NovelQuest: An AI- powered Book Recommendation System

Devam Tanwani

Abstract

This report introduces NovelQuest, an innovative book recommendation website that enhances book discovery through an AI-powered recommendation system. Addressing the inefficiencies of traditional genre-based filters, NovelQuest allows users to input their own book descriptions that they want to read, leveraging machine learning and natural language processing (NLP) algorithms to match these descriptions with relevant books. This approach recommends books that resonate with individual preferences, saving readers time and frustration. This report also explores the limitations of current methods, target market, competition, and NovelQuest's unique features like user-driven descriptions and NLP matching. It also details potential revenue models and the system's functionalities. By prioritizing personalization at the next level, NovelQuest aims to revolutionize book discovery and reignite a love for reading.

1.0 Problem Statement

Many book readers face the challenge of finding the perfect book that aligns with their specific interests and preferences. Existing filters (like genres, authors, date) are helpful, but they often leave readers wanting more context about the book's content and themes. After filtering with genres, readers still have to click on each book to read the description of the book and determine if it intrigues them. Users have to manually read book descriptions (back blurbs) to understand the plot, which is time-consuming and inefficient. Readers also look for other features like the number of pages, publishing date, number of reviews, and honest opinions from fellow readers.

This is pretty time consuming as a reader has to go through this for each book and then decide which book is calling to them.
All this decision making process comes with consequences such as:

- *Wasted Time:* Users spend significant time browsing irrelevant books, leading to frustration and discouragement.
- *Ineffective Search Queries:* Keyword searches often miss the mark, failing to capture the essence of the desired plot or character development.
- *Unexplored Options:* Users might miss out on fantastic books that don't neatly fit into traditional genre categories but align with their specific interests.
- *Disengagement from Reading:* The difficulty in finding desired books can lead to decreased engagement with reading

This report proposes NovelQuest: a book recommendation website where users can not only filter books by genre but can also input a detailed description of the book they're looking for. The app would use machine learning and natural language processing (NLP) to analyze this description and search the internet to find books or related books that match the user's specified criteria. This would provide a more personalized and nuanced book recommendation experience, helping users find books that align closely with their specific interests and preferences.

By providing a more effective and personalized search experience, this platform will help users discover books they will truly enjoy and maintain their passion for reading.

2.0 Market/Customer/Business Need Assessment

2.1 Market Size

The global book market is substantial and consistently growing, with diverse offerings across genres, formats, and languages. Key statistics highlighting the market's size include:

- *Global Market Value:* The global book market was valued at approximately $151.8 billion in 2023, with projections indicating continued growth in the coming years. The global book market recorded a rise in revenues in many markets in the first eight months of 2023. There was a considerable increase in the emerging book markets of India (+8.7%).
- *E-Books and Digital Reading:* The increasing popularity of e-books and digital reading platforms has further expanded the market. The global ebook market alone was valued at around $18 billion in 2020 to $132.4bn in 2023, and is expected to grow at a compound annual growth rate (CAGR) of 6.5% from 2021 to 2028.
- *Reading Population:* An estimated 65% of adults read at least one book per year. Avid readers, who form the core target audience for NovelQuest, typically read an average of 12 books annually.
- *Community Engagement:* Online book communities and social platforms where readers share reviews and recommendations are becoming increasingly popular. Leveraging community insights and integrating social features can enhance the recommendation process.

2.2 Customer Needs and Pain Points

Despite the vast availability of books, readers often struggle with:

- *Time-Consuming Search:* The process of browsing through numerous book descriptions to find a suitable match is tedious and inefficient.
- *Limited Filter Options:* Current filtering systems based on genres, authors, or publication dates do not capture the nuances of the book's content, plots and theme.
- *Missed Opportunities:* Many great books are overlooked because they don't fit neatly into traditional categories but may align perfectly with a reader's specific interests.

