

## Программа курса «Анализ данных и искусственный интеллект»

Номер	Название темы	Количество часов	Описание темы
1	Данные и системы их хранения	10	Технологии обработки данных. Большие данные: инструменты и технологии. Применение методов машинного обучения и искусственного интеллекта. Методы обработки данных в промышленной аналитике. Искусственный интеллект в задачах бизнес-аналитики. Базы данных. Визуальное моделирование баз данных. Графическая и физическая модель базы данных. Реляционные базы данных. SQL, синтаксис обращения к базам данных. PostgreSQL. Практическая работа (7 часов). Разбор практических кейсов по применению методов обработки данных, методов машинного обучения и искусственного интеллекта в промышленной аналитике и задачах бизнес-аналитики: разбор задач при обработке «сырых» данных, примеры источников данных. Инструменты для моделирования. Draw.io, DB design, PGAdmin.
2	Разработка компьютерных программ	21,5	Управляющие конструкции в Python. Ветвления, циклы, исключения. Типы данных. Коллекции в Python. Списки, кортежи, словари. Функции. Синтаксис, аргументы, именованные аргументы, область видимости переменных, возвращаемое значение. Классы. Создание объектов своих классов, инкапсуляция, наследование классов. Создание простого окна, взаимодействие с пользователем, Qt Designer. Основные виджеты окон, их свойства и методы. Сигналы и события, диалоговое окна. Различные способы компоновки виджетов. Различные способы компоновки виджетов. Сборка оконного десктопного приложения. Связь приложения с внешним миром. Запись и чтение из файла, работа с настройками приложения, выход в Интернет. Методы управления проектированием сложных систем искусственного интеллекта. Управление проектированием информационных систем. Объектная парадигма. Основы языка UML. Основные принципы применения CASE-средств. Функциональные требования к ПО - UML Use-Case. Структура и поведение программ. Диаграммы классов, последовательностей и активностей UML. Проектирование реализации - диаграммы компонентов, диаграммы развертывания в UML. Понятие жизненного цикла программного обеспечения (ПО). Стадии и модели жизненного цикла ПО. Стандарты жизненного цикла ПО. Методы описания объекта автоматизации. Стандарты IDEF0, BPMN. Управление требованиями к ПО, модели описания требований, системы управления требованиями. Проектирование архитектуры ПО. Фреймворки проектирования архитектуры. Язык ArchiMate. Обеспечение качества ПО. Системы автоматизированного тестирования. Практическая работа: Разбор практических кейсов по применению Python при разработке приложений
3	Модельно-алгоритмический аппарат интеллектуальных систем	23,5	Корреляционно-регрессионный анализ. Корреляционное поле. Коэффициент линейной корреляции. t-критерий Стьюдента. Регрессия. Метод наименьших квадратов. Множественная и многофакторная регрессия. Метод Брандона. Полный факторный эксперимент. Расчет коэффициентов уравнения регрессии. Дробный факторный эксперимент. Нейросетевые технологии анализа данных. Искусственные нейронные сети и глубокое обучение. Нейро-нечеткие системы. Проблема снижения размерности данных. Метод главных компонент (Principal Component Analysis). Алгоритмы кластеризации. Метод k-means, c-means, hierarchical clustering. Выбор алгоритма кластеризации. Задача классификации. Матрица ошибок. Точность классификации. Модель анализа ROC-AUC. Методы классификации. Метод ближайших соседей k-NN. Метод опорных векторов. Понятие временного ряда. Предварительная обработка временных рядов. Методы семейства ARMA. Применение LSTM нейронных сетей к анализу временных рядов. Теоретические основы имитационного моделирования сложных систем. Методика построения

			<p>имитационных моделей сложных систем. Моделирование в PowerSim  Моделирование в среде Anylogic.  Практическая работа: Разбор практических кейсов по применению методов корреляционно-регрессионного анализа для решения учебных и практических задач</p>
4	Практическое применение интеллектуальных методов	20	<p>Обработка естественного языка. Базовые понятия. Понятия токенизации, стемминга, лемматизации. Модель Bag of Words, ее применение для классификации текста. Применение машинного обучения для сентимент-анализа предложений. Тематическое моделирование. Задачи компьютерного зрения. Сегментация изображений, детекция объектов, классификация изображений, отслеживание движущихся объектов в видеопотоке, распознавание лиц, оптическое распознавание символов (OCR). Генерация изображений. Базовые методы CV. Точечные операторы. Линейная фильтрация. Операторы соседства. Преобразования Фурье. Пирамиды и вейвлеты. Геометрические преобразования. Глобальная оптимизация. Установка библиотеки OpenCV. Отслеживание объектов. Вычисление разности между кадрами, использование цветowych пространств, метод вычитания фоновых изображений, алгоритм CAMShift, использование оптических потоков. Обнаружение и отслеживание лиц с использованием каскадов Хаара. Использование интегральных изображений для отслеживания глаз и направления взгляда. Классификация изображений с помощью однослойных нейронных сетей. Сверточные нейросети (CNN), классификация изображений с их помощью. Однослойная и многослойная нейронная сеть. Визуализация символов с использованием базы данных оптического распознавания символов, системы оптического распознавания символов. Теоретические основы принятия управленческих решений. Методы принятия управленческих решений. Системы поддержки принятия управленческих решений. Автоматизированные системы поддержки принятия управленческих решений.  Практическая работа: Решение кейсов по теме</p>
5	Итоговая аттестация	2	Зачёт

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение высшего образования  
«Московский государственный технический университет имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

Центр НТИ «Цифровое материаловедение: новые материалы и вещества»

УТВЕРЖДАЮ  
Первый проректор –  
проректор по учебной работе  
МГТУ им. Н.Э. Баумана  
Б.В. Падалкин  
2021 г.



