REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technologies

REINFORCEMENT LEARNING WITH MULTI-ARMED BANDITS FOR DYNAMIC CONTENT OPTIMIZATION OF WEB PERFORMANCE

Hanan Qahtan Hussein Al-ZUHAIRI

Master's Thesis

Supervisor

Prof. Dr. Osman Nuri UÇAN

İstanbul, 2024

REINFORCEMENT LEARNING WITH MULTI-ARMED BANDITS FOR DYNAMIC CONTENT OPTIMIZATION OF WEB PERFORMANCE

Hanan Qahtan Hussein Al-ZUHAIRI

Information Technologies

Master’s Thesis

ALTINBAŞ UNIVERSITY

2024

The thesis titled REINFORCEMENT LEARNING WITH MULTI-ARMED BANDITS FOR DYNAMIC CONTENT OPTIMIZATION OF WEB PERFORCMENT prepared by HANAN QAHTAN HUSSEIN AL-ZUHAIRI and submitted on 24/01/2024 has been **accepted unanimously** for the degree of Master of Science in Electrical and Computer Engineering.

_____

Prof. Dr. Osman Nuri UÇAN

Supervisor

Thesis Defense Committee Members:

Prof. Dr. Osman Nuri UÇAN	Department of Electrical and Electronic Engineering, Altınbaş University	_____
Assoc. Prof. Dr. Sefer KURNAZ	Department of Computer Engineering, Altınbaş University	_____
Asst. Prof. Dr. Serdar KARGIN	Department of Biomedical Engineering, İstanbul Arel University	_____

I hereby declare that this thesis meets all format and submission requirements of a Master's Thesis.

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Hanan AL-ZUHAIRI

Signature

DEDICATION

I dedicate and promise this research project to my supervisor, who is significant for guiding me during the entire research project, as well as to my family, who has always supported me during difficult times.

PREFACE

First and foremost, I would like to express my gratitude to my dissertation supervisor, Prof. Dr. Osman Nuri UÇAN, for his guidance and assistance. I got a lot better at grasping the reasoning behind the study by talking about my progress, issues. The technical need of this study activity became clearer to me as a result.

ABSTRACT

REINFORCEMENT LEARNING WITH MULTI-ARMED BANDITS FOR DYNAMIC CONTENT OPTIMIZATION OF WEB PERFORMANCE

AL-ZUHAIRI, Hanan

M.Sc. Information Technologies, Altinbaş University,

Supervisor: Prof. Dr. Osman Nuri UÇAN

Date: January / 2024

Pages: 70

The field of web performance optimization faces the challenge of dynamically optimizing web content to enhance user experience and maximize engagement. Traditional approaches to content optimization, such as static A/B testing or rule-based algorithms, often fall short in adapting to the dynamic nature of web content and fail to provide personalized experiences that align with user preferences. As a result, organizations struggle to achieve optimal web performance, leading to lower user satisfaction, decreased retention rates, and missed conversion opportunities. Additionally, the rapid growth of internet usage and the increasing demand for fast and efficient web experiences further intensify the need for effective content optimization strategies. Websites must continuously evaluate and adjust their content configuration to deliver engaging experiences that meet user expectations. However, manually identifying the most rewarding content variants in real-time becomes a daunting task as the number of potential combinations increases. Hence, there is a pressing need for an intelligent and data-driven approach that can dynamically optimize web content to maximize user engagement and improve web performance. this research aims to explore the application of reinforcement learning techniques with multi-armed bandits for dynamic content optimization of web performance.

Keywords: Dynamic Content Optimization, Web Performance, Python, User Experience, Engagement.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

ASEO	:	Advanced Search Engine Optimization
Q	:	Quantities Of Knowledge Unites
DQN	:	Deep Q learning
ITS	:	Information Technology Service
MAB	:	Multi-Armed Bandit
POMDP	:	Partially Observable Markov Decision Process
PR	:	PageRank
RL	:	Reinforced Learning
SEO	:	Search Engine Optimization
SERPs	:	Search Engine Results Pages
ZPD	:	Zone of Proximal Development
HTML	:	Hyper Text Markup Language
KPI	:	Key Performance Indicator
MLE	:	Maximum Likelihood Estimation
CTR	:	Click Through Rate
AM	:	Automation Maximum
AUC	:	Area Under Curve
BP	:	Back Propagation
CE	:	Cross Entropy
CDF	:	Cumulative Distribution Function
UI	:	User Interface
URL	:	Uniform Resource Locator

PR	:	Page Rank
KCs	:	Knowledge Tracing
KPI	:	Key Performance Indicator
STI	:	Science Technology Innovative

1. INTRODUCTION

1.1 BACKGROUND

The field of web performance optimization has been a subject of significant interest and research in recent years. With the exponential growth of internet usage and the increasing demand for fast and efficient web experiences, organizations are constantly seeking innovative approaches to enhance the performance of their websites. One prominent area of study within this domain is the application of reinforcement learning techniques, specifically with multi-armed bandits, to dynamically optimize web content and improve overall web performance. To understand the background of this topic, it is essential to explore the concepts of reinforcement learning and multi-armed bandits. Reinforcement learning is a branch of machine learning that focuses on training an agent to make sequential decisions in an environment to maximize a long-term reward. Unlike supervised learning, where the agent is provided with labeled data, reinforcement learning relies on trial-and-error interactions with the environment to learn optimal decision-making policies. This makes it particularly suitable for scenarios where the agent must adapt and optimize its actions based on feedback received from the environment.

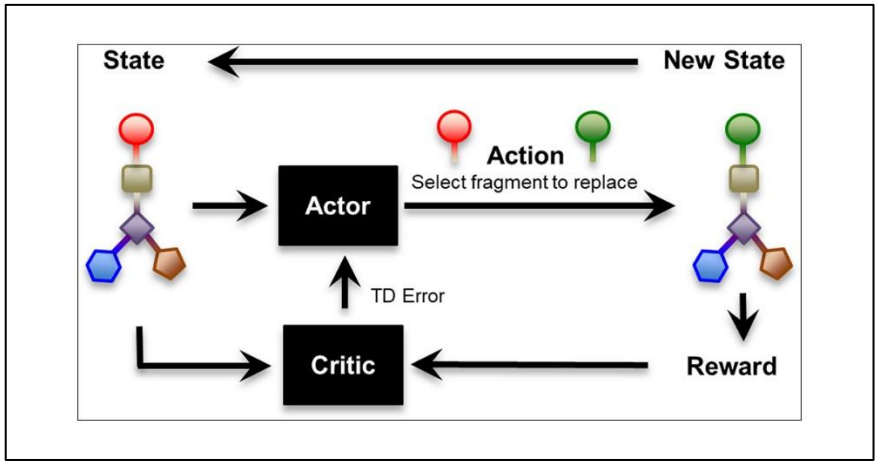

Figure 1.1: Feedback Learning in RL [1].

Multi-armed bandits, a classic problem in the field of reinforcement learning, refer to a collection of slot machines (or "one-armed bandits") with different payout probabilities. The goal is to determine the best strategy for allocating resources to these machines to maximize the cumulative reward over a series of plays. Each play represents a decision, such as displaying a particular variant of web content to a user, and the reward corresponds to the

user's response or engagement level. By employing multi-armed bandit algorithms, organizations can dynamically explore and exploit different content variants to find the most effective combination for optimizing web performance. In the context of web performance optimization, the dynamic nature of web content necessitates a continual evaluation of various strategies to identify the most effective content configuration. The primary objective is to provide website visitors with a personalized and engaging experience that aligns with their preferences, thereby increasing user satisfaction, retention, and conversion rates. Traditional approaches to content optimization often rely on static A/B testing or rule-based algorithms, which may not efficiently adapt to changing user preferences or traffic patterns. In contrast, reinforcement learning with multi-armed bandits offers a more adaptive and data-driven approach to content optimization. By combining the principles of reinforcement learning and multi-armed bandits, organizations can leverage data-driven decision-making to dynamically optimize web performance. The agent, in this case, could be an intelligent algorithm that continuously explores different content options and learns from user interactions to identify the most rewarding content variants. Through an iterative process of exploration and exploitation, the algorithm can adapt its content allocation strategy in real-time, maximizing user engagement and achieving the desired web performance objectives. Furthermore, the application of reinforcement learning with multi-armed bandits for web content optimization aligns with the growing trend towards personalization and dynamic content delivery. In an era where users expect tailored experiences, websites that can dynamically adjust their content based on real-time feedback can gain a competitive edge. By intelligently optimizing web performance through reinforcement learning, organizations can deliver content that is not only highly relevant but also aligns with the evolving preferences and needs of their users. In summary, the concept of using reinforcement learning with multi-armed bandits for dynamic content optimization of web performance represents a cutting-edge approach in the field of web performance optimization. By leveraging data-driven decision-making and adaptive strategies, organizations can enhance user experiences, improve engagement metrics, and ultimately achieve their web performance objectives in an ever-evolving digital landscape.

