

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 25.08.2025 12:11:10

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

**Автономная некоммерческая образовательная организация высшего образования  
«Европейский университет в Санкт-Петербурге»**

УТВЕРЖДАЮ:

Ректор  /В.В. Волков

« 24 » сентября 2024 г.

Протокол УС № 3 от 27 августа 2024 г.



Рабочая программа дисциплины  
**Машинное обучение: введение**

образовательная программа  
направление подготовки  
**09.04.03 Прикладная информатика**

направленность (профиль)  
**«Прикладной анализ данных»**  
программа подготовки – магистратура

язык обучения – русский  
форма обучения - очная

квалификация (степень) выпускника  
**Магистр**

**Санкт-Петербург**

**Автор:**

Тушканова О.Н., к.т.н., доцент факультета социологии АНООВО «ЕУСПб»

**Рецензент:**

Левшун Д.С., к.т.н., доцент факультета социологии АНООВО «ЕУСПб»

Рабочая программа дисциплины «**Машинное обучение: введение**», входящей в образовательную программу уровня магистратуры «Прикладной анализ данных», утверждена на заседании Совета факультета социологии.

Протокол заседания № 8 от 29 января 2024 года.

## АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ «Машинное обучение: введение»

Дисциплина «**Машинное обучение: введение**» является дисциплиной обязательной части основной профессиональной образовательной программы высшего образования «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика.

Дисциплина «**Машинное обучение: введение**» дает магистрантам представление о теоретических основах алгоритмов машинного обучения, типах задач анализа данных и методов анализ данных. В рамках дисциплины магистранты развивают навыки применения алгоритма машинного обучения в соответствии с поставленной задачей анализа данных, а также навыки реализации программ и применения современных методов и библиотек машинного обучения.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточный контроль в форме зачета (в третьем модуле), зачета с оценкой (в четвертом модуле).

Общая трудоемкость освоения дисциплины составляет 6 зачетных единиц, 216 часов.

## Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ .....	8
4. ОБЪЕМ ДИСЦИПЛИНЫ .....	8
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ .....	9
5.1 Содержание дисциплины .....	9
5.2 Структура дисциплины .....	12
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ .....	13
6.1 Общие положения.....	13
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины.....	13
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине.....	14
6.4 Перечень литературы для самостоятельной работы обучающегося:.....	15
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	15
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ .....	16
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации .....	16
7.2 Контрольные задания для текущей аттестации .....	20
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации .....	21
7.4 Типовые задания к промежуточной аттестации .....	24
7.5 Средства оценки индикаторов достижения компетенций .....	37
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА .....	38
8.1. Основная литература .....	38
8.2 Дополнительная литература .....	39
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	39
9.1 Программное обеспечение .....	39
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины: .....	39
9.3 Лицензионные электронные ресурсы библиотеки Университета .....	40
9.4 Электронная информационно-образовательная среда Университета.....	40
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА .....	41
ПРИЛОЖЕНИЕ 1 .....	42

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель** освоения дисциплины «**Машинное обучение: введение**» — ознакомление с теоретическими основами алгоритмов машинного обучения, типами задач анализа данных и методов анализ данных.

### Задачи:

1. Применение алгоритма машинного обучения в соответствии с поставленной задачей анализа данных.
2. Реализация программ и применение современных методов и библиотек машинного обучения.

## 2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: универсальными (УК), общепрофессиональными (ОПК) и профессиональными (ПК). Планируемые результаты формирования компетенций и индикаторы их достижения в результате освоения дисциплины представлены в Таблице 1.

### Планируемые результаты освоения дисциплины, соотнесенные с индикаторами достижения компетенций обучающихся

Таблица 1

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	ИД.УК-1.1. Анализирует проблемную ситуацию как систему, выявляя её составляющие и связи между ними	Знать: методы научного познания, в основе которых лежит рассмотрение объекта как системы: целостного комплекса взаимосвязанных элементов, методы и модели стратегического планирования З (УК-1)
	ИД.УК-1.2. Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению	Уметь: с использованием методов системного подхода анализировать альтернативные варианты решения исследовательских задач, вырабатывать стратегию действий и оценивать социальную эффективность реализации стратегических планов У (УК-1)
	ИД.УК-1.3. Критически оценивает надёжность источников информации, работает с противоречивой информацией из разных источников ИД.УК-1.4. Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов ИД.УК-1.5. Строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения	Владеть: целостной системой навыков методологического использования системного подхода при решении проблем, возникающих при выполнении исследовательских работ, навыками отстаивания своей точки зрения при выработке стратегических планов выполнения исследовательских работ В (УК-1)
УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	ИД.УК-6.1. Оценивает свои ресурсы и их пределы (личностные, ситуативные, временные) для успешного выполнения порученного задания	Знать: содержание процесса формирования целей профессионального и личностного развития З (УК-6)
	ИД.УК-6.2. Определяет образовательные потребности и способы совершенствования собственной (в т.ч. профессиональной) деятельности на основе самооценки ИД.УК-6.3. Выбирает и реализует с использованием инструментов	Уметь: применять полученные знания для формирования собственной жизненной стратегии с учётом индивидуально-личностных особенностей У (УК-6)

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
	непрерывного образования возможности развития профессиональных компетенций и социальных навыков ИД.УК-6.4. Выстраивает гибкую профессиональную траекторию, с учётом накопленного опыта профессиональной деятельности, изменяющихся требований рынка труда и стратегии личного развития	Владеть: приёмами и технологиями формирования целей саморазвития на основе самооценки В (УК-6)
ОПК-1 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	ИД.ОПК-1.1. Решает нестандартные задачи профессиональной деятельности с применением математических, естественнонаучных, социально-экономических методов ИД.ОПК-1.2. В условиях неопределённости и риска способен выработать эффективную стратегию действий, используя математические, естественнонаучные, социально-экономические методы науки	Знать: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности З (ОПК-1)
		Уметь: решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных социально-экономических и профессиональных знаний У (ОПК-1)
		Владеть: навыком применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач В (ОПК-1)
ОПК-4 Способен применять на практике новые научные принципы и методы исследований	ИД.ОПК-4.1. На основе современных теорий и концепций обосновывает актуальность постановки целей и задач научных исследований в профессиональной области знаний ИД.ОПК-4.2. Анализирует новые научные принципы и методы исследований в профессиональной области знаний ИД.ОПК-4.3. Применяет новые научные принципы и методы исследований в профессиональной области знаний ИД.ОПК-4.4. Разрабатывает предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний	Знать: актуальные направления применения новых научных принципов и методов исследований в профессиональной деятельности З (ОПК-4)
		Уметь: самостоятельно формировать планы и программы научных исследований с применением новых принципов и методов, характерных для выбранной отрасли науки У (ОПК-4)
		Владеть: навыками системного использования различных новых научных принципов и методов исследований для различных направлений науки В (ОПК-4)
ОПК-6 Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества	ИД.ОПК-6.1. Обосновывает актуальность постановки целей и задач исследования современных проблем и методов прикладной информатики и развития информационного общества ИД.ОПК-6.2. На основе актуальных теорий и концепций научных исследований формулирует задачи и гипотезы для поиска вариантов	Знать: содержание, объекты и субъекты информационного общества, критерии эффективности его функционирования; структуру интеллектуального капитала, проблемы инвестиций в экономику информатизации и методы оценки эффективности; правовые, экономические, социальные и психологические аспекты

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
	<p>решения современных проблем и методов прикладной информатики ИД.ОПК-6.3. Анализирует современные проблемы и методы прикладной информатики, а также направления развития информационного общества</p> <p>ИД.ОПК-6.4. Разрабатывает предложения и рекомендации для исследования современных проблем и методов прикладной информатики и развития информационного общества</p>	<p>информатизации; теоретические проблемы прикладной информатики, в том числе семантической обработки информации, развитие представлений об оценке качества информации в информационных системах; современные методы, средства, стандарты информатики для решения прикладных задач различных классов; правовые, экономические, социальные и психологические аспекты информатизации деятельности организационно-экономических систем; З (ОПК-6)</p> <p>Уметь: самостоятельно проводить анализ современных методов и средств информатики для решения прикладных задач различных классов У (ОПК-6)</p> <p>Владеть: навыками системного использования различных групп методов исследований современных проблем прикладной информатики и развития информационного общества В (ОПК-6)</p>
ОПК-7 Способен использовать методы научных исследований и математического моделирования в области проектирования и управления информационными системами	<p>ИД.ОПК-7.1. Обосновывает актуальность выбора определенных методов научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</p> <p>ИД.ОПК-7.2. Осуществляет методологическое обоснование научного исследования</p> <p>ИД.ОПК-7.3. Анализирует направления и методы научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</p>	<p>Знать: логические методы и приемы научного исследования; методологические принципы современной науки, направления, концепции, моделирования в области проектирования и управления информационными системами, источники знания и приемы работы с ними; основные особенности научного метода познания; программно-целевые методы решения научных проблем; основы моделирования управленческих решений; динамические оптимизационные модели; математические модели оптимального управления для непрерывных и дискретных процессов, их сравнительный анализ; многокритериальные методы принятия решений З (ОПК-7)</p> <p>Уметь: самостоятельно осуществлять методологическое обоснование научного исследования У (ОПК-7)</p> <p>Владеть: навыками системного использования различных групп методов научных исследований и математического моделирования в области</p>

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
		проектирования и управления информационными системами В (ОПК-7)
ПК-6 Способен применить технологии машинного обучения к реальным общественным задачам	ИД.ПК-6.1. Анализирует новые направления развития технологий машинного обучения в различных областях	Знать: методологический аппарат и направления развития технологий машинного обучения, в том числе для решения общественных задач З (ПК-6)
	ИД.ПК-6.2. Обосновывает перспективы применения технологий машинного обучения к решению конкретных общественных задач	Уметь: внедрять технологии машинного обучения в процесс решения различных общественных задач У (ПК-6)
	ИД.ПК-6.3. Формирует программы внедрения машинного обучения для решения определенных общественных задач ИД.ПК-6.4. Непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения	Владеть: навыками применения актуальных технологий машинного обучения к реальным общественным задачам В (ПК-6)

В результате освоения дисциплины магистрант должен:

— **знать:** методы сбора и обработки данных; особенности применения различных теоретико-методологических концепций с использованием технологий прикладного анализа данных; современные методы машинного обучения; современные библиотеки машинного обучения.

— **уметь:** использовать полученные знания и умения в профессиональной деятельности, деловой коммуникации и межличностном общении; корректно ставить задачи анализа данных; выбирать эффективные алгоритмы машинного обучения для решения поставленных задач; правильно оформлять и представлять результаты работы.

– **владеть:** навыками применение алгоритма машинного обучения в соответствии с поставленной задачей анализа данных; навыками реализации программ и применения современных методов и библиотек машинного обучения; способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности.

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «**Машинное обучение: введение**» является обязательной дисциплиной Блока 1 «Дисциплины (модули)» обязательной части образовательной программы «Прикладной анализ данных». Код дисциплины по Учебному плану Б1.О.09. Курс читается в третьем и четвертом модулях, форма промежуточной аттестации в третьем модуле – зачет, в четвертом модуле – зачет с оценкой.

Для успешного освоения данной дисциплины требуются знания, полученные в рамках прохождения обучения на уровне бакалавриата/ специалитета.

Знания, умения и навыки, полученные при освоении данной дисциплины, применяются магистрантами в процессе прохождения Б2.О.01(У) Технологической (проектно-технологической) практики и выполнения выпускной квалификационной работы.

### 4. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость освоения дисциплины составляет 6 (шесть) зачетных единиц, 216 часов.

Таблица 2

Объем дисциплины

Типы учебных занятий	Объем дисциплины
----------------------	------------------



и самостоятельная работа		Всего	Модуль									
			1	2	3	4	5	6	7	8	9	10
<b>Контактная работа обучающихся с преподавателем в соответствии с УП:</b>		<b>56</b>	-	-	28	28	-	-	-	-	-	-
Лекции (Л)		28	-	-	14	14	-	-	-	-	-	-
Лабораторные занятия (ЛЗ)		28	-	-	14	14	-	-	-	-	-	-
<b>Самостоятельная работа (СР)</b>		<b>160</b>	-	-	80	80	-	-	-	-	-	-
<b>Промежуточная аттестация</b>	<b>форма</b>	<b>Зачет с оценкой</b>	-	-	Зачет с оценкой	Зачет с оценкой	-	-	-	-	-	-
	<b>час.</b>	<b>-</b>	-	-	-	-	-	-	-	-	-	-
<b>Общая трудоемкость дисциплины (час./з.е.)</b>		<b>216/6</b>	-	-	108/3	108/3	-	-	-	-	-	-

## 5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

### 5.1 Содержание дисциплины

Таблица 3

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
<b>1</b>	Задачи машинного обучения. Важность данные в машинном обучении.	Формальная постановка задачи машинного обучения. Набор данных, выборка, объекты, признаки, переменные. Неструктурированные и структурированные данные. Типы признаков. Качество данных. Подходы к предобработке данных.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)
<b>2</b>	Основы машинного обучения. Линейная регрессия.	Обучение с учителем. Типы задач машинного обучения. Линейная регрессия. Полиномиальная регрессия. Регуляризация. Основы	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
		валидации модели (RMSE, MAE, тренировочная и тестовая выборки). Линейная регрессия для анализа временных рядов.		ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)
3	Логистическая регрессия. Линейная классификация.	Разделяющая гиперплоскость. Линейный классификатор. Логистическая регрессия. Градиентный спуск. Стратегии многоклассовой классификации. Основы валидации модели (частота ошибок, точность - accuracy).	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)
4	Валидация, оценка и сравнение моделей машинного обучения с учителем.	«Недообучение» и переобучение. Матрица неточностей, точность (precision), полнота, F-мера. Кривые ROC и PRC. Перекрестная проверка.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
				ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)
5	Деревья решений и ансамбли классификаторо в.	Жадное формирование деревьев. Критерии информативности. Бинарные и N-арные деревья. Рандомизация подвыборок. Бэггинг и бутстрап. Компоненты ошибки. Композиции деревьев. Случайный лес. Градиентный бустинг. Голосование классификаторов.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)
6	Обучение без учителя. Кластеризация.	Постановка задачи кластеризации. Основные алгоритмы и методы кластеризации. Меры сходства и различия. Алгоритм k- means. Иерархические алгоритмы. Расстояние между кластерами. Метрики оценки качества кластеризации.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
				ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	
7	Понижение размерности и отбор признаков.	Вычислительная сложность и избыточность модели. Оценка информативности признаков. Метод главных компонент. Другие способы отбора признаков.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)