Дополнительное профессиональное образование

ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА  
ПРОГРАММА ПОВЫШЕНИЯ КВАЛИФИКАЦИИ  
«Анализ данных и искусственный интеллект»

Регистрац. № 05.2224.12.10

Москва, 2021

**АВТОР ПРОГРАММЫ:**

Доцент  
Центра НТИ «Цифровое материаловедение:  
новые материалы и вещества»

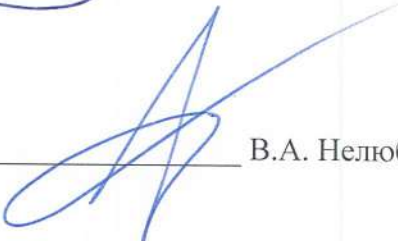
  
В.С. Тынченко

**СОГЛАСОВАНО:**

Начальник УСП  
МГТУ им. Н.Э. Баумана

  
Т.А. Гузева

Директор  
Центра НТИ «Цифровое материаловедение:  
новые материалы и вещества»

  
В.А. Нелюб



7

## Оглавление

<b>1. ОБЩАЯ ХАРАКТЕРИСТИКА ДПП</b>	<b>4</b>
1.1. Цель ДПП	4
1.2. Планируемые результаты обучения	4
1.3. Дополнительные характеристики ДПП	4
1.4. Перечень профессиональных компетенций в рамках имеющейся квалификации, качественное изменение которых осуществляется в результате обучения	5
1.5. Соответствие видов деятельности профессиональным компетенциям и их составляющих	5
<b>2. УЧЕБНЫЙ ПЛАН ДПП</b>	<b>7</b>
2.1. Категория слушателей ДПП	7
2.2. Общая трудоёмкость программы, аудиторная и самостоятельная работа	7
2.3. Форма обучения	7
2.4. Учебный план	7
<b>3. КАЛЕНДАРНЫЙ УЧЕБНЫЙ ГРАФИК</b>	<b>8</b>
<b>4. РАБОЧАЯ ПРОГРАММА ДПП</b>	<b>9</b>
4.1 Рабочая программа модуля «Анализ данных и искусственный интеллект»	9
<b>5. УСЛОВИЯ РЕАЛИЗАЦИИ ДПП</b>	<b>18</b>
5.1. Организационные условия реализации ДПП	18
5.2. Педагогические условия реализации ДПП	18
5.3. Учебно-методическое обеспечение ДПП	18
5.4. Методические рекомендации	19
<b>6. ФОРМЫ ИТОГОВОЙ АТТЕСТАЦИИ ДПП</b>	<b>21</b>
<b>7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ИТОГОВОЙ АТТЕСТАЦИИ</b>	<b>22</b>
7.1. Паспорт комплекта оценочных средств	22
7.2. Комплект оценочных средств	22

## **1. ОБЩАЯ ХАРАКТЕРИСТИКА ДПП**

Программа подготовлена на основе:

- Федерального закона от 29 декабря 2012 года № 273-ФЗ «Об образовании в Российской Федерации»;

- требований Приказа Минобрнауки России от 01.07.2013 года № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»;

- методических рекомендаций-разъяснений Минобрнауки России по разработке дополнительных профессиональных программ на основе профессиональных стандартов от 22 апреля 2015 года № ВК-1030/06;

- модель компетенций в сфере искусственного интеллекта Письмо Минобрнауки России «О направлении модели компетенций» от 01.10.2021 №МН-5/20376 .

Реализация программы ДПП направлена на получение новой компетенции, необходимой для профессиональной деятельности.

### **1.1. Цель ДПП**

Получить знания, навыки и умения, необходимые для создания информационных технологий нового поколения, обеспечивающих экономически эффективное извлечение полезной информации из больших объемов разнообразных данных путем высокой скорости их сбора, обработки и анализа, и применение этих технологий в информационно-аналитической деятельности, в системах управления и принятия решений, а также для разработки на их основе новых продуктов и услуг.

### **1.2. Планируемые результаты обучения**

Планируемые результаты обучения по ДПП:

- освоение профессиональных компетенций в процессе изучения перечисленных тем в учебном плане;

- успешное освоение программы повышения квалификации;

- успешное прохождение итоговой аттестации (зачет в форме тестирования).

Обучающимся, успешно прошедшим обучение, выполнившим текущие контрольные задания и выдержавшим предусмотренный учебным планом зачет, выдается удостоверение о повышении квалификации по ДПП «Анализ данных и искусственный интеллект».

### **1.3. Дополнительные характеристики ДПП**

Характеристики новой квалификации определены в приказе Минтруда России от 06.07.2020 №405н «Об утверждении профессионального стандарта «Специалист по большим данным».

