

Аннотация рабочей программы дисциплины

Б1.О.26 «Аналитика данных»

Объем трудоемкости: 2 зачетных единицы

Цель дисциплины: формирование у студентов системного представления о методах и инструментах аналитики данных, применяемых для исследования, интерпретации, визуализации данных и поддержки принятия решений в различных предметных областях.

Задачи дисциплины: 1) освоение современных методологий аналитики данных и инструментов Python для анализа и визуализации; 2) приобретение навыков разведочного анализа данных (EDA), построения признаков и применения методов статистического анализа; 3) формирование умений строить и интерпретировать модели анализа данных, выявлять закономерности и аномалии; 4) освоение инструментов автоматизации аналитических вычислений и визуализации, создание аналитических отчетов и дашбордов.

Место дисциплины в структуре образовательной программы

Дисциплина «Аналитика данных» относится к обязательной части Блока 1 "Дисциплины (модули)" учебного плана.

При освоении данной дисциплины необходимым является предварительное изучение курсов «Теория вероятностей и математическая статистика», «Программирование на Python», «Основы анализа данных». Сведения дисциплины используются при изучении курсов «Интеллектуальные методы анализа данных», «Бизнес-аналитика и большие данные», «Системы поддержки принятия решений».

Требования к уровню освоения дисциплины

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код и наименование индикатора*	Результаты обучения по дисциплине
BD-1 Способен осуществлять поиск сбор очистку и предварительный анализ данных	
BD-1.3 Применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных.	Применяет методы снижения размерности (PCA, t-SNE и др.) для упрощения анализа и визуализации многомерных данных.
BD-1.4 Отбирает признаки данных, значимые для исследования.	Отбирает и обосновывает выбор значимых признаков для последующего анализа и моделирования
BD-2 Способен определять требования к наборам данных для решения задач машинного обучения проводить разметку и анализ наборов данных оценивать качество данных обеспечивать непрерывную интеграцию данных	
BD-2.1 Определяет требования к наборам и качеству данных для решения задач машинного обучения.	Определяет критерии качества и требования к наборам данных для корректного применения методов машинного обучения
ML-2 Способен применять фундаментальные принципы и методы машинного обучения включая подготовку данных оценку качества моделей и работу с признаками	
ML-2.2 Применяет методы предварительной обработки данных и работы с признаками.	Применяет методы очистки, нормализации и кодирования признаков для подготовки данных к анализу.
ML-4 Способен применять методы обучения без учителя для анализа структуры данных и выявления скрытых закономерностей.	
ML-4.1 Применяет алгоритмы кластеризации и понижения размерности для решения практических задач.	Использует алгоритмы кластеризации и снижения размерности для группировки объектов и упрощения анализа данных.
ML-4.2 Выявляет аномалии и применяет методы поиска ассоциативных правил.	Выявляет аномалии в данных и применяет методы поиска ассоциативных правил для анализа взаимосвязей.
O-1 Способен осуществлять управление знаниями в том числе с применением алгоритмов интеллектуального поиска решений и формирования стратегий	

Код и наименование индикатора*	Результаты обучения по дисциплине
О-1.2 Способен преобразовывать неформализованные и слабо-формализованные данные предприятия в семантические единицы баз знаний.	Преобразует неструктурированные и слабо формализованные данные в семантические единицы, пригодные для хранения и анализа в базах знаний.
ОПК-2 Способен применять современный математический аппарат, связанный с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности	
ОПК-2.2 Применяет современный математический аппарат при построении моделей в различных областях человеческой деятельности	ОПК-2.2 Применяет современный математический аппарат при построении моделей в различных областях человеческой деятельности

Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа СРС
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Основы аналитики данных и виды анализа (описательная, диагностическая, предсказательная, предписывающая).	8	2		2	4
2.	Сбор, очистка и подготовка данных.	8	2		2	4
3.	Разведочный анализ данных (EDA) и визуализация.	8	2		2	4
4.	Создание новых признаков, кодирование категориальных переменных, методы отбора информативных признаков.	8	2		2	4
5.	Статистический анализ данных и A/B-тестирование.	8	2		2	4
6.	Методы анализа, тренды, сезонность, прогнозирование.	8	2		2	4
7.	Кластеризация, выявление аномалий, ассоциативные правила.	10	2		2	6
8.	Создание дашбордов, storytelling с данными, подготовка аналитических отчетов.	11,8	2		2	7,8
ИТОГО по разделам дисциплины		69,8	16		16	37,8
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0,2				
Подготовка к текущему контролю						
Общая трудоемкость по дисциплине		72				

Курсовые работы: не предусмотрены

Форма проведения аттестации по дисциплине: зачет.

Автор: старший преподаватель кафедры анализа данных и искусственного интеллекта, Т.М. Панеш.