

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ им. В.Г.ШУХОВА»**
(БГТУ им. В.Г. Шухова)



УТВЕРЖДАЮ
Директор института

[Handwritten signature]
« 20 » / 05 2021 г.

РАБОЧАЯ ПРОГРАММА
дисциплины

Методы анализа данных

направление подготовки:

09.03.04 «Программная инженерия»

Направленность программы (профиль):

Разработка программно-информационных систем

Квалификация

Бакалавр

Форма обучения

Очная

Институт энергетики, информационных технологий и управляющих систем

Кафедра программного обеспечения вычислительной техники и
автоматизированных систем

Белгород 2021

Рабочая программа составлена на основании требований:

- Федерального государственного образовательного стандарта высшего образования – бакалавриат по направлению подготовки 09.03.04 «Программная инженерия», утвержденного приказа Минобрнауки России от 19.09.2017 № 920
- учебного плана, утвержденного ученым советом БГТУ им. В.Г. Шухова в 2021 году.

Составитель : к.т.н., доцент  (ученая степень и звание, подпись) (Кабалянц П.С.)
(инициалы, фамилия)

Рабочая программа обсуждена на заседании кафедры

« 14 » 05 2021 г., протокол № 8

Заведующий кафедрой: к.т.н., доцент  (ученая степень и звание, подпись) (Поляков В.М.)
(инициалы, фамилия)


Рабочая программа согласована с выпускающей кафедрой программного обеспечения вычислительной техники и автоматизированных систем
(наименование кафедры/кафедр)

Заведующий кафедрой: к.т.н., доцент  (ученая степень и звание, подпись) (Поляков В.М.)
(инициалы, фамилия)

« 14 » 05 2021 г.

Рабочая программа одобрена методической комиссией института

« 20 » 05 2021 г., протокол № 9

Председатель к.т.н., доцент  (ученая степень и звание, подпись) (Семернин А.Н.)
(инициалы, фамилия)

1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Категория (группа) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Наименование показателя оценивания результата обучения по дисциплине
Разработка программного обеспечения	ПК-2. Способен использовать различные технологии разработки программного обеспечения автоматизированных систем	ПК-2.1. Анализирует и выбирает необходимую технологию разработки программного обеспечения для решения профессиональных задач	Знания
		ПК-2.2. Использует современные технологии разработки программного обеспечения для решения прикладных задач	Умения
		ПК-2.4. Применяет языки программирования различного уровня для написания компонентов программных продуктов	Навыки

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1. Компетенция ПК-2. Способен использовать различные технологии разработки программного обеспечения автоматизированных систем

Данная компетенция формируется следующими дисциплинами:

Стадия	Наименования дисциплины
1.	Архитектура вычислительных систем
2.	Алгоритмы и структуры данных
3.	Объектно-ориентированное программирование
4.	Компьютерная графика
5.	Методы анализа данных
6.	Теория информации
7.	Технологии Web-программирования
8.	Проектирование клиент-серверных приложений
9.	Параллельное программирование
10.	Программирование микроконтроллеров
11.	Основы искусственного интеллекта
12.	Безопасность программно-информационных систем
13.	Основы построения трансляторов
14.	Программирование на языке Java
15.	Программирование на языке Python
16.	Производственная преддипломная практика
17.	Выполнение и защита выпускной квалификационной работы

3. ОБЪЁМ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины составляет 4 зач. ед., 144 часа.

Дисциплина реализуется в рамках практической подготовки: 2 зач. единиц.

Форма промежуточной аттестации: экзамен.

Вид учебной работы	Всего часов	Семестр № 6
Общая трудоемкость дисциплины, час	144	144
Контактная работа (аудиторные занятия), в т.ч.:	55	55
лекции	17	17
лабораторные	34	34
практические		
групповые консультации в период теоретического обучения и промежуточной аттестации	4	4
Самостоятельная работа студентов, включая индивидуальные и групповые консультации, в том числе:	93	93
Курсовой проект		
Курсовая работа		
Расчетно-графическое задание		
Индивидуальное домашнее задание	9	9
Самостоятельная работа на подготовку к аудиторным занятиям (лекции, практические занятия, лабораторные занятия)	44	44
Экзамен	36	36

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Наименование тем, их содержание и объем

