

УДК 004.5

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА И ВИЗУАЛИЗАЦИИ АКТИВНОСТИ ПОЛЬЗОВАТЕЛЕЙ ВЕБ-ПРИЛОЖЕНИЯ

Автор: Афанасьева Е.А.

Научный руководитель: доцент, кандидат экономических наук Васильева Т.В.

Северо-Западный институт управления Российской академии народного хозяйства и государственной службы при Президенте РФ

г. Санкт-Петербург

Аннотация. Работа посвящена использованию инструментов машинного обучения для совершенствования системы веб-аналитики, которая ориентирована на анализ и визуализацию данных, и повышения эффективности деятельности организации. В качестве практического примера рассматривается активность покупателей онлайн-магазина. Все решения реализуются с помощью языка программирования Python.

Введение. С каждым днём количество веб-приложений коммерческих и некоммерческих организаций увеличивается. Первые же вынуждены вести борьбу за потенциальных клиентов онлайн-услуг и непрерывно совершенствовать используемые технологии. Как следствие, всё большее внимание уделяется взаимодействию между человеком и компьютером, а также развитию и внедрению современных инструментов веб-аналитики, позволяющих делать процесс использования веб-приложений более удобным и восприимчивым к потребностям пользователя.

Основная часть. В данный момент помимо использования широко распространённых инструментов, таких как Google Analytics или Яндекс.Метрика, необходимо также оперативно узнавать о новых IT-решениях на рынке и внедрять их в работу. Анализ активности пользователей на сайте может быть качественно улучшен с помощью реализации алгоритмов машинного обучения, которые позволяют быстро анализировать большие, сложные данные, создавать гибкие решения и визуализировать результаты, что облегчает их интерпретацию.

В ходе работы была проанализирована активность пользователей онлайн-магазина, контент которого был разделён на три типа: административный, информационный и относящийся к продукту. Были предоставлены данные о частоте посещений различных разделов сайта, продолжительности сеанса, переходах между страницами, коэффициент выхода и выполнение целевого действия. Была проведена обработка первичных данных, нормализация, кластерный анализ, классификация полученных результатов. Для этого были использованы встроенные общедоступные библиотеки языка программирования Python, такие как numpy, pandas, scikit-learn, plotly и другие.

Выводы. В результате практического применения вышеописанных инструментов были проанализированы данные об активности пользователей и их характеристики, что помогло составить общий портрет пользователей и определить тенденции их поведения, а также найдены уязвимые места в интерфейсе веб-приложения, ключевые факторы, влияющие на отток клиентов с сайта, что в дальнейшем даст возможность исправить проблемы и сделать работу более эффективной. Все результаты были представлены с помощью широкого спектра визуальных средств, предоставляемых библиотеками языка Python.

Афанасьева Е.А.

Васильева Т.В.

