

Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»
(Финуниверситет)
Калужский филиал Финуниверситета
Кафедра «Бизнес-информатика и высшая математика»

«УТВЕРЖДАЮ»

Директор Калужского филиала
Финуниверситета



В.А. Матчинов

«30» июня 2025 г.

И.В. Винокуров

ТЕХНОЛОГИИ ОБРАБОТКИ И АНАЛИЗА БОЛЬШИХ ДАННЫХ

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки
38.03.01 «Бизнес-анализ, налоги и аудит», профиль: «Учёт, анализ и аудит»

*Рекомендовано Ученым советом Калужского филиала Финуниверситета
(протокол № 30 от 30.06.2025 г.)*

Одобрено кафедрой «Бизнес – информатика и высшая математика»
Калужского филиала Финуниверситета
(протокол № 10 от 13 мая 2025 г.)


Калуга 2025

Рабочая программа предназначена для преподавания дисциплины «Технологии обработки и анализа больших данных» студентам, обучающимся по направлению подготовки 38.03.01 «Бизнес-анализ, налоги и аудит», профиль: «Учёт, анализ и аудит», по очной форме обучения.

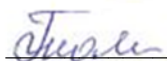
В рабочей программе излагаются планируемые результаты освоения дисциплины, содержание дисциплины, тематика и содержание семинаров и практических занятий, технологии их проведения. Приводится перечень учебно-методического обеспечения для самостоятельной работы обучающихся, фонд оценочных средств для проведения промежуточной аттестации обучающихся, перечень основной и дополнительной литературы, а также ресурсов информационно-телекоммуникационной сети Интернет.

СОГЛАСОВАНО:

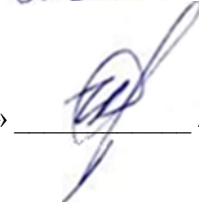
Заместитель директора
по учебно-методической работе
«30» июня 2025 г.

 /Орловцева О.М./

Начальник учебно-методического отдела
«30» июня 2025 г.

 /Толстикова В.С./

Заведующий кафедрой
«Бизнес-информатика и высшая математика»
«30» июня 2025 г.

 /Дробышева И.В./

Оглавление

1. Наименование дисциплины	4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения, со- отнесённых с планируемыми результатами обучения по дисциплине.....	4
3. Место дисциплины в структуре образовательной программы	4
4. Объем дисциплины в зачётных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	5
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий.	5
5.1 Содержание дисциплины	5
5.2 Учебно-тематический план	6
5.3 Содержание семинаров, практических занятий	7
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	8
6.1 Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	8
6.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю успеваемости.....	9
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	9
7.1 Перечень компетенций с указанием индикаторов их достижения в процессе освоения дисциплины.....	9
7.2 Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний	9
8. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины	10
Основная литература:	10
Дополнительная литература:	11
9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины.....	11
10. Методические указания для обучающихся по освоению дисциплины	11
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)	13
11.1 Комплект лицензионного программного обеспечения	13
11.2 Современные профессиональные базы данных и информационные справочные системы.....	13
11.3 Сертифицированные программные и аппаратные средства защиты информации.....	13
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	13

1. Наименование дисциплины

«Технологии обработки и анализа больших данных»

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения, со- отнесённых с планируемыми результатами обучения по дисциплине

В результате изучения дисциплины у студентов должны быть сформированы следующие компетенции:

Таблица 1

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесённые с компетенциями/ индикаторами достижения компетенции
ПКП-5	Способность к использованию специальных программных продуктов, применяемых для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте	1. Использует специальные программные продукты для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте	Знать специализированные программные продукты для анализа данных Уметь применять специализированные программные продукты для анализа данных
		2. Демонстрирует владение специальными программными продуктами, применяемых для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте	Знать специальные программные продукты, применяемые для выполнения бухгалтерско-аналитических и контрольных функций Уметь решать задачи подготовки, обработки и анализа полученных результатов с использованием современных программных продуктов

3. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии обработки и анализа больших данных» относится к элективной дисциплине модуля «Технологии анализа и прогнозирования данных» по направлению 38.03.01 «Бизнес-анализ, налоги и аудит», профиль: «Учёт, анализ и аудит».

4. Объем дисциплины в зачётных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Таблица 2

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 6 (в часах)
Общая трудоемкость дисциплины	3 з/е, 108 ч.	108 ч.
Контактная работа – аудиторные занятия	34	34
Лекции	16	16
Семинары, практические занятия	18	18
Самостоятельная работа	74	74
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	зачёт	зачёт

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1 Содержание дисциплины

Тема 1. Программные продукты для создания аналитических моделей

Среда для формирования и исследования аналитических моделей Google Colaboratory. Интегрированная среда разработки приложений на языке Python – Visual Studio Code. Способы представления данных аналитических моделей и их визуализация с использованием библиотеки Matplotlib. Работа с Google Disc. Сохранение и загрузка моделей.

