Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 13.08.2024 12:18:51
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖДЕНО
Декан факультета
Информационных технологий
_____ / Демидов Д.Г. /
«15» февраля 2024 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Методы работы с большими данными»

Направление подготовки
09.03.01 Информатика и вычислительная техника

Профиль подготовки
«Интеллектуальные беспилотные системы»

Квалификация (степень) выпускника
Бакалавр

Форма обучения
Очная

Москва 2024 г.

Рабочая программа составлена на основе ФГОС ВО и учебного плана Московского политехнического университета по направлению (специальности) 09.03.01 Информатики и вычислительная техника, по профилю подготовки Интеллектуальные беспилотные системы.

Составители рабочей программы:

доцент кафедры «СМАРТ технологии», к.ф.-м.н.		Т. Т. Идиатуллов
(должность, ученое звание, степень)	(подпись)	(Ф.И.О.)

доцент кафедры «СМАРТ технологии», к.т.н., доцент		Д. И. Давлетчин
(должность, ученое звание, степень)	(подпись)	(Ф.И.О.)

Рабочая программа утверждена на заседании кафедры

«СМАРТ технологии»
(наименование кафедры)

Заведующий кафедрой к.т.н., доцент		Е. В. Петрунина
	(подпись)	(Ф.И.О.)

СОГЛАСОВАНО:

Заведующий кафедрой «СМАРТ технологии», к.т.н., доцент		Е. В. Петрунина
	(подпись)	(Ф.И.О.)

1. Цели, задачи и планируемые результаты обучения по дисциплине

К основным целям освоения дисциплины «Методы работы с большими данными» относится: формирование у студентов необходимой теоретической базы и практических навыков, которые позволят всесторонне и системно понимать современные проблемы прикладной математики и информатики, проблемы обработки и анализа информации, а также разрабатывать и анализировать концептуальные и теоретические модели при решении научных и прикладных задач в области информационных технологий.

К основным задачам дисциплины «Методы работы с большими данными» относятся:

1. Изучение основных методов решения на ЭВМ задач анализа и интерпретации данных, получаемых с помощью различного рода информационно-измерительных систем. Знание базовых алгоритмов анализа и интерпретации данных.
2. Формирование навыков применения методов решения на ЭВМ задач анализа и интерпретации данных при разработке алгоритмов анализа и обработки измерительной информации. Умение использовать стандартную терминологию, определения и обозначения.
3. Освоение основных тенденций развития теории и практики анализа и интерпретации данных. Приобретение практических навыков работы с современными пакетами прикладных программ для решения задач анализа и интерпретации данных.

Перечень планируемых результатов обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

В результате освоения дисциплины у обучающихся формируются следующие компетенции и должны быть достигнуты следующие результаты обучения как этап формирования соответствующих компетенций.

Код компетенции	В результате освоения образовательной программы обучающийся должен обладать	Перечень планируемых результатов обучения по дисциплине
ПК-2	ПК-2. Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и	ПК-2.1. Знает: методы целеполагания; теорию ключевых показателей деятельности; методы концептуального проектирования; стандарты оформления

	крупного масштаба и сложности.	технических заданий; теорию тестирования; методы оценки качества программных систем; методы тестирования; международные стандарты на структуру документов требований; нормативные и методические материалы по созданию документов требований к системам. ПК-2.2. Умеет: формулировать цели, исходя из анализа проблем, потребностей и возможностей; разрабатывать технико-экономическое обоснование; декомпозировать функции на подфункции; алгоритмизировать деятельность; разрабатывать структуры типовых документов; исполнять ручные тесты. ПК-2.3. Владеет: навыками логического мышления; средствами автоматизации проектирования ПО.
ПК-3	ПК-3 Способен разрабатывать требования и проектировать программное обеспечение	ИПК-3.1 Знает методы разработки требований и проектирования программного обеспечения в рамках задач, поставленных на практику. ИПК-3.2 Умеет формировать требования и выполнять работы по проектированию программного обеспечения в рамках задач производственной практики. ИПК-3.3 Владеет навыками разработки требований и проектирования прикладного программного обеспечения в рамках задач производственной практики.
ПК-5	ПК-5 Способен	ИПК-5.1 Знает методы разработки

	разрабатывать и применять системы на базе технологий искусственного интеллекта и беспилотной робототехники	и применения систем на базе технологий искусственного интеллекта и беспилотной робототехники. ИПК-5.2 Умеет разрабатывать и применять системы на базе технологий искусственного интеллекта и беспилотной робототехники. ИПК-5.3 Владеет навыками интеграции и применения систем на базе технологий искусственного интеллекта и беспилотной робототехники.