Курс 3 Семестр №6

№ п/п	Наименование раздела (краткое содержание)	Объем на тематический раздел по видам учебной нагрузки, час			
		Лекции	Практические занятия	Лабораторные занятия	Самостоятельная работа на подготовку к аудиторным
1	Типы данных и их предварительная обработка. Качественные и количественные данных, дискретные и непрерывные, ранжированные и номинальные. Случай бинарных данных. Задачи анализа данных для разных типов данных: случаи табличных данных, коллекций текстов и коллекций цифровых изображений. Предварительная обработка табличных данных: методы кодирования, шкалирования и нормирования данных. Преобразование данных. Объекты и признаки. Задача выбора описания объектов.	2	-	2	3
2	Многомерный линейный регрессионный анализ и нелинейная модель регрессии. Экспоненциальная модель регрессии. Регрессионная модель как модель обучения. Оценки качества регрессионной модели.	1	-	2	2
3	Задача кластеризации. Выбор расстояния между объектами и расстояния между кластерами. Агломеративные и дивизимные алгоритмы кластеризации. Теоретико-графовая модель агломеративной кластеризации. Алгоритмы кластеризации K-средних, K-медиан и K-центроидов. Особенности и недостатки алгоритмов кластеризации. Оценки качества кластеризации. Задача оптимальной визуализации результатов кластеризации: алгоритм SNE и его модификации.	3	-	6	3
4	Метод главных компонент и его модификации. Преобразование Карунена-Лоева, его связь с сингулярным разложением. Задача оптимального описания объектов. Проблема интерпретации полученных компонент. Факторный анализ.	2	-	4	2
5	Данные как многомерные временные ряды. Модели авторегрессии и модели скользящего среднего. Прогнозирование временных рядов. Модели временных рядов: тренд, случайная и периодические компоненты. Задача определения аномалий временного ряда.	2	-	2	2
6	Моделирование данных. EM-алгоритм. Задача восстановления распределения классов. Задача обучения по прецедентам. Примеры прикладных задач. Тематическое моделирование как пример работы EM-	1	-	2	7

	алгоритма.				
7	Задача классификации. Разметка данных. Постановка задачи, выбор модели. Проблема переобучения, внешние и внутренние критерии качества классификации. Метрические классификаторы, метод k-соседей и окно Э.Парзена.	2	-	4	7
8	Логический алгоритм классификации. Понятие закономерности и информативности. Разновидности закономерностей: конъюнкции пороговых предикатов (гиперпараллелепипед), синдромно правила, шары, гиперплоскости. Бинаризация признаков. Алгоритм разбиения области значений признака на информативные зоны. Решающие списки и деревья.	1	-	4	5
9	Линейные алгоритмы классификации. Метод опорных векторов. Оптимальная гиперплоскость, разделяющая классы. Понятие зазора между классами (margin). Случаи, когда возможно линейно разделить классы, и случаи, когда линейно разделить классы невозможно. Кусочно-линейная функция потерь. Задача квадратичного программирования и сопряженная задача. Понятие опорных векторов. Функция ядра. Байесовские методы классификации. Линейный дискриминантный анализ. Нейронные сети как язык описания линейных алгоритмов классификации.	2	-	6	5
10	Логистическая регрессия. Обзор методов классификации. Модель агентов и среды, обучение с подкреплением. Онлайн классификация и итерационная классификация.	1	-	2	8
	ВСЕГО	17	-	34	44

4.2. Содержание практических (семинарских) занятий

Практические занятия учебным планом не предусмотрены.

4.3. Содержание лабораторных занятий

Курс 3 Семестр №6

№ п/п	Наименование раздела дисциплины	Тема лабораторного занятия	К-во часов	Самостоятельная работа на подготовку к аудиторным занятиям
1	Типы данных. Задачи анализа данных.	Работа с основными библиотеками для обработки данных языка Python.	2	2
2	Многомерный линейный регрессионный анализ.	Лабораторная работа по регрессионному анализу.	2	2
3	Задача кластеризации.	Лабораторная работа на кластеризацию данных. Возможности библиотек scipy и skikit-learn для решения этой задачи.	6	2
4	Метод главных компонент.	Лабораторная работа на кластеризацию данных с использованием компонентного анализа.	4	2

5	Данные как временные ряды.	Лабораторная работа по прогнозированию значений временного ряда.	2	4
6	Моделирование данных. Тематическое моделирование.	Работа с текстами. Возможности библиотеки nltk языка Python.	2	4
7	Метрические классификаторы.	Лабораторная работа на классификацию табличных данных с использованием окна Э.Парзена.	4	4
8	Логические алгоритмы классификации.	Лабораторная работа на классификацию табличных данных с использованием решающих деревьев.	4	4
9	Линейные алгоритмы классификации.	Простая распознающая нейронная сеть.	6	6
10	Логистическая регрессия.	Лабораторная работа на бинарную классификацию табличных данных с использованием логистической регрессии.	2	4
ИТОГО			34	34

4.4. Содержание курсового проекта/работы

Не предусмотрено учебным планом.

