

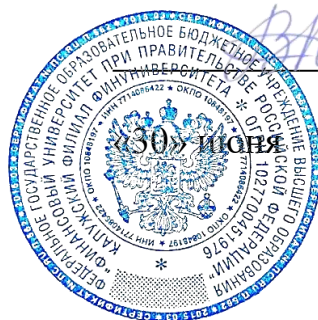
**«Финансовый университет при Правительстве Российской Федерации»  
(Финуниверситет)**

**Калужский филиал Финуниверситета**

**Кафедра «Бизнес – информатика и высшая математика»**

**«УТВЕРЖДАЮ»**

Директор Калужского филиала  
Финуниверситета



*В.А. Матчинов* - В.А. Матчинов

30 июня 2022 г.

Винокуров И.В.

**ТЕХНОЛОГИИ ОБРАБОТКИ И АНАЛИЗА БОЛЬШИХ  
ДАННЫХ**

**Рабочая программа дисциплины**

для студентов, обучающихся по направлению подготовки

38.03.01 «Экономика»

Образовательная программа «Бизнес-анализ, налоги и аудит»

Очная форма обучения

*Рекомендовано Ученым советом Калужского филиала Финуниверситета  
(протокол №56 от 30.06. 2022 г.)*

Одобрено кафедрой «Бизнес – информатика и высшая математика»  
Калужского филиала Финуниверситета  
(протокол № 12 от 28 июня 2022 г.)


КАЛУГА 2022


Рабочая программа предназначена для преподавания дисциплины «Технологии обработки и анализа больших данных» студентам, обучающимся по направлению подготовки 38.03.01 «Экономика», образовательная программа «Бизнес-анализ, налоги 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. В рабочей программе дисциплины 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  /Орловцева О.М./  
«28» июня 2022 г.

Начальник учебно-методического отдела  /Толстикова В.С./  
«28» июня 2022 г.

Заведующий кафедрой  
«Бизнес-информатика и высшая математика»  /Дробышева И.В./  
«28» июня 2022 г.

## Содержание

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

### **1.Наименование дисциплины**

Б.1.2.2.2.1.1. Технологии обработки и анализа больших данных

## 2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции <sup>1</sup>	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции
<b>ПКП-5</b>	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.	Знание: Специальных программных продуктов для выполнения различных функций в экономическом субъекте Умение: Использовать и применять специальные ПП для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте
		2. Демонстрирует владение специальными программными продуктами, применяемых для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте.	Знание: возможные варианты организации работы бухгалтерской службы Умение: владеть специальными программными продуктами в бухгалтерском учете; формулировать и распределять функциональные обязанности между сотрудниками бухгалтерской службы.

### Место дисциплины в структуре образовательной программы

*Дисциплина «Технологии обработки и анализа больших данных» является дисциплиной модуля «Технологии анализа и прогнозирования данных»*

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

Для очной формы обучения

Таблица 1

Вид учебной работы по дисциплине	Всего (в часах и зач.ед.)	Семестр 6 (в часах)
<b>Общая трудоемкость дисциплины</b>	<b>108</b>	<b>108</b>
<b>Контактная работа - Аудиторные занятия</b>	<b>34</b>	<b>34</b>
<i>Лекции</i>	16	16
<i>Семинары, практические</i>	18	18

<sup>1</sup> Заполняется при реализации актуализированных ОС ВО ФУ и ФГОС ВО3++

занятия		
<b>Самостоятельная работа</b>	<b>74</b>	<b>74</b>
Вид текущего контроля	к\р	к\р
Вид промежуточной аттестации	зачет	зачет

**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. От данных к ценности: Основы наук о данных**

Основные определения: 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 Big Data. Цикл обработки данных: поиск данных, сбор данных, очистка данных, трансформация данных, интеллектуальный анализ данных, интерпретация и практическое применение результатов. Статистические основы обработки данных: краткий обзор выборочного метода, методов описательной статистики, построения интервальных оценок и проверки гипотез. Обзор современных технологий машинного обучения: R, Python, Spark, Microsoft Azure ML.