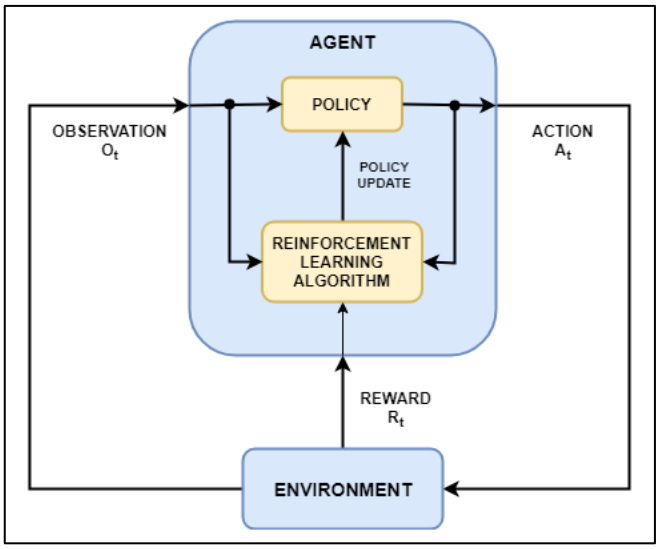

Figure 1.2: Agent Based RL System [2].

1.2 PROBLEM STATEMENT

The field of web performance optimization faces the challenge of dynamically optimizing web content to enhance user experience and maximize engagement. Traditional approaches to content optimization, such as static A/B testing or rule-based algorithms, often fall short in adapting to the dynamic nature of web content and fail to provide personalized experiences that align with user preferences. As a result, organizations struggle to achieve optimal web performance, leading to lower user satisfaction, decreased retention rates, and missed conversion opportunities. Additionally, the rapid growth of internet usage and the increasing demand for fast and efficient web experiences further intensify the need for effective content optimization strategies. Websites must continuously evaluate and adjust their content configuration to deliver engaging experiences that meet user expectations. However, manually identifying the most rewarding content variants in real-time becomes a daunting task as the number of potential combinations increases. Hence, there is a pressing need for an intelligent and data-driven approach that can dynamically optimize web content to maximize user engagement and improve web performance. This approach should consider the dynamic nature of web content, adapt to changing user preferences and traffic patterns, and make informed decisions based on real-time feedback. In light of these challenges, this research aims to explore the application of reinforcement learning techniques with multi-armed bandits for dynamic content optimization of web performance.

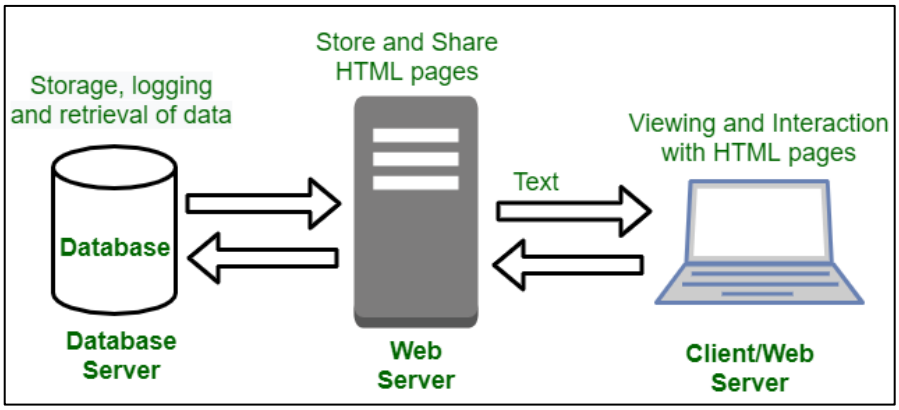

Figure 1.3: Dynamic Web Content[2].

By leveraging the principles of reinforcement learning, organizations can develop intelligent algorithms capable of continuously exploring different content options, learning from user interactions, and dynamically allocating resources to maximize user engagement. The objective is to identify an effective strategy that adapts to changing circumstances and provides personalized, relevant content to website visitors. The successful implementation of reinforcement learning with multi-armed bandits for web content optimization would contribute to improving web performance metrics such as user engagement, retention rates, conversion rates, and overall user satisfaction. Additionally, it would enable organizations to deliver tailored experiences that align with the evolving preferences and needs of their users, thereby gaining a competitive advantage in today's digital landscape. Overall, this research addresses the critical problem of dynamically optimizing web content to enhance web performance and user experience. By integrating reinforcement learning techniques with multi-armed bandits, this study aims to provide a solution that empowers organizations to achieve optimal content configuration, maximize user engagement, and improve web performance in a rapidly evolving digital environment.

1.3 THESIS OBJECTIVES

The main objective of this thesis is to explore the application of reinforcement learning techniques with multi-armed bandits for dynamic content optimization of web performance. To achieve this overarching objective, the following specific objectives will be pursued.

Conduct a comprehensive literature review: This objective involves conducting an in-depth review of existing literature related to web performance optimization, reinforcement learning, and multi-armed bandits. The review will encompass relevant theories, methodologies, algorithms, and case studies to establish a strong theoretical foundation for

the research. Analyze the requirements and challenges of web content optimization: This objective focuses on identifying the specific requirements and challenges associated with dynamically optimizing web content for improved performance. It involves analyzing factors such as user preferences, traffic patterns, content variability, and real-time feedback to gain a comprehensive understanding of the complexities involved in web content optimization. Develop a reinforcement learning framework with multi-armed bandits: This objective involves designing and implementing a reinforcement learning framework that integrates multi-armed bandits algorithms. The framework will be tailored to address the dynamic nature of web content optimization, enabling continuous exploration and exploitation of content variants based on user interactions and feedback. Collect and analyze real-time data: This objective focuses on gathering relevant data from real-world web performance scenarios. Data collection methods may include web analytics tools, user tracking mechanisms, and experimental setups. The collected data will be analyzed to gain insights into user behavior, engagement metrics, and the effectiveness of different content configurations. Evaluate and compare performance metrics: This objective aims to evaluate the effectiveness of the proposed reinforcement learning framework with multi-armed bandits for web content optimization. Various performance metrics such as user engagement, retention rates, conversion rates, and overall user satisfaction will be considered to assess the impact of the optimization strategy. Conduct comparative analysis with traditional approaches: This objective involves comparing the performance of the reinforcement learning framework with multi-armed bandits to traditional approaches, such as static A/B testing or rule-based algorithms. The comparative analysis will help highlight the advantages and limitations of the proposed approach in terms of adaptability, scalability, and performance improvement. Provide recommendations and guidelines: This objective focuses on deriving practical recommendations and guidelines based on the findings of the research. These recommendations will provide insights for organizations seeking to implement dynamic content optimization strategies using reinforcement learning with multi-armed bandits. The guidelines will encompass best practices, potential challenges, and implementation considerations to facilitate successful adoption.

Contribute to the field of web performance optimization: The final objective of this thesis is to contribute to the existing body of knowledge in the field of web performance optimization. The research outcomes, including the developed framework, comparative analysis, and

recommendations, will serve as valuable contributions to both academia and industry. The findings will help advance the understanding of dynamic content optimization and inform future research in the area. By accomplishing these objectives, this thesis aims to provide valuable insights into the application of reinforcement learning with multi-armed bandits for dynamic content optimization of web performance. The research outcomes will contribute to the improvement of web performance, user experience, and overall satisfaction in the digital landscape.

1.4 THESIS MOTIVATION

The motivation behind this thesis is driven by several factors and considerations that highlight the importance and relevance of exploring reinforcement learning with multi-armed bandits for dynamic content optimization of web performance. The following motivations have shaped the research direction: Evolving web landscape: The web landscape is continuously evolving, with a growing number of websites and increasing user expectations for seamless and personalized experiences. This necessitates the development of innovative strategies to optimize web performance and deliver content that aligns with user preferences. The motivation to explore reinforcement learning with multi-armed bandits stems from the need to leverage intelligent algorithms that can adapt to changing circumstances and dynamically allocate resources to maximize user engagement. Limitations of traditional approaches: Traditional approaches to web content optimization, such as static A/B testing or rule-based algorithms, often face limitations in adapting to dynamic content configurations and real-time user feedback. These approaches may not effectively capture the complexity of user preferences or optimize content allocation for optimal performance. The motivation to explore reinforcement learning with multi-armed bandits arises from the desire to overcome these limitations and develop a more adaptive and data-driven approach to content optimization. Potential for personalized experiences: Personalization has become a key driver in user engagement and satisfaction. By delivering tailored experiences, websites can enhance user engagement, improve conversion rates, and establish stronger connections with their audience. The motivation to explore reinforcement learning with multi-armed bandits for dynamic content optimization stems from the potential to create personalized experiences by continually exploring and exploiting content variants based on real-time feedback, thus meeting user expectations and preferences more

effectively. Data-driven decision-making: In the era of big data and advanced analytics, organizations have access to vast amounts of data that can inform decision-making processes. Leveraging this data through reinforcement learning with multi-armed bandits can enable intelligent content allocation decisions based on real-time user interactions and feedback.

