

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.03.03 «Прикладная информатика» (уровень бакалавриата), утвержденный приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 г. № 922, зарегистрированный в Минюсте России «12» октября 2017 г. № 48531 с изменениями и дополнениями от 26.11.2020 г., 8.02.2021 г.

Программа рассмотрена на заседании кафедры алгебраических и информационных систем ИМИТ ИГУ «04» апреля 2023 г.

Протокол № 9 Зав. кафедрой \_\_\_\_\_  


*Настоящая программа не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.*