

# Final Report on Used Book Sales System powered by Blockchain

Names

May 15, 2024

## Abstract

This report presents the final outcomes and findings of the project focusing on the development of a used book sales system powered by blockchain technology. The system aims to streamline the process of buying and selling used books while ensuring transparency, security, and traceability through blockchain integration. Leveraging content-based search algorithms, the system provides personalized recommendations to users, facilitating the discovery of books tailored to their interests and preferences. The report discusses the project's objectives, methodology, requirements analysis, design, development, testing, deployment, maintenance, challenges encountered, evaluation of results, and future development prospects. Through innovative use of blockchain and advanced search algorithms, the system offers a novel approach to enhancing the efficiency and trustworthiness of the used book marketplace.

## 1 Introduction

The digital age has revolutionized the way we interact with literature, providing unprecedented access to a vast array of books through online platforms. However, despite the convenience of digital bookstores, users often encounter challenges in navigating the plethora of available titles, finding books that match their interests, and ensuring the authenticity of transactions, particularly in the realm of used book sales. Addressing these challenges, our project focuses on the development of a used book sales system powered by blockchain technology. By integrating blockchain, we aim to enhance transparency, security, and trust in the used book marketplace, mitigating concerns related to counterfeit books and fraudulent transactions. Furthermore, by incorporating content-based search algorithms, our system offers personalized recommendations to users, guiding them towards books that align with their tastes and preferences. This report provides a comprehensive overview of our project, detailing the objectives, methodology, design, development, testing, deployment, and future prospects of our innovative solution. Through this endeavor, we seek to revolutionize the used book marketplace, fostering a seamless and trustworthy experience for book enthusiasts worldwide.

## 2 Problem statement

The domain of online bookstores, encompassing both new and used books, constitutes a significant segment of the e-commerce industry. However, the current landscape of online book sales platforms, particularly in the realm of used books, presents several challenges that hinder user experience and trustworthiness. Existing solutions often lack robust mechanisms for facilitating efficient search and discovery processes, resulting in user frustration and suboptimal outcomes. Users encounter difficulties in navigating through the vast array of available titles, and concerns regarding the authenticity and provenance of books persist, casting doubt on the reliability of transactions.

Inadequate search algorithms exacerbate the problem, failing to provide personalized recommendations tailored to individual user preferences. As a result, users struggle to find relevant titles amidst the extensive catalog, hampering their overall experience. Furthermore, the management of bookstores faces challenges in efficiently handling inventory and ensuring the quality and authenticity of used books.

To address these challenges, our project proposes a comprehensive solution that leverages blockchain technology and advanced search algorithms. By integrating blockchain, we aim to enhance transparency, security, and trust in the used book marketplace. Blockchain ensures immutable records of transactions, mitigating concerns related to counterfeit books and fraudulent activities. Additionally, our solution focuses on improving the performance of search algorithms to provide users with personalized recommendations, facilitating smoother transactions and enhancing user satisfaction.

In summary, our project seeks to revolutionize the way users interact with online bookstores by addressing the inefficiencies and challenges prevalent in the current landscape. By providing a robust and user-friendly platform for buying and selling used books, we aim to foster a seamless and trustworthy experience for both book enthusiasts and bookstore management. Used references are BookSwap.lk and UsedBooks.lk

## 3 Literature Review

In this section, we review the existing literature on online book sales platforms, focusing on the problems with the current systems and the solutions provided by recent research and industry developments. We organize the review in reverse chronological order, starting from the latest publications and advancements.

### 3.1 Recent Developments (2019-present)

#### 3.1.1 Amazon

Amazon, as a dominant player in the online book sales market globally, offers a comprehensive platform for buying and selling both new and used books. However, despite its extensive catalog and user-friendly interface, Amazon faces

criticism for its lack of transparency in the used book marketplace. Users often encounter difficulties in verifying the authenticity and quality of used books, leading to trust issues and dissatisfaction.

### **3.1.2 Examples from Sri Lanka**

In Sri Lanka, online bookstores like *Sarasavi.lk* and *MD Gunasena* have emerged as popular platforms for purchasing both new and used books. However, similar to global platforms, users in Sri Lanka also face challenges related to the authenticity and quality of used books. Trust issues arise due to the lack of transparency in the transaction process and the inability to verify the condition of books before purchase.