## **Тема 2. Практическое использование моделей классификации**

Принципы разработки и оценки систем машинного обучения. Основные классы моделей машинного обучения: обучение с учителем (классификация и регрессия) и без учителя (кластеризация и поиск аномалий). Задача классификации с обучением. Задача кредитного скоринга. Модель логистической регрессии и ее компьютерная реализация. Методы оценки качества моделей классификации: доля правильных ответов, точность, полнота, F1, AUC. Цены ошибок первого и второго рода. Модели деревьев и лесов решений. Усиление (бустинг) деревьев решений. Нейронные сети. Компьютерная реализация методов классификации. Недообучение и переобучение в моделях классификации. 8 Основы технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными. Работа с несбалансированными выборками.

## **Тема 3. Практическое использование моделей регрессии**

Модель множественной линейной регрессии. Методы оценки качества моделей регрессии. Коэффициент детерминации, средняя абсолютная ошибка предсказания, средняя относительная ошибка предсказания. Методы классификации и их компьютерная реализация. Недообучение и переобучение в моделях регрессии. Основы технологии улучшения моделей машинного обучения в задачах регрессии. Задача прогнозирования продаж.

## **Тема 4. Практическое использование моделей кластерного анализа и поиска аномалий**

Кластерный анализ и его компьютерная реализация. Задача сегментирования потребителей. Иерархические агломеративные алгоритмы. Метод К-средних. Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами внутри кластеров. Характеристики центров кластеров. Задача снижения размерности факторного пространства. Метод главных компонент и его компьютерная реализация. Геометрическое и экономическое содержание метода главных компонент. Использование методов снижения размерности для улучшения качества моделей машинного обучения. Использование методов снижения размерности для выявления латентных факторов. Поиск аномалий. Задача поиска мошеннических транзакций.

#### **Тема 5. Анализ текстов, изображений, эмоций и рекомендательные системы**

Принципы анализа текстовой информации. Принципы анализа графической информации. Принципы анализа 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. Современные практические исследования Финансового университета в области обработки данных и машинного обучения: индикаторы смены состояний финансовых рынков, анализ сложных сетей в управлении городами, энергетике, на транспорте, предиктивный ремонт оборудования, системы скоринга.

## **5.2 Учебно-тематический план**

Таблица 2

№ п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Все го	Контактная работа- Аудиторная работа			Самостоятел ьная работа	
			Обща я, в т.ч.:	Лекц ии	Семинары, практичес кие занятия		
1.	От данных к ценности: Основы наук о данных	23	8	4	4	15	Аудиторные самостоятель ные работы.
2.	Практическое использование моделей классификации	22	7	3	4	15	Участие в решении задач на

3.	Практическое использование моделей регрессии	22	7	3	4	15	практически х занятиях. Собеседован ия по домашним заданиям.
4.	Практическое использование моделей кластерного анализа и поиска аномалий	21	6	3	3	15	
5.	Анализ текстов, изображений, эмоций и рекомендательные системы	20	6	3	3	14	
	<b>В целом по дисциплине</b>					к\р	<b>Согласно учебному плану:</b>
	<b>Итого</b>	108	34	16	18	74	

### 5.3 Содержание семинаров, практических занятий

Таблица 3

	Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники	Формы проведения занятий
1	От данных к ценности: Основы наук о данных	Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов. Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов Рекомендуемые источники: п.8, [1], [2],[6]; п.9, [4]-[7]	Лабораторная работа
2	Практическое использование моделей классификации	Построение системы прогнозирования продаж. Построение системы кредитного скоринга. Оптимизация системы кредитного скоринга. Рекомендуемые источники: п.8, [1]-[8]; п.9,[4]-[9]	Лабораторная работа
3	Практическое использование моделей регрессии	Построение системы прогнозирования продаж. Построение системы прогнозирования продаж» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов). Рекомендуемые источники: п.8, [1]-[8]; п.9, [4]-[9]	Лабораторная работа
4	Практическое использование моделей кластерного анализа и поиска аномалий	Кластерный анализ и его компьютерная реализация. Иерархические агломеративные алгоритмы. Метод К-средних. Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами. Поиск аномалий. Задача поиска мошеннических транзакций. Сегментирование потребителей. Рекомендуемые источники: п.8, [2]-[5]; п.9,[4]-[9]	Лабораторная работа
5	Анализ текстов, изображений, эмоций и рекомендательные системы	Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей. Построение рекомендательного сервиса. Анализ эмоциональной окраски записей в социальных сетях.	Лабораторная работа