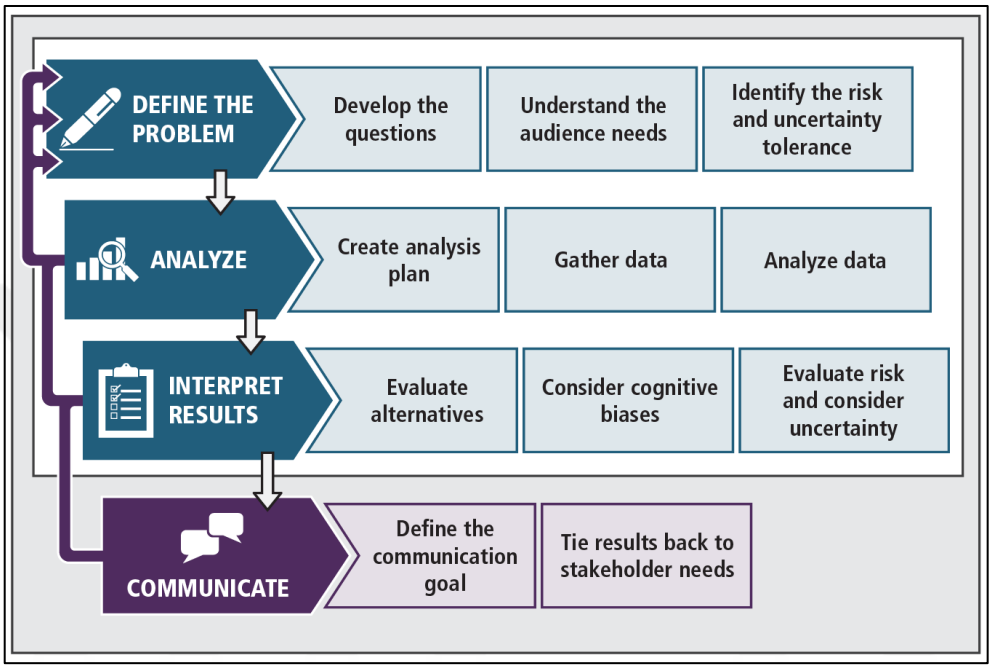

Figure 1.4: Data-Driven Decision-Making [4].

The motivation behind this research is to harness the power of data-driven decision-making to optimize web performance, improve user engagement, and drive better business outcomes. Advancements in reinforcement learning: Reinforcement learning has gained significant attention and advancements in recent years, demonstrating its potential in various domains. The motivation to explore its application in web performance optimization arises from the desire to leverage the advancements in reinforcement learning algorithms, techniques, and frameworks to address the challenges of dynamic content optimization. By adapting these techniques to the context of web performance, organizations can benefit from more efficient and effective content allocation strategies. Industry demand and practical implications: Web performance optimization is a critical concern for organizations across industries. The motivation to explore reinforcement learning with multi-armed bandits for dynamic content optimization stems from the practical implications and potential benefits it offers to organizations seeking to improve web performance metrics, enhance user

experiences, and achieve their business goals. The research outcomes can provide valuable insights, recommendations, and guidelines that can be directly applied by practitioners in the field to enhance their content optimization strategies. Overall, the motivations behind this thesis lie in the need for adaptive, data-driven, and personalized approaches to web performance optimization. By exploring reinforcement learning with multi-armed bandits, the research aims to address the limitations of traditional approaches, leverage advancements in reinforcement learning, and contribute to the field of web performance optimization, ultimately benefiting organizations and users in the ever-evolving digital landscape.

1.5 THESIS CONTRIBUTIONS

This thesis makes several contributions to the field of web performance optimization and dynamic content optimization through the application of reinforcement learning with multi-armed bandits. The key contributions are as follows:

Novel framework for dynamic content optimization: The thesis presents a novel framework that integrates reinforcement learning techniques with multi-armed bandits for dynamic content optimization in web performance. The framework offers a data-driven and adaptive approach to continuously explore and exploit different content variants based on real-time user interactions and feedback. This contribution provides a practical and innovative solution to address the challenges of optimizing web content in a dynamic and evolving digital landscape.

Comparative analysis with traditional approaches: The thesis conducts a comprehensive comparative analysis between the proposed reinforcement learning framework with multi-armed bandits and traditional approaches, such as static A/B testing or rule-based algorithms. The analysis highlights the advantages and limitations of the proposed approach in terms of adaptability, scalability, and performance improvement. This contribution provides valuable insights into the effectiveness and superiority of the reinforcement learning framework in dynamic content optimization for web performance.

Evaluation of performance metrics: The thesis evaluates the performance of the reinforcement learning framework with multi-armed bandits using various performance metrics, including user engagement, retention rates, conversion rates, and overall user satisfaction. The evaluation provides empirical evidence of the impact of the optimization strategy on web performance metrics, demonstrating the effectiveness of the proposed

approach in enhancing user experiences and achieving web performance objectives. Practical recommendations and guidelines: Based on the research findings, the thesis provides practical recommendations and guidelines for organizations seeking to implement dynamic content optimization strategies using reinforcement learning with multi-armed bandits. These recommendations encompass best practices, potential challenges, and implementation considerations, enabling practitioners to apply the research outcomes effectively in their own web performance optimization efforts. This contribution facilitates the practical implementation and adoption of the proposed approach in real-world scenarios.

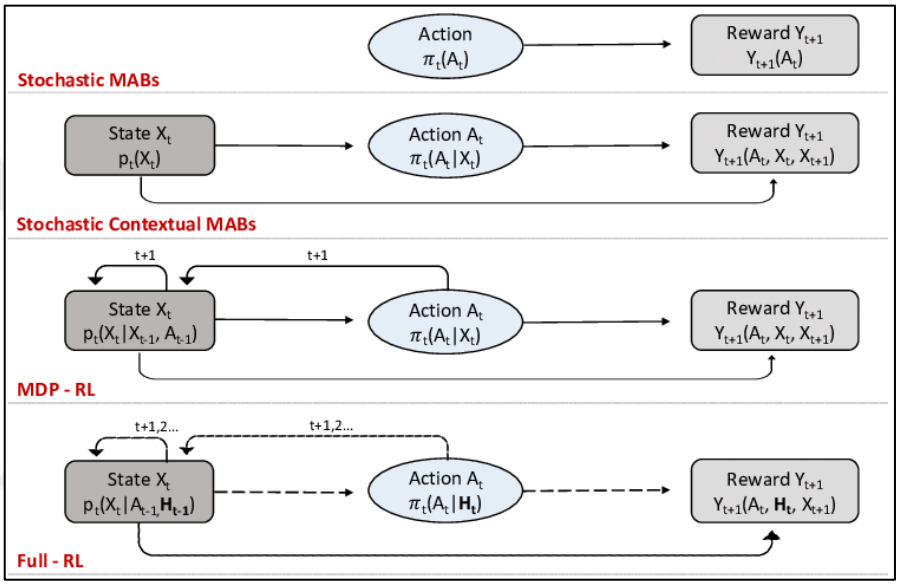

Figure 1.5: RL with MAB [5].

Contribution to academic knowledge: The thesis contributes to the academic knowledge in the field of web performance optimization by exploring the application of reinforcement learning techniques with multi-armed bandits. The research outcomes, including the developed framework, comparative analysis, and empirical evaluation, contribute to the understanding of dynamic content optimization and its impact on web performance metrics. This contribution advances the knowledge base and provides a foundation for future research and studies in the field. Impact on industry practices: The practical implications of the research are significant, as they provide industry practitioners with insights and strategies to optimize web performance through dynamic content allocation. By incorporating the findings of the thesis, organizations can improve user engagement, retention rates, and conversion rates, leading to enhanced business outcomes. The research outcomes have the potential to positively impact industry practices by driving the adoption of intelligent and

data-driven approaches to web performance optimization. In summary, this thesis contributes to the field of web performance optimization by introducing a novel framework, conducting a comparative analysis, evaluating performance metrics, providing practical recommendations, advancing academic knowledge, and influencing industry practices. These contributions collectively enhance the understanding and implementation of dynamic content optimization strategies using reinforcement learning with multi-armed bandits, ultimately benefiting organizations, users, and the broader field of web performance optimization.

2. LITERATURE REVIEW

2.1 INTRODUCTION

There has been a large amount of research done into the many ranking variables and how they influence the position of a website inside the results of a search engine. This research has been done in order to better understand how search engines place websites in their results. The placement of websites in the results provided by search engines has been the primary focus of this research. The overwhelming majority of search engines only provide a small portion of the criteria that are taken into account by their individual ranking algorithms. This has resulted in a wealth of research that aims to identify particular ranking variables and investigate the impact that these variables have on the search engine results pages. As a direct result of this, a significant amount of investigation is being carried out at the moment.

2.2 RELATED WORKS

The authors of the article [21], discovered a connection between the ranks that a website achieves in a search engine and the degree to which it is user-friendly on the whole. They came to the conclusion that one of their key aims, which was to boost the website's usability, could be accomplished by incorporating SEO best practices in a website. This was one of their primary goals [22]. In order to have a better understanding of the dynamic relationship between these components, research was conducted to explore the impact, on search engine rankings, of keyword density as well as keyword position within the text of web pages. This was done with the goal of determining how these factors interact with one another. It has been hypothesized that the optimum number of times a keyword should be used is. In addition to that, they developed a method for predicting rankings on other websites located all across the internet by using an algorithm. The authors of [3], took rank prediction one step further by applying techniques from "reverse engineering" to examine important websites in order to discover which aspects of those websites contribute to their high search engine ranks. This allowed them to take rank prediction one step closer to being a reality. They were able to determine, as a result of this, which characteristics of well-known websites contribute to those sites' high rankings in search engines. The strategy that was utilized in order to accomplish this objective was to examine each website in light of its counterpart. There are various on-page ranking factors, all of which are addressed in length in [4][5].

