

УДК 004.62

## АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ВЛИЯНИЯ ХАРАКТЕРИСТИК ХУДОЖЕСТВЕННОЙ ЛИТЕРАТУРЫ НА ВЕРОЯТНОСТЬ ПОПАДАНИЯ В ТОПЫ ПРОДАЖ

Янов Ф.А. (Университет ИТМО)

Научный руководитель – доцент факультета инфокоммуникационных технологий,  
Валитова Ю.О. (Университет ИТМО)

Доклад посвящен решению проблемы определения характеристик художественной литературы, влияющих на ее попадание в тренды. В ходе исследования собирались данные о популярности произведений с сайтов Book24, Лабиринт, Ozon. В ходе анализа выявлялось влияние на попадание в топы продаж таких характеристик художественной литературы как стоимость издания, количество страниц, жанр, также было проведено тематическое моделирование краткого описания сюжета.

**Введение.** Попадание товара в тренды определяет его маркетинговую привлекательность. При этом в современном изменчивом мире крайне сложно уследить за трендами, в частности, за трендами в художественной литературе. Современные книги могут, как резко взлететь в тренды, так и не составить конкуренцию классической литературе 19 или 20 века. Например, продажи произведения молодого писателя Александра Полярного «Мятная сказка» стремятся к 300 тысячам продаж, в то время как новый роман Бориса Акунина, являющегося известным и признанным автором, «Мир и война» с трудом сможет набрать 50 тысяч.

Исходя из трендов и предположений о востребованности произведения художественной литературы, продавцы определяют, какие книги и в каком количестве экземпляров им лучше закупить, чтобы получить максимальную выгоду. С помощью понимания тем, привлекающих аудиторию, легче определить, какие произведения выгодно закупать. Одновременно с этим многие писатели пытаются определить, на какую тему им лучше написать своё произведение. Для решения этих проблем необходимо выявить характеристики произведения художественной литературы, влияющие на его попадание в топы продаж. Это подтверждает актуальность исследования.

**Основная часть.** В ходе проведения анализа предметной области были выбраны магазины, продающие книги в Российской Федерации, а именно Book24, Лабиринт, Ozon. После этого средствами Python и Web-scraping с этих ресурсов собираются данные по самым популярным художественным произведениям – хитам продаж. С помощью библиотеки Pandas данные были подготовлены к дальнейшему анализу. В ходе предобработки были переименованы столбцы, очищены от лишних знаков. Средний рейтинг был приведён к одинаковому формату. После этого проводился их анализ, направленный на выявление влияния на попадания в топы продаж таких характеристик художественной литературы как стоимость издания, количество страниц, жанр, также было проведено тематическое моделирование краткого описания сюжета. Результаты анализа визуализированы в инфографику и дашборд, доступный также в мобильном виде. Полученные в ходе исследования результаты позволят писателю определить, что будет читать современный читатель, а магазинам – актуальны ли произведения писателя для закупки.

**Выводы.** На основе полученных результатов исследования будут получены практические рекомендации, помогающие принимать решения для молодых писателей. Магазины смогут определить книги, которые перспективнее закупать. Решение будет актуально для всех регионов Российской Федерации, а использовать его результаты смогут и писатели из других

стран. К тому же данная информация будет полезна и обычным людям, следящим за современной литературой и желающим прочитать самые популярные произведения рынка.

Янов Ф.А. (автор)

Валитова Ю.О. (научный руководитель)