

Государственное образовательное учреждение
«Приднестровский государственный университет им. Т.Г. Шевченко»

Физико-технический институт

Физико-математический факультет

Кафедра высшей и прикладной математики и информатики

УТВЕРЖДАЮ

Директор физико-технического института

Д.Н. Калошин

2024 г.



РАБОЧАЯ ПРОГРАММА

по дисциплине

Б1.В.09 «Интеллектуальный анализ данных»

на 2025/2026 учебный год

Направление

01.03.02 «Прикладная математика и информатика»

Профиль

Системное программирование и компьютерные технологии

Квалификация

Бакалавр

Форма обучения

Очная

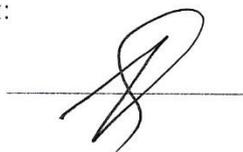
ГОД НАБОРА 2022

Тирасполь 2024 г.

Рабочая программа дисциплины «Интеллектуальный анализ данных» разработана в соответствии с требованиями Государственного образовательного стандарта ВО по направлению подготовки 01.03.02 «Прикладная математика и информатика» и основной профессиональной образовательной программы (учебного плана) по профилю подготовки «Системное программирование и компьютерные технологии».

Составитель рабочей программы:

Доцент



Коровай А.В.

Рабочая программа утверждена на заседании кафедры высшей и прикладной математики и информатики

« 30 » 08 2024 г. протокол № 1

Зав. кафедрой, отвечающей за реализацию дисциплины

« 30 » 08 2024 г.



Коровай А.В.

Зав. выпускающей кафедрой высшей и прикладной математики и информатики

« 30 » 08 2024 г.



Коровай А.В.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Интеллектуальный анализ данных» является формирование у студентов теоретических знаний, практических умений и навыков по применению методов интеллектуального анализа данных с использованием современных программных средств.

Задачами освоения дисциплины «Интеллектуальный анализ данных» являются:

- овладение теоретическими основами наиболее распространенных методов интеллектуального анализа данных и условиями их применения;
- изучение основных методов поиска в данных внутренних закономерностей, взаимосвязей, тенденций;
- изучение основ количественных методов оценки адекватности и точности построенных моделей;
- изучение принципов работы программных средств, предназначенных для решения задач интеллектуального анализа данных;
- формирование практических навыков применения программных средств, применяемых для интеллектуального анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина Б1.В.09 «Интеллектуальный анализ данных» относится к дисциплинам вариативной части Блока 1.

Для освоения дисциплины «Интеллектуальный анализ данных» обучающиеся используют знания, умения и навыки, сформированные в ходе изучения дисциплин «Математический анализ», «Алгебра и аналитическая геометрия», «Теория вероятностей и математическая статистика», «Введение в компьютерный анализ данных».

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Изучение дисциплины направлено на формирование компетенций, приведенных в таблице ниже

Категория (группа) компетенций	Код и наименование	Код и наименование индикатора достижения компетенции
<i>Универсальные компетенции и индикаторы их достижения</i>		
Учебным планом не предусмотрены		
<i>Общепрофессиональные компетенции и индикаторы их достижения</i>		
Учебным планом не предусмотрены		
<i>Обязательные профессиональные компетенции и индикаторы их достижения</i>		
	ПК-1 Способен демонстрировать общенаучные базовые знания естественных наук, математики и информатики, понимание основных фактов, концепций, принципов теорий, связанных с прикладной математикой и информатикой	ИД-1 _{ПК-1} Обладает базовыми знаниями, полученными в области математических и (или) естественных наук, программирования и информационных технологий.
		ИД-2 _{ПК-1} Умеет находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в области математических и (или) естественных наук, программирования и информационных технологий.
		ИД-3 _{ПК-1} Имеет практический опыт научно-исследовательской деятельности в области