Within these references, there is also a comprehensive description of the manner in which the position of a page in the search engine results is affected by these components. The authors of the study carried out research to examine whether or not there was a connection between the rank of a page in the search engine results pages (SERPs) and the number of times specific keywords appeared in the page's title and text. According to the conclusions of their analysis, the ideal number of times to put a keyword in the title of a page is three times. This number was determined by looking at how well different keyword combinations performed. This is the case since it has been proved to have the greatest influence on the ranking of the page. Consequently, this is the case. When a page's title had more common words, the page's performance as a whole was impacted in a negative way. Web pages that utilized keywords more frequently throughout the entirety of their texts had a higher score in the keyword analysis. This was done to ensure that relevant content was being returned to users. The writers of [5], conducted research with the goal of studying the influence that meta data has on the position that a website retains in the results that are produced by search engines. They feel that the content of the HTML "title" tag, the content of the HTML "meta subject" element, and the content of the HTML "meta description" tag are the three bits of meta data that are the most important. According to the findings, pages that featured a sizeable quantity of meta data performed noticeably better in comparison to pages that were functionally comparable but did not contain such information. This was the case regardless of whether or not the pages were identical in any other way. In addition to this, quite a few further variations of the analysis were carried out on the meta data that was provided on the pages. Pages that featured all three varieties of meta data fared considerably better in search engine rankings in comparison to pages that included only two varieties of meta data or just one variety of meta data. The examinations that are discussed in [22], went through additional development in a wide variety of different directions. The authors demonstrated that it was possible, using only a small number of variables, to accurately estimate the positions that seven out of ten pages would hold in the search engine results. This was nevertheless accomplished despite the fact that there were just a few variables to work with. Their methodology placed a significant amount of emphasis on the off-page component known as "PR" (PageRank), which was followed by the on-page components known as "HOST" (domain name), "TITLE" (title tag in HTML), "M_DES" (meta-description of the HTML content), and "PATH" (the URL path that contains the file name of the document).

The authors of the same study investigated the ratings offered by regular users in addition to the ratings provided by SEO professionals. Following are some of the factors that were identified as ranking factors based on the responses of 37 industry experts: keywords included in the "title" tag, anchor text included in inbound links, global popularity of the web page URL, web page age, popularity of internal links, topic relevance of incoming links, popularity of links within topic groups, keywords included in the web page body, global popularity of linked pages, tense, and keywords included in the web page body. The authors of [3], used reverse engineering to determine the five most important ranking variables, which are as follows: the length of the URL, the presence of the keyword in the domain name, the frequency of the keyword in the H1 tag, and the frequency of the keyword in the page title ("title" tag), along with the number of URL layers. Similar findings about the significance of ranking variables were reported in [23][24][25][26]. As a result, researchers came to the conclusion that the majority of ranking criteria had not materially altered over the course of the last several years. This was shown by the fact that [23][24][25][26], comprised the bulk of the ranking criteria that required to be satisfied to be considered. The author of [27], suggests for a metric that takes into consideration not only the conventional ranking signals but also the degree to which a page conforms to the best practices for SEO. This metric would take into account both the traditional ranking signals as well as the degree to which a page conforms to these best practices. This metric would take into consideration not just the conventional indicators of ranking but also the degree to which a page adheres to these best practices. The measure that has been proposed is a linear combination of ranking components that are not learned but are, rather, predetermined. This combination does not take into account any new information. To provide a greater level of detail, this suggests that the standards for ranking have been decided upon in advance. In order to arrive at their findings, the authors of [28], carried out a comprehensive literature analysis of the SEO criteria that influence SERP ranks. After being analyzed, the 24 aspects of a website were broken down into two categories: on-page elements and off-page factors. The on-page components made up the majority of the list, followed by the off-page factors. It was discovered that the most important on-page features are the keywords that are used in a page's title tag as well as the keyword density that is used throughout the content of the page. This was true despite the fact that the page may or may not have been enhanced for search engine crawlers. Even though just two of [29], fourteen SEO criteria were off-site, which

refers to backlinks, eight of our on-site characteristics were the same as those that were used by that study. [29], study was conducted by [29]. This suggests that backlinks are an essential component of search engine optimization. When it comes to conducting research on SEO, the writers working in this subject have a tendency to reach a consensus on the on-page characteristics that should be considered to be of the utmost importance. This is one of the most common occurrences. The authors of [30], conducted interviews with industry professionals and surveys among those professionals in order to get expert opinions on the parameters that determine whether or not a website is search engine optimized. These interviews and surveys were conducted in order to gain expert opinions on the criteria. It was determined that the overarching structure of the website as well as the distribution of keywords strategically across the content of the website had the most influence on the results provided by the search engines. After this step, the application of meta tags was performed. This particular information may be found in the master's thesis [31], which is where extensive research on the factors that determine rankings are to be found. The author, in a method that was comparable to what we performed, surveyed a sizeable number of digital marketing companies in order to collect the perspectives of those businesses regarding the factors that impact ranking. It was revealed, based on the data that was collected, that the keyword frequency on the page rated third and fourth among the eight most essential ranking variables (with link-related criteria coming in first and social metrics coming in eighth). This was discovered after it was discovered that link-related criteria came in first. This was found out after it was revealed that the amount of times a keyword appears on a page was ranked third and fourth among the eight most important factors that affect ranking. Instead of conducting polls in order to identify which features are the most important, we came to the conclusion that the best way to find out which aspects are the most important would be to apply machine learning classification algorithms to manually annotated web pages. We arrived to this decision after coming to the realization that the best approach to find out which aspects are the most important would be to apply these algorithms. In addition, the search engine optimization (SEO), which is also sometimes referred to as "search engine optimization," was a key focus of the PhD dissertation [32]. This is because SEO is also sometimes referred to as "search engine optimization." In order to build an on-page SEO model based on previously successful SEO strategies, this thesis heavily relied on an experimental website case study as its primary source of data. This study was carried out as

part of the thesis. This was done in order to build an on-page SEO model based on successful SEO approaches that had been used in the past. In addition to that, this inquiry made use of methods that were predicated on keywords that were found in a variety of different areas of HTML code. According to the information that was presented in article [33], a machine learning algorithm was used to categorize websites into the aforementioned categories by making use of data gathered from text and HTML elements. This information was applied to classify websites. These attributes were taken just from the different websites themselves. The investigation arrived at the conclusion that the classification procedure does not give HTML tags nearly enough consideration. This is the case in spite of the fact that HTML tags are an essential component of techniques that are based on the frequency of keywords to identify the themes that are covered by web sites. An investigation of the classification of web pages without regard to the language they were written in was carried out in [34], which also addressed concerns of a like nature. They came up with a strategy that successfully classifies websites in accordance with the aims that such websites serve by utilizing deep learning as the method of classification. This approach is effective despite the fact that the websites in question may be written in a variety of languages. The authors of the paper [35], made use of a wide array of machine learning techniques in order to hazard an educated estimate as to where websites associated with the online gift business will go on search engine results pages (SERP). For the aim of this investigation, an examination was carried out on thirty different blogs that had been ranked on the first page of Google and were pertinent to the topic under investigation. Following an analysis of the machine learning models LightGBM and XGBoost, the data revealed that connections, domain security, and H3 headers had the greatest impact. On the other hand, the results of the analysis of the frequency of keywords revealed that there had not been any major shifts. In addition to this, we examined the total length of the text, the number of H1, H2, and H3 tags, the total lengths of the tags, as well as the frequency with which particular keywords were used in the tags, URLs, and meta descriptions. Over the course of the past several years, a large amount of research has been conducted on the subject of categorizing websites according to the myriad of functions that they provide for their users. A classic illustration of a widespread and typical categorization activity is the work that has been done [15][16][17], to categorize spam on the internet. Extraction of characteristics for the aim of accomplishing this goal has also been the focus of a significant amount of research [36]. This piece of work is a

contribution to the body of literature that already exists on the subject of the categorization of web pages. This is because classification that is based on SEO suggestions does not have enough research to support it. In [18], where the phrase "black hat" is used, unethical search engine optimization strategies are referred to as "black hat," which is an abbreviation for "black hat." Black hat SEO strategies include a wide variety of different tactics, some of which are keyword stuffing (increasing the phrase density), doorway sites (automatically generated pages to target certain keywords), content copying, hiding text or links, cloaking (creating unique page versions for search engines and humans), and many others. Black hat SEO strategies are considered unethical and should be avoided at all costs. Black hat search engine optimization is regarded unethical and is looked down upon by organizations who run search engines. As a direct result of the ever-evolving algorithms deployed by search engines, "black hat" search engine optimization is currently experiencing more competition than it has ever faced before. This is a significant shift from previous years. The research methods that were useful in the investigation of one subject can also be helpful in the investigation of this other subject. In recent years, search engine optimization has also been implemented in scholarly crawlers and indices. This practice is reaching ever-greater proportions of the population. When expanded out, the abbreviation "ASEO" refers to the full version of the term "Academic Search Engine Optimization." In addition to the rankings of scholarly research papers, the academic search engines such as Google scholar, Microsoft academia, Scopus, and other academic search engines are currently receiving the majority of the attention that is currently being directed at these types of resources. These difficulties were investigated by the authors of [37], who carried out research in the form of a study on the connection that exists between rankings and citations. As a consequence of their research, they arrived at the conclusion that one of the most important factors that is considered throughout the ranking process is the total number of citations received by a piece of work. Traditional SEO and advanced search engine optimization (ASEO) are in no manner comparable to one another, as all of us are aware. It requires the study of more issues and the adoption of further tactics, both of which are beyond the scope of this work. Neither of these can be covered in this particular piece of research. In modern SEO research, website audits carried out with the help of SEO tools are applied on a fairly regular basis, and the results are compared to rankings discovered in other areas. For example, in [38], the authors connected the search engine optimization (SEO) performance of institutions with their ranks

in the Shanghai edition of the Academic Ranking of World institutions. This was done in order to better understand the relationship between the two. According on the criteria used for this ranking, Shanghai has been determined to be the most important city in the world.

