

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«Национальный исследовательский ядерный университет «МИФИ»  
**Обнинский институт атомной энергетики –**  
филиал федерального государственного автономного образовательного учреждения высшего образования  
«Национальный исследовательский ядерный университет «МИФИ»  
**(ИАТЭ НИЯУ МИФИ)**

## **ОТДЕЛЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ**

Утверждено на заседании  
УМС ИАТЭ НИЯУ МИФИ  
Протокол от 30.08.2021 № 2-8/2021

### **РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**

*Методы интеллектуального анализа данных*

*название дисциплины*

---

для направления подготовки

09.04.01 «Информатика и вычислительная техника»  
Профиль «Большие данные и машинное обучение (для Атомной энергетики)»

---

---

Форма обучения: очная

**г. Обнинск 2022 г.**

## **1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

- формирование у магистров представлений о современных понятиях Науки о Данных (Data Science), позиционирование методов интеллектуального анализа данных и Data Mining в структуре Науки о Данных,
- изучение магистрами современных методов интеллектуального анализа данных, основанных на понятиях прикладной статистики, эвристических методах, методах искусственного интеллекта и машинного обучения, и реализованных в виде программных продуктов для обработки и статистического анализа данных, для разведки данных (Data Mining);
- приобретение студентами знаний и навыков в области методов и средств прикладного интеллектуального анализа данных и разведки данных (Data Mining), и знаний и навыков реализации этих методов и средств в вычислительных средах,
- создание основы для изучения студентами в последующих семестрах технологий Больших Данных, методов и алгоритмов машинного обучения, визуальной статистики и других разделов программы.

## **2. ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

- закрепить понятия прикладной статистики, полученные магистрами на предыдущих стадиях обучения;
- познакомить студентов с программными средствами прикладной статистики и интеллектуального анализа данных, средствами реализации на ЭВМ методов и подходов Data Mining, выработать навыки работы в среде программных средств;
- познакомить студентов с терминологией прикладного анализа данных, применяемой в программных средствах, в том числе с англоязычной терминологией;
- детально рассмотреть основные методы прикладного статистического анализа данных, представления результатов в виде отчетности и статистической графики, закрепить полученные знания при работе в среде современных программных средств и инструментариев анализа данных и Data Mining.

## **3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ (далее – ОП) МАГИСТРАТУРЫ**

Дисциплина реализуется в рамках части, формируемой участниками образовательных отношений и относится к профессиональному модулю.

Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин:

М.01.01. Иностранный язык

Дисциплины и/или практики, для которых освоение данной дисциплины необходимо как предшествующее:

М.02.15. Дополнительные разделы исследования данных

Дисциплина изучается на 1 курсе в 1 семестре .

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

В результате освоения ОП магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине:

Код компетенций	Наименование компетенции	Код и наименование индикатора достижения компетенции
ОПК-1	Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте;	З- Знать современные методы моделирования процессов и объектов с применением стандартных пакетов для проведения исследований и проектирования. У- Уметь применять знания в области интеллектуального анализа данных, геоинформационных систем и технологий, параллельных и многопоточных вычислений с использованием стандартных.
ОПК-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач;	В- Владеть методами моделирование процессов и объектов на базе стандартных пакетов для проведения исследований втоматизированного проектирования.
СПК-1	Способен использовать и развивать методы научных исследований и инструментарий в области интеллектуального анализа данных	З- Знать современное состояние в области нейронных сетей и генетических алго-ритмов, нечеткой логики, нереляционных баз данных, Data mining, методов и систем поддержки принятия решений. У- Уметь применять методы искусственного интеллекта с использованием класси-ческой и нечеткой логики в научно-практических исследованиях и задачах. В- Владеть методами современных интеллектуальных кибернетических систем для решения научно-прикладных задач и поддержки принятия управленческих решений.

#### 5. ОБЪЕМ ДИСЦИПЛИНЫ В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Вид работы	Количество часов на вид работы:
<b>Контактная работа обучающихся с преподавателем</b>	
<b>Аудиторные занятия (всего)</b>	<b>48</b>
В том числе:	
<i>лекции</i>	16
<i>практические занятия</i>	16
<i>лабораторные занятия</i>	16
<b>Промежуточная аттестация</b>	
В том числе:	
<i>зачет</i>	8
<b>Самостоятельная работа обучающихся</b>	
<b>Самостоятельная работа обучающихся</b>	<b>60</b>
<b>Всего (часы):</b>	<b>144</b>
<b>Всего (зачетные единицы):</b>	<b>4</b>