## 5.2 Структура дисциплины

Таблица 4

Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП		СР	
			Л	ЛЗ		
Очная форма обучения						
Тема 1	Задачи машинного обучения и почему данные важны.	26	4	2	20	ПЗ, тест
Тема 2	Основы машинного обучения. Линейная регрессия.	28	4	4	20	ПЗ, тест
Тема 3	Логистическая регрессия. Линейная классификация.	28	4	4	20	ПЗ, тест
Тема 4	Валидация, оценка и сравнение моделей машинного обучения с учителем.	26	2	4	20	ПЗ, тест
Промежуточная аттестация		-	-	-	-	Зачет
Итого в 3 модуле:		108/3	14	14	80	-

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП		СР	
			Л	ЛЗ		
Очная форма обучения						
Тема 5	Деревья решений и ансамбли классификаторов.	34	4	4	26	ПЗ, тест
Тема 6	Обучение без учителя. Кластеризация.	34	4	4	26	ПЗ, тест
Тема 7	Понижение размерности и отбор признаков.	40	6	6	28	ПЗ, тест
Промежуточная аттестация		-	-	-	-	Зачет с оценкой
Итого в 4 модуле:		108/3	14	14	80	-
Всего:		216/6	28	28	160	

*\*Примечание: формы текущего контроля успеваемости: практическое задание (ПЗ).*

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 6.1 Общие положения

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответственным образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим лекциям/семинарским занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к семинарским занятиям, контрольному тесту также является важной формой работы магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

### 6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины

#### Тема 1. Задачи машинного обучения. Важность данные в машинном обучении:

1.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

1.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 20 часов.

## **Тема 2. Основы машинного обучения. Линейная регрессия:**

2.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

2.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 20 часов.

## **Тема 3. Логистическая регрессия. Линейная классификация:**

3.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

3.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 20 часов.

## **Тема 4. Валидация, оценка и сравнение моделей машинного обучения с учителем:**

4.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

4.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 20 часов.

## **Тема 5. Деревья решений и ансамбли классификаторов:**

5.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

5.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 16 часов. Итого: 26 часов.

## **Тема 6. Обучение без учителя. Кластеризация:**

6.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

6.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 16 часов. Итого: 26 часов.

## **Тема 7. Понижение размерности и отбор признаков:**

7.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 10 часов.

7.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 18 часов. Итого: 28 часов.

## **6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине**

Вопросы для самостоятельной подготовки по темам дисциплины:

1. Масштабирование и нормализация признаков. Бинаризация признаков. Выбросы. Пропущенные значения. Дублирование. Мета-данные. Форматы хранения данных.
2. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber. Метрики оценки регрессии – MAPE, SMAPE, R2.
3. Максимизация зазора и опорные вектора. Метод опорных векторов и его геометрический смысл. Функции ядра. Kernel trick.
4. Борьба с переобучением. Анализ эффективности (cost-benefit analysis). Выбор оптимального порога классификации для решающей функции.
5. Информационный выигрыш, критерий Джини. Метод случайных подпространств. Метод out-of-bag. Алгоритм CatBoost. Блендинг и стеккинг.
6. Графовые алгоритмы кластеризации. Вероятностные алгоритмы. EM-алгоритм. Алгоритмы, основанные на плотности. Алгоритм DBSCAN. Формула Ланса-Вильямса.
7. Метод случайных проекций. T-sne и визуализация. Permutation importance .
8. Регулярные выражения. Другие метрики оценки важности токенов. Эмбединги.

#### **6.4 Перечень литературы для самостоятельной работы обучающегося:**

1. Платонов, А.В. Машинное обучение: учебное пособие для вузов / А.В. Платонов. — Москва: Издательство Юрайт, 2024. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/544780>.
2. Вьюгин В.В. Математические основы теории машинного обучения и прогнозирования. М.: 2013. - 387 с.
3. Информационные системы и технологии / О. В. Лимановская, Т. И. Алферьева ; научный редактор И. Н. Обабков ; Министерство науки и высшего образования Российской Федерации, Уральский федеральный университет им. первого Президента России Б. Н. Ельцина. — Екатеринбург : Издательство Уральского университета, 2020. — 88 с. — ISBN 978-5-7996-3015-7.
4. Горбаченко, В. И. Машинное обучение : учебное пособие / В. И. Горбаченко, К. Е. Савенков, М. А. Малахов. — Москва : Ай Пи Ар Медиа, 2023. — 217 с. — ISBN 978-5-4497-1860-0. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/125886.html>

#### **6.5 Перечень учебно-методического обеспечения для самостоятельной работы**

Для обеспечения самостоятельной работы магистрантов по дисциплине **«Машинное обучение: введение»** разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

## 7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому лабораторному занятию, участие в подготовке практических заданий, написание тестов, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по обсуждаемым вопросам.

Текущий контроль проводится в форме оценивания выполненных практических заданий, оценка результатов тестов, демонстрирующих степень знакомства с дополнительной литературой.

Таблица 5

#### Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Задачи машинного обучения. Важность данные в машинном обучении.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 1, тест 1	зачтено/ не зачтено
Основы машинного обучения. Линейная регрессия.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1)	Практическое задание 2, тест 2	зачтено/ не зачтено



Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
Логистическая регрессия. Линейная классификация.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2, тест 3	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Валидация, оценка и сравнение моделей машинного обучения с учителем.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2, тест 4	зачтено/ не зачтено
Деревья решений и ансамбли классификаторов.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3, тест 5	зачтено/ не зачтено
Обучение без учителя. Кластеризация.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1)	Практическое задание 3, тест 6	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
Понижение размерности и отбор признаков.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3, тест 7	зачтено/ не зачтено

Таблица 6

## Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Практическое задание	магистрант выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо

Формы текущего контроля успеваемости	Критерии оценивания
	выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено
Тест	верные ответы менее, чем на 50% вопросов теста — не зачтено верные ответы на 30 и более % вопросов теста — зачтено

## 7.2 Контрольные задания для текущей аттестации

### Примерные вопросы для тестов по темам 1-7

1. Предположим, у вас есть набор данных с  $m = 10$  объектов и  $n = 2$  признаками для каждого объекта. Вы хотите построить модель многомерной линейной регрессии для этого набора данных. Что вы предпочтёте, градиентный спуск или аналитическое решение уравнения регрессии?

○ Градиентный спуск, так как он всегда будет сходиться к оптимальному вектору весов  $w$ .

○ Градиентный спуск, так как обращение матрицы ( $X$  транспонированное  $\cdot X$ ), которое необходимо для аналитического решения, будет выполняться очень медленно на такой выборке.

○ Аналитическое решение.

2. На рисунке ниже в столбце  $y$  приведены правильные ответы для объектов из некоторой выборки, а в столбце  $a(x)$  - предсказания некоторого алгоритма  $a(x)$  для тех же объектов. Посчитайте среднеквадратичную ошибку (MSE) алгоритма  $a(x)$  на выборке, представленной в таблице. Ответ округлите до целого числа, используя общепринятые правила округления.

3. Пусть для некоторого уравнения регрессии:

$$y(x_1, x_2) = w_0 + w_1 \cdot x_1 + w_2 \cdot x_2$$

заданы следующие веса:  $w_0 = 12$ ,  $w_1 = 1$ ,  $w_2 = -1$ .

Какое значение примет  $y$  при  $x_1 = 5$ ,  $x_2 = 8$ ?

### Примерный материал практических заданий:

#### Тема 1. Задачи машинного обучения. Важность данные в машинном обучении.

##### Практическое задание 1:

Исследуйте набор данных, представленный в файле «credit.csv», с помощью изученных подходов и методов предобработки данных и ответьте на следующие вопросы, сопровождая ответы необходимыми манипуляциями и вычислениями:

##### Вопросы:

1. В какой шкале представлены все признаки в наборе?
2. Укажите признак с наименьшей дисперсией, если это возможно.
3. Медиана какого признака отличается для двух классов сильнее всего?
4. В каком классе (признак class) больше всего unskilled работников мужского рода? Какой у них средний возраст?
5. Есть ли в наборе признаки с выбросами? Какие?
6. Сколько пропущенных значений в наборе? Как они обозначены
7. Сколько пропущенных значений в наборе? Примените минимум две техники заполнения пропущенных значений. Обоснуйте свое решение.
8. Как минимум для одного признака каждого типа постройте диаграмму, которая может помочь его исследовать.

9. Проанализируйте взаимосвязи признаков с помощью «scatter plot» матрицы и тепловой карты корреляций. Какие выводы можно сделать относительно взаимосвязи признаков в этом наборе данных?

10. Какой(-ие) признак(-и) данного набора имеет смысл стандартизировать? Почему?

11. Какой признак(и) (если такой(ие) есть) кажется вам проблематичным? Что бы вы предложили с ним сделать?

12. Придумайте один новый признак, попытайтесь объяснить, почему он может быть информативен.

**Темы 2-4. Основы машинного обучения. Линейная регрессия. Логистическая регрессия. Линейная классификация. Валидация, оценка и сравнение моделей машинного обучения с учителем.**

Практическое задание 2:

Практическое задание проводится в форме приватного соревнования на платформе Kaggle. Студентам в малых группах предлагается построить наиболее точную модель регрессии для заданного набора данных. Ответ студентов должен включать весь разработанный код, а решение должно быть воспроизводимо. Точность итоговой модели проверяется на отложенной выборке с помощью метрики RMSE (среднеквадратичное отклонение). Зачет выставляется студентам, которые сумели преодолеть базовый уровень точности (англ. baseline). Базовый уровень точности выбирается в зависимости от общего уровня группы.

**Темы 5-7. Деревья решений и ансамбли классификаторов. Обучение без учителя. Кластеризация. Понижение размерности и отбор признаков.**

Практическое задание 3:

Практическое задание проводится в форме приватного соревнования на платформе Kaggle. Студентам в малых группах предлагается построить наиболее точный классификатор для заданного набора данных.

Пример такого набора - это набор Churn, который содержит данные от телекоммуникационной компании. Данные содержат информацию о пользователях компании, их демографических характеристиках, услугах, которыми они пользуются, продолжительности пользования услугами компании, способе оплаты и размере оплаты. Задача состоит в том, чтобы спрогнозировать отток пользователей (выявить людей, которые будут и не будут продлевать свой контракт с компанией).

После завершения соревнования студентам необходимо представить исполняемый код, который был использован для обучения их классификаторов.

Зачет выставляется студентам, которые сумели преодолеть базовый уровень точности (англ. baseline). Базовый уровень точности выбирается в зависимости от общего уровня группы.

### **7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

**Форма промежуточной аттестации в третьем модуле – зачет**, при выставлении которого учитываются результаты текущего контроля успеваемости обучающегося по дисциплине и результаты тестирования, выраженные в 100-балльной шкале.

Перед зачетом проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

**Форма промежуточной аттестации в четвертом модуле – зачет с оценкой**, при выставлении которого учитываются результаты текущего контроля успеваемости

обучающегося по дисциплине и результаты тестирования, выраженные в 100-балльной шкале.

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают оценку по дисциплине.

Тест включает 20 вопросов по всем компетенциям дисциплины, 10 из них вопросы закрытого типа, 10 – открытого типа, все вопросы разного уровня сложности.

Тест оценивается в баллах в соответствии со следующими критериями:

#### **Задания закрытого типа**

*Базовый уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

*Повышенный уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балла; если на любой одной позиции ответа записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

#### **Задания открытого типа**

*Высокий уровень сложности:* магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left( \frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где F – итоговое количество баллов за тест,

K – количество осваиваемых в рамках дисциплины компетенций,

$k_n$  – максимально возможное количество баллов за вопросы по компетенции,

$x_n$  – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Таблица 7

### **Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
Зачет / тестирование	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4)	41-100% правильных ответов	Зачтено
				0-40% правильных ответов	Не зачтено

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
		ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
<b>Зачет оценкой/тестирование</b>	с УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	81-100% правильных ответов	Зачтено, отлично
				61-80% правильных ответов	Зачтено, хорошо
				41-60% правильных ответов	Зачтено, удовлетворительно
				0-40% правильных ответов	Не зачтено, неудовлетворительно

Результаты сдачи промежуточной аттестации по направлениям подготовки уровня магистратуры оцениваются по стобалльной системе оценки в соответствии с Положением о формах, периодичности и порядке организации и проведения текущего контроля успеваемости и промежуточной аттестации обучающихся в АНООВО «ЕУСПб» следующим образом согласно таблице 7а.

Таблица 7а

**Система оценки знаний обучающихся**

Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «не зачтено», показывают несформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, отлично», «зачтено, хорошо», «зачтено, удовлетворительно», показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценках «не зачтено, неудовлетворительно», показывают несформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

#### **7.4 Типовые задания к промежуточной аттестации**

##### **УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий**

###### **Задания закрытого типа**

*Базовый уровень сложности*

###### **1. Прочитайте задание, выберите правильный ответ.**

Что из перечисленного является правильным определением машинного обучения согласно Т. Митчеллу?

- 1) Программирование поведения системы с нуля
- 2) Поиск закономерностей в данных без использования меток
- 3) Программа учится на опыте Е, выполняя задание Т, измеряемое по метрике Р
- 4) Создание искусственного интеллекта любой ценой
- 5) Прямое копирование поведения человека

Поле для ответа:

###### **2. Прочитайте задание, выберите правильные ответы.**

Какие характеристики описывают «большие данные» (Big Data)?