Вид профессиональной деятельности:

- Создание и применение технологий больших данных (Код 06.042).

Трудовые функции:

- проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика (А/04.6);

- разработка продуктов на основе встроенной аналитики больших данных (С/01.8);

#### **1.4. Перечень профессиональных компетенций в рамках имеющейся квалификации, качественное изменение которых осуществляется в результате обучения**

Получаемые компетенции базируются на основании модели компетенций в сфере искусственного интеллекта Письмо Минобрнауки России «О направлении модели компетенций» от 01.10.2021 №МН-5/20376, закрепленный в СУОС МГТУ им. Н.Э.Баумана 09.04.04 Программная инженерия:

ОПК-10. Способен разрабатывать оригинальные алгоритмы и программные средства с использованием современных интеллектуальных компьютерные технологий, для решения задач в области создания и применения искусственного интеллекта.

ОПК-12. Способен адаптировать и применять на практике для решения задач в области создания и применения технологий и систем искусственного интеллекта классические и новые научные принципы и методы исследований.

#### **1.5. Соответствие видов деятельности профессиональным компетенциям и их составляющих**

Профессиональные компетенции	Практический опыт	Умения	Знания
Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика (А/04.6)			
ОПК-10. Способен разрабатывать оригинальные алгоритмы и программные средства с использованием современных интеллектуальных компьютерные технологий, для решения задач в области создания и применения искусственного интеллекта	Выбор методов и инструментальных средств анализа больших данных для проведения аналитических работ	Планировать аналитические работы с использованием технологий больших данных	Основы управления аналитическими работами
Разработка продуктов на основе встроенной аналитики больших данных (С/01.8)			

<p>ОПК-12. Способен адаптировать и применять на практике для решения задач в области создания и применения технологий и систем искусственного интеллекта классические и новые научные принципы и методы исследований</p>	<p>Проведение научно-исследовательских и опытно-конструкторских работ по созданию вариантов технического воплощения концепции продукта, создание прототипа продукта на основе встроенной аналитики больших данных</p>	<p>Разрабатывать научно-техническую документацию</p>	<p>Основы инновационной деятельности и управления инновациями в сфере информационных технологий</p>
--	---	--	---



## 2. УЧЕБНЫЙ ПЛАН ДПП

### 2.1. Категория слушателей ДПП

Имеющаяся квалификация (требования к слушателям) – к освоению ДПП допускаются лица, имеющие высшее образование.

### 2.2. Общая трудоёмкость программы, аудиторная и самостоятельная работа

Общая трудоёмкость программы 150 академических часа, из них 75 академических часа аудиторной работы, 73 академических часа самостоятельной работы и 2 академических часа итоговой аттестации.

### 2.3. Форма обучения

Форма обучения по ДПП – дистанционная.

### 2.4. Учебный план

ДПП «Анализ данных и искусственный интеллект» реализуется одним модулем.

№ п/п	Наименование темы, модуля	Форма контроля	Всего, час	В том числе			
				Лекции	Практ. занятия	Самост. работа	Итоговая аттестация
1.	Данные и системы их хранения	Тест, Решение практ. кейса	20	3	7	10	-
2.	Разработка компьютерных программ	Тест, Решение практ. кейса	43	4,5	17	21,5	-
3.	Модельно-алгоритмический аппарат интеллектуальных систем	Тест, Решение практ. кейса	45	7,5	16	21,5	-
4.	Практическое применение интеллектуальных методов	Тест, Решение практ. кейса	40	4	16	20	-
5.	Итоговая аттестация	Тест	2	-	-	-	2
	<b>ИТОГО</b>	-	150	19	56	73	2

### 3. КАЛЕНДАРНЫЙ УЧЕБНЫЙ ГРАФИК

№ п/п	Наименование темы, модуля	1 неделя	2 неделя	3 неделя	4 неделя	5 неделя	6 неделя	7 неделя	8 неделя	9 неделя	10 неделя	11 неделя	12 неделя
1.	Данные и системы их хранения	■	■										
2.	Разработка компьютерных программ			■	■	■							
3.	Модельно-алгоритмический аппарат интеллектуальных систем						■	■	■				
4.	Практическое применение интеллектуальных методов									■	■	■	
5.	Итоговая аттестация												За чет

Минимальный срок освоения ДПП — 12 недель.

## 4. РАБОЧАЯ ПРОГРАММА ДПП

### 4.1. Рабочая программа модуля «Анализ данных и искусственный интеллект»

4.1.1. Цель изучения модуля: получить знания, навыки и умения, необходимые для создания информационных технологий нового поколения, обеспечивающих экономически эффективное извлечение полезной информации из больших объемов разнообразных данных путем высокой скорости их сбора, обработки и анализа, и применение этих технологий в информационно-аналитической деятельности, в системах управления и принятия решений, а также для разработки на их основе новых продуктов и услуг.

4.1.2. Задача изучения модуля: ознакомить обучающихся с технологиями обработки и управления данными, методами обработки данных в промышленной аналитике и задачах бизнес-аналитики; дать умения и практические навыки по построению реляционных баз данных в среде PostgreSQL.