Тема 2. Регрессионные модели

Прогнозирование значений наблюдаемой величины с использованием линейной и полиномиальной регрессий. Регрессор LinearRegression библиотеки Scikit-Learn. Оценка точности полученных результатов. Построение графиков.

Тема 3. Модели кластеризации

Прогнозирование принадлежности одному или нескольким классам с использованием логистической регрессии. Регрессор LogisticRegression библиотеки Scikit-Learn. Кластеризация методом k-средних. Регрессор KMeans библиотеки Scikit-Learn. Кластеризация с использованием методов дерева принятия решений и случайного леса. Регрессоры DecisionTreeClassifier и RandomForestClassifier библиотеки Scikit-Learn. Классификация методом опорных векторов. Регрессор SVM библиотеки Scikit-Learn.

5.2 Учебно-тематический план

Таблица 3

№	Наименование тем(разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Аудиторная работа			Самостоятельная работа	
			Общая	Лекции	Семинары, практические занятия		
Тема 1. Программные продукты для создания аналитических моделей							
1	Изучение среды для создания и исследования аналитических моделей Google Colaboratory	12	4	2	2	8	Выполнение и защита практических работ
2	Изучение среды разработки программных моделей на языке Python – Visual Studio Code	12	4	2	2	8	Выполнение и защита практических работ
Тема 2. Регрессионные модели							
3	Формирование и исследование моделей линейной и полиномиальной регрессий.	16	4	2	2	12	Выполнение и защита практических работ
4	Оценка точности регрессионных моделей. Построение графиков	16	4	2	2	12	Выполнение и защита практических работ
Тема 3. Модели кластеризации							
5	Проектирование и реализация модели логистической регрессии	14	4	2	2	10	Выполнение и защита практических работ
6	Проектирование и реализация модели кластеризации методом k-средних	18	6	2	4	12	Выполнение и защита практических работ
7	Проектирование и реализация моделей дерева принятия решений и случайного леса	20	8	4	4	12	Выполнение и защита практических работ
	В целом по дисциплине	108	34	16	18	74	Согласно учебному плану: контрольная работа
	В % к итогу	100	31	15	16	69	

5.3 Содержание семинаров, практических занятий

Таблица 4

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8, 9	Формы проведения занятий
Тема 1. Программные продукты для создания аналитических моделей	Изучение сред разработки и исследования аналитических моделей Google Colaboratory и Visual Studio Code Основная литература: 1-7 Дополнительная литература: 8,9	Компьютер- ный практикум
Тема 2. Регрессионные модели	Проектирование и реализация регрессионных моделей для обработки экспериментальных данных Оценка точности моделей Основная литература: 1-7 Дополнительная литература: 8,9	Компьютер- ный практикум
Тема 3. Модели кластеризации	Проектирование и реализация моделей кластеризации для обработки эксперименталь ных данных ▪ Вычисление и визуализация метрик точности Основная литература: 1-7 Дополнительная литература: 8,9	Компьютер- ный практикум

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6. 1 Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Таблица 5

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 1 Программные продукты для создания аналитических моделей	Изучение основных функциональных возможностей библиотек Matplotlib и Plotly для визуализации результатов исследования аналитических моделей в средах Google Colaboratory и Visual Studio Code Основная литература: 1,2 Дополнительная литература: 8,9	Изучение методических материалов по теме в электрон- ном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам
Тема 2. Регрессионные модели	Матрица ошибок confusion matrix Вычисление и способы визуализации Основная литература: 1,2 Дополнительная литература: 8,9	Изучение методических материалов по теме в электрон- ном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам
Тема 3. Модели кластеизации	Метрики точности регрессионных моделей и моделей кластеризации – Accuracy, Precision, F1-score и AUC- ROC Основная литература: 1,2 Дополнительная литература: 8,9	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам

6.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю успеваемости

Примерные темы для контрольной работы:

Проектирование и исследование модели аналитики для <заданной предметной области>.

Критерии балльной оценки по контрольной работе содержатся в соответствующих методических рекомендациях кафедры.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

7.1 Перечень компетенций с указанием индикаторов их достижения в процессе освоения дисциплины

Перечень компетенций представлен в разделе 2, который характеризует перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.

7.2 Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Таблица 6

Компетенция	Типовые задания
ПКП-5 Способность к использованию специальных программных продуктов, применяемых для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте	Использует специальные программные продукты для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте Задание 1. Создайте модель кластерного анализа в среде Google Colaboratory Задание 2. Реализуйте классификацию по экспериментальным данным методом оперных векторов
	Демонстрирует владение специальными программными продуктами, применяемых для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте Задание 1. Постройте график линейной регрессии для заданной предметной области Задание 2. Вычислите метрики точности моделей кластеризации для заданной предметной области

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях кафедры.

