



Федеральное государственное бюджетное образовательное учреждение
высшего образования

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
АРХИТЕКТУРНО-СТРОИТЕЛЬНЫЙ УНИВЕРСИТЕТ

Кафедра Информатики

УТВЕРЖДАЮ
Начальник учебно-методического управления

«31» октября 2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Введение в машинное обучение

направление подготовки/специальность 38.03.05 Бизнес-информатика

направленность (профиль)/специализация образовательной программы Бизнес-аналитика

Форма обучения очная

Санкт-Петербург, 2024

1. Цели и задачи освоения дисциплины (модуля)

Целью освоения дисциплины «Введение в машинное обучение» является изучение основных моделей и методов машинного обучения, развитие навыков применения современных информационно-коммуникационных технологий в профессиональной деятельности при решении прикладных задач в области бизнес-аналитики.

Задачи дисциплины:

- ознакомление обучающихся с основными методами интеллектуального анализа данных, основными методами решения задач машинного обучения;
- изучение методов и алгоритмов поиска классификации, кластеризации, регрессии, используемых при решении прикладных задач;
- обучение навыкам решения прикладных задач в области бизнес-аналитики с применением методов машинного обучения.

2. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов освоения ОПОП
ОПК-3 Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации	ОПК-3.1 Осуществляет управление процессами разработки и использования продуктов и услуг в сфере информационно-коммуникационных технологий	знает – основные классы задач машинного обучения; умеет – свести прикладную бизнес-задачу к известному типу, формализовать её и построить математическую модель; владеет – навыками решения основных классов задач машинного обучения; – навыками осуществления управления проектом по построению комплексных систем на основе аналитики больших данных.
ОПК-3 Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации	ОПК-3.2 Применяет алгоритмы, языки программирования и управления базами данных, современные программные среды разработки информационных систем для решения прикладных бизнес-задач	знает – классы методов и алгоритмов машинного обучения; умеет – применять различные методы машинного обучения для решения прикладных бизнес-задач, а также оценивать полученный результат; владеет – навыками решения практических задач с использованием методов машинного обучения.
ОПК-3 Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации	ОПК-3.3 Разрабатывает алгоритмы и программный код для практического применения в информационных системах управления бизнесом	знает – основные библиотеки, применяемые в решении задач из области машинного обучения; умеет – создавать эффективные программы, реализующие алгоритмы машинного обучения, для практического применения в информационных системах управления бизнесом.

		владеет – навыками разработки алгоритмов и программ для решения комплекса задач предметной области с применением методов машинного обучения.
--	--	--

3. Указание места дисциплины (модуля) в структуре образовательной программы

Данная дисциплина (модуль) включена в Блок «Дисциплины, модули» Б1.О.29 основной профессиональной образовательной программы 38.03.05 Бизнес-информатика и относится к обязательной части учебного плана.

№ п/п	Предшествующие дисциплины	Код и наименование индикатора достижения компетенции
1	Высшая математика	УК-2.1, УК-2.3, УК-2.4
2	Анализ данных на языке Python	ОПК-4.1, ОПК-4.2, ОПК-4.3
3	Основы программирования	ОПК-3.1, ОПК-3.2, ОПК-3.3

Успешное освоение данной дисциплины базируется на знаниях, умениях и навыках, полученных обучающимися при изучении предшествующих дисциплин.

№ п/п	Последующие дисциплины	Код и наименование индикатора достижения компетенции
1	Информационно-аналитическая деятельность поддержки принятия решений	ОПК-4.1, ОПК-4.2, ОПК-4.3
2	Информационные технологии финансового и экономического анализа	ПК-2.3, ПК-4.1, ПК-1.1, ПК-1.2

4. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Вид учебной работы	Всего часов	Из них часы на практическую подготовку	Семестр
			6
Контактная работа	48		48
Практические занятия (Пр)	48	0	48
Иная контактная работа, в том числе:			
консультации по курсовой работе (проекту), контрольным работам (РГР)			
контактная работа на аттестацию (сдача зачета, зачета с оценкой; защита курсовой работы (проекта); сдача контрольных работ (РГР))			
контактная работа на аттестацию в сессию (консультация перед экзаменом и сдача			
Часы на контроль	4		4
Самостоятельная работа (СР)	56		56
Общая трудоемкость дисциплины (модуля)			
часы:	108		108
зачетные единицы:	3		3

5. Содержание дисциплины (модуля), структурированное по разделам (темам) с указанием отведенного на них количества академических часов и видов учебных занятий

5.1. Тематический план дисциплины (модуля)

