

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



РАБОЧАЯ ПРОГРАММА

дисциплины

Анализ больших данных

Направление подготовки

09.04.01 Информатика и вычислительная техника

Направленность программы

Интеллектуальные системы

Квалификация

Магистр

Форма обучения

очная

Институт энергетики, информационных технологий и управляющих систем

Кафедра программного обеспечения вычислительной техники и
автоматизированных систем

Белгород – 2022

Рабочая программа составлена на основании требований:

- Федерального государственного образовательного стандарта высшего образования по направлению 09.04.01 «Информатика и вычислительная техника», утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. N 918
- учебного плана, утвержденного ученым советом БГТУ им. В.Г. Шухова в 2022 году.

Составители: к.ф.-м.н.  (С.В.Зуев)
(ученая степень и звание, подпись) (инициалы, фамилия)


Рабочая программа обсуждена на заседании кафедры
Программного обеспечения вычислительной техники и
автоматизированных систем

« 7 » апреля 20 22 г., протокол № 5

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


Рабочая программа согласована на заседании кафедры
Программного обеспечения вычислительной техники и
автоматизированных систем

« 7 » апреля 20 22 г., протокол № 5

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

Рабочая программа одобрена методической комиссией института

« 28 » апреля 20 22 г., протокол № 8

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

1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Категория (группа) компетенций	Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Наименование показателя оценивания
	ПК-1 Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях	ПК-1.1. Использует принципы управления инфраструктурой коллективной разработки программных средств	Знания
		ПК-1.2. Осуществляет поиск и подбор ресурсов для организации разработки и эксплуатации программного обеспечения	Умения
		ПК-1.3. Руководит персоналом и прочими ресурсами для решения поставленной профессиональной задачи	Навыки

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1. Компетенция ПК-1

Данная компетенция формируется следующими дисциплинами, практиками.

Стадия	Наименования дисциплины
1.	Управление проектированием информационных систем
2.	Анализ больших данных
3.	Инструменты анализа данных
4.	Производственная преддипломная практика
5.	Государственная итоговая аттестация

3. ОБЪЕМ ДИСЦИПЛИНЫ

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

Форма промежуточной аттестации: экзамен.

Вид учебной работы	Всего часов	Семестр № 2
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, в том числе:	89	89
Курсовой проект		
Курсовая работа		
Расчетно-графическое задания		
Индивидуальное домашнее задание		
Самостоятельная работа по подготовке к аудиторным занятиям (лекции, практические занятия, лабораторные занятия)	53	53
Экзамен	36	36

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Наименование тем, их содержание и объем

Курс1 Семестр 2

№ п/п	Наименование раздела (краткое содержание)	Объем на тематический раздел по видам учебной нагрузки, час			
		Лекции	Практические занятия	Лабораторные занятия	Самостоятельная работа
1	Многомерный регрессионный анализ. Постановка основной задачи линейного регрессионного анализа. Метод наименьших квадратов оценивания коэффициентов регрессии и оценка их значимости. Интервальные оценки коэффициентов регрессии. Оптимальный выбор матрицы плана. Задача статистического прогноза.	3		8	10
2	Дисперсионный анализ. Постановка задачи. Однофакторный дисперсионный анализ. Понятие о многофакторном анализе.	4		6	10
3	Факторный анализ. Общая и матричная постановка задачи. Алгоритм метода главных компонент. Проблема интерпретации	4		6	9

	факторов.				
4	Дискриминантный анализ. Постановка задачи классификации. Задача классификации в случае двух классов. Линейное различающее правило. Задача классификации в случае, когда количество классов больше двух.	6		6	10
5	Временные ряды. Общие положения. Критерии случайности. Тренд и сезонность.	4		8	14
	ВСЕГО	17		34	53

4.2. Содержание практических (семинарских) занятий

Учебным планом не предусмотрены

4.3. Содержание лабораторных занятий

№ п/п	Наименование раздела дисциплины	Тема лабораторного занятия	Кол-во часов	Кол-во часов СРС
семестр №2				
1	Многомерный регрессионный анализ	Метод наименьших квадратов оценивания коэффициентов регрессии.	4	5
2	Многомерный регрессионный анализ	Задача статистического прогноза	4	5
3	Дисперсионный анализ	Однофакторный дисперсионный анализ	6	6
4	Факторный анализ	Алгоритм метода главных компонент.	6	6
5	Дискриминантный анализ	Классификация сердечных больных	6	6
6	Временные ряды	Прогнозирование уровня моря	8	7
ИТОГО:			34	35

4.4. Содержание курсового проекта/работы

Учебным планом не предусмотрены.

4.5. Содержание расчетно-графического задания, индивидуальных домашних заданий

Учебным планом не предусмотрены.

5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1. Реализация компетенций

1. Компетенция ПК-1. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях.

