

**Титульный лист материалов по дисциплине**  
(заполняется по каждому виду учебного материала)

ДИСЦИПЛИНА **Модели и методы предиктивной аналитики**  
(полное наименование дисциплины без сокращений)

ИНСТИТУТ **информационных технологий**

КАФЕДРА **Прикладной математики**  
(полное наименование кафедры)

ВИД УЧЕБНОГО  
МАТЕРИАЛА **Практические занятия**  
(в соответствии с пп. I-III)

ПРЕПОДАВАТЕЛЬ **Парамонов Александр Александрович**  
(фамилия, имя, отчество)

СЕМЕСТР **7, 2025-2026**  
(указать семестр обучения, учебный год)

## **Задание 1. Парсинг данных**

Цель:

Сформировать три отдельных набора данных, представляющих собой временные ряды, многомерные данные и наборы текста из открытых источников. Приоритетным инструментом для сбора данных является использование API.

Требования к данным/источникам:

### **Временные ряды.**

Источник данных: открытые API, предоставляющие данные временных рядов (например, финансовые данные, метеорологические данные, данные о заболевании).

Формат данных: даты и соответствующие им значения (например, цена акций, температура, количество случаев заболевания).

### **Многомерные данные.**

Источник данных: открытые API, предоставляющие многомерные данные (например, данные о параметрах транспортных средств, данные о здоровье, спортивная статистика).

Формат данных: набор данных с несколькими измерениями (например, страна, год, показатель здоровья).

### **Наборы текста.**

Источник данных: открытые API, предоставляющие текстовые данные (например, новости, статьи, социальные медиа).

Формат данных: текстовые записи с метаданными (например, дата публикации, автор, заголовок, текст).

Этапы выполнения задания:

#### **1. Исследование и выбор API.**

Найти подходящие API для каждого типа данных. Зарегистрироваться и получить ключи доступа к API, если это требуется.

#### **2. Разработка скриптов для парсинга.**

Написать скрипты на Python, использующие библиотеки requests или http.client для обращения к API. Обработать ответы от API и преобразовать их в удобный для анализа формат (например, pandas DataFrame для временных рядов и многомерных данных, список словарей для текстовых данных).

#### **3. Сохранение данных.**

Сохранить данные в подходящих форматах (например, CSV для временных рядов и многомерных данных, JSON или CSV для текстовых

данных). Обеспечить сохранение метаданных (наличие ссылок на исходные данные).

4. Составление отчёта о работе.

В отчёте описать процесс парсинга, включая используемые API, процесс обработки данных и запросов.

Результат работы:

После выполнения задания у студента должны быть три отдельных набора данных, готовых для дальнейшего анализа и отчёт о проделанной работе.