

Федеральное государственное образовательное бюджетное учреждение  
высшего образования

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ  
РОССИЙСКОЙ ФЕДЕРАЦИИ»  
(Финансовый университет)**

**Краснодарский филиал Финуниверситета**

Кафедра «Математика и информатика»

СОГЛАСОВАНО

ООО «Портал-Юг»  
Генеральный директор



Е.В. Мостовой

«20» февраля 2024 г.

УТВЕРЖДАЮ

Краснодарский филиал  
Финуниверситета

Директор

Э.В.Соболев

«20» февраля 2024 г.

Калайдин Е.Н.

**Б1.В.02.07 Технологии машинного обучения**

**Рабочая программа дисциплины**

для студентов, обучающихся по направлениям подготовки  
38.03.05 «Бизнес-информатика» профиль «Бизнес-аналитика»  
(очно-заочная форма обучения)

*Рекомендовано Ученым советом Краснодарского филиала Финуниверситета  
(протокол № 12 от 20.02.2024)*

*Одобрено кафедрой «Математика и информатика»  
(протокол № 13 от 27.02.2024)*

**Краснодар 2024**

**УДК 004.62**

**ББК: 16.63**

**К-17**

Рецензенты: доктор физико-математических наук, профессор кафедры «Математика и информатика» Калайдин Е.Н., кандидат педагогических наук, доцент кафедры «Математика и информатика» Пьянкова Н.Г.

Калайдин Е.Н. «Технологии машинного обучения». Рабочая программа дисциплины для студентов, обучающихся по направлениям 38.03.05 «Бизнес-информатика». — Краснодар: Краснодарский филиал Финуниверситета, кафедра «Математика и информатика», 2024 г.

Рабочая программа определяет требования к результатам освоения дисциплины, ее содержание, тематику практических и семинарских занятий и технологии их проведения, формы самостоятельной работы, учебно-методическое обеспечение дисциплины.

38.03.05 Бизнес-информатика

*Учебное издание*

Калайдин Е.Н.

## **ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ**

Компьютерный набор, верстка: Калайдин Е.Н.

Гарнитура *Times New Roman*

© Калайдин Е.Н. 2024

© Краснодарский филиал Финуниверситета, 2024

## СОДЕРЖАНИЕ

1. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине .....	4
2. Место дисциплины в структуре образовательной программы .....	6
3. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся .....	6
4. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий .....	7
4.1 Содержание дисциплины .....	7
4.2 Учебно-тематический план .....	10
4.3 Содержание семинаров, практических занятий .....	11
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине .....	12
5.1 Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы .....	12
5.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю .....	15
6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине .....	17
7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины .....	17
8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины .....	26
9. Методические указания для обучающихся по освоению дисциплины .....	27
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем .....	27
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине .....	27

## 1. Наименование дисциплины

Б1.В.02.06 «Технологии машинного обучения»

## 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	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ПКН-1	Способность внедрять транзакционные системы и консультировать по вопросам систем сбора, накопления и хранения транзакционных данных	1. Проводит анализ рынка систем сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Методы и способы сбора, накопления и хранения транзакционных данных. Современный рынок систем сбора, накопления и хранения транзакционных данных, особенности этих систем. <b>Уметь:</b> Применять методы и инструменты для сбора, накопления и хранения транзакционных данных.
		2. Внедряет системы сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Методы и способы сбора, накопления и хранения транзакционных данных. Современный рынок систем сбора, накопления и хранения транзакционных данных, особенности этих систем. <b>Уметь:</b> Внедрять системы сбора, накопления и хранения транзакционных данных.
		3. Консультирует по вопросам применения систем сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Систему ключевых особенностей систем сбора, накопления и хранения транзакционных данных <b>Уметь:</b> Представлять ключевые особенности применения систем сбора, накопления и хранения транзакционных данных
ПКН-2	Способность	1. Анализирует	<b>Знать:</b>

	анализировать и проектировать информационные потоки организации	информационные потоки организации.	Ключевые требования к продуктам ИТ-предпринимательства. Способы их выявления. <b>Уметь:</b> Выявлять и оценивать ключевые требования к продуктам ИТ-предпринимательства.
		2. Создают модели «как есть» и «как должно быть» информационных потоков организации.	<b>Знать:</b> Ключевые требования к продуктам ИТ-предпринимательства и показатели для их разработки и продвижения. <b>Уметь:</b> Использовать ключевые требования к продуктам ИТ-предпринимательства для консультирования по вопросу разработки и продвижения стартапов в ИТ и других результатов деятельности в сфере предпринимательства.
<b>ПКН-3</b>	Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применяет аналитические системы работы с данными.	<b>Знать:</b> методы и принципы анализа данных, особенности применения актуальных продуктов для анализа данных. <b>Уметь:</b> выполнять анализ данных с применением подходящих программных продуктов.
		2. Проводит анализ рынка аналитических систем работы с данными.	<b>Знать:</b> назначение и свойства решений, используемых для анализа больших наборов данных. <b>Уметь:</b> выполнять анализ рынка аналитических систем работы с данными в соответствии с потребностями организации.
		3. Консультирует по вопросам применения аналитических систем	<b>Знать:</b> основные принципы применения аналитических систем работы с данными в организациях. <b>Уметь:</b>

		работы с данными.	формировать предложения и рекомендации по применению аналитических систем работы с данными для решения соответствующих организационных задач.
<b>ПКП-2</b>	Способность формировать требования для проектов по созданию продуктов ИТ-предпринимательства	1. Выявляет ключевые требования к продуктам ИТ-предпринимательства.	<b>Знать:</b> - Основные особенности текущей бизнес-модели компании <b>Уметь:</b> - Разрабатывать и обосновывать варианты целевых бизнес-моделей компании путем внедрения корпоративной системы
		2. Консультирует по вопросу разработки и продвижения стартапов в ИТ и других результатов деятельности в сфере предпринимательства.	<b>Знать:</b> - функциональные возможности корпоративных систем; - базовые принципы ведения проектов и внедрения корпоративных систем. <b>Уметь:</b> - консультировать заказчиков по вопросам внедрения и ведения проектов корпоративных систем на практике, ее базовых принципов и технологий работы.

