

Аннотация рабочей программы дисциплины «Основы анализа данных и интеллектуальные системы»

Целью освоения дисциплины «Основы анализа данных и интеллектуальные системы» является формирование комплекса знаний об основах анализа данных и интеллектуальных системах в цифровой экономике.

Задачи дисциплины

–сформировать у обучающихся знания об анализе проблемных ситуаций на основе системного подхода, применения методов анализа данных, математического моделирования и интеллектуальных систем для решения профессиональных задач,

–научить обучающихся проводить анализ данных с применением интеллектуальных систем для извлечения информации и получения новых знаний о закономерностях развития цифровой экономики.

Содержание дисциплины:

Системология анализа данных

Данные. Информация. Знания. Обнаружение в данных ранее неизвестных, нетривиальных, практически полезных и доступных для интерпретации знаний, необходимых для принятия решений в различных сферах цифровой экономики. Категоризация данных.

Основы анализа данных.

Уровни анализа (синтаксический, семантический, прагматический). Типы выявляемых закономерностей: ассоциация (идентификация), последовательность, классификация и распознавание образов, кластеризация, прогнозирование.

Примеры бизнес-приложений Data Mining в различных сферах цифровой экономики.

Этапы анализа данных.

Сбор и ввод (загрузка) данных. Первичная обработка, трансформация и преобразование данных, очистка, борьба с шумами и выбросами, устранение пропусков и ошибок, нормализация данных. Визуализация. Выбор модели для анализа. Тестирование и валидация модели. Интерпретация и представление результатов.

Методы data mining.

Градиентный метод. Алгоритмы построения деревьев решений. Бэггинг. Бустинг, градиентный бустинг. Кластерный анализ, алгоритм k-means. Ассоциативные правила, методы и алгоритмы их поиска.

Анализ данных с применением машинного обучения.

Обучение с учителем, обучение без учителя, обучение с подкреплением, генетический алгоритм. Нейронные и глубокие сети. Минимизация участия человека в выявлении закономерностей.

Интеллектуальные системы, фреймворки автоматического машинного обучения. (AutoMachineLearning): Azure Machine Learning Studio, H2O.ai, RapidMiner, Knime, Trifacta. Анализ данных без написания программного кода людьми, не имеющими специальной математической подготовки. Распределенные облачные вычислительные системы в машинном обучении.

Применение машинного обучения в цифровой экономике.

Кредитный скоринг. Прогнозирование ухода (оттока) клиентов. Биржевой анализ. Анализ рыночной корзины. Финансовый надзор. Прогнозирование качества продукции.

Визуализация и инфографика в анализе данных.

Графическое отображение данных. Визуализации данных с помощью вычислительных машин. Визуализация на различных этапах обработки данных: визуализация исходных данных, выборки, результатов. Интерактивные дашборды. Презентационная, исследовательская и гибридная визуализация данных. Инструменты и библиотеки визуализации данных.

Основы искусственного интеллекта.

Подходы к пониманию и разработке искусственного интеллекта. Исследовательские центры искусственного интеллекта. Развитие искусственного интеллекта в России и мире. Национальная стратегия развития искусственного интеллекта в РФ. Система регулирования общественных отношений в связи с развитием интеллектуальных систем. Индикаторы развития искусственного интеллекта в цифровой экономике.

Перспективы развития анализа данных и интеллектуальных систем.

Анализ слабоструктурированных (полуструктурированных) и неструктурированных данных. Моделирование сложных экономических процессов и их анализа с использованием теории нелинейной динамики и хаоса.

Необходимость разработки стройной теории искусственного интеллекта.

Объем дисциплины – 8 з.е.

Форма промежуточного контроля – экзамены