4.5. Содержание расчетно-графического задания, индивидуальных домашних заданий

Учебным планом предусмотрено выполнение 1 ИДЗ, на которое отводится 9 часов самостоятельной работы студента.

Тематика ИДЗ:

1. Многомерный линейный регрессионный анализ.
2. Кластеризация данных.
3. Классификация табличных данных.

5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1. Реализация компетенций

1. Компетенция ПК-2. Способен использовать различные технологии разработки программного обеспечения автоматизированных систем

Наименование индикатора достижения компетенции	Используемые средства оценивания
ПК-2.1. Анализирует и выбирает необходимую технологию разработки программного обеспечения для решения профессиональных задач	защита лабораторной работы; защита ИДЗ, экзамен
ПК-2.2. Использует современные технологии разработки программного обеспечения для решения прикладных задач	защита лабораторной работы, защита ИДЗ, экзамен
ПК-2.4. Применяет языки программирования различного уровня для написания компонентов программных продуктов	защита лабораторной работы, защита ИДЗ, экзамен

5.2. Типовые контрольные задания для промежуточной аттестации

5.2.1. Перечень контрольных вопросов (типовых заданий) для экзамена

№ п/п	Наименование раздела дисциплины	Содержание вопросов (типовых заданий)
1	Типы данных. Задачи анализа данных. (ПК-2.1)	Основные типы данных. Матрица ковариаций. Многомерный нормальный закон распределения. Что такое шкалирование, кодирование и нормирование данных?
2	Многомерный линейный регрессионный анализ. (ПК-2.1)	Постановка задачи множественного линейного регрессионного анализа. Оценивание коэффициентов регрессии по методу наименьших квадратов. Исследование значимости коэффициентов регрессии. Особенности нелинейной модели регрессии.
3	Задача кластеризации. (ПК-2.1)	Типы расстояний между объектами и особенности их использования. Расстояния между кластерами. Агломеративные и дивизимные алгоритмы кластеризации. Алгоритмы кластеризации К-средних, К-медиан и К-центроидов. Оценки качества кластеризации. Оптимальная визуализация результатов кластеризации: алгоритм SNE и его модификации.
4	Метод главных компонент. (ПК-2.1)	Постановка задачи факторного анализа. Стандартизованная матрица. Метод главных компонент.
5	Данные как временные ряды. (ПК-2.2)	Временной ряд и задача его анализа как задача обучения. Модель скользящего среднего. Авторегрессионная модель. Различные модели тренда.
6	Моделирование данных. Тематическое моделирование. (ПК-2.2)	Оценка плотности распределения класса. Смеси распределений классов. EM-алгоритм. Модель порождения коллекции документов. Использование EM-алгоритма для задачи тематического моделирования. Параметрический анализ модели Latent Dirichlet Allocation.
7	Метрические классификаторы. (ПК-2.2)	Задача классификации. Разметка данных. Проблема переобучения, внешние и внутренние критерии качества классификации. Метрические классификаторы, метод k-соседей и окно Э.Парзена.

8	Логические алгоритмы классификации. (ПК-2.3)	Понятие закономерности и информативности. Разновидности закономерностей: конъюнкции пороговых предикатов (гиперпараллелепипед), синдромно правила, шары, гиперплоскости. Бинаризация признаков. Алгоритм разбиения области значений признака на информативные зоны. Решающие списки и деревья.
9	Линейные алгоритмы классификации. (ПК-2.3)	Задача дискриминантного анализа. Линейная дискриминантная функция. Дискриминантная задача для числа классов больше двух. Метод опорных векторов. Нейронная сеть как модель линейной классификации.
10	Логистическая регрессия. (ПК-2.3)	Алгоритм классификации с помощью логистической регрессии. Модель обучения с подкреплением. Алгоритмы инкрементного и онлайн-обучения.