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

### 6.1. Разделы дисциплины и трудоемкость по видам учебных занятий (в академических часах)

Неделя	Наименование раздела / темы дисциплины	Виды учебной работы				
		Лек	Пр	Лаб	Внеауд	СРО
1-8	<b>1.</b>					
1	1.1. Тема: «Цели и задачи Интеллектуального Анализа Данных (ИАД). Этапы проведения ИАД и их содержание. Подготовка данных и управление данными для ИАД».	2	2			10
2	1.2. Тема: «Элементарные статистические расчеты. Статистическая графика в ИАД. Проверка гипотез. Дисперсионный анализ (ANOVA). Оценки связи между числовыми переменными. Понятия Разведки данных (Data Mining)»	2	2	4		10
3-5	1.3. Тема: «Основные	2	4	4		10

	программные средства для проведения ИАД. Основные программные средства для проведения Data Mining. Виды выходной продукции ИАД и Data Mining».....					
6-8	1.4. Тема: «Регрессионный анализ в ИАД и в Data Mining»	2	2	4		2
9-12	1.5. . Тема: «Многомерный анализ: классификация, дискриминантный анализ, обучение с учителем, без учителя»	2				12
13	1.6. Тема: «Многомерный анализ: снижение размерности, метод главных компонент»	2	2	2		6
14-15	1.7. Тема: «Кластерный анализ»	2	2			4
16	1.8. Тема: «Реализация методов Data Mining»	-	2	2		6
	<b>Итого за 1 семестр:</b>	<b>16</b>	<b>16</b>	<b>16</b>		<b>60</b>
	<b>Всего:</b>	<b>16</b>	<b>16</b>	<b>16</b>		<b>60</b>

Прим.: Лек – лекции, Пр – практические занятия / семинары, Лаб – лабораторные занятия, Внеауд – внеаудиторная контактная работа, СРО – самостоятельная работа.

## 6.2. Содержание дисциплины, структурированное по разделам (темам)

### Лекционный курс

Неделя	Наименование раздела / темы дисциплины	Содержание
<b>1-8</b>	<b>1. Название раздела 1</b>	
1	1.1. Тема: «Цели и задачи Интеллектуального Анализа Данных (ИАД). Этапы проведения ИАД и их содержание. Подготовка данных и управление данными для ИАД».	Наука о Данных – основные сведения. Этапы: доступ к данным, управление данными, статистические расчеты, методы статистики, представление результатов ИАД (отчет, таблица, статистическая графика), виды статистических графиков.
2	1.2. Тема: «Элементарные статистические расчеты. Статистическая графика в ИАД. Проверка гипотез. Дисперсионный анализ (ANOVA). Оценки связи между числовыми переменными. Понятия Разведки данных (Data Mining)»	Одномерные статистические характеристики. Моменты. Матричные графики для представления результатов одномерных и многомерных аналитических процедур. Порядковые статистики. Устойчивые и робастные методы статистики. Меры связи между переменными (числовыми). Меры связи на основе порядковых статистик. Этапы разведки данных (Data Mining). Роли переменных для Data Mining. Терминология.
	1.3. Тема: «Основные программные средства для проведения ИАД.	Сведения о ПО для статистических расчетов. Сведения о ПО для Data Mining. Работа в среде

	Основные программные средства для проведения Data Mining. Виды выходной продукции ИАД и Data Mining».....	ПО для статистических расчетов. Работа в среде ПО для разведки данных Data Mining. Освоение программных средств для ИАД, особенности программного языка SAS, написание программ, практическая работа в среде SAS STUDIO. Приемы работы в англоязычных программных средах ИАД
	1.4. Тема: «Регрессионный анализ в ИАД и в Data Mining»	Восстановление зависимостей. Оптимизационная задача о восстановлении зависимостей для задач классификации и регрессии. Регрессионный анализ. Терминология регрессионного анализа. Критерии качества регрессии. Линейный и нелинейный регрессионный анализ. Пошаговая регрессия, стратегии отбора предикторов. Специальные разделы регрессионного анализа (робастная регрессия, квантильная регрессия, диагностика и анализ остатков, логистическая регрессия). Нахождение оптимума между длиной выборки и сложностью статистической модели. Обобщенные линейные модели (GLM).
	1.5. . Тема: «Многомерный анализ: классификация, дискриминантный анализ, обучение с учителем, без учителя»	Задача классификации. Линейный классификатор, оптимальный линейный классификатор. Разделы дискриминантного анализа: пошаговый дискриминантный анализ, канонический дискриминантный анализ.
	1.6. Тема: «Многомерный анализ: снижение размерности, метод главных компонент»	Линейное преобразование переменных в задаче снижения размерности. Главные компоненты. Метки главных компонент.
	1.7. Тема: «Кластерный анализ»	Кластеризация – постановка задачи, типы алгоритмов кластерного анализа. Алгоритмы агломеративные и дивизивные. Методы иерархической классификации. Методы, основанные на плотности (метод DBSCAN).
	1.8. Тема: «Реализация методов Data Mining»	Лекционные занятия не предусмотрены