В процессе освоения образовательной программы данные компетенции, в том числе их отдельные компоненты, формируются поэтапно в ходе освоения обучающимися дисциплин (модулей), практик в соответствии с учебным планом и календарным графиком учебного процесса.

2. Место дисциплины в структуре ОПОП

Дисциплина относится к числу учебных дисциплин формируемые участниками образовательных отношений части Дисциплины по выбору студента «Элективные дисциплины» основной профессиональной образовательной программы.

Дисциплина взаимосвязана логически и содержательно-методически со следующими дисциплинами и практиками ОПОП:

- Алгоритмическое программирование;
- Искусственные нейронные сети;
- Численные методы в задачах управления

3. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, 144 академических часа (из них 72 часа – самостоятельная работа студентов).

Разделы дисциплины изучаются на четвертом курсе в седьмом семестре, форма контроля - экзамен.

3.1. Виды учебной работы и трудоемкость (по формам обучения)

Очная форма обучения

№ п/п	Вид учебной работы	Количество часов	Семестры
			7
1	Аудиторные занятия	72	72
	В том числе:		
1.1	Лекции	36	36
1.2	Семинарские/практические занятия		
1.3	Лабораторные занятия	36	36
2	Самостоятельная работа	72	72
	В том числе:		
2.1	Выполнение самостоятельных практических занятий	70	70
2.2	Тестирование	2	2
3	Промежуточная аттестация		
	Экзамен		
	Итого:	144/4	144/4

3.2. Тематический план изучения дисциплины (по формам обучения)
Очная форма обучения

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/практические занятия	Лабораторные занятия	Практическая подготовка	
1	**Тема 1. Введение. Предмет, задачи и содержание дисциплины**. Основные положения, понятия и определения. Краткие сведения о развитии теории анализа и интерпретации данных. Структура, содержание дисциплины, ее связи с другими дисциплинами учебного плана специальности.		2		2		4
2	**Тема 2. Основные понятия дисциплины** Введение в анализ данных. Проблема обработки данных. Матрица данных. Гипотезы компактности и скрытых факторов. Структура матрицы данных и задачи обработки. Матрица объект-объект и признак-признак		4		4		8
3	**Тема 3. Классификация данных с использованием детерминированных моделей** Решающие поверхности и дискриминантные функции. Линейные дискриминантные функции. Линейная разделимость. Кусочно-линейные дискриминантные функции**.**		4		4		12
4	**Тема 4. Кластерный анализ** Основные типы задач кластер-анализа. Меры подобия и функции расстояния Выбор критерия кластеризации. Кластерные методы, основанные на евклидовой метрике. Иерархическая кластеризация. Метод Квнутригрупповых средних.		4		4		12

	Использование методов теории графов в задачах кластеризации. Кластеризация на основе анализа плотностей вероятностей						
5	**Тема 5. . Методы снижения размерностей данных** Анализ матриц исходных данных. Метод главных компонент. Корреляционная матрица и ее основные свойства		6		6		12
6	**Тема 6. Системы DATA MINING в задачах анализа и интерпретации данных** (Понятие об интеллектуальных системах анализа и интерпретации данных. DATA MINING - системы извлечения новых знаний из данных.		8		8		12
7	**Тема 7. Современные пакеты прикладных программ для решения задач обработки экспериментальных данных**. Библитеки Sklearn, Seaborn, Keras, Tensorflow.		8		8		12
Итого			**36**		**36**		**72**

3.3. Содержание дисциплины

Введение

Предмет дисциплины и ее задачи. Краткие сведения о развитии теории анализа и интерпретации данных. Структура, содержание дисциплины, ее связи с другими дисциплинами учебного плана специальности. Перечень дисциплин и их разделов, усвоение которых необходимо студентам для изучения курса «Анализ и интерпретация данных».