- 1) Высокая точность данных
- 2) Объём данных
- 3) Скорость поступления данных
- 4) Ручная обработка данных

Поле для ответа:

###### **3. Прочитайте задание, выберите правильный ответ.**

Что происходит при слишком большой скорости обучения (learning rate) в градиентном спуске?

- 1) Алгоритм быстро находит минимум



- 2) Модель начинает учитывать больше признаков
- 3) Возможен перескок через минимум и расходимость
- 4) Обучение становится точнее

Поле для ответа:

*Повышенный уровень сложности*

4. **Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое модель в контексте машинного обучения?

- 1) Набор признаков, описывающих объект
- 2) Метод визуализации данных
- 3) Решающая функция, отображающая объект в ответ
- 4) Процесс сбора данных
- 5) Таблица объект-признак

Поле для ответа:

Обоснование: \_\_\_\_\_.

5. **Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что означает термин «переобучение» в контексте деревьев решений?

- 1) Алгоритм не может обучиться на данных
- 2) Дерево имеет слишком много параметров и плохо обобщает новые данные
- 3) Алгоритм игнорирует часть данных
- 4) Модель обучается быстрее, чем нужно

Поле для ответа:

6. **Прочитайте задание и установите соответствие.**

Соотнесите типы нейронных сетей (слева) с их применением (справа).

- А) Полносвязная сеть
- Б) Сверточная сеть
- В) Рекуррентная сеть
- Г) Генеративная сеть

- 1) Хорошо работает с изображениями
- 2) Используется для обработки последовательностей
- 3) Применяется для генерации новых данных
- 4) Содержит связи между каждым нейроном соседних слоёв

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

7. **Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие ошибки используют для оценки качества модели линейной регрессии?

- 1) Среднеквадратичная ошибка (MSE)

- 2) Точность (ассигасу)
- 3) Средняя абсолютная ошибка (MAE)
- 4) F-мера

Поле для ответа:

Обоснование: \_\_\_\_\_.

**8. Прочитайте текст и напишите обоснованный ответ.**

Чем машинное обучение отличается от традиционного программирования?

Поле для ответа:

**УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки**

**Задания закрытого типа**

*Базовый уровень сложности*

**9. Прочитайте задание, выберите правильные ответы.**

Какие этапы входят в стандартную схему анализа данных?

- 1) Извлечение признаков
- 2) Обработка естественного языка
- 3) Оптимизация нейросети
- 4) Очистка данных

Поле для ответа:

**10. Прочитайте задание, выберите правильные ответы.**

Какие из следующих функций могут использоваться как функция потерь в задачах регрессии?

- 1) MSE
- 2) Log-Loss
- 3) Accuracy
- 4) Cross-Entropy
- 5) MSLE

Поле для ответа:

**11. Прочитайте задание, выберите правильный ответ.**

Что делает логистическая регрессия?

- 1) Кластеризует объекты
- 2) Предсказывает числа
- 3) Даёт вероятность принадлежности к классу
- 4) Убирает лишние признаки
- 5) Строит дерево

Поле для ответа:

*Повышенный уровень сложности*

**12. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какой тип задачи машинного обучения относится к обучению без учителя?

- 1) Регрессия
- 2) Многоклассовая классификация
- 3) Кластеризация
- 4) Прогнозирование
- 5) Ранжирование

Поле для ответа:

Обоснование: \_\_\_\_\_.

**13. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое гиперплоскость в линейной классификации?

- 1) Случайный вектор признаков
- 2) Линия, делящая данные на обучающую и тестовую выборки
- 3) Поверхность, разделяющая классы
- 4) Среднее значение признаков
- 5) Граница между признаками

Поле для ответа:

Обоснование: \_\_\_\_\_.

**14. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что делает метод Random Forest при обучении?

- 1) Строит одно дерево и использует все признаки
- 2) Использует нейронную сеть для финального предсказания
- 3) Объединяет множество деревьев, построенных на случайных подвыборках
- 4) Удаляет признаки с низкой корреляцией

Поле для ответа:

Обоснование: \_\_\_\_\_.

*Высокий уровень сложности*

**15. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие величины используются в метриках качества классификации?

- 1) TP (True Positive)
- 2) FP (False Positive)
- 3) MSE (среднеквадратичная ошибка)
- 4) Logits

Поле для ответа:

Обоснование: \_\_\_\_\_.

**16. Прочитайте текст и напишите обоснованный ответ.**

Какие типы признаков используются в машинном обучении и как они различаются?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**ОПК-1 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте**

**Задания закрытого типа**

*Базовый уровень сложности*

**17. Прочитайте задание, выберите правильный ответ.**

Что представляет собой целевая переменная  $y$  в задачах машинного обучения?

- 1) Массив признаков
- 2) Ответ на задачу, предсказываемый моделью
- 3) Метка ошибки
- 4) Алгоритм построения модели
- 5) Подбор гиперпараметров

Поле для ответа:

Обоснование: \_\_\_\_\_.

**18. Прочитайте задание, выберите правильные ответы.**

Какие из перечисленных функций относятся к оценке качества модели?

- 1) Среднеквадратичная ошибка (MSE)
- 2) Функция потерь
- 3) Бинаризация
- 4) Регуляризация
- 5) One-hot кодирование

Поле для ответа:

Обоснование: \_\_\_\_\_.

**19. Прочитайте задание, выберите правильные ответы.**

Какие из перечисленных методов являются модификациями градиентного спуска?

- 1) Метод Нестерова
- 2) Метод  $k$ -ближайших соседей
- 3) Метод моментов
- 4) Метод наименьших квадратов
- 5) Метод опорных векторов

Поле для ответа:

Обоснование: \_\_\_\_\_.

*Повышенный уровень сложности*

**20. Прочитайте задание и установите соответствие.**

Установите соответствие между типами задач машинного обучения и их описаниями.

Тип задачи:

- А) Кластеризация
- Б) Регрессия
- В) Бинарная классификация
- Г) Визуализация

Описание:

- 1) Группировка объектов без меток
- 2) Предсказание числового значения
- 3) Сопоставление объекта одному из двух классов
- 4) Представление данных в понятной форме на плоскости или в пространстве

Поле для ответа:

**21. Прочитайте задание и установите последовательность.**

Установите правильную последовательность этапов обучения нейронной сети.

- А) Расчёт ошибки
- Б) Обратное распространение ошибки
- В) Прямой проход
- Г) Обновление весов

Поле для ответа:

**22. Прочитайте задание и установите последовательность.**

Расположите этапы построения дерева решений в правильном порядке (слева – буквы, справа – правильная последовательность).

- 1) Выбор признака для разбиения
- 2) Определение критерия останова
- 3) Разбиение данных по выбранному признаку
- 4) Построение листа (конечного узла)

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**23. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответа.**

Какие типы признаков относятся к категориальным (дискретным)?

- 1) Бинарный
- 2) Количественный
- 3) Номинальный
- 4) Порядковый
- 5) Вещественный

Поле для ответа:

Обоснование: \_\_\_\_\_.

**24. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какая метрика показывает долю правильно классифицированных объектов?

- 1) Precision
- 2) Recall
- 3) Accuracy
- 4) Loss
- 5) Margin

Поле для ответа:

Обоснование: \_\_\_\_\_.

**ОПК-4 Способен применять на практике новые научные принципы и методы исследований**

**Задания закрытого типа**

*Базовый уровень сложности*

**25. Прочитайте задание, выберите правильный ответ.**

Какой тип задачи относится к обучению без учителя?

- 1) Бинарная классификация
- 2) Регрессия
- 3) Кластеризация
- 4) Ранжирование
- 5) Прогнозирование

Поле для ответа:

**26. Прочитайте задание, выберите правильный ответ.**

Какой критерий используется при обучении линейной регрессии?

- 1) Средняя абсолютная ошибка
- 2) Среднеквадратичная ошибка
- 3) Максимальная правдоподобность
- 4) Коэффициент детерминации

Поле для ответа:

**27. Прочитайте задание, выберите правильные ответы.**

Какие утверждения справедливы для нормализации признаков?

- 1) Min-Max нормализация масштабирует значения в интервал  $[0,1]$
- 2) Standard scaling приводит признак к среднему 1 и стандартному отклонению 0
- 3) Standard scaling позволяет использовать правило трёх сигм в случае нормального распределения
- 4) Нормализация всегда ухудшает качество модели

Поле для ответа:

*Повышенный уровень сложности*

**28. Прочитайте задание и установите соответствие.**

Установите соответствие между типами признаков и примерами значений.

Тип признака:

- А) Бинарный
- Б) Номинальный
- В) Порядковый
- Г) Количественный

Пример значения:

- 1) Уровень образования: начальное, среднее, высшее
- 2) Пол: 0 – муж., 1 – жен.
- 3) Цвет: красный, синий, зелёный
- 4) Рост в сантиметрах

Поле для ответа:

**29. Прочитайте задание и установите соответствие.**

Соотнесите тип метрики с её назначением:

- А) MSE
- Б) MAE
- В)  $R^2$  (коэффициент детерминации)

- 1) Показывает, насколько модель объясняет дисперсию целевой переменной
- 2) Оценивает среднеквадратичную ошибку
- 3) Оценивает среднюю абсолютную ошибку

**30. Прочитайте задание и установите последовательность.**

Расположите шаги оценки качества классификатора по матрице ошибок:

- 1) Построение confusion matrix (матрицы ошибок)
- 2) Подсчёт TP, FP, TN, FN
- 3) Вычисление метрик (Accuracy, Precision, Recall)
- 4) Интерпретация результатов

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**31. Прочитайте текст и напишите обоснованный ответ.**

Что такое дерево решений и для чего оно используется в машинном обучении?

Поле для ответа:

**32. Прочитайте текст и напишите обоснованный ответ.**

Что такое нейрон в контексте искусственной нейронной сети?

Поле для ответа:

**ОПК-6 Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества**

**Задания закрытого типа**

*Базовый уровень сложности*

**33. Прочитайте задание, выберите правильный ответ.**

Какая функция ошибки используется для оценки качества модели в задаче регрессии?

- 1) Log-loss
- 2) Mean Squared Error (MSE)
- 3) Precision
- 4) Accuracy
- 5) AUC

Поле для ответа:

**34. Прочитайте задание, выберите правильный ответ.**

Для чего используется регуляризация в линейной регрессии?

- 1) Для увеличения сложности модели
- 2) Для улучшения качества на обучающей выборке
- 3) Для борьбы с переобучением
- 4) Для оценки значимости признаков
- 5) Для ускорения градиентного спуска

Поле для ответа:

**35. Прочитайте задание, выберите правильные ответы.**

Что происходит при переобучении модели линейной регрессии?

- 1) Ошибка на тестовой выборке ниже, чем на обучающей
- 2) Модель хорошо аппроксимирует как обучающие, так и тестовые данные
- 3) Весовые коэффициенты могут иметь большие значения
- 4) Качество на новых данных ухудшается

Поле для ответа:

*Повышенный уровень сложности*

**36. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что означает термин «антиградиент» в контексте градиентного спуска?

- 1) Максимальное значение функции
- 2) Направление наискорейшего убывания функции
- 3) Набор весовых коэффициентов
- 4) Сумма всех значений признаков

Поле для ответа:

Обоснование: \_\_\_\_\_.

**37. Прочитайте задание и установите последовательность.**

Установите правильную последовательность этапов процесса анализа данных.

- 1) Очистка данных



- 2) Сбор данных
- 3) Построение модели
- 4) Оценка модели на тестовых данных

Поле для ответа:

**38. Прочитайте задание и установите соответствие.**

Соотнесите термин с его определением:

- А) Регуляризация
- Б) Градиентный спуск
- В) Предобработка данных
- Г) MAE

- 1) Метод оптимизации параметров модели
- 2) Подготовка признаков перед обучением
- 3) Средняя абсолютная ошибка
- 4) Метод борьбы с переобучением через ограничение весов

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**39. Прочитайте текст и напишите обоснованный ответ.**

Что такое ансамбль моделей и какую основную пользу он приносит?

Поле для ответа:

**40. Прочитайте задание и установите соответствие.**

Соотнесите метрики расстояния с их описанием:

- А) Евклидово
- Б) Манхэттенское
- В) Косинусное
- Г) Хэммингово

- 1) Измеряет угол между векторами
- 2) Сумма абсолютных разностей координат
- 3) Расстояние "по прямой линии" между точками
- 4) Количество позиций с разными символами

Поле для ответа:

**41. Прочитайте текст и напишите обоснованный ответ.**

Зачем делят данные на обучающую и тестовую выборки?

Поле для ответа:

**ОПК-7 Способен использовать методы научных исследований и математического моделирования в области проектирования и управления информационными системами**

**Задания закрытого типа**

*Базовый уровень сложности*

**42. Прочитайте задание, выберите правильные ответы.**

Какие утверждения верны для признаков в машинном обучении?

- 1) Бинарные признаки принимают только значения 0 и 1
- 2) Порядковые признаки не имеют смысла в обучении
- 3) Количественные признаки могут принимать любые вещественные значения
- 4) Номинальные признаки всегда упорядочены

Поле для ответа:

**43. Прочитайте задание, выберите правильный ответ.**

Что такое свободный коэффициент в линейной регрессии?

- 1) Коэффициент перед самым важным признаком
- 2) Произвольное число в формуле
- 3) Смещение (bias), добавляемое ко всем предсказаниям
- 4) Среднее значение целевой переменной

Поле для ответа:

*Повышенный уровень сложности*

**44. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое переобучение (overfitting) в модели линейной регрессии?

- 1) Модель показывает хорошие результаты на новых данных
- 2) Модель плохо обучается на тренировочных данных
- 3) Модель слишком точно подстроилась под обучающую выборку и плохо работает на новых данных
- 4) Модель не использует все признаки

Поле для ответа:

Обоснование: \_\_\_\_\_.

