

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Костромской государственный университет»  
(КГУ)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**ИНФОРМАЦИОННЫЙ АНАЛИЗ БОЛЬШИХ ДАННЫХ В СИСТЕМАХ  
УПРАВЛЕНИЯ**

направление подготовки «(27.04.04) Управление в технических системах»  
направленность «Интеллектуальные системы адаптивного управления»  
Квалификация выпускника: магистр

Кострома  
2021

Рабочая программа дисциплины «Информационный анализ больших данных в системах управления» разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования (утвержден приказом Минобрнауки России от 11.08.2020 № 942) по направлению подготовки 27.04.04 «Управление в технических системах (уровень магистратуры)».

Разработал:

В.В. Олоничев, к.т.н., доцент кафедры АМТ

Рецензент:

Л.В. Воронова, к.т.н., доцент кафедры АМТ

**ПРОГРАММА УТВЕРЖДЕНА:**

Заведующий кафедрой автоматике и микропроцессорной техники:

Староверов Борис Александрович, д.т.н., профессор

Протокол заседания кафедры № 9 от 12.05.2021г.

**ПРОГРАММА ПЕРЕУТВЕРЖДЕНА:**

На заседании кафедры Автоматики и микропроцессорной техники:

Протокол заседания кафедры № 8 от 04.03.2022 г.

Заведующий кафедрой автоматике и микропроцессорной техники:

Староверов Борис Александрович, д.т.н., профессор

**ПРОГРАММА ПЕРЕУТВЕРЖДЕНА:**

На заседании кафедры Автоматики и микропроцессорной техники:

Протокол заседания кафедры № 6 от 21.04.2023 г.

Заведующий кафедрой автоматике и микропроцессорной техники:

Староверов Борис Александрович, д.т.н., профессор

## **1. Цели и задачи освоения дисциплины**

Цель изучения дисциплины — формирование у студентов знаний, умений и навыков создания и применения программ для систем управления, работающих с большими данными и использующие элементы машинного обучения..

### **2. Перечень планируемых результатов обучения по дисциплине**

В результате освоения дисциплины обучающийся должен:

#### **знать:**

типы данных и операции с ними в библиотеке Pandas;  
основные виды моделей, используемых в машинном обучении;  
этапы построения и проверки модели при машинном обучении.

#### **уметь:**

использовать язык программирования Python и библиотеку Pandas для анализа данных и формирования отчетов;

использовать язык программирования Python и библиотеки yellowbrick и scikitlearn для задач машинного обучения

#### **владеть:**

навыками анализа больших данных и генерации отчетов;

навыками использования машинного обучения для построения и практического использования математических моделей изучаемых объектов.

#### **Освоить компетенцию:**

ПК-1 - Способен проводить непосредственное руководство процессами разработки программного обеспечения, программно-техническими, технологическими и человеческими ресурсами.

#### **Код и содержание индикаторов освоения компетенций:**

ИПК1.1. Знает основные понятия, определения, свойства формализованных моделей и требования к типовым объектам управления.

ИПК1.2. Умеет формализовать поставленную профессиональную задачу.

ИПК1.3. Владеет способами формализации моделей объектов и систем управления; практическими навыками разработки формализованных моделей объектов и систем управления.

## **3. Место дисциплины в структуре ОП ВО**

Дисциплина Б1.В.04 относится к части учебного плана, формируемой участниками образовательных отношений. Изучается в 1 семестре очной формы обучения. Изучение дисциплины основывается на ранее освоенных дисциплинах: «Программирование» и «Объектно-ориентированное программирование». Изучение дисциплины является основой для освоения последующих дисциплин/практик: Междисциплинарный проект.

## **4. Объем дисциплины**

### **4.1. Объем дисциплины в зачетных единицах с указанием академических часов и виды учебной работы**

Виды учебной работы,	Очная
Общая трудоемкость в зачетных	3
Общая трудоемкость в часах	108
Аудиторные занятия в часах, в том	32
Лекции	16
Практика	-
Лабораторные занятия	16
Самостоятельная работа в часах	73.65
Контроль	2.35
Форма промежуточной аттестации	Экзамен

#### 4.2. Объем контактной работы на 1 обучающегося

Виды учебных занятий	Очная форма
Лекции	16
Практические занятия	-
Лабораторные занятия	16
Консультации	-
Зачет/зачеты	-
Экзамен/экзамены	0.35
Курсовые работы	-
Курсовые проекты	-
Практическая подготовка	-
<b>Всего</b>	<b>32.35</b>

### 5. Содержание дисциплины, структурированное по темам (разделам), с указанием количества часов и видов занятий

#### 5.1. Тематический план учебной дисциплины

№	Название раздела, темы	Всего час	Аудиторные занятия			Самостоятельная работа
			Лекц.	Практ.	Лаб.	
1	Обработка больших данных в задачах формирования	24	4	-	-	20
2	Использование библиотеки Pandas в задачах анализа данных	27	4	-	8	15
3	Машинное обучение и его использование в системах автоматического и автоматизированного управления.	19	4	-	-	15

4	Использование библиотек yellowbrick и scikitlearn для построения систем машинного	35.65	4	-	8	23.65
	Итого:	105.65	16	-	16	73.65

## 5.2. Содержание

### 1. Обработка больших данных в задачах формирования отчетов.

Большие данные и дата-аналитика. Технология map-reduce. Форматы хранения больших данных.