### 3. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии машинного обучения» является дисциплиной Профиля Высокопроизводительные вычисления в цифровой экономике по направлению подготовки «Бизнес-аналитика» направления подготовки 38.03.05 «Бизнес-информатика».

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

Вид учебной работы по дисциплине	Всего (в з.е. и часах)	Семестр 7 (в часах)
Общая трудоёмкость дисциплины (в том числе курсовая работа)	5/180	5/180
Контактная работа- Аудиторные занятия	28	28

Лекции	12	12
Семинары, практические занятия	16	16
Самостоятельная работа	152	152
Вид промежуточной аттестации	экзамен	экзамен

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

### **5.1 Содержание дисциплины**

#### **Тема 1. Информационные технологии анализа данных**

Задачи анализа данных, понятие набора данных (dataset). Подготовительные операции для выполнения анализа данных: загрузка данных, трансформация данных, изучение данных, очистка данных, визуализация данных.

Технологический стек анализа данных, построенный на базе языка программирования Python. Язык программирования Python: основные характеристики, возможности языка для решения задач анализа данных и машинного обучения. Версии языка программирования Python, дистрибутивы и библиотеки Python. Знакомство с дистрибутивом Anaconda и составом инструментов для задач анализа данных и машинного обучения, входящих в дистрибутив. Интерактивная оболочка IPython notebook: принципы работы и применение для решения задач анализа данных и машинного обучения.

Знакомство с библиотеками numpy и pandas и решением базовых задач подготовительных операций для выполнения анализа данных с помощью этих библиотек.

**Тема 2. Введение в машинное обучение и знакомство с технологиями машинного обучения.**

Постановки задач машинного обучения. Объекты и признаки. Типы признаков: бинарные, номинальные, порядковые, количественные. Типы задач машинного обучения: классификация, регрессия, прогнозирование, кластеризация. Примеры задач решаемых методами машинного обучения. Проблема недообучения / переобучения.

Оценка качества результатов: тестирование, разбиение набора данных для обучения и тестирования, кроссвалидация. Метрики качества результатов: Precision, Recall, F1, ROC. Критерии качества классификации: чувствительность и специфичность, ROC-кривая и AUC, точность и полнота, AUCPR. Внутренние и внешние критерии. Эмпирические и аналитические критерии.

Библиотека Scikit-Learn для решения задач машинного обучения. Обзор основных возможностей библиотеки. Примеры решения задач машинного обучения с помощью библиотеки. Работа со справочными материалами по библиотеке.

#### **Тема 3. Методы обучения с учителем.**

Метрические методы классификации и специфика их области применения. Гипотеза компактности и непрерывности. Метод ближайших соседей (kNN) и его обобщения. Подбор числа k. Методы метрической

классификации в библиотеке Scikit-Learn.

Логические методы классификации. Решающие списки и деревья. Решающее правило. Алгоритм синтеза списка. Решающее дерево и алгоритм ID3. Редукция решающих деревьев. Преимущества и недостатки решающих списков и деревьев. Решающий лес. Свойства решающего леса. Логические методы классификации в библиотеке Scikit-Learn.

Байесовские классификаторы. Вероятностная постановка задачи классификации. Принцип максимума апостериорной вероятности. Допущения наивного байесовского классификатора. Наивный байесовский классификатор. Байесовские классификаторы в библиотеке Scikit-Learn.

Метод опорных векторов. Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Задача квадратичного программирования и двойственная задача. Понятие опорных векторов. Функция ядра и примеры функций ядер. Регуляризация для отбора признаков. Метод опорных векторов в библиотеке Scikit-Learn.

Тема 4. Методы восстановления регрессии.

Задача регрессии, многомерная линейная регрессия. Метод наименьших квадратов, его вероятностный смысл и геометрический смысл, сингулярное разложение. Проблемы мультиколлинеарности и переобучения. Штрафы L1 и L2, Elastic Net. Сравнение различных штрафов. Линейная и многомерная регрессия в библиотеке Scikit-Learn.

Гипотеза экспоненциальности функций правдоподобия классов. Логистическая регрессия. Принцип максимума правдоподобия и логарифмическая функция потерь. Метод стохастического градиента для логарифмической функции потерь. Бинаризация признаков. Логистическая регрессия в библиотеке Scikit-Learn.

Тема 5. Обучение без учителя, кластеризация.

Постановка задачи обучения без учителя. Примеры прикладных задач обучения без учителя.

Задача кластеризации. Прикладные задачи кластеризации. Типы кластерных структур. Алгоритм k средних (k-means). Другие методы кластеризации. Метрики в задачах кластеризации. Алгоритм построения дендрограммы. Определение числа кластеров. Методы кластеризации в библиотеке Scikit-Learn.

Задача снижения размерности. Метод главных компонент. Прикладные задачи снижения размерности. Методы снижения размерности и метод главных компонент в библиотеке Scikit-Learn.

Тема 6. Коллаборативная фильтрация, отбор признаков.

Задачи коллаборативной фильтрации, исходные данные и их матричное представление. Корреляционные методы user-based, item-based. Различные меры сходства субъектов и объектов. Проблемы практического применения корреляционных методов для решения задач большой размерности. Реализации методов коллаборативной фильтрации на языке Python.