**45. Прочитайте задание и установите последовательность.**

Расположите шаги градиентного спуска в правильном порядке:

- 1) Проверка условия останова
- 2) Выбор начальной точки
- 3) Обновление весов по антиградиенту
- 4) Вычисление градиента

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**46. Прочитайте текст и напишите обоснованный ответ.**

Что такое выбросы в данных?

Поле для ответа:

**47. Прочитайте текст и напишите обоснованный ответ.**

Что происходит на этапе прямого прохода в нейронной сети?

Поле для ответа:

**ПК-6 Способен применить технологии машинного обучения к реальным общественным задачам**

**Задания закрытого типа**

*Базовый уровень сложности*

**48. Прочитайте задание, выберите правильный ответ.**

Какой тип признаков предполагает упорядоченное конечное множество значений?

- 1) Бинарные
- 2) Номинальные
- 3) Порядковые
- 4) Количественные
- 5) Категориальные

Поле для ответа:

**49. Прочитайте задание, выберите правильные ответы.**

Какие задачи относятся к обучению с учителем?

- 1) Кластеризация
- 2) Регрессия
- 3) Бинарная классификация
- 4) Визуализация

Поле для ответа:

**50. Прочитайте задание, выберите правильный ответ.**

Что делает метод one-hot encoding?

- 1) Удаляет ненужные признаки
- 2) Преобразует числовые признаки в категориальные
- 3) Кодировать категориальные признаки в набор бинарных
- 4) Удаляет строки с пропущенными значениями

Поле для ответа:

*Повышенный уровень сложности*

**51. Прочитайте задание и установите последовательность.**

Расположите этапы оценки качества модели линейной регрессии:

- 1) Разделение выборки на тренировочную и тестовую
- 2) Обучение модели на тренировочной выборке
- 3) Применение модели к тестовой выборке
- 4) Подсчёт метрики качества (например, MSE)

Поле для ответа:

**52. Прочитайте задание и установите соответствие.**

Установите соответствие между метриками и их назначением:

- А) Accuracy
- Б) Precision
- В) Recall
- Г) F1-мера

- 1) Среднее между точностью и полнотой
- 2) Доля правильно классифицированных объектов
- 3) Доля верно угаданных среди всех угаданных
- 4) Доля верно угаданных среди всех объектов нужного класса

Поле для ответа:

**53. Прочитайте задание и установите соответствие.**

Соотнесите функции активации (слева) с их особенностями (справа).

- А) ReLU
- Б) Сигмоида
- В) tanh
- Г) Линейная функция

- 1) Возвращает значение от 0 до 1
- 2) Возвращает максимум между 0 и входным значением
- 3) Использует гиперболический тангенс
- 4) Возвращает входное значение без изменений

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**54. Прочитайте текст и напишите обоснованный ответ.**

Почему важно нормализовать входные данные при обучении нейронной сети?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**55. Прочитайте текст и напишите обоснованный ответ.**

Как работает обучение с подкреплением?

Поле для ответа:

Обоснование: \_\_\_\_\_.

## 7.5 Средства оценки индикаторов достижения компетенций

Таблица 8

**Средства оценки индикаторов достижения компетенций**

<b>Коды компетенций</b>	<b>Индикаторы компетенций</b> (в соот.с Таблицей 1)	<b>Средства оценки (в соот. с Таблицами 5, 7)</b>
УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	Практическое задание, тест
УК-6	ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4.	Практическое задание, тест
ОПК-1	ИД.ОПК-1.1. ИД.ОПК-1.2.	Практическое задание, тест
ОПК-4	ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	Практическое задание, тест
ОПК-6	ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4.	Практическое задание, тест
ОПК-7	ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3.	Практическое задание, тест
ПК-6	ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	Практическое задание, тест

Таблица 9

**Описание средств оценки индикаторов достижения компетенций**

<b>Средства оценки</b> (в соот. С Таблицами 5, 7)	<b>Рекомендованный план выполнения работы</b>
Практическое задание	<p>Магистрант в ходе подготовки и выполнения практических заданий показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> <li>1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</li> <li>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</li> <li>3. Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</li> <li>4. На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</li> </ol>

Средства оценки (в соот. С Таблицами 5, 7)	Рекомендованный план выполнения работы
	<p>5. Обосновывать актуальность постановки целей и задач исследования современных проблем и методов прикладной информатики и развития информационного общества, формулировать задачи и гипотезы для поиска вариантов решения современных проблем и методов прикладной информатики, на основе анализа направлений развития информационного общества разрабатывать предложения и рекомендации для решения современных проблем прикладной информатики и развития информационного общества</p> <p>6. Обосновывать актуальность выбора определенных методов научных исследований, анализировать направления и методы научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</p> <p>7. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения</p>
Тест	<p>Магистрант в ходе подготовки и выполнения тестов показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</p> <p>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</p> <p>3. Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</p> <p>4. На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</p> <p>5. Обосновывать актуальность постановки целей и задач исследования современных проблем и методов прикладной информатики и развития информационного общества, формулировать задачи и гипотезы для поиска вариантов решения современных проблем и методов прикладной информатики, на основе анализа направлений развития информационного общества разрабатывать предложения и рекомендации для решения современных проблем прикладной информатики и развития информационного общества</p> <p>6. Обосновывать актуальность выбора определенных методов научных исследований, анализировать направления и методы научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</p> <p>7. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения</p>

## 8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

### 8.1. Основная литература

1. Бруссард, М. Искусственный интеллект: пределы возможного / Мерedit Бруссард ; пер. с англ. - Москва : Альпина нон-фикшн, 2020. - 362 с. - ISBN 978-5-00139-

080-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1220958> . – Режим доступа: по подписке

## **8.2 Дополнительная литература**

1. Андрейчиков, А. В. Интеллектуальные информационные системы и методы искусственного интеллекта : учебник / А.В. Андрейчиков, О.Н. Андрейчикова. — Москва : ИНФРА-М, 2021. — 530 с. + Доп. материалы [Электронный ресурс]. — (Высшее образование: Магистратура). — DOI 10.12737/1009595. - ISBN 978-5-16-014883-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1009595> . – Режим доступа: по подписке.

3. Болотова, Ю. А. Методы и алгоритмы интеллектуальной обработки цифровых изображений : учеб. пособие / Ю.А. Болотова, А.А. Друки, В.Г. Спицын ; Томский политехнический университет. - Томск : Изд-во Томского политехнического университета, 2016. - 208 с. - ISBN 978-5-4387-0710-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1043928> . – Режим доступа: по подписке

4. Трегуб, И. В. Имитационные модели принятия решений : учебное пособие / И. В. Трегуб, Т. А. Горошникова. — Москва : ИНФРА-М, 2020. — 193 с. — (Высшее образование: Магистратура). - ISBN 978-5-16-015393-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1030572> . – Режим доступа: по подписке

## **9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

### **9.1 Программное обеспечение**

При осуществлении образовательного процесса магистрантами и профессорско-преподавательским составом используется следующее лицензионное программное обеспечение:

1. OS Microsoft Windows (OVS OS Platform)
2. MS Office (OVS Office Platform)
3. Adobe Acrobat Professional 11.0 MLP AOO License RU
4. Adobe CS5.5 Design Standart Win IE EDU CLP
5. ABBYY FineReader 11 Corporate Edition
6. ABBYY Lingvo x5
7. Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
8. Adobe Acrobat Reader DC /Pro – бесплатно
9. Google Chrome – бесплатно
10. Opera – бесплатно
11. Mozilla – бесплатно
12. VLC – бесплатно

### **9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:**

#### **Информационно-справочные системы**

1. Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
2. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
3. Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npod.ru>
4. Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
5. Официальный интернет-портал правовой информации. Государственная

система правовой информации: <http://pravo.gov.ru>

6. Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
7. Российское образование. Федеральный портал: <http://www.edu.ru>

### **Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:**

1. Google. Книги: <https://books.google.com>
2. Internet Archive: <https://archive.org>
3. Koob.ru. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>
4. [SOC.LIB.RU](http://soc.lib.ru/). Социология, психология, управление: <http://soc.lib.ru/>
5. Socioline.ru. Учебники, монографии по социологии: <http://socioline.ru>
6. Библиотека Гумер – гуманитарные науки: <http://www.gumer.info>
7. ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
8. Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
9. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
10. Неприкосновенный запас: <http://magazines.russ.ru/nz/>
11. Президентская библиотека: <http://www.prlib.ru>
12. Российская государственная библиотека: <http://www.rsl.ru/>
13. Российская национальная библиотека: <http://www.nlr.ru/poisk/>

## **9.3 Лицензионные электронные ресурсы библиотеки Университета**

### **Профессиональные базы данных:**

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronic-resources>, включая следующие базы данных:

1. **eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: <http://elibrary.ru>;
2. **Университетская информационная система РОССИЯ** — база электронных ресурсов для учебных программ и исследовательских проектов в области социально-гуманитарных наук: <http://www.uisrussia.msu.ru/>;
3. Электронные журналы по подписке (текущие номера научных зарубежных журналов)

### **Электронные библиотечные системы:**

1. **Znanium.com** — Электронная библиотечная система (ЭБС) — <http://znanium.com/>;
2. Университетская библиотека онлайн — Электронная библиотечная система (ЭБС) — <http://biblioclub.ru/>

## **9.4 Электронная информационно-образовательная среда Университета**

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — Sakai@EU, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://eusp.org/>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:



— доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;

— фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;

— формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;

— взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине.

## **10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

**Для лиц с ограниченными возможностями здоровья и инвалидов** предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ**

**«Машинное обучение: введение»**

# ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

## 1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому лабораторному занятию, участие в подготовке практических заданий, написание тестов, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по обсуждаемым вопросам.

Текущий контроль проводится в форме оценивания выполненных практических заданий, оценка результатов тестов, демонстрирующих степень знакомства с дополнительной литературой.

Таблица 1

**Показатели, критерии и оценивание компетенций и индикаторов их  
достижения в процессе текущей аттестации**

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Задачи машинного обучения. Важность данные в машинном обучении.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 1, тест 1	зачтено/ не зачтено
Основы машинного обучения. Линейная регрессия.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1)	Практическое задание 2, тест 2	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
Логистическая регрессия. Линейная классификация.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2, тест 3	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Валидация, оценка и сравнение моделей машинного обучения с учителем.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 2, тест 4	зачтено/ не зачтено
Деревья решений и ансамбли классификаторов.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4) В (ОПК-4) 3 (ОПК-6) У (ОПК-6) В (ОПК-6) 3 (ОПК-7) У (ОПК-7) В (ОПК-7) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3, тест 5	зачтено/ не зачтено
Обучение учителя. Кластеризация.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1)	Практическое задание 3, тест 6	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
Понижение размерности и отбор признаков.	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 3, тест 7	зачтено/ не зачтено

Таблица 2

## Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Практическое задание	магистрант выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо

Формы текущего контроля успеваемости	Критерии оценивания
	выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено
Тест	верные ответы менее, чем на 50% вопросов теста — не зачтено верные ответы на 30 и более % вопросов теста — зачтено

## 2 Контрольные задания для текущей аттестации

### Перечень вопросов для тестов по темам 1-7

1. Предположим, у вас есть набор данных с  $m = 10$  объектов и  $n = 2$  признаками для каждого объекта. Вы хотите построить модель многомерной линейной регрессии для этого набора данных. Что вы предпочтёте, градиентный спуск или аналитическое решение уравнения регрессии?

- Градиентный спуск, так как он всегда будет сходиться к оптимальному вектору весов  $w$ .
- Градиентный спуск, так как обращение матрицы ( $X$  транспонированное  $\cdot X$ ), которое необходимо для аналитического решения, будет выполняться очень медленно на такой выборке.
- Аналитическое решение.

2. На рисунке ниже в столбце  $y$  приведены правильные ответы для объектов из некоторой выборки, а в столбце  $a(x)$  - предсказания некоторого алгоритма  $a(x)$  для тех же объектов. Посчитайте среднеквадратичную ошибку (MSE) алгоритма  $a(x)$  на выборке, представленной в таблице. Ответ округлите до целого числа, используя общепринятые правила округления.

3. Пусть для некоторого уравнения регрессии:

$$y(x_1, x_2) = w_0 + w_1 \cdot x_1 + w_2 \cdot x_2$$

заданы следующие веса:  $w_0 = 12$ ,  $w_1 = 1$ ,  $w_2 = -1$ .

Какое значение примет  $y$  при  $x_1 = 5$ ,  $x_2 = 8$ ?

### Материал практических заданий:

**Тема 1. Задачи машинного обучения. Важность данные в машинном обучении.**

#### Практическое задание 1:

Исследуйте набор данных, представленный в файле «credit.csv», с помощью изученных подходов и методов предобработки данных и ответьте на следующие вопросы, сопровождая ответы необходимыми манипуляциями и вычислениями:

#### Вопросы:

1. В какой шкале представлены все признаки в наборе?
2. Укажите признак с наименьшей дисперсией, если это возможно.
3. Медиана какого признака отличается для двух классов сильнее всего?
4. В каком классе (признак class) больше всего unskilled работников мужского рода? Какой у них средний возраст?
5. Есть ли в наборе признаки с выбросами? Какие?
6. Сколько пропущенных значений в наборе? Как они обозначены
7. Сколько пропущенных значений в наборе? Примените минимум две техники заполнения пропущенных значений. Обоснуйте свое решение.
8. Как минимум для одного признака каждого типа постройте диаграмму, которая может помочь его исследовать.

9. Проанализируйте взаимосвязи признаков с помощью «scatter plot» матрицы и тепловой карты корреляций. Какие выводы можно сделать относительно взаимосвязи признаков в этом наборе данных?

10. Какой(-ие) признак(-и) данного набора имеет смысл стандартизировать? Почему?

11. Какой признак(и) (если такой(ие) есть) кажется вам проблематичным? Что бы вы предложили с ним сделать?

12. Придумайте один новый признак, попытайтесь объяснить, почему он может быть информативен.