		математических и (или) естественных наук, программирования и информационных технологий.
ПК-2 Способен понимать и применять в исследовательской и прикладной деятельности современный математический аппарат.	ИД-1 _{ПК-2} Знает современный математический аппарат.	
	ИД-2 _{ПК-2} Умеет применять методы, алгоритмы и приёмы современного математического аппарата.	
ПК-5 Способен разрабатывать и применять алгоритмические и программные решения в области системного и прикладного программного обеспечения.	ИД-3 _{ПК-2} Владеет практическими навыками применения современного математического аппарата в исследовательской и прикладной деятельности.	
	ИД-1 _{ПК-5} Знает разработку архитектуры, алгоритмических и программных решений системного и прикладного программного обеспечения.	
	ИД-2 _{ПК-5} Умеет использовать языки программирования, алгоритмы, библиотеки и пакеты программ, продукты системного и прикладного программного обеспечения.	
	ИД-3 _{ПК-5} Владеет навыками решения практических задач с применением языков программирования, алгоритмов, библиотек и пакетов программ.	
Рекомендуемые профессиональные компетенции и индикаторы их достижения		
Учебным планом не предусмотрены		

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Распределение трудоемкости в з.е./часах по видам аудиторной и самостоятельной работы студентов по семестрам

Семестр	Трудоемкость, з.е./часы	Количество часов					Самостоятельная работа (СР)	Форма контроля
		В том числе						
		Аудиторных						
		Всего	Лекций (Л)	Практических занятий (ПЗ)	Лабораторных занятий (ЛЗ)			
7	4/144	72	36		36	36	экзамен/36	
Итого:	4/144	72	36		36	36	36	

4.2. Распределение видов учебной работы и их трудоемкости по разделам дисциплины

№ раздела	Наименование разделов	Всего	Аудиторная работа			СР
			Л	ПЗ	ЛЗ	
1	Введение в интеллектуальный анализ данных. Инструментальные средства интеллектуального анализа данных.	18	4		6	8
2	Предварительная обработка данных.	8	2		4	2

№ раздела	Наименование разделов	Всего	Аудиторная работа			СР
			Л	ПЗ	ЛЗ	
3	Техники Data Mining: классификация, кластеризация, прогнозирование и визуализация.	64	24		20	20
4	Нейросетевые алгоритмы.	18	6		6	6
Итого:		108	36		36	36

4.3. Тематический план по видам учебной деятельности

Лекции

№ п/п	Номер раздела дисциплины	Объем часов	Тема лекции	Учебно-наглядные пособия
<i>Введение в интеллектуальный анализ данных.</i>				
<i>Инструментальные средства интеллектуального анализа данных</i>				
1	1	2	Основы понятия и определения интеллектуального анализа данных. Области применения Data Mining. Этапы интеллектуального анализа данных. Основные задачи анализа данных. Классификация методов анализа данных.	презентация
2	2	2	Основные инструменты Data Mining. Библиотеки Python для анализа данных.	презентация
Итого по разделу часов:		4		
<i>Предварительная обработка данных</i>				
3	2	2	Подготовка данных для анализа. Разведочный анализ данных (EDA). Очистка данных. Преобразование данных. Масштабирование.	презентация
Итого по разделу часов:		2		
<i>Техники Data Mining: классификация, кластеризация, прогнозирование и визуализация</i>				
4	3	2	Постановка задачи линейной регрессии. Метод наименьших квадратов. Метод градиентного спуска.	презентация
5		2	Метрики качества линейных регрессионных моделей. Переобучение и недообучение. Основные виды регуляризации: Ridge, Lasso, ElasticNet.	презентация
6		2	Задача классификации. Процесс классификации. Бинарная классификация и метрики качества ее моделей.	презентация
7		2	Метод k ближайших соседей.	презентация
8		2	Бинарная логистическая регрессия. Многоклассовая логистическая регрессия.	презентация
9		2	Метод опорных векторов.	презентация
10		2	Байесовский классификатор.	презентация
11		2	Деревья решений.	презентация
12	2	Ансамблевые модели и их типы. Бэггинг. Случайные леса. Бустинг.	презентация	