5.3. Типовые контрольные задания (материалы) для текущего контроля в семестре

Текущий контроль осуществляется в течение семестра в форме защиты лабораторных работ и защиты ИДЗ.

Защита лабораторной работы проводится в форме устного опроса студента и направлена на проверку степени усвоения материала и понимания теоретических сведений, используемых в процессе выполнения работы; для защиты необходимо представить в печатной (рукописной) форме отчет по лабораторной работе, выполненный самостоятельно и в соответствии со всеми требованиями, приведёнными в методических указаниях к выполнению лабораторных работ. Примерный перечень контрольных вопросов для защиты лабораторных работ приведен в таблице:

Тематика лабораторной работы	Контрольные вопросы
1. Работа с основными библиотеками для обработки данных языка Python.	1. Как подключаются библиотеки в языке Python? 2. Какая библиотека в языке Python используется для матричных вычислений? 3. Какая библиотека в языке Python используется для статистического анализа? 4. Какие библиотеки в языке Python используются для анализа данных?
2. Лабораторная работа по регрессионному анализу.	1. Привести условия применения линейной регрессионной модели. 2. Как связаны экспоненциальная и линейная модели регрессии? 3. Как проверить адекватность модели регрессии? 4. Назовите основные модели регрессии.

3. Лабораторная работа на кластеризацию данных. Возможности библиотек scіru и skіkit-learn для решения этой задачи.	<ol style="list-style-type: none"> 1. Что необходимо минимизировать в методе К-средних? 2. В чем идея иерархической кластеризации? 3. Какие метрики оценки качества кластеризации Вы знаете?
4. Лабораторная работа на кластеризацию данных с использованием компонентного анализа.	<ol style="list-style-type: none"> 1. Что такое главные компоненты? 2. В чем состоит метод главных компонент? 3. Как использовать метод главных компонент при кластеризации данных?
5. Лабораторная работа по прогнозированию значений временного ряда.	<ol style="list-style-type: none"> 1. Назовите основные методы сглаживания временного ряда? 2. Как выбрать степень полиномиального тренда? 3. Какие условия существования тренда у временного ряда? 4. Как оценить адекватность авторегрессионной модели?
6. Работа с текстами. Возможности библиотеки nltk языка Python.	<ol style="list-style-type: none"> 1. В чем состоит задача кластеризации? 2. Какое расстояние между объектами используется в алгоритме кластеризации К-средних? 3. Какие основные недостатки алгоритма К-средних?
7. Лабораторная работа на классификацию табличных данных с использованием окна Э.Парзена.	<ol style="list-style-type: none"> 1. Каким условиям должны удовлетворять данные для корректности применения наивного байесовского классификатора? 2. Как оценить качество классификации? 3. В чем состоит алгоритм обучения наивного байесовского классификатора?
8. Лабораторная работа на классификацию табличных данных с использованием решающих деревьев.	<ol style="list-style-type: none"> 1. В чем состоит постановка задачи парсинга? 2. Как связана задача парсинга с регулярными выражениями? 3. Какие библиотеки языка Python связаны с задачей парсинга?
9. Простая распознающая нейронная сеть.	<ol style="list-style-type: none"> 1. Приведите несколько основных архитектур нейронных сетей. 2. Приведите примеры различных функций активаций нейронов. 3. В чем состоит обучение нейронной сети?
10. Лабораторная работа на бинарную классификацию табличных данных с использованием логистической регрессии.	<ol style="list-style-type: none"> 1. Что означает переобучение? 2. Какие метрики используются для оценки качества классификации? 3. В чем особенности использования логистической регрессии?

Критерии оценки лабораторной работы: лабораторная работа считается защищенной, если студент выполнил задание к работе полностью и во время устного опроса по работе правильно ответил на заданные преподавателем дополнительные вопросы.

5.4. Описание критериев оценивания компетенций и шкалы оценивания

При промежуточной аттестации в форме экзамена используется следующая шкала оценивания: 2 – неудовлетворительно, 3 – удовлетворительно, 4 – хорошо, 5 – отлично.