**Темы 2-4. Основы машинного обучения. Линейная регрессия. Логистическая регрессия. Линейная классификация. Валидация, оценка и сравнение моделей машинного обучения с учителем.**

Практическое задание 2:

Практическое задание проводится в форме приватного соревнования на платформе Kaggle. Студентам в малых группах предлагается построить наиболее точную модель регрессии для заданного набора данных. Ответ студентов должен включать весь разработанный код, а решение должно быть воспроизводимо. Точность итоговой модели проверяется на отложенной выборке с помощью метрики RMSE (среднеквадратичное отклонение). Зачет выставляется студентам, которые сумели преодолеть базовый уровень точности (англ. baseline). Базовый уровень точности выбирается в зависимости от общего уровня группы.

**Темы 5-7. Деревья решений и ансамбли классификаторов. Обучение без учителя. Кластеризация. Понижение размерности и отбор признаков.**

Практическое задание 3:

Практическое задание проводится в форме приватного соревнования на платформе Kaggle. Студентам в малых группах предлагается построить наиболее точный классификатор для заданного набора данных.

Пример такого набора - это набор Churn, который содержит данные от телекоммуникационной компании. Данные содержат информацию о пользователях компании, их демографических характеристиках, услугах, которыми они пользуются, продолжительности пользования услугами компании, способе оплаты и размере оплаты. Задача состоит в том, чтобы спрогнозировать отток пользователей (выявить людей, которые будут и не будут продлевать свой контракт с компанией).

После завершения соревнования студентам необходимо представить исполняемый код, который был использован для обучения их классификаторов.

Зачет выставляется студентам, которые сумели преодолеть базовый уровень точности (англ. baseline). Базовый уровень точности выбирается в зависимости от общего уровня группы.

### **3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

**Форма промежуточной аттестации в третьем модуле – зачет**, при выставлении которого учитываются результаты текущего контроля успеваемости обучающегося по дисциплине и результаты тестирования, выраженные в 100-балльной шкале.

Перед зачетом проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

**Форма промежуточной аттестации в четвертом модуле – зачет с оценкой**, при выставлении которого учитываются результаты текущего контроля успеваемости



обучающегося по дисциплине и результаты тестирования, выраженные в 100-балльной шкале.

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают оценку по дисциплине.

Тест включает 20 вопросов по всем компетенциям дисциплины, 10 из них вопросы закрытого типа, 10 – открытого типа, все вопросы разного уровня сложности.

Тест оценивается в баллах в соответствии со следующими критериями:

#### **Задания закрытого типа**

*Базовый уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте - 1 балл; ответ отличен от эталонного - 0 баллов.

*Повышенный уровень сложности:* задание считается выполненным верно, если ответ полностью совпадает с эталоном ответа: каждый символ в ответе стоит на своём месте, лишние символы в ответе отсутствуют - 2 балла; если на любой одной позиции ответа записан не тот символ, который представлен в эталоне ответа - 1 балл; во всех других случаях выставляется 0 баллов

#### **Задания открытого типа**

*Высокий уровень сложности:* магистрант демонстрирует умение применять знания в нестандартной ситуации, решать нетиповые задачи, приводит корректные обоснования и доказательства, ответ полный, в ответе отсутствуют фактические ошибки, изложение связное, структура прозрачная, логика изложения прослеживается - 3 балла; ответ значительно отличается от эталонного, имеются фактические ошибки, искажающие его смысл или ответ сформулирован неверно или не сформулирован - 0 баллов.

Итоговый балл за тест рассчитывается по формуле:

$$F = \frac{100}{K} * \left( \frac{x_1}{k_1} + \frac{x_2}{k_2} + \dots + \frac{x_n}{k_n} \right),$$

где F – итоговое количество баллов за тест,

K – количество осваиваемых в рамках дисциплины компетенций,

$k_n$  – максимально возможное количество баллов за вопросы по компетенции,

$x_n$  – количество баллов, набранное магистрантом, за правильные ответы на вопросы по соответствующей компетенции.

Таблица 7

### **Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации**

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
Зачет / тестирование	УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ОПК-1) У (ОПК-1) В (ОПК-1) 3 (ОПК-4) У (ОПК-4)	41-100% правильных ответов	Зачтено
				0-40% правильных ответов	Не зачтено

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. с Таблицей 1)	Коды ЗУВ (в соот. с Таблицей 1)	Критерии оценивания	Оценка
		ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)		
<b>Зачет оценкой/тестирование</b>	с УК-1 УК-6 ОПК-1 ОПК-4 ОПК-6 ОПК-7 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ОПК-1.1. ИД.ОПК-1.2. ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4. ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4. ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ОПК-1) У (ОПК-1) В (ОПК-1) З (ОПК-4) У (ОПК-4) В (ОПК-4) З (ОПК-6) У (ОПК-6) В (ОПК-6) З (ОПК-7) У (ОПК-7) В (ОПК-7) З (ПК-6) У (ПК-6) В (ПК-6)	81-100% правильных ответов	Зачтено, отлично
				61-80% правильных ответов	Зачтено, хорошо
				41-60% правильных ответов	Зачтено, удовлетворительно
				0-40% правильных ответов	Не зачтено, неудовлетворительно

Результаты сдачи промежуточной аттестации по направлениям подготовки уровня магистратуры оцениваются по стобалльной системе оценки в соответствии с Положением о формах, периодичности и порядке организации и проведения текущего контроля успеваемости и промежуточной аттестации обучающихся в АНООВО «ЕУСПб» следующим образом согласно таблице 3а.

Таблица 3а

**Система оценки знаний обучающихся**

Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «не зачтено», показывают несформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, отлично», «зачтено, хорошо», «зачтено, удовлетворительно», показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценках «не зачтено, неудовлетворительно», показывают несформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

#### **4 Задания к промежуточной аттестации**

##### **УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий**

##### **Задания закрытого типа**

*Базовый уровень сложности*

##### **1. Прочитайте задание, выберите правильный ответ.**

Что из перечисленного является правильным определением машинного обучения согласно Т. Митчеллу?

- 1) Программирование поведения системы с нуля
- 2) Поиск закономерностей в данных без использования меток
- 3) Программа учится на опыте Е, выполняя задание Т, измеряемое по метрике Р
- 4) Создание искусственного интеллекта любой ценой
- 5) Прямое копирование поведения человека

Поле для ответа:

##### **2. Прочитайте задание, выберите правильные ответы.**

Какие характеристики описывают «большие данные» (Big Data)?

- 1) Высокая точность данных
- 2) Объём данных
- 3) Скорость поступления данных
- 4) Ручная обработка данных

Поле для ответа:

##### **3. Прочитайте задание, выберите правильный ответ.**

Что происходит при слишком большой скорости обучения (learning rate) в градиентном спуске?

- 1) Алгоритм быстро находит минимум
- 2) Модель начинает учитывать больше признаков

- 3) Возможен перескок через минимум и расходимость
- 4) Обучение становится точнее

Поле для ответа:

**4. Прочитайте задание, выберите правильный ответ.**

Что показывает точность (ассигасу) классификации?

- 1) Сколько признаков у объекта
- 2) Сколько времени работает алгоритм
- 3) Долю правильно угаданных ответов
- 4) Количество классов в задаче
- 5) Размер обучающей выборки

Поле для ответа:

**5. Прочитайте задание, выберите правильный ответ.**

Какой метод максимизирует ширину разделяющей полосы между классами?

- 1) Логистическая регрессия
- 2) Метод ближайших соседей
- 3) Метод опорных векторов (SVM)
- 4) Дерево решений
- 5) Метод главных компонент

Поле для ответа:

**6. Прочитайте задание, выберите правильный ответ.**

Что происходит в узле дерева решений?

- 1) Случайный выбор следующего признака
- 2) Расчет среднего значения признака
- 3) Разделение данных по условию на признаке
- 4) Построение линейной модели

Поле для ответа:

**7. Прочитайте задание, выберите правильный ответ.**

Что оценивают метрики качества кластеризации?

- 1) Количество признаков в данных
- 2) Насколько хорошо объекты внутри кластера похожи друг на друга
- 3) Количество данных в обучающей выборке
- 4) Скорость работы алгоритма

Поле для ответа:

**8. Прочитайте задание, выберите правильный ответ.**

Какой из алгоритмов кластеризации основан на выборе центров кластеров и их обновлении?

- 1) DBSCAN

- 2) K-means
- 3) Агломеративная кластеризация
- 4) PCA

Поле для ответа:

*Повышенный уровень сложности*

**9. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое модель в контексте машинного обучения?

- 1) Набор признаков, описывающих объект
- 2) Метод визуализации данных
- 3) Решающая функция, отображающая объект в ответ
- 4) Процесс сбора данных
- 5) Таблица объект-признак

Поле для ответа:

Обоснование: \_\_\_\_\_.

**10. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что означает термин «переобучение» в контексте деревьев решений?

- 1) Алгоритм не может обучиться на данных
- 2) Дерево имеет слишком много параметров и плохо обобщает новые данные
- 3) Алгоритм игнорирует часть данных
- 4) Модель обучается быстрее, чем нужно

Поле для ответа:

**11. Прочитайте задание и установите соответствие.**

Соотнесите типы нейронных сетей (слева) с их применением (справа).

- А) Полносвязная сеть
- Б) Сверточная сеть
- В) Рекуррентная сеть
- Г) Генеративная сеть

- 1) Хорошо работает с изображениями
- 2) Используется для обработки последовательностей
- 3) Применяется для генерации новых данных
- 4) Содержит связи между каждым нейроном соседних слоёв

Поле для ответа:

**12. Прочитайте задание и установите соответствие.**

Соотнесите понятия, связанные с деревьями решений (слева), с их определениями (справа).

- А) Узел (Node)
- Б) Лист (Leaf)

- В) Разбиение (Split)
- Г) Глубина дерева (Depth)

- 1) Конечная точка дерева, где принимается решение
- 2) Процесс деления данных на подмножества по признаку
- 3) Максимальное количество уровней в дереве
- 4) Точка внутри дерева, где данные разделяются

Поле для ответа:

**13. Прочитайте задание и установите соответствие.**

Соотнесите методы кластеризации с их характеристиками:

- А) K-means
- Б) DBSCAN
- В) Агломеративная кластеризация

- 1) Иерархический метод, объединяющий объекты
- 2) Основан на плотности, выделяет шум
- 3) Требуется задать число кластеров заранее

Поле для ответа:

**14. Прочитайте задание и установите соответствие.**

Соотнесите понятия, связанные с типами машинного обучения (слева), с их описанием (справа).

- А) Обучение с учителем
- Б) Обучение без учителя
- В) Обучение с подкреплением
- Г) Обучение по частичному обучающему набору

- 1) Метод, где агент обучается на основе награды
- 2) Метод, в котором используются размеченные данные
- 3) Метод, где алгоритм сам ищет структуры в неразмеченных данных
- 4) Используется ограниченное количество размеченных примеров

Поле для ответа:

**15. Прочитайте задание и установите соответствие.**

Соотнесите слои нейронной сети (слева), с их назначением (справа).

- А) Входной слой
- Б) Скрытый слой
- В) Выходной слой
- Г) Слой нормализации

- 1) Подаёт исходные данные в модель
- 2) Производит предсказание
- 3) Обработывает данные внутри модели
- 4) Преобразует данные к общему масштабу

Поле для ответа:

**16. Прочитайте задание и установите соответствие.**

Соотнесите функции активации (слева), с их характеристиками (справа).

- А) ReLU
- Б) Sigmoid
- В) Tanh
- Г) Softmax

- 1) Выдаёт значения от 0 до 1
- 2) Используется для классификации по классам
- 3) Обнуляет отрицательные значения
- 4) Возвращает значения от  $-1$  до 1

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**17. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие ошибки используют для оценки качества модели линейной регрессии?

- 1) Среднеквадратичная ошибка (MSE)
- 2) Точность (accuracy)
- 3) Средняя абсолютная ошибка (MAE)
- 4) F-мера

Поле для ответа:

Обоснование: \_\_\_\_\_.

**18. Прочитайте текст и напишите обоснованный ответ.**

Чем машинное обучение отличается от традиционного программирования?

Поле для ответа:

**19. Прочитайте текст и напишите обоснованный ответ.**

Что показывает значение сигмоидной функции в логистической регрессии?

Поле для ответа:

**20. Прочитайте текст и напишите обоснованный ответ.**

Что такое кластеризация в машинном обучении?

Поле для ответа:

**21. Прочитайте текст и напишите обоснованный ответ.**

Что такое обучение с учителем и в чём его основная особенность?

Поле для ответа:

**22. Прочитайте текст и напишите обоснованный ответ.**

Для чего используется функция активации в нейронной сети?

Поле для ответа:

**УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки**

**Задания закрытого типа**

*Базовый уровень сложности*

**23. Прочитайте задание, выберите правильные ответы.**

Какие этапы входят в стандартную схему анализа данных?

- 1) Извлечение признаков
- 2) Обработка естественного языка
- 3) Оптимизация нейросети
- 4) Очистка данных

Поле для ответа:

**24. Прочитайте задание, выберите правильные ответы.**

Какие из следующих функций могут использоваться как функция потерь в задачах регрессии?

- 1) MSE
- 2) Log-Loss
- 3) Accuracy
- 4) Cross-Entropy
- 5) MSLE

Поле для ответа:

**25. Прочитайте задание, выберите правильный ответ.**

Что делает логистическая регрессия?

- 1) Кластеризует объекты
- 2) Предсказывает числа
- 3) Даёт вероятность принадлежности к классу
- 4) Убирает лишние признаки
- 5) Строит дерево

Поле для ответа:

**26. Прочитайте задание, выберите правильные ответы.**

Какие методы относятся к линейной классификации?

- 1) Логистическая регрессия
- 2) Метод опорных векторов (SVM)



- 3) Дерево решений
- 4) Случайный лес
- 5) Метод k ближайших соседей

Поле для ответа:

**27. Прочитайте задание, выберите правильный ответ.**

Какой критерий используется для оценки качества разбиения в задаче классификации?