Table 2.1: Summary of Related Works.

Reference	Contribution
[21]	Identifies the connection between search engine ranks and user-friendly websites, emphasizing the importance of incorporating SEO best practices for usability enhancement.
[22]	Investigates the impact of keyword density and position on search engine rankings, contributing to the understanding of how these factors interact.
[3]	Utilizes reverse engineering to determine key ranking variables, including URL length, keyword presence in the domain name, H1 tag frequency, and title tag frequency.
[4][5]	Provides a comprehensive analysis of on-page ranking factors and their influence on search engine result page positions.
[28]	Conducts a literature analysis on SEO criteria influencing SERP ranks, categorizing factors into on-page elements and off-page factors.
[29]	Examines the importance of backlinks as an essential component of search engine optimization and highlights the consensus on on-page characteristics among SEO researchers.
[30]	Gathers expert opinions on website optimization parameters through interviews and surveys, focusing on website structure and keyword distribution.
[31]	Explores factors influencing rankings through machine learning classification algorithms applied to manually annotated web pages, emphasizing the practical approach.
[32]	Relies on an experimental website case study to build an on-page SEO model based on successful SEO strategies, incorporating meta tags and HTML elements.
[33]	Uses machine learning algorithms and data from text and HTML elements to classify websites based on their characteristics, revealing the importance of HTML tags.
[35]	Applies machine learning techniques to estimate search engine result page positions of online gift-related websites, highlighting the impact of connections, domain security, and header tags.
[37]	Investigates the relationship between rankings and citations, identifying the significance of the total number of citations received for search engine rankings.
[38]	Connects SEO performance with ranks in the Academic Ranking of World Institutions, providing insights into the relationship between SEO and institutional ranks.

3. MATERIALS AND METHODS

3.1 WEB CONTENT OPTIMIZATION

Web content optimization, also known as search engine optimization (SEO), is a comprehensive strategy employed to improve the visibility and ranking of a website on search engine result pages (SERPs). It involves a wide range of techniques and practices aimed at enhancing the website's performance, user experience, and relevance to search queries. In this extensive explanation, we will delve into the various aspects of web content optimization, including keyword research, on-page optimization, and off-page optimization, along with the importance of mobile optimization and user engagement. Keyword Research: Keyword research forms the foundation of web content optimization. It involves identifying the specific words and phrases that users are likely to type into search engines when looking for information or products related to a particular topic or industry. Effective keyword research requires a comprehensive understanding of the target audience and their search intent.

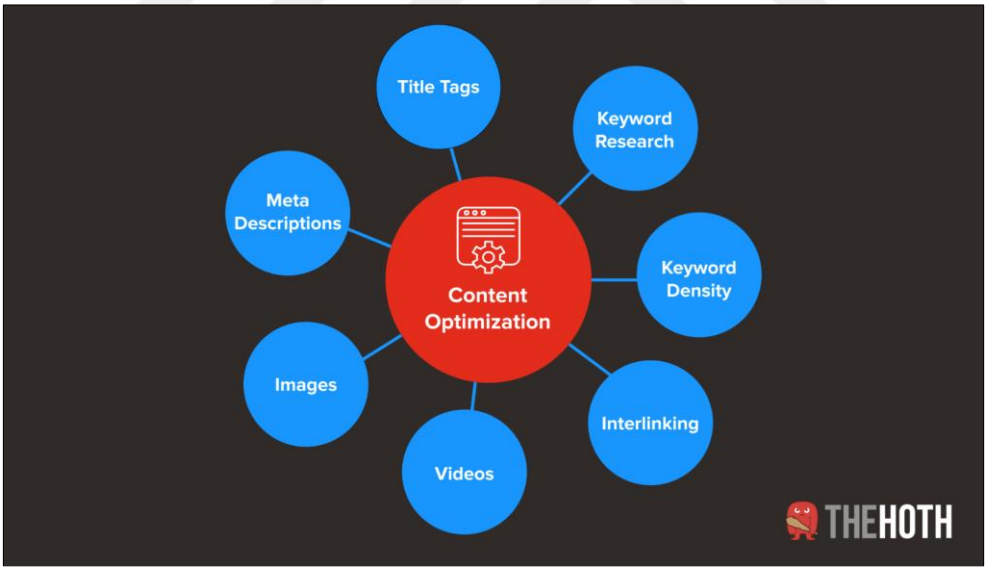

Figure 3.1: Web Content Optimization [7].

By conducting keyword research, website owners can identify high-volume and relevant keywords to integrate into their content, thereby increasing the likelihood of appearing in relevant search results. On-Page Optimization: On-page optimization refers to the practices implemented directly on the website's pages to improve its visibility and relevance. This includes optimizing meta tags (title, description) to accurately represent the content,

incorporating keywords strategically within the content, and organizing the information in a user-friendly and search engine-friendly manner. On-page optimization also involves ensuring that the website's URLs, headings, images, and internal links are optimized for better search engine crawling and indexing. By optimizing these on-page elements, website owners can improve their website's visibility, click-through rates, and overall user experience. Off-Page Optimization: Off-page optimization encompasses the techniques employed outside of the website to improve its visibility and reputation. The primary focus of off-page optimization is acquiring high-quality backlinks from authoritative and relevant websites. Backlinks act as votes of confidence from other websites, indicating the website's credibility and value. Building a diverse and natural backlink profile is crucial for search engines to perceive the website as trustworthy and authoritative.

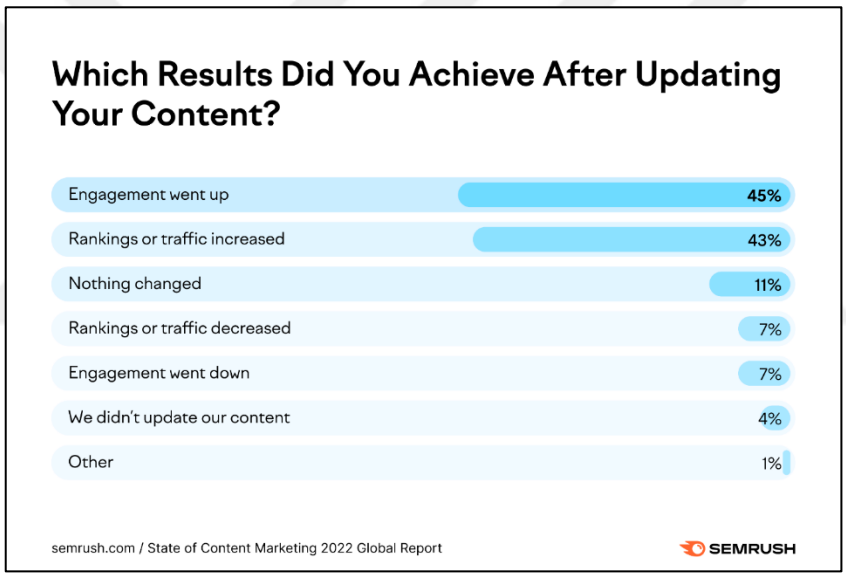

Figure 3.2: Off-Page SEO Based on User Voting [7].

Additionally, off-page optimization involves social media marketing, content marketing, influencer outreach, and other promotional activities to increase the website's exposure and attract more visitors. Mobile Optimization: With the widespread use of smartphones and tablets, mobile optimization has become a vital aspect of web content optimization. Websites that are not optimized for mobile devices may suffer from poor user experience, slow loading times, and lower search engine rankings. Mobile optimization involves creating responsive designs that adapt to different screen sizes, optimizing images and multimedia elements for mobile devices, and ensuring fast loading speeds. By prioritizing mobile optimization, website owners can provide an optimal browsing experience to mobile users and improve

their chances of ranking higher in mobile search results. User Engagement and Experience. User engagement and experience are increasingly critical factors in web content optimization. Search engines analyze user behavior metrics, such as bounce rate, time on site, and click-through rates, to determine the relevance and quality of a website. Engaging and user-friendly content that caters to the needs and preferences of the target audience is crucial. This involves creating compelling headlines, well-structured content, clear calls-to-action, and interactive elements that encourage users to stay longer on the website and explore further. By focusing on user engagement and experience, website owners can enhance their website's reputation, improve rankings, and drive more organic traffic. In conclusion, web content optimization is a multifaceted process that involves keyword research, on-page optimization, off-page optimization, mobile optimization, and prioritizing user engagement and experience. By implementing these strategies effectively, website owners can increase their website's visibility, attract more organic traffic, and ultimately achieve their online goals. Remember, the world of SEO is ever-evolving, and it's essential to stay updated with the latest trends and best practices to maintain a competitive edge in the digital landscape.