Тема 1. Основные понятия дисциплины

Введение в анализ данных. Проблема обработки данных. Матрица данных. Гипотезы компактности и скрытых факторов. Структура матрицы данных и задачи обработки. Матрица объект-объект и признак-признак. Расстояние и близость. Измерение признаков. Отношения и их представление. Основные проблемы измерений. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных .

Тема 2. Классификация данных с использованием детерминированных моделей

Решающие поверхности и дискриминантные функции. Линейные дискриминантные функции. Линейная разделимость. Кусочно-линейные дискриминантные функции. Нелинейные дискриминантные функции. Фимашины. Потенциальные функции как дискриминантные функции. Пространство весов. Процедуры обучения с коррекцией ошибок: правило с фиксированным приращением, правило абсолютной коррекции, частично корректирующее правило. Обобщенные градиентные методы. Персептронный критерий. Процедуры обучения на основе минимальной среднеквадратичной ошибки: псевдоинверсный метод, метод Хо-Кашпа

Тема 3. Классификация данных на основе статистических моделей

Функция потерь. Байесовская дискриминантная функция. Принятие решение по максимуму правдоподобия. Оптимальная дискриминантная функция для нормально распределенных образов. Дискриминантная функция Фишера. Множественный дискриминантный анализ. Пошаговый дискриминантный анализ. Ошибки классификации. Примеры построения статистических дискриминантных функций для различных статистических моделей данных. Обучение для статистических дискриминантных функций. Оценки максимального правдоподобия, байесовские оценки. Непараметрическое оценивание. Парзеновские окна, метод непараметрического оценивания на основе К-ближайшего соседства

Тема 4. Кластер-анализ

Основные типы задач кластер-анализа. Меры подобия и функции расстояния Выбор критерия кластеризации. Кластерные методы, основанные на евклидовой метрике. Иерархическая кластеризация. Метод Квнутригрупповых средних. Использование методов теории графов в задачах кластеризации. Кластеризация на основе анализа плотностей вероятностей

Тема 5. Методы снижения размерностей данных

Анализ матриц исходных данных. Метод главных компонент. Корреляционная матрица и ее основные свойства. Собственные векторы и собственные числа корреляционной матрицы. Приведение корреляционной матрицы к диагональной форме. Геометрическая интерпретация главных компонент на плоскости. Модели факторного анализа. Оценка факторных нагрузок методом максимального правдоподобия и центроидным методом. Вращение факторов и их интерпретация. Использование кластеризации признаков для снижения размерности. Многомерное шкалирование. Метрический и неметрический подход к многомерному шкалированию. Методы ортогонального проектирования. Нелинейные методы многомерного шкалирования. Многомерное шкалирование неметрических данных. Многомерные развертки

Тема 6. Системы DATA MINING в задачах анализа и интерпретации данных (Понятие об интеллектуальных системах анализа и интерпретации

данных. DATA MINING - системы извлечения новых знаний из данных. Типы систем DATA MINING -предметно-ориентированные аналитические системы, статистические пакеты, нейронные сети, деревья решений, обнаружение логических закономерностей, генетические алгоритмы, системы визуализации многомерных данных

Тема 7. Современные пакеты прикладных программ для решения задач обработки экспериментальных данных. Библитеки Sklearn, Seaborn, Keras, Tensorflow. Табличные процессоры и базы данных в задачах обработки данных. Виды статистических пакетов.

3.4. Тематика семинарских/практических и лабораторных занятий

Лабораторные работы
Лабораторная работа №1 «Ознакомление с фреймворком Scilearn. Препроцессинг данных. Нормализация. Предобработка данных.»
Лабораторная работа №2 «.Работа с библиотеками Numpy, Scipy, Pandas, Matplotlib, Seaboarn, Scikit-learn. Решение задачи линейной регрессии»
Лабораторная работа №3 « Решение задачи классификации и кластеризации»
Лабораторная работа №4 «Tensorflow. Ознакомление с фреймворком Keras. Решение задачи классификации изображений. MLP»
Лабораторная работа №5 «Обучение нейронной сети. Гиперпараметры. Подбор гиперпараметров с помощью KerasTuner. Переобучение, недообучение.»
Лабораторная работа №6 «Создание модели сверточной нейронной мети для решения задачи распознавания букв латинского алфавита»
Лабораторная работа №7 «CNN. Архитектуры RCNN, FastCNN, ResNet.»
Лабораторная работа №8 «Глубокое обучение. Использование предобученных моделей для решения задачи классификации»
Лабораторная работа №9 «Автокодировщики. VAE, CVAE. Генератиные состязательные сети.»