13		2	Постановка задачи кластерного анализа. Иерархические методы кластеризации. Агломеративная иерархическая кластеризация.	презентация
14		2	Метод k-средних. Метод DBSCAN. Метрики качества кластеризации.	презентация
15		2	Анализ временных рядов. Моделирование и построение прогноза.	презентация
Итого по разделу часов:		24		
<i>Нейросетевые алгоритмы</i>				
16	4	2	Искусственные нейронные сети: основные понятия. Классификация нейронных сетей. Архитектуры нейронных сетей. Функции активации.	презентация
17		2	Однослойная нейронная сеть. Обучение нейронной сети. Многослойная нейронная сеть. Метод обратного распространения ошибки.	презентация
18		2	Сверточные нейронные сети. Рекуррентные нейронные сети.	презентация
Итого по разделу часов:		6		
Итого:		36		

Лабораторные занятия

№ п/п	Номер раздела дисциплины	Объем часов	Тема лабораторных занятий	Учебно-наглядные пособия
<i>Введение в интеллектуальный анализ данных.</i>				
<i>Инструментальные средства интеллектуального анализа данных</i>				
1	1	2	Библиотека NumPy.	метод.рекомендации
2		2	Библиотека Matplotlib.	метод.рекомендации
3		2	Библиотека Pandas.	метод.рекомендации
Итого по разделу часов:		6		
<i>Предварительная обработка данных</i>				
4	2	4	Разведочный анализ данных.	метод.рекомендации
Итого по разделу часов:		4		
<i>Техники Data Mining: классификация, кластеризация, прогнозирование и визуализация</i>				
5	3	2	Линейная регрессия.	метод.рекомендации
6		2	Регуляризация линейной регрессии.	метод.рекомендации
7		2	Метод k ближайших соседей.	метод.рекомендации
8		2	Логистическая регрессия.	метод.рекомендации
9		2	Метод опорных векторов.	метод.рекомендации
10		2	Байесовский классификатор.	метод.рекомендации
11		2	Деревья решений. Случайный лес.	метод.рекомендации
12		2	Методы кластеризации.	метод.рекомендации
13		2	Метод k-средних. Метод DBSCAN.	метод.рекомендации
14		2	Анализ и прогнозирование временных рядов.	метод.рекомендации

Итого по разделу часов:	20			
<i>Нейросетевые алгоритмы</i>				
15	4	2	Построение и обучения однослойного персептрона.	метод.рекомендации
16		4	Классификация изображений на основе нейронных сетей.	метод.рекомендации
Итого по разделу часов:		6		
Итого:		36		

Самостоятельная работа обучающегося

Раздел дисциплины	№ п/п	Тема и вид самостоятельной работы обучающегося	Трудоемкость (в часах)
Раздел 1	1	Основы понятия и определения интеллектуального анализа данных. Этапы интеллектуального анализа данных. (1,2)	1
	2	Основные задачи анализа данных. Классификация методов анализа данных. (1,2)	1
	3	Библиотека NumPy. (1,2,3)	2
	4	Библиотека Matplotlib. (1,2,3)	2
	5	Библиотека Pandas. (1,2,3)	2
Итого по разделу часов:			8
Раздел 2	1	Подготовка данных для анализа. Разведочный анализ данных. Преобразование данных. (1,2,3)	2
Итого по разделу часов:			2
Раздел 3	1	Линейная и нелинейная регрессия. Метод наименьших квадратов. Метод градиентного спуска.	2
	2	Метрики качества линейных регрессионных моделей. Переобучение и недообучение. Регуляризация.	2
	3	Задача классификации. Процесс классификации. Бинарная классификация и метрики качества ее моделей. (1,2,3)	2
	4	Метод k ближайших соседей. (1,2,3)	2
	5	Логистическая регрессия (1,2,3)	2
	6	Метод опорных векторов. (1,2,3)	2
	7	Байесовский классификатор. (1,2,3)	2
	8	Деревья решений. Ансамблевые модели. Бэггинг. Случайные леса. Бустинг. (1,2,3)	2
	9	Кластерный анализ. Иерархические и неиерархические методы кластеризации. Метод k-средних. Метод DBSCAN. (1,2,3)	2
	10	Анализ временных рядов. Моделирование и построение прогноза.	2
Итого по разделу часов:			20
Раздел 4	1	Искусственные нейронные сети: основные понятия. Классификация нейронных сетей. Архитектуры нейронных сетей. (1,2)	1
	2	Однослойная нейронная сеть. Обучение нейронной сети. (1,2,3)	1
	3	Многослойная нейронная сеть. Метод обратного распространения ошибки. (1,2,3)	2
	4	Сверточные нейронные сети. Рекуррентные нейронные сети. (1,2)	2
Итого по разделу часов:			6
Итого:			36