#### 4.1.3. Планируемые результаты обучения

Процесс изучения раздела направлен на формирование следующих компетенций

Код компетенции	Перечень планируемых результатов обучения по модулю	Формы и методы обучения, способствующие формированию и развитию компетенции
ОПК-10 ОПК-12	<b>Знать:</b> - Основы управления аналитическими работами. - Основы инновационной деятельности и управления инновациями в сфере информационных технологий <b>Уметь:</b> - Планировать аналитические работы с использованием технологий больших данных. - Разрабатывать научно-техническую документацию <b>Владеть:</b> - Выбор методов и инструментальных средств анализа больших данных для проведения аналитических работ. - Проведение научно-исследовательских и опытно-конструкторских работ по созданию вариантов технического воплощения концепции продукта, создание прототипа продукта на основе встроеной аналитики больших данных	Формы обучения: Фронтальная. Методы обучения: Лекции; Практические занятия; Самостоятельная работа.

#### 4.1.4 Содержание курса

##### **Тема 1. Данные и системы их хранения (20 часов)**

Лекции (3 часа). Технологии обработки данных. Большие данные: инструменты и технологии. Применение методов машинного обучения и искусственного интеллекта. Методы обработки данных в промышленной аналитике. Искусственный интеллект в задачах бизнес-аналитики. Базы данных. Визуальное моделирование баз данных. Графическая и физическая модель базы данных. Реляционные базы данных. SQL, синтаксис обращения к базам данных. PostgreSQL.

Практическая работа (7 часов). Разбор практических кейсов по применению методов обработки данных, методов машинного обучения и искусственного интеллекта в промышленной аналитике и задачах бизнес-аналитики: разбор задач при обработке «сырых» данных, примеры источников данных. Инструменты для моделирования. Draw.io, DB desgn, PGAdmin. Разбор практического кейса: построение графической модели базы данных, на примере магазина; построение физической модели базы данных на примере магазина.

Самостоятельная работа (10 часов). Проработка материала дополнительной литературы по теме. Решение практического кейса, в процессе которого необходимо сформировать сводную таблицу данных из нескольких источников, а также сформировать описательную статистику полученных данных. Решение практического кейса, в процессе которого необходимо: построить графическую модель с суммарным количеством сущностей не менее 10 и суммарным количеством атрибутов не менее 30 (тема свободная, но отличная от рассмотренной на занятии); построить физическую модель с количеством таблиц не менее 4, суммарным количеством атрибутов не менее 10 и суммарным количеством связей не менее 5 (тема свободная, но отличная от рассмотренной на занятии).

##### **Тема 2. Разработка компьютерных программ (43 часа)**

Лекции (4,5 часа). Управляющие конструкции в Python. Ветвления, циклы, исключения. Типы данных. Коллекции в Python. Списки, кортежи, словари. Функции. Синтаксис, аргументы, именованные аргументы, область видимости переменных, возвращаемое значение. Классы. создание объектов своих классов, инкапсуляция, наследование классов. Создание простого окна, взаимодействие с пользователем, Qt Designer. Основные виджеты окон, их свойства и методы. Сигналы и события, диалоговое окно. Различные способы компоновки виджетов. Различные способы компоновки виджетов. Сборка оконного десктопного приложения. Связь приложения с внешним миром. Запись и чтение из файла, работа с настройками приложения, выход в Интернет.

Методы управления проектированием сложных систем искусственного интеллекта. Управление проектированием информационных систем. Объектная парадигма. Основы языка UML. Основные принципы применения CASE-средств. Функциональные требования к ПО - UML Use-Case. Структура и поведение программ. Диаграммы классов, последовательностей и активностей UML. Проектирование реализации - диаграммы компонентов, диаграммы развертывания в UML. Понятие жизненного цикла программного обеспечения (ПО). Стадии и модели жизненного цикла ПО. Стандарты жизненного цикла ПО. Методы описания объекта автоматизации. Стандарты IDEF0, BPMN. Управление требованиями к ПО, модели описания требований, системы управления требованиями. Проектирование архитектуры ПО. Фреймворки проектирования архитектуры. Язык ArchiMate. Обеспечение качества ПО. Системы автоматизированного тестирования.

Практическая работа (17 часов). Разбор практических кейсов по применению Python при разработке приложений с консольным интерфейсом. Разбор практических кейсов по применению Python при разработке приложений. Решение практического кейса «Компьютерный ассистент с графической оболочкой пользователя». Разбор практических кейсов по применению языка UML для проектирования автоматизированных систем обработки информации и управления. Разбор учебного кейса по проектированию информационной системы бизнес-единицы с использованием CASE-средства Modelio Open Source. Разбор практических кейсов по применению стандартов IDEF0, BPMN для разработки программного обеспечения. языка UML для проектирования автоматизированных систем обработки информации и управления. Разбор учебного кейса разработки комплекта требований к ПО в формате бэклога продукта по методологии SCRUM с использованием Atlassian Jira. Разработка архитектуры приложения на языке ArchiMate с использованием системы Archi. Разработка плана управления жизненным циклом.

Самостоятельная работа (21,5 часа). Проработка материала дополнительной литературы по теме. Решение практического кейса, в процессе которого необходимо реализовать консольное приложение с выводом информации на экран, проверкой пользовательского ввода, обработкой исключений. Решение практического кейса, в процессе которого необходимо: создать оконное приложение с возможностью добавления команд; реализовать запуск добавленных команд. Решение практического кейса проектирования информационной системы бизнес-единицы с использованием CASE-средства Modelio Open Source. Решение практического кейса формирования бэклога продукта, архитектуры решения и плана управления жизненным циклом.