- 1) Среднеквадратичная ошибка
- 2) Энтропия
- 3) Евклидово расстояние
- 4) Коэффициент корреляции

Поле для ответа:

**28. Прочитайте задание, выберите правильный ответ.**

Какая из метрик расстояния измеряет расстояние "по прямой линии" между точками?

- 1) Манхэттенское расстояние
- 2) Евклидово расстояние
- 3) Косинусное расстояние
- 4) Хэммингово расстояние

Поле для ответа:

**29. Прочитайте задание, выберите правильный ответ.**

Что такое агломеративная кластеризация?

- 1) Алгоритм, который делит данные на кластеры сверху вниз
- 2) Метод кластеризации, основанный на плотности
- 3) Иерархический метод кластеризации, начинающий с отдельных объектов и объединяющий их в кластеры
- 4) Метод, который выбирает случайные центры кластеров

Поле для ответа:

**30. Прочитайте задание, выберите правильный ответ.**

Какой тип нейронной сети чаще всего используется для обработки изображений?

- 1) Рекуррентная нейронная сеть
- 2) Сверточная нейронная сеть
- 3) Байесовская сеть
- 4) Генетический алгоритм

Поле для ответа:

*Повышенный уровень сложности*

**31. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какой тип задачи машинного обучения относится к обучению без учителя?

- 1) Регрессия

- 2) Многоклассовая классификация
- 3) Кластеризация
- 4) Прогнозирование
- 5) Ранжирование

Поле для ответа:

Обоснование: \_\_\_\_\_.

**32. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое гиперплоскость в линейной классификации?

- 1) Случайный вектор признаков
- 2) Линия, делящая данные на обучающую и тестовую выборки
- 3) Поверхность, разделяющая классы
- 4) Среднее значение признаков
- 5) Граница между признаками

Поле для ответа:

Обоснование: \_\_\_\_\_.

**33. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что делает метод Random Forest при обучении?

- 1) Строит одно дерево и использует все признаки
- 2) Использует нейронную сеть для финального предсказания
- 3) Объединяет множество деревьев, построенных на случайных подвыборках
- 4) Удаляет признаки с низкой корреляцией

Поле для ответа:

Обоснование: \_\_\_\_\_.

**34. Прочитайте задание и установите соответствие.**

Соотнесите этапы работы нейронной сети (слева) с их описаниями (справа).

- А) Прямой проход
- Б) Вычисление ошибки
- В) Обратное распространение
- Г) Обновление весов

- 1) Сравнение предсказания с правильным ответом
- 2) Пересчёт связей на основе градиента
- 3) Распространение входных данных вперёд по сети
- 4) Корректировка параметров нейронов

Поле для ответа:

**35. Прочитайте задание и установите соответствие.**

Соотнесите методы ансамблей (слева) с их характеристиками (справа).

- А) Бэггинг (Bagging)
- Б) Бустинг (Boosting)
- В) Стекинг (Stacking)
- Г) Случайный лес (Random Forest)

- 1) Объединение нескольких моделей для уменьшения переобучения с помощью случайного отбора признаков
- 2) Последовательное обучение моделей, каждая из которых исправляет ошибки предыдущих
- 3) Обучение нескольких моделей на разных выборках с последующим усреднением результатов
- 4) Комбинирование различных моделей с обучением мета-модели для итогового предсказания

Поле для ответа:

**36. Прочитайте задание и установите последовательность.**

Расположите этапы работы метода K-means в правильном порядке:

- 1) Назначение объектов ближайшему центроиду
- 2) Инициализация центроидов
- 3) Обновление центроидов
- 4) Оценка сходимости алгоритма

Поле для ответа:

**37. Прочитайте задание и установите соответствие.**

Соотнесите термины, связанные с переобучением (слева), с их определениями (справа).

- А) Переобучение
- Б) Недообучение
- В) Регуляризация
- Г) Проверочная выборка

- 1) Недостаточная сложность модели
- 2) Используется для оценки качества модели
- 3) Слишком хорошее запоминание обучающих данных
- 4) Метод предотвращения переобучения

Поле для ответа:

**38. Прочитайте задание и установите соответствие.**

Соотнесите типы задач машинного обучения (слева), с примерами применения (справа).

- А) Классификация
- Б) Регрессия
- В) Кластеризация
- Г) Ранжирование

- 1) Разделение клиентов по типу поведения
- 2) Определение вероятности болезни
- 3) Предсказание стоимости жилья

4) Упорядочивание товаров по предпочтению

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**39. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие величины используются в метриках качества классификации?

- 1) TP (True Positive)
- 2) FP (False Positive)
- 3) MSE (среднеквадратичная ошибка)
- 4) Logits

Поле для ответа:

Обоснование: \_\_\_\_\_.

**40. Прочитайте текст и напишите обоснованный ответ.**

Какие типы признаков используются в машинном обучении и как они различаются?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**41. Прочитайте текст и напишите обоснованный ответ.**

В чём идея метода опорных векторов (SVM)?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**42. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие два метода относятся к бустингу?

- 1) Bagging
- 2) AdaBoost
- 3) Random Forest
- 4) Gradient Boosting
- 5) KNN

Поле для ответа:

Обоснование: \_\_\_\_\_.

**43. Прочитайте текст и напишите обоснованный ответ.**

Для чего используют метрику косинусного расстояния?

Поле для ответа:

Обоснование: \_\_\_\_\_.

44. Прочитайте текст и напишите обоснованный ответ.

Что такое переобучение и чем оно опасно?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**ОПК-1 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте**

**Задания закрытого типа**

*Базовый уровень сложности*

45. Прочитайте задание, выберите правильный ответ.

Что представляет собой целевая переменная  $y$  в задачах машинного обучения?

- 1) Массив признаков
- 2) Ответ на задачу, предсказываемый моделью
- 3) Метка ошибки
- 4) Алгоритм построения модели
- 5) Подбор гиперпараметров

Поле для ответа:

Обоснование: \_\_\_\_\_.

46. Прочитайте задание, выберите правильные ответы.

Какие из перечисленных функций относятся к оценке качества модели?

- 1) Среднеквадратичная ошибка (MSE)
- 2) Функция потерь
- 3) Бинаризация
- 4) Регуляризация
- 5) One-hot кодирование

Поле для ответа:

Обоснование: \_\_\_\_\_.

47. Прочитайте задание, выберите правильные ответы.

Какие из перечисленных методов являются модификациями градиентного спуска?

- 1) Метод Нестерова
- 2) Метод  $k$ -ближайших соседей
- 3) Метод моментов
- 4) Метод наименьших квадратов
- 5) Метод опорных векторов

Поле для ответа:

Обоснование: \_\_\_\_\_.

48. Прочитайте задание, выберите правильный ответ.

Что делает метод опорных векторов (SVM)?

- 1) Объединяет похожие объекты в группы
- 2) Делает случайный выбор
- 3) Максимально разделяет классы с помощью полосы
- 4) Уменьшает количество признаков
- 5) Переводит текст в числа

Поле для ответа:

Обоснование: \_\_\_\_\_.

**49. Прочитайте задание, выберите правильные ответы.**

Что делает функция потерь в машинном обучении?

- 1) Измеряет ошибку
- 2) Увеличивает скорость
- 3) Помогает обучать модель
- 4) Сортирует признаки
- 5) Добавляет новые данные

Поле для ответа:

Обоснование: \_\_\_\_\_.

**50. Прочитайте задание, выберите правильный ответ.**

Что такое переобучение дерева решений?

- 1) Дерево слишком простое
- 2) Дерево не может обучиться
- 3) Дерево имеет слишком много листьев и запоминает данные
- 4) Дерево использует неправильные признаки

Поле для ответа:

Обоснование: \_\_\_\_\_.

**51. Прочитайте задание, выберите правильный ответ.**

Что такое кластеризация в машинном обучении?

- 1) Метод регрессии
- 2) Метод классификации с учителем
- 3) Метод группировки данных по сходству
- 4) Метод снижения размерности

Поле для ответа:

Обоснование: \_\_\_\_\_.

**52. Прочитайте задание, выберите правильный ответ.**

Какой метод кластеризации основывается на плотности и может находить кластеры произвольной формы?

- 1) K-means
- 2) DBSCAN
- 3) Агломеративная кластеризация
- 4) Линейная регрессия

Поле для ответа:

Обоснование: \_\_\_\_\_.

**53. Прочитайте задание, выберите правильный ответ.**

Что означает «слой» в нейронной сети?

- 1) Группа связанных компьютеров
- 2) Набор нейронов, работающих на одном уровне
- 3) Отдельный файл с данными
- 4) Программа для работы с текстом

Поле для ответа:

Обоснование: \_\_\_\_\_.

*Повышенный уровень сложности*

**54. Прочитайте задание и установите соответствие.**

Установите соответствие между типами задач машинного обучения и их описаниями.

Тип задачи:

- А) Кластеризация
- Б) Регрессия
- В) Бинарная классификация
- Г) Визуализация

Описание:

- 5) Группировка объектов без меток
- 6) Предсказание числового значения
- 7) Сопоставление объекта одному из двух классов
- 8) Представление данных в понятной форме на плоскости или в пространстве

Поле для ответа:

**55. Прочитайте задание и установите последовательность.**

Установите правильную последовательность этапов обучения нейронной сети.

- А) Расчёт ошибки
- Б) Обратное распространение ошибки
- В) Прямой проход
- Г) Обновление весов

Поле для ответа:

**56. Прочитайте задание и установите последовательность.**

Расположите этапы построения дерева решений в правильном порядке (слева – буквы, справа – правильная последовательность).

- 1) Выбор признака для разбиения
- 2) Определение критерия останова
- 3) Разбиение данных по выбранному признаку
- 4) Построение листа (конечного узла)

Поле для ответа:

**57. Прочитайте задание и установите последовательность.**

Расположите стадии агломеративной кластеризации в правильном порядке:

- 1) Объединение двух ближайших кластеров
- 2) Инициализация с каждым объектом как отдельным кластером
- 3) Повторение объединения до достижения нужного числа кластеров
- 4) Построение дендрограммы

Поле для ответа:

**58. Прочитайте задание и установите соответствие.**

Соотнесите понятия, связанные с обучением нейронных сетей (слева), с их описанием (справа).

- А) Эпоха
- Б) Пакет (batch)
- В) Градиентный спуск
- Г) Потери (loss)

- 1) Один проход по всем обучающим данным
- 2) Мера ошибки модели
- 3) Метод обновления весов
- 4) Набор примеров, обрабатываемый одновременно

Поле для ответа:

**59. Прочитайте задание и установите соответствие.**

Соотнесите методы повышения качества модели (слева), с их описанием (справа).

- А) Dropout
- Б) Аугментация
- В) K-fold кросс-валидация
- Г) Ensemble

- 1) Искусственное увеличение обучающего набора
- 2) Случайное отключение нейронов во время обучения
- 3) Объединение нескольких моделей
- 4) Разбиение данных на несколько частей для оценки модели

Поле для ответа:

**60. Прочитайте задание и установите соответствие.**

Соотнесите архитектуры нейронных сетей (слева), с их особенностями (справа).

- А) Полносвязная сеть
- Б) Сверточная сеть
- В) Рекуррентная сеть
- Г) Трансформер

- 1) Использует внутреннее внимание и работает параллельно



- 2) Используется для обработки изображений
- 3) Каждый нейрон соединён со всеми
- 4) Работает с последовательностями данных

Поле для ответа:

**61. Прочитайте задание и установите соответствие.**

Соотнесите метрики классификации (слева), с их определениями (справа).

- А) Точность (accuracy)
- Б) Полнота (recall)
- В) Точность по классу (precision)
- Г) F-мера

- 1) Доля правильно угаданных объектов среди всех
- 2) Среднее между полнотой и точностью
- 3) Доля верных положительных среди всех предсказанных положительных
- 4) Доля верных положительных среди всех истинных положительных

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**62. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответа.**

Какие типы признаков относятся к категориальным (дискретным)?

- 1) Бинарный
- 2) Количественный
- 3) Номинальный
- 4) Порядковый
- 5) Вещественный

Поле для ответа:

Обоснование: \_\_\_\_\_.

**63. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какая метрика показывает долю правильно классифицированных объектов?

- 1) Precision
- 2) Recall
- 3) Accuracy
- 4) Loss
- 5) Margin

Поле для ответа:

Обоснование: \_\_\_\_\_.

**64. Прочитайте текст и напишите обоснованный ответ.**

Какие задачи относятся к обучению без учителя, и в чём его основное отличие от обучения с учителем?

Поле для ответа:

**65. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какой алгоритм кластеризации требует заранее задать число кластеров?

- 1) DBSCAN
- 2) K-means
- 3) Агломеративная кластеризация
- 4) Метод ближайших соседей

Поле для ответа:

Обоснование: \_\_\_\_\_.

**66. Прочитайте текст и напишите обоснованный ответ.**

Для чего используется функция активации в нейронной сети?

Поле для ответа:

**67. Прочитайте текст и напишите обоснованный ответ.**

Зачем делят данные на обучающую и тестовую выборки?

Поле для ответа:

**ОПК-4 Способен применять на практике новые научные принципы и методы исследований**

**Задания закрытого типа**

*Базовый уровень сложности*

**68. Прочитайте задание, выберите правильный ответ.**

Какой тип задачи относится к обучению без учителя?

- 1) Бинарная классификация
- 2) Регрессия
- 3) Кластеризация
- 4) Ранжирование
- 5) Прогнозирование

Поле для ответа:

**69. Прочитайте задание, выберите правильный ответ.**

Какой критерий используется при обучении линейной регрессии?

- 1) Средняя абсолютная ошибка
- 2) Среднеквадратичная ошибка
- 3) Максимальная правдоподобность

4) Коэффициент детерминации

Поле для ответа:

**70. Прочитайте задание, выберите правильные ответы.**

Какие утверждения справедливы для нормализации признаков?