Примерные вопросы к зачету:

1. Опишите основные этапы работы с моделями данных в Google Colaboratory и Visual Studio Code.
2. Раскройте понятие линейной регрессии. Приведите типы задач, в которых используется линейная регрессия. Опишите модель линейной регрессии, реализованную в библиотеке Scikit-Learn.
3. Раскройте понятие логистической регрессии. Приведите типы задач, в которых используется логистическая регрессия. Опишите модель логистической регрессии, реализованную в библиотеке Scikit-Learn.
4. Раскройте понятие дерева принятия решений. Приведите типы задач, в которых используется дерево принятия решений. Опишите модель дерева принятия решений, реализованную в библиотеке Scikit-Learn.
5. Раскройте понятие кластеризации. Приведите типы задач, в которых используется кластеризация. Опишите реализацию модели кластеризации в библиотеке Scikit-Learn.
6. Раскройте понятие бинарной классификации. Приведите типы задач, в которых используется бинарная классификация. Опишите реализации моделей бинарной классификации в библиотеке Scikit-Learn.
7. Раскройте понятие многоклассовой классификации. Приведите типы задач, в которых используется многоклассовая классификация. Опишите модели многоклассовой классификации, реализованные в библиотеке Scikit-Learn.
8. Раскройте основные достоинства и недостатки известных вам моделей кластеризации.
9. Раскройте понятие метрики точности модели. Приведите известные вам метрики точности и опишите реализацию их вычисления на языке Python.
10. Приведите основные библиотеки Python, используемые для обработки экспериментальных данных. Опишите их назначения и основные функциональные возможности.

8. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

Основная литература:

1. Протоdjаконов, А. В. Алгоритмы Data Science и их практическая реализация на Python: учебное пособие / Протоdjаконов А. В., Пылов П. А., Садовников В. Е. – М., Вологда: Инфра-Инженерия, 2022. – 392 с. – URL: <https://www.iprbookshop.ru/124000.html>
2. Маккинли, У. Python и анализ данных / Уэс Маккинли. – Саратов: Профобразование, 2019. – 482 с.
3. Чару, А. Нейронные сети и глубокое обучение. Учебный курс / Аггарвал Чару. – СПб.: Вильямс, 2020. – 752 с.

4. Гудфеллоу, Я., Бенджио, И., Курвилль, А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с. – URL: <https://e.lanbook.com/book/107901.html>
5. Хливненко, Л. В. Практика нейросетевого моделирования: учебное пособие для вузов / Л. В. Хливненко, Ф. А. Пятакович. – СПб.: Лань, 2021. – 200 с. – URL: <https://e.lanbook.com/book/173811.html>.
6. Мюллер, А, Гвидо, С. Введение в машинное обучение с помощью Python. Руководство для специалистов по обработке данных. / Андреас Мюллер, Сара Гвидо. – СПб.: ООО “Альфа-книга”, 2017. – 480 с.
7. Николаенко, С., Кадури, А., Архангельская, Е. Глубокое обучение. Погружение в мир нейронных сетей. – СПб.: Питер, 2018. – 480 с. – URL: <https://djvu.online/file/JA192EZJDBIE>

Дополнительная литература:

8. Грас, Дж. Data Science. Наука о данных с нуля / Джоэл Грас. – СПб: БХВ-Петербург, 2020. – 411 с.
9. Chollet, Francois Deep Learning with Python, Second Edition. – NY: Simon and Schuster, 2021. – 478 с.
10. Клетте, Р. Компьютерное зрение. Теория и алгоритмы: учебник / Рейнхальд Клетте; перевод с англ. А. А. Слинкина. – М.: ДМК Пресс, 2019. – 506 с. – URL: <https://e.lanbook.com/book/131691.html>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

1. Электронная библиотека Финансового университета <http://elib.fa.ru/>
2. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
4. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.urait.ru/>
5. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
6. Библиотека Scikit-Learn <https://scikit-learn.org>
7. Библиотека Keras <https://keras.io>

10. Методические указания для обучающихся по освоению дисциплины

Обучающимся в рамках самостоятельной работы следует использовать Методические рекомендации по планированию и организации внеаудиторной самостоятельной работы студентов по образовательным программам бакалавриата и магистратуры в Финансовом университете, утвержденные Приказом ректора №1040/о от 11.05.2021 г.

Самостоятельная работа студентов проходит внеаудиторно. Организации самостоятельной работы служит учебно-тематический план изучения дисциплины. В данном плане указана тематика лекций, семинаров, вопросы и задания для самостоятельного изучения. Во время лекций необходимо конспектировать содержание лекции. После лекции необходимо отредактировать записи, оформить

конспект, дополняя его содержание дополнительной информацией. При оформлении конспекта целесообразно выделять названия тем и формулировки вопросов, основные определения, примеры.