### **Тема 3. Модельно-алгоритмический аппарат интеллектуальных систем (45 часов)**

Лекции (7,5 часа). Корреляционно-регрессионный анализ. Корреляционное поле. Коэффициент линейной корреляции. t-критерий Стьюдента. Регрессия. Метод наименьших квадратов. Множественная и многофакторная регрессия. Метод Брандона. Полный факторный эксперимент. Расчет коэффициентов уравнения регрессии. Дробный факторный эксперимент. Нейросетевые технологии анализа данных. Искусственные нейронные сети и глубокое обучение. Нейро-нечеткие системы. Проблема снижения размерности данных. Метод главных компонент (Principal Component Analysis). Алгоритмы кластеризации. Метод k-means, c-means, hierarchical clustering. Выбор алгоритма кластеризации. Задача классификации. Матрица ошибок. Точность классификации. Модель анализа ROC-AUC. Методы классификации. Метод ближайших соседей k-NN. Метод опорных векторов. Понятие временного ряда. Предварительная обработка временных рядов. Методы семейства ARMA. Применение LSTM нейронных сетей к анализу временных рядов. Теоретические основы имитационного моделирования сложных систем. Методика построения имитационных моделей сложных систем. Моделирование в PowerSim. Моделирование в среде Anylogic.

Практическая работа (16 часов). Разбор практических кейсов по применению методов корреляционно-регрессионного анализа для решения учебных и практических задач, таких как «Регрессионный анализ по результатам дробного факторного эксперимента для процесса выщелачивания восстановленного барийсодержащего сырья». Разбор практических кейсов по применению нейросетевых технологий к анализу данных. Решение учебного кейса аппроксимации гладкой функции многомерной нейронной сетью, реализованной в Python с использованием фреймворка Keras. Решение учебного кейса аппроксимации двумерной функции при помощи нейро-нечеткой модели в Python с использованием библиотеки ANFIS. Разбор практических кейсов по применению методов k-means, c-means, hierarchical clustering для решения задач кластеризации. Разбор учебных кейсов «Снижение размерности в задаче определения кредитного потенциала заемщика», «Кластеризация городов Российской Федерации по типу прироста (убыли) населения». Разбор практических кейсов по применению методов ближайших соседей и опорных векторов для решения задач классификации. Решение практического кейса: «Оценка и классификация кредитного потенциала заемщиков». Разбор учебных и практических кейсов по применению методов предварительной обработки временных рядов, методов прогнозирования из семейства ARMA, а также LSTM нейронных сетей к анализу временных рядов. Решение практического кейса: «Прогнозирование объемов

авиаперевозок». Разбор практических кейсов по применению методов имитационного моделирования в системах PowerSim и Anylogic. Решение практического кейса: «Модель Ф. Басса».

Самостоятельная работа (11,5 часа). Проработка материала дополнительной литературы по теме. Решение практического кейса, в процессе которого необходимо провести регрессионный анализ по результатам дробного факторного эксперимента для процесса при реакции, протекающей в спиртовом растворе в присутствии катализатора. Решение практического кейса, в процессе которого необходимо провести аппроксимацию функции при помощи нейронных и нейро-нечетких систем. Решение практического кейса «Кластеризация контрольных карт технологического процесса». Решение практического кейса «Классификация изображений». Решение практического кейса «Прогнозирование объемов продаж на торговом предприятии». Решение практического кейса: «Выплата Кредита».

#### **Тема 4. Практическое применение интеллектуальных методов (40 часов)**

Лекции (4 часа). Обработка естественного языка. Базовые понятия. Понятия токенизации, стемминга, лемматизации. Модель Bag of Words, ее применение для классификации текста. Применение машинного обучения для сентимент-анализа предложений. Тематическое моделирование. Задачи компьютерного зрения. Сегментация изображений, детекция объектов, классификация изображений, отслеживание движущихся объектов в видеопотоке, распознавание лиц, оптическое распознавание символов (OCR). Генерация изображений. Базовые методы CV. Точечные операторы. Линейная фильтрация. Операторы соседства. Преобразования Фурье. Пирамиды и вейвлеты. Геометрические преобразования. Глобальная оптимизация. Установка библиотеки OpenCV. Отслеживание объектов. Вычисление разности между кадрами, использование цветовых пространств, метод вычитания фоновых изображений, алгоритм CAMShift, использование оптических потоков. Обнаружение и отслеживание лиц с использованием каскадов Хаара. Использование интегральных изображений для отслеживания глаз и направления взгляда. Классификация изображений с помощью однослойных нейронных сетей. Сверточные нейросети (CNN), классификация изображений с их помощью. Однослойная и многослойная нейронная сеть. Визуализация символов с использованием базы данных оптического распознавания символов, системы оптического распознавания символов. Теоретические основы принятия управленческих решений. Методы принятия управленческих решений. Системы поддержки принятия

управленческих решений. Автоматизированные системы поддержки принятия управленческих решений.