*Практические/семинарские занятия*

Неделя	Наименование раздела / темы дисциплины	Содержание
<b>1-16</b>	<b>1. Название раздела 1</b>	
	1.1. Тема: «Цели и задачи Интеллектуального Анализа Данных (ИАД). Этапы проведения ИАД и их содержание. Подготовка данных и управление данными для ИАД».	Реализация ИАД в среде SAS STUDIO ON DEMAND. Сведения о программном языке SAS. Роли переменных. Классифицирующие (группирующие) переменные. BY-переменные. Процедуры управления данными. Импорт-экспорт данных.
	1.2. Тема: «Элементарные	Реализация элементарных расчетов в среде

	статистические расчеты. Статистическая графика в ИАД. Проверка гипотез. Дисперсионный анализ (ANOVA). Оценки связи между числовыми переменными. Понятия Разведки данных (Data Mining)»	SAS STUDIO. Реализации расчетов корреляционных связей.
	1.3. Тема: «Основные программные средства для проведения ИАД. Основные программные средства для проведения Data Mining. Виды выходной продукции ИАД и Data Mining».....	Представление результатов статистических расчетов и Data Mining. Среда SAS on Demand – SAS STUDIO Среда ORANGE- компоненты анализа
	1.4. Тема: «Регрессионный анализ в ИАД и в Data Mining»	Листинг процедур регрессионного анализа (интерпретация результатов).
1	1.5. . Тема: «Многомерный анализ: классификация, дискриминантный анализ, обучение с учителем, без учителя»	
2	1.6. Тема: «Многомерный анализ: снижение размерности, метод главных компонент»	
...	1.7. Тема: «Кластерный анализ»	Практические приемы кластерного анализа. Характеристики качества кластеризации. Выбор оптимального числа кластеров.
...	1.8. Тема: «Реализация методов Data Mining»	

#### Лабораторные занятия

Неделя	Наименование раздела / темы дисциплины	Содержание
<b>1-16</b>		
	1.1. Тема: «Цели и задачи Интеллектуального Анализа Данных (ИАД). Этапы проведения ИАД и их содержание. Подготовка данных и управление данными для ИАД».	Реализация ИАД в среде SAS STUDIO ON DEMAND. Сведения о программном языке SAS/ Роли переменных. Классифицирующие (группирующие) переменные. ВУ-переменные. Процедуры управления данными. Импорт-экспорт данных.
	1.2. Тема: «Элементарные статистические расчеты. Статистическая графика в ИАД. Проверка гипотез. Дисперсионный анализ (ANOVA). Оценки связи между числовыми переменными. Понятия Разведки данных (Data Mining)»	
	1.3. Тема: «Основные программные средства для проведения ИАД. Основные программные средства для проведения Data Mining. Виды выходной продукции ИАД и Data Mining».....	
	1.4. Тема: «Регрессионный анализ в ИАД и в Data Mining»	Интерпретация результатов регрессионного анализа. Диагностика остатков в

		регрессионном анализе. . Регрессионный анализ данных при наличии выбросов. Альтернативы МНК-оценкам Практическая реализация отдельных методов регрессионного анализа
	1.5. . Тема: «Многомерный анализ: классификация, дискриминантный анализ, обучение с учителем, без учителя»	
	1.6. Тема: «Многомерный анализ: снижение размерности, метод главных компонент»	Интерпретация результатов расчета главных компонент. Разбор конкретных примеров.
...	1.7. Тема: «Кластерный анализ»	Особенности реализации алгоритмов кластерного анализа в программных средствах. Определение оптимального количества кластеров для иерархических алгоритмов и для алгоритмов метода К-средних (эвристический подход).
...	1.8. Тема: «Реализация методов Data Mining»	