3.2 SUMMARY OF THE METHODOLOGY

In this article, we will investigate a method for increasing the effectiveness of intelligent tutoring systems by boosting their capacity for time management and motivation by means of the process of modifying and personalizing learning sequences. Specifically, we will focus on how this method might be implemented. This will make it possible for the systems to be more effective in terms of preserving the students' motivation and maximizing the amount of time students spend working on their academics. The system will constantly give the user with the action that, should they choose to carry it out, will considerably expedite their advancement in the event that they elect to carry it out. In the event that they decide to participate in that activity. We provide DQN, an approach that is dependent on exercise data, in addition to RL, a method that is dependent on a lot less prior information than DQN is. Both of these approaches are available to our customers. We recommend making use of both of these strategies because they both are founded on genuine estimations of the progress that the student has made, and we are recommending that you use both of them. In order for it to function properly, the framework makes use of three different methodological techniques.

To begin, it takes contemporary ideas of learning that is organically driven and adapts them so that they may be used to active learning. This is the first step in the process. Utilizing the individualized activities that are made available to users in order to generate an empirical estimate of the progress that the user is making can help achieve this goal. Second, it handles the exploration and exploitation phases of the optimization process in an effective and efficient manner by making use of cutting-edge MAB (Multi-Armed Bandit) algorithms. This allows it to manage these aspects of the process. The achievement of this goal is made possible through the utilization of the term MAB as the identifying label. Thirdly, despite the fact that the initial study on MAB has a fairly broad scope, it is vital to make use of past information in order to direct the inquiry. This is because the investigation will not be successful without it. Following the conclusion of a simulation with simulated users, the findings of an experiment involving 400 first graders were then carefully examined. The purpose of this experiment is to teach middle school students (those in grades 7-8) how to effectively budget their weekly allowances using a spreadsheet. The experiment consisted of two independent stages, each of which was conducted in a distinctive manner. It is necessary to make use of cutting-edge MAB (Multi-Armed Bandit) algorithms in order to properly manage the exploration and exploitation phases of the optimization process. This guarantees that the best possible outcomes will be obtained.

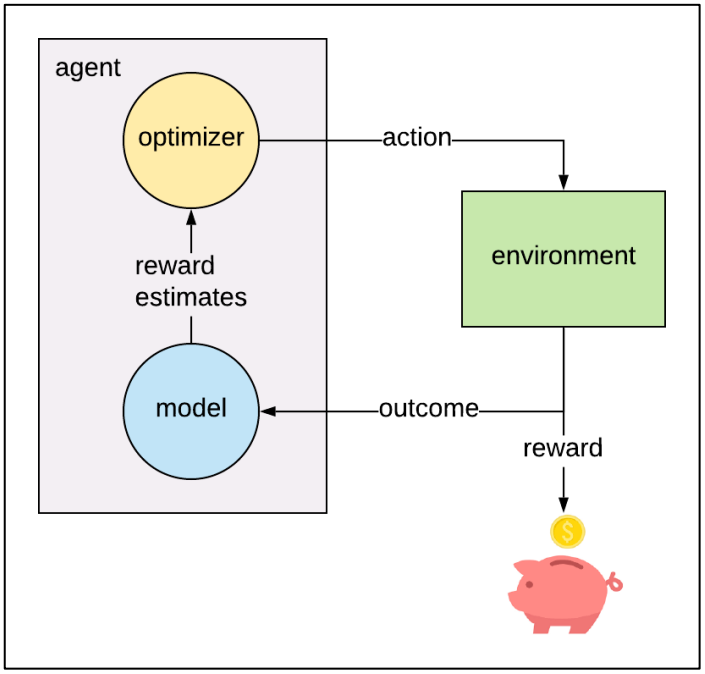

Figure 3.3: MAB Problem Based on Competitive Agents [12].

Thirdly, despite the fact that the initial study on MAB has a rather wide scope, it is necessary to make use of previous information in order to direct the investigation. For the aim of this experiment, middle school kids (those in grades 7-8) are instructed on how to properly budget their weekly allowances using a spreadsheet. Following the completion of a simulation with simulated users, the findings of an experiment involving 400 first graders were then carefully examined. The experiment had two distinct phases that were each carried out separately. In order to accomplish this, it makes use of cutting-edge MAB (Multi-Armed Bandit) algorithms. These algorithms grant it the ability to effectively manage the exploration and exploitation stage of the optimization process. Thirdly, despite the fact that the initial study on MAB has a rather wide scope, it is necessary to make use of previous information in order to direct the investigation. For the aim of this experiment, middle school kids (those in grades 7-8) are instructed on how to properly budget their weekly allowances using a spreadsheet. Following the completion of a simulation with simulated users, the findings of an experiment involving 400 first graders were then carefully examined. The experiment had two distinct phases that were each carried out separately. For the aim of this experiment, middle school kids (those in grades 7-8) are instructed on how to properly budget their weekly allowances using a spreadsheet. Following the completion of a simulation with simulated users, the findings of an experiment involving 400 first graders were then carefully examined. The experiment had two distinct phases that were each carried out separately. For the aim of this experiment, middle school kids (those in grades 7-8) are instructed on how to properly budget their weekly allowances using a spreadsheet. Following the conclusion of a simulation in which simulated users played a role, the findings of an experiment in which 400 kids in the first grade took part were analyzed in a study. The simulation was completed before anything else.

3.3 WEB OPTIMIZATION

uncharacteristically scant attention paid to a variety of educational and cognitive theories It can be difficult for users to understand all of the challenges that a user faces, precisely identify his strengths and weak points, and determine which activities can bring him the maximum gain in his learning because of the large number of users, the differences between them, and the amount of time that is available to the user. This is because of the amount of time that is available to the user. This is as a result of the difficulties that are encountered by

a user. It has been established that a sequence that is optimal for a typical user may not be useful for a lot of different people. Even when using computerized methods of analysis, it is difficult to select which parameters best reflect each individual user because of the fact that each user is unique. In some cases, it may even be impossible. Because of this, there is a possibility that cognitive and learner models will not be appropriate in the real world. This is a direct result of the previous point. As a result, it is to one's greatest advantage to rely on them as little as possible, since this will provide way for alternatives that require less precision and can be produced more rapidly. Consequently, it is necessary for the STI to try out a number of different methodologies in order to identify which ones are most likely to be of use to the student in terms of boosting their learning. Optimization strategies that have been demonstrated to work effectively we use approaches that don't require any prior knowledge about the learner and instead rely solely on estimates of how far along they are in each exercise. This allows us to avoid the need for any kind of testing. Because of this, there is no requirement for us to conduct any form of background research. To begin, let's take a look at the fundamental idea that the priority ought to be placed on selecting those activities and facets that are recognized to provide the most overall advantage in terms of one's educational experience. Effective online strategies will be used for this purpose since they can research a variety of learning activities, evaluate the growth they can offer to each particular user, and then take advantage of those online strategies that are most likely to encourage learning. As the user progresses, they will learn by doing so on their own. Increased amounts of drive and determination our strategy requires that, in every instance in which it is possible to do so, the activity that is chosen is the one that will be of the most educational benefit to the user. This must be done in order to ensure that our method is successful. This not only sets the way for the use of optimization strategies that are more effective, but it also gives individuals with a road forward that is more inspirational to go forward to move forward. Researchers in the domains of psychology and neuroscience have discovered that people derive the most pleasure out of life when they challenge themselves to accomplish things that are just a little bit above their capabilities but are still within their reach. This tactic is founded on the notion of the user's "zone of proximal development," which denotes the area in which they are able to make advancements if they are given some guidance. This is the region in which this tactic seeks to be implemented. Students will be more engaged in their studies and will learn more effectively, according to popular theories

of motivation and instructional design. These ideas believe that this will occur if the projects that students are given are slightly more difficult than what they are currently working on. This finding is in line with these theories, which maintain that students will learn more effectively if the assignments they are given are just a little bit more challenging than what they are tackling at the moment. This finding supports these ideas. The most important contribution that we have made to it is the implementation of multi-arm bandit algorithms. These algorithms make it possible to have a learning experience that is truly individualized by lowering the threshold of required prior knowledge in the area of study that is now being investigated. This information might be as broad as an evaluation of the pedagogical restrictions, or it can be as particular as an examination into the correlations that exist between the suggested tasks and the knowledge elements, in accordance with the approach that has been presented. According to the findings of our research, utilizing this tactic can lead to learning results that are on par with, and in some cases even superior to, those that are produced by a learning sequence that has been built by a human expert. This is because this tactic allows for the creation of learning pathways that are more efficient than those produced by traditional learning sequences. Following that, we will go over the specifics of our methodology and algorithms, and then we will finish up with a discussion of the present state of the art as well as other initiatives that are comparable. mics being contemplated for application in algorithms.

