

**Г О У В П О Российско-Армянский (Славянский)
университет**

Утверждено
Директор Института Математики и Информатики



Арамян Р.Г.

« 21 » марта 2025, протокол № 9.1

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС ДИСЦИПЛИНЫ

Наименование дисциплины: Data Mining

Автор: Аветисян Карен Ишханович

Направление подготовки: 01.04.02 Прикладная математика и информатика

Наименование образовательной программы: Искусственный интеллект и машинное обучение (Artificial Intelligence and Machine Learning)

1. АННОТАЦИЯ

1.1. Краткое описание содержания данной дисциплины

Целью освоения дисциплины «Data Mining» является формирование представления о типах задач, возникающих в области интеллектуального анализа данных (Data Mining) и методах их решения, которые помогут обучающимся выявлять, формализовать и успешно решать практические задачи анализа данных, возникающие в процессе их профессиональной деятельности. В ходе изучения дисциплины перед обучающимися ставятся следующие задачи:

- изучение методов и моделей Data Mining;
- изучение алгоритмов классификации и регрессии;
- изучение алгоритмов поиска ассоциативных правил;
- изучение методов кластеризации.

1.2. Трудоемкость в академических кредитах и часах, формы итогового контроля

Трудоемкость курса составляет 5 академических кредитов, что эквивалентно **180** часам (1 академический кредит равен 36 часам). Итоговый контроль будет проведен в форме **экзамена**.

1.3. Взаимосвязь дисциплины с другими дисциплинами учебного плана специальности

При изучении дисциплины «Data Mining» используются понятия и методы структур данных, нейронных сетей и Big Data.

1.4. Результаты освоения программы дисциплины:

Код компетенции	Наименование компетенции	Код индикатора достижения компетенций	Наименование индикатора достижений компетенций
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	УК-1.1	Знать теории и методы критического анализа и системного подхода
		УК-1.2	Уметь применять методы поиска, анализа и синтеза информации для решения конкретных задач

		УК-1.3	Владеть способностью выработки стратегических решений на основе комплексного анализа ситуации и прогнозирования исходов
УК-2	Способен управлять проектом на всех этапах его жизненного цикла	УК-2.1	Знать методы управления проектами, включая планирование, исполнение, контроль и завершение проектов
		УК-2.2	Уметь координировать различные этапы проекта, обеспечивая соблюдение сроков и качества выполнения
		УК-2.3	Владеть умениями по эффективному руководству проектными командами и решению конфликтных ситуаций на всех этапах реализации проекта
ПК-2	Способностью разрабатывать и анализировать концептуальные и теоретические модели решаемых научных проблем и задач	ПК-2.1	Знать теоретические основы и концептуальные модели в научных исследованиях
		ПК-2.2	Уметь анализировать и разрабатывать концептуальные и теоретические модели для проектной и производственной деятельности
		ПК-2.3	Владеть глубоким пониманием теоретических подходов и их практического применения в создании новых исследовательских проектов

ПК-13	Способностью осознавать корпоративную политику в области повышения социальной ответственности бизнеса перед обществом, принимать участие в ее развитии	ПК-13.1	Знать корпоративную политику и стандарты социальной ответственности
		ПК-13.2	Уметь участвовать в разработке и реализации корпоративной политики
		ПК-13.3	Владеть лидерскими качествами для продвижения социальной ответственности и устойчивого развития в бизнесе

2. УЧЕБНАЯ ПРОГРАММА

2.1. Трудоемкость дисциплины и виды учебной работы

Виды учебной работы	Всего, в акад. часах	Распределение по семестрам					
		— сем	— сем	3 — сем	— сем.	— сем	— сем.
1	2	3	4	5	6	7	8
1.Общая трудоемкость изучения дисциплины по семестрам, в т. ч.:	32			32			
1.1.Аудиторные занятия, в т. ч.:	32			32			
1.1.1.Лекции							
1.1.2.Практические занятия, в т. ч.	32			32			
Итоговый контроль				Экзамен			

2.2. Содержание дисциплины

2.2.1. Тематический план и трудоемкость аудиторных занятий (модули, разделы дисциплины и виды занятий) по рабочему учебному плану

Разделы и темы дисциплины	Всего часов	Лекции, часов	Практ. занятия, часов	Семинары, часов	Лабор, часов	Другие виды занятий, часов
1	2	3	4	5	6	7
II курс	32		32			
<i>МОДУЛЬ 1</i>	32		32			

Введение в Data Mining	1		1			
Задачи Data Mining: интеллектуальный анализ данных	2		2			
Инструменты Data Mining	2		2			
Задачи классификации	4		4			
Линейные модели для классификации и регрессии	4		4			
Методы снижения размерности пространства	4		4			
Кластерный анализ	4		4			
Поиск ассоциативных правил	4		4			
Рекомендательные системы	4		4			
Методы классификации и прогнозирования	3		3			
ИТОГО	32		32			

2.2.2. Краткое содержание разделов дисциплины в виде тематического плана

Тема 1. Введение в Data Mining

- **Обзор дисциплины:** Представление основных понятий и истории развития Data Mining. Освещение целей и значимости интеллектуального анализа данных в современных условиях.
- **Задачи Data Mining:** Обсуждение основных задач, которые решает Data Mining, включая классификацию, кластеризацию, и поиск ассоциативных правил.
- **Инструменты Data Mining:** Перечень и характеристика основных инструментов и программных решений, используемых для анализа данных.