- 1) Min-Max нормализация масштабирует значения в интервал  $[0,1]$
- 2) Standard scaling приводит признак к среднему 1 и стандартному отклонению 0
- 3) Standard scaling позволяет использовать правило трёх сигм в случае нормального распределения
- 4) Нормализация всегда ухудшает качество модели

Поле для ответа:

**71. Прочитайте задание, выберите правильный ответ.**

Для чего нужна функция потерь в обучении модели?

- 1) Чтобы случайно выбирать ответы
- 2) Чтобы измерять ошибки модели
- 3) Чтобы увеличивать размер данных
- 4) Чтобы рисовать графики
- 5) Чтобы сортировать признаки

Поле для ответа:

**72. Прочитайте задание, выберите правильные ответы.**

Какие из приведенных метрик оценки качества относятся к классификации?

- 1) Accuracy
- 2) Recall
- 3) Gini Index
- 4) MAE (Mean Absolute Error)
- 5) MSE (Mean Squared Error)

Поле для ответа:

**73. Прочитайте задание, выберите правильный ответ.**

Как работает метод случайного леса?

- 1) Строит одно большое дерево
- 2) Обучает несколько деревьев и усредняет их предсказания
- 3) Ищет лучший признак один раз
- 4) Использует только один случайный признак

Поле для ответа:

**74. Прочитайте задание, выберите правильные ответы.**

В чём преимущества деревьев решений?

- 1) Интерпретируемость
- 2) Не склонны к переобучению

- 3) Простота реализации
- 4) Работают только с числовыми признаками

Поле для ответа:

**75. Прочитайте задание, выберите правильные ответы.**

Какие из перечисленных метрик относятся к метрикам расстояния?

- 1) Евклидово
- 2) Среднеквадратичная ошибка
- 3) Манхэттенское
- 4) Косинусное
- 5) Точность

Поле для ответа:

**76. Прочитайте задание, выберите правильный ответ.**

Какой из следующих терминов относится к этапу обучения нейронной сети?

- 1) Компиляция кода
- 2) Передача данных в интернет
- 3) Обратное распространение ошибки
- 4) Вывод изображения на экран

Поле для ответа:

*Повышенный уровень сложности*

**77. Прочитайте задание и установите соответствие.**

Установите соответствие между типами признаков и примерами значений.

Тип признака:

- А) Бинарный
- Б) Номинальный
- В) Порядковый
- Г) Количественный

Пример значения:

- 5) Уровень образования: начальное, среднее, высшее
- 6) Пол: 0 – муж., 1 – жен.
- 7) Цвет: красный, синий, зелёный
- 8) Рост в сантиметрах

Поле для ответа:

**78. Прочитайте задание и установите соответствие.**

Соотнесите тип метрики с её назначением:

- А) MSE
- Б) MAE
- В)  $R^2$  (коэффициент детерминации)

- 4) Показывает, насколько модель объясняет дисперсию целевой переменной
- 5) Оценивает среднеквадратичную ошибку
- 6) Оценивает среднюю абсолютную ошибку

**79. Прочитайте задание и установите последовательность.**

Расположите шаги оценки качества классификатора по матрице ошибок:

- 1) Построение confusion matrix (матрицы ошибок)
- 2) Подсчёт TP, FP, TN, FN
- 3) Вычисление метрик (Accuracy, Precision, Recall)
- 4) Интерпретация результатов

Поле для ответа:

**80. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Какое из расстояний измеряет длину пути, если двигаться только по осям координат (по прямоугольной сетке)?

- 1) Евклидово
- 2) Манхэттенское
- 3) Косинусное
- 4) Хэммингово

Поле для ответа:

Обоснование: \_\_\_\_\_.

**81. Прочитайте задание и установите последовательность.**

Установите правильную последовательность прохождения данных через нейрон.

- А) Умножение входных данных на веса
- Б) Суммирование полученных значений
- В) Применение функции активации
- Г) Передача результата на следующий слой

Поле для ответа:

**82. Прочитайте задание и установите последовательность.**

Расположите шаги алгоритма бустинга в правильном порядке.

- 1) Обучение базовой модели на взвешенных данных
- 2) Инициализация весов объектов
- 3) Обновление весов с учетом ошибок предыдущей модели
- 4) Комбинирование всех базовых моделей для итогового предсказания

Поле для ответа:

**83. Прочитайте задание и установите соответствие.**

Соотнесите этапы работы с данными (слева), с их описанием (справа).

- А) Сбор данных
- Б) Очистка данных
- В) Визуализация
- Г) Разделение на выборки

- 1) Удаление пропусков и выбросов
- 2) Сбор информации из различных источников
- 3) Создание графиков для анализа
- 4) Деление данных на обучающую и тестовую части

Поле для ответа:

### **Задания открытого типа**

*Высокий уровень сложности*

84. **Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответа.**

Какие из перечисленных задач относятся к обучению с учителем?

- 1) Кластеризация
- 2) Регрессия
- 3) Многоклассовая классификация
- 4) Снижение размерности
- 5) Визуализация

Поле для ответа:

Обоснование: \_\_\_\_\_.

85. **Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какой из перечисленных алгоритмов относится к методам кластеризации?

- 1) Линейная регрессия
- 2) K-means
- 3) DBSCAN
- 4) Метод опорных векторов

Поле для ответа:

Обоснование: \_\_\_\_\_.

86. **Прочитайте текст и напишите обоснованный ответ.**

Что такое дерево решений и для чего оно используется в машинном обучении?

Поле для ответа:

87. **Прочитайте текст и напишите обоснованный ответ.**

Что такое нейрон в контексте искусственной нейронной сети?

Поле для ответа:

**88. Прочитайте текст и напишите обоснованный ответ.**  
Что делает алгоритм К-ближайших соседей при классификации?

Поле для ответа:

**89. Прочитайте текст и напишите обоснованный ответ.**  
Чем отличается обучение без учителя от обучения с учителем?

Поле для ответа:

**ОПК-6 Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества**

**Задания закрытого типа**

*Базовый уровень сложности*

**90. Прочитайте задание, выберите правильный ответ.**

Какая функция ошибки используется для оценки качества модели в задаче регрессии?

- 1) Log-loss
- 2) Mean Squared Error (MSE)
- 3) Precision
- 4) Accuracy
- 5) AUC

Поле для ответа:

**91. Прочитайте задание, выберите правильный ответ.**

Для чего используется регуляризация в линейной регрессии?

- 1) Для увеличения сложности модели
- 2) Для улучшения качества на обучающей выборке
- 3) Для борьбы с переобучением
- 4) Для оценки значимости признаков
- 5) Для ускорения градиентного спуска

Поле для ответа:

**92. Прочитайте задание, выберите правильные ответы.**

Что происходит при переобучении модели линейной регрессии?

- 1) Ошибка на тестовой выборке ниже, чем на обучающей
- 2) Модель хорошо аппроксимирует как обучающие, так и тестовые данные
- 3) Весовые коэффициенты могут иметь большие значения
- 4) Качество на новых данных ухудшается

Поле для ответа:

**93. Прочитайте задание, выберите правильный ответ.**

Что такое функция сигмоиды в логистической регрессии?

- 1) Функция, строящая деревья решений
- 2) Функция, преобразующая значение в вероятность

- 3) Метод измерения расстояния
- 4) Правило голосования между моделями
- 5) Функция, удаляющая выбросы

Поле для ответа:

**94. Прочитайте задание, выберите правильные ответы.**

Какие действия выполняет метод опорных векторов (SVM)?

- 1) Строит разделяющую гиперплоскость
- 2) Максимизирует ширину разделяющей полосы
- 3) Кластеризует данные
- 4) Использует деревья решений
- 5) Обучается без меток классов

Поле для ответа:

**95. Прочитайте задание, выберите правильный ответ.**

Какой из следующих методов является бустингом?

- 1) KNN
- 2) Decision Tree
- 3) AdaBoost
- 4) SVM
- 5) PCA

Поле для ответа:

**96. Прочитайте задание, выберите правильные ответы.**

Какие методы относятся к ансамблям?

- 1) Random Forest
- 2) Decision Tree
- 3) AdaBoost
- 4) K-Means
- 5) PCA

Поле для ответа:

**97. Прочитайте задание, выберите правильные ответы.**

Какие методы кластеризации являются иерархическими?

- 1) K-means
- 2) Агломеративная кластеризация
- 3) DBSCAN
- 4) Дивизивная кластеризация
- 5) Линейная регрессия

Поле для ответа:



**98. Прочитайте задание, выберите правильный ответ.**

Что делает функция активации в нейронной сети?

- 1) Преобразует входные данные в изображение
- 2) Заменяет все значения на нули
- 3) Определяет выходное значение нейрона
- 4) Сохраняет модель на диск

Поле для ответа:

*Повышенный уровень сложности*

**99. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что означает термин «антиградиент» в контексте градиентного спуска?

- 1) Максимальное значение функции
- 2) Направление наискорейшего убывания функции
- 3) Набор весовых коэффициентов
- 4) Сумма всех значений признаков

Поле для ответа:

Обоснование: \_\_\_\_\_.

**100. Прочитайте задание и установите последовательность.**

Установите правильную последовательность этапов процесса анализа данных.

- 1) Очистка данных
- 2) Сбор данных
- 3) Построение модели
- 4) Оценка модели на тестовых данных

Поле для ответа:

**101. Прочитайте задание и установите соответствие.**

Соотнесите термин с его определением:

- |   |  |
|---|--|
| А) Регуляризация<br>Б) Градиентный спуск<br>В) Предобработка данных<br>Г) MAE | 5) Метод оптимизации параметров модели<br>6) Подготовка признаков перед обучением<br>7) Средняя абсолютная ошибка<br>8) Метод борьбы с переобучением через ограничение весов |
|---|--|

Поле для ответа:

**102. Прочитайте задание и установите соответствие.**

Соотнесите компоненты нейронной сети (слева) с их описанием (справа).

- А) Нейрон
- Б) Входной слой
- В) Скрытый слой
- Г) Выходной слой

- 1) Первый слой, принимающий данные
- 2) Элемент сети, выполняющий простые вычисления
- 3) Слой, передающий результаты наружу
- 4) Внутренний слой, обрабатывающий информацию

Поле для ответа:

**103. Прочитайте задание и установите последовательность.**

Установите порядок построения нейронной сети.

- А) Определение числа слоёв
- Б) Выбор функции активации
- В) Настройка параметров обучения
- Г) Запуск процесса обучения

Поле для ответа:

**104. Прочитайте задание и установите последовательность.**

Расположите этапы обучения нейронной сети в правильном порядке:

- 1) Вычисление потерь
- 2) Прямой проход
- 3) Обратное распространение ошибки
- 4) Обновление весов

Поле для ответа:

**105. Прочитайте задание и установите последовательность.**

Расположите этапы подготовки данных в правильном порядке:

- 1) Очистка данных
- 2) Сбор данных
- 3) Масштабирование признаков
- 4) Разделение на выборки

Поле для ответа:

**106. Прочитайте задание и установите последовательность.**

Расположите этапы обучения модели с учителем в правильном порядке:

- 1) Обучение модели на тренировочной выборке
- 2) Разделение данных
- 3) Оценка модели
- 4) Применение модели

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**107. Прочитайте текст и напишите обоснованный ответ.**

Что такое ансамбль моделей и какую основную пользу он приносит?

Поле для ответа:

**108. Прочитайте задание и установите соответствие.**

Соотнесите метрики расстояния с их описанием:

А) Евклидово

Б) Манхэттенское

В) Косинусное

Г) Хэммингово

5) Измеряет угол между векторами

6) Сумма абсолютных разностей координат

7) Расстояние "по прямой линии" между точками

8) Количество позиций с разными символами

Поле для ответа:

**109. Прочитайте текст и напишите обоснованный ответ.**

Зачем делят данные на обучающую и тестовую выборки?

Поле для ответа:

**110. Прочитайте текст и напишите обоснованный ответ.**

Для чего нужна проверочная (валидационная) выборка?

Поле для ответа:

**111. Прочитайте текст и напишите обоснованный ответ.**

Что такое градиентный спуск и какова его роль в обучении нейронной сети?

Поле для ответа:

**112. Прочитайте текст и напишите обоснованный ответ.**

Зачем нормализуют или масштабируют данные перед обучением модели?

Поле для ответа:

**ОПК-7 Способен использовать методы научных исследований и математического моделирования в области проектирования и управления информационными системами**

**Задания закрытого типа**

*Базовый уровень сложности*

**113. Прочитайте задание, выберите правильные ответы.**

Какие утверждения верны для признаков в машинном обучении?

- 1) Бинарные признаки принимают только значения 0 и 1
- 2) Порядковые признаки не имеют смысла в обучении
- 3) Количественные признаки могут принимать любые вещественные значения
- 4) Номинальные признаки всегда упорядочены

Поле для ответа:

**114. Прочитайте задание, выберите правильный ответ.**

Что такое свободный коэффициент в линейной регрессии?

- 1) Коэффициент перед самым важным признаком
- 2) Произвольное число в формуле
- 3) Смещение (bias), добавляемое ко всем предсказаниям
- 4) Среднее значение целевой переменной

Поле для ответа:

**115. Прочитайте задание, выберите правильные ответы.**

Какие подходы используются для оценки качества модели линейной регрессии?

- 1) Кросс-валидация
- 2) Регуляризация
- 3) Разделение на обучающую и тестовую выборки
- 4) Выбор функции активации

Поле для ответа:

**116. Прочитайте задание, выберите правильные ответы.**

Что делает сигмоидная функция в логистической регрессии?

- 1) Преобразует любое число в значение от 0 до 1
- 2) Делит данные на кластеры
- 3) Помогает вычислить вероятность класса
- 4) Строит дерево решений
- 5) Сортирует признаки по важности

Поле для ответа:

**117. Прочитайте задание, выберите правильный ответ.**

Какой подход используется при построении деревьев в градиентном бустинге?

