

МИНОБРНАУКИ РОССИИ  
Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Астраханский государственный университет имени В. Н. Татищева»  
(Астраханский государственный университет им. В. Н. Татищева)

СОГЛАСОВАНО  
Руководитель ОПОП

Ю.А. Головки

«5» мая 2025 г.

УТВЕРЖДАЮ  
И.о. заведующего кафедрой  
информационных технологий  
О.Н. Выборнова

«5» мая 2025 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

**«Машинное обучение»**

Составитель(и)	<b>Выборнова О.Н., доцент, к.т.н., доцент кафедры ИТ; Соболевский В.В., ст. преподаватель кафедры ИТ;</b>
Направление подготовки / специальность	<b>09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ</b>
Направленность (профиль) / специализация ОПОП	<b>Технологии разработки и администрирования информационных систем</b>
Квалификация (степень)	<b>бакалавр</b>
Форма обучения	<b>Очно-заочная</b>
Год приёма	<b>2022</b>
Курс	<b>4</b>
Семестр(ы)	<b>7</b>

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

**1.1. Целями освоения дисциплины (модуля) «Машинное обучение»** являются формирование у студентов практических навыков применения инструментальных средств при решении профессиональных задач в области машинного обучения

### 1.2. Задачи освоения дисциплины (модуля):

- сформировать теоретические знания по основам машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования;
- выработать умения по практическому применению методов машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования при решении прикладных задач в различных прикладных областях;
- выработать умения и навыки использования различных программных инструментов анализа баз данных и систем машинного обучения.

## 2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОПОП

**2.1. Учебная дисциплина (модуль) «Машинное обучение»** относится к части, формируемой участниками образовательных отношений, и осваивается в 7 семестре.

**2.2. Для изучения данной учебной дисциплины (модуля) необходимы следующие знания, умения, навыки, формируемые предшествующими учебными дисциплинами (модулями):**

- математические основы информационных технологий и вычислительной техники
- математические основы искусственного интеллекта
- системы искусственного интеллекта
- основы программирования

Знания: линейной алгебры, дифференциального исчисления и теории вероятностей, язык программирования Python, численные методы, принципы векторизации

Умения: применять математический аппарат, применять знания для подготовки обучающего набора данных, составлять программы на языке программирования Python

Навыки: использования технологии векторизации для обработки векторных и матричных данных, математического мышления, определения алгоритмов и структур данных при изучении существующих моделей машинного обучения

**2.3. Последующие учебные дисциплины (модули) и (или) практики, для которых необходимы знания, умения, навыки, формируемые данной учебной дисциплиной (модулем):**

- интеллектуальные системы и технологии
- глубокое обучение
- компьютерное зрение

## 3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Процесс освоения дисциплины (модуля) направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО и ОПОП ВО по данному направлению подготовки / специальности:

- а) универсальной (УК): УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач
- б) профессиональной (ПК): ПК-3 Способен обеспечения эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных,

являющихся частью различных информационных систем

**Таблица 1. Декомпозиция результатов обучения**

Код и наименование компетенции	Планируемые результаты обучения по дисциплине (модулю)					
	Знать (1)		Уметь (2)		Владеть (3)	
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Знать: принципы сбора, отбора и обобщения информации.	УК-1.2. Уметь: соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности.	УК-1.3. Владеть: практическим опытом работы с информационными источниками, опыт научного поиска, создания научных текстов.			
ПК-3 Способен обеспечения эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем	ПК-3.1. Знать технологии и методы по обеспечению эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем	ПК-3.2. Уметь выполнять комплекс работ по обеспечению эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем	ПК-3.3. Владеть инструментальными средствами обеспечения эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем			

#### 4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Объем дисциплины (модуля) составляет 4 зачетные единицы, в том числе 54 час, выделенных на контактную работу обучающихся с преподавателем (из них 18 часов – лекции, 36 часов – лабораторные работы), и 90 часов – на самостоятельную работу обучающихся:

**Таблица 2.2. Структура и содержание дисциплины (модуля)**

Раздел, тема дисциплины (модуля)	Семестр	Контактная работа (в часах)			Самост. работа		Формы текущего контроля успеваемости, форма промежуточной аттестации
		Л	ПЗ	ЛР	КР	СР	
Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.	7	2		4		11	Тест
Тема 2. Линейная регрессия. Логистическая регрессия		2		4		11	Тест
Тема 3. Введение в машинное обучение		4		8		22	Тест, отчет по лабораторной работе
Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.		2		6		10	Тест, отчет по лабораторной работе
Тема 5. Обработка естественного языка.		2		4		12	Тест, отчет по лабораторной работе
Тема 6. Модели последовательностей. Рекуррентные		4		6		12	Тест, лабораторная работа

модели.						
Тема 7. Практические аспекты обучения глубоких нейронных сетей.		2		4		12
		<b>18</b>		<b>36</b>		<b>90</b>
						Тест, отчет по лабораторной работе
						Диф. зачёт (зачёт с оценкой)

*Примечание:* Л – лекция; ПЗ – практическое занятие, семинар; ЛР – лабораторная работа; КР – курсовая работа; СР – самостоятельная работа.