Для проведения лабораторных работ требуется компьютерный класс. объеденный в локальную сеть с выходом в Интернет. Компьютеры должны быть объединены локальной сетью. Необходим выход в глобальную сеть Интернет. Требуемое программное обеспечением: компилятор языка Python, Tensorflow с пакетом Keras, текстовый редактор, офисный пакет LibreOffice.

Компьютерный класс должен иметь возможность обновления и установки дополнительного свободно распространяемого программного обеспечения.

3.5. Тематика вопросов для самостоятельного изучения

• Изучение тенденции применения различных языков программирования при решении разных практических задач.

• Изучение сред разработки, систем управления версиями.

• Изучение методов коллективной разработки.

• Изучение средств автоматизированного тестирования приложений.

4. Учебно-методическое и информационное обеспечение

4.1. Нормативные документы и ГОСТы

1. ГОСТ Р 43.0.12-2018 Базы знаний в технической деятельности.
2. ГОСТ Р 57321.2-2018 Менеджмент знаний. Менеджмент знаний в области инжиниринга. Часть 2. Проектирование на основе баз знаний.
3. ГОСТ Р 43.0.28-2022 Информационное обеспечение техники и операторской деятельности. Базы знаний в интеллектуализации деятельности.
4. ГОСТ Р 59869-2021 Интеллектуальные системы обучения. Общие положения.

4.2. Основная литература

1. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 495 с. — (Высшее образование). — ISBN 978-5-534-16238-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/536688 (дата обращения: 26.05.2024).
2. Яхъяева Г. Э. Основы теории нейронных сетей. – Национальный Открытый Университет «ИНТУИТ», 2016 г. – 200 с. (http://www.knigafund.ru/books/178963).
3. Воронов, М. В. Системы искусственного интеллекта : учебник и практикум для вузов / М. В. Воронов, В. И. Пименов, И. А. Небаев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 268 с. — (Высшее образование). — ISBN 978-5-534-17032-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/544161 (дата обращения: 26.05.2024).

4.3. Дополнительная литература

1. Емельянов В.В., Курейчик В.М., Курейчик В.В. Теория и практика эволюционного моделирования. – М.: ФИЗМАТЛИТ, 2003. – 432 с. – ISBN 5-9221-0337-7. (www.knigafund.ru/books/207330)

2. Бречка, Д.М. Алгоритмы машинных вычислений: учебно-методическое пособие. [Электронный ресурс] — Электрон. дан. — Омск : ОмГУ, 2014. — 64 с. — Режим доступа: http://e.lanbook.com/book/75387 — Загл. с экрана.
3. Окулов, С.М. Алгоритмы компьютерной арифметики. [Электронный ресурс] / С.М. Окулов, С.М. Лялин, О.А. Пестов, Е.В. Разова. — Электрон. дан. — М. : Издательство "Лаборатория знаний", 2015. — 288 с. — Режим доступа: http://e.lanbook.com/book/66112 — Загл. с экрана.
4. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И. Д. Рудинского. – М.: Горячая линия –Телеком, 2006. – 452 с.
5. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с. (Науки об искусственном). – ISBN 5-8360-0330-0.
6. Гасанов, Э. Э. Интеллектуальные системы. Теория хранения и поиска информации : учебник для вузов / Э. Э. Гасанов, В. Б. Кудрявцев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2024. — 271 с. — (Высшее образование). — ISBN 978-5-534- 08684-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/537938 (дата обращения: 26.05.2024).
7. Галушкин А.И. Теория нейронных сетей. Кн.1: Учебное пособие для вузов. – М.: ИПЖР, 2000. – 416 с.
8. Башмаков А.И., Башмаков И.А. Интеллектуальные информационные технологии: Учеб. пособие. – М.: Изд-во МГТУ им. Н.Э.Баумана, 2005. – 304 с. – (Серия «Информатика в техническом университете»).