Примечание:

- 1 – проработка учебного материала (по конспектам лекций, учебной и научной литературе);
 2 – подготовка к лабораторным занятиям, к тестированию;
 3 – выполнение индивидуальных заданий.

5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ ПРОЕКТОВ (РАБОТ)

По данной дисциплине курсовые проекты не предусмотрены.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Обеспечение обучающихся учебниками, учебными пособиями

№ п/п	Наименование учебника, учебного пособия	Автор	Год издания	Кол-во экз.	Электр. версия	Место размещения электронной версии
Основная литература						
1	Интеллектуальный анализ данных: учебное пособие	Замятин А. В.	2020	–	+	http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000722107
2	Базовые алгоритмы машинного обучения на языке Python: учебно-методическое пособие	Долганов А. Ю., Ронкин М. В., Созыкин А. В.	2023	–	+	https://elar.urfu.ru/handle/10995/122740
3	Основы машинного обучения: учебное пособие	Лимановская О.В., Алферьева Т.И.	2020	–	+	https://elar.urfu.ru/handle/10995/88687
4	Искусственные нейронные сети и приложения: учеб. пособие	Гафаров Ф.М., Галимянов А.Ф.	2018	–	+	https://kpfu.ru/staff_files/F1493580427/NejronGafGaf.pdf
Дополнительная литература						
1	Интеллектуальный анализ данных: учебное пособие	Сапрыкин О.Н.	2020	–	+	https://e.lanbook.com/book/188906
2	Обработка данных в Python. Основы работы с библиотекой Pandas: учебно-методическое пособие	Титов А. Н., Тагиева Р. Ф.	2022	–	+	https://e.lanbook.com/book/331013
3	Визуализация данных в Python. Работа с библиотекой Matplotlib: учебно-методическое пособие	Титов А.Н., Тагиева Р. Ф.	2022	–	+	https://e.lanbook.com/book/331025
4	Кластерный анализ. Python: учеб. пособие	Демидова Л. А.	2022	–	+	https://e.lanbook.com/book/240092
Итого по дисциплине: % печатных изданий – 0%; % электронных – 100%						

6.2. Программное обеспечение и Интернет-ресурсы

1. Google Colab
URL: <https://colab.research.google.com>
2. Информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных
URL: <http://www.machinelearning.ru>
3. Вводный курс машинного обучения (автор Дмитрий Макаров)
URL: https://www.youtube.com/playlist?list=PLX9afROa9MG19JK0JMoy4tIrHvx_z6dHZ
URL: <https://www.dmitrymakarov.ru/intro/>
4. Машинное обучение (автор Сергей Балакирев)
URL: <https://proproprogs.ru/ml>
URL: <https://www.youtube.com/playlist?list=PLA0M1Bcd0w8zxDIDOTQHsX68MCDOAJDtj>
5. YouTube-канал: machine learrning
https://www.youtube.com/@machine_learrning/playlists