- 1) Каждое дерево обучается независимо
- 2) Каждое дерево обучается на тех же данных
- 3) Новое дерево обучается на ошибках предыдущих
- 4) Используются только признаки с наименьшей дисперсией

Поле для ответа:

**118. Прочитайте задание, выберите правильные ответы.**

Что характерно для метода бэггинга?

- 1) Использование одного дерева
- 2) Bootstrap-подвыборки
- 3) Случайные признаки в каждом узле
- 4) Усреднение предсказаний
- 5) Последовательное обучение моделей

Поле для ответа:

**119. Прочитайте задание, выберите правильные ответы.**

Какие из перечисленных утверждений верны про DBSCAN?

- 1) Позволяет находить кластеры произвольной формы
- 2) Требуется заранее задать число кластеров
- 3) Основан на плотности точек
- 4) Использует центроиды для кластеров

Поле для ответа:

**120. Прочитайте задание, выберите правильный ответ.**

Что такое нейрон в искусственной нейронной сети?

- 1) Блок, хранящий все данные
- 2) Единица, выполняющая простые вычисления
- 3) Программа для визуализации данных
- 4) Метод сортировки информации

Поле для ответа:

*Повышенный уровень сложности*

**121. Прочитайте задание, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа.**

Что такое переобучение (overfitting) в модели линейной регрессии?

- 1) Модель показывает хорошие результаты на новых данных
- 2) Модель плохо обучается на тренировочных данных
- 3) Модель слишком точно подстроилась под обучающую выборку и плохо работает на новых данных
- 4) Модель не использует все признаки

Поле для ответа:

Обоснование: \_\_\_\_\_.

**122. Прочитайте задание и установите последовательность.**

Расположите шаги градиентного спуска в правильном порядке:

- 1) Проверка условия останова
- 2) Выбор начальной точки
- 3) Обновление весов по антиградиенту
- 4) Вычисление градиента

Поле для ответа:

**123. Прочитайте задание и установите соответствие.**

Установите соответствие между методами и их описанием:

- А) Логистическая регрессия
- Б) Метод опорных векторов (SVM)
- В) One-vs-all
- Г) One-vs-one

- 1) Строит K бинарных классификаторов, каждый отличает один класс от остальных
- 2) Использует полосу максимального зазора между классами
- 3) Преобразует скалярное значение в вероятность от 0 до 1
- 4) Обучает классификаторы для каждой пары классов

Поле для ответа:

**124. Прочитайте задание и установите соответствие.**

Соотнесите термины, связанные с обучением нейронной сети (слева), с их значениями (справа).

- А) Обучающая выборка
- Б) Эпоха
- В) Ошибка
- Г) Вес

- 1) Один проход по всем обучающим данным
- 2) Разность между предсказанием и правильным ответом
- 3) Данные, на которых обучается сеть
- 4) Число, регулирующее силу связи между нейронами

Поле для ответа:

**125. Прочитайте задание и установите последовательность.**

Установите порядок обработки изображения в сверточной нейронной сети.

- А) Свертка
- Б) Функция активации
- В) Пулинг
- Г) Полносвязный слой

Поле для ответа:

**126. Прочитайте задание и установите последовательность.**

Расположите этапы работы алгоритма градиентного спуска в правильном порядке:

- 1) Вычисление градиента
- 2) Выбор начальных весов
- 3) Обновление весов
- 4) Проверка условия остановки

Поле для ответа:

**127. Прочитайте задание и установите последовательность.**

Расположите этапы работы сверточной нейронной сети в правильном порядке:

- 1) Свертка
- 2) Подвыборка (пулинг)
- 3) Полносвязный слой
- 4) Выход модели

Поле для ответа:

**128. Прочитайте задание и установите последовательность.**

Расположите этапы работы обучения с подкреплением в правильном порядке:

- 1) Выполнение действия агентом
- 2) Получение состояния среды
- 3) Получение награды
- 4) Обновление стратегии

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**129. Прочитайте текст и напишите обоснованный ответ.**

Что такое выбросы в данных?

Поле для ответа:

**130. Прочитайте текст и напишите обоснованный ответ.**

Что происходит на этапе прямого прохода в нейронной сети?

Поле для ответа:

**131. Прочитайте текст и напишите обоснованный ответ.**

Что такое эпоха в обучении нейронной сети?

Поле для ответа:

**132. Прочитайте текст и напишите обоснованный ответ.**

Для чего используется метод K-means?

Поле для ответа:

**133. Прочитайте текст и напишите обоснованный ответ.**

Что делает слой свёртки в сверточной нейронной сети?

Поле для ответа:

**134. Прочитайте текст и напишите обоснованный ответ.**

Что такое функция потерь и почему она важна?

Поле для ответа:

**ПК-6 Способен применить технологии машинного обучения к реальным общественным задачам**

**Задания закрытого типа**

*Базовый уровень сложности*

**135. Прочитайте задание, выберите правильный ответ.**

Какой тип признаков предполагает упорядоченное конечное множество значений?

- 1) Бинарные
- 2) Номинальные
- 3) Порядковые
- 4) Количественные
- 5) Категориальные

Поле для ответа:

**136. Прочитайте задание, выберите правильные ответы.**

Какие задачи относятся к обучению с учителем?

- 1) Кластеризация
- 2) Регрессия
- 3) Бинарная классификация
- 4) Визуализация

Поле для ответа:

**137. Прочитайте задание, выберите правильный ответ.**

Что делает метод one-hot encoding?

- 1) Удаляет ненужные признаки
- 2) Преобразует числовые признаки в категориальные
- 3) Кодировывает категориальные признаки в набор бинарных
- 4) Удаляет строки с пропущенными значениями

Поле для ответа:

**138. Прочитайте задание, выберите правильный ответ.**

Что является целью линейной классификации?

- 1) Построение дерева решений



- 2) Разделение объектов на классы с помощью прямой (гиперплоскости)
- 3) Уменьшение размерности признаков
- 4) Построение нейронной сети
- 5) Максимизация дисперсии признаков

Поле для ответа:

**139. Прочитайте задание, выберите правильные ответы.**

Какие задачи можно решать с помощью деревьев решений?

- 1) Классификация
- 2) Регрессия
- 3) Кластеризация
- 4) Поиск ближайших соседей
- 5) Детектирование выбросов

Поле для ответа:

**140. Прочитайте задание, выберите правильные ответы.**

Какие критерии используются в алгоритме ID3 для выбора признака?

- 1) Длина названия признака
- 2) Случайный выбор
- 3) Энтропия
- 4) Прирост информации (Gain)
- 5) Корреляция с целевой переменной

Поле для ответа:

**141. Прочитайте задание, выберите правильные ответы.**

Какие из этих характеристик описывают метрики качества кластеризации?

- 1) Количество признаков
- 2) Разделимость кластеров
- 3) Время обучения модели
- 4) Компактность кластеров

Поле для ответа:

**142. Прочитайте задание, выберите правильные ответы.**

Какие из следующих утверждений верны для метода K-means?

- 1) Использует центры кластеров (центроиды)
- 2) Может находить кластеры произвольной формы
- 3) Требуется заранее задать число кластеров
- 4) Основан на плотности точек
- 5) Объединяет кластеры иерархически

Поле для ответа:

*Повышенный уровень сложности*

**143. Прочитайте задание и установите последовательность.**

Расположите этапы оценки качества модели линейной регрессии:

- 1) Разделение выборки на тренировочную и тестовую
- 2) Обучение модели на тренировочной выборке
- 3) Применение модели к тестовой выборке
- 4) Подсчёт метрики качества (например, MSE)

Поле для ответа:

**144. Прочитайте задание и установите соответствие.**

Установите соответствие между метриками и их назначением:

- А) Accuracy
- Б) Precision
- В) Recall
- Г) F1-мера

- 5) Среднее между точностью и полнотой
- 6) Доля правильно классифицированных объектов
- 7) Доля верно угаданных среди всех угаданных
- 8) Доля верно угаданных среди всех объектов нужного класса

Поле для ответа:

**145. Прочитайте задание и установите соответствие.**

Соотнесите функции активации (слева) с их особенностями (справа).

- А) ReLU
- Б) Сигмоида
- В) tanh
- Г) Линейная функция

- 1) Возвращает значение от 0 до 1
- 2) Возвращает максимум между 0 и входным значением
- 3) Использует гиперболический тангенс
- 4) Возвращает входное значение без изменений

Поле для ответа:

**146. Прочитайте задание и установите последовательность.**

Установите порядок работы рекуррентной нейронной сети при анализе последовательности.

- А) Получение текущего входа
- Б) Учет предыдущего состояния
- В) Вычисление нового состояния
- Г) Вывод результата

Поле для ответа:

**147. Прочитайте задание и установите последовательность.**

Расположите этапы обработки текста перед подачей в модель в правильном порядке:

- 1) Токенизация
- 2) Приведение к нижнему регистру
- 3) Удаление стоп-слов
- 4) Векторизация

Поле для ответа:

**148. Прочитайте задание и установите последовательность.**

Расположите этапы построения дерева решений в правильном порядке:

- 1) Выбор признака для разбиения
- 2) Деление выборки
- 3) Построение поддеревьев
- 4) Определение выходного значения в листьях

Поле для ответа:

**149. Прочитайте задание и установите последовательность.**

Расположите этапы процесса кросс-валидации в правильном порядке:

- 1) Деление данных на K блоков
- 2) Обучение модели на K-1 блоках
- 3) Тестирование на оставшемся блоке
- 4) Усреднение результатов

Поле для ответа:

**150. Прочитайте задание и установите последовательность.**

Расположите этапы работы модели K-ближайших соседей (KNN) в правильном порядке:

- 1) Расчёт расстояний до всех объектов
- 2) Выбор K ближайших соседей
- 3) Подсчёт голосов классов
- 4) Присвоение класса объекту

Поле для ответа:

**Задания открытого типа**

*Высокий уровень сложности*

**151. Прочитайте задание, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов.**

Какие из следующих способов помогают бороться с переобучением?

- 1) Увеличение количества признаков
- 2) Регуляризация
- 3) Удаление выбросов
- 4) Использование более сложной модели

Поле для ответа:

Обоснование: \_\_\_\_\_.

**152. Прочитайте текст и напишите обоснованный ответ.**

Зачем нужно делить данные на обучающую и тестовую выборки?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**153. Прочитайте текст и напишите обоснованный ответ.**

Почему важно нормализовать входные данные при обучении нейронной сети?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**154. Прочитайте текст и напишите обоснованный ответ.**

Как работает обучение с подкреплением?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**155. Прочитайте текст и напишите обоснованный ответ.**

Что такое точность (ассигасу) в задачах классификации?

Поле для ответа:

Обоснование: \_\_\_\_\_.

**156. Прочитайте текст и напишите обоснованный ответ.**

Что такое признаки (features) в машинном обучении?

Поле для ответа:

Обоснование: \_\_\_\_\_.

## 5 Средства оценки индикаторов достижения компетенций

Таблица 4

<b>Средства оценки индикаторов достижения компетенций</b>		
<b>Коды компетенций</b>	<b>Индикаторы компетенций (в соот.с Таблицей 1)</b>	<b>Средства оценки (в соот. с Таблицами 5, 7)</b>
УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	Практическое задание, тест
УК-6	ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4.	Практическое задание, тест

<b>Коды компетенций</b>	<b>Индикаторы компетенций</b> (в соот.с Таблицей 1)	<b>Средства оценки</b> (в соот. с Таблицами 5, 7)
ОПК-1	ИД.ОПК-1.1. ИД.ОПК-1.2.	Практическое задание, тест
ОПК-4	ИД.ОПК-4.1. ИД.ОПК-4.2. ИД.ОПК-4.3. ИД.ОПК-4.4.	Практическое задание, тест
ОПК-6	ИД.ОПК-6.1. ИД.ОПК-6.2. ИД.ОПК-6.3. ИД.ОПК-6.4.	Практическое задание, тест
ОПК-7	ИД.ОПК-7.1. ИД.ОПК-7.2. ИД.ОПК-7.3.	Практическое задание, тест
ПК-6	ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	Практическое задание, тест

Таблица 5

### Описание средств оценки индикаторов достижения компетенций

<b>Средства оценки</b> (в соот. С Таблицами 5, 7)	<b>Рекомендованный план выполнения работы</b>
Практическое задание	<p>Магистрант в ходе подготовки и выполнения практических заданий показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> <li>1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</li> <li>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</li> <li>3. Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</li> <li>4. На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</li> <li>5. Обосновывать актуальность постановки целей и задач исследования современных проблем и методов прикладной информатики и развития информационного общества, формулировать задачи и гипотезы для поиска вариантов решения современных проблем и методов прикладной информатики, на основе анализа направлений развития информационного общества разрабатывать предложения и рекомендации для решения современных проблем прикладной информатики и развития информационного общества</li> <li>6. Обосновывать актуальность выбора определенных методов научных исследований, анализировать направления и методы научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</li> <li>7. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных</li> </ol>

Средства оценки (в соот. С Таблицами 5, 7)	Рекомендованный план выполнения работы
	задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения
Тест	<p>Магистрант в ходе подготовки и выполнения тестов показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> <li>1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</li> <li>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</li> <li>3. Делать обоснованный выбор методов разработки алгоритмов и программных средств для решения профессиональных задач, в том числе с использованием современных информационных технологий</li> <li>4. На основе современных теорий и концепций обосновывать актуальность, анализировать, применять и разрабатывать предложения и рекомендации по использованию новых научных принципов и методов исследований в профессиональной области знаний</li> <li>5. Обосновывать актуальность постановки целей и задач исследования современных проблем и методов прикладной информатики и развития информационного общества, формулировать задачи и гипотезы для поиска вариантов решения современных проблем и методов прикладной информатики, на основе анализа направлений развития информационного общества разрабатывать предложения и рекомендации для решения современных проблем прикладной информатики и развития информационного общества</li> <li>6. Обосновывать актуальность выбора определенных методов научных исследований, анализировать направления и методы научных исследований и математического моделирования в области проектирования и управления информационными системами в различных областях</li> <li>7. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения</li> </ol>