


Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

УТВЕРЖДЕНО
решением Ученого совета факультета математики,
информационных и авиационных технологий
от «16» июня 20__20__ г., протокол № 5/20
Председатель _____ Волков М.А.
(подпись, расшифровка подписи)
«16» июня 20__20__ г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина	Методы и системы обработки больших данных
Факультет	Факультет математики, информационных и авиационных технологий
Кафедра	Информационных технологий
Курс	3

Направление (специальность) 02.03.03. — «Математическое обеспечение и администрирование информационных систем»
код направления (специальности), полное наименование

Направленность (профиль/специализация) Технология программирования
полное наименование

Форма обучения очная

очная, заочная, очно-заочная (указать только те, которые реализуются)

Дата введения в учебный процесс УлГУ: « 01 » _____ сентября 20__20__ г.

Программа актуализирована на заседании кафедры: протокол № _____ от _____ 20__ г.

Программа актуализирована на заседании кафедры: протокол № _____ от _____ 20__ г.

Программа актуализирована на заседании кафедры: протокол № _____ от _____ 20__ г.

Сведения о разработчиках:

ФИО	Кафедра	Должность, ученая степень, звание
Волков Максим Анатольевич	Информационных технологий	Зав. кафедрой, к.ф.-м.н., доцент
Котков Александр Валерьевич	Информационных технологий	Ассистент

СОГЛАСОВАНО	СОГЛАСОВАНО
Заведующий кафедрой информационных технологий, реализующей дисциплину	Заведующий выпускающей кафедрой информационных технологий
( / М.А.Волков / Подпись ФИО « 16 » июня 20__20__ г.	( / М.А.Волков / Подпись ФИО « 16 » июня 20__20__ г.

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целями освоения дисциплины «Методы и системы обработки больших данных» являются изучение основных технологий решения задач обработки больших по объему данных, умение применять методы анализа больших данных на практике и реализовывать приложения для аналитики больших данных.

Задачи освоения дисциплины:

- изучить задачи классификации и кластеризации больших объемов данных;
- изучить критерии аналитических задач, решение которых предпочтительно с использованием технологий Big Data;
- изучить интеллектуальные системы для решения аналитических задач;
- сформировать навыки работы с большими массивами данных;
- изучить технологии и программные средства обработки больших данных и методы машинного обучения для решения прикладных задач;
- изучить языки программирования для работы с большими объемами данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина «Методы и системы обработки больших данных» является дисциплиной по выбору Блока 1 Дисциплины (модули) Основной Профессиональной Образовательной Программы по направлению подготовки 02.03.03. – Математическое обеспечение и администрирование информационных систем.

Для изучения данной дисциплины необходимы знания основных понятий и методов математики, информатики и программирования, полученные в ходе изучения дисциплин: Технология разработки программного обеспечения, Высокоуровневые методы информатики и программирования, Операционные системы и оболочки, Программирование в среде Windows, Базы данных.

Дисциплина закладывает информационные знания необходимые для изучения таких курсов, как: Системы реального времени, Методы машинного обучения, Параллельное программирование, Системы искусственного интеллекта, Программирование для Интернет, Современные системы автоматизации разработки информационных систем, а также при прохождении практики и подготовке к государственной итоговой аттестации.

3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОПОП

Код и наименование реализуемой компетенции	Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций
ОПК-5. Способен устанавливать и сопровождать программное обеспечение для информационных систем и баз данных, в том числе отечественного производства.	Знать: виды программного обеспечения для обработки больших объемов данных; Уметь: устанавливать и сопровождать программное обеспечение для обработки больших объемов данных; Владеть: навыками работы с программным обеспечением для обработки больших объемов данных;
ПК-2. Способен использовать основные методы и средства автоматизации проектирования, реализации, испытаний и	Знать: основные понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных; Уметь: применять методы обработки больших объемов

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

оценки качества при создании конкурентоспособного программного продукта и программных комплексов, а также способен использовать методы и средства автоматизации, связанные с сопровождением, администрированием и модернизацией программных продуктов и программных комплексов.	данных и использовать программные средства с технологией Big Data при решении практических задач; Владеть: понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов; навыками применения современных методов разработки технического, информационного и алгоритмического обеспечения Big Data.
ПК-6. Способен принимать участие в управлении работами по созданию (модификации) и сопровождению ПО, программных систем и комплексов	Знать: основные методы интеллектуального анализа больших данных и машинного обучения, основные принципы поиска, сбора, очистки, хранения, обработки, анализа и визуализации больших данных; Уметь: находить, собирать и хранить большие объемы данных; применять модели машинного обучения для решения практических задач в различных областях; Владеть: профессиональной терминологией в области больших данных и машинного обучения; навыками практического применения методов интеллектуального анализа больших данных и машинного обучения.