	Рекомендуемые источники : п.8, [2]-[5]; п.9, [4]-[9]	
--	--	--

## 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 4

Наименование разделов, тем дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
От данных к ценности: Основы наук о данных	Статистические основы обработки данных: краткий обзор выборочного метода, методов описательной статистики, построения интервальных оценок и проверки гипотез.	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение и оформление лабораторной работы
Практическое использование моделей классификации	Технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными. Работа с несбалансированными выборками	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение и оформление лабораторной работы
Практическое использование моделей регрессии	Технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение и оформление лабораторной работы
Практическое использование моделей кластерного анализа и поиска аномалий	Использование методов снижения размерности для улучшения качества моделей машинного обучения. Использование методов снижения размерности для выявления латентных факторов.	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение и оформление лабораторной работы
Анализ текстов, изображений, эмоций и рекомендательные системы	Современные практические исследования Финансового университета в области обработки данных и машинного обучения: индикаторы смены состояний финансовых рынков, анализ сложных сетей в управлении городами, энергетике, на транспорте, предиктивный ремонт оборудования, системы скоринга.	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение и оформление лабораторной работы

### 6.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю (согласно таблице 2)

#### Примерный вариант контрольной работы

Контрольная работа проводится в виде соревнования на платформе kaggle. Участники соревнования строят модели кредитного скоринга. Речь идет о следующем: когда клиент обращается в банк с заявлением о предоставлении кредита, банк принимает решение о выдаче кредита или об отказе в предоставлении кредита с использованием статистических моделей на основании информации о тех клиентах, которые уже брали кредит (кто-то из них выполнил свои обязательства по кредитному договору, а кто-то не выполнил). На вероятность возврата кредита может влиять много факторов, причем сложным

образом, и для прогнозирования результатов по каждому отдельному случаю участникам соревнования необходимо построить модель машинного обучения, которая на основании данных из заявления о выдаче кредита предсказывает, вернет ли заемщик этот кредит. Ссылка на чемпионат выдается участникам преподавателем. Набор данных `credit_train.csv`, который предоставляется участникам соревнования, содержит ряд особенностей, включая пропущенные значения, дублирующиеся строки, выбросы и т. п. При построении модели машинного обучения целесообразно попробовать создать новые синтетические факторные признаки, являющиеся функциями от исходных факторов, преобразовать типы данных исходных факторов и т. п. Исходный полный набор данных о кредитах разделен на три набора:

- набор данных `credit_train.csv`, предназначенный для обучения модели, включающий 100000 записей о кредитах, относительно каждого из которых известно значение признака «Loan Status» - «Fully Paid» (погашен полностью) или «Charged Off» (не погашен);
- публичный тестовый набор данных, содержащий 5000 записей о кредитах (участникам соревнования недоступна информация о значении признака «Loan Status»);
- закрытый тестовый набор данных, содержащий 5000 записей о кредитах; вычисление (участникам соревнования недоступна информация о значении признака «Loan Status»).

Участник соревнования должен загрузить на платформу чемпионата файл с прогнозами по 10000 кредитам, данные по которым представлены в файле `credit_test.csv`, половина из этих кредитов относится к публичному тестовому набору данных, другая половина - к закрытому тестовому набору данных. Сразу после загрузки результатов автоматически рассчитывается доля правильных ответов (Accuracy) и рейтинг участников на основании публичного тестового набора данных. В момент окончания соревнования для каждого участника происходит перерасчет доли правильных ответов и рейтинга на основании закрытого тестового набора данных.

#### Примеры тестовых заданий

1. Как не используют выборки из генеральной совокупности аналитики больших данных:

a) Как метод формирования комплексного суждения о генеральной совокупности случайной величины

b) Как метод тестирования полученных моделей

c) Как метод верификации исходных данных

2. Укажите лишний этап построения статистической модели:

a) Сбор и верификация исходных данных

b) Выбор факторов

c) Построение модели

d) Получение оценок

e) Согласование полученных результатов с заинтересованными лицами

f) Проверка статистической значимости модели

3. Глубокое обучение включает в себя:

- а) Регрессионные модели
- б) Совокупность различных нейросетевых моделей
- в) Методы классификации
- г) Градиентный бустинг
- д) Обучение с подкреплением

4. Какой метод верификации исходных данных не применяется для верификации данных о стоимости активов:

- а) семантические анализаторы
- б) матрицы граничных значений
- в) конверторы отраслевых классификаторов
- г) наборы решающих правил
- д) проверка данных с использованием колл-центра
- е) тестовые и валидационные выборки

5. Какие нейронные сети лучше подходят для задач поиска аналога исследуемого объекта?