Латентные методы на основе матричных разложений. Метод главных компонент для разреженных данных. Метод стохастического градиента. Неотрицательные матричные разложения. Разреженный SVD. Метод



чередующихся наименьших квадратов ALS. Реализации методов коллаборативной фильтрации на языке Python.

Скольльзящий контроль, разновидности эмпирических оценок скользящего контроля. Критерий непротиворечивости.

Разновидности аналитических оценок. Регуляризация. Критерий Акаике (AIC), байесовский информационный критерий (BIC). Статистические критерии: коэффициент детерминации, критерий Фишера, анализ регрессионных остатков.

Сложность задачи отбора признаков. Полный перебор. Метод добавления и удаления, шаговая регрессия. Поиск в глубину, метод ветвей и границ, усечённый поиск в ширину. Генетический алгоритм отбора признаков. Реализация алгоритмов отбора признаков на языке Python.

Тема 7. Генетические алгоритмы.

Эволюционное моделирование. Технология генетических алгоритмов. Задачи, решаемые методами генетических алгоритмов. Задача коммивояжера. Задачи множественной оптимизации.

Основные понятия генетического алгоритма. Популяция, хромосома, ген, поколение, функция приспособленности.

Основные этапы генетического алгоритма: формирование популяции, скрещивание или почкование, мутация, отбор. Критерии остановки алгоритма.

Особенности алгоритма для различного типа задач.

Преимущества и ограничения технологий генетических алгоритмов.

Тема 8. Анализ сетевых структур.

Знакомство с представлением данных из реальных приложений в виде сетевых структур (графов). Знакомство с технологическим стеком для работы с данными в виде сети: графовые базы данных (на примере neo4j), библиотека python NetowrkX для работы с графами, пакет Gephi для интерактивной работы с сетевыми структурами.

Работа с графовыми базами данных, язык запросов для графовой СУБД neo4j Cypher. Бессхемная модель данных. Вопросы эффективного представления данных в графовых базах данных, сравнение графовой и реляционной модели.

Основные свойства сетей: степени узлов, распределение степеней узлов, коэффициенты кластеризации, мотивы в сети, диаметр и средняя длина пути в сети. Основные модели теории сложных сетей (ТСС): модель тесного мира, модель безмасштабной сети, самоподобная сеть, случайная сеть.

Форматы файлов для работы с сетевыми структурами. Базовые операции с сетями с помощью библиотеки NetowrkX. Визуализация сетей. Использование Gephi для визуализации и интерактивной работы с сетями.

Тема 9. Анализ текстов на естественном языке.

Основные задачи анализа текстов на естественном языке. Основные подходы к анализу текста. Библиотеки nltk и rumorphy2 - основные возможности для решения базовых задач.

Подходы к решению задач классификации текстов. Алгоритм TF-IDF.

Тема 10. Визуализация результатов анализа данных.

Проблемы представления результатов анализа. Формы представления

результатов анализа: табличная, иерархическая, графическая, многомерная.

Визуальные способы оценки совокупности данных: гистограмма, полигон частот, диаграмма размаха («ящик с усами»), скрипичная диаграмма, тепловая карта.

Специфические формы представления: деревья, наборы правил, наборы карт, диаграммы рассеяния.

Способы визуализации многомерной информации.

Представление и интерпретация результатов, выявленных различными методами анализа данных.

## 5.2 Учебно-тематический план

№ п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа- Аудиторная работа			Самостоят ельная работа	
			Общая, в т.ч.:	Лекции	Семинары, практические занятия		
1	Информационные технологии анализа данных	36	4	2	2	32	Участие в решении задач на практических занятиях. Обсуждения по результатам самостоятельно й работы Участие в решении задач на практических занятиях. Обсуждения по результатам самостоятельно й работы
2	Введение в машинное обучение и знакомство с технологиями машинного обучения.	14	4	2	2	10	
3	Методы обучения с учителем.	12	2	1	1	10	
4	Методы восстановления регрессии.	12	2	1	1	10	
5	Обучение без учителя, кластеризация.	12	2	1	1	10	
6	Коллаборативная фильтрация, отбор признаков.	13	3	1	2	10	
7	Генетические алгоритмы.	23	3	1	2	20	
8	Анализ сетевых структур.	23	3	1	2	20	Участие в решении задач на практических занятиях. Обсуждения по результатам
9	Анализ текстов на естественном языке.	23	3	1	2	20	

10	Визуализация результатов анализа данных.	12	2	1	1	10	самостоятельно й работы
	В целом по дисциплине	180	28	12	16	152	Согласно учебному плану

### 5.3 Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Информационные технологии анализа данных	Входной контроль. Изучение технологического стека анализа данных, построенного на базе языка программирования Python. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Введение в машинное обучение знакомство технологиями машинного обучения.	Решение задач машинного обучения с использованием библиотеки Scikit-Learn. Оценка качества результатов: тестирование, разбиение набора данных для обучения и тестирования, кроссвалидация. Расчет метрик качества результатов. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Методы обучения с учителем.	Решение задач машинного обучения с использованием метрических методов классификации при помощи библиотеки ScikitLearn. Решение задач машинного обучения с использованием решающих списков и деревьев при помощи библиотеки Scikit-Learn. Решение задач машинного обучения с использованием байесовских классификаторов при помощи библиотеки Scikit-Learn. Решение задач машинного обучения с использованием метода опорных векторов при помощи библиотеки Scikit- Learn. Рекомендуемые источники: 8[1-5]	Решение и обсуждение задач
Методы восстановления регрессии.	Решение задач восстановления регрессии при помощи библиотеки ScikitLearn. Решение задач машинного обучения с использованием логистической регрессии при помощи библиотеки ScikitLearn. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Обучение без учителя, кластеризация.	Решение задач кластеризации при помощи библиотеки Scikit-Learn (алгоритм k средних (kmeans), другие методы кластеризации). Решение задач снижения размерности при помощи библиотеки ScikitLearn (метод главных компонент). Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Коллаборативная фильтрация, отбор признаков.	Решение задач коллаборативной фильтрации на языке Python (корреляционные методы, латентные методы). Реализация скользящего контроля на Python. Решение задачи отбора признаков на Python. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач

Генетические алгоритмы.	Выполнение аналитических задач с использованием технологий генетических алгоритмов. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Анализ сетевых структур.	Решение задач создания (в т.ч. загрузки из внешних источников и генерации) сетевых структур. Извлечение данных, выполнение запросов к графовой СУБД. Определение свойств сети, визуализация сети. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Анализ текстов на естественном языке.	Токенизация, нормализация текста. Векторное представление текста, решение задач с помощью алгоритма TFIDF. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач
Визуализация результатов анализа данных.	Выполнение заданий по визуализации результатов анализа. Тематические дискуссии на темы: какие методы визуализации помогают улучшить понимание выявленных закономерностей; можно ли считать саму визуализацию инструментом анализа данных. Рекомендуемые источники: 8.[1-5]	Решение и обсуждение задач

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

### 6.1 Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Информационные технологии анализа данных	Знакомство с интерактивной оболочкой IPython notebook. Изучение принципов работы и применения для решения задач анализа данных и машинного обучения. Знакомство с библиотеками numpy и pandas и решением базовых задач подготовительных операций для выполнения анализа данных с помощью этих библиотек	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Введение в машинное обучение и знакомство с технологиями машинного обучения.	Расчет метрик качества Precision, Recall, F1, RUC. Критерии качества классификации: чувствительность и специфичность, ROC-кривая и AUC, точность и полнота, AUC-PR	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.

Методы обучения с учителем.	Решение задач машинного обучения с использованием метрических методов классификации при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами метрических алгоритмов). Решение задач машинного обучения с использованием решающих списков и деревьев при помощи библиотеки ScikitLearn (сравнение решения задач с различными параметрами и видами алгоритмов решающих списков и деревьев). Решение задач машинного обучения с использованием байесовских классификаторов при помощи библиотеки ScikitLearn (сравнение решения задач с различными параметрами и видами байесовских алгоритмов). Решение задач машинного обучения с использованием метода опорных векторов при помощи библиотеки Scikit-Learn (сравнение решения задач с различными параметрами и видами алгоритмов метода опорных векторов).	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Методы восстановления регрессии.	Решение задач восстановления регрессии при помощи библиотеки Scikit-Learn с штрафами L1 и L2, Elastic Net. Решение задач машинного обучения с использованием логистической регрессии при помощи библиотеки ScikitLearn.	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Обучение без учителя, кластеризация.	Решение задач кластеризации при помощи библиотеки Scikit-Learn (использование методов, альтернативных алгоритму k средних). Решение задач снижения размерности при помощи библиотеки ScikitLearn (использование методов, альтернативных методу главных компонент).	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.

Коллаборативная фильтрация, отбор признаков.	Решение задач коллаборативной фильтрации на языке Python корреляционными методами с различными метриками. Тестирование альтернативных алгоритмов решения задачи отбора признаков.	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Генетические алгоритмы.	Генетические алгоритмы в прикладных задачах. Преимущества и ограничения технологий генетических алгоритмов.	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Анализ сетевых структур.	Выявление базовых моделей теории сложных сетей (безмасштабная сеть, тесный мир, случайных граф) на реальных данных сетевых структур существующих в приложениях.	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Анализ текстов на естественном языке.	Решение задач выявления свойств в тексте на естественном языке. Классификация текстов	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий.
Визуализация результатов анализа данных.	Современные методы визуализации аналитической информации. Краткое описание данных на основе визуализации	Разбор теоретического материала по учебникам, учебным пособиям, конспектам лекций или видео лекциям. Самостоятельное изучение указанных теоретических вопросов. Решение задач по темам практических занятий;

## **6.2 Перечень вопросов, заданий, тем для подготовки к текущему контролю**

### **Примерная тематика курсовой работы**

1. Продвинутая аналитика крупномасштабного графового датасета.
2. Решение задач машинного обучения на графовом датасете.
3. Предварительный анализ данных и построение признаков в задачах визуализации информации.
4. Методы и модели машинного обучения в экспертных системах.
5. Построение рекомендательной системы в области закупок с учетом характеристик контракта.
6. Машинное обучение в задачах автоматизации наукометрической деятельности.
7. Предварительный анализ данных и построение признаков в задачах сжатия информации.
8. Предварительный анализ данных и построение признаков в задачах идентификации личности по изображению.
9. Дообучение модели GPT-3 на корпусе текстов юридического характера
10. Машинное обучение в задачах анализа изображений.
11. Разработка диалоговой системы с применением обучения с подкреплением.
12. Предварительный анализ данных и построение признаков в задачах машинного перевода.
13. Предварительный анализ данных и построение признаков в задачах распознавания текста.
14. Применение методов машинного обучения в процессе создания и эксплуатации баз знаний.
15. Машинное обучение в задачах автоматизации наукометрической деятельности.
16. Использование методов машинного обучения (ML) в скоринговых моделях анализа кредитоспособности заемщика банка - физического лица.
17. Обработка данных о научных изданиях и статистики цитирования.
18. Использование методов машинного обучения для решения задачи выделения атрибутов товаров из текста.
19. Использование методов машинного обучения для классификации текстов описаний товаров.
20. Машинное обучение в задачах оценки импакт-фактора научных журналов
21. Кластеризация информации о научной аффилиации в системах научного цитирования.
22. Машинное обучение для автоматизации сбора информации о научных изданиях и цитировании.
23. Предварительный анализ данных и построение признаков в задачах распознавания объектов на фотографии.
24. Кластеризация наукометрических данных в системах научного

цитирования.