**Таблица 3. Матрица соотнесения разделов, тем учебной дисциплины (модуля) и формируемых компетенций**

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции		Общее количество компетенций
		УК-1	ПК-3	
<i>Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.</i>	<b>17</b>	+		<b>1</b>
<i>Тема 2. Линейная регрессия. Логистическая регрессия</i>	<b>17</b>	+		1
<i>Тема 3. Введение в машинное обучение</i>	<b>34</b>	+	+	2
<i>Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.</i>	<b>18</b>	+	+	2
<i>Тема 5. Обработка естественного языка.</i>	<b>18</b>	+	+	2
<i>Тема 6. Модели последовательностей. Рекуррентные модели.</i>	<b>22</b>	+	+	2
<i>Тема 7. Практические аспекты обучения глубоких нейронных сетей.</i>	<b>18</b>	+	+	2
<b>Итого</b>	<b>144</b>			

### **Краткое содержание каждой темы дисциплины (модуля)**

#### **Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.**

Типы матриц. Операции над матрицами. Свойства матричных операций. Вектор. Действия над векторами. Разложение вектора на базис.

#### **Тема 2. Линейная регрессия. Логистическая регрессия**

Линейная регрессия. Парная и множественная регрессия. Матричное представление. Методы оценки. Логистическая регрессия. Прогнозирование вероятности возникновения события. Подбор параметров. Регуляризация.

#### **Тема 3. Введение в машинное обучение.**

Градиентный спуск. Обучение с учителем и без учителя. Задачи классификации и кластеризации. Бинарная и множественная классификации. Виды функций активации нейрона. Глубокие нейронные сети и метод обратного распространения ошибки.

#### **Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.**

Компьютерное зрение. Сверточные сети. Аугментация. Перенос результатов обучения нейронной сети. Проблема переобучения.

#### **Тема 5. Обработка естественного языка.**

Обработка естественного языка. Токенизация и кодирование слов методом встраивания.

#### **Тема 6. Модели последовательностей. Рекуррентные модели.**

Последовательные и рекуррентные сети. Ячейки длинной краткосрочной памяти.

## **Тема 7. Практические аспекты обучения глубоких нейронных сетей.**

Методы определения переобучения. Регуляризация. Проблема взрыва/затухания градиентов. Способы оптимизации обучения нейронных сетей.

### **5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)**

#### **5.1. Указания для преподавателей по организации и проведению учебных занятий по дисциплине (модулю)**

Учебная деятельность студента в процессе изучения строится из контактных форм работы с преподавателем (аудиторные занятия, зачет с оценкой) и самостоятельной работы.

Для успешного освоения дисциплины является обязательным посещение всех занятий, выполнение домашнего задания и иных форм самостоятельной работы, которые назначаются преподавателем.

Методическая поддержка дисциплины обеспечивается использованием дистанционных технологий. Студентам предлагается информационный ресурс «Электронное образование». Доступ студентов к учебным ресурсам осуществляется по учетной записи и паролю после регистрации на курс «Машинное обучение» на период обучения по данной дисциплине.

На сервере размещен методический материал по данной дисциплине, в содержание которого входит:

- теоретический материал;
- задания и указания по выполнению лабораторно-практических работ, требования к содержанию и их оформлению, рекомендации по их защите;
- тестовые вопросы, предназначенные всех видов контроля, включая самоконтроль освоения учебного материала.

Аудиторные занятия проводятся на основе теоретического материала, опубликованного на образовательном портале, это позволяет студентам изучить пропущенный материал или самостоятельно разобраться с темой, не освоенной на занятии.

Для исключения отрыва студентов от учебного процесса проводится учет посещаемости аудиторных занятий.

#### **5.2. Указания для обучающихся по освоению дисциплины (модулю)**

В рамках дисциплины «Машинное обучение» предполагается организация следующих видов

- самостоятельной работы студентов (таблица 4):
- работа с лекционным материалом, учебно-методическим информационным обеспечением;
- подготовка к лабораторно-практическим работам, подготовка отчетов к защите отчетов;
- подготовка к компьютерному тестированию.

В качестве форм и методов контроля внеаудиторной самостоятельной работы используются: электронные отчеты по выполнению лабораторных работ; устный опрос, протоколы компьютерного тестирования.

Задания к лабораторно-практическим занятиям размещены на образовательном портале «Электронное образование». Рекомендуется заранее ознакомиться с темой, основными вопросами, рекомендациями, требованиями к представлению отчета и критериями оценивания заданий.

