

## **Аннотация дисциплины**

### **СОВРЕМЕННЫЕ ТЕХНОЛОГИИ ПРИКЛАДНОГО ПРОГРАММИРОВАНИЯ И ОБРАБОТКИ ДАННЫХ**

**Цель дисциплины:** формирование у студентов умений и навыков применения современных технологий прикладного программирования и обработки данных для поиска, обработки, анализа информации.

#### **Краткое содержание:**

Введение в программирование на языке Python. Задачи анализа данных, понятие набора данных (dataset). Подготовительные операции для выполнения анализа данных. Язык программирования Python: основные характеристики, возможности языка для решения задач анализа данных и машинного обучения. Версии языка программирования Python, дистрибутивы и библиотеки Python. Основные синтаксические конструкции Python. Знакомство с типами данных и операциями, переменными. Работа со словарями в Python. Множества. Постановки задач машинного обучения. Объекты и признаки. Типы признаков: бинарные, номинальные, порядковые, количественные. Типы задач машинного обучения: классификация, регрессия, прогнозирование, кластеризация. Примеры задач решаемых методами машинного обучения. Проблема недообучения / переобучения. Обзор технологий хранения данных: файловых систем, реляционных СУБД, OLAP, Data Warehouses, не реляционных (“NoSQL”) баз данных. Сравнительный анализ и области применения различных технологий хранения информации. Работа с файлами. Работа с реляционными базами данных на примере SQLite. Краткий обзор основных видов не реляционных баз данных: хранилищ «ключ-значение», хранилище семейств колонок, документо-ориентированных СУБД, баз данных на основе графов. Сравнительный анализ и области применения не реляционных баз данных. Знакомство с различными классами информационно-аналитических систем. Технологии Data Mining. Технологии анализа больших объемов данных (Big Data): причины возникновения, основные особенности функционирования и специфика создания приложений.