- а) Сети Кохонена
- б) Сети встречного распространения
- в) RBF сети на радиальных базисных функциях
- г) Любые MLP нейросети
- д) Все выше перечисленное

6. Какая проблема решается путем логарифмического шкалирования исходных данных?

- а) мультиколлинеарности
- б) робастности
- в) гетероскедастичности
- г) гомоскедастичности

7. Какие требования к факторам предъявляют классические статистические модели:

- а) значимость
- б) независимость
- в) внятная экономическая интерпретация
- г) все вышеперечисленное

8. При возникновении новых, не описанных ранее ситуаций, какая технология машинного обучения реагирует на них, получая данные из внешней среды?

- а) Обучение с подкреплением
- б) Обучение с противником
- в) Вероятностное прогнозирование
- г) Распознавание образов

9. Как не используют выборки из генеральной совокупности аналитики больших данных:

- а) Как метод формирования комплексного суждения о генеральной совокупности случайной величины
- б) Как метод тестирования полученных моделей
- в) Как метод верификации исходных данных

10.Метод главных компонент применяется для решения проблемы:

- a) Робастности
- b) Мультиколлинеарности
- c) Гомоскедастичности
- d) Гетероскедастичности

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

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

Перечень компетенций, формируемых в процессе освоения дисциплины, содержится в разделе 2. Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения, соотнесенных с планируемыми результатами обучения по дисциплине.

Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
<b>ПКП-5</b> 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.	Знание: Специальных программных продуктов для выполнения различных функций в экономическом субъекте Умение: Использовать и применять специальные ПП для выполнения бухгалтерско-аналитических и контрольных функций в экономическом субъекте	Осуществите сбор статистической информации по выбранной самостоятельно компании для построения системы прогнозирования продаж. Представьте эти данные в виде файла формата csv. Загрузите полученный файл в студию машинного обучения Microsoft Azure ML. Проведите первичную обработку данных. Постройте систему прогнозирования продаж.
	2.Демонстрирует владение специальными программными продуктами, применяемых для выполнения бухгалтерско-аналитических и контрольных	Знание: возможные варианты организации работы бухгалтерской службы Умение: владеть специальными программными продуктами в	Осуществите сбор статистической информации по выбранной самостоятельно компании для построения системы прогнозирования продаж. Представьте эти данные в виде файла формата csv. Загрузите полученный файл в студию машинного обучения Microsoft Azure ML. Проведите первичную обработку данных. Осуществите визуализацию полученных результатов. Постройте связанную с данными модель

	функций в экономическом субъекте.	бухгалтерском учете; формулировать и распределять функциональные обязанности между сотрудниками бухгалтерской службы.	машинного обучения. Создайте соответствующий web-сервис.
--	-----------------------------------	---	--

Теоретические вопросы для подготовки к зачету

1. Интеллектуальный анализ данных, большие данные и машинное обучение: основные понятия и методы.
2. Области применения методов и технологий интеллектуального анализа данных, машинного обучения и обработки больших данных. Примеры задач.
3. Основные характеристики больших данных и их влияние на сбор, хранение, обработку и анализ данных (4V).
4. Критерии аналитических задач, решение которых предпочтительно с использованием технологий Big Data.
5. Цикл обработки данных.
6. Статистические основы обработки данных.
7. Современные технологии машинного обучения.
8. Принципы разработки и оценки систем машинного обучения.
9. Основные классы моделей машинного обучения.
10. Модель множественной линейной регрессии.
11. Модель логистической регрессии.
12. Методы оценки качества моделей классификации.
13. Модели деревьев и лесов решений и их компьютерная реализация. Усиление (бустинг) деревьев решений.
14. Нейронные сети и их компьютерная реализация.
15. Машины опорных векторов и их компьютерная реализация.
16. Технологии улучшения моделей машинного обучения.
17. Кластерный анализ. Методы оценки качества моделей кластерного анализа.
18. Методы снижения размерности и их использование.
19. Метод главных компонент. Его геометрическое и экономическое содержание.
20. Поиск аномалий.
21. Принципы анализа текстовой информации.
22. Принципы анализа графической информации.
23. Принципы анализа эмоциональной окраски текстов.
24. Принципы создания рекомендательных систем.
25. Интеллектуальные сервисы и чат-боты.
26. Перспективы развития систем обработки больших данных и машинного обучения.
27. Финансовые технологии, основанные на обработке данных и машинном обучении.