№	Разделы дисциплины	Семестр	Контактная работа (по учебным занятиям), час.						СР	Всего, час.	Код индикатора достижения компетенции
			лекции		ПЗ		ЛР				
			всего	из них на практическую подготовку	всего	из них на практическую подготовку	всего	из них на практическую подготовку			
1.	1 раздел. Базовые алгоритмы										
1.1.	Введение в анализ данных и машинное обучение	6			2				2	4	ОПК-3.1, ОПК-3.2, ОПК-3.3
1.2.	Линейная и логистическая регрессия, регуляризация	6			6				5	11	ОПК-3.1, ОПК-3.2, ОПК-3.3
1.3.	Методы классификации в машинном обучении	6			6				5	11	ОПК-3.1, ОПК-3.2, ОПК-3.3
1.4.	Кластеризация	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
1.5.	Методы понижения размерности	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
1.6.	Деревья и градиентный бустинг	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
2.	2 раздел. Глубокое обучение										
2.1.	Введение в Pytorch	6			2				4	6	ОПК-3.1, ОПК-3.2, ОПК-3.3
2.2.	Введение в машинное зрение	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
2.3.	Введение в обработку текста	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
2.4.	Обработка временных рядов	6			4				5	9	ОПК-3.1, ОПК-3.2, ОПК-3.3
2.5.	Обучение с подкреплением	6			8				10	18	ОПК-3.1, ОПК-3.2, ОПК-3.3
3.	3 раздел. Контроль										
3.1.	Зачет	6								4	ОПК-3.1, ОПК-3.2, ОПК-3.3

5.1. Практические занятия

№ разд	Наименование раздела и темы практических занятий	Наименование и содержание практических занятий
1	Введение в анализ данных и машинное обучение	Введение в анализ данных и машинное обучение Машинное обучение и науки о данных. История машинного обучения. Общая схема решения задач машинного обучения. Типы задач машинного обучения. Примеры прикладных задач. Основные понятия и определения.
2	Линейная и логистическая регрессия, регуляризация	Линейная и логистическая регрессия, регуляризация Обучение с учителем. Методы обучения с учителем. Регрессионные модели. Постановка задачи обучения регрессионной модели. Риск модели с квадратичной функцией потерь. Функция регрессии. Оптимальность регрессионных моделей. Линейная и логистическая регрессия. Регуляризованные регрессионные модели. Гребневая регрессия. LASSO-регрессия. Пошаговая регрессия.
3	Методы классификации в машинном обучении	Методы классификации в машинном обучении Постановка задачи классификации. Бинарная классификация. Функция потерь и эмпирический риск классификатора. Виды функций потерь. Метод опорных векторов (SVM). Наивный байесовский классификатор.
4	Кластеризация	Кластеризация Понятие машинного обучения без учителя. Виды и задачи обучения без учителя. Примеры практических задач. Кластерный анализ данных: цели, история развития, постановка задачи. Практические применения кластерного анализа. Подходы к кластеризации данных. Виды кластерного анализа. Общая схема решения задачи кластеризации. Метод К-средних. Алгоритм K-means++. Сходимость метода К-средних. Модификации метода. Визуализация результатов кластеризации. Методы расчета расстояния между кластерами.
5	Методы понижения размерности	Методы понижения размерности Постановка задачи сокращения размерности данных. Подходы к сокращению размерности. Метод главных компонент. Выбор числа главных компонент.
6	Деревья и градиентный бустинг	Деревья и градиентный бустинг Понятие дерева решений. Регрессионные и классификационные деревья. Типы вершин. Обучение деревьев. Обучение ансамблей. Подходы к обучению ансамблей: бэггинг, бустинг, стеккинг. Случайный лес. Техники бустинга моделей машинного обучения. Градиентный бустинг деревьев решений.
7	Введение в Pytorch	Введение в Pytorch. Нейронные сети Математическая модель искусственного нейрона, активационные характеристики нейронов, нейронные сети прямого распространения, математическая модель многослойной нейронной сети. Постановка задачи обучения многослойной нейронной сети. Простой градиентный метод. Виды функций потерь. Метод обратного распространения ошибки. Сеть обратного распространения.
8	Введение в машинное зрение	Введение в машинное зрение Математическая модель сверточной нейронной сети. Виды слоев сети. Постановка задачи обучения сверточной нейронной сети. Применение сверточных нейронных сетей для решения задач

		классификации, регрессии и трансформации изображений.
9	Введение в обработку текста	Введение в обработку текста Задачи обработки естественного языка (NLP). Общая архитектура NLP-системы. Уровни обработки естественного языка. Морфологический анализ текста. Синтаксический анализ текста. Виды синтаксических деревьев. Лексико-семантический анализ.
10	Обработка временных рядов	Обработка временных рядов Нейронные сети для прогнозирования временных рядов. Рекуррентные нейронные сети. Нейронные сети с долгой краткосрочной памятью.
11	Обучение с подкреплением	Обучение с подкреплением на нейронных сетях Решение задач с применением глубоких нейронных сетей на реальных наборах данных.