2.3 Business Need

There is a significant business need for NovelQuest as it addresses key pain points in the book discovery process:

- *Enhanced User Experience:* By providing highly personalized book recommendations, NovelQuest enhances the user experience and satisfaction.
- *Increased Engagement:* Helping readers find books they truly enjoy can lead to increased engagement and loyalty.
- *Market Differentiation:* NovelQuest's unique approach sets it apart from competitors, providing a competitive edge in the book recommendation market.

2.4 Value Proposition

NovelQuest addresses customer needs by offering:

- *User-driven Descriptions:* Users can write descriptions of their ideal book, outlining desired plot elements, character types, or thematic nuances.
- *NLP and AI Matching:* The platform utilizes natural language processing (NLP) and artificial intelligence (AI) to understand user descriptions and identify relevant keywords or themes.
- *Matching Book Recommendations:* Based on the user's description, the platform recommends books with descriptions that closely match the user's preferences.

3.0 Target Specification and Characterization

3.1 Customer Characteristics

3.1.1 Avid Readers

- *Reading Frequency:* Regularly read books, typically finishing multiple books per month.
- *Genres of Interest:* Wide range of genres, from fiction (mystery, romance, sci-fi) to non-fiction (biographies, self-help).
- *Demographics:* Typically aged 18-60, with a higher concentration of young adults and middle-aged individuals.
- *Digital Savvy:* Comfortable using digital platforms for reading and book discovery.

3.1.2 Book Enthusiasts

- *Reading Habits:* Enjoy exploring various genres and discovering new authors.
- *Community Engagement:* Active in book clubs, online reading groups, and literary discussions.
- *Demographics:* Varied age group, often with a passion for literature and continuous learning.

3.1.3 Book Clubs and Reading Groups

- *Group Dynamics:* Small to medium-sized groups that read and discuss books collectively.
- *Genre Variety:* Diverse preferences within the group, seeking books that spark engaging discussions.
- *Engagement:* Active participants in discussions, looking for new and interesting book recommendations.

3.1.4 Libraries and Bookstores

- *Cataloging Needs:* Require advanced cataloging and recommendation systems to serve their patrons better.
- *Genre Preferences:* Broad spectrum of genres to cater to a diverse audience.
- *Demographics:* Serving all age groups, from young readers to senior citizens.
- *Technological Integration:* Interested in integrating AI-powered recommendation systems to enhance user experience.

3.2 Target Specifications

- *Intuitive Design:* The website should have a clean, user-friendly interface that is easy to navigate for all age groups.
- *Search Functionality:* Advanced search filters for book title, author, genre, availability, and library location.
- *Personalized Recommendations:* AI-driven suggestions based on user preferences, history, and reading habits.

3.3 Performance Metrics

- *Search Speed:* Quick response time for providing book recommendations.
- *Accuracy Rate:* High accuracy in matching user descriptions with appropriate books.
- *User Satisfaction:* Positive feedback from users regarding the relevance and usefulness of recommendations.

4.0 External Search

4.1 Source:
https://wordsrated.com/book-sales-statistics/

Book sales distribution by year ($ million)

Physical sales values (top, magenta): 2011: $18,552 · 2012: $17,873 · 2013: $16,828 · 2014: $17,211 · 2015: $16,242 · 2016: $13,571 · 2017: $12,913 · 2018: $11,386 · 2019: $10,804 · 2020: $7,514

Online sales values (bottom, dark purple): 2011: $8,808 · 2012: $9,847 · 2013: $10,242 · 2014: $10,749 · 2015: $11,558 · 2016: $12,699 · 2017: $13,317 · 2018: $14,244 · 2019: $14,966 · 2020: $18,586

■ Physical sales ■ Online sales (e-book and audiobooks included)

Since 2017, the majority of book sales in the United States have occurred online. The rise of e-commerce, along with the growing popularity of ebooks and audiobooks, has fueled a shift from physical bookstores to online book discovery. This shift is driven by the convenience of online platforms that offer search filters, user reviews, and personalized recommendations, making it easier for readers to find the books they're looking for.