Наименование индикатора достижения компетенции	Используемые средства оценивания
--	----------------------------------

ПК-1.1. Использует принципы управления инфраструктурой коллективной разработки программных средств	Защита лабораторной работы, устный опрос, экзамен
ПК-1.2. Осуществляет поиск и подбор ресурсов для организации разработки и эксплуатации программного обеспечения	Защита лабораторной работы
ПК-1.3. Руководит персоналом и прочими ресурсами для решения поставленной профессиональной задачи	Защита лабораторной работы, устный опрос, 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	Многомерный регрессионный анализ	Как ставится задача многомерного линейного регрессионного анализа? Как находятся оценки коэффициентов регрессии методом наименьших квадратов? Что требуется для оценки значимости коэффициентов регрессии? Как строятся доверительные интервалы для них? В чем заключается задача статистического прогноза? Что такое оптимальный предиктор? Что такое эмпирический предиктор?
2	Дисперсионный анализ	В чем состоит основная идея дисперсионного анализа? Как классифицируются задачи дисперсионного анализа? На каком тождестве основан одномерный дисперсионный анализ? Перечислите основные предпосылки обоснованного дисперсионного анализа. Как произвести их статистическую проверку?
3	Факторный анализ	Как строится стандартизованная матрица данных в факторном анализе? В чем состоит общая постановка задачи факторного анализа? Что такое общие и характерные латентные факторы? Как формулируется основная задача факторного анализа в матричной форме? В чем заключается неоднозначность ее решения? В чем заключается метод главных компонент? Опишите алгоритм этого метода. Трудности факторного анализа. Интерпретация факторов.
4	Дискриминантный анализ	В чем заключается общая постановка задачи дискриминации? Постановка задачи дискриминации для двух классов в случае нормального распределения вектора признаков с одинаковой ковариационной матрицей для разных классов.

		Линейная дискриминантная функция. Дискриминантная функция для случая, когда ковариационные матрицы вектора признаков для разных классов различны. Эвристические правила дискриминации. Метод k ближайших соседей.
5	Временные ряды	Что такое временной ряд? На какие составляющие можно разложить величину, определяющую временной ряд? Что такое тренд временного ряда? Что такое эффект сезонности временного ряда?

5.2.2. Перечень контрольных материалов для защиты курсового проекта/ курсовой работы

Учебным планом не предусмотрены.

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

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

В методических указаниях к выполнению лабораторных работ по дисциплине представлен перечень лабораторных работ, обозначены цель и задачи, представлены индивидуальные варианты заданий и перечень контрольных вопросов.

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. Примерные перечень контрольных вопросов для защиты лабораторных работ приведен в таблицах.

1. Компетенция ПК-1. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях.

Наименование индикатора (показателя оценивания)	Используемые средства оценивания
Использует принципы управления инфраструктурой коллективной разработки программных средств	Вопросы при защите лабораторной работы: 1. Как формулируется общая задача регрессионного анализа? 2. В какой форме представляются исходные данные для оценивания коэффициентов регрессии методом наименьших квадратов? 3. Приведите доказательство основной матричной формулы для коэффициентов регрессии.

	<p>4. Как производится проверка значимости коэффициентов регрессии?</p> <p>5. Как определяются доверительные интервалы для коэффициентов регрессии?</p>
<p>Осуществляет поиск и подбор ресурсов для организации разработки и эксплуатации программного обеспечения</p>	<p>Вопросы при защите лабораторной работы:</p> <ol style="list-style-type: none"> 1. Постановка задачи статистического прогноза. 2. Дайте определение оптимального предиктора. Существование оптимального предиктора. Функция регрессии 3. Определение оптимального предиктора в случае линейной функции регрессии. Оптимальный линейный предиктор. 4. Что такое эмпирический предиктор? 5. Оценивание качества регрессионной модели с помощью корреляционного отношения.
<p>Руководит персоналом и прочими ресурсами для решения поставленной профессиональной задачи</p>	<p>Вопросы при защите лабораторной работы:</p> <ol style="list-style-type: none"> 1. Постановка задач дисперсионного анализа и их классификация. Предпосылки применимости метода. 2. Однофакторный дисперсионный анализ. Критерий проверки основной гипотезы. 3. Как обосновать применимость метода дисперсионного анализа?

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:

Показатель оценивания	Критерий оценивания
Знания	Знание основных терминов, понятий, структур
	Объем усвоенного материала
	Полнота ответов на поставленные вопросы
	Четкость изложения ответов на вопросы
Умения	Умение полно ответить на поставленный вопрос по лабораторной работе
	Умение дополнить и изменить свою программу в соответствии с поставленной задачей
	Умение разработать алгоритм по поставленному заданию
Навыки	Навык самостоятельной разработки приложения
	Навык работы в разработанных приложениях
	Навык правильности выбора алгоритма

Оценка преподавателем выставляется интегрально по всем показателям и критериям оценивания.

