МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Забайкальский государственный университет»
(ФГБОУ ВО «ЗабГУ»)

Энергетический факультет
Кафедра Информатики, вычислительной техники и прикладной математики

УТВЕРЖДАЮ:

Декан факультета

Энергетический факультет

Батухтин Андрей Геннадьевич

«____»_____20____ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.03.01 Python, обработка данных и машинное обучение
на 144 часа(ов), 4 зачетных(ые) единиц(ы)
для направления подготовки (специальности) 09.03.01 - Информатика и вычислительная техника

составлена в соответствии с ФГОС ВО, утвержденным приказом
Министерства образования и науки Российской Федерации от
«___» _____ 20____ г. №____

Профиль – Автоматизированные системы и вычислительные машины в промышленных комплексах (для набора 2024)
Форма обучения: Очная

1. Организационно-методический раздел

1.1 Цели и задачи дисциплины (модуля)

Цель изучения дисциплины:

Подготовить студентов к разработке алгоритмов и систем анализа данных и машинного обучения, используя язык программирования Python, и сформировать компетенции, необходимые для применения методов анализа данных, машинного обучения и искусственного интеллекта в автоматизированных системах и программно-аппаратных комплексах

Задачи изучения дисциплины:

Ознакомить студентов с основными инструментами и библиотеками Python для анализа данных и машинного обучения. Развить навыки предобработки, визуализации и анализа данных, включая использование методов статистического анализа. Научить проектировать, реализовывать и оценивать модели машинного обучения для решения прикладных задач. Продемонстрировать интеграцию алгоритмов машинного обучения в программно-аппаратные комплексы. Подготовить студентов к освоению более сложных тем в области нейронных сетей.

1.2. Место дисциплины (модуля) в структуре ОП

Курс "Python, обработка данных и машинное обучение" рассчитан на 16 лекций и 16 практик. Успешное прохождение курса важно для изучения курса Нейросетевого анализа данных. Курс будет проводится на 3 году обучения для студентов которые знакомы с программированием на языках Pascal и C++. Знакомы с ООП. Изучали теориею вероятностей и математическую статистику, математический анализ, численные методы и линейную алгебру.

1.3. Объем дисциплины (модуля) с указанием трудоемкости всех видов учебной работы

Общая трудоемкость дисциплины (модуля) составляет 4 зачетных(ые) единиц(ы), 144 часов.

Виды занятий	Семестр 5	Всего часов
Общая трудоемкость		144
Аудиторные занятия, в т.ч.	68	68
Лекционные (ЛК)	34	34
Практические (семинарские) (ПЗ, СЗ)	0	0
Лабораторные (ЛР)	34	34
Самостоятельная работа	40	40

студентов (СРС)		
Форма промежуточной аттестации в семестре	Экзамен	36
Курсовая работа (курсовой проект) (КР, КП)		

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Планируемые результаты освоения образовательной программы		Планируемые результаты обучения по дисциплине
Код и наименование компетенции	Индикаторы достижения компетенции, формируемые в рамках дисциплины	Дескрипторы: знания, умения, навыки и (или) опыт деятельности
ПК-5	Способен применять методы анализа данных, машинного обучения и искусственного интеллекта при разработке информационных систем	Знать: Знания: Основные этапы анализа данных: сбор,предобработка,визуализация и интерпретация. Принципы работы алгоритмов машинного обучения (регрессия, классификация, кластеризация). Методы оценки качества моделей машинного обучения (метрики, валидация, переобучение). Основные библиотеки Python для анализа данных и машинного обучения (NumPy, pandas, Matplotlib, scikit-learn и др.). Уметь: Умения: Применять Python для анализа и обработки данных (загрузка, преобразование, очистка и визуализация). Разрабатывать и реализовывать алгоритмы машинного обучения для анализа данных. Настраивать гиперпараметры моделей для улучшения их производительности. Выполнять интерпретацию результатов анализа и машинного обучения для принятия решений.

		Владеть: Работа с реальными наборами данных, включая их сбор и обработку. Интеграция алгоритмов анализа данных в простые информационные системы. Написание структурированного и документированного кода для анализа данных.
ПК-7	Способен разрабатывать алгоритмы анализа данных и искусственного интеллекта для программно-аппаратных комплексов	Знать: Принципы разработки алгоритмов анализа данных и их особенности в контексте программно-аппаратных систем. Подходы к оптимизации производительности алгоритмов машинного обучения для внедрения в ограниченных вычислительных средах. Основы взаимодействия Python с оборудованием (например, через библиотеки, такие как PySerial, pyModbus). Уметь: Разрабатывать алгоритмы анализа данных с учетом ограничений аппаратного обеспечения. Прототипировать программные модули, интегрирующие машинное обучение в программно-аппаратные комплексы. Оценивать производительность алгоритмов и проводить их оптимизацию для внедрения в промышленные системы. Владеть: Разработка кода, совместимого с программно-аппаратными комплексами. Реализация и тестирование систем, использующих машинное обучение для автоматизации процессов. Использование современных инструментов для создания

		производительных и надежных решений.