5.2. Самостоятельная работа обучающихся

№ разд	Наименование раздела дисциплины и темы	Содержание самостоятельной работы
1	Введение в анализ данных и машинное обучение	Введение в анализ данных и машинное обучение Изучение материала; подготовка к практическим занятиям и тестированию.
2	Линейная и логистическая регрессия, регуляризация	Линейная и логистическая регрессия, регуляризация Изучение материала; подготовка к практическим занятиям и тестированию.
3	Методы классификации в машинном обучении	Методы классификации в машинном обучении Изучение материала; подготовка к практическим занятиям и тестированию.
4	Кластеризация	Кластеризация Изучение материала; подготовка к практическим занятиям и тестированию.
5	Методы понижения размерности	Методы понижения размерности Изучение материала; подготовка к практическим занятиям и тестированию.
6	Деревья и градиентный бустинг	Деревья и градиентный бустинг Изучение материала; подготовка к практическим занятиям и тестированию.
7	Введение в Pytorch	Введение в Pytorch Изучение материала; подготовка к практическим занятиям и тестированию.
8	Введение в машинное зрение	Введение в машинное зрение Изучение материала; подготовка к практическим занятиям и тестированию.
9	Введение в обработку текста	Введение в обработку текста Изучение материала; подготовка к практическим занятиям и тестированию.
10	Обработка временных рядов	Обработка временных рядов Изучение материала; подготовка к практическим занятиям и тестированию.
11	Обучение с подкреплением	Обучение с подкреплением на нейронных сетях Изучение материала; подготовка к практическим занятиям и тестированию.

6. Методические материалы для самостоятельной работы обучающихся по дисциплине (модулю)

Программой дисциплины предусмотрено проведение практических занятий, на которых дается основной систематизированный материал, а также практические задания, предполагающие закрепление изученного материала и формирование у обучающихся необходимых знаний, умений и навыков. Кроме того, важнейшим этапом изучения дисциплины является самостоятельная работа обучающихся с использованием всех средств и возможностей современных образовательных технологий.

В объем самостоятельной работы по дисциплине включается следующее:

- изучение теоретических вопросов по всем темам дисциплины;
- подготовка к практическим занятиям;
- подготовка к текущему контролю успеваемости обучающихся;
- подготовка к зачету.

Залогом успешного освоения дисциплины является обязательное посещение практических занятий, так как пропуск одного (тем более, нескольких) занятий может осложнить освоение разделов курса. На практических занятиях материал, изложенный на лекциях, закрепляется при выполнении практических заданий.

Приступая к изучению дисциплины, необходимо в первую очередь ознакомиться с содержанием РПД, а также методическими указаниями по организации самостоятельной работы.

При подготовке к практическим занятиям и в рамках самостоятельной работы по изучению дисциплины обучающимся необходимо:

- ознакомиться с соответствующей темой занятия;
- осмыслить круг изучаемых вопросов и логику их рассмотрения;
- изучить рекомендуемую рабочей программой литературу по данной теме.
- при самостоятельном изучении теоретической темы сделать конспект, используя

рекомендованные в РПД источники;

- выполнить практические задания в рамках изучаемой темы;
- ответить на контрольные вопросы по теме, используя материалы ФОС;
- подготовиться к проверочной работе, предусмотренной в контрольных точках;
- подготовиться к промежуточной аттестации.

Работы, выполняемые на практических занятиях, сдаются только лично на занятиях преподавателю, который ведет группу. Задания, выполняемые на компьютере, студенты сначала показывают только в электронном виде в соответствующих программах. При необходимости, при преподавателе доделывают или исправляют ошибки. Если требуется распечатать выполненные работы и сдать их в бумажном виде, преподаватель говорит об этом на занятиях.

Итогом изучения дисциплины является зачет. Зачет проводится по расписанию. Форма проведения занятия может быть устная, письменная и в электронном виде. Обучающиеся, не прошедшие аттестацию, должны ликвидировать задолженность в установленном порядке.

7. Оценочные материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине (модулю)

7.1. Перечень компетенций с указанием этапов их формирования в процессе освоения дисциплины

№ п/п	Контролируемые разделы дисциплины (модуля)	Код и наименование индикатора контролируемой компетенции	Вид оценочного средства
1	Введение в анализ данных и машинное обучение	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
2	Линейная и логистическая регрессия, регуляризация	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические занятия, тесты
3	Методы классификации в машинном обучении	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
4	Кластеризация	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты

5	Методы понижения размерности	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
6	Деревья и градиентный бустинг	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
7	Введение в Pytorch	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
8	Введение в машинное зрение	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические занятия, тесты
9	Введение в обработку текста	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
10	Обработка временных рядов	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
11	Обучение с подкреплением	ОПК-3.1, ОПК-3.2, ОПК-3.3	Практические задания, тесты
12	Зачет	ОПК-3.1, ОПК-3.2, ОПК-3.3	

7.2. Типовые контрольные задания или иные материалы текущего контроля успеваемости, необходимые для оценки знаний, умений и навыков и (или) опыта профессиональной деятельности, характеризующих этапы формирования компетенций в процессе освоения дисциплины

Комплект заданий для проверки сформированности индикаторов достижений компетенций ОПК-3.1, ОПК-3.2, ОПК-3.3

Тестовые и контрольные задания размещены: ЭИОС / СДО СПбГАСУ Moodle / Кафедры / Информатики / БАКАЛАВРИАТ и СПЕЦИАЛИТЕТ - кафедра информатики / Введение в машинное обучение (<https://moodle.spbgasu.ru/course/view.php?id=3051>), а также в Приложении.