Оценка сформированности компетенций по показателю Знания.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Знание терминов, определений, понятий	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	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 дополнительными 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	Не умеет использовать	Использование теоретических	Умеет использовать	Умело использует

теоретические знания для выбора методики решения профессиональных задач	теоретические знания для выбора методики решения профессиональных задач	знаний для выбора методики решения профессиональных задач вызывает затруднения	теоретические знания для выбора методики решения профессиональных задач	теоретические знания для выбора методики решения профессиональных задач
---	---	--	---	---

Оценка сформированности компетенций по показателю Навыки.

Критерий	Уровень освоения и оценка			
	2	3	4	5
Владение навыками теоретического и экспериментального исследования объектов профессиональной деятельности	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	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	Не может самостоятельно выполнять исследования объектов профессиональной деятельности	Выполняет исследования объектов профессиональной деятельности с посторонней помощью	При выполнении исследования 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	Специализированная мебель.

	лабораторных занятий	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	Microsoft Windows 10 Корпоративная	(Соглашение Microsoft Open Value Subscription V6328633 Соглашение действительно с 02.10.2017 по 31.10.2020). Договор поставки ПО 0326100004117000038-0003147-01 от 06.10.2017.
2	Microsoft Office Professional Plus 2016	Соглашение Microsoft Open Value Subscription V6328633 Соглашение действительно с 02.10.2017 по 31.10.2020
3	JavaJDK, NetBeansIDE, EclipseIDE, JavaDevC++ – пакеты для разработки программ на языке Java	Свободно распространяемое ПО согласно условиям лицензионного соглашения
4	CodeBlocks (компиляторы gcc),	Свободно распространяемое ПО согласно условиям лицензионного соглашения
5	Открытая система анализа данных DLP (MyDLP)	Свободно распространяемое ПО согласно условиям лицензионного соглашения
56	Microsoft Visual Studio	Подписка Microsoft Imagine Premium id: 6f22ecb4-6882-420b-a39b-afba0ace820c.

6.3. Перечень учебных изданий и учебно-методических материалов

1. Анализ данных качественных исследований [Электронный ресурс] : практикум / . — Электрон. текстовые данные. — Ставрополь: Северо-Кавказский федеральный университет, 2016. — 94 с. – (Доступ: <http://www.iprbookshop.ru/66014> - ЭБС «IPRbooks», по паролю).
2. Жуковский О.И. Информационные технологии и анализ данных [Электронный ресурс] : учебное пособие / О.И. Жуковский. — Электрон. текстовые данные. — Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2014. — 130 с. – (Доступ: <http://www.iprbookshop.ru/72106> - ЭБС «IPRbooks», по паролю).
3. Афонин П.Н. Статистический анализ с применением современных программных средств [Электронный ресурс] : учебное пособие / П.Н. Афонин, Д.Н. Афонин. — Электрон. текстовые данные. — СПб. : Интермедия, 2015. — 100 с. – (Доступ: <http://www.iprbookshop.ru/28030> - ЭБС «IPRbooks», по паролю).
4. Уэс Маккинли Python и анализ данных [Электронный ресурс] / Маккинли Уэс. — Электрон. текстовые данные. — Саратов: Профобразование, 2017. — 482 с. – (Доступ: <http://www.iprbookshop.ru/64058> - ЭБС «IPRbooks», по паролю).
5. Боровиков В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов.–ЗАО Издательский дом «Питер», 2003 – 686 с.

6. Брусенцев А.Г. Анализ данных и процессов. Часть I: методы статистического анализа данных. – Учебное пособие. – Белгород: Издательство БГТУ им. В.Г. Шухова, 2017. – 63 с.
7. Симчера В.М. Методы многомерного анализа статистических данных [Электронный ресурс]: учеб. пособие. – Москва: Финансы и статистика, 2008. – 400 с. – Режим доступа: <https://e.lanbook.com/book/1005>
8. Маглеванный И.И. Математические основы первичной обработки экспериментальных данных [Электронный ресурс]: методические материалы по прикладной статистике / И.И. Маглеванный, Т.И. Карякина. – Волгоград: Волгоградский государственный социально-педагогический университет, «Перемена», 2015. – 42 с. – 2227-8397. – Режим доступа: <http://www.iprbookshop.ru/40738.html>
9. Прикладная математическая статистика [Электронный ресурс]: учебное пособие / Томск: Томский государственный университет систем управления и радиоэлектроники, 2016. – 113 с. – 2227-8397. – Режим доступа: <http://www.iprbookshop.ru/72166.html>
10. Барсегян А. Анализ данных и процессов / А. Барсегян. – 3-е изд. – СПб.: БХВ-Петербург, 2009.
11. Марц Н. Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени. / Н. Марц, Д. Уоррен. – М.: Вильямс, 2016.

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/>

7. УТВЕРЖДЕНИЕ РАБОЧЕЙ ПРОГРАММЫ

Рабочая программа утверждена на 20____ /20____ учебный год без изменений

Протокол № _____ заседания кафедры от « ____ » _____ 20__ г.

Заведующий кафедрой _____ (Поляков В.М.)
подпись, ФИО

Директор института _____ (Белоусов А.В.)
подпись, ФИО