3. Содержание дисциплины

3.1. Разделы дисциплины и виды занятий

3.1 Структура дисциплины для очной формы обучения

Модуль	Номер раздела	Наименование раздела	Темы раздела	Всего часов	Аудиторные занятия			СРС
					ЛК	ПЗ (СЗ)	ЛР	
1	1.1	Основы Python и работа с данными	Основы синтаксиса Python, работа с файлами, библиотеками NumPy и pandas для обработки данных.	26	8	0	8	10
	1.2	Визуализация данных	Библиотеки Matplotlib и Seaborn, создание графиков, интерпретация визуализации.	12	2	0	4	6
	1.3	Предобработка данных	Методы очистки, заполнения пропущенных данных, нормализации и масштабирования.	14	4	0	4	6
2	2.1	Введение в машинное обучение	Принципы машинного обучения, основные термины, библиотеки scikit-learn.	10	2	0	2	6
	2.2	Модели классификации и регрессии	Методы линейной регрессии, логистической регрессии, деревья решений и SVM.	26	10	0	10	6
	2.3	Метрики качества, переобучение, регуляризация, настройка ги	Метрики (MAE, MSE, F1-score), кросс-валидация. Метрики качества, переобучение, регуляризация,	16	6	0	4	6

		перпараметро в.	настройка гиперпараметров.					
Итого				104	32	0	32	40

3.2. Содержание разделов дисциплины

3.2.1. Лекционные занятия, содержание и объем в часах

Модуль	Номер раздела	Тема	Содержание	Трудоемкость (в часах)
1	1.1	Основы синтаксиса Python, работа с файлами, библиотеками NumPy и pandas для обработки данных.	Введение в Python, структуры данных, работа с таблицами.	8
	1.2	Библиотеки Matplotlib и Seaborn, создание графиков, интерпретаци я визуализации.	Основные типы графиков и их применение.	2
	1.3	Методы очистки, заполнения пропущенных данных, нормализации и масштабиро вания.	Методы очистки, заполнения пропущенных данных, нормализации и масштабирования. Этапы предобработки данных.	4
2	2.1	Принципы машинного обучения, основные термины, библиотеки scikit-learn.	Концепции обучения с учителем и без учителя. Принципы машинного обучения, основные термины, библиотеки scikit-learn.	2
	2.2	Методы линейной	Методы линейной регрессии, логистической регрессии, деревья	10

		регрессии, логистической регрессии, деревья решений и SVM.	решений и SVM. Теория работы моделей.	
	2.3	Метрики качества, переобучение, регуляризация, настройка гиперпараметров.	Метрики (MAE, MSE, F1-score), кросс-валидация. Метрики качества, переобучение, регуляризация, настройка гиперпараметров.	6

3.2.2. Практические занятия, содержание и объем в часах

Модуль	Номер раздела	Тема	Содержание	Трудоемкость (в часах)

3.2.3. Лабораторные занятия, содержание и объем в часах

Модуль	Номер раздела	Тема	Содержание	Трудоемкость (в часах)
1	1.1	Основы синтаксиса Python, работа с файлами, библиотеками NumPy и pandas для обработки данных.	Решение задач на манипуляцию данными.	8
	1.2	Библиотеки Matplotlib и Seaborn, создание графиков, интерпретация визуализации.	Построение графиков по данным из практических задач.	4
	1.3	Методы очистки, заполнения	Решение задач на предобработку реальных данных.	4

		пропущенных данных, нормализации и масштабирования.		
2	2.1	Принципы машинного обучения, основные термины, библиотеки scikit-learn.	Установка scikit-learn, применение базовых алгоритмов.	2
	2.2	Методы линейной регрессии, логистической регрессии, деревья решений и SVM.	Методы линейной регрессии, логистической регрессии, деревья решений и SVM. Реализация моделей на реальных данных.	10
	2.3	Метрики качества, переобучение, регуляризация, настройка гиперпараметров.	Метрики (MAE, MSE, F1-score), кросс-валидация. Метрики качества, переобучение, регуляризация, настройка гиперпараметров.	4