7.3. Система оценивания результатов обучения по дисциплине (модулю) при проведении текущего контроля успеваемости

<p>Оценка «отлично» (зачтено)</p>	<p>знания:</p> <ul style="list-style-type: none"> - систематизированные, глубокие и полные знания по всем разделам дисциплины, а также по основным вопросам, выходящим за пределы учебной программы; - точное использование научной терминологии, систематически грамотное и логически правильное изложение ответа на вопросы; - полное и глубокое усвоение основной и дополнительной литературы, рекомендованной рабочей программой по дисциплине (модулю) <p>умения:</p> <ul style="list-style-type: none"> - умеет ориентироваться в теориях, концепциях и направлениях дисциплины и давать им критическую оценку, используя научные достижения других дисциплин <p>навыки:</p> <ul style="list-style-type: none"> - высокий уровень сформированности заявленных в рабочей программе компетенций; - владеет навыками самостоятельно и творчески решать сложные проблемы и нестандартные ситуации; - применяет теоретические знания для выбора методики выполнения заданий; - грамотно обосновывает ход решения задач; - безупречно владеет инструментарием учебной дисциплины, умение его эффективно использовать в постановке научных и практических задач; - творческая самостоятельная работа на практических/семинарских/лабораторных занятиях, активно участвует в групповых обсуждениях, высокий уровень культуры исполнения заданий
-----------------------------------	---

<p>Оценка «хорошо» (зачтено)</p>	<p>знания: - достаточно полные и систематизированные знания по дисциплине; - усвоение основной и дополнительной литературы, рекомендованной рабочей программой по дисциплине (модулю) умения: - умеет ориентироваться в основных теориях, концепциях и направлениях дисциплины и давать им критическую оценку; - использует научную терминологию, лингвистически и логически правильно излагает ответы на вопросы, умеет делать обоснованные выводы; - владеет инструментарием по дисциплине, умение его использовать в постановке и решении научных и профессиональных задач навыки: - самостоятельная работа на практических занятиях, участие в групповых обсуждениях, высокий уровень культуры исполнения заданий; - средний уровень сформированности заявленных в рабочей программе компетенций; - без затруднений выбирает стандартную методику выполнения заданий; - обосновывает ход решения задач без затруднений</p>
<p>Оценка «удовлетворительно» (зачтено)</p>	<p>знания: - достаточный минимальный объем знаний по дисциплине; - усвоение основной литературы, рекомендованной рабочей программой; - использование научной терминологии, стилистическое и логическое изложение ответа на вопросы, умение делать выводы без существенных ошибок умения: - умеет ориентироваться в основных теориях, концепциях и направлениях по дисциплине и давать им оценку; - владеет инструментарием учебной дисциплины, умение его использовать в решении типовых задач; - умеет под руководством преподавателя решать стандартные задачи навыки: - работа под руководством преподавателя на практических занятиях, допустимый уровень культуры исполнения заданий; - достаточный минимальный уровень сформированности заявленных в рабочей программе компетенций; - испытывает затруднения при обосновании алгоритма выполнения заданий</p>
<p>Оценка «неудовлетворительно» (не зачтено)</p>	<p>знания: - фрагментарные знания по дисциплине; - отказ от ответа (выполнения письменной работы); - знание отдельных источников, рекомендованных рабочей программой по дисциплине; умения: - не умеет использовать научную терминологию; - наличие грубых ошибок навыки: - низкий уровень культуры исполнения заданий; - низкий уровень сформированности заявленных в рабочей программе компетенций; - отсутствие навыков самостоятельной работы; - не может обосновать алгоритм выполнения заданий</p>

7.4. Теоретические вопросы и практические задания для проведения промежуточной аттестации обучающихся, необходимые для оценки знаний, умений и навыков и (или) опыта профессиональной деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

7.4.1. Теоретические вопросы для проведения промежуточной аттестации обучающихся

Примерные вопросы

1. Машинное обучение и науки о данных. История машинного обучения.
2. Общая схема решения задач машинного обучения. Типы задач машинного обучения.
3. Примеры прикладных задач. Основные понятия и определения.
4. Обучение с учителем. Методы обучения с учителем.
5. Регрессионные модели. Постановка задачи обучения регрессионной модели. Риск модели с квадратичной функцией потерь.
6. Функция регрессии. Оптимальность регрессионных моделей.
7. Линейная и логистическая регрессия.
8. Регуляризованные регрессионные модели. Гребневая регрессия. LASSO-регрессия.

Пошаговая регрессия.