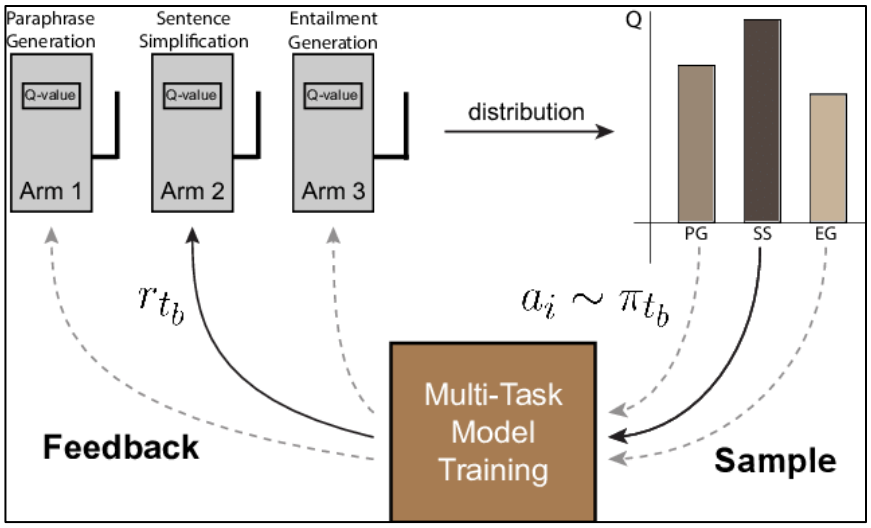

Figure 3.4: MAB for Multitask Training [19].

3.4 STATE OF THE ART: OPTIMIZATION OF LEARNING SEQUENCES WITH MACHINE LEARNING

There is a wide variety of methodology that may be utilized in order to optimize training sequences. This is something that can be done by hand, particularly if the individual who is providing the teaching has previous experience in the relevant field in addition to the essential didactic understanding for it. In this particular situation, the optimization is as automatic as it possibly can be, calling on the assumption of as few user- and domain-specific parameters as is physically possible. In other words, the optimization is as close to perfect as it possibly can be. This is a research field that is fast developing, and many different avenues of investigation are currently being proposed. In order to provide users with experiences that are most related to their anticipated skill levels, the Partially Observable Markov Decision Process (POMDP) was examined as a reference framework. This was done in order to give consumers the most relevant experiences possible. Finding a solution to a POMDP is famously difficult to do for a number of reasons. As a direct consequence of this, approximatively approaches that are predicated on the idea of "envelope states" have been developed. These methods, rather than aiming to accurately identify the individual units of comprehensive knowledge, consider simply the groups of these individual units. In the vast majority of situations, a single system will be able to support both the learner model and the instructor model at the same time. For instance, in POMDP-based techniques, an in-house learner model is applied in order to increase learning sequences. Because POMDP defines the appropriate path to pursue dependent on the cognitive and learning models of the users, these techniques have the potential to be the most effective because of this. Utilizing the Knowledge Tracing methods and the various permutations of those approaches allows for the estimation of the model's parameters to be carried out in a number of different contexts.

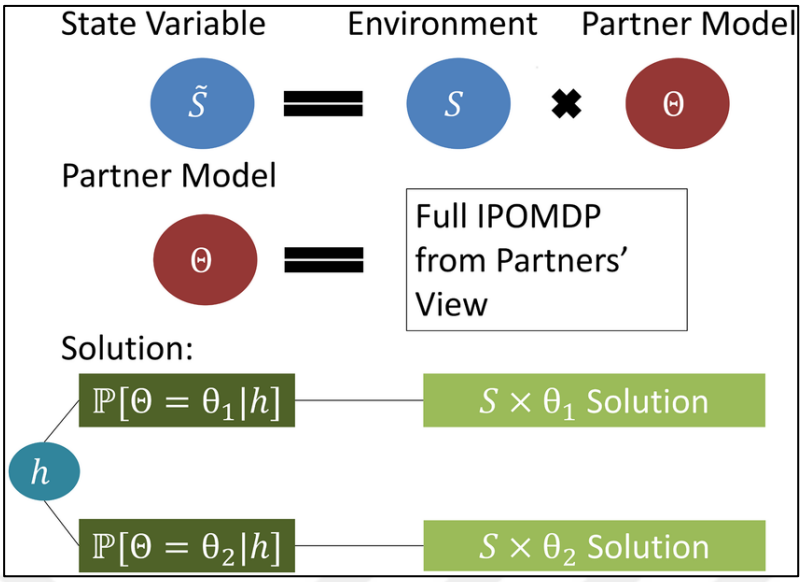

Figure 3.5: The Partially Observable Markov Decision Process (POMDP)[15].

Because there is a lack of data and because it is difficult to identify a large number of factors for a single user, it is difficult to solve this problem. These models often include a significant number of parameters, and it is difficult to solve this problem. On the other hand, finding a solution to this problem is difficult because the models in question have a substantial number of parameters. This directly contributes to the generation of inaccurate or unreliable models in practical applications, which is a common occurrence. Another difficulty is that these planning methods are not tailored to the requirements of a specific user, rather, they are broad, which significantly lessens the usefulness of the strategies. This is one of the obstacles that must be overcome. Alternately, networks Bayesians can model and decide how to best aid users, while reinforcement learning can lead users through activities and improve the acceptance of instructional approaches. Both of these options are viable alternatives. Bayesians might also decide how users can be helped in the most effective way. Other methods, such as those that are based on ant colony optimization algorithms, take into consideration a global optimization of the learning sequence based on data from all users. However, these methods are unable to provide a sequence that is tailored to the specific requirements of the individual user. Several researchers have already used the concept of Proximal Zone of Development (ZPD) to design STI learning strategies, based on the fundamental principles of pedagogy.

3.5 LEARNING SCENARIO

In this section, we will provide a more in-depth description of the experimental setting, as well as the learning scenario that was used throughout the entirety of our research. For the purpose of determining whether or not our algorithms are effective, we make use of an educational sequence that has been meticulously planned out and for which the pedagogical validity has been verified by a number of different studies. Because of this, we are able to guarantee that our algorithms are as efficient as they possibly can be. This scenario, which educates children about financial literacy, is aimed at children between the ages of 6 and 9, and the target audience for the scenario is comprised of children in that age range. The topic of money was selected because it is readily understood by people in all parts of the world, and because it comprises a varied enough variety of mathematical operations that it can be easily altered to meet the specific requirements of any individual user. In addition to the analysis of numbers and the myriad mathematical operations that may be carried out on them, the central topic also revolves around the practical application of monetary currency and the manner in which it is utilized in day-to-day life. Users interact with scenario-integrated software by means of a web browser to practice a range of tasks. These tasks include, but are not limited to, compounding quantities of money, paying for items and transferring change, and other similar activities. The student is presented with a product, along with the price that is shown next to it on the screen. After that, they will be required to choose which combination of bills and coins to create in order to pay for the item or make change for a customer, all while adhering to a variety of constraints based on the numbers and the objects that are available inside the parameters of the exercise. As the focal point of these activities, the following specialized areas will take center stage: The ability to carry out the following mathematical operations: a) add and subtract whole numbers, b) subtract whole numbers, c) decompose integers into groups of ten and units, d) add and subtract decimals, e) add and subtract decimals (in cents), f) decompose decimal numbers (hundreds) into groups of ten and units, g) decompose decimal numbers (tens) into groups of ten and units. By taking part in all of the different activities, the participants will have the opportunity to raise their level of financial literacy. Within a structure that is somewhat analogous to a tree, one can find a total of eleven parameters. The user has the choice of acting as a buyer and making payment for a purchase of either one item (type M) or two items (type MM), or acting as a seller and delivering the proper change for a purchase of either one item (type R)

or two items (type RM). The amount of work that is required to separate the sums into their component parts is precisely proportional to the degree of difficulty presented by each distinct type of workout. If there is already a coin or note in circulation with the proper value, then it is not difficult to write out an amount such as item a = 1, 2, or 5 euros in the appropriate format. On the other hand, creating a total that requires the addition of a number of separate things all at once could be more challenging, for example.

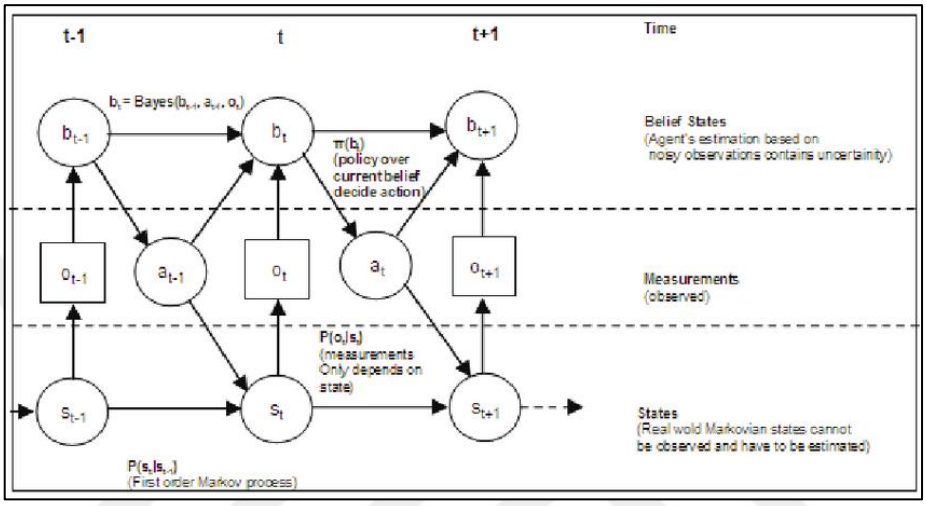

Figure 3.6: Graph Representation of the Markovian Chain [23].