В процессе подготовки к лабораторно-практическим занятиям, необходимо обратить особое внимание на самостоятельное изучение рекомендованной литературы. Самостоятельная работа с учебниками, учебными пособиями, научной, справочной литературой, материалами периодических изданий и Интернета является наиболее эффективным методом получения

дополнительных знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала.

**Таблица 4. Содержание самостоятельной работы обучающихся**

Темы/вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
<i>Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.</i> Разложение вектора на базис	11	Изучение теоретического материала
<i>Тема 2. Линейная регрессия. Логистическая регрессия</i> Нормальное уравнение прямой Прогнозирование и подбор параметров	11	Изучение теоретического материала
<i>Тема 3. Введение в машинное обучение</i> Задачи прогнозирования, распознавания	22	Изучение теоретического материала, написание программного кода
<i>Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.</i> Создание собственного классификатора изображений	10	Изучение теоретического материала, написание программного кода
<i>Тема 5. Обработка естественного языка.</i> Самостоятельная работа по организации текста BBC News	12	Изучение теоретического материала, написание программного кода
<i>Тема 6. Модели последовательностей. Рекуррентные модели.</i> Генератор текста с помощью RNN на основе символов	12	Изучение теоретического материала, написание программного кода
<i>Тема 7. Практические аспекты обучения глубоких нейронных сетей.</i> Проблема взрывающихся и исчезающих градиентов	12	Изучение теоретического материала, написание программного кода

### **5.3. Виды и формы письменных работ, предусмотренных при освоении дисциплины (модуля), выполняемые обучающимися самостоятельно**

#### **Отчет по лабораторным работам.**

Результатом работы, выполняемой студентами, является электронный отчет по выполнению лабораторно-практической работы, тематика которых представлена в таблице 4,

Электронный отчет представляет собой файл формата Jupiter Notebook, содержащий программный код, результаты выполнения программы и текстовые пояснения. Код должен быть исполняемым и его запуск должен приводить к тем же результатам, что и в предоставленном отчете. Файл передается на проверку преподавателю путем загрузки на платформу «Электронное образование» в соответствующий заданию раздел.

## **6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

### **6.1. Образовательные технологии**

Учебные занятия по дисциплине могут проводиться с применением информационно-телекоммуникационных сетей при опосредованном (на расстоянии) взаимодействии обучающихся и преподавателя в режимах on-line в формах: видеолекций, лекций-презентаций, видеоконференции, собеседования в режиме чат, форума, чата, выполнения виртуальных практических и/или лабораторных работ и др.

**Таблица 5. Образовательные технологии, используемые при реализации учебных занятий**

Раздел, тема дисциплины (модуля)	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.	Обзорная лекция	Не предусмотрено	Решение задач, тест
Тема 2. Линейная регрессия. Логистическая регрессия	Обзорная лекция	Не предусмотрено	Решение задач, тест
Тема 3. Введение в машинное обучение	Обзорная лекция	Не предусмотрено	Выполнение лабораторной работы, тест
Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.	Лекция-диалог	Не предусмотрено	Выполнение лабораторной работы, тест
Тема 5. Обработка естественного языка.	Лекция-диалог	Не предусмотрено	Выполнение лабораторной работы, тест
Тема 6. Модели последовательностей. Рекуррентные модели.	Лекция-диалог	Не предусмотрено	Выполнение лабораторной работы, тест
Тема 7. Практические аспекты обучения глубоких нейронных сетей.	Лекция-диалог	Не предусмотрено	Выполнение лабораторной работы, тест

## 6.2. Информационные технологии

При реализации учебной и внеучебной работы используются следующие информационные технологии:

- образовательный сайт <http://moodle.asu-edu.ru> (размещение учебно-методического материала, публикация заданий для предоставления студентами выполненных отчетов по всем видам работ, ознакомление учащихся с оценками и т.д., размещение объявлений, on-line консультации, обсуждение вопросов в форуме и т.д.);
- среда разработки моделей машинного обучения <https://colab.research.google.com>
- веб-сервис для хостинга IT-проектов и их совместной разработки <https://github.com>
- онлайн-визуализатор n-мерных векторов <https://projector.tensorflow.org>
- ресурсы ЭБС и сети Internet, как источников информации.