25. Предварительный анализ данных и построение признаков в задачах выявления механических повреждений клубней картофеля.

26. Построение онтологических эмбедингов графов знаний.

27. Использование методов машинного обучения (ML) для анализа финансового состояния корпоративного заемщика банка.

28. Использование методов машинного обучения (ML) при анализе поведения клиента в целях повышения объема продаж товаров и услуг.

29. Использование методов машинного обучения (ML) при формировании сегментов клиентов в CRM системах.

30. Регрессионные модели и методы для прогнозирования биржевых курсов

31. Применение методов и моделей машинного обучения для построения индекса недвижимости.

32. Машинное обучение в задачах прогнозирования транспортного трафика

33. Машинное обучение в задачах кредитного скоринга.

34. Проверка статистических гипотез на больших наборах данных о биржевой активности.

35. Поиск заимствований в больших наборах данных текстовой информации.

36. Дообучение модели GPT-3 на корпусе текстов юридического характера.

37. Использование методов машинного обучения для классификации изображений автомобилей.

38. Использование методов машинного обучения для решения задачи классификации автомобилей такси.

39. Система для сбора и генерации постов в социальных сетях.

40. Предварительный анализ данных и построение признаков в задачах выявления болезней комнатных растений. Примечания: сорт растения обсуждается.

41. Предварительный анализ данных и построение признаков в задачах выявления наличия соланина в клубнях картофеля.

42. Применение нейросетевых технологий в задачах машинного обучения.

43. Кластеризация наукометрических данных в системах научного цитирования.

44. Кластеризация информации о научной аффилиации в системах научного цитирования.

45. Предварительный анализ данных и построение признаков в задаче классификации эмоций человека по голосу.

46. Кластеризация наукометрических данных в системах научного цитирования.

47. Машинное обучение в задачах обработки наукометрической информации.

48. Машинное обучения для анализа и парсинга новостей.



49. Машинное обучение в рекомендательных системах.
50. Машинное обучение в задачах анализа изображений.
51. Машинное обучение в задачах анализа текста.
52. Методы машинного обучения для поиска паттернов в видео.
53. Анализ текстовых признаков в задачах машинного обучения.
54. Машинное обучение в оптимизации инвестиционных проектов.

### **Примеры практических заданий**

#### **Задание 1.**

Дан набор данных "Boston house prices". Напишите скрипт с использованием библиотеки `scikit learn`, который определяет относительную важность признаков для определения средней цены дома в городе. Признаками взять все атрибуты датасета, кроме целевой переменной.

#### **Задание 2.**

Дан набор данных "Boston house prices". Напишите скрипт с использованием библиотеки `scikit learn`, который строит парную линейную регрессию между средней ценой дома и уровнем преступности на душу населения. Оцените коэффициент детерминации такой регрессии и сделать вывод о ее статистической значимости.

#### **Задание 3.**

Дан набор данных "Boston house prices". Напишите скрипт с использованием библиотеки `scikit learn`, который предсказывает среднюю цену дома методом множественной линейной регрессии по совокупности остальных атрибутов. Сделайте вывод о статистической значимости такой регрессии.

#### **Задание 4.**

Дан набор данных "Iris". Напишите скрипт с использованием библиотеки `scikit learn`, который определяет наилучшие метапараметры какого-либо алгоритма классификации вида растения (алгоритм классификации - на выбор студента) методов перекрестной проверки.

#### **Задание 5.**

Дан набор данных "Iris". Напишите скрипт с использованием библиотеки `scikit learn`, который определяет относительную важность признаков и на основании результатов его выполнения постройте модель классификации на основании двух наиболее значимых признаках.

*Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Департамента анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных.*

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

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Технологии машинного обучения».

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля и промежуточной аттестации.

## 7. 1. Описание показателей и критериев оценивания компетенций

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
<b>ПKN-1</b>	Способность внедрять транзакционные системы и консультировать по вопросам систем сбора, накопления и хранения транзакционных данных	1. Проводит анализ рынка систем сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Методы и способы сбора, накопления и хранения транзакционных данных. Современный рынок систем сбора, накопления и хранения транзакционных данных, особенности этих систем. <b>Уметь:</b> Применять методы и инструменты для сбора, накопления и хранения транзакционных данных.
		2. Внедряет системы сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Методы и способы сбора, накопления и хранения транзакционных данных. Современный рынок систем сбора, накопления и хранения транзакционных данных, особенности этих систем. <b>Уметь:</b> Внедрять системы сбора, накопления и хранения транзакционных данных.
		3. Консультирует по вопросам применения систем сбора, накопления и хранения транзакционных данных.	<b>Знать:</b> Систему ключевых особенностей систем сбора, накопления и хранения транзакционных данных <b>Уметь:</b> Представлять ключевые особенности применения систем сбора, накопления и хранения транзакционных данных
<b>ПKN-2</b>	Способность анализировать и проектировать информационные потоки организации	1. Анализирует информационные потоки организации.	<b>Знать:</b> Ключевые требования к продуктам ИТ-предпринимательства. Способы их выявления. <b>Уметь:</b> Выявлять и оценивать