## 7. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

1. Крянев А.В., Лукин Г.В., Удумян Д.К. Метрический анализ и обработка данных – М.Физматлит. 2012 – 308 с. ([http://e.lanbook.com/books/element.php?pl1\\_id=59523](http://e.lanbook.com/books/element.php?pl1_id=59523))
2. Салмин А.А. Анализ данных. Конспект лекций. – Самара.: ФГОБУ ВПО «ПГУТИ», 2013. - 111 с. (<http://window.edu.ru/resource/815/78815>)
3. Вьюгин В.В. «Математические основы теории машинного обучения и прогнозирования» М.: 2013. - 387 с.
4. ALAN C. ELLIOTT, WAYNE A. WOODWARD. SAS ESSENTIALS. A Guide to Mastering SAS for Research. Wiley, 2011
5. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. - СПб.: Питер, 2001. - 656 с.
6. Боровиков В.П. Популярное введение в современный анализ данных и машинное обучение на Statistica. Бином, 2019, 394 стр.
7. Домингос П. Верховный алгоритм. «МАНН, ИВАНОВ И ФЕРБЕР», 2016, 334 с.,
8. BURGES, CHRISTOPHER J.C.. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2, 121-167, 1998
9. Уиттон Д., Хасты Т., Тибширани Р. Введение в статистическое обучение с примерами на языке R. Пер. с англ. С. Э. Мاستицкого - М.: ДМК Пресс, 2016. - 450 с.: ил.
10. Хасты Т., Тибширани Р. , Фридман, Дж., Основы статистического обучения. М., Диалектика, 2020, 759 с.
11. Червоненкис А.Я. Компьютерный анализ данных. ШАД Яндекс, 2009, 259 с.
12. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. Москва, Наука 1974
13. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. Москва, Наука 1979
14. Воронцов К. В.. Математические методы обучения по прецедентам (теория обучения



машин). <http://www.ccas.ru/voron>, voron@ccas.ru

## 8. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 8.1. Связь между формируемыми компетенциями и формами контроля их освоения

№ п/п	Контролируемые разделы (темы) дисциплины	Индикатор достижения компетенции	Наименование оценочного средства текущей и промежуточной аттестации
<b>Текущая аттестация, 1 семестр</b>			
1.	1.1. Тема: «Цели и задачи Интеллектуального Анализа Данных (ИАД). Этапы проведения ИАД и их содержание. Подготовка данных и управление данными для ИАД».	ОПК-1, ОПК-2, СПК-1	Зд
2.	1.2. Тема: «Элементарные статистические расчеты. Статистическая графика в ИАД. Проверка гипотез. Дисперсионный анализ (ANOVA). Оценки связи между числовыми переменными. Понятия Разведки данных (Data Mining)»	ОПК-1, ОПК-2, СПК-1	Зд
3	ОПК-1, ОПК-2, СПК-1	ОПК-2, СПК-1	Зд
4	1.4. Тема: «Регрессионный анализ в ИАД и в Data Mining»	ОПК-2, СПК-1	БДЗ
5	1.5. . Тема: «Многомерный анализ: классификация, дискриминантный анализ, обучение с учителем, без учителя»	ОПК-2, СПК-1	Зд
6	1.6. Тема: «Многомерный анализ: снижение размерности, метод главных компонент»	ОПК-2, СПК-1	Зд
7	1.7. Тема: «Кластерный анализ»	ОПК-2, СПК-1	Зд
8	1.8. Тема: «Реализация методов Data Mining»	ОПК-1, ОПК-2, СПК-1	БДЗ
<b>Аттестация, 1 семестр</b>			
	<b>ЗАЧЕТ</b>		<b>экзамен</b>

### 8.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков, характеризующих этапы формирования компетенций

Фонд оценочных средств по дисциплине обеспечивает проверку освоения планируемых результатов обучения (компетенций и их индикаторов) посредством мероприятий текущей и промежуточной аттестации по дисциплине.

Оценочные средства приведены в Приложении «Фонд оценочных средств».

### 8.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

- Итоговая аттестация по дисциплине является интегральным показателем качества теоретических и практических знаний и навыков обучающихся по дисциплине и складывается из оценок, полученных в ходе текущей и промежуточной аттестации.
- Текущая аттестация в семестре проводится с целью обеспечения своевременной обратной связи, для коррекции обучения, активизации самостоятельной работы обучающихся.
- Промежуточная аттестация предназначена для объективного подтверждения и оценивания достигнутых результатов обучения после завершения изучения дисциплины.
- Текущая аттестация осуществляется два раза в семестр:
  - контрольная точка № 1 (КТ № 1) – выставляется в электронную ведомость не позднее 8 недели учебного семестра. Включает в себя оценку мероприятий текущего контроля аудиторной и самостоятельной работы обучающегося по разделам/темам учебной дисциплины с 1 по 8 неделю учебного семестра.
  - контрольная точка № 2 (КТ № 2) – выставляется в электронную ведомость не позднее 16 недели учебного семестра. Включает в себя оценку мероприятий текущего контроля аудиторной и самостоятельной работы обучающегося по разделам/темам учебной дисциплины с 9 по 16 неделю учебного семестра.
- Результаты текущей и промежуточной аттестации подводятся по шкале балльно-рейтинговой системы.

