

ПРОФЕССИЯ DATA SCIENTIST

Столяревская А.Л.

Международный Соломонов университет, г. Харьков

Наше время характеризуется ростом объёма данных, порождаемых во всех областях жизни современного общества. Прогнозируемым количеством данных к 2020 году являются 35 зеттабайт. Стремительный рост данных обеспечивают основные виды средств массовой информации и коммуникации – телевидение, радио, печать, интернет, мобильные сети.

IT-индустрия в целях анализа растущего объёма данных поднимает проблематику больших данных (Big Data), а академическое сообщество формирует науку о данных (Data Science). Наука о данных [1] - это раздел информатики, изучающий проблемы анализа, обработки и представления данных в цифровой форме. Более емкое определение: наука о данных занимается извлечением знаний из данных. Она объединяет следующие методы: обработку больших объемов данных, статистические методы, интеллектуальный анализ данных и методы искусственного интеллекта, а также методы проектирования и разработки баз данных. С 2010 года наука о данных рассматривается и как академическая дисциплина, и как практическая, межотраслевая сфера деятельности. Специализация «учёный по данным» (Data Scientist) с начала 2010-х годов считается одной из самых привлекательных, высокооплачиваемых и перспективных профессий [2].

Анализ данных не изучают в университете на уровне учебных программ по информатике. Наука о данных предполагает глубокое аналитическое мышление и творческий подход. Специалисты в этой сфере должны уметь не только решать комплексные проблемы с помощью анализа данных, но и задавать правильные вопросы, чтобы делать полезные заключения о данных.

Образовательные программы по дисциплине наука о данных предусматривают подготовку в области математики, статистики и информационных технологий, интеллектуальной обработки и анализа данных. Специалист по науке о данных обучается решать задачи поиска, сбора, хранения, подготовки, анализа данных и интерпретации результатов в области своей специализации. Примерами являются специализация университета Berkeley (курс Master of Information and Data Science) и общедоступная дистанционная специализация Data Science [3] на платформе Coursera.

Литература:

1. Википедия. Data Science [Электронный ресурс] // Режим доступа: http://en.wikipedia.org/wiki/Data_science.
2. Davenport T.H., Patil D.J. Data Scientist: The Sexiest Job of the 21st Century [Электронный ресурс] // Режим доступа: <http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/ar/1>
3. Специализация Data Science [Электронный ресурс] // Режим доступа: <http://knowledgeforit.com/content/view/375585/388/>