			ключевые требования к продуктам ИТ-предпринимательства.
		2. Создают модели «как есть» и «как должно быть» информационных потоков организации.	<b>Знать:</b> Ключевые требования к продуктам ИТ-предпринимательства и показатели для их разработки и продвижения. <b>Уметь:</b> Использовать ключевые требования к продуктам ИТ-предпринимательства для консультирования по вопросу разработки и продвижения стартапов в ИТ и других результатов деятельности в сфере предпринимательства.
<b>ПКН-3</b>	Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применяет аналитические системы работы с данными.	<b>Знать:</b> методы и принципы анализа данных, особенности применения актуальных продуктов для анализа данных. <b>Уметь:</b> выполнять анализ данных с применением подходящих программных продуктов.
		2. Проводит анализ рынка аналитических систем работы с данными.	<b>Знать:</b> назначение и свойства решений, используемых для анализа больших наборов данных. <b>Уметь:</b> выполнять анализ рынка аналитических систем работы с данными в соответствии с потребностями организации.
		3. Консультирует по вопросам применения аналитических систем работы с данными.	<b>Знать:</b> основные принципы применения аналитических систем работы с данными в организациях. <b>Уметь:</b> формировать предложения и рекомендации по применению аналитических систем работы с данными для решения соответствующих

			организационных задач.
<b>ПКП-2</b>	Способность формировать требования для проектов по созданию продуктов ИТ-предпринимательства	1. Выявляет ключевые требования к продуктам ИТ-предпринимательства.	<b>Знать:</b> - Основные особенности текущей бизнес-модели компании <b>Уметь:</b> - Разрабатывать и обосновывать варианты целевых бизнес-моделей компании путем внедрения корпоративной системы
		2. Консультирует по вопросу разработки и продвижения стартапов в ИТ и других результатов деятельности в сфере предпринимательства.	<b>Знать:</b> - функциональные возможности корпоративных систем; - базовые принципы ведения проектов и внедрения корпоративных систем. <b>Уметь:</b> - консультировать заказчиков по вопросам внедрения и ведения проектов корпоративных систем на практике, ее базовых принципов и технологий работы.

## 7.2. Вопросы для оценки знаний и умений, характеризующих формирование компетенций

Шифр компетенции	Вопросы	Правильный ответ
<b>ПКН-1</b>	1. Какие основные задачи решаются в рамках машинного обучения?	Задачи классификации, регрессии и кластеризации
	2. Что такое набор данных (датасет)?	Это коллекция данных
	3. Для чего используется линейная регрессия?	Для предсказания непрерывной целевой переменной
	4. Что называют обучением с подкреплением?	Обучение, основанное на собственном опыте
	5. В чем заключается основная идея генетических алгоритмов?	Использование эволюционных принципов для поиска оптимальных решений
<b>ПКН-2</b>	6. Какие методы имеет предварительный анализ данных?	Сплошное наблюдение, выборочное наблюдение, сравнение данных
	7. Какие три типа пропусков данных существуют?	MCAR, MAR, MNAR

	8. Что такое NearMiss?	Метод недостаточной выборки
	9. Как называются параметры, значения которых задается до начала обучения модели и не изменяется в процессе обучения?	Гиперпараметры модели
	10. Какие модели машинного обучения являются сетевыми?	Рекуррентная, сверточная нейронная сети, наивный Байесовский классификатор
<b>ПКН-3</b>	11. Напишите формулу такого показателя эффективности диагностики модели машинного обучения	Правильность = количество правильных прогнозов / общее количество прогнозов
	12. Какая функция активации обычно используется в выходном слое нейронной сети для задачи регрессии?	Линейная
	13. Какую роль играет переобучение в задаче машинного обучения и разработке прикладных средств анализа текста?	Снижение точности модели
	14. Что такое FastText?	Расширение Word2Vec
	15. Какие существуют методы визуализации данных?	Гистограммы, круговые и столбчатые диаграммы
<b>ПКП-2</b>	16. В каком типе задач машинного обучения используется метод опорных векторов?	Классификация
	17. Какую роль играет регуляризация в задаче машинного обучения?	Помогает улучшить обобщение модели
	18. Какие есть принципы нормирования данных?	Масштабирование, стандартизация, нормализация
	19. Какие методы используются для преобразования категориальных признаков в числовые?	One-Hot Encoding, Label Encoding
	20. Какие есть виды ансамблевых методов?	Стекинг, бэггинг, бустинг

### 7.3. Практико-ориентированные задания

Шифр компетенции	Практико-ориентированные задания	Правильный ответ
------------------	----------------------------------	------------------

ПКН-1	1. В базе данных имеется 10 000 записей. Сколько информации (в битах) необходимо для кодирования одной записи, если на каждый атрибут выделяется по 3 бита?	29 бит информации.
	2. В задаче регрессии имеется 500 точек данных. Сколько параметров требуется настроить в линейной регрессии, если каждая точка представлена двумя координатами (x, y)?	Необходимо настроить 1000 параметров.
	3. Имеется набор данных для классификации, состоящий из 50 000 примеров. Сколько раз необходимо выполнить итерацию обучения алгоритма k-ближайших соседей (k-NN), если в каждой итерации используется 10% набора данных для проверки?	Необходимо выполнить 10 итераций.
	4. Ваша задача разработать систему рекомендаций для электронной коммерции. На основе какой фильтрации будет оптимально это сделать?	Коллаборативной фильтрации
ПКН-2	5. Перед вами стоит задача построить модель прогнозированных временных рядов. Какие модельные формы вы будете для этого использовать?	ARIMA или LSTM
	6. Дан набор данных, состоящий из последовательностей действий пользователей на веб-сайте. Требуется классифицировать эти действия, то есть определить, какие из них являются целевыми, а какие – нет. Что для этого нужно сделать?	Определить цели, собрать данные, создать модель классификации и внедрить её на веб-сайте.
	7. Доктор хочет построить модель, прогнозирующую возможность развития осложнений у пациента после операции - все ли будет благополучно в течение нескольких следующих месяцев или нет. К какому типу относится данная задача?	Классификация
	8. Астроном хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача?	Кластеризации
ПКН-3	9. Заполните пропуск: «Для выпуклой функции .... указывает сторону наискорейшего убывания»	Градиент
	10. Что может являться критерием останова в градиентном спуске?	Достижение определенной точности
	11. Рассмотрим два объекта: у первого отступ линейного классификатора равен -10, у второго -1000. Как соотносятся два этих объекта?	Второй объект находится дальше от разделяющей поверхности, чем первый.
	12. Рассмотрим пользователя социальной	Предсказание пола