Этап рейтинговой системы / Оценочное средство	Неделя	Балл	
		Минимум*	Максимум**
<b>Текущая аттестация</b>	<b>1-16</b>	<b>36 - 60% от максимума</b>	<b>60</b>
<b>Контрольная точка № 1</b>	<b>7-8</b>	<b>18 (60% от 30)</b>	<b>30</b>
<i>Оценочное средство № 1.1</i>	БДЗ	60% от М1	М1
<b>Контрольная точка № 2</b>	<b>15-16</b>	<b>18 (60% от 30)</b>	<b>30</b>
<i>Оценочное средство № 2.1</i>	БДЗ	60% от Т1	Т1
<i>Оценочное средство № 2.У</i>		60% от ТУ	ТУ
<b>Промежуточная аттестация</b>	-	<b>24 – (60% 40)</b>	<b>40</b>
<b>ИТОГО по дисциплине</b>		<b>60</b>	<b>100</b>

\* - Минимальное количество баллов за оценочное средство – это количество баллов, набранное обучающимся, при котором оценочное средство засчитывается, в противном случае обучающийся должен ликвидировать появившуюся академическую задолженность по текущей или промежуточной аттестации. Минимальное количество баллов за текущую аттестацию, в т.ч. отдельное оценочное средство в ее составе, и промежуточную аттестацию составляет 60% от соответствующих максимальных баллов.

Студент считается аттестованным по разделу, зачету или экзамену, если он набрал не менее 60% от максимального балла, предусмотренного рабочей программой.

Студент может быть аттестован по дисциплине, если он аттестован по каждому разделу, зачету/экзамену и его суммарный балл составляет не менее 60.

### 8.4. Шкала оценки образовательных достижений

Итоговая аттестация по дисциплине оценивается по 100-балльной шкале и представляет сумму баллов, заработанных обучающимся при выполнении заданий в рамках текущей и промежуточной аттестации

<i>Сумма баллов</i>	<i>Оценка по 4-х балльной шкале</i>	<i>Оценка ECTS</i>	<i>Требования к уровню освоения учебной дисциплины</i>
<b>90-100</b>	5- «отлично»/ «зачтено»	A	Оценка «отлично» выставляется обучающемуся, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, использует в ответе материал монографической литературы
<b>85-89</b>	4 - «хорошо»/ «зачтено»	B	Оценка «хорошо» выставляется обучающемуся, если он твёрдо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос
<b>75-84</b>		C	
<b>70--74</b>		D	
<b>65-69</b>	3 - «удовлетворительно»/ «зачтено»	D	Оценка «удовлетворительно» выставляется обучающемуся, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала
<b>60-64</b>		E	
<b>0-59</b>	2 - «неудовлетворительно»/ «не зачтено»	F	Оценка «неудовлетворительно» выставляется обучающемуся, который не знает значительной части программного материала, допускает существенные ошибки. Как правило, оценка «неудовлетворительно» ставится обучающимся, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине

## **9. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

### **а) основная учебная литература:**

1. Боровиков В.П. Популярное введение в современный анализ данных и машинное обучение на Statistica. Бином, 2019, 394 стр.

2. Домингос П. Верховный алгоритм. «МАНН, ИВАНОВ И ФЕРБЕР», 2016, 334 с.,
3. BURGES, CHRISTOPHER J.C.. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2, 121-167, 1998
4. Уиттон Д., Хасти Т., Тибширани Р. Введение в статистическое обучение с примерами на языке R. Пер. с англ. С. Э. Мاستицкого - М.: ДМК Пресс, 2016. - 450 с.: ил.
5. Хасти Т., Тибширани Р., Фридман, Дж., Основы статистического обучения. М., Диалектика, 2020, 759 с.
6. Червоненкис А.Я. Компьютерный анализ данных. ШАД Яндекс, 2009, 259 с.
7. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. Москва, Наука 1974
8. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. Москва, Наука 1979
9. Воронцов К. В.. Математические методы обучения по прецедентам (теория обучения машин). <http://www.ccas.ru/voron>, [voron@ccas.ru](mailto:voron@ccas.ru)
10. Mohammed J. Zaki and Wagner Meira, Jr. Data Mining and Machine Learning: Fundamental Concepts and Algorithms. Second Edition. Cambridge University Press, March 2020 ISBN: 978-1108473989. 660 p.