4. ОБЩАЯ ТРУДОЕМКОСТЬ ДИСЦИПЛИНЫ

4.1. Объем дисциплины в зачетных единицах (всего): 2 з.е.

4.2. Объем дисциплины по видам учебной работы (в часах): 72 часа

Вид учебной работы	Количество часов (форма обучения: очная)	
	Всего по плану	в т.ч. по семестрам
		б
1	2	3
Контактная работа обучающихся с преподавателем	36/36*	36/36*
Аудиторные занятия:	36/36*	36/36*
Лекции	18/18*	18/18*
практические и семинарские занятия	-	-
лабораторные работы (лабораторный практикум)	18/18*	18/18*
Самостоятельная работа	36	36
Текущий контроль (количество и вид: контрольн. работа, коллоквиум, реферат)	Проверка лабораторных работ, тестирование	Проверка лабораторных работ, тестирование
Курсовая работа	-	-
Виды промежуточной аттестации (экзамен, зачет)	зачет	зачет

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

Всего часов по дисциплине	72	72
---------------------------	----	----

*Количество часов работы ППС с обучающимися в дистанционном формате с применением электронного обучения

В случае необходимости использования в учебном процессе частично/исключительно дистанционных образовательных технологий в таблице через слеш указывается количество часов работы ППС с обучающимися для проведения занятий в дистанционном формате с применением электронного обучения.

4.3. Содержание дисциплины (модуля.) Распределение часов по темам и видам учебной работы:

Форма обучения: очная

Название разделов и тем	Всего	Виды учебных занятий					Форма текущего контроля знаний
		Аудиторные занятия			Занятия в интерактивной форме	Самостоятельная работа	
		Лекции	Практические занятия, семинары	Лабораторные работы, практикумы			
1	2	3	4	5	6	7	
Раздел 1. Большие данные и машинное обучение							
Тема 1. Интеллектуальный анализ данных, большие данные, машинное обучение.	15	4		4		6	Устный опрос, тестирование, защита лабораторных работ
Тема 2. Принципы анализа текстовой и графической информации, эмоциональной окраски текстов.	15	4		4		6	Устный опрос, тестирование, защита лабораторных работ
Раздел 2. Искусственный интеллект							
Тема 3. Понятие искусственного интеллекта и области его применения.	14	4		4		8	Устный опрос, тестирование, защита лабораторных работ
Тема 4. Технологии интеллектуального анализа	16	4		4		8	Устный опрос, тестирование,

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

данных.							защита лабораторных работ
Тема 5. Этапы создания и сферы применения экспертных систем.	12	2		2		8	Устный опрос, тестирование, защита лабораторных работ
Итого	72	18	-	18	-	36	

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Раздел 1. Большие данные и машинное обучение.

Тема 1. Интеллектуальный анализ данных, большие данные, машинное обучение. Методы и задачи интеллектуального анализа данных, машинного обучения и обработки больших данных. Области применения методов и технологий интеллектуального анализа данных, машинного обучения и обработки больших данных. Примеры задач машинного обучения: поиск информации в интернете, распознавание изображений, лиц, эмоций, пола, возраста, распознавание речи, языка, эмоциональной окраски текстов, прогнозирование продаж, прогнозирование оттока клиентов, кредитный скоринг, рекомендательные системы и др. Основные характеристики больших данных и их влияние на сбор, хранение, обработку и анализ данных (4V). Критерии аналитических задач, решение которых предпочтительно с использованием технологий BigData.

Тема 2. Принципы анализа текстовой и графической информации, эмоциональной окраски текстов. Принципы создания рекомендательных систем. Интеллектуальные сервисы и чат-боты. Перспективы развития систем обработки больших данных и машинного обучения. Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей.

Раздел 2. Искусственный интеллект.


Тема 3. Понятие искусственного интеллекта и области его применения. Признаки интеллектуальности информационных систем. Структура исследований в области искусственного интеллекта. Основные классы интеллектуальных информационных систем. Знания как особая форма информации. Методы и средства представления знаний. Модели знаний. Системы представления знаний и базы знаний. Приобретение знаний от экспертов. Извлечение знаний из документов. Технологии OLAP и многомерные модели данных.

Тема 4. Технологии интеллектуального анализа данных (Data Mining). Согласование и интеграция знаний. Экспертные системы, их виды, области использования.

Тема 5. Этапы создания и сферы применения экспертных систем. Нейросетевые технологии. Проблемы, решаемые искусственными нейронными сетями. Основные направления применения нейросетевых технологий в экономике.

6. ТЕМЫ ПРАКТИЧЕСКИХ И СЕМИНАРСКИХ ЗАНЯТИЙ

Данный вид работы не предусмотрен УП.