4.2 Source:

https://www.gfk.com/press/global-book-market-revenues-up-in-2023-thanks-to-rising-prices

The above source showcases the trend of book markets in different countries. The revenue produced by book markets is on an increase every year, thanks to higher prices as well as more user engagement in buying books, especially online.

Datasets:

1. https://www.kaggle.com/datasets/arpansri/books-summary
2. https://www.kaggle.com/datasets/ymaricar/cmu-book-summary-dataset

Combined the 2 datasets using the concat() function in pandas.

	book_name	genre	summary
0	Animal Farm	Roman à clef, Satire, Children's literature, S...	old major, the old boar on the manor farm, ca...
1	A Clockwork Orange	Science Fiction, Novella, Speculative fiction,...	alex, a teenager living in near-future englan...
2	The Plague	Existentialism, Fiction, Absurdist fiction, Novel	the text of the plague is divided into five p...
3	An Enquiry Concerning Human Understanding	NaN	the argument of the enquiry proceeds by a ser...
4	A Fire Upon the Deep	Hard science fiction, Science Fiction, Specula...	the novel posits that space around the milky ...
...	...	...	...
1225	Your Move: The Underdog's Guide to Building Yo...	business, marketing, money, work	is ramit sethi's no-bs guide to starting your...
1226	You'll See It When You Believe It	motivation, productivity, psychology, work	shows you how to discover your true, best sel...
1227	You're Not Listening	business, communication, happiness, management...	is a book that will improve your communicatio...
1228	Zero To One	business, management, marketing, motivation, p...	is an inside look at peter thiel's philosophy...
1229	iWoz	biography, business, technology, work	is steve wozniak's autobiography, detailing h...

17789 rows × 3 columns

The resulting dataset includes book names, along with their genres and descriptions, which will enable us to analyze and generate book recommendations that closely match the user's description.

5.0 Benchmarking alternate products

Benchmarking is crucial for understanding the competition and identifying ways for NovelQuest to distinguish itself in the market. This analysis helps uncover key features, strengths, and potential gaps in current offerings, providing insights into how NovelQuest can offer unique value to its users.

Current book recommendation platforms include:

- *Goodreads:* Provides recommendations based on user ratings and reviews but lacks deep personalization.
- *Amazon:* Offers suggestions based on purchase history but may not capture specific thematic preferences.
- *LibraryThing:* Focuses on cataloging and community reviews but doesn't utilize advanced NLP techniques for personalized recommendations.

NovelQuest sets itself apart by:

- *Personalized Descriptions:* Allowing users to input detailed descriptions of their ideal book.
- *Advanced AI and NLP:* Utilizing cutting-edge technology to match user descriptions with relevant books.
- *Comprehensive Search:* Combining multiple data sources to provide accurate and tailored recommendations.

6.0 Applicable Patents
1. [Content based recommendation system](#)
2. [System and method for providing personalized recommendations](#)

7.0 Applicable Regulations
1. Rules against False Marketing
2. License for the open-source codes that might be used in the model implementation.
3. Laws related to AI

8.0 Applicable Constraints
1. Lack of initial data to perform algorithms.
2. Data collection from open source APIs such as google books, goodreads.
3. Limited availability of trend data for accurate prediction and analysis.

9.0 Business Model (Monetization Idea)
The NovelQuest website will be free to use. We would not charge the users any money so that they can get their perfect book to read and make an informed decision.
Some ways to monetize the NovelQuest website can be done using:

9.1 Partnerships

Collaborate with libraries and specific online stores to feature their books prominently on NovelQuest. This partnership ensures that users have access to a wide variety of books and that libraries and stores can reach a larger audience. For example, once the website gains a substantial number of users, we can approach Amazon. If they see value in our platform, we can arrange for Amazon links to appear first when users search for books. This prioritization can drive traffic to Amazon and serve as a revenue source for NovelQuest.

9.2 Affiliate Marketing

Affiliate marketing is a way to earn money by promoting products or services from other companies. Partner with online bookstores like Amazon, Flipkart, and Delhi Book Market to earn commissions on book purchases made through NovelQuest's links. By integrating with these platforms, NovelQuest can provide users with a seamless purchase experience. Users can read about a book, get recommendations, and buy it directly through the platform. This model not only generates revenue but also adds value by simplifying the book buying process for users.