**б) дополнительная учебная литература:**

1. Крянев А.В., Лукин Г.В., Удумян Д.К. Метрический анализ и обработка данных – М.Физматлит. 2012 – 308 с. ([http://e.lanbook.com/books/element.php?pl1\\_id=59523](http://e.lanbook.com/books/element.php?pl1_id=59523))
2. Салмин А.А. Анализ данных. Конспект лекций. – Самара.: ФГОБУ ВПО «ПГУТИ», 2013. - 111 с. (<http://window.edu.ru/resource/815/78815>)
3. Вьюгин В.В. «Математические основы теории машинного обучения и прогнозирования» М.: 2013. - 387 с.
4. ALAN C. ELLIOTT, WAYNE A. WOODWARD. SAS ESSENTIALS. A Guide to Mastering SAS for Research. Wiley, 2011
5. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. - СПб.: Питер, 2001. - 656 с.

**10. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ» (ДАЛЕЕ - СЕТЬ «ИНТЕРНЕТ»), НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

**Открытое образование:** <https://openedu.ru/course/hse/INTRML/>

**Хабр:** <https://habr.com/ru/post/319288/>

**Машинное обучение. Курс лекций. К.В. Воронцов.**  
[http://www.machinelearning.ru/wiki/index.php?title=Машинное\\_обучение\\_%28курс\\_лекций%2C\\_К.В.Воронцов%29](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_%28курс_лекций%2C_К.В.Воронцов%29)

**Основной ресурс получения «помощи» и исчерпывающих справочных сведений о системе SAS.** Включает удобные средства поиска нужной информации, детальные описания всех процедур с примерами: [www.Support.sas.com](http://www.Support.sas.com)

## **11. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Методические указания для обучающихся по освоению дисциплины (рекомендуемый режим и характер учебной работы, в том числе в части выполнения самостоятельной работы) – комплекс рекомендаций и разъяснений, позволяющий обучающимся оптимальным образом организовать процесс изучения как теоретического учебного материала дисциплины, так и подготовки к практическим занятиям и/или лабораторным работам, в том числе проводимым с использованием активных и интерактивных технологий обучения.

Содержание методических рекомендаций может включать:

- общие рекомендации по самостоятельной работе обучающихся;
- советы по планированию и организации времени, отведенного на изучение дисциплины;
- описание последовательности действий студента или «сценарий изучения дисциплины»;
- пожелания по изучению отдельных тем курса;
- рекомендации по освоению лекционного материала, подготовке к лекциям;
- перечень тем, выносимых на самостоятельное изучение, с рекомендациями по использованию учебно-методических материалов с указанием вида контроля;

## **12. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ (ПРИ НЕОБХОДИМОСТИ)**

Использование информационных технологий при осуществлении образовательного процесса по дисциплине осуществляется в соответствии с утвержденным Положением об Электронной информационно-образовательной среде ИАТЭ НИЯУ МИФИ.

Электронная система управления обучением (LMS) используется для реализации образовательных программ при очном, дистанционном и смешанном режиме обучения. Система реализует следующие основные функции:

- 1) Создание и управление классами,
- 2) Создание курсов,
- 3) Организация записи учащихся на курс,
- 4) Предоставление доступа к учебным материалам для учащихся,
- 5) Публикация заданий для учеников,
- 6) Оценка заданий учащихся, проведение тестов и отслеживание прогресса обучения,
- 7) Организация взаимодействия участников образовательного процесса.

Система интегрируется с дополнительными сервисами, обеспечивающими возможность использования таких функций как рабочий календарь, видео связь, многопользовательское редактирование документов, создание форм опросников, интерактивная доска для рисования. Авторизация пользователей в системе осуществляется посредством корпоративных аккаунтов, привязанных к домену oiate.ru.

### ***12.1. Перечень информационных технологий***

При осуществлении образовательного процесса по дисциплине используются следующие информационные технологии:

- проведение лекций и практических занятий с использованием слайд-презентаций;
- Проведение занятий в с использованием Дистанционных Обучающих Технологий (ДОТ) с использованием Google Meet, Google Classroom;
- использование текстового редактора Microsoft Word;
- использование табличного редактора Microsoft Excel;

- использование текстового редактора NoteBook (Блокнот);
- организация взаимодействия с обучающимися посредством электронной почты и ЭИОС.

.....

## **12.2. Перечень программного обеспечения**

1. Текстовый редактор Microsoft Word;
2. Табличный редактор Microsoft Excel;
3. Редактор презентаций Microsoft PowerPoint;
4. Текстовый редактор NoteBook (Блокнот);
5. Браузеры: Google Chrome, Internet Explorer, Yandex, Mozilla Firefox, Opera.
7. Локальная компьютерная сеть и глобальная сеть Интернет.