Практическая работа (16 часов). Последовательное решение при помощи Python на учебном примере текста задач токенизации, стемминга, лемматизации, чанкинга, частотного анализа при помощи Bag of Words. Создание прогнозатора категорий и анализатора грамматических родов. Создание сентимент-анализатора на базе наивного байесовского классификатора. Установка OpenCV в Python. Отработка загрузки, отображения и сохранения изображений, доступ и манипулирование пикселями, манипулирование изображениями (сдвиг, масштабирование, повороты), арифметика изображений, разделение и слияние каналов, различные методы размытия, базовые методы распознавания. Расчет разницы между кадрами в видеопотоке. Отслеживание объектов в видеопотоке с веб-камеры устройства с помощью цветowych пространств, с помощью вычитания фоновых изображений. Создание интерактивного трекера объектов в видеопотоке с веб-камеры устройства с помощью алгоритма CAMShift. Отслеживание объектов в видеопотоке с использованием оптических потоков. Реализация детекции лиц и оценка направления взгляда в видеопотоке с веб-камеры устройства с использованием каскадов Хаара. Разбор практических кейсов по применению методов принятия управленческих решений. Разбор учебных кейсов по применению систем поддержки принятия решений в маркетинге.

Самостоятельная работа (20 часов). Проработка материала дополнительной литературы по теме. Решение практического кейса, в процессе которого необходимо провести анализ собственного текста обучающегося, в частности решить задачи: токенизации, стемминга, лемматизации, чанкинга, частотного анализа при помощи Bag of Words; создания прогнозатора категорий и анализатора грамматических родов; создания сентимент-анализатора на базе наивного байесовского классификатора. Решение практического кейса, в процессе которого необходимо на собственные устройства обучающихся решить задачи: отслеживания объектов в видеопотоке с веб-камеры устройства; создания интерактивного трекера объектов в видеопотоке с веб-камеры устройства с помощью алгоритма CAMShift; детекции лиц и оценка направления взгляда в видеопотоке с веб-камеры устройства с использованием каскадов Хаара. Решение практического кейса по применению систем поддержки принятия управленческих решений в кредитно-финансовых операциях.

4.1.5. Оценочное средство для текущего контроля (примерные вопросы для теста):

*Тема 1.*



Вопрос: Какие данные могут считаться Big Data?

- а) Каталог товаров гипермаркета
- б) Список всех студентов университета
- в) Данные ежедневных метеонаблюдений за год
- г) Логи пользовательской активности портала mail.ru

Вопрос: Система управления базами данных представляет собой программный продукт, входящий в состав:

- а) прикладного программного обеспечения.
- б) операционной системы;
- в) уникального программного обеспечения;
- г) системного программного обеспечения;
- д) систем программирования;

## *Тема 2.*

Вопрос: Неизменяемый тип данных в Python, позволяющий хранить значения в виде последовательности:

- а) Словарь (dict)
- б) Список (list)
- в) Кортеж (tuple)
- г) Цикл (for)

Вопрос: Как взаимодействуют друг с другом сигналы и слоты?

- а) Это никак не связанные элементы интерфейса
- б) Слот вызывается, когда срабатывает его сигнал
- в) Сигнал вызывается, когда срабатывает его слот
- г) Слот вызывает метод сигнала (события)

Вопрос: В случае, когда при построении диаграммы вариантов использования (диаграммы прецедентов, use-case diagram) необходимо выделить общую часть функционала нескольких вариантов использования в отдельный прецедент, какой спецификатор связи необходимо указывать между базовыми прецедентами и вновь создаваемым?

- а) conclude
- б) extend
- в) suspect
- г) include
- д) intend

Тема 3.

Вопрос: Для проверки значимости коэффициента парной корреляции применяется:

- а) Математическое ожидание
- б) Коэффициент регрессии
- в) t-критерий Стьюдента
- г) Критерий Фишера

Вопрос: По характеру распространения сигнала нейронные сети подразделяют на:

- а) монотонные
- б) сети с прямым распространением
- в) сети с обратными связями
- г) бинарные
- д) асинхронные

Вопрос: Расстояние, широко используемое для характеристики объектов с бинарными признаками:

- а) Расстояние Хэмминга
- б) Расстояние Махаланобиса
- в) Расстояние Евклида
- г) Расстояние Гаусса

Вопрос: Предсказательная способность гипотезы классификации обычно оценивается по:

- а) Вероятности ошибочной классификации
- б) Вероятности верной классификации
- в) Вероятности обеспечения требуемого уровня классификации
- г) Вероятности ошибки 2 рода

Вопрос: Модель ARMA учитывает:

- а) Авторегрессию
- б) Скользящее среднее
- в) Разность необработанных наблюдений
- г) Сезонные тренды

Вопрос: Система Anylogic реализует следующие три современных подхода:

- а) Продолжительное моделирование
- б) Дискретно-событийное моделирование
- в) Агентное моделирование
- г) Контр-агентное моделирование
- д) Системная динамика

Тема 4.

Вопрос: Выберите утверждения, соответствующие подходу токенизации:

- а) не изменяет длину данных
- б) нематематический подход
- в) изменяет тип данных
- г) математический подход
- д) не изменяет тип данных
- е) изменяет длину данных

Вопрос: Метрика в задачах компьютерного зрения, вычисляемая как площадь пересечения двух областей, деленная на общую площадь регионов.