### **3.1.3 Solutions Proposed**

Recent research and industry developments have focused on addressing the shortcomings of existing online book sales platforms. Proposed solutions include the integration of blockchain technology to enhance transparency and traceability in the used book marketplace. By leveraging blockchain, researchers aim to provide immutable records of book transactions, ensuring the authenticity and provenance of used books.

## **3.2 Previous Studies (2009-2018)**

### **3.2.1 eBay**

Before the rise of specialized online bookstores, platforms like eBay served as popular venues for buying and selling used books globally. While eBay offered a wide range of products and a decentralized marketplace, users faced challenges in navigating through listings and verifying the quality of books.

### **3.2.2 Examples from Sri Lanka**

In Sri Lanka, online marketplaces like *ikman.lk* and *LankaAds* have been utilized for buying and selling various goods, including books. However, similar to eBay, users encounter difficulties in assessing the condition and authenticity of used books listed on these platforms. Trust issues arise due to the lack of standardized processes for book transactions and the absence of user reviews/ratings specific to book sellers.

### **3.2.3 Solutions Implemented**

To address the limitations of platforms like eBay, researchers and industry practitioners introduced features such as seller ratings and buyer reviews to enhance trust and transparency in online transactions. Additionally, advancements in search algorithms aimed to improve the discoverability of relevant books and provide personalized recommendations to users.

### 3.3 Comparison of Features

Platform	Key Features
Amazon	Extensive catalog, user reviews, personalized recommendations
eBay	Decentralized marketplace, seller ratings, buyer reviews

Table 1: Comparison of Features in Online Book Sales Platforms

### 3.4 End of Literature Review

In summary, existing literature on online book sales platforms highlights the challenges faced by users in verifying the authenticity and quality of used books. While platforms like Amazon and eBay offer various features to enhance user experience, trust issues persist in the used book marketplace. Our project aims to address these challenges by leveraging blockchain technology and advanced search algorithms to provide a seamless and trustworthy platform for buying and selling used books.

### 3.5 Expected Features

The features we aim to bring to our platform include:

- Blockchain integration for transparent and traceable transactions
- Advanced search algorithms for personalized recommendations
- User reviews and ratings for enhanced trust and transparency

### 3.6 Research Gap

Our project fills a research gap by providing a comprehensive solution that addresses the shortcomings of existing online book sales platforms. By integrating blockchain technology and advanced search algorithms, we aim to enhance transparency, security, and user satisfaction in the used book marketplace. Additionally, by incorporating examples from Sri Lanka, we strive to make our solution relevant and applicable to diverse global contexts.

## 4 Project Objectives

## 5 Methodology

In this section, we outline the methodology employed in our project, including the features incorporated, expected results, and the rationale behind the decisions made.

## 5.1 Features and Implementation

We draw inspiration from successful implementations in platforms like Amazon, incorporating features such as decentralized data management and federated access. Our project aims to merge these concepts and adapt them to the specific needs of the online book sales domain.

### 5.1.1 Decentralized Data Management

Decentralization provides federated access, ensuring that no single party is solely responsible for updating the database. This approach enhances transparency in data entry and modification, reducing the risk of censorship. Additionally, it allows for the sharing of physical resources among consortium members, promoting collaboration and efficiency.

### 5.1.2 Data Ingest

To automate the process of updating the database, we will implement an automated system that constantly updates the database with new information. We plan to utilize IPFS as an open metadata database, providing an API endpoint for accessing bookstore data and training AI models.

## 5.2 API Development

The API will be a central component of our project, trained on and utilizing data from the Metadata database. It will be accessible to anyone, with rate limits in place to prevent abuse. The API will provide relevant results tailored to user preferences, which can be used to customize recommended views.

### 5.2.1 Recommendation API

Our recommendation API will serve two main functions:

1. User Receives relevant results (List of best matched books, ranked in order of preference).
2. User makes requests for recommendations, providing user preference data and receiving similar book suggestions.

## 5.3 Search API

The search API will incorporate Natural Language Processing (NLP) techniques to extract tags, understand intent, and match tags with relevant results in multiple languages. It will utilize content-based filtering methods to provide accurate and personalized search results.

## 5.4 Frontend and Backend Development

For the frontend, we will primarily use React, employing the MERN (MongoDB, Express.js, React, Node.js) stack for seamless integration. The backend will be developed using Python, TensorFlow, and Hugging Face for NLP recommendation processes.

### 5.4.1 NLP Recommendation Process

The NLP recommendation process will involve several steps, including:

- iText Similarity: Analyzing synopses to determine textual similarity.
- Named Entity Recognition: Identifying genres and authors from book descriptions.
- Keyword Extraction: Extracting keywords from synopses and reviews.
- Topic Extraction: Identifying topics from extracted keywords.