	сети как объект в задаче машинного обучения. Что может являться задачей классификации?	пользователя
	13. Ансамблевые методы обычно делают более качественные предсказания, чем отдельные алгоритмы (у ансамблей ниже ошибка на тестовой выборке). А какой показатель всегда ухудшается при ансамблировании?	Скорость выполнения предсказаний
ПКП-2	14. Алгоритм бэггинг подразумевает выбор случайных частей данных, обучение алгоритма (например, решающего дерева) на каждой части и составление ансамбля из обученных алгоритмов. Для чего нужно обучать алгоритмы на разных частях данных?	Для того, чтобы обученные алгоритмы выполняли несовпадающие (различные) предсказания
	15. Метеоролог хочет построить модель, прогнозирующую температуру воздуха на завтра. К какому типу относится данная задача?	Прогнозирования
	16. Верно ли суждение, что элементами обучающей выборки являются объекты, характеристики которых являются значениями признаков?	Да, верно
	17. Рассмотрим признак “Образовательная программа” при анализе данных по студентам университета. Этот признак может принимать три значения: “Экономика”, “Математика”, “Философия”. Воспользуемся one-hot кодированием и заменим этот признак на три бинарных, которые будут соответствовать категориям в том порядке, в котором они перечислены выше. Как будет закодирован признак со значением “Философия”?	(0,0,1)
	18. Чему будет равен корень из среднеквадратичной ошибки для набора из 3 наблюдений, где отклонение предсказания линейной регрессии от реальных значений равны: -1, 2, 2?	2

#### 7.4. Тесты

Шифр компетенции	Тестовые задания	Правильный ответ
ПКН-1	1. Какой из следующих методов машинного обучения используется для решения задачи классификации? а. Метод k-средних б. Наивный Байесовский метод с. Линейная регрессия d. Метод опорных векторов	d
	2. Какая из следующих моделей машинного обучения	d

	является регрессионной моделью? a. Наивный Байесовский метод b. Метод k-ближайших соседей c. Деревья решений d. Линейная регрессия	
	3. Какой из следующих алгоритмов машинного обучения основан на использовании деревьев решений? a. Метод опорных векторов b. Случайный лес c. Наивный Байесовский метод d. k-ближайших соседей	b
ПКН-2	4. Что является основным недостатком наивного Байесовского метода? a. Сложность в масштабировании b. Низкая точность c. Высокая вычислительная сложность d. Зависимость от распределения данных	d
	5. Что используется в методе k-ближайших соседей для определения расстояния между точками? a. Евклидово расстояние b. Манхэттенское расстояние c. Расстояние Чебышёва d. Расстояние Хемминга	a
	6. Какой из перечисленных методов относится к группе методов глубокого обучения? a. Метод k-means b. Метод опорных векторов c. Сверточная нейронная сеть d. Дерево решений	c
ПКН-3	7. Что такое предобучение в машинном обучении? a. Процесс обучения модели на большом наборе данных b. Процесс улучшения качества модели после ее обучения c. Процесс получения начальных весов модели для дальнейшего обучения d. Процесс отбора признаков для обучения модели	c
	8. Что такое метод градиентного спуска? a. Это метод оптимизации функции потерь в машинном обучении b. Это метод выбора признаков в модели c. Это метод кластеризации данных d. Это метод оценки качества модели	a
	9. Что такое ранняя остановка в обучении нейронных сетей? a. Остановка обучения после определенного количества эпох b. Остановка обучения, если качество модели на валидационной выборке не улучшается в течение определенного количества эпох c. Остановка обучения, если качество модели на обучающей выборке достигло максимума d. Остановка обучения, если модель достигла нулевого значения функции потерь	b
ПКП-2	10. В чем заключается принцип работы автоэнкодера? a. В сжатии данных b. В восстановлении данных	a



	с. В классификации данных d. В предсказании временных рядов	
	11. Что из следующего является особенностью сверточной нейронной сети? a. Наличие слоев с большим количеством нейронов b. Наличие слоёв, состоящих из сверточных фильтров c. Наличие только одного полносвязного слоя d. Отсутствие слоёв с нелинейными функциями активации	b
	12. Каким образом DropConnect улучшает обобщающую способность модели? a. Путем случайного отключения нейронов b. Путем случайного изменения весов между нейронами c. Путем уменьшения количества слоёв d. Путем увеличения количества скрытых слоев	a

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

### ***Основная литература:***

1. Моделирование систем и процессов. Практикум : учебное пособие для вузов / В. Н. Волкова [и др.] ; под редакцией В. Н. Волковой. — Москва : Юрайт, 2023. — 295 с. — (Высшее образование). — ЭБС Юрайт. — URL: <https://urait.ru/bcode/512216> (дата обращения: 10.05.2023). — Текст : электронный.