## 6.3. Программное обеспечение, современные профессиональные базы данных и информационные справочные системы

### 6.3.1. Программное обеспечение

Наименование программного обеспечения	Назначение
Adobe Reader	Программа для просмотра электронных документов
Платформа дистанционного обучения LMS Moodle	Виртуальная обучающая среда
Google Chrome	Браузер
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Офисная программа
7-zip	Архиватор

Microsoft Windows 10 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты
MATLAB R2014a	Пакет прикладных программ для решения задач технических вычислений

### 6.3.2. Современные профессиональные базы данных и информационные справочные системы

1. Универсальная справочно-информационная полнотекстовая база данных периодических изданий ООО «ИВИС» <http://dlib.eastview.com>
2. Электронные версии периодических изданий, размещенные на сайте информационных ресурсов [www.polpred.com](http://www.polpred.com)
3. Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: <https://library.asu-edu.ru/catalog/>.
4. Электронный каталог «Научные журналы АГУ»: <http://journal.asu.edu.ru/issledovaniya-i-innovacii/11745-nauchnye-jurnaly-agu.html>.
5. Корпоративный проект Ассоциации региональных библиотечных консорциумов (АРБИКОН) «Межрегиональная аналитическая роспись статей» (МАРС) <http://mars.arbicon.ru>

## 7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

### 7.1. Паспорт фонда оценочных средств

При проведении текущего контроля и промежуточной аттестации по дисциплине (модулю) «Машинное обучение» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин (модулей) и прохождением практик, а в процессе освоения дисциплины (модуля) – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

**Таблица 6. Соответствие разделов, тем дисциплины (модуля), результатов обучения по дисциплине (модулю) и оценочных средств**

Контролируемый раздел, тема дисциплины (модуля)	Код контролируемой компетенции	Наименование оценочного средства
Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.	УК-1	Тест
Тема 2. Линейная регрессия. Логистическая регрессия	УК-1	Тест
Тема 3. Введение в машинное обучение	УК-1, ПК-3	Тест, отчет по лабораторной работе
Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.	УК-1, ПК-3	Тест, отчет по лабораторной работе
Тема 5. Обработка естественного языка.	УК-1, ПК-3	Тест, отчет по лабораторной работе
Тема 6. Модели последовательностей. Рекуррентные модели.	УК-1, ПК-3	Тест, лабораторная работа
Тема 7. Практические аспекты обучения глубоких нейронных сетей.	УК-1, ПК-3	Тест, отчет по лабораторной работе

## 7.2. Описание показателей и критериев оценивания компетенций, описание шкал оценивания

**Таблица 7. Показатели оценивания результатов обучения в виде знаний**

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует глубокое знание теоретического материала, умение обоснованно излагать свои мысли по обсуждаемым вопросам, способность полно, правильно и аргументированно отвечать на вопросы, приводить примеры
4 «хорошо»	демонстрирует знание теоретического материала, его последовательное изложение, способность приводить примеры, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует неполное, фрагментарное знание теоретического материала, требующее наводящих вопросов преподавателя, допускает существенные ошибки в его изложении, затрудняется в приведении примеров и формулировке выводов
2 «неудовлетворительно»	демонстрирует существенные пробелы в знании теоретического материала, не способен его изложить и ответить на наводящие вопросы преподавателя, не может привести примеры

**Таблица 8. Показатели оценивания результатов обучения в виде умений и владений**

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы
4 «хорошо»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует отдельные, несистематизированные навыки, испытывает затруднения и допускает ошибки при выполнении заданий, выполняет задание по подсказке преподавателя, затрудняется в формулировке выводов
2 «неудовлетворительно»	не способен правильно выполнить задания

## 7.3. Контрольные задания и иные материалы, необходимые для оценки результатов обучения по дисциплине (модулю)

### Тема 1. Линейная алгебра, краткий обзор. Векторы и матрицы.

#### 1. Тест «Векторы и матрицы»

Даны две матрицы:

$$A = \begin{bmatrix} 1 & -4 \\ -2 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 3 \\ 5 & 8 \end{bmatrix}$$

Найдите значение  $A - B$

Выберите один ответ:

- $\begin{bmatrix} 1 & -3 \\ 1 & -7 \end{bmatrix}$
- $\begin{bmatrix} 1 & 7 \\ 7 & 9 \end{bmatrix}$
- $\begin{bmatrix} 1 & -7 \\ -7 & -7 \end{bmatrix}$
- $\begin{bmatrix} 1 & -7 \\ -7 & 7 \end{bmatrix}$

...

## **Тема 2. Линейная регрессия. Логистическая регрессия.**

### 1. Тест «Градиентный спуск»

Для  $n$  признаков мы определяем функцию стоимости как:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Какие определения линейной регрессии из перечисленных ниже также будут правильными?

- $J(\theta) = \frac{1}{2m} \sum_{i=1}^m ((\sum_{j=0}^n \theta_j x_j^{(i)}) - y^{(i)})^2$
- $J(\theta) = \frac{1}{2m} \sum_{i=1}^m ((\sum_{j=0}^n \theta_j x_j^{(i)}) - \sum_{j=0}^n y_j^{(i)})^2$
- $J(\theta) = \frac{1}{2m} \sum_{i=1}^m (\theta^T x^{(i)} - y^{(i)})^2$
- $J(\theta) = \frac{1}{2m} \sum_{i=1}^m ((\sum_{j=1}^n \theta_j x_j^{(i)}) - y^{(i)})^2$

...