We will implement content-based filtering to ensure accurate and diverse book recommendations tailored to user preferences.

## 5.5 Conclusion

By implementing these methodologies, we aim to develop a robust and user-friendly platform for online book sales, offering personalized recommendations and enhancing the overall user experience.

# 6 Requirements Analysis

# 7 Design

# 8 Development

## 8.1 Introduction (Our Initial Scope)

In our initial proposal, we aimed to develop a comprehensive book recommendation system with three main components:

1. An API server that hosts the recommender model.
2. A test application to showcase the model's capabilities.
3. A structured dataset of books for training and evaluation.

This section recaps our initial objectives and the scope of the project as presented in our first proposal.

## 8.2 Results (Finished Scope)

We explored various AI/ML techniques to provide accurate book recommendations. After thorough experimentation, we settled on a combination of models to deliver the best results. The key outcomes are:

- A robust model server that hosts our recommender system.
- A demo application demonstrating the recommendation capabilities.
- A sample real-world application for a bookstore, integrating our recommendation system.

However, we faced challenges in handling the data, which limited our ability to accomplish all our goals.

## 8.3 What We Didn't Do (Unfinished Tasks)

While we made significant progress, there were several tasks we couldn't complete:

- Comprehensive testing of the recommender system.
- Converting the dataset from the SLNB to a suitable AI dataset.
- Implementing a seamless data handling pipeline.

These shortcomings are primarily due to the complexity and time required for these tasks.

## 8.4 Why We Didn't Do It (Issues Faced)

Despite our best efforts, certain obstacles hindered our progress:

- The conversion of the dataset from SLNB format to an AI-compatible format was more challenging than anticipated. We tried several methods, including manual extraction, automated scripts, and third-party tools, but each approach had its own set of issues, such as data inconsistency and integration difficulties.
- Limited resources and time constraints prevented us from conducting extensive testing. We implemented basic tests but couldn't cover all edge cases and scenarios.
- Integrating diverse data sources into a cohesive dataset proved to be complex, leading to delays and integration issues.

These challenges highlighted the need for more robust data handling and pre-processing strategies.

## 8.5 Implementation Details (How We Did It)

Our implementation process involved several key steps:

### 1. AI Modeling:

- We experimented with three different feature extraction models: TF-IDF, BERT, and Word2Vec. TF-IDF yielded the best results, while BERT presented performance issues and Word2Vec was primarily effective for string matches.

### 2. Demo App:

- A Flask-based API server (`app.py`) was developed to serve the recommendation model. The model, pre-trained and saved using `joblib`, is loaded at runtime.

### 3. Bookshop App:

- We integrated the recommendation system into a sample bookstore application, providing users with personalized book suggestions.

### 4. Packaging & Deployment:

- The system was packaged using virtual environments to manage dependencies and prevent version conflicts.

### 5. Testing and Statistics:

- Basic tests were conducted to ensure the functionality of API routes, error handling, and model predictions.

## 8.6 Application Demo (Guided Tour of the API and Services)

In this section, we provide a guided tour of our system, highlighting the polished parts and showcasing the functionality.

### 8.6.1 Intro

We proposed three main components:

1. An API server containing the recommender model.
2. A test app to showcase the model.
3. A structured book dataset.



### 8.6.2 AI Modeling

#### Content-based Filtering Model:

##### 1. Data Preprocessing:

- Loaded book data from the goodbooks-10k repository.
- Selected relevant columns: authors, average rating, genres, language code, title, and description.
- Handled missing values in description and title columns.

##### 2. Feature Extraction:

- Used TF-IDF vectorization for title and description columns.
- Binarized genres using MultiLabelBinarizer.
- One-hot encoded the language code column.

##### 3. Similarity Calculation:

- Computed cosine similarity between composite feature vectors.

##### 4. Recommendation Function:

- The `recommend_items` function returns the top 10 similar book titles based on cosine similarity.

### 8.6.3 Demo App

#### Flask Application:

- Created an API endpoint at `/predict` for POST requests with a JSON payload containing a book title.
- The model is loaded from `model.joblib`, and predictions are returned as a JSON response.

### 8.6.4 Bookshop App

#### Implementation Pipeline:

- Import Data → Analyze Data → Preprocessing → Feature Extraction → Similarity Measures → Final Model
- Focused on preprocessing, feature extraction, and similarity measures.