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

7. ЛАБОРАТОРНЫЕ РАБОТЫ, ПРАКТИКУМЫ


- Лабораторная работа № 1. Предварительный анализ больших наборов данных.
Лабораторная работа № 2. Визуальный анализ данных.
Лабораторная работа № 3. Обучение с учителем. Методы классификации.
Лабораторная работа № 4. Линейные модели классификации и регрессии.
Лабораторная работа № 5. Задача восстановления регрессии.

8. ТЕМАТИКА КУРСОВЫХ, КОНТРОЛЬНЫХ РАБОТ, РЕФЕРАТОВ

Данный вид работы не предусмотрен УП

9. ПЕРЕЧЕНЬ ВОПРОСОВ К ЗАЧЕТУ

1. Уровни понимания. Методы решения задач.
2. Решение задач методом поиска в пространстве состояний.
3. Фреймы. Исчисления предикатов.
4. Системы продукций. Семантические сети.
5. Нечеткая логика.
6. Алгоритмы эвристического поиска.
7. Поиск решений на основе исчисления предикатов.
8. Переход от Базы данных к Базе знаний. Особенности знаний.
9. Генетический алгоритм.
10. Стратегия решений организации поиска.
11. Назначение экспертных систем.
12. Структура экспертных систем.
13. Этапы разработки экспертных систем.
14. Представление знаний в экспертных системах.
15. Методы работа со знаниями.
16. Основная модель нейросетевой технологии.
17. Методы извлечения знаний
18. Цепи Маркова
19. Вероятностный подход
20. Случайный лес
21. Машинное обучение. Основные понятия, связи, теоремы
22. Обучение с учителем: регрессия
23. Обучение с учителем: классификация
24. Обучение без учителя: кластеризация
25. Обучение без учителя: уменьшение размерности
26. Обучение с подкреплением
27. Задачи, которые решает машинное обучение
28. Выбор методологии для проекта с машинным обучением. Примеры, причины.
29. SMART-цель
30. Этапы решения задач МО
31. Метод имитация отжига
32. Метод роя частиц
33. Генетический алгоритм
34. Модель МО. Выбор модели. Процесс обучения. Валидация
35. Тестирование МО
36. Проблемы разработки МО

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

10. САМОСТОЯТЕЛЬНАЯ РАБОТА ОБУЧАЮЩИХСЯ

Форма обучения: очная

Название разделов и тем	Вид самостоятельной работы	Объем в часах	Форма контроля
Тема 1. Интеллектуальный анализ данных, большие данные, машинное обучение.	чтение основной и дополнительной литературы, самостоятельное изучение материала по литературным источникам, выполнение лабораторных работ	6	Устный опрос, тестирование, защита лабораторных работ
Тема 2. Принципы анализа текстовой и графической информации, эмоциональной окраски текстов.	чтение основной и дополнительной литературы, самостоятельное изучение материала по литературным источникам, выполнение лабораторных работ	6	Устный опрос, тестирование, защита лабораторных работ
Тема 3. Понятие искусственного интеллекта и области его применения.	чтение основной и дополнительной литературы, самостоятельное изучение материала по литературным источникам, выполнение лабораторных работ	8	Устный опрос, тестирование, защита лабораторных работ
Тема 4. Технологии интеллектуального анализа данных.	чтение основной и дополнительной литературы, самостоятельное изучение материала по литературным источникам, выполнение лабораторных работ	8	Устный опрос, тестирование, защита лабораторных работ
Тема 5. Этапы создания и сферы применения экспертных систем.	чтение основной и дополнительной литературы, самостоятельное изучение материала по литературным источникам, выполнение лабораторных работ	8	Устный опрос, тестирование, защита лабораторных работ

11. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

а) Список рекомендуемой литературы

основная

1. Воронов В.И. Data Mining - технологии обработки больших данных : учебное пособие / Воронов В.И., Воронова Л.И., Усачев В.А.. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/81324.html>
2. Воронова Л.И. Big Data. Методы и средства анализа : учебное пособие / Воронова Л.И., Воронов В.И.. — Москва : Московский технический университет связи и информатики, 2016. — 33 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/61463.html>

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

- среды программирования на языках Python/Java.

в) Профессиональные базы данных, информационно-справочные системы

1. Электронно-библиотечные системы:

1.1. IPRbooks : электронно-библиотечная система : сайт / группа компаний Ай Пи Ар Медиа. - Саратов, [2020]. – URL: <http://www.iprbookshop.ru>. – Режим доступа: для зарегистрир. пользователей. - Текст : электронный.

1.2. ЮРАЙТ : электронно-библиотечная система : сайт / ООО Электронное издательство ЮРАЙТ. – Москва, [2020]. - URL: <https://www.biblio-online.ru>. – Режим доступа: для зарегистрир. пользователей. - Текст : электронный.