### 2. Использование библиотеки Pandas в задачах анализа данных.

Библиотека Pandas и типы данных этой библиотеки. Функции Pandas для манипулирования данными и построения графиков.

### 3. Машинное обучение и его использование в системах автоматического и автоматизированного управления.

Машинное обучение и бинарная модель объектов управления. Области применения машинного обучения. Экономическая эффективность использования машинного обучения и моделей объектов в системах автоматического и автоматизированного управления.

### 4. Использование библиотек yellowbrick и scikitlearn для построения систем машинного обучения.

Этапы машинного обучения: подготовка данных, выбор вида модели, верификация модели. ROC-кривая, матрица ошибок (confusion matrix) и отчет о бинарной классификации (binary classification report): полнота модели и точность модели.

## 6. Методические материалы для обучающихся по освоению дисциплины

### 6.1. Самостоятельная работа обучающихся по дисциплине

№ п/п	Раздел (тема) дисциплин	Задание	Часы	Методические рекомендации по выполнению задания	Форма контроля
1	Обработка больших данных в задачах формирования отчетов.	Изучение лекционного материала.	20	Изучить: - материалы лекции; - учебник [1]; - методические указания [1];	Вопросы по темам/разделам дисциплины Тестовое задание
2	Использование библиотеки Pandas в задачах анализа данных.	Изучение лекционного материала.	15	Изучить: - материалы лекции; - учебник [1]; - методические указания [1]; Отчет по	Вопросы по темам/разделам дисциплины Тестовое задание. Защита лабораторных работ

				лабораторной работе подготовить в соответствии с методическими указаниями [4].	
3	Машинное обучение и его использование в системах автоматического и автоматизированного управления.	Изучение лекционного материала. Оформление отчета по лабораторной работе	15	Изучить: - материалы лекции; - учебник [1]; - методические указания [1];	Вопросы по темам/разделам дисциплины Тестовое задание
4	Использование библиотек yellowbrick и scikitlearn для построения систем машинного обучения.		23.6 5	Изучить: - материалы лекции; - учебник [1]; - методические указания [1]; Отчет по лабораторной работе подготовить в соответствии с методическими указаниями [4].	Вопросы по темам/разделам дисциплины Тестовое задание. Защита лабораторных работ
	Итого		73.6 5		

## 6.2. Тематика и задания для практических занятий

### 6.3. Тематика и задания для лабораторных занятий

Лабораторная работа 1. Загрузка данных в DataFrame Pandas, первичная обработка, фильтрация, преобразование типов и форматов. Виды входных и выходных файлов.

Лабораторная работа 2. Аналитическая обработка данных в DataFrame Pandas, функции для обработки числовых данных в колонках и столбцах, функции для перебора данных, группировка числовых данных и построение сводного отчета.

Лабораторная работа 3. Загрузка исходных данных для машинного обучения, их первичная обработка и приведение к числовому виду и масштабирование. Разделение выборки на обучающую и тестируемую

Лабораторная работа 4. Выбор модели для бинарной классификации. Построение графика ROC и диаграмм для confusion matrix и binary classification report.

### 6.4. Тематика и методические указания по выполнению курсовой

## работы

### 7. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

#### а) основная:

1. Макшанов, А. В. Большие данные. Big Data : учебник для вузов / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — Санкт-Петербург : Лань, 2021. — 188 с. — ISBN 978-5-8114-6810-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165835> (дата обращения: 24.06.2021).

2. Адлер, Ю. П. Статистическое управление процессами. «Большие данные» : учебное пособие / Ю. П. Адлер, Е. А. Черных. - Москва : Изд. Дом МИСиС, 2016. - 52 с. - ISBN 978-5-87623-969-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1232190> (дата обращения: 24.06.2021).

#### б) дополнительная:

1. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ – МГСУ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/145102> (дата обращения: 24.06.2021).

#### в) методические указания

1. Информационный анализ больших данных в системах управления: метод. указания / сост. В.В. Олоничев. URL: <ftp://amt401/pub/ovv/>.

### 8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

Информационно-образовательные ресурсы:

Основы технологии MPI на примерах. URL: <https://pro-prof.com/archives/4386>

Электронные библиотечные системы:

1. ЭБС «Лань».
2. ЭБС «Университетская библиотека online».
3. ЭБС Znanium.com.
4. Научная электронная библиотека ELIBRARY.

### 9. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Лекционная аудитория должна быть оснащена презентационным оборудованием (персональный компьютер, мультимедийный проектор, экран).

Компьютерный класс Б-403

лицензионное проприетарное программное обеспечение не используется.