- а) mAP
- б) Accuracy
- в) F-score
- г) IoU

Вопрос: Какие из следующих элементов содержатся в экспертной системе:

- а) база знаний по конкретной предметной области
- б) блок синтетического вывода
- в) лингвистический процессор
- г) блок логического вывода
- д) центральный процессор

## 5. УСЛОВИЯ РЕАЛИЗАЦИИ ДПП

### 5.1. Организационные условия реализации ДПП

Наименование аудитории	Вид занятия	Наименование оборудования, программного обеспечения
Лекционная аудитория	Лекции	Компьютер, мультимедийный проектор, экран, доска, пишущий инструмент, Microsoft Office, доступ к сети интернет
Компьютерный класс	Практические занятия	Компьютер, мультимедийный проектор, экран, доска, пишущий инструмент, Microsoft Office, доступ к сети интернет
Компьютерный класс	Самостоятельная работа	Компьютер, мультимедийный проектор, экран, доска, пишущий инструмент, Microsoft Office, доступ к сети интернет
Лекционная аудитория	Итоговая аттестация	Компьютер, мультимедийный проектор, экран, доска, пишущий инструмент, Microsoft Office, доступ к сети интернет

### 5.2. Педагогические условия реализации ДПП

Реализация программы обеспечивается преподавательским составом из числа штатных преподавателей, совместителей или привлеченных на условиях почасовой оплаты труда, удовлетворяющим следующим условиям:

- наличие высшего профессионального образования, соответствующего профилю программы;
- значительный опыт практической деятельности в соответствующей сфере.

### 5.3. Учебно-методическое обеспечение ДПП

Основная литература:

1. Воронов, В. И. Data Mining – технология обработки больших объемов данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. – Москва : Московский технический университет связи и информатики, 2018. – 47 с. – ISBN 2227-8397, - Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – URL: <http://www.iprbookshop.ru/81324.html>.
2. Федин, Ф. О. Анализ данных. Часть 2. Инструменты Data Mining : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. – Москва : Московский городской педагогический университет, 2012. – 308 с. – ISBN 2227-8397, - Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – URL: <http://www.iprbookshop.ru/26445.html>.

Дополнительная литература:

1. Чубукова, И. А. Data Mining : учебное пособие / И. А. Чубукова. – 3-е изд. – Москва, САРАТОВ : Интернет-университет информационных технологий

- (ИНТУИТ), Ай Пи Ар Медиа, 2020. – 469 с. – ISBN 978-5-4497-0289-0. – Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – URL: <http://www.iprbookshop.ru/89404.html>.
2. Воронова, Л. И. Big Data. Методы и средства анализа : учебное пособие / Л. И. Воронова, В. И. Воронов. – Москва : Московский технический университет связи и информатики, 2016. – 33 с. – ISBN 2227-8397.
  3. Сараев, П. В. Методы машинного обучения : методические указания и задания к лабораторным работам по курсу / П. В. Сараев. – Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2017. – 48 с. – ISBN 2227-8397. - Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – URL: <http://www.iprbookshop.ru/83183.html>.
  4. Воронова, Л. И. Machine Learning : регрессионные методы интеллектуального анализа данных : учебное пособие / Л. И. Воронова, В. И. Воронов. – Москва : Московский технический университет связи и информатики, 2018. – 82 с. – ISBN 2227-8397. - Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. – URL: <http://www.iprbookshop.ru/81325.html>.

#### **5.4. Методические рекомендации**

ДПП построена по тематическому принципу, каждый раздел представляет собой логически заверченный материал.

Преподавание программы основано на лично-ориентированной технологии образования, сочетающей два равноправных аспекта этого процесса: обучение и учение. Лично-ориентированный подход развивается при участии слушателей в активной работе на практических занятиях. Лично-ориентированный подход направлен, в первую очередь, на развитие индивидуальных способностей обучающихся, создание условий для развития творческой активности слушателя и разработке инновационных идей, а также на развитие самостоятельности мышления при решении учебных задач разными способами, нахождение рационального варианта решения, сравнения и оценки нескольких вариантов их решения и т.п. Это способствует формированию приемов умственной деятельности по восприятию новой информации, ее запоминанию и осознанию, созданию образов для сложных понятий и процессов, приобретению навыков поиска решений в условиях неопределенности.

Практические занятия проводятся для приобретения навыков решения практических задач в предметной области модуля. Задания, выполняемые на

практических занятиях, выполняются с использованием активных и интерактивных методов обучения.

Самостоятельная работа слушателей предназначена для проработки дополнительной литературы. Результаты практических заданий слушателей учитываются на итоговой аттестации.

При изучении курса предусмотрены следующие методы организации и осуществления учебно-познавательной деятельности:

- объяснительно-иллюстративный метод;
- репродуктивный метод;
- частично-поисковый метод.

## 6. ФОРМЫ ИТОГОВОЙ АТТЕСТАЦИИ ДПП

Итоговая аттестация проводится в форме зачета. Зачет проводится в формате тестирования. Результатом зачета служат правильные ответы на вопросы теста, состоящего из 15 (пятнадцати) вопросов.