Критериями оценивания достижений показателей являются:

Наименование показателя оценивания результата обучения по дисциплине	Критерий оценивания
Знания	Знание терминов, определений, понятий анализа данных
	Знание основных закономерностей, соотношений, принципов работы истоков анализа данных
	Объем освоенного материала
	Полнота ответов на вопросы
	Четкость изложения и интерпретации знаний
Умения	Умение решать стандартные профессиональные задачи с применением методов анализа данных
	Умение использовать теоретические знания для выбора методики решения профессиональных задач
	Умение проверять решение и анализировать результаты
Навыки	Владение навыками теоретического и экспериментального исследования объектов профессиональной деятельности в области анализа данных
	Качество выполнения исследований объектов профессиональной деятельности в области анализа данных
	Самостоятельность выполнения исследований объектов профессиональной деятельности в области анализа данных

Оценка преподавателем выставляется интегрально с учётом всех показателей и критериев оценивания.

Оценка сформированности компетенций по показателю Знания.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Знание терминов, определений, понятий анализа данных	Не знает терминов и определений	Знает термины и определения, но допускает неточности формулировок	Знает термины и определения	Знает термины и определения, может корректно сформулировать их самостоятельно
Знание основных закономерностей, соотношений, принципов работы истоков анализа данных	Не знает основные закономерности и соотношения, принципы построения знаний	Знает основные закономерности, соотношения, принципы построения знаний	Знает основные закономерности, соотношения, принципы построения знаний, их интерпретирует и использует	Знает основные закономерности, соотношения, принципы построения знаний, может самостоятельно их получить и использовать
Объем освоенного материала	Не знает значительной части материала дисциплины	Знает только основной материал дисциплины, не усвоил его деталей	Знает материал дисциплины в достаточном объеме	Обладает твердым и полным знанием материала дисциплины, владеет дополни-

				тельными знаниями
Полнота ответов на вопросы	Не дает ответы на большинство вопросов	Дает неполные ответы на все вопросы	Дает ответы на вопросы, но не все - полные	Дает полные, развернутые ответы на поставленные вопросы
Четкость изложения и интерпретации знаний	Излагает знания без логической последовательности	Излагает знания с нарушениями в логической последовательности	Излагает знания без нарушений в логической последовательности	Излагает знания в логической последовательности, самостоятельно их интерпретируя и анализируя
	Не иллюстрирует изложение поясняющими схемами, рисунками и примерами	Выполняет поясняющие схемы и рисунки небрежно и с ошибками	Выполняет поясняющие рисунки и схемы корректно и понятно	Выполняет поясняющие рисунки и схемы точно и аккуратно, раскрывая полноту усвоенных знаний
	Неверно излагает и интерпретирует знания	Допускает неточности в изложении и интерпретации знаний	Грамотно и по существу излагает знания	Грамотно и точно излагает знания, делает самостоятельные выводы

Оценка сформированности компетенций по показателю Умения.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Умение решать стандартные профессиональные задачи с применением методов анализа данных	Не умеет решать стандартные профессиональные задачи с применением методов анализа данных	Допускает неточности в решении стандартных профессиональных задач с применением методов анализа данных	Умеет решать стандартные профессиональные задачи с применением методов анализа данных	Безошибочно решает стандартные профессиональные задачи с применением методов анализа данных
Умение использовать теоретические знания для выбора методики решения профессиональных задач	Не умеет использовать теоретические знания для выбора методики решения профессиональных задач	Использование теоретических знаний для выбора методики решения профессиональных задач вызывает затруднения	Умеет использовать теоретические знания для выбора методики решения профессиональных задач	Умело использует теоретические знания для выбора методики решения профессиональных задач
Умение проверять решение и анализировать результаты	Не умеет проверять решение и анализировать результаты	Умеет проверять решение некоторых задач	Умеет проверять решение некоторых задач и анализировать результаты	Умеет проверять решение и анализировать результаты

Оценка сформированности компетенций по показателю Навыки.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Владение навыками теоретического и экспериментального исследования объектов профессиональной деятельности в области анализа данных	Не владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности	Не достаточно хорошо владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности	Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности	Профессионально владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности
Качество выполнения исследований объектов профессиональной деятельности в области анализа данных	Не качественно выполняет исследования объектов профессиональной деятельности, допускает грубые ошибки	Не достаточно качественно выполняет исследования объектов профессиональной деятельности, допускает и исправляет ошибки с посторонней помощью	Не достаточно качественно выполняет исследования объектов профессиональной деятельности, допускает и исправляет ошибки самостоятельно	Качественно выполняет исследования объектов профессиональной деятельности
Самостоятельность выполнения исследований объектов профессиональной деятельности в области анализа данных	Не может самостоятельно выполнять исследования объектов профессиональной деятельности	Выполняет исследования объектов профессиональной деятельности с посторонней помощью	При выполнении исследований объектов профессиональной деятельности иногда требуется посторонняя помощь	Самостоятельно выполняет исследования объектов профессиональной деятельности