1.3. Консультант студента : электронно-библиотечная система : сайт / ООО Политехресурс. – Москва, [2020]. – URL: http://www.studentlibrary.ru/catalogue/switch_kit/x2019-128.html. – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

1.4. Лань : электронно-библиотечная система : сайт / ООО ЭБС Лань. – Санкт-Петербург, [2020]. – URL: <https://e.lanbook.com>. – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

1.5. Znanium.com : электронно-библиотечная система : сайт / ООО Знаниум. - Москва, [2020]. - URL: <http://znanium.com>. – Режим доступа : для зарегистрир. пользователей. - Текст : электронный.

1.6. Clinical Collection : коллекция для медицинских университетов, клиник, медицинских библиотек // EBSCOhost : [портал]. – URL: <http://web.a.ebscohost.com/ehost/search/advanced?vid=1&sid=e3ddfb99-a1a7-46dd-a6eb-2185f3e0876a%40sessionmgr4008>. – Режим доступа : для авториз. пользователей. – Текст : электронный.

2. КонсультантПлюс [Электронный ресурс]: справочная правовая система. /ООО «Консультант Плюс» - Электрон. дан. - Москва : КонсультантПлюс, [2020].

3. Базы данных периодических изданий:

3.1. База данных периодических изданий : электронные журналы / ООО ИВИС. - Москва, [2020]. – URL: <https://dlib.eastview.com/browse/udb/12>. – Режим доступа : для авториз. пользователей. – Текст : электронный.

3.2. eLIBRARY.RU: научная электронная библиотека : сайт / ООО Научная Электронная Библиотека. – Москва, [2020]. – URL: <http://elibrary.ru>. – Режим доступа : для авториз. пользователей. – Текст : электронный

3.3. «Grebennikon» : электронная библиотека / ИД Гребенников. – Москва, [2020]. – URL: <https://id2.action-media.ru/Personal/Products>. – Режим доступа : для авториз. пользователей. – Текст : электронный.

4. Национальная электронная библиотека : электронная библиотека : федеральная государственная информационная система : сайт / Министерство культуры РФ ; РГБ. – Москва, [2020]. – URL: <https://нэб.рф>. – Режим доступа : для пользователей научной библиотеки. – Текст : электронный.

5. SMART Imagebase // EBSCOhost : [портал]. – URL: <https://ebsco.smartimagebase.com/?TOKEN=EBSCO-1a2ff8c55aa76d8229047223a7d6dc9c&custid=s6895741>. – Режим доступа : для авториз. пользователей. – Изображение : электронные.

6. Федеральные информационно-образовательные порталы:

6.1. Единое окно доступа к образовательным ресурсам : федеральный портал / учредитель ФГАОУ ДПО ЦРГОП и ИТ. – URL: <http://window.edu.ru/>. – Текст : электронный.

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

6.2. Российское образование : федеральный портал / учредитель ФГАОУ ДПО ЦРГОП и ИТ. – URL: <http://www.edu.ru>. – Текст : электронный.


7. Образовательные ресурсы УлГУ:

7.1. Электронная библиотека УлГУ : модуль АБИС Мега-ПРО / ООО «Дата Экспресс». – URL: <http://lib.ulsu.ru/MegaPro/Web>. – Режим доступа : для пользователей научной библиотеки. – Текст : электронный.

7.2. Образовательный портал УлГУ. – URL: <http://edu.ulsu.ru>. – Режим доступа : для зарегистр. пользователей. – Текст : электронный.

Согласовано:

Заместитель начальника УИТиТ /Клочкова А.В. _____

 08.06.2020

12. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Аудитории для проведения лекций, семинарских занятий, для выполнения лабораторных работ и практикумов, для проведения текущего контроля и промежуточной аттестации, курсового проектирования, групповых и индивидуальных консультаций.

Аудитории укомплектованы специализированной мебелью, учебной доской. Аудитории для проведения лекций оборудованы мультимедийным оборудованием для предоставления информации большой аудитории. Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде, электронно-библиотечной системе.

13. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ДЛЯ ОБУЧАЮЩИХСЯ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Обучение по ОПОП ВО обучающихся с ограниченными возможностями здоровья осуществляется с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся. Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и отдельно. В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) могут предлагаться одни из следующих вариантов восприятия информации с учетом их индивидуальных психофизических особенностей:

– для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); в печатной форме на языке Брайля; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации.

– для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации.

– для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации».

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

В случае необходимости использования в учебном процессе частично/исключительно дистанционных образовательных технологий, организация работы ППС с обучающимися с ОВЗ и инвалидами предусматривается в электронной информационно-образовательной среде с учетом их индивидуальных психофизических особенностей.

Разработчик _____ Зав.кафедрой ИТ Волков М.А.
подпись ФИО

Разработчик _____ ассистент Котков А.А.
подпись ФИО