4.4. Электронные образовательные ресурсы

1. ЭОР в разработке.

4.5. Лицензионное и свободно распространяемое программное обеспечение

1. Среда разработки Microsoft VisualStudio с установленным пакетом расширения языка Python
3. Офисный пакет Libre Office или Microsoft Office

4.6. Современные профессиональные базы данных и информационные справочные системы

1. Справочно-правовая системы «КонсультантПлюс: Некоммерческая интернетверсия» https://www.consultant.ru/online/

2. Официальный интернет-портал правовой информации http://pravo.gov.ru
3. Российская национальная библиотека http://www.nlr.ru
4. ЭБС «Университетская библиотека онлайн» https://biblioclub.ru/index.php
5. Единое окно доступа к образовательным ресурсам Федеральный портал http://window.edu.ru
6. Научная электронная библиотека http://www.elibrary.ru
7. Российская государственная библиотека http://www.rsl.ru

5. Материально-техническое обеспечение

1. Компьютерные классы с оснащением: столы, стулья, аудиторная доска, использование переносного мультимедийного комплекса (переносной проектор, персональный ноутбук).
2. Персональные компьютеры, мониторы, мышки, клавиатуры. Рабочее место преподавателя: стол, стул.
3. Аудитория для самостоятельной работы.
4. Библиотека, читальный зал.

6. Методические рекомендации

6.1. Методические рекомендации для преподавателя по организации обучения

1. При подготовке к занятиям следует предварительно проработать материал занятия, предусмотрев его подачу точно в отведенное для этого время занятия. Следует подготовить необходимые материалы – теоретические сведения, задачи и др. При проведении занятия следует контролировать подачу материала и решение заданий с учетом учебного времени, отведенного для занятия.
2. При проверке работ и отчетов следует учитывать не только правильность выполнения заданий, но и оптимальность выбранных методов решения, правильность выполнения всех его шагов.
3. При организации и проведения экзаменов в практико-ориентированной форме следует использовать утвержденные кафедрой Методические рекомендации.

6.2. Методические указания для обучающихся по освоению дисциплины

Получение углубленных знаний по дисциплине достигается за счет активной самостоятельной работы обучающихся. Выделяемые часы целесообразно использовать для знакомства с учебной и научной литературой по проблемам дисциплины, анализа научных концепций.

Промежуточная аттестация обучающихся в форме зачѐта проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине проводится преподавателем, ведущим занятия по дисциплине методом экспертной оценки. По итогам промежуточной аттестации по дисциплине выставляется оценка «отлично», «хорошо», «удовлетворительно» или «не удовлетворительно».

К промежуточной аттестации допускаются только обучающиеся, выполнившие все виды учебной работы, предусмотренные рабочей программой по дисциплине «Искусственные нейронные сети».

6.3. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья: - создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;

- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников, например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и ассимиляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для

организации форм текущего и промежуточного контроля;

- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи: - продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут; - продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления, обучающегося при защите курсовой работы - не более чем на 15 минут.

Для обеспечения подготовки людей в формате очной аудиторной работы с ограниченными возможностями движения выбираются аудитории с доступностью в рамках требований по организации безбарьерной среды движения.

7. Фонд оценочных средств

7.1. Методы контроля и оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине (модулю).

Показатель	**Критерии оценивания**			
	2	**3**	**4**	**5**
ПК-2. Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности.				
знает: основы высшей математики, информатики и программирования.	Обучающийся демонстрирует полное отсутствие знаний основных характеристик, методов и алгоритмов машинного обучения..	Обучающийся демонстрирует неполное соответствие знаний основных характеристик, методов и алгоритмов областей применения знаний основных характеристик, методов и алгоритмов машинного	Обучающийся демонстрирует частичное соответствие знаний основных характеристик, основных характеристик, методов и алгоритмов областей применения знаний основных характеристик, методов и	Обучающийся демонстрирует полное соответствие знаний основных характеристик, областей применения основных характеристик, методов и алгоритмов областей применения знаний основных