9. Постановка задачи классификации.
10. Бинарная классификация.
11. Функция потерь и эмпирический риск классификатора. Виды функций потерь.
12. Метод опорных векторов (SVM).
13. Наивный байесовский классификатор.
14. Понятие машинного обучения без учителя. Виды и задачи обучения без учителя. Примеры практических задач.
15. Кластерный анализ данных: цели, история развития, постановка задачи. Практические применения кластерного анализа.
16. Подходы к кластеризации данных. Виды кластерного анализа. Общая схема решения задачи кластеризации.
17. Метод К-средних. Алгоритм K-means++. Сходимость метода К-средних. Модификации метода.
18. Визуализация результатов кластеризации. Методы расчета расстояния между кластерами.
19. Постановка задачи сокращения размерности данных. Подходы к сокращению размерности.
20. Метод главных компонент. Выбор числа главных компонент.
21. Понятие дерева решений. Регрессионные и классификационные деревья. Типы вершин.
22. Обучение деревьев. Обучение ансамблей. Подходы к обучению ансамблей: бэггинг, бустинг, стеккинг.
23. Случайный лес.
24. Техники бустинга моделей машинного обучения. Градиентный бустинг деревьев решений.
25. Математическая модель искусственного нейрона, активационные характеристики нейронов.
26. нейронные сети прямого распространения, математическая модель многослойной нейронной сети.
27. Постановка задачи обучения многослойной нейронной сети. Простой градиентный метод.
28. Виды функций потерь.
29. Метод обратного распространения ошибки. Сеть обратного распространения.
30. Математическая модель сверточной нейронной сети. Виды слоев сети. Постановка задачи обучения сверточной нейронной сети.
31. Применение сверточных нейронных сетей для решения задач классификации, регрессии и трансформации изображений.
32. Задачи обработки естественного языка (NLP). Общая архитектура NLP-системы. Уровни обработки естественного языка.
33. Морфологический анализ текста. Синтаксический анализ текста. Виды синтаксических деревьев. Лексико-семантический анализ.
34. Нейронные сети для прогнозирования временных рядов. Рекуррентные нейронные сети.
35. Нейронные сети с долгой краткосрочной памятью.

7.4.2. Практические задания для проведения промежуточной аттестации обучающихся

Практические задания для проведения промежуточной аттестации размещены: ЭИОС / СДО СПбГАСУ Moodle / Кафедры / Информатики / БАКАЛАВРИАТ и СПЕЦИАЛИТЕТ - кафедра

информатики / Введение в машинное обучение (<https://moodle.spbgasu.ru/course/view.php?id=3051>), а также в Приложении.

7.4.3. Примерные темы курсовой работы (проекта) (при наличии)

Курсовые работы (проекты) учебным планом не предусмотрены.

7.5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта профессиональной деятельности, характеризующие этапы формирования компетенций

Процедура проведения промежуточной аттестации и текущего контроля успеваемости регламентируется локальным нормативным актом, определяющим порядок организации и проведения текущего контроля успеваемости и промежуточной аттестации обучающихся.

Процедура оценивания формирования компетенций при проведении текущего контроля приведена в п. 7.3. Типовые контрольные задания или иные материалы текущего контроля приведены в п. 7.2.

Промежуточная аттестация по дисциплине проводится в форме зачета. Зачет проводится в форме практического контрольного задания и собеседования.

7.6. Критерии оценивания сформированности компетенций при проведении промежуточной аттестации

Критерии оценивания	Уровень освоения и оценка			
	Оценка «неудовлетворительно»	Оценка «удовлетворительно»	Оценка «хорошо»	Оценка «отлично»
	«не зачтено»	«зачтено»		
	Уровень освоения компетенции «недостаточный». Компетенции не сформированы. Знания отсутствуют, умения и навыки не сформированы	Уровень освоения компетенции «пороговый». Компетенции сформированы. Сформированы базовые структуры знаний. Умения фрагментарны и носят репродуктивный характер. Демонстрируется низкий уровень самостоятельности практического навыка.	Уровень освоения компетенции «продвинутый». Компетенции сформированы. Знания обширные, системные. Умения носят репродуктивный характер, применяются к решению типовых заданий. Демонстрируется достаточный уровень самостоятельности устойчивого практического навыка.	Уровень освоения компетенции «высокий». Компетенции сформированы. Знания аргументированные, всесторонние. Умения успешно применяются к решению как типовых, так и нестандартных творческих заданий. Демонстрируется высокий уровень самостоятельности, высокая адаптивность практического навыка

знания	<p>Обучающийся демонстрирует:</p> <ul style="list-style-type: none"> -существенные пробелы в знаниях учебного материала; -допускаются принципиальные ошибки при ответе на основные вопросы билета, отсутствует знание и понимание основных понятий и категорий; -непонимание сущности дополнительных вопросов в рамках заданий билета. 	<p>Обучающийся демонстрирует:</p> <ul style="list-style-type: none"> -знания теоретического материала; -неполные ответы на основные вопросы, ошибки в ответе, недостаточное понимание сущности излагаемых вопросов; -неуверенные и неточные ответы на дополнительные вопросы. 	<p>Обучающийся демонстрирует:</p> <ul style="list-style-type: none"> -знание и понимание основных вопросов контролируемого объема программного материала; - знания теоретического материала -способность устанавливать и объяснять связь практики и теории, выявлять противоречия, проблемы и тенденции развития; -правильные и конкретные, без грубых ошибок, ответы на поставленные вопросы. 	<p>Обучающийся демонстрирует:</p> <ul style="list-style-type: none"> -глубокие, всесторонние и аргументированные знания программного материала; -полное понимание сущности и взаимосвязи рассматриваемых процессов и явлений, точное знание основных понятий, в рамках обсуждаемых заданий; -способность устанавливать и объяснять связь практики и теории, -логически последовательные, содержательные, конкретные и исчерпывающие ответы на все задания билета, а также дополнительные вопросы экзаменатора.
умения	<p>При выполнении практического задания билета обучающийся продемонстрировал недостаточный уровень умений. Практические задания не выполнены. Обучающийся не отвечает на вопросы билета при дополнительных наводящих вопросах преподавателя.</p>	<p>Обучающийся выполнил практическое задание билета с существенными неточностями. Допускаются ошибки в содержании ответа и решении практических заданий. При ответах на дополнительные вопросы было допущено много неточностей.</p>	<p>Обучающийся выполнил практическое задание билета с небольшими неточностями. Показал хорошие умения в рамках освоенного учебного материала. Предложенные практические задания решены с небольшими неточностями. Ответил на большинство дополнительных вопросов.</p>	<p>Обучающийся правильно выполнил практическое задание билета. Показал отличные умения в рамках освоенного учебного материала. Решает предложенные практические задания без ошибок. Ответил на все дополнительные вопросы.</p>