## **12.3. Перечень информационных справочных систем**

Доступ к электронным библиотечным ресурсам и электронной библиотечной системе (ЭБС) осуществляется посредством специальных разделов на официальном сайте ИАТЭ НИЯУ МИФИ. Обеспечен доступ к электронным каталогам библиотеки ИАТЭ НИЯУ МИФИ, а также электронным образовательным ресурсам (ЭИОС), сформированным на основании прямых договоров с правообладателями учебной и учебно-методической литературы, методических пособий:

- 1) Электронно-библиотечная система BOOK.ru, [www.book.ru](http://www.book.ru);
- 2) Базы данных «Электронно-библиотечная система elibrary» (ЭБС elibrary);
- 3) Базовая версия ЭБС IPRbooks, [www.iprbooks.ru](http://www.iprbooks.ru);
- 4) <http://ibooks.ru/home.php?routine=bookshelf>

## **13. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ**

Аудитории, где имеются проектор с компьютером, компьютеры для выполнения практических работ, соединенные в сеть и имеющие выход в Интернет (аудитории 2-613, 2-604, 2-606 или аналоги). Для аудитории с лекционным оборудованием желательны доска с необходимыми атрибутами, стенд с бумагой, а также светодиодная панель на стенке, подключенная к компьютеру.

Желательны видеовходы к проектору для подключения стороннего (преподавательского) компьютера, с переходниками:

Аудитория 2-610 ИАТЭ НИЯУ МИФИ  
 Аудитория 2-613 ИАТЭ НИЯУ МИФИ  
 Аудитория 2-604 ИАТЭ НИЯУ МИФИ  
 Аудитория 2-606 ИАТЭ НИЯУ МИФИ

## **14. ИНЫЕ СВЕДЕНИЯ И (ИЛИ) МАТЕРИАЛЫ**

### **14.1. Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине**

№ пп	Наименование темы дисциплины	Вид занятий (лекция, семинары, практические занятия)	Количество ак. ч.	Наименование активных и интерактивных форм проведения занятий
1	Темы 1-8.	Лекции	16	<p>Обратная связь с аудиторией с целью активизации процессов восприятия, объяснение терминологии.</p> <p>Приведение примеров конкретных предметных областей, где могут применяться изучаемые методы и алгоритмы (примеры из области экономики, производства, медицины, финансов и банковской сферы,</p> <p>Разъяснение терминологии, происхождения терминов, перевода</p>
2	Темы 1-8.	Лабораторные	16	<p>Обратная связь с аудиторией с целью активизации процессов восприятия, объяснение терминологии.</p> <p>Приведение примеров конкретных предметных областей, где могут применяться изучаемые методы и алгоритмы (примеры из области экономики, производства, медицины, финансов и банковской сферы,</p> <p>Разъяснение терминологии, происхождения терминов, перевода</p>
3	Темы 1-8.	Практические занятия	16	<p>Обратная связь с аудиторией с целью активизации процессов восприятия, объяснение терминологии.</p> <p>Приведение примеров конкретных предметных областей, где могут применяться изучаемые методы и алгоритмы (примеры из области экономики, производства, медицины, финансов и банковской сферы,</p> <p>Разъяснение терминологии, происхождения терминов, перевода</p>

## **14.2. Формы организации самостоятельной работы обучающихся (темы, выносимые для самостоятельного изучения; вопросы для самоконтроля; типовые задания для самопроверки)**

Для самостоятельной работы – самостоятельное изучение тем:

Виды статистических графиков

Логистическая регрессия

Язык программирования SAS 4GL

ORANGE Nodes

## **14.3. Краткий терминологический словарь)**

*Переменная* — атрибут физической или абстрактной системы, который может изменять своё, как правило численное, значение. Понятие переменной широко используется в таких областях как математика, естественные науки, техника и программирование. Примерами переменных могут служить: температура воздуха, параметр функции и многое другое. Переменная характеризуется только множеством значений, которые она может принимать. Переменную обозначают символом, общим для каждого из её значений.

*Статистическая модель* – математически выраженная модель, которая описывает взаимосвязь между случайными переменными.

*Эвристическая модель* проблемы – интеллектуальное средство, отражающее основные понятия предметной области и формализующее анализ

*Фактор* – нечто, оказывающее влияние на решение проблемы; обобщенное обозначение таких понятий как переменная, характеристика, свойство, параметр, признак и т.д.

*Измерение* – алгоритмическая операция, которая данному объекту (процессу, явлению, состоянию, фактору и т.д.) ставит в соответствие определенное обозначение: число, номер или символ

## **15. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ**

В соответствии с методическими рекомендациями Минобрнауки РФ (утв. 8 апреля 2014 г. № АК-44/05вн) в курсе предполагается использовать социально-активные и рефлексивные методы обучения, технологии социокультурной реабилитации обучающихся с ОВЗ с целью оказания помощи в установлении полноценных межличностных отношений.

Обучение лиц с ограниченными возможностями здоровья осуществляется с учетом индивидуальных психофизических особенностей, а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида.