		обучения. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	алгоритмов машинного обучения.. Допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	характеристик, методов и алгоритмов машинного обучения.. Свободно оперирует приобретенными знаниями.
умеет: решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и моделирования	Обучающийся не умеет пользоваться инструментальными средствами разработки алгоритмов машинного обучения, решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и	Обучающийся демонстрирует неполное соответствие следующих умений: пользоваться инструментальными средствами разработки алгоритмов машинного обучения, решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерн	Обучающийся демонстрирует частичное соответствие следующих умений: пользоваться пользоваться инструментальными средствами разработки алгоритмов машинного обучения, решать стандартные профессиональные задачи с применением естественнонаучных и	Обучающийся демонстрирует полное соответствие следующих умений: пользоваться инструментальными средствами разработки алгоритмов машинного обучения, решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерн

	моделировани я	ых знаний, методов математическо го анализа и моделировани я . Допускаются значительные ошибки, проявляется недостаточнос ть умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании умениями при их переносе на новые ситуации.	общеинженерн ых знаний, методов математическо го анализа и моделировани я. Умения освоены, но допускаются незначительны е ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	ых знаний, методов математическо го анализа и моделировани я. Свободно оперирует приобретенны ми умениями, применяет их в ситуациях повышенной сложности.
владеет: методами теоретическо го и эксперимент ального исследовани я объектов профессиона льной деятельност и.	Обучающийся не владеет навыками методами теоретическог о и экспериментал ьного исследования объектов профессионал ьной деятельности с использование м алгоритмов МО.	Обучающийся в неполном объеме навыками методами теоретическог о и экспериментал ьного исследования объектов профессиональ ной деятельности с использование м алгоритмов МО..	Обучающийся частично владеет навыками методами теоретическог о и экспериментал ьного исследования объектов профессиональ ной деятельности с использование м алгоритмов МО.. Навыки	Обучающийся в полном объеме владеет навыками и методами теоретическог о и экспериментал ьного исследования объектов профессиональ ной деятельности с использование м алгоритмов МО..

		Допускаются значительные ошибки, проявляется недостаточность владения навыками по ряду показателей. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	Свободно применяет полученные навыки в ситуациях повышенной сложности.
ПК-5. Способен разрабатывать и применять системы на базе технологий искусственного интеллекта и беспилотной робототехники				
Знает основы искусственного интеллекта;.	Обучающийся демонстрирует полное отсутствие знаний основ и методов искусственного интеллекта	Обучающийся демонстрирует неполное соответствие знаний основ и методов искусственного интеллекта Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании	Обучающийся демонстрирует частичное соответствие знаний основных принципов основ и методов искусственного интеллекта Допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие знаний основ и методов искусственного интеллекта. Свободно оперирует приобретенными знаниями.

		знаниями при их переносе на новые ситуации.		
Умеет применять искусственный интеллект в инженерных задачах;	Обучающийся не умеет пользоваться инструментальными и средствами разработки программного обеспечения, не показывает умения проектировани я алгоритмы МО для решения инженерных задач.	Обучающийся демонстрирует неполное соответствие следующих умений: пользоваться инструментальными и средствами разработки программного обеспечения, не показывает умения проектировани я алгоритмы МО для решения инженерных задач. Допускаются значительные ошибки, проявляется недостаточнос ть умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании	Обучающийся демонстрирует частичное соответствие следующих умений: пользоваться инструментальными и средствами разработки программного обеспечения, не показывает умения проектировани я алгоритмы МО для решения инженерных задач я. Умения освоены, но допускаются незначительны е ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные	Обучающийся демонстрирует полное соответствие следующих умений: пользоваться инструментальными и средствами разработки программного обеспечения, не показывает умения проектировани я алгоритмы МО для решения инженерных задач Свободно оперирует приобретенны ми умениями, применяет их в ситуациях повышенной сложности.