По результатам итоговой аттестации слушателю выставляется оценка «ЗАЧТЕНО/НЕ ЗАЧТЕНО»:

Оценка «ЗАЧТЕНО» выставляется слушателю, который:

- правильно ответил не менее чем на 60% вопросов в тесте;
- продемонстрировал необходимые систематизированные знания и достаточную степень владения принципами предметной области программы, понимание их особенностей и взаимосвязь между ними в течение всего срока обучения по ДПП.

Оценка «НЕ ЗАЧТЕНО» ставится слушателю, который:

- ответил правильно менее чем на 60% полученных вопросов в билете;
- имеет крайне слабые теоретические и практические знания, обнаруживает неспособность к построению самостоятельных заключений.

## 7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ИТОГОВОЙ АТТЕСТАЦИИ

### 7.1. Паспорт комплекта оценочных средств

Предметы оценивания	Объекты оценивания	Показатели оценки
ОПК-10. Способен разрабатывать оригинальные алгоритмы и программные средства с использованием современных интеллектуальных компьютерных технологий, для решения задач в области создания и применения искусственного интеллекта	Ответы на вопросы теста	Количество правильных ответов
ОПК-12. Способен адаптировать и применять на практике для решения задач в области создания и применения технологий и систем искусственного интеллекта классические и новые научные принципы и методы исследований	Ответы на вопросы теста	Количество правильных ответов

### 7.2. Комплект оценочных средств

#### 7.2.1. Темы для подготовки к зачету:

1. Данные и системы их хранения
2. Разработка компьютерных программ
3. Модельно-алгоритмический аппарат интеллектуальных систем
4. Практическое применение интеллектуальных методов

#### 7.2.2. Примерные задания для проведения зачета:

1. Какие данные могут считаться Big Data?
  - а) Каталог товаров гипермаркета
  - б) Список всех студентов университета
  - в) Данные ежедневных метеонаблюдений за год
  - г) Логи пользовательской активности портала mail.ru
2. Система управления базами данных представляет собой программный продукт, входящий в состав:
  - а) прикладного программного обеспечения.



- б) операционной системы;
  - в) уникального программного обеспечения;
  - г) системного программного обеспечения;
  - д) систем программирования;
3. Неизменяемый тип данных в Python, позволяющий хранить значения в виде последовательности:
- а) Словарь (dict)
  - б) Список (list)
  - в) Кортеж (tuple)
  - г) Цикл (for)
4. Как взаимодействуют друг с другом сигналы и слоты?
- а) Это никак не связанные элементы интерфейса
  - б) Слот вызывается, когда срабатывает его сигнал
  - в) Сигнал вызывается, когда срабатывает его слот
  - г) Слот вызывает метод сигнала (события)
5. В случае, когда при построении диаграммы вариантов использования (диаграммы прецедентов, use-case diagram) необходимо выделить общую часть функционала нескольких вариантов использования в отдельный прецедент, какой спецификатор связи необходимо указывать между базовыми прецедентами и вновь создаваемым?
- а) conclude
  - б) extend
  - в) suspect
  - г) include
  - д) intend
6. Основными изобразительными элементами процессных диаграмм IDEF0 в SADT модели являются:
- а) прямоугольники (процессы)
  - б) стрелки (линии потока)
  - в) ромбы (ветвления процесса)
  - г) параллелограммы (ввод-вывод информации процесса)

7. Для проверки значимости коэффициента парной корреляции применяется:
- а) Математическое ожидание
  - б) Коэффициент регрессии
  - в) t-критерий Стьюдента
  - г) Критерий Фишера
8. По характеру распространения сигнала нейронные сети подразделяют на:
- а) монотонные
  - б) сети с прямым распространением
  - в) сети с обратными связями
  - г) бинарные
  - д) асинхронные
9. Расстояние, широко используемое для характеристики объектов с бинарными признаками:
- а) Расстояние Хэмминга
  - б) Расстояние Махаланобиса
  - в) Расстояние Евклида
  - г) Расстояние Гаусса
10. Предсказательная способность гипотезы классификации обычно оценивается по:
- а) Вероятности ошибочной классификации
  - б) Вероятности верной классификации
  - в) Вероятности обеспечения требуемого уровня классификации
  - г) Вероятности ошибки 2 рода
11. Модель ARMA учитывает:
- а) Авторегрессию
  - б) Скользящее среднее
  - в) Разность необработанных наблюдений
  - г) Сезонные тренды
12. Система Anylogic реализует следующие три современных подхода:
- а) Продолжительное моделирование
  - б) Дискретно-событийное моделирование

- в) Агентное моделирование
- г) Контр-агентное моделирование
- д) Системная динамика

13. Выберите утверждения, соответствующие подходу токенизации:

- а) не изменяет длину данных
- б) нематематический подход
- в) изменяет тип данных
- г) математический подход
- д) не изменяет тип данных
- е) изменяет длину данных

14. Метрика в задачах компьютерного зрения, вычисляемая как площадь пересечения двух областей, деленная на общую площадь регионов.

- а) mAP
- б) Accuracy
- в) F-score
- г) IoU

15. Какие из следующих элементов содержатся в экспертной системе:

- а) база знаний по конкретной предметной области
- б) блок синтетического вывода
- в) лингвистический процессор
- г) блок логического вывода
- д) центральный процессор