владение навыками	Не может выбрать методику выполнения заданий. Допускает грубые ошибки при выполнении заданий, нарушающие логику решения задач. Делает некорректные выводы. Не может обосновать алгоритм выполнения заданий.	Испытывает затруднения по выбору методики выполнения заданий. Допускает ошибки при выполнении заданий, нарушения логики решения задач. Испытывает затруднения с формулированием корректных выводов. Испытывает затруднения при обосновании алгоритма выполнения заданий.	Без затруднений выбирает стандартную методику выполнения заданий. Допускает ошибки при выполнении заданий, не нарушающие логику решения задач. Делает корректные выводы по результатам решения задачи. Обосновывает ход решения задач без затруднений.	Применяет теоретические знания для выбора методики выполнения заданий. Не допускает ошибок при выполнении заданий. Самостоятельно анализирует результаты выполнения заданий. Грамотно обосновывает ход решения задач.
-------------------	--	---	---	--

Оценка по дисциплине зависит от уровня сформированности компетенций, закрепленных за дисциплиной, и представляет собой среднее арифметическое от выставленных оценок по отдельным результатам обучения (знания, умения, владение навыками).

Оценка «отлично»/«зачтено» выставляется, если среднее арифметическое находится в интервале от 4,5 до 5,0.

Оценка «хорошо»/«зачтено» выставляется, если среднее арифметическое находится в интервале от 3,5 до 4,4.

Оценка «удовлетворительно»/«зачтено» выставляется, если среднее арифметическое находится в интервале от 2,5 до 3,4.

Оценка «неудовлетворительно»/«не зачтено» выставляется, если среднее арифметическое находится в интервале от 0 до 2,4.

8. Учебно-методическое и материально-техническое обеспечение дисциплины (модуля)

8.1. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины (модуля)

№ п/п	Автор, название, место издания, издательство, год издания учебной и учебно-методической литературы	Количество экземпляров/электронный адрес ЭБС
<u>Основная литература</u>		
1	Барский А. Б., Введение в нейронные сети, Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2024	https://www.iprbooks.hop.ru/133929.html
2	Платонов А. В., Машинное обучение, Москва: Юрайт, 2024	https://urait.ru/bcode/544780
3	Шарден Б., Массарон Л., Боскетти А., Крупномасштабное машинное обучение вместе с Python, Москва: ДМК Пресс, 2018	https://e.lanbook.com/book/105836
4	Ростовцев В. С., Искусственные нейронные сети, Санкт-Петербург: Лань, 2024	https://e.lanbook.com/book/364517
<u>Дополнительная литература</u>		
1	Седов В. А., Седова Н. А., Введение в нейронные сети, Саратов: Ай Пи Эр Медиа, 2018	http://www.iprbookshop.ru/69319.html

2	Рашка С., Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения, Москва: ДМК Пресс, 2017	https://e.lanbook.com/book/100905
---	---	---

Обучающиеся из числа инвалидов и лиц с ОВЗ обеспечиваются печатными и (или) электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

8.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

Наименование ресурса сети «Интернет»	Электронный адрес ресурса
Система дистанционного обучения СПбГАСУ Moodle	https://moodle.spbgasu.ru/

8.3. Перечень современных профессиональных баз данных и информационных справочных систем

Наименование	Электронный адрес ресурса
Образовательные интернет-ресурсы СПбГАСУ	https://www.spbgasu.ru/university/obrazovatelnye-internet-resursy/
Российская государственная библиотека	www.rsl.ru
Научная электронная библиотека eLIBRARY.RU	https://www.elibrary.ru/
Электронно-библиотечная система издательства "IPRsmart"	http://www.iprbookshop.ru/
Электронно-библиотечная система издательства "ЮРАЙТ"	https://www.biblio-online.ru/
Электронно-библиотечная система издательства "Лань"	https://e.lanbook.com/
Электронная библиотека Ирбис 64	http://ntb.spbgasu.ru/irbis64r_plus/

8.4. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень лицензионного и свободно распространяемого программного обеспечения

Наименование	Способ распространения (лицензионное или свободно распространяемое)
Python версия 3.7.6386.10	Свободно распространяемое
PyCharm Community	Свободно распространяемое
Anaconda	Свободно распространяемое