### 8.6.5 Packaging & Deployment

#### Virtual Environments:

- Managed package versions and dependencies using virtual environments.

### 8.6.6 Testing and Statistics

#### Tests Conducted:

- Verified API route functionality.
- Checked for descriptive error messages.
- Ensured accurate model predictions.

### 8.6.7 Feature Extraction Models

#### Three Feature Extractors:

1. **feature\_extractor\_distilbert.py:**

- Uses DistilBERT for document embeddings.
- Combines document embeddings with genres.

2. **feature\_extractor\_tf\_idf.py:**

- Uses TF-IDF for title and description columns.
- Binarizes genres and one-hot encodes language codes.

3. **feature\_extractor\_word2vec.py:**

- Trains Word2Vec on tokenized text.
- Combines word embeddings with other features.

### 8.6.8 Preprocessor Classes

#### DataPreprocessor:

- Selects columns of interest and preprocesses book data.

#### KerasPreprocessor:

- Prepares book titles for Keras models.

### 8.6.9 Tokenization

#### Notebooks:

- **data\_exploration.ipynb:** Explores tokenization-related data.
- **ELM0.ipynb:** Covers ELMo embeddings and contextual word representations.

## **9 Maintenance and Support**

### **9.1 Regular Dataset Updates**

At the start, we aim to use a structured database for our system. However, the initial database we have is in PDF format, which requires significant time and effort to clean and integrate into our system. This process involves extracting the data from the PDFs, ensuring its accuracy, and formatting it appropriately for database entry. Given these challenges, our goal for the future is to use a dedicated database from the outset. This will streamline data management, making it easier to update and maintain regularly, thus improving the efficiency and reliability of our system.

### **9.2 Expanding Payment Methods**

To enhance our system, we plan to expand it by adding multiple payment methods. Integrating diverse payment options, such as credit/debit cards, PayPal, and cryptocurrencies, will provide users with the flexibility and convenience to choose their preferred payment method, thereby improving their overall experience. Additionally, to build trust and reliability with our users, we will implement a tracking system. This system will allow users to monitor the status of their orders in real-time, from purchase to delivery. Providing transparency and real-time updates through tracking not only enhances user satisfaction but also strengthens their confidence in our platform's reliability and commitment to excellent service.

### **9.3 Blockchain for Security Backups**

To enhance the security and integrity of our book buying and selling system, we will use blockchain technology to keep backups. Blockchain provides a decentralized and tamper-proof way to store data, ensuring that our backups are secure from unauthorized access and alterations. By utilizing blockchain, each backup is cryptographically linked to the previous one, creating an immutable chain of records. This approach not only protects the data from potential breaches but also guarantees data integrity and transparency. In the event of data loss or corruption, we can reliably restore the system from these secure backups, ensuring continuous and trustworthy operation.

### **9.4 Blockchain-Enhanced Search Methods**

To enhance our book buying and selling system, we plan to expand the searching method using blockchain technology. By leveraging blockchain, we can create a decentralized search index that ensures the integrity and authenticity of search results. Each entry in the search index is recorded on the blockchain, making it immutable and transparent. This prevents manipulation and ensures users receive accurate and trustworthy search results. Additionally, blockchain can

improve the search efficiency by distributing the search processing across multiple nodes, reducing the load on any single server and potentially increasing the speed and reliability of search queries. Implementing a blockchain-based search method not only enhances transparency and trust but also provides a robust and scalable solution for handling large volumes of search queries.

## **9.5 Conclusion**

By focusing on regular dataset updates, expanding payment methods, utilizing blockchain for security backups, and enhancing search methods with blockchain technology, we aim to maintain and support our system effectively. These measures ensure our platform remains reliable, secure, and user-friendly, meeting the evolving needs of our users and maintaining their trust and satisfaction.

## **10 Challenges**

In the development of our project, we have encountered several challenges that have required innovative solutions and careful consideration.

### **10.1 Collaborative Filtering Without User Data**

One of the primary challenges we faced was implementing collaborative filtering without access to user data. Collaborative filtering relies on user behavior and preferences to generate recommendations. However, as we did not have access to extensive user data, we needed to explore alternative methods for generating accurate recommendations. Considering the time constraints we halted our efforts in creating the collaborative filtering model.

### **10.2 Multilingual Filtering**

Another challenge we encountered was implementing filtering mechanisms for multiple languages. While we initially developed our filtering algorithms in English, expanding these mechanisms to accommodate additional languages presented unique challenges. Adapting our algorithms to effectively process and analyze text in multiple languages required significant research and development efforts.

