

И. В. Неред, Н. Б. Осипенко
(УО «ГГУ им. Ф. Скорины», Гомель)

ЛАТЕНТНО-СЕМАНТИЧЕСКИЙ АНАЛИЗ ТЕКСТА В ТЕСТИРОВАНИИ ИНДИВИДУАЛЬНЫХ ЧЕРТ ХАРАКТЕРА ЧЕЛОВЕКА

Особенность современного психологического тестирования состоит в том, что существует огромное разнообразие тестов, оценивающих человека с разных сторон. Тем не менее, используя для конкретного человека даже десяток тестов очень трудно свести интерпретации результатов этих тестов в единую картину описания человека. Связано это с тем, что каждый тест создавался на основе уникальной картины мира автора теста и используемого при этом языка.

В связи с этим возникла проблема синтеза психологических тестов. Таким образом, создаваемое программное средство ориентировано на унифицированный синтез интерпретации результатов различных психологических тестов.

Латентно-семантический анализ – это метод обработки информации на естественном языке, анализирующий взаимосвязь между коллекцией документов и терминами в них встречающимися, сопоставляющий некоторые факторы (тематики) всем документам и терминам.

В качестве исходной информации латентно-семантического анализа использует матрицу термины-на-документы, описывающую набор данных, используемый для обучения системы. Элементы этой матрицы содержат, как правило, веса, учитывающие частоты использования каждого термина в каждом документе и участие термина во всех документах. Наиболее распространенный вариант латентно-семантический анализ основан на использовании разложения матрицы по сингулярным значениям. С помощью SVD-разложения любая матрица раскладывается во множество ортогональных матриц, линейная комбинация которых является достаточно точным приближением к исходной матрице.

Разработанный макет компонента латентно-семантического анализа позволяет проецировать введенный пользователем текст на матрицу заранее подготовленных текстов, каждый из которых определяет некоторую тематику. Компонент будет использован для расширения функциональности приложения, развернутого на хостинге и доступного желающим в Интернете по ссылке <http://gsu-psychoanalysis.tk/>, позволяющего выдавать наиболее и наименее рекомендуемые пользователю сферы деятельности и профессии на основе известного в кругах психологов алгоритма квадрата Пифагора и серии психологических тестов.

Для облегчения разработки и следования стандартам веб-разработки использован фреймворк Angular 5; при создании интерфейса приложения – CSS фреймворк Twitter Bootstrap 4; а для реализации приложения – платформа NodeJS и модуль Electron, которые позволяют создавать платформонезависимые приложения при помощи языка JavaScript.