При подготовке к семинару необходимо изучить вопросы семинара, соответствующий теоретический материал, делая для себя необходимые записи в рабочей тетради. После занятий необходимо просмотреть записанные решения и восстановить в решениях имеющиеся пробелы.

При затруднении в решении практических вопросов (задач), можно обратиться за консультацией (помощью) к преподавателю. Семинары проходят, как правило, в интерактивной форме и преподаватель учитывает активность обучающихся, направленную на решение предложенных вопросов (вариантов задач), а также вариантов ответов на решаемые вопросы (проблемы).

Не следует бояться дать неверный ответ или допустить иную ошибку: исправление и анализ ошибок в режиме общения с преподавателем и сокурсниками в ходе семинара способствует более глубокому освоению учебного материала и предупреждает возникновение ошибок в дальнейшем. Домашние задания (подготовку к занятиям) следует осуществлять регулярно. Если то или иное задание, при подготовке к семинару вызвало затруднение, необходимо обратиться к преподавателю за консультацией. Регулярность в выполнении домашних заданий (подготовке к занятиям) - важный фактор качественного освоения дисциплины.

Методические рекомендации по обучению лиц с ограниченными возможностями здоровья

Профессорско-преподавательский состав знакомится с психолого-физиологическими особенностями обучающихся инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ). При необходимости осуществляется дополнительная поддержка преподавания тьюторами, психологами, социальными работниками, прошедшими подготовку ассистентами.

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 с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе. Подбор и разработка учебных материалов производятся с учётом предоставления материала в различных формах: аудиальной, визуальной, с использованием специальных технических средств и информационных систем. Освоение дисциплины лицами с ОВЗ осуществляется с использованием средств обучения общего и специального назначения (персонального и коллективного использования). Материально-техническое обеспечение предусматривает приспособление аудиторий к нуждам лиц с ОВЗ. Форма проведения аттестации для студентов-инвалидов устанавливается с учётом индивидуальных психофизических особенностей.

Для студентов с ОВЗ предусматривается доступная форма предоставления заданий оценочных средств, а именно:

- в печатной или электронной форме (для лиц с нарушениями опорно-двигательного аппарата);

- в печатной форме или электронной форме с увеличенным шрифтом и контрастностью (для лиц с нарушениями слуха, речи, зрения);
- методом чтения ассистентом задания вслух (для лиц с нарушениями зрения). Студентам с инвалидностью увеличивается время на подготовку ответов на контрольные вопросы.

Для таких студентов предусматривается доступная форма предоставления ответов на задания, а именно:

- письменно на бумаге или набором ответов на компьютере (для лиц с нарушениями слуха, речи);
 - выбором ответа из возможных вариантов с использованием услуг ассистента (для лиц с нарушениями опорно-двигательного аппарата);
 - устно (для лиц с нарушениями зрения, опорно-двигательного аппарата).
- При необходимости для обучающихся с инвалидностью процедура оценивания результатов обучения может проводиться в несколько этапов.

Методические рекомендации по выполнению контрольной работы

Методические рекомендации по выполнению контрольной работы предусмотрены в «Методических рекомендациях по подготовке написанию и оформлению контрольной работы», разрабатываемой преподавателем кафедры на учебный год, в котором реализуется учебная дисциплины

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

11.1 Комплект лицензионного программного обеспечения

1. Антивирусная защита Windows defender
2. Astra Linux, Libre Office

11.2 Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Консультант Плюс»;
2. Информационно-правовая система «Гарант»;

11.3 Сертифицированные программные и аппаратные средства защиты информации

Сертифицированные программные и аппаратные средства защиты информации не предусмотрены

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

- аудитории для проведения лекционных и семинарских занятий, оборудованные видеопроекционным оборудованием для презентаций, средствами звуковоспроизведения, экраном, доской меловой/интерактивной;
- библиотеку, имеющую рабочие места для студентов, оснащенные

компьютерами с доступом к базам данных и сети Интернет

- компьютерные классы с набором лицензионного базового программного обеспечения для проведения практических занятий и выходом в глобальную сеть Internet;

Образовательные технологии

В процессе изучения дисциплины «Технология обработки и анализа больших данных» предполагается:

- сопровождение курса лекций наглядной презентацией, включающей практические примеры, схемы, графики, табличный материал;
- рассмотрение на семинарских занятиях интерактивных ситуационных задач по проблематике дисциплины;
- деловые игры;
- разбор конкретных ситуаций, коллективное обсуждение проблем российской и зарубежной практики по изучаемым темам;
- виртуальное общение в течение срока изучения курса в целях обеспечения лекций и практических занятий необходимым материалом и также контроля самостоятельной работы студентов.