



Recommendation System: New Reads for Good Readers

Using GoodReads Data to Suggests New Book Titles



Intro



Opportunity:

- There were 857.9 MM Books Sold in 2021 (up 8.9% YoY) presenting a larger space in the Market
- The Pandemic created huge interest in written word again that can be capitalized on

Goal:

- Suggest books a reader may like based on title input
- Increase time spent on Websites like GoodReads (Now Owned by Amazon)

Background

Data

- Good Reads Dataset of 500K Books Published in Recent Years
- NLP of Book Descriptions
- Most Under 500 Words in Length

Cleaning:

- 100,000 books considered once subset for English Language and Book Validity
- Descriptions Cleaned of Punctuation, stopwords, and html tags

Topic Modeling: Grouping Book Types

Topic Number	Top Terms (20)	Genre/Classification
Topic 4	WAR, WORLD, HISTORY, AMERICAN, CENTURY, FIRST, STATE, AMERICA, BATTLE, YEAR, POLITICAL, CITY, GREAT, II, CIVIL, ARMY, MILITARY, ONE, ACCOUNT, BRITISH	American/World History
Topic 8	LIFE, LOVE, WOMAN, MAN, FAMILY, HEART, YOUNG, WOULD, FATHER, COULD, MARRIAGE, MOTHER, CHILD, PASSION, NEVER, DAUGHTER, WIFE, HUSBAND, MEN, BEAUTIFUL	Romance/Family Drama
Topic 2	TEXT, WORK, NEW, STUDENT, STUDY, VOLUME, GUIDE, PROVIDES, CHAPTER, THEORY, INFORMATION, ALSO, RESEARCH, INTRODUCTION, WELL, LANGUAGE, INCLUDES, SOCIAL, DEVELOPMENT, OFFER	Textbook/Reference
Topic 9	BOOK, CHILD, SERIES, PAGE, COLOR, READ, READER, ILLUSTRATION, READING, PICTURE, QUALITY, FULL, PART, ORIGINAL, COVER, FIRST, ILLUSTRATED, WORLD, AUTHOR, CLASSIC	Children's Novels

Recommender System

Input ISBN of Book You Have Read and Return:

You read: Breakfasts & Brunches

You Should Read These Next:

		Name	Language
10160		How My Parents Learned to Eat	eng
14944	The Nodwick Chronicles V: Tour of Doodie (The ...		eng
32095	Star Trek: The Classic Episodes, Volume 1		eng

Future Work

Completion of Stream Lit App for Interactive Recommending

Continued Scraping of Good Reads as new Publishing occurring rapidly, but Amazon has shut down the API

Similar Analysis on Foreign Language books

Recommendations for other Media beyond Book for Books (i.e. if you enjoyed this book, then you may love this podcast, tv show, or movie)

Appendix

Topic Outputs

Topic 0:

BR, THE, IN, AND, THIS, BUT, FROM, IT, WHEN, MAN, WOULD, AS, WITH, AN, JOHN, COULD, FOR, ROBERT, WHAT, YOU

Topic 1:

FIND, ONE, FRIEND, NEW, GET, OLD, TIME, KNOW, MURDER, HE, TAKE, THING, WANT, DAY, GO, LIKE, HELP, EVEN, YEAR, BACK

Topic 2:

TEXT, WORK, NEW, STUDENT, STUDY, VOLUME, GUIDE, PROVIDES, CHAPTER, THEORY, INFORMATION, ALSO, RESEARCH, INTRODUCTION, WELL, LANGUAGE, INCLUDES, SOCIAL, DEVELOPMENT, OFFER

Topic 3:

PURCHASE, RELEASED, MAY, MICHAEL, BABY, JAMES, ELIZABETH, WIFE, NIGHT, CHRISTMAS, HUSBAND, HEART, LADY, MAN, FAMILY, FIRE, YORK, MARRIAGE, HUNTER, MARY

Topic 4:

WAR, WORLD, HISTORY, AMERICAN, CENTURY, FIRST, STATE, AMERICA, BATTLE, YEAR, POLITICAL, CITY, GREAT, II, CIVIL, ARMY, MILITARY, ONE, ACCOUNT, BRITISH

Topic 5:

AMP, COMPANY, THE, FIRST, SCIENCE, PHILOSOPHY, GREEK, HUMAN, GOD, RELIGION, WORK, ROMAN, ANCIENT, MIND, INTRODUCTION, TRANSLATION, BIOGRAPHY, WRITTEN, JOHN, ALIEN

Topic 6:

STORY, THE, TALE, COLLECTION, SHORT, WRITER, FICTION, AUTHOR, TELL, CLASSIC, FIRST, READER, CHARACTER, NOVEL, ADVENTURE, ONE, NEW, TIME, YOUNG, AWARD

Topic 7:

EM, STRONG, THE, NEW, YORK, TIMES, BESTSELLING, AMERICAN, ONE, COM, PHILOSOPHY, CLASSIC, AUTHOR, IN, AWARD, FILM, READER, WRITING, READ, LANGUAGE

Topic 8:

LIFE, LOVE, WOMAN, MAN, FAMILY, HEART, YOUNG, WOULD, FATHER, COULD, MARRIAGE, MOTHER, CHILD, PASSION, NEVER, DAUGHTER, WIFE, HUSBAND, MEN, BEAUTIFUL

Topic 9:

BOOK, CHILD, SERIES, PAGE, COLOR, READ, READER, ILLUSTRATION, READING, PICTURE, QUALITY, FULL, PART, ORIGINAL, COVER, FIRST, ILLUSTRATED, WORLD, AUTHOR, CLASSIC