## **Тема 3. Введение в машинное обучение**

### 1. Тест «Введение в машинное обучение»

Вопрос 1

Диаграмма традиционного программирования на входе имеет Правила и Данные. А что у нее на выходе? Выберите один ответ:

- 1) Двоичные числа
- 2) Машинное обучение
- 3) Ошибки
- 4) Ответы

...

2. Лабораторная работа «Прогнозирование цен на недвижимость»
3. Лабораторная работа «Распознавание рукописных цифр»
4. Лабораторная работа «Распознавание положения рук жестового языка»

## **Тема 4. Компьютерное зрение. Сверточные сети на больших наборах данных.**

### 1. Тест

Вопрос 1

Какой параметр вы задали в методе fit, чтобы она использовала обратные вызовы? Выберите один ответ:

- 1) callbacks=
- 2) oncallbacks=
- 3) callback=
- 4) oncallback=

### Вопрос 2

Как свертки улучшают распознавание изображений? Выберите один ответ:

- 1) выделяют признаки в изображениях
- 2) делают изображение меньше
- 3) делают изображение более четким
- 4) ускоряют обработку изображений

### Вопрос 3

После уменьшения размера изображений результаты обучения стали другими. Почему? Выберите один ответ:

- 1) На изображениях было меньше информации
- 2) На изображениях было больше сжатой информации
- 3) Мы удалили некоторые свертки для обработки небольших изображений
- 4) Тренировка была быстрее

...

## 2. Тест

### Вопрос 1

Как дополнение изображений помогает решить проблему переобучения? Выберите один ответ:

- 1) Он автоматически подгоняет признаки к изображениям, находя их с помощью методов обработки изображений
- 2) манипулирует набором проверки, чтобы генерировать больше сценариев для признаков в изображениях.
- 3) манипулирует обучающим набором, чтобы генерировать больше сценариев для признаков в изображениях.
- 4) Замедляет процесс обучения

### Вопрос 2

При использовании переноса обучения как изменить количество классов, которые исходная модель может классифицировать? (т.е. исходная модель обрабатывает 1000 классов, но вам нужны только 2). Выберите один ответ:

- 1) Использовать все классы, но установить их вес в 0
- 2) Игнорировать все классы кроме вашего (т.е. номера 2 и выше, если я классифицирую только 2)
- 3) использовать Dropout для устранения нежелательных классов.
- 4) При добавлении DNN с мультиклассификацией в начало сети указывается выходной слой с числом нужных классов

### Вопрос 3

Почему переобучение с большей вероятностью случается на меньших наборах данных? Выберите один ответ:

- 1) Потому что там меньше вероятность того, что все возможные признаки обнаружатся в процессе обучения
- 2) Потому что не хватает данных для активации всех сверток или нейронов
- 3) Потому что в меньшем наборе данных ваши данные проверки с большей вероятностью будут похожи на ваши данные обучения
- 4) Поскольку при меньшем количестве данных обучение будет проходить быстрее, а некоторые признаки могут быть упущены

### Вопрос 4

Можно ли использовать аугментацию одновременно с методом Transfer Learning? Выберите один ответ:

- 1) Да. Это предварительно обученные слои замораживаются. Таким образом, вы можете аугментировать изображения, когда обучаете на них последние слои DNN.
- 2) Нет - так как если слои заморожены, они не могут быть аугментированы

...

3. *Лабораторная работа «Обучение классификатора изображений на 25 тысячах примеров»*

**Тема 5. Обработка естественного языка.**

1. *Тест*

Вопрос 1

Если вы не используете специальный токен для неизвестных слов, что произойдет при кодировании текста? Выберите один ответ:

- 1) Слово не закодируется и будет пропущено в предложении
- 2) слово не закодируется и составление последовательности остановится
- 3) Новое слово будет заменено наиболее используемым токеном
- 4) Слово не закодируется и будет заменено нулем в предложении

Вопрос 2

Что делает параметр `num_words = n` при токенизации корпуса текстов? Выберите один ответ:

- 1) число `n` определяет максимально возможное количество токенизируемых слов, и токенайзер выбирает `n` наиболее употребимых слов
- 2) если количество уникальных слов в корпусе больше `n`, будет выведено сообщение об ошибке
- 3) число `n` определяет максимально возможное количество токенизируемых слов, и токенайзер выбирает первые `n` слов, которые были токенизированы
- 4) это число определяет максимально возможное количество токенизируемых слов, и токенизация прекращается при его достижении

...