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

### **Основная литература:**

1. Миркин, Б.Г. Введение в анализ данных: учебник и практикум для бакалавриата и магистратуры / Б.Г. Миркин. - Москва: Юрайт, 2019. - 174 с. — Текст : непосредственный. —То же -2020. —ЭБС Юрайт. —URL: <https://urait.ru/bcode/450262> (дата обращения: 09.09.2021). —Текст : электронный
2. Соловьев, В.И. Анализ данных в экономике: теория вероятностей, прикладная статистика, обработка и визуализация данных в Microsoft Excel: учебник для направления бакалавриата "Экономика и управление" / В.И. Соловьев; Финуниверситет. - Москва: Кнорус, 2019. - 498 с. - Текст : непосредственный. - То же. - 2021. - ЭБС BOOK.ru. -URL: <https://book.ru/book/938856> (дата обращения: 09.09.2021). – Текст : электронный.

### **Дополнительная литература:**

3. Калинина, В.Н. Анализ данных: Компьютерный практикум: учебное пособие / В.Н. Калинина, В.И. Соловьев. — Москва: КНОРУС, 2017. — 166 с. — Текст: непосредственный. — То же. — ЭБС BOOK.ru. — URL: <https://www.book.ru/book/929386> (дата обращения: 09.09.2021). — Текст: электронный.
4. Колдаев, В. Д. Структуры и алгоритмы обработки данных : учебное пособие / В. Д. Колдаев. -Москва : РИОР : ИНФРА-М, 2020. - 296 с. -Текст : электронный. - URL: <https://znanium.com/catalog/product/1054007> (дата обращения: 09.09.2021). – Режим доступа: по подписке
5. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва: Альпина Пабл., 2016. - 461с.- Текст: электронный - URL: <https://znanium.com/catalog/product/551044> (дата обращения: 09.09.2021). – Режим доступа: по подписке
6. Ростовцев, В.С. Искусственные нейронные сети : учебник / В.С. Ростовцев. — Санкт-Петербург : Лань, 2021. — 216 с. — ЭБС Лань. — URL: <https://e.lanbook.com/reader/book/160142> (дата обращения: 09.09.2021). —Текст: электронный.
7. Галушкин, А. И. Нейронные сети: основы теории / А.И. Галушкин. - Москва : Гор. линия-Телеком, 2012. - 496 с. – ЭБС ZNANIUM.com. -URL: <https://znanium.com/catalog/product/353660> (дата обращения: 09.09.2021). - Текст: электронный.

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

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>
2. Сайт кафедры Департамента анализа данных и машинного обучения. <http://findata.fa.ru>
3. Электронная библиотека Финансового университета (ЭБ) [http://elib.fa.ru/\(http://library.fa.ru/files/elibfa.pdf\)](http://elib.fa.ru/(http://library.fa.ru/files/elibfa.pdf))
4. Документация по Azure ML <https://docs.microsoft.com/ruru/azure/machine-learning/studio/>

5. Applied Machine Learning / Microsoft. – <https://www.edx.org/course/appliedmachine-learning-microsoft-dat203-3x-1>
6. Data Science Essentials / Microsoft. – <https://www.edx.org/course/datascience-essentials-microsoft-dat203-1x-3>
7. Principles of Machine Learning / Microsoft. – <https://www.edx.org/course/principles-machine-learning-microsoft-dat203-2x-3>
8. Профессиональный ресурс по машинному обучению. – <https://machinelearning.ru>
9. Специализация «Машинное обучение и анализ данных» / МФТИ и Яндекс. – <https://www.coursera.org/specializations/machine-learning-dataanalysis>
10. Электронно-библиотечная система Znanium <http://www.znanium.com>

## **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 психологофизиологическими 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 социальноактивные 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 опорнодвигательного 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 ESET NOD32
2. Windows, Microsoft 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 «Технологии 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.