Тема 2. Задачи классификации

- **Общие принципы классификации:** Определение классификации и её роль в анализе данных. Примеры применения классификации в различных областях.
- **Линейные модели для классификации и регрессии:** Изучение линейных моделей, их свойств и областей применения. Обсуждение регрессионного анализа как метода предсказания.

Тема 3. Методы снижения размерности пространства

- **Понятие и цели снижения размерности:** Обзор методов снижения размерности, таких как PCA (анализ главных компонент) и t-SNE. Применение этих методов для упрощения моделей и улучшения интерпретации данных.

Тема 4. Кластерный анализ

- **Основы кластеризации:** Описание методов кластерного анализа, включая K-means и иерархическую кластеризацию. Практические примеры использования кластеризации для сегментации рынка и анализа социальных сетей.

Тема 5. Поиск ассоциативных правил

- **Методы и применение:** Изучение методов нахождения ассоциативных правил в больших наборах данных. Примеры использования ассоциативных правил для анализа потребительского поведения.

Тема 6. Рекомендательные системы

- **Основные алгоритмы и подходы:** Рассмотрение механизмов работы рекомендательных систем, включая коллаборативную фильтрацию и контент-ориентированные системы. Примеры применения в электронной коммерции и медиа.

Тема 7. Методы классификации и прогнозирования

- **Расширенные техники классификации:** Обсуждение более сложных методов классификации, таких как случайные леса, градиентный бустинг и нейронные сети. Применение этих методов для прогнозирования в различных секторах экономики.

2.2.3. Материально-техническое обеспечение дисциплины

Доска:

Традиционная черная или маркерная доска для визуального представления учебного материала, формул и схем во время лекций и семинаров.

Проектор:

Современный проектор, подключаемый к компьютеру для демонстрации презентаций, видеоматериалов и программного кода. Проектор необходим для визуализации сложных концепций и детализации примеров анализа данных.

2.3. Модульная структура дисциплины с распределением весов по формам контролей

Формы контролей	Вес формы (форм) текущего контроля в результирующей оценке текущего контроля (по модулям)		Вес формы промежуточного контроля в итоговой оценке промежуточного контроля		Вес итоговой оценки промежуточного контроля в результирующей оценке промежуточных контролей		Вес итоговой оценки промежуточного контроля в результирующей оценке промежуточных контролей (семестровой оценке)		Веса результирующей оценки промежуточных контролей и оценки итогового контроля в результирующей оценке итогового контроля
	М1	М2	М1	М2	М1	М2			
Вид учебной работы/контроля									
Контрольная работа <i>(при наличии)</i>				0,7					
Устный опрос <i>(при наличии)</i>									
Тест <i>(при наличии)</i>									
Лабораторные работы <i>(при наличии)</i>									
Письменные домашние задания <i>(при наличии)</i>				0,3					
Реферат <i>(при наличии)</i>									
Эссе <i>(при наличии)</i>									
Проект <i>(при наличии)</i>									
<i>Другие формы (при наличии)</i>									
Веса результирующих оценок текущих контролей в итоговых оценках промежуточных контролей						0,3			
Веса оценок промежуточных контролей в итоговых оценках промежуточных контролей						0,7			

Вес итоговой оценки 1-го промежуточного контроля в результирующей оценке промежуточных контролей								
Вес итоговой оценки 2-го промежуточного контроля в результирующей оценке промежуточных контролей							1	
Вес результирующей оценки промежуточных контролей в результирующей оценке итогового контроля								0,4
Вес итогового контроля (Экзамен/зачет) в результирующей оценке итогового контроля								0,6
	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$

3. Теоретический блок

3.1. Материалы по теоретической части курса

3.1.1. Учебники

- 1) Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004
- 2) Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Технологии анализа данных : Data Mining, Visual Mining, Text Mining, OLAP- СПб.: БХВ-Петербург, 2008
- 3) Чубукова И.А. Data Mining. Учебное пособие – М.: Изд-во «Интернет-университет информационных технологий – ИНТУИТ.ру, 2006
- 4) Jones H. Data Analytics: An Essential Beginner's Guide To Data Mining, Data Collection, Big Data Analytics For Business, And Business Intelligence Concepts. 2018.

4. Фонды оценочных средств (указываются материалы, необходимые для проверки уровня знаний в соответствии с содержанием учебной программы дисциплины).

4.1. Планы практических и семинарских занятий

Оценочные материалы: Включают контрольные работы, тесты, индивидуальные и групповые проекты, которые проверяют знания студентов по ключевым аспектам курса. Также могут использоваться портфолио, включающие все выполненные студентами работы, что позволяет оценить их прогресс на протяжении всего курса.

5. Методический блок

5.1. Методика преподавания

5.1.1. Методические рекомендации для студентов по подготовке к семинарским, практическим или лабораторным занятиям, по организации самостоятельной работы студентов при изучении конкретной дисциплины.

- **Изучение материалов:** Необходимо изучать предоставленные учебные материалы перед занятиями. Это поможет лучше понять сложные концепции и активно участвовать в обсуждениях.
- **Практические задачи:** Регулярно выполнять практические задания для закрепления теоретических знаний. Это включает написание кода, использование программных инструментов и анализ данных.
- **Групповые проекты и обсуждения:** Участвовать в групповых проектах и обсуждениях. Групповая работа способствует развитию коммуникативных навыков и обмену знаниями.