2. *Лабораторная работа «Организация структуры в текстах BBC News»*

**Тема 6. Модели последовательностей. Рекуррентные модели.**

1. *Тест «Рекуррентные сети»*

Вопрос 1

Если в предложении 120 токенов, и он проходит через Conv1D с 128 фильтрами и размером фильтра 5, какова форма вывода? Выберите один ответ:

- 1) (None, 116, 128)
- 2) (None, 120, 128)
- 3) (None, 116, 124)
- 4) (None, 120, 124)

2. *Лабораторная работа «Распознавание положения рук «камень ножницы бумага»»*

**Тема 7. Практические аспекты обучения глубоких нейронных сетей**

1. *Тест «Практические аспекты обучения сетей»*

Вопрос 1

Что такое распад весов? Выберите один ответ:

- 1) Постепенное искажение весов в нейронной сети, если она обучается на зашумленных данных.
- 2) Техника, позволяющая избежать исчезновения градиента путем наложения ограничений на значения весов.
- 3) Способ регуляризации (как при регуляризации L2), который во время градиентного спуска уменьшает веса на каждой итерации.
- 4) Процесс постепенного снижения скорости обучения в процессе обучения.

2. *Лабораторная работа «Обучение генератора стихов на корпусе пушкинских поэм»*

Лабораторные работы представляют собой код программы на языке Python в котором студенту предлагается дописать пропущенные фрагменты в соответствии с поставленной задачей. Выполненной лабораторная работа считается в том случае, когда весь код после запуска выдает требуемый результат.

**Пример задания лабораторной работы:** Напишите классификатор MNIST, который обучается до точности 99% или выше и делает это без фиксированного числа эпох - то есть вы должны прекратить обучение, как только достигнете этого уровня точности.

1. Этого надо достигнуть менее чем за 10 эпох, поэтому можно установить `epochs=10`, но не более.
2. Когда точность станет 99% или больше, надо распечатать строку «Достигнута точность 99%, поэтому обучение закончено!»

```
import tensorflow as tf
from os import path, getcwd, chdir
# заберите mnist.npz из github
# поместите в локальную папку и пропишите на нее путь
path = f"{getcwd()}../tmp2/mnist.npz"
def train_mnist():
    # Здесь ваш код
    import tensorflow as tf
    mnist = tf.keras.datasets.mnist
    (x_train, y_train),(x_test, y_test) = mnist.load_data()
    # Здесь ваш код
    model = tf.keras.models.Sequential([
    # Здесь ваш код
    ])
    model.compile(optimizer='adam',
                  loss='sparse_categorical_crossentropy',
                  metrics=['accuracy'])
    # Здесь ваш код
    return history.epoch, history.history['acc'][-1]
train_mnist()
```

### Перечень вопросов и заданий, выносимых на дифференцированный зачёт

1. Линейная регрессия.
2. Градиентный спуск
3. Линейная регрессия многих переменных.
4. Введение в машинное обучение
5. Новая парадигма программирования.
6. Обучение с учителем и обучение без учителя. Нейронные сети.
7. Концепция компьютерного зрения
8. Знакомство с понятием сверток
9. Построение модели сверточной сети для распознавания реальных изображений
10. Логистическая регрессия.
11. Классификаторы.
12. Модель логистической регрессии.
13. Сверточные сети на больших наборах данных
14. Исследование большого набора данных

15. Аугментация: расширение набора данных путем матричных преобразований
16. Трансферринг: Перенос готовых результатов обучения на свою модель
17. От двоичной ко множественной классификации.
18. Практические аспекты обучения сетей
19. Как выбрать способ улучшения производительности модели.
20. Регуляризация
21. проблема взрывающихся и исчезающих градиентов.
22. Обработка естественного языка
23. Определение эмоциональной составляющей фрагмента текста
24. Кластеризация, встраивание слов, многомерные векторы слов.
25. модели последовательных нейронных сетей
26. Генератор текстов.

**Таблица 9. Примеры оценочных средств с ключами правильных ответов**

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
<b><i>УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</i></b>				
1.	Задание закрытого типа	Блок-схема машинного обучения на входе получает Ответы и Данные. А что у нее на выходе? 1 Двоичные числа 2 Модели 3 Правила 4 Ошибки	2	1
2.		Что делает оптимизатор? 1 Генерирует новое улучшенное предположение 2 Выясняет, насколько эффективно скомпилирован ваш код 3 Принимает решение об остановке обучения нейронной сети 4 Измеряет, насколько хорошо текущее предположение	1	1
3.		С какой целью данные делят на обучающий и тестовый наборы? 1 чтобы обучить сеть с ранее неизвестными данными 2 чтобы тестировать сеть с ранее неизвестными данными 3 сделать тестирование быстрее 4 сделать обучение быстрее	2	1
4.		После уменьшения размера изображений результаты обучения стали другими. Почему? 1 На изображениях было меньше информации	1	1