		умениями при их переносе на новые ситуации.	ситуации.	
Владеет: Навыками использования систем автоматизированного проектирования и специализированного программного обеспечения для инженерных задач, в том числе с использованием ИИ; навыками разработки графических библиотек, программных модулей для САПР и специализированного программного обеспечения, в том числе с использованием ИИ.	Обучающийся не владеет навыками использования систем автоматизированного проектирования и специализированного программного обеспечения для инженерных задач, в том числе с использованием ИИ; Не владеет навыками разработки графических библиотек, программных модулей для САПР и специализированного программного обеспечения, в том числе с использованием ИИ.	Обучающийся в неполном объеме владеет навыками использования систем автоматизированного проектирования и специализированного программного обеспечения для инженерных задач, в том числе с использованием ИИ; навыками разработки графических библиотек, программных модулей для САПР и специализированного программного обеспечения, в том числе с использованием ИИ.. Допускаются значительные ошибки,	Обучающийся частично владеет навыками использования систем автоматизированного проектирования и специализированного программного обеспечения для инженерных задач, в том числе с использованием ИИ; навыками разработки графических библиотек, программных модулей для САПР и специализированного программного обеспечения, в том числе с использованием ИИ.. Навыки освоены, но допускаются незначительные ошибки,	Обучающийся в полном объеме владеет навыками использования систем автоматизированного проектирования и специализированного программного обеспечения для инженерных задач, в том числе с использованием ИИ; навыками разработки графических библиотек, программных модулей для САПР и специализированного программного обеспечения, в том числе с использованием ИИ.. Свободно применяет полученные навыки в

		проявляется недостаточность владения навыками по ряду показателей. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	ситуациях повышенной сложности.

7.2. Шкала и критерии оценивания результатов обучения

Критерии оценки ответа на экзамене

Шкала оценивания	Описание
Отлично	*Выполнены все виды учебной работы, предусмотренные рабочей программой. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателей, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.*
Хорошо	*Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует неполное, правильное соответствие знаний, умений, навыков приведенным в таблицах показателей, либо если при этом были допущены 2-3 несущественные ошибки.*
Удовлетворительно	*Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, в котором освещена основная, наиболее важная часть*

	материала, но при этом допущена одна значительная ошибка или неточность.
Неудовлетворительно	*Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателей, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.*

Критерии оценки работы обучающегося на лабораторных занятиях

Шкала оценивания	**Описание**
Отлично	*Выполнены все практические задания, предусмотренные практическими занятиями, обучающийся четко и без ошибок ответил на все контрольные вопросы, активно работал на практических занятиях.*
Хорошо	*Выполнены все практические задания, предусмотренные практическими занятиями, обучающийся с корректирующими замечаниями преподавателя ответил на все контрольные вопросы, достаточно активно работал на практических занятиях.*
Удовлетворительно	*Выполнены все практические задания, предусмотренные практическими занятиями с замечаниями преподавателя; обучающийся ответил на все контрольные вопросы с замечаниями.*
Неудовлетворительно	*Обучающийся не выполнил или выполнил неправильно практические задания, предусмотренные практическими занятиями; студент ответил на контрольные вопросы с ошибками или не ответил на контрольные вопросы.*

Критерии оценки тестирования

Тестирование оценивается в соответствии с процентом правильных ответов, данных студентом на вопросы теста. Стандартная шкала соответствия результатов тестирования выставляемой балльной оценке:

Шкала оценивания	**Описание**
Отлично	*Свыше 85% правильных ответов (включительно);*
Хорошо	*От 70 % до 84,9 % правильных ответов;*
Удовлетворительно	*От 55 % до 69,9 % правильных ответов;*
Неудовлетворительно	*Менее 54,9 % правильных ответов.*

7.3. Оценочные средства

7.3.01. Текущий контроль на лабораторных занятиях

Пример заданий текущего контроля:

Текущий контроль. Перечень примерных вопросов для защиты лабораторных работ:

1. CRISP-DM — межотраслевой стандарт ведения проектов интеллектуального анализа данных.
2. Библиотеки Numpy, Scipy, Pandas, Matplotlib, Seaboarn, Scikit-learn
3. Аналитические возможности библиотеки NumPy.
4. Функционал библиотеки NumPy для работы с данными.
5. Использование библиотеки Pandas для хранения и анализа данных. Класс pandas.Series. Атрибуты и методы. Класс pandas.DataFrame.
6. Методы объединения и группировки данных в pandas.
7. Визуализация данных в pandas.
8. Визуализация данных в библиотеке matplotlib. Основные элементы и параметры визуализаций в matplotlib. Использование matplotlib.pyplot.
9. Использование библиотеки Scikit-learn для реализации методов машинного обучения.
10. Метрики в задачае регрессии: MAE, MSE, MAPE, R2. Метрики в задаче классификации: кросс-энтропия, precision, recall, F-мера, ROC-кривая, AUC ROC.
11. Обучение линейного алгоритма бинарной классификации образов с помощью градиентного алгоритма
12. Реализация алгоритма метода опорных векторов для задачи бинарной классификации
13. Библиотека Sklearn. Метрические методы классификации Knn

14. Сегментация рынка и выявление профилей клиентов с использованием кластеризации методом k-средних.
15. Библиотека Sklearn. Кластеризация. Kmeans
16. Библиотека Sklearn. Линейная регрессия
17. Деревья решений. Алгоритмы построения деревьев: ID3, C4.5, CART.

7.3.02. Промежуточная аттестация (экзамен)

Пример экзаменационного билета

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

Факультет информационных технологий, Кафедра СМАРТ-технологии
Дисциплина: Методы работы с большими данными
направление подготовки 09.03.01 Информатика и вычислительная техника
БИЛЕТ № 1

1. Персептрон Розенблатта. Теорема об обучении персептрона.

2. Проблемы обучения: ошибка аппроксимации; переобучение; ошибка, связанная со сложностью модели.

Зав. Кафедрой _____ /
_____

Типовые вопросы к экзамену

1. Искусственный интеллект. Основные понятия.
2. Интеллектуальные информационные системы.
3. Интеллектуальный анализ данных и машинное обучение.
4. Постановка задачи машинного обучения.
5. Метрики в задачае регрессии: MAE, MSE, MAPE, R2. Метрики в задаче классификации: кросс-энтропия, precision, recall, F-мера, ROC-кривая, AUC ROC.
6. Обучение линейного алгоритма бинарной классификации образов с помощью градиентного алгоритма
7. Реализация алгоритма метода опорных векторов для задачи бинарной классификации
8. Библиотека Sklearn. Метрические методы классификации Knn
9. Библиотека Sklearn. Кластеризация. Kmeans
10. Библиотека Sklearn. Линейная регрессия
11. Деревья решений. Алгоритмы построения деревьев: ID3, C4.5, CART.
12. Биологические аспекты нервной деятельности. Биологический нейрон.
13. Формальный нейрон Маккалока-Питтса.
14. Персептрон Розенблатта. Теорема об обучении персептрона.
15. Персептронная представляемость.
16. Проблема функции ИСКЛЮЧАЮЩЕЕ ИЛИ.
17. Линейная разделимость. Преодоление проблемы линейной разделимости.
18. Обучение с учителем: классификация образов.
19. Обучение с учителем: аппроксимация многомерных функций
20. Алгоритм обратного распространения ошибки.
21. Проблемы обучения: ошибка аппроксимации; переобучение; ошибка, связанная со сложностью модели.
22. Оптимизация размера сети.
23. Адаптивная оптимизации архитектуры сети. Валидация обучения. Ранняя остановка обучения.
24. Предобработка данных. Максимизация энтропии как цель предобработки.
25. Предобработка данных. Кодирование нечисловых переменных.
26. Предобработка данных. Понижение размерности входов методом главных компонент.
27. CRISP-DM — межотраслевой стандарт ведения проектов интеллектуального анализа данных.
28. Современные архитектуры нейронных сетей для решения инженерных задач обработки изображений.

Прикладной вопрос:

1. Перцептрон. Обучение с учителем: классификация образов.
2. Перцептрон. Обучение с учителем: аппроксимация многомерных функций
3. Алгоритм обратного распространения ошибки.
4. Проблемы обучения: ошибка аппроксимации; переобучение; ошибка, связанная со сложностью модели.
5. Оптимизация размера сети.
6. Адаптивная оптимизации архитектуры сети. Валидация обучения. Ранняя остановка обучения.
7. Прореживание связей.
8. Сети встречного распространения. Структура сети.