

УДК 004.8

**АНАЛИЗ РОССИЙСКОГО РЫНКА СПЕЦИАЛИСТОВ
В ОБЛАСТИ АНАЛИТИКИ**

Ларионова А.В. (ИТМО)

**Научный руководитель – кандидат физико-математических наук Бойцев А.А.
(ИТМО)**

Введение. Анализ данных – это процесс исследования, фильтрации, преобразования и моделирования данных с целью извлечения полезной информации. Одной из методологий для работ по исследованию данных является CRISP-DM [1], но основная проблема заключается в том, что не всегда выполняются все необходимые для работы этапы данной методологии. Тем самым, это влияет на качество выводов, сделанных при анализе данных.

Основная часть. Проведение анализа данных в соответствии с методологией CRISP-DM (в частности – определение цели работы, сбор и предобработка данных) решает проблему повышения точности выводов, которые можно сделать о российском рынке специалистов. Сбор данных при помощи API служит для получения наиболее актуальных вакансий от крупнейшей компании в интернет-рекрутменте в России. Предобработка данных позволяет сделать отбор значимых критериев для сравнительного анализа и дальнейшего применения методов машинного обучения. Предположительно, выполнение этих этапов может обеспечить высокое качество выводов о российском рынке специалистов в области аналитики [2].

Выводы. Проведенная работа демонстрирует сравнительный анализ собранных и предобработанных данных – разницу критериев, рассматриваемых в вакансиях на HeadHunter в области аналитики. Также, представлен обзор методов машинного обучения, применимых для данных критериев.

Список использованных источников:

1. Evaluating CRISP-DM for Data Science. Data Science Process Alliance, 2021
2. В. Мионов. Профессия «Бизнес-аналитик». Издательство «Олимп-Бизнес», 2022

Автор: Ларионова А.В.

Научный руководитель: Бойцев А.А.