2. Колдаев, В. Д. Структуры и алгоритмы обработки данных : учебное пособие / В. Д. Колдаев. - Москва : РИОР : ИНФРА-М, 2021. - 296 с. - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1230215> (дата обращения: 10.05.2023). - Текст : электронный.

### ***Дополнительная литература:***

3. Гуриков, С. Р. Основы алгоритмизации и программирования на Python : учебное пособие / С. Р. Гуриков. — Москва : ИНФРА-М, 2022. — 343 с. — (Высшее образование: Бакалавриат). - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1356003> (дата обращения: 10.05.2023). - Текст : электронный.

4. Шрамко, Е. С. Эволюция искусственного интеллекта: от машинного обучения к глубокому обучению / Е. С. Шрамко, А. А. Пахомов, Л. В. Гаев // Кооперация науки и общества - путь к модернизации и инновационному развитию : Сборник статей Всероссийской научно-практической конференции с международным участием, Новосибирск, 28 мая 2024 года. — Уфа: ООО "Аэтерна", 2024. — С. 63-65. — EDN BJEKO.

5. 5. Городничев, Д. Ю. Машинное обучение и глубокое обучение / Д. Ю. Городничев // Современные проблемы лингвистики и методики преподавания русского языка в ВУЗе и школе. — 2022. — № 38. — С. 278-281. — EDN PFFNFS.

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

1. Pyru 1.0.9 [Электронный ресурс]:сайт. - Режим доступа: <https://pypi.python.org/pypi/pyru/40>
2. Python Data Analysis Library [Электронный ресурс]: сайт. - Режим доступа: <http://pandas.pydata.org/>
3. Python Documentation [Электронный ресурс]: сайт. - Режим доступа: <http://python.org/doc/>
4. Python Standard Library [Электронный ресурс]: сайт. - Режим доступа: <https://docs.python.org/2/library/>
5. Scikit-learn Machine Learning in Python [Электронный ресурс]: сайт. - Режим доступа: <http://scikit-learn.org>
6. Аналитические технологии для прогнозирования и анализа данных [Электронный ресурс]: сайт. - Режим доступа: <http://www.neuroproject.ru/neuro.php>
7. Deductor. Продвинутая аналитика без программирования [Электронный ресурс]: сайт. - Режим доступа: <https://basegroup.ru/deductor/description>
8. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
9. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
10. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
11. Электронно-библиотечная система Znanium <http://www.znanium.com>
12. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru/>
13. Деловая онлайн-библиотека Alpina Digital <http://lib.alpinadigital.ru/>
14. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
15. Национальная электронная библиотека [http://i \[эб.рф/](http://i[эб.рф/)
16. <https://campus.fa.ru/course/view.php?id=11164> (ЭУК)
17. <https://campus.fa.ru/course/view.php?id=12181> (ЭУК)

**Онлайн-курсы с получением сертификата или иного документа об окончании, которые разработаны другими образовательными организациями:**

18. <https://stepik.org/course/102480/promo>
19. [https://karpov.courses/analytics?utm\\_source=yandex&utm\\_medium=cpc&utm\\_campaign=Y\\_Data\\_analytics\\_general\\_search\\_msk&utm\\_term=анализ%20данных%20обучение&utm\\_content=k50id0100000032526554133\\_32526554133\\_cid\\_6297\\_3419\\_gid\\_4603485637\\_aid\\_10837803152\\_adp\\_no\\_pos\\_premium3\\_src\\_sear\\_ch\\_none\\_dvc\\_desktop\\_main&openstat=ZGlyZWN0LnJhbmRleC5ydT\\_s2Mjk3MzQxOTsxMDgzNzgwMzE1Mjt5YW5kZXguU6cHJlbW11bQ&yclid=9990705705238396927](https://karpov.courses/analytics?utm_source=yandex&utm_medium=cpc&utm_campaign=Y_Data_analytics_general_search_msk&utm_term=анализ%20данных%20обучение&utm_content=k50id0100000032526554133_32526554133_cid_6297_3419_gid_4603485637_aid_10837803152_adp_no_pos_premium3_src_sear_ch_none_dvc_desktop_main&openstat=ZGlyZWN0LnJhbmRleC5ydT_s2Mjk3MzQxOTsxMDgzNzgwMzE1Mjt5YW5kZXguU6cHJlbW11bQ&yclid=9990705705238396927)

## **10. Методические указания для обучающихся по освоению дисциплины**

При изучении теоретического материала необходимо опираться на рабочую программу дисциплины и литературу из основного списка. Кроме этого материала необходимо активно работать с Интернет-источниками и пособиями других авторов, помогающими усвоить материал отдельных разделов программы.

Практические занятия проводятся по следующей схеме: Студенты в индивидуальном режиме выполняют задания в компьютерном классе. Существуют две основные формы заданий. В первом случае всем студентам предлагаются одинаковые исходные данные. После выполнения заданий устраивается обсуждение результатов в форме тематической дискуссии. Во втором случае каждому из студентов предлагается индивидуальное задание. В этом случае после выполнения задания организуется работа в группах: каждый из студентов анализирует и критически оценивает или же интерпретирует результаты одного или нескольких членов группы.

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

Комплект лицензионного программного обеспечения:

1. Пакет офисных программ
2. Антивирус Kaspersky

Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
4. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

Сертифицированные программные и аппаратные средства защиты информации - не используются

Дистрибутив языка Python 3.6 (или более поздней версии) Anaconda  
Графовая СУБД Neo4j.

Пакет для работы с графами Gephi.

Для манипулирования с файлами файловый менеджер Far.

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

Для проведения лекций и практических занятий необходима аудитория, оснащенная проектором и компьютерами с постоянным подключением к сети Интернет.