3.3. Содержание материалов, выносимых на самостоятельное изучение

Модуль	Номер раздела	Содержание материалов, выносимого на самостоятельное изучение	Виды самостоятельной деятельности	Трудоемкость (в часах)
1	1.1	Основы синтаксиса Python, работа с файлами, библиотеками NumPy и pandas для обработки данных. Решение задач на манипуляцию данными.	Создание компьютерной программы	12
	1.2	Визуализация реальных наборов данных.	Создание компьютерных программ	6
	1.3	Предобработка больших	Создание компьютерных	8

		датасетов из открытых источников.	программ	
2	2.1	Обзор задач, решаемых машинным обучением, подготовка данных.	Создание компьютерных программ	6
	2.2	Использование различных моделей для данных из открытых источников.	Создание компьютерных программ	14
	2.3	Метрики (MAE, MSE, F1-score), кросс-валидация. Метрики качества, переобучение, регуляризация, настройка гиперпараметров.	Создание компьютерных программ	6

4. Фонд оценочных средств для проведения текущей и промежуточной аттестации обучающихся по дисциплине

Фонд оценочных средств текущего контроля и промежуточной аттестации по итогам освоения дисциплины представлен в приложении.

Фонд оценочных средств

5. Учебно-методическое и информационное обеспечение дисциплины

5.1. Основная литература

5.1.1. Печатные издания

1.

5.1.2. Издания из ЭБС

1. Федоров, Д. Ю. Программирование на python : учебное пособие для вузов / Д. Ю. Федоров. — 6-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 187 с. — (Высшее образование). — ISBN 978-5-534-19666-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/556864
2. 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

5.2. Дополнительная литература

5.2.1. Печатные издания

1.
2.

5.2.2. Издания из ЭБС

1.
2.

5.3. Базы данных, информационно-справочные и поисковые системы

Название	Ссылка
https://urait.ru/	https://urait.ru/

6. Перечень программного обеспечения

Программное обеспечение общего назначения: ОС Microsoft Windows, Microsoft Office, ABBYY FineReader, ESET NOD32 Smart Security Business Edition, Foxit Reader, АИБС "МегаПро".

Программное обеспечение специального назначения:

1) Anaconda

2) Apache OpenOffice

3) Atom

4) JetBrain PyCharm

5) LibreOffice

6) Python

7. Материально-техническое обеспечение дисциплины

Наименование помещений для проведения учебных занятий и для самостоятельной работы обучающихся	Оснащенность специальных помещений и помещений для самостоятельной работы
Учебные аудитории для проведения занятий лекционного типа	Состав оборудования и технических средств обучения указан в паспорте аудитории, закрепленной расписанием по факультету
Учебные аудитории для проведения лабораторных занятий	
Учебные аудитории для промежуточной	

аттестации	
Учебные аудитории для проведения групповых и индивидуальных консультаций	Состав оборудования и технических средств обучения указан в паспорте аудитории, закрепленной расписанием по кафедре
Учебные аудитории для текущей аттестации	

8. Методические рекомендации по организации изучения дисциплины

Лекционные занятия. Рекомендуется вести краткий конспект со слайдов во время занятия. Перед занятием рекомендуется повторить пройденный материал, с использованием конспектов, слайдов лекций, рекомендованных и самостоятельно найденных источников; познакомится с планом предстоящего занятия. Подготовить вопросы по пройденным темам.
Лабораторные занятия. Цель занятий – углубление и закрепление теоретических знаний, формирование у них определенных умений и навыков. Перед занятием студент должен выполнить задания. Подготовится к защите работы, знать ответы на вопросы к заданию. Рекомендуется доработать конспект. Желательно использовать несколько источников.
Самостоятельная работа. Приводит студента к получению нового знания, упорядочению и углублению имеющихся знаний, формированию у него профессиональных навыков и умений.
Рекомендуется с самого начала освоения курса работать с рекомендуемыми источниками и предлагаемыми заданиями в форме подготовки к очередному аудиторному занятию, дорабатывать конспект во время самостоятельной работы.
При этом актуализируются имеющиеся знания, а также создается база для усвоения нового материала, возникают вопросы, ответы на которые студент получает в аудитории.
По заданиям, предлагаемым для решения на лабораторных занятиях, студент должен отчитаться до наступления сессии.

Разработчик/группа разработчиков:
Сергей Владимирович Ветров

Типовая программа утверждена

Согласована с выпускающей кафедрой
Заведующий кафедрой
_____ «___»_____20____г.