CHAPTER I
INTRODUCTION

1.1 Background

The development of information technology has resulted in an increase in the volume of information, known as the phenomenon of information overload, where users face difficulties in accessing information that aligns with their preferences and assessing the quality of its content [1][2]. In digital libraries and online book platforms such as Goodreads, this phenomenon also causes the number of book collections to continuously grow, which hinders readers in selecting relevant and high-quality reading materials [1][3]. Search system approaches are commonly used to find books using keywords, alongside recommendation systems to enhance user experience through content personalization [4][5]. One of the challenges in search systems is the system's ability to provide high-quality and reliable results.

A commonly used approach in search systems is Best Matching 25 (BM25), which works by calculating the relevance between the query text and the document [6]. As a foundational function in information retrieval, BM25 rapidly generates candidate items during searches across large document collections, but it cannot consider the quality of the items [5][6]. The re-ranking technique is a stage used to alter the order of candidates from the initial ranking results by calculating input from other signals that the search engine cannot capture [5]. The utilization of other signals for re-ranking, such as boosting features like the number of book copies or publication year on BM25 book search candidates, can help improve the quality of items in its search results [7]. However, this still employs manual feature score weighting on the search results, which relies on weight values and does not learn feature patterns such as the interaction of rating counts, publication years, and genres on Goodreads. Learning book metadata patterns can produce a consistent book quality score function that does not depend on manual guesswork of feature values.

In addition to the need for relevant and high-quality search results, there is an emerging requirement for the system's ability to explain why or how a model generates its predictions. Black-box models, such as gradient boosting, generally offer high performance and accuracy; however, their internal processes tend to be complex and difficult for users or practitioners to interpret, which leads to a decline in system trust and reliability [8][9]. In information retrieval, the aspect of explanation is required to build transparent and trustworthy systems, which is essential for system development and auditing [10]. The application of eXplainable AI (XAI) in machine learning systems can enhance system transparency, thereby assisting users or practitioners in understanding how the model yields its prediction results. This capability is beneficial for increasing trust in the system as well as gaining insights into the model's behavior in processing data, which is useful for model auditing or debugging [11][12].

Approaches to explaining black-box models generally utilize post-hoc methods that generate explanations from the outputs of the model's processes. XAI research focusing on black-box model explanations using methods such as LIME demonstrates reasonably good local explanations regarding the model's decision outcomes; however, limitations exist as these explanations take the form of approximations of the model's actual outputs [13][14]. ost-hoc explanations produce interpretations that do not accurately reflect the model's internal logic, which can cause misinterpretation and provide explanations that do not align with the true black-box logic [14][15][16]. These limitations encourage the implementation of transparent or glass-box models, whose internal processes can be interpreted inherently without supplementary post-hoc models, thereby making it easier for users or practitioners to understand how the model learns from data and makes decisions [15] The aspect of faithfulness (honesty) in an explanation is valuable for determining whether the generated explanation accurately reflects the model's true logic [17]. In digital libraries managing thousands of book collections, this explanatory capability is useful for observing how the model learns the utilized book data, such as detecting any model bias toward specific features within the search results.

The Explainable Boosting Machine (EBM) model was developed to address the limitations of black-box models regarding transparency and system interpretability. The internal processes of EBM are based on the Generalized Additive Model (GAM) with an additive structure, which clearly demonstrates the model's workflow through the summation of each utilized feature calculation. Furthermore, EBM's computational outcomes can be explained by decomposing the contribution score of each feature for a given item prediction. This capability illustrates the influence and contribution of each feature to the prediction score both locally, to explain individual item prediction scores, and globally, to clarify the impact of each feature on predictions across the entire dataset [18][19][20].

In addition to its interpretability capabilities, EBM functions as a prediction model whose final score can be utilized to re-rank BM25 candidates by learning the interaction patterns of general user ratings toward book features This two-stage practice aligns with the Learning to Rank (LtR) paradigm employed in search engine systems, where the LtR model learns feature signals from initial candidates to improve the ordering of more relevant items [5][7]. The application of re-ranking on BM25 retrieval by incorporating other metadata feature signals, along with the integration of black-box models like XGBoost, demonstrates that BM25 is still widely used as a foundational search function due to its strong performance; however, a glass-box reranker model has not yet been implemented [5][6] Meanwhile, EBM delivers prediction accuracy that often approaches or matches that of XGBoost, while providing inherent explanations that are valuable for accurate system transparency [21][22].