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		2 На изображениях было больше сжатой информации 3 Мы удалили некоторые свертки для обработки небольших изображений 4 Тренировка была быстрее		
5.		Если размер изображения 150x150, и к нему применили свертку 3x3, какой размер получится в результате? Поясните ответ 1 450x450 2 148x148 3 153x153 4 150x150	2 $W' = W - F + 1$ , где: - $W'$ — новый размер ширины (или высоты), - $W$ — исходный размер ширины (или высоты), - $F$ — размер фильтра (стандартно - 1).	3
6.	Задание открытого типа	Что делает функция ReLU?	Для положительных значений аргумента ( $x > 0$ ) функция возвращает сам аргумент ( $x$ ). Для отрицательных значений аргумента ( $x \leq 0$ ) функция возвращает ноль.	3
7.		Если у вас предложения разной длины, каков будет результат использования <code>pad_sequences</code> ?	Они будут выровнены по самому длинному предложению, путем добавления нулей в конец	3
8.		У <code>embedding</code> есть параметр "размерность". Какое у него назначение?	это количество размерностей вектора, представляющего код слова	3
9.		Что такое шум, в контексте временных рядов?	Непредсказуемые изменения в данных временных рядов	3
10.		Что такое автокорреляция?	Автокорреляция — это мера корреляции временного ряда с самим собой, сдвинутого на некоторое количество временных интервалов назад или вперед. Говоря простыми словами, это степень, в которой значения ряда зависят друг от друга через временные	3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			промежутки.	
<b><i>ПК-3 Способен обеспечения эффективной работы баз данных, включая развертывание, сопровождение, оптимизация функционирования баз данных, являющихся частью различных информационных систем</i></b>				
11.	Задание закрытого типа	Какой метод вызывается, когда заканчивается эпоха обучения? 1 On_epoch_end 2 On_training_complete 3 On_end 4 On_epoch_finished	1	1
12.		Что такое свертка? 1 Механизм фильтрации нежелательных изображений 2 Механизм уменьшения изображений 3 Механизм выделения признаков на изображениях 4 Механизм увеличения изображений	3	1
13.		Почему переобучение с большей вероятностью случается на меньших наборах данных? Выберите один ответ: 1) Потому что там меньше вероятность того, что все возможные признаки обнаружатся в процессе обучения 2) Потому что не хватает данных для активации всех сверток или нейронов 3) Потому что в меньшем наборе данных ваши данные проверки с большей вероятностью будут похожи на ваши данные обучения 4) Поскольку при меньшем количестве данных обучение будет проходить быстрее, а некоторые признаки могут быть упущены	3	2
14.		Сверточные нейронные сети лучше подходят для классификации изображений, таких как лошади и люди, потому что: 1 Существует большое разнообразие людей 2 Существует большое	4	2

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		<p>разнообразие лошадей</p> <p>3 На этих изображениях признаки могут находиться в разных частях кадра</p> <p>4 Все перечисленное верно</p>		
15.		<p>Если вы не используете специальный токен для неизвестных слов, что произойдет при кодировании текста?</p> <p>1 Новое слово будет заменено наиболее используемым токеном</p> <p>2 Слово не закодируется и будет пропущено в предложении</p> <p>3 Слово не закодируется и будет заменено нулем в предложении</p> <p>4 Слово не закодируется и составление последовательности остановится</p>	2	2
16.	Задание открытого типа	Для каких задач применяется обучение с учителем?	Обучение с учителем (supervised learning) применяется для решения задач, в которых известны правильные ответы (метки классов или целевые переменные) для каждого примера в обучающем наборе. Основные типы задач, решаемых методом обучения с учителем: классификация, регрессия	3
17.		Для каких задач применяется обучение без учителя?	Обучение без учителя (unsupervised learning) применяется в ситуациях, когда отсутствуют размеченные данные, и целью является выявление скрытых структур или закономерностей в необработанных данных. Основные задачи, решаемые методами обучения без	3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			учителя: кластеризация, редукция размерности, ассоциативные правила	
18.		Как называется библиотека Tensorflow, содержащая общедоступные датасеты, которые можно использовать для тренировки и тестирования нейронных сетей?	TensorFlow Datasets	2
19.		Как рекуррентные нейронные сети помогают понять влияние последовательности на смысл?	Они сохраняют смысл от одной ячейки к другой	2
20.		Если потребовалось изменить скорость обучения оптимизатора на лету, после каждой эпохи, что нужно сделать?	В библиотеке Keras (которая входит в состав TensorFlow) предусмотрена специальная функциональность для динамического изменения скорости обучения на каждом этапе обучения. Эта возможность реализуется с помощью обратных вызовов (callbacks), одним из которых является LearningRateScheduler.	3

Полный комплект оценочных материалов по дисциплине (модулю) (фонд оценочных средств) хранится в электронном виде на кафедре, утверждающей рабочую программу дисциплины (модуля).

#### **7.4. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине (модулю)**

Итоговая оценка по промежуточной аттестации выставляется в соответствии с Положением АГУ о балльно-рейтинговой системе (БАРС). Итоговая оценка складывается из баллов, полученных студентом за текущую успеваемость. Для получения положительной оценки студенту необходимо набрать в семестре минимально 60 баллов. В течение семестра студент может набрать максимально 90 баллов за выполнение аудиторной и самостоятельной работы.