9.3 Targeted Advertisements

Display ads tailored to users preferences, such as promotions for upcoming books, literary events, and reading accessories. By leveraging user data, NovelQuest can show relevant advertisements that align with the interests and reading habits of its users.

9.4 Data Analytics

Provide anonymized user data and reading trends to publishers and authors for a fee. This data can help them understand market demands, reader preferences, and emerging trends. Publishers and authors can use this information for improving book content, and strategic decision-making. By offering these insights, NovelQuest can position itself as a valuable partner in the publishing industry,

10.0 Concept generation

As an avid book lover and an overthinker, I spend a lot of time selecting which book to read next. This process often involves reading reviews, watching videos, and comparing multiple options to ensure that I choose a book that truly resonates with me. Recently, I tried to search for a book on google based on a small description of a book that I would want to read. However, the search results did not really match with the description that I inputted. It was vague and irrelevent. This experience sparked the idea for NovelQuest, a website dedicated to accurately recommending books that match user-generated descriptions. By leveraging machine learning and with the help of a large corpus of available novels on the internet, NovelQuest aims to make searching for the perfect book more efficient, personalized and satisfying for readers like myself.

11.0 Concept development

NovelQuest is a book recommendation platform designed to offer a personalized and efficient book searching experience. It is a minimalistic website where users can input descriptions of the type of book they want to read, including plot elements and themes. Using machine learning and NLP algorithms, NovelQuest will analyze these descriptions and search an extensive database of books to find those that match the user's preferences. This approach will save users time and enhance their reading experience by helping them find books they genuinely want to read.

The platform will allow users to filter books by genre, author, date, and other traditional categories as well. Additionally, NovelQuest will provide detailed book information, including summaries, reviews, and ratings, to assist users in making smart informed decisions. It will also provide links to online shopping stores like amazon, flipkart, where the book is available.

12.0 Final product prototype

NovelQuest is an innovative book recommendation platform designed to enhance the book discovery process. Users can input detailed descriptions of their ideal book by specifying plot elements, character traits, and themes. The platform leverages machine learning and natural language processing (NLP) to analyze these descriptions and match them with books from an extensive database, providing highly personalized recommendations.

Schematic diagram:
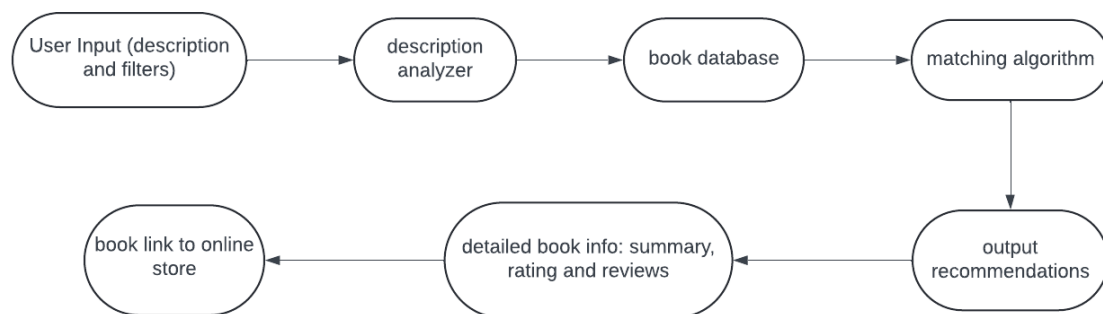
Steps in the Process:

1. *User Input (Description and Filters):* Users enter descriptions and select filters for the books they are looking for.
2. *Description Analyzer:* Utilizes NLP and AI to interpret user descriptions and identify key elements.
3. *Book Database:* Contains extensive information about various books, including summaries, genres, and reviews.
4. *Matching Algorithm:* Compares user descriptions with the database to find the best matches.