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

6.1. Материально-техническое обеспечение

№	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
1.	Учебная аудитория для проведения лекционных занятий	Специализированная мебель. Мультимедийная установка, экран, доски
2.	Учебная аудитория для проведения лабораторных занятий	Специализированная мебель. Компьютеры на базе процессоров Intel или AMD.
3.	Читальный зал библиотеки для самостоятельной работы	Специализированная мебель. Компьютерная техника, подключенная к сети интернет и имеющая доступ в электронно-образовательную среду

6.2. Лицензионное и свободно распространяемое программное обеспечение

№	Перечень лицензионного программного обеспечения.	Реквизиты подтверждающего документа
1.	Python 3.8 (CPython)	Свободно распространяемое ПО
2.	Anaconda (дистрибутив Python)	Свободно распространяемое ПО
3.	ОС Ubuntu 20.1	Свободно распространяемое ПО
4.	Google Chrome	Свободно распространяемое ПО
5.	Mozilla Firefox	Свободно распространяемое ПО
6.	Интегрированная среда разработки Microsoft Visual Studio	Подписка Microsoft Imagine Premium id: 6f22ecb4-6882-420b-a39b-afba0ace820c.

6.3. Перечень учебных изданий и учебно-методических материалов

1. Аггарвал Ч. Нейронные сети и глубокое обучение: учебный курс. - СПб. : ООО "Диалектика", 2020. - 752 с.
2. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: классификация и снижение размерности. Факторный, дискриминантный и кластерный анализ. - М. 1989.
3. Барсегян А.А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.
4. Брусенцев А. Г. Анализ данных и процессов. Ч. 1: Методы статистического анализа данных. - Белгород : Изд-во БГТУ, 2017. - 62 с.
5. Грас Дж. Data Science. Наука о данных с нуля. — СПб.: БХВ-Петербург, 2017. — 336 с.
6. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы для экономистов и менеджеров. М., 2000.
7. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. – М.: Наука. – 1976. – 736 с.
8. Маккинли У. Python и анализ данных. — Саратов: Профобразование, 2017. — 482 с.
9. Федин Ф.О. Анализ данных. Часть 1. Подготовка данных к анализу [Электронный ресурс]: учебное пособие / Ф.О. Федин, Ф.Ф. Федин. — Электрон. текстовые данные. — М.: Московский городской педагогический университет, 2012. — 204 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/26444.html>
10. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining [Электронный ресурс] : учебное пособие / Ф.О. Федин, Ф.Ф. Федин. — Электрон. текстовые данные. — М. : Московский городской педагогический университет, 2012. — 308 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/26445.html>

6.4. Перечень интернет ресурсов, профессиональных баз данных, информационно-справочных систем

1. Электронная библиотека (на базе ЭБС «БиблиоТех») — Режим доступа: <http://ntb.bstu.ru>
2. Электронно-библиотечная система IPRbooks — Режим доступа: <http://www.iprbookshop.ru>
3. Электронно-библиотечная система «Университетская библиотека ONLINE» — Режим доступа: <http://www.biblioclub.ru/>
4. https://www.youtube.com/playlist?list=PLJb04_psp36AfotU - VWozLueaoD_GPZV5 – курс лекций по Python и анализу данных.
7. <https://kaggle.com/> - библиотека датасетов

8. <http://www.machinelearning.ru> - Информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных, профессора МФТИ, д.ф.-м.н. Воронцова К.В.
9. <https://dyakonov.org> - Блог профессора МГУ, д.ф.-м.н. Дьяконова А.Г.

7. УТВЕРЖДЕНИЕ РАБОЧЕЙ ПРОГРАММЫ

Рабочая программа утверждена на 2022 / 2023 учебный год
без изменений / с изменениями, дополнениями

Протокол № 10 заседания кафедры от «20» 05 2022 г.

Заведующий кафедрой _____ Поляков В.М.
подпись, ФИО

Директор института _____ Белоусов А.В.
подпись, ФИО