Для текущего контроля знаний, умений, навыков и (или) опыта деятельности, необходимых для формирования компетенции дисциплины «Машинное обучение», используется инструментарий системы Moodle: *Тест, Задание*.

Для стимулирования развития творческого и научно-исследовательского потенциала студентов при промежуточном оценивании предусмотрена система дополнительных баллов, а именно: начисление до 10 поощрительных баллов за участие в конференциях, семинарах, выставках и т.п. в области машинного обучения, программировании с представлением индивидуальных проектов в области машинного обучения.

Начисление баллов зависит от статуса мероприятия и статуса участия в нем студента. Начисление баллов происходит при предоставлении диплома, сертификата, грамоты, материалов конференции, опубликованной статьи, тезисов и т.п.

На зачетном занятии студенту выставляется итоговое количество баллов, полученные им в течение семестра за выполнение и отчет лабораторно-практических работ.

**Таблица 10. Технологическая карта рейтинговых баллов по дисциплине (модулю)**

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
<b>Основной блок</b>				
1.	<i>Тест</i>	8/5	40	По расписанию
2.	<i>Лабораторная работа</i>	6/7 1/8	50	
<b>Всего</b>			<b>90</b>	-
<b>Блок бонусов</b>				
3.	<i>Посещение занятий</i>		2	
4.	<i>Своевременное выполнение всех заданий</i>		2	
5.	<i>Участие в профильных мероприятиях</i>		6	
<b>Всего</b>			<b>10</b>	-
<b>ИТОГО</b>			<b>100</b>	-

**Таблица 11. Система штрафов (для одного занятия)**

Показатель	Балл
<i>Нарушение учебной дисциплины</i>	-1
<i>Неготовность к занятию</i>	-1
<i>Пропуск занятия без уважительной причины</i>	-1
<i>Списывание</i>	-5

**Таблица 12. Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине (модулю)**

Сумма баллов	Оценка по 4-балльной шкале	
90–100	5 (отлично)	Зачтено
85–89	4 (хорошо)	
75–84		
70–74		
65–69	3 (удовлетворительно)	Зачтено
60–64		
Ниже 60	2 (неудовлетворительно)	Не зачтено

При реализации дисциплины (модуля) в зависимости от уровня подготовленности обучающихся могут быть использованы иные формы, методы контроля и оценочные средства, исходя из конкретной ситуации.

## **8. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

### **8.1. Основная литература**

1. Мирзоев М.С., Основы математической обработки информации / М.С. Мирзоев - М. : Прометей, 2016. - 316 с. - ISBN 978-5-906879-01-1 - // ЭБС "Консультант студента": [сайт]. - URL : <http://www.studentlibrary.ru/book/ISBN9785906879011.html>.
2. Рыбина Г.В., Основы построения интеллектуальных систем : учеб. пособ./ Г.В. Рыбина. - М. : Финансы и статистика, 2014. - 432 с. - ISBN 978-5-279-03412-3 - // ЭБС "Консультант студента" : URL : <http://www.studentlibrary.ru/book/ISBN9785279034123.html>.
3. Системы искусственного интеллекта. Модуль "Модели и методы извлечения знаний" [Электронный ресурс] / Яковина И.Н. - Новосибирск : Изд-во НГТУ, 2014. - <http://www.studentlibrary.ru/book/ISBN9785778225879.html>.

### **8.2. Дополнительная литература**

1. Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах П. - М. : ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970602737.html>.
2. Рашка, С. Python и машинное обучение / С. Рашка; пер. с англ. А. В. Логунова. - Москва : ДМК Пресс, 2017. - 418 с. - ISBN 978-5-97060-409-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970604090.html>

### **8.3. Интернет-ресурсы, необходимые для освоения дисциплины (модуля)**

1. Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента». Многопрофильный образовательный ресурс «Консультант студента». [www.studentlibrary.ru](http://www.studentlibrary.ru).
2. среда разработки моделей машинного обучения <https://colab.research.google.com>
3. веб-сервис для хостинга IT-проектов и их совместной разработки <https://github.com>
4. онлайн-визуализатор n-мерных векторов <https://projector.tensorflow.org>

## **9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Учебные аудитории, библиотеки АГУ, компьютерные классы, мультимедийные аудитории.

Рабочая программа дисциплины (модуля) при необходимости может быть адаптирована для обучения (в том числе с применением дистанционных образовательных технологий) лиц с ограниченными возможностями здоровья, инвалидов. Для этого требуется заявление обучающихся, являющихся лицами с ограниченными возможностями здоровья, инвалидами, или их законных представителей и рекомендации психолого-медико-педагогической комиссии. Для инвалидов содержание рабочей программы дисциплины (модуля) может определяться также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).