8.5. Материально-техническое обеспечение дисциплины

Сведения об оснащённости учебных аудиторий и помещений для самостоятельной работы

Наименование учебных аудиторий и помещений для самостоятельной работы	Оснащённость оборудованием и техническими средствами обучения
72. Учебные аудитории для проведения практических занятий, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации	Учебная аудитория для проведения практических занятий, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации – комплект мультимедийного оборудования (персональный компьютер, мультимедийный проектор, экран, аудиосистема), доска маркерная белая эмалевая, комплект учебной мебели, подключение к компьютерной сети СПбГАСУ, выход в Интернет.
72. Компьютерный класс	Рабочие места с ПК (стол компьютерный, системный блок, монитор, клавиатура, мышь), стол рабочий, подключение к компьютерной сети СПбГАСУ, выход в Internet.

72. Помещения для самостоятельной работы	Помещение для самостоятельной работы (читальный зал библиотеки, ауд. 217): ПК-23 шт., в т.ч. 1 шт.- ПК для лиц с ОВЗ (системный блок, монитор, клавиатура, мышь) с подключением к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду СПбГАСУ.
--	---

Для инвалидов и лиц с ОВЗ обеспечиваются специальные условия для получения образования в соответствии с требованиями нормативно-правовых документов.

Рабочая программа составлена на основе ФГОС ВО - бакалавриат по направлению подготовки 38.03.05 Бизнес-информатика (приказ Минобрнауки России от 29.07.2020 № 838).

Программу составил:

доцент Инф, к.ф.-м.н. Мовсесова Л.В.

Программа обсуждена и рекомендована на заседании кафедры Информатики

30.08.2024, протокол № 1

Заведующий кафедрой к.ф.-м.н., доцент Л.В. Мовсесова

Программа одобрена на заседании учебно-методической комиссии факультета

19.09.2024, протокол № 2.

Председатель УМК д.э.н., профессор Г.Ф. Токунова

Тестовые задания для проверки сформированности индикаторов достижения компетенций ОПК-3.1, ОПК-3.2, ОПК-3.3

Вопрос 1. Назовите функции активации, применяющиеся в нейронных сетях (установите соответствие):

- Функция, которая обеспечивает значения всех выходных узлов от 0 до 1, а сумма всех значений выходных узлов всегда равна 1.
- Функция, которая приводит любое реальное значение к диапазону значений от 0 до 1.

Варианты ответа:

- Softmax
- Sigmoid
- Momentum
- Adam

Вопрос 2. Терминология нейронных сетей (установите соответствие):

– Весь датасет прошел через нейронную сеть в прямом и обратном направлении один раз

- Разбиение датасета на пакеты
- Число батчей, необходимых для завершения одной эпохи

Варианты ответа:

- Батч (batch)
- Итерация
- Эпоха (epoch)

Вопрос 3. Назовите оптимизаторы, применяющиеся в нейронных сетях (установите соответствие):

- Стохастический градиентный спуск
- Оптимизатор импульса
- Среднеквадратичное распространение
- Адаптивная оценка момента

Варианты ответа:

- SGD
- Momentum
- RMSProp
- Adam

Вопрос 4. Функция потерь в машинном обучении – это ...

- алгоритм, который используется для минимизации ошибки
- функция, которая используется для оценки качества модели во время обучения
- функция, которая определяет сложность решаемой задачи
- функция, интерпретируемая в зависимости от области применения

Вопрос 5. Оцениваемая в процентах метрика, которая наиболее часто используемая для оценки качества бинарной классификации – это ...

- MSE
- Precision
- Accuracy
- R^2

Вопрос 6. Что происходит в процессе обучения модели в машинном обучении?

- Модель получает входные данные и использует их для предсказания выходных данных
- Результаты предсказания сравниваются с реальными данными, используя функцию потерь
- Ошибку, которую необходимо минимизировать, вычисляет функция потерь
- Все перечисленное

Вопрос 7. Метод машинного обучения, используемый для задачи классификации – это...

- Линейная регрессия
- Градиентный спуск
- Логистическая регрессия
- Метод ближайших соседей
- Все перечисленные

Вопрос 8. Метод машинного обучения, используемый для задачи кластеризации – это...

- Линейная регрессия
- К-средних
- Логистическая регрессия
- Метод опорных векторов
- Все перечисленные

Вопрос 9. Метод машинного обучения, используемый для уменьшения размерности данных – это...

- Линейная регрессия
- Логистическая регрессия
- Метод ближайших соседей
- Метод главных компонент
- Все перечисленные

Вопрос 10. Для обработки временных рядов используется

- Деревья решений
- Линейная регрессия
- Метод ближайших соседей
- Рекуррентные нейронные сети
- Все перечисленное

Вопрос 11. Что такое переобучение модели?

- Модель плохо обучается на тренировочных данных, она переобучается на тестовых
- Модель обучается без использования данных тренировочных данных
- Модель слишком сложная, хорошо работает на тренировочных данных, но плохо на тестовых
- Все вышеперечисленное

Вопрос 12. Метод, используемый для борьбы с переобучением – это ...