**Для лиц с нарушением слуха** возможно предоставление информации визуально (краткий конспект лекций, основная и дополнительная литература), на лекционных и практических занятиях допускается присутствие ассистента, а так же, сурдопереводчиков и тифлосурдопереводчиков.

Оценка знаний обучающихся на практических занятиях осуществляется на основе письменных конспектов ответов на вопросы, письменно выполненных практических заданий.

Доклад так же может быть предоставлен в письменной форме (в виде реферата), при этом требования к содержанию остаются теми же, а требования к качеству изложения материала (понятность, качество речи, взаимодействие с аудиторией и т. д) заменяются на соответствующие требования, предъявляемые к письменным работам (качество оформления текста и списка литературы, грамотность, наличие иллюстрационных материалов и т.д.)

С учетом состояния здоровья просмотр кинофильма с последующим анализом может быть проведен дома (например, при необходимости дополни-тельной звукоусиливающей аппаратуры



(наушники)). В таком случае обучающийся предоставляет письменный анализ, соответствующий предъявляемым требованиям.

Промежуточная аттестация для лиц с нарушениями слуха проводится в письменной форме, при этом используются общие критерии оценивания. При необходимости, время подготовки на зачете может быть увеличено.

Для **лиц с нарушением зрения** допускается аудиальное предоставление информации (например, с использованием программ-синтезаторов речи), а так же использование на лекциях звукозаписывающих устройств (диктофонов и т.д.). Допускается присутствие на занятиях ассистента (помощника), оказывающего обучающимся необходимую техническую помощь.

Оценка знаний обучающихся на практических занятиях осуществляется в устной форме (как ответы на вопросы, так и практические задания). При необходимости анализа фильма может быть заменен описанием ситуации межэтнического взаимодействия (на основе опыта респондента, художественной литера-туры и т.д.), позволяющим оценить степень сформированности навыков владения методами анализа и выявления специфики функционирования и развития психики, позволяющими учитывать влияние этнических факторов. При проведении промежуточной аттестации для лиц с нарушением зрения тестирование может быть заменено на устное собеседование по вопросам.

**Лица с нарушениями опорно-двигательного аппарата** не нуждаются в особых формах предоставления учебных материалов. Однако, с учетом состояния здоровья часть занятий может быть реализована дистанционно (при помощи сети «Интернет»). Так, при невозможности посещения лекционного занятия обучающийся может воспользоваться кратким конспектом лекции.

При невозможности посещения практического занятия обучающийся должен предоставить письменный конспект ответов на вопросы, письменно выполненное практическое задание.

Доклад так же может быть предоставлен в письменной форме (в виде реферата), при этом требования к содержанию остаются теми же, а требования к качеству изложения материала (понятность, качество речи, взаимодействие с аудиторией и т. д) заменяются на соответствующие требования, предъявляемые к письменным работам (качество оформления текста и списка литературы, грамотность, наличие иллюстрационных материалов и т.д.).

Промежуточная аттестация для лиц с нарушениями опорно-двигательного аппарата проводится на общих основаниях, при необходимости процедура зачета может быть реализована дистанционно (например, при помощи программы Skype).

Для этого по договоренности с преподавателем обучающийся в определенное время выходит на связь для проведения процедуры зачета. В таком случае зачет сдается в виде собеседования по вопросам (см. формы проведения промежуточной аттестации для лиц с нарушениями зрения). Вопрос и практическое задание выбираются самим преподавателем.

Примечание: Фонды оценочных средств, включающие типовые задания и методы оценки, критерии оценивания, позволяющие оценить результаты освоения данной дисциплины обучающимися с ОВЗ могут входить в состав РПД на правах отдельного документа.

**Программу составил (а) (и):**

И.О. Фамилия, должность, ученая степень, ученое звание

А. М. Стерин, профессор, доктор физ.-мат. наук, старший научный сотрудник

....

**Рецензент (ы):**

И.О. Фамилия, должность, ученая степень, ученое звание

....



## ЛИСТ СОГЛАСОВАНИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

*Интеллектуальный анализ данных и Data mining (дополнительные главы)*

<p>Программа рассмотрена на заседании отделения Наименование отделения (протокол № ____ от «__» _____ 20__ г.)</p>	<p>Руководитель образовательной программы 00.00.00 Наименование специальности/направления подготовки «__»____20__ г. ____ И.О.Фамилия</p> <p>Начальник отделения Название отделения «__»____20__ г. ____ И.О.Фамилия</p> <p>Научный руководитель магистерской программы (<b>при необходимости</b>) 00.00.00 Наименование специальности/направления подготовки «__»____20__ г. ____ И.О.Фамилия</p>
--	--