5. *Output Recommendation:* Provides a list of books that closely match user preferences.
6. *Detailed Book Information:* Offers summaries, themes, user reviews, ratings, and links to purchase the books online.
7. *User Feedback (Optional):* Users can provide feedback on the recommendations and share their own reviews and experiences.

Desired Outcome:

An simple and minimalistic book recommendation website that addresses these challenges by leveraging user-written descriptions and AI-powered matching:

- *User-driven Descriptions:* Users can write descriptions of their ideal book, outlining desired plot elements, character types, or thematic nuances.
- *NLP and AI Matching:* The platform utilizes natural language processing (NLP) and artificial intelligence (AI) to understand user descriptions and identify relevant keywords or themes.
- *Matching Book Recommendations:* Based on the user's description and NLP analysis, the platform recommends books with descriptions that closely match the user's preferences.

Example Scenario:
Let's say you write this description: "A suspenseful thriller set in a college campus, where the protagonist is wrongly accused of a murder. Amidst the investigation, he discovers unexpected love."
NovelQuest would recommend books like "The Secret History" by Donna Tartt or "Gone Girl" by Gillian Flynn.
The system would consider elements like college life, suspense, crime, and romance.

```
# Example usage
user_description = "a deadly virus has gotten over the city with a child protagonist"
recommended_books = recommend_books(user_description, df, X, vectorizer)
recommended_books
```

	book_name	genre	summary
7254	Just a Couple of Days	NaN	dr flake fountain approached military develop ...
11518	Whiteout	thriller, fiction	rabbit tested find cure deadly virus dangerous...
2905	The Hot Zone	hard science fiction, fiction, non-fiction	book four section shadow mount elgon delf hist...
7943	The Skinner	science fiction, speculative fiction	skinner tell story three individual journeyed ...

13.0 Financial Modelling

13.1 Market Identification

NovelQuest will enter the personalized content recommendation market, a sector that is experiencing rapid growth driven by increasing consumer demand for tailored experiences.

13.2 Data Collection

To create a reliable financial model, data was collected on the growth trends of similar markets, such as streaming services and online education platforms. These data points provide a basis for forecasting the potential growth of NovelQuest.

13.3 Financial Equation

The financial equation for NovelQuest is based on an exponential growth model, reflecting the rapid adoption typically seen in tech-driven markets:

$$Revenue(t) = P \times N(t) - C$$

Where:

- P is the monthly subscription fee (e.g., $5).
- $N(t)$ is the number of users at time t, modeled as an exponential function: $N(t) = N0 \times e^{rt}$
- N0 is the initial number of users.
- r is the monthly growth rate (assumed to be 15%).
- C represents fixed and variable costs, including development, maintenance, and server costs.

Example:

Assuming an initial user base of 1,000 and a 15% growth rate:

$$N = 1000 \times e^{0.2} \times 12 \approx 6049.6$$

At 12 months, the revenue is:

$$Revenue(12) = 5 \times 6049.6 - 3000 \approx \$27,248$$

This model indicates that NovelQuest could achieve a monthly profit by its first year, depending on user growth and cost management.

13.4 Sensitivity Analysis

The financial model's sensitivity to changes in growth rate and subscription pricing was analyzed. This analysis helps to understand the potential risks and returns under different market conditions, allowing for strategic adjustments as needed.

13.0 Conclusion

NovelQuest aims to revolutionize book discovery by offering a personalized search experience. Unlike traditional filters, it leverages user-written descriptions and AI to understand reader intent and recommend books that truly resonate. This not only saves time and frustration, but also fosters a deeper connection with reading by connecting readers with their perfect match. By building a strong foundation with a functional prototype, data acquisition strategy, and a skilled team, NovelQuest can establish itself as a valuable tool for readers and a key player in the book discovery landscape. Imagine a world where you can describe your ideal book's atmosphere, characters, or plot twists, and NovelQuest curates a list of recommendations that perfectly capture your desired reading experience. This is the future NovelQuest envisions, a future where book discovery is no longer a guessing game, but a personalized journey of literary exploration.