### **10.3 Uploading Metadata to IPFS**

Integrating IPFS as an open metadata database posed challenges related to data management and synchronization. Uploading metadata to IPFS required careful consideration of data formats, storage mechanisms, and synchronization protocols to ensure efficient and reliable access to information.

## **10.4 Security Mechanisms in Blockchain**

Ensuring the security of data stored on the blockchain was another significant challenge. While blockchain technology offers inherent security features, vulnerabilities such as data tampering and unauthorized access needed to be addressed. Implementing robust security mechanisms to safeguard against potential threats and vulnerabilities was crucial to maintaining the integrity and trustworthiness of our platform.

## **10.5 Development Challenges**

In addition to the specific challenges outlined above, we also encountered various development challenges throughout the project lifecycle. These challenges included technical complexities, integration issues, and resource constraints. Addressing these challenges required a collaborative and iterative approach, with continuous adaptation and refinement of our strategies and methodologies.

## **10.6 Conclusion**

Despite the challenges encountered, our team remained committed to overcoming obstacles and delivering a high-quality solution. Through creative problem-solving and diligent effort, we successfully navigated through various challenges to achieve our project objectives.

# **11 Evaluation and Results**

## **12 Future Development**

As we move forward with the project, several avenues for future development and enhancement present themselves. These include:

### **12.1 Expansion of Language Support**

One key area for future development is the expansion of language support within the platform. While our initial efforts focused on English language support, expanding our algorithms and databases to accommodate additional languages will broaden the platform's accessibility and appeal to a more diverse user base.

### **12.2 Integration of User Feedback Mechanisms**

Incorporating user feedback mechanisms into the platform will be essential for continuous improvement and refinement. By soliciting feedback from users, we can gain insights into their preferences, identify areas for improvement, and tailor recommendations to better meet their needs and expectations.

### **12.3 Enhancement of Recommendation Algorithms**

Continued research and development efforts will be directed towards enhancing the accuracy and effectiveness of our recommendation algorithms. This includes exploring advanced machine learning techniques, such as deep learning and reinforcement learning, to further refine the recommendation process and provide more personalized and relevant suggestions to users.

### **12.4 Integration of Social Features**

Integrating social features into the platform, such as user reviews, ratings, and recommendations from friends or social networks, can enhance user engagement and foster a sense of community among users. By leveraging social data, we can further enhance the accuracy and relevance of recommendations and provide users with valuable insights and recommendations from their peers.

### **12.5 Expansion of Blockchain Applications**

Exploring additional applications of blockchain technology within the platform presents opportunities for enhanced security, transparency, and trust. This includes implementing smart contracts for automated transactions, establishing decentralized governance mechanisms, and leveraging blockchain for digital rights management and intellectual property protection.

## **13 Conclusion**

Finally, the development of a used book sales system powered by blockchain technology represents an important step forward in addressing the challenges faced by users of the online book market. Our project aims to streamline the process of buying and selling used books, ensuring transparency, security and traceability. By using content-based search algorithms, the system provides users with personalized recommendations, improving their overall experience and facilitating the discovery of books tailored to their interests and preferences. Throughout the project, we have meticulously analyzed the existing literature on online bookselling platforms, identified the shortcomings of current systems, and proposed innovative solutions to address these challenges. We aim to reduce issues related to counterfeit books and fraudulent transactions, improve transparency and trust in the used book market. Additionally, by incorporating advanced search algorithms, our system offers personalized recommendations, directing users to books that align with their tastes and preferences. The development process included requirements analysis, design, development, testing, deployment, and maintenance phases, and everything was executed meticulously to ensure the successful implementation of our solution. Despite the challenges of collaborative filtering without user data, multilingual filtering, and uploading metadata to IPFS, our team was committed to overcoming obstacles through creative problem solving and agile efforts. Moving forward, our project has

tremendous potential for further innovation and improvement. Future development prospects include expanding language support, integrating user feedback mechanisms, improving recommendation algorithms, integrating social features, and exploring additional blockchain applications. By focusing on these areas, we aim to create a platform that provides unparalleled value and utility to users in the online bookselling domain. In summary, the development of a used book sales system powered by blockchain technology represents a significant advance in the online book market. Through our project, we strive to provide a seamless and trusted platform for buying and selling used books, fostering a positive and rewarding experience for book lovers worldwide.

## **14 Appendices**

## **15 References**