6.3. Методические указания и материалы по видам занятий

1. Демидова Л. А. Разведочный анализ данных. Python. Часть 1: учебно-методическое пособие / Л. А. Демидова. – Москва: РТУ МИРЭА, 2022 – 107 с. (URL: <https://e.lanbook.com/book/310970>).
2. Демидова Л. А. Разведочный анализ данных. Python. Часть 2: учебно-методическое пособие / Л. А. Демидова. – Москва: РТУ МИРЭА, 2023 – 92 с. (URL: <https://e.lanbook.com/book/382691>).
3. Виноградов Д. В. Интеллектуальный анализ данных : учеб. пособие / Д. В. Виноградов; Владим. гос. ун-т им. А. Г. и Н. Г. Столетовых. – Владимир: Изд-во ВлГУ, 2021. – 260 с. (URL: <http://dspace.www1.vlsu.ru/handle/123456789/9029>)
4. Коротеев, М. В. Практикум по машинному обучению на Python : учебное пособие / М. В. Коротеев, В. А. Одинцова, Е. С. Плешакова. – Москва : Финансовый университет, 2023. – 161 с. (URL: <https://e.lanbook.com/book/431102>).
5. Коротеев, М. В. Практикум по машинному обучению на Python : учебное пособие / М. В. Коротеев, В. А. Одинцова, Е. С. Плешакова. – Москва : Финансовый университет, 2023 – Часть 2 – 2024. – 77 с. (URL: <https://e.lanbook.com/book/431105>).

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Реализация данной учебной дисциплины осуществляется с использованием материально-технической базы, обеспечивающей проведение всех видов учебных занятий, предусмотренных программой учебной дисциплины и соответствующей действующим санитарным и противопожарным правилам и нормам.

Компьютерные классы оснащены современными персональными компьютерами и программным обеспечением в соответствии с тематикой изучаемого материала. Число рабочих мест достаточно, чтобы обеспечивалась индивидуальная работа студента на отдельном персональном компьютере.

8. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ

Данный курс предполагает овладение студентами теоретическими основами наиболее распространенных методов интеллектуального анализа данных, приобретение ими практических навыков применения программных средств, применяемых для анализа данных. Для успешного изучения материала необходимы знания таких основополагающих понятий как: виды случайных

величин, характеристики случайных величин, основные законы распределения случайных величин, способы оценки выборочных характеристик, основы статистического вывода с использованием статистических критериев и проверки гипотез. Полученные в рамках курса знания являются основой для принятия эффективных управленческих решений в профессиональной деятельности.

Различные виды учебных занятий: лекции и лабораторные занятия – тесно связаны друг с другом. Материал, законспектированный на лекциях, необходимо регулярно дополнять сведениями из литературных и интернет-источников, представленных в рабочей программе дисциплины. Подготовка к лабораторным занятиям заключается в предварительном изучении соответствующего материала по конспекту лекций или по рекомендованной литературе.

Пропуск занятий, невыполнение лабораторных работ или неуспевание материала требуют компенсации путем самостоятельной работы студента. В случае необходимости следует обращаться к преподавателю за консультацией.

Планирование времени на самостоятельную работу, необходимого на изучение настоящей дисциплины, студентам лучше всего осуществлять на весь семестр, предусматривая при этом регулярное повторение пройденного материала.

9. ТЕХНОЛОГИЧЕСКАЯ КАРТА ДИСЦИПЛИНЫ

Курс 4 группа ФМ22ДР62ПФ (403) семестр 7

Преподаватель-лектор: *ст. преподаватель Калинин Е.В.*

Преподаватель, ведущий лабораторные занятия: *ст. преподаватель Калинин Е.В.*

Кафедра *Высшей и прикладной математики и информатики*

Семестр	Количество часов						Форма контроля
	Трудо- емкость, з.е./часы	В том числе					
		Аудиторных				СР	
		Всего	Л	ПЗ	ЛЗ		
7	4/144	72	36		36	36	экзамен/36

Форма текущей аттестации	Расшифровка	Минимальное кол-во баллов	Максимальное кол-во баллов
Посещение лекционных занятий		0	10
Выполнение и защита лабораторных работ	За каждую лабораторную работу от 2 до 3 баллов	0	45
Тестирование		0	15
Итого количество баллов по текущей аттестации:		45	70
Промежуточная аттестация	Экзамен	10	30
Итого по дисциплине:		55	100