- Регуляризация
- Увеличение размера тренировочной выборки
- Увеличение числа слоев в нейронной сети
- Уменьшение количества итераций обучения
- Все перечисленные

Практические задания

При выполнении заданий необходимо произвести **предобработку данных**.

1. Загрузить набор данных (по своему варианту) как датафрейм библиотеки *pandas*.

2. Получить представления о наборе данных с помощью методов библиотеки *pandas* (*shape*, *head*, *tail*, *describe*, *info*). Посчитать количество пропущенных значений в признаках.

3. Произвести разведочный анализ данных – визуализировать данные, используя библиотеки *matplotlib*, *seaborn*, *plotly*.

Для этого построить (в зависимости от конкретного датасета и решаемой задачи):

- 1) парные диаграммы;
- 2) тепловую карту по матрице корреляции;
- 3) гистограммы;
- 4) диаграммы размаха («ящик с усами»);
- 5) графики распределения;
- 6) точечные 3D-графики распределения признаков.

4. Сделать выводы на основе проведенного анализа.

5. Произвести предобработку данных (в зависимости от решаемой задачи):

1) удалить из набора данных признаки, слабо коррелирующие между собой и с анализируемым признаком;

2) закодировать категориальные признаки в дискретные величины (метод *get_dummies* библиотеки *scikit-learn*);

3) объединить значения по периодам;

4) произвести нормализацию данных (метод *RobustScaler* библиотеки *scikit-learn*).

6. Разбить обработанный набор данных на обучающую и тестовую выборки (метод *train_test_split* библиотеки *scikit-learn*).

1. Задача бинарной классификации сбалансированного набора данных

– Произвести обучение следующих моделей библиотеки *scikit-learn*:

- 1) метод опорных векторов (*Support vector machine*)
- 2) логистическая регрессия (*LogisticRegression*)
- 3) дерево решений (*Decision Tree*)
- 4) случайный лес (*RandomForest*)

– Отобразить корректность работы каждой модели на тестовой и обучающей выборках с помощью метрики *accuracy* библиотеки *scikit-learn*.

2. Задача многоклассовой классификации сбалансированного набора данных

– Произвести обучение следующих моделей библиотеки *scikit-learn*:

- 1) метод опорных векторов (*Support vector machine*)
- 2) логистическая регрессия (*LogisticRegression*)

- 3) дерево решений (*Decision Tree*)
- 4) наивный Байес (*Naive Bayes*)
 - Отобразить корректность работы каждой модели на тестовой и обучающей выборках с помощью метрики *accuracy* библиотеки *scikit-learn*.

3. Задача классификации несбалансированного набора данных

- Определить веса классов (*compute_class_weight*)
- Для учета несбалансированности набора данных использовать:
 - 1) метод *взвешенной* классификации (*class_weight = 'balanced'*);
 - 2) метод *oversampling* (*RandomOverSampler*);
 - 3) метод *undersampling* (*RandomUnderSampler*).
- Произвести обучение следующих моделей библиотеки *scikit-learn*:
 - 1) Логистическая регрессия (*LogisticRegression*)
- Отобразить точность работы каждой модели на тестовой и обучающей выборках с помощью метрик *precision*, *recall*, *f1-score* и ROC-кривой библиотеки *scikit-learn*.
- Повторно отобразить метрики после проведения кросс-валидации.

4. Задача восстановления регрессии

- Произвести обучение следующих моделей библиотеки *scikit-learn*:
 - 1) линейная регрессия (*Linear Regression*);
 - 2) регрессия LASSO (*Lasso Regression*);
 - 3) гребневая регрессия (*Ridge Regression*);
 - 4) полиномиальная регрессия (*Polynomial Regression*).
- Отобразить точность работы каждой модели на тестовой и обучающей выборках с помощью метрик *r squared* (коэффициент детерминации), *mean absolute error* (средняя абсолютная ошибка), *mean squared error* (средняя квадратичная ошибка) библиотеки *scikit-learn*.
- Повторно отобразить метрики после проведения кросс-валидации.
- Произвести поиск оптимальных гиперпараметров с помощью метода *GridSearchCV* по модели с наилучшими показателями.

5. Задача кластеризации

- Произвести снижение размерности набора данных
- Произвести обучение следующих моделей библиотеки *scikit-learn*:
 - 1) метод *k-средних* (*k-means clustering*)
- Произвести разведочный анализ данных по кластерам для оценки качества обучения модели.

6. Задача выявления аномалий

- Провести кластеризацию данных

- Произвести обучение следующих моделей библиотеки *scikit-learn*:
 - 1) метод k-средних (*k-means clustering*)
 - 2) цепь Маркова (*Markov Chain*)
 - 3) изолирующий лес (*Isolation forest*)
 - 4) одноклассовый метод опорных векторов (*One class SVM*)
- Произвести разведочный анализ данных по набору данных для оценки качества обучения модели.

7. Задача прогнозирования временных рядов

- Провести кластеризацию данных
- Произвести обучение следующих моделей библиотеки *scikit-learn*:
 - 1) *XGBoost Regressor*
- Отобразить точность работы модели на тестовой и обучающей выборках с помощью метрики *r squared* (коэффициент детерминации) библиотеки *scikit-learn*.