The Goodreads platform provides interaction data that can be utilized as a proxy for general quality because it represents real user evaluations and offers abundant interaction data. However, utilizing the book's average rating value directly as a prediction target is limited, as the value conceals the distribution of user ratings, such as differing rating patterns across genres and varying user opinions within individual rating values. Furthermore, a target approach based purely on a book's average rating cannot transparently explain the contribution of book metadata features, making it difficult for system developers to audit.

In this study, EBM is utilized to learn the interaction patterns of Goodreads users toward book feature characteristics to re-rank BM25 candidates based on general community book quality predictions. Simultaneously, it decomposes the book feature contributions toward the prediction score, which developers can leverage to gain model insights for debugging, auditing, and utilizing book features to enhance future system relevance. Therefore, this study proposes a combination of BM25 as the search engine and EBM as the LtR model to improve candidate ordering, ensuring higher quality based on rating predictions as an estimation of general Goodreads book quality. This architecture also includes inherent interpretation within the system to explain feature contributions without compromising accuracy.

1.2 Research Questions

This study addresses the following research questions:

1. How does the application of Explainable Boosting Machine (EBM) to BM25 search results affect the ranking quality of candidate books by re-ranking them based on predicted rating scores learned from Goodreads user rating interaction patterns and book metadata?
2. How do the contributions of individual book features and their feature interactions influence the predicted rating scores of BM25 candidate books?
3. How are local and global explanations in Explainable Boosting Machine related in explaining the contribution scores of individual book features to the final prediction score?

1.3 Research Objectives

This study aims to implement a two-stage ranking architecture that combines Best Matching 25 (BM25) as the retrieval stage and Explainable Boosting Machine (EBM) as the re-ranking model to improve the ranking quality of book search results based on the predicted quality (ratings) learned from the Goodreads community.

Furthermore, this study aims to analyze the contribution of individual book metadata features and their feature interactions to the prediction scores generated by Explainable Boosting Machine. It also investigates the relationship between local and global explanations in interpreting feature contributions and validates the faithfulness of the model through an ablation study.

1.4 Research Contributions

This study is expected to provide the following contributions:

1 To contribute to Explainable Artificial Intelligence (XAI) research by applying a glass-box model, namely Explainable Boosting Machine (EBM), as a re-ranking model that predicts book quality on top of the BM25 retrieval function in digital book search.

2 To provide a book search framework enriched with a quality-aware re-ranking module based on community rating predictions and interpretable feature explanations. In addition, the proposed framework facilitates the analysis of model behavior on the dataset, including the identification of potential feature bias through model explanations.

1.5 Scope and Limitations

1. The study utilizes the publicly available UCSD Goodreads Book dataset, which contains Goodreads user rating interactions and book metadata, including book titles, average ratings, rating counts, publication years, and genre information required by the Explainable Boosting Machine (EBM). In addition, external book data obtained from the Open Library API are used to evaluate the prediction capability of the trained EBM model on unseen data. The study focuses exclusively on English-language books to ensure consistency with the majority of the dataset, maintain uniform text representation during BM25 indexing, and reduce the complexity introduced by multilingual data.
2. The developed system is a book search system that accepts simple English-language queries and provides quality-aware recommendations based on Goodreads community ratings. Given a textual query, the system retrieves candidate books using BM25 and automatically predicts their quality scores using the EBM model. The prediction model is trained using historical user ratings from the Goodreads dataset as the target variable to estimate book quality from metadata features.
3. The explainability component of EBM is presented through visualization of shape functions that illustrate the contribution of metadata features, including popularity, publication year, genre, and feature interactions, to the predicted rating score. The consistency between feature values and their contribution scores is analyzed. Furthermore, both local and global explanations are employed to investigate how individual feature contributions influence the re-ranking of BM25 search results.
4. The proposed system is evaluated using both prediction and ranking metrics. Prediction performance is measured using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Ranking quality is evaluated using Precision@K and Normalized Discounted Cumulative Gain (NDCG@K) with relevance labels derived from a quality-based proxy, where books with an average rating (avg_rating) of at least 4.0 are considered relevant, while graded relevance levels are applied for NDCG evaluation. The explainability component is evaluated by examining the consistency of feature explanations across different testing scenarios and by analyzing the contribution of metadata features through both local and global explanation visualizations.