The display of the price may either be done in the form of text or it can be done using a voice synthesizer, and this choice is determined by a third component. Another issue that should be taken into consideration is the so-called "Price Notation," which may be interpreted in a wide range of ways based on regional and national standards, not all of which conform to a single standard. This is something that should be taken into consideration. We take into account, as a very last step but certainly not the least important one, the medium in which the payment items are shown or symbolically represented. This medium includes things like coins, banknotes, and tokens, just to name a few examples. One or two goods, together with the pricing that are related with those things, are shown to the user when they initially begin the task. The participant will first decide which banknotes and coins he would want to use for the activity, and then he will deposit those materials in the section of the room that has been set aside specifically for that function. Simply tapping a button on the screen will allow him to submit a request for help. You will need to choose the OK button before you can send in your answer. After that, the replies will be shown in their respective boxes. If you have successfully finished the exercise, you will get the message that "Congratulations, you can

go to the next exercise" if you are able to go on to the subsequent activity. The user is given three opportunities to complete the exercise, and new suggestions are made available to them with each opportunity. This is done so that the user may develop the most efficient and effective sequence that is realistically achievable. If, after three separate tries, the solution is still found to be incorrect, the right answer will be sent to you after the issue has been fixed. The next events are going to develop in a sequence that will result in the task becoming increasingly more challenging. The prices are derived from the criteria that we have previously investigated, and as we have seen, it gets increasingly more difficult to create them using real dollars as we continue up the pricing ladder. The prices are produced based on the criteria that we have already investigated. Prices may be written in the form of numbers that can be read directly, such as those that are seen on coins or banknotes, or they can be written in the form of numbers that have to be formed from a variety of different components in order for them to be read properly. Either way, prices can be written in the form of numbers that can be read directly. The introduction of centimes occurs at a moment in the progression that significantly ups the ante in terms of complexity, on a variety of different levels. As a result of this, it is essential to have an understanding of the concept of decimal as well as the ability to read and build numbers using the decimal system. Using tokens, which are abstracted from real cash, might make it more difficult to operate with decimals. This is due to the fact that tokens are not actual currency.

3.6 INTELLIGENT SYSTEMS WITH MULTI-ARMED BANDITS

You won't be able to create a STI unless you have first specified a series of activities, designated with the letter A, that a user may participate in to gain the required skills and knowledge chunks. Once you have done this, you will be able to create a STI. Information that is pertinent to the current matter at hand or the user's progress in their own education may be taken into consideration, depending on the circumstances. Each of our many algorithms has a unique set of requirements and expectations about the breadth and depth of this information. A STI should regularly adjust the activities that it delivers for users depending on the results of prior users' involvement in order to improve users' overall understanding of the subject matter as well as their ability to do tasks relating to all aspects of the topic. Following this, we will expound on the notion that was first presented by in order to improve the efficiency of online instructional sequences by using multi-arm bandit

algorithms. This was done in order to boost the efficacy of online educational sequences. Then, we will discuss two algorithms: RL, which makes its exercise selections based on previous successes and failures regardless of the users' skill levels, and DQN, which explicitly assesses the user's skill level in order to create its exercise selections (by a method that is comparable to knowledge tracing). RL makes its exercise selections based on previous successes and failures, and DQN makes its exercise selections based on previous successes and failures. When deciding which exercises to recommend to users, none of these algorithms takes into consideration the users' existing levels of expertise.

3.6.1 Multi-Arm Bandits for Optimized Learning Paths

In order to address the problem of optimization for ITS, we make use of cutting-edge methods from the field of multi-armed bandits (MAB, which is an abbreviation for "Multi-Armed Bandits"). A multi-armed bandit is a kind of slot machine that has numerous possibilities, each of which may or may not offer a large payoff. Casinos will occasionally use analogies with multi-armed bandits to explain how difficult it is to choose one slot machine out of all of the other alternatives. These comparisons are meant to illustrate how difficult it may be to decide which computer to buy. You will need to play each of the slot machines using real money in order to assess how well they each perform. Only then will you be able to decide which one of the machines is the finest. Before we can decide which options are superior to others, we need to experiment with a wide variety of possibilities. However, we also need to choose the best of those items in a manner that the user really learns, and this amounts to what the area of machine learning refers to as an "explore/exploit" technique. In this part of the guide, we will use a similar strategy to ITS, however, this time we will consider of the user as the player.

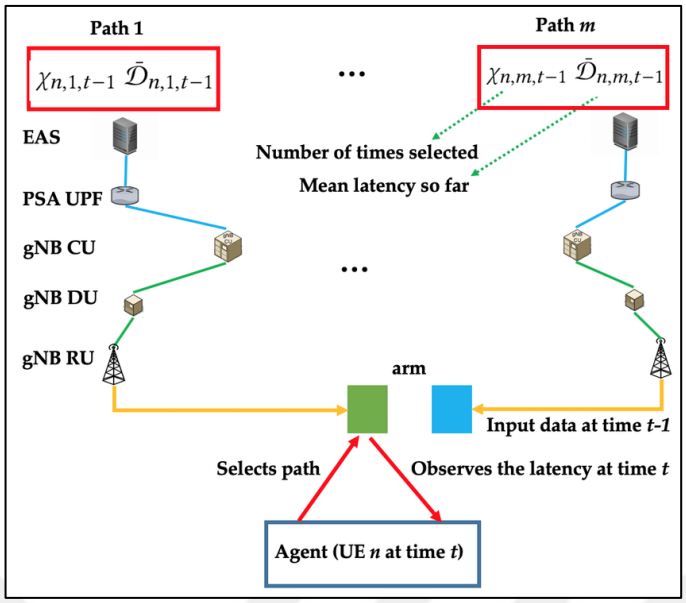

Figure 3.7: MAB for Learning Optimization [26].

We function on the presumption that users ought to put a high importance on engaging in the kinds of activities that have been proved to generate fruitful learning, where fruitful learning is defined as a training offer that is tailored to individual users and results in significant advancement. It is essential to apply specialist techniques in order to check the expansion of the incentive (in this case, one's own learning development), due to the fact that the incentive does not remain constant. Because of this, the use of a learning management system is required. When a user has reached a specific level of experience, the activity should no longer be rewarding, instead, it should be dull. This is because the user has achieved the level of knowledge necessary to make the activity uninteresting. It is not possible for us to make the assumption that the incentives are "iid," which stands for "independent and equally distributed." This is as a result of the fact that many users may have varying tastes, as well as the fact that diverse human variables, such as preoccupation or mistakes made when using the system, may have parasitic effects. As a result of this, we depend on a customized version of the EXP4 algorithm that was supplied to us. This version takes into account the suggestions provided by a number of experts and then decides what actions to take based on those recommendations rather than relying just on the advice of a single expert. Regarding us in specifically, we figure out the number wa that allows us to keep track of the most recent rewards (a correlate of learning progress) for every action a. This helps us estimate how much we have learned. Wa = wa + r, where r is a reward that reflects the advantage that the activity has given for learning, is the value that is updated

every time such an activity is utilized. This is because r is a reward that represents the benefit that the activity has provided for learning. This is due to the fact that r is a reward that represents the learning advantage that the activity has delivered to the participant. With the aid of and, it is possible to determine the tracking dynamics of the estimate. Following on from the previous sentence, we are going to present a few other alternatives to calculating this reward so that we can keep track of how far along the learning process we currently are. To generate a probability distribution that is uniform throughout, we will choose an activity to carry out at each and every given time in the way as follows: The value of pi is equal to w a(1) + u, where w and u refer to the normalized versions of the variables. The pace of exploration is represented by wa, and u stands for a uniform distribution that makes certain a significant amount of research is conducted on the activities. If you make all of your choices on the basis of the statistics that I just presented, the possibilities for what you are capable of doing are almost limitless. On the other hand, it can have the opposite effect if there are an overwhelming number of variations in the exercises that are provided, or if they often go from being too simple to being too challenging. As a consequence of all of these variables, users may become less interested and immersed in the activity. In order to prevent the level of challenge from becoming overwhelming, we are going to impose certain restrictions on how far players are allowed to go in the game. We have included the possibility of an expert producing a ZPD that is adaptive to the users' previous performance in order to make it simpler for users to use the ZPD. This was done so that users would find it easier to utilize ZPD. An increase in motivation, a lessened necessity for quantitative assessments by the instructional design expert, and a more trustworthy pool of alternative activities from which to draw are the outcomes of using the zone of proximal development. These ideas will be put into action in two distinct algorithms, and each of those algorithms will have its own one-of-a-kind collection of particulars and specialized field of study. In the next paragraph, you will be introduced to two distinct algorithms, each of which has its own one-of-a-kind technique for calculating rewards based on the basic assumptions it makes about the user's learning. These approaches are based on the fact that each algorithm makes its own set of fundamental assumptions about the user's learning. The resulting algorithm, denoted by Alg. 1, is shown in the following.

3.6.2 The RL (Reinforced Learning) Algorithm

We were able to incorporate the insights that we gained from the empirical estimate of learning progression as well as the zone of proximal growth into our methodology. Both of these concepts were provided to us by others. The zone of proximal development may include instances of both of these ideas at the same time. As we've seen, it's probable that one of the most effective ways to motivate people to learn is to reward behaviors that result in the greatest levels of performance. This is because rewarding behaviors tends to motivate individuals to continue engaging in those behaviors in the first place. As we've progressed, it's possible that this has been one of the more successful strategies we've used. This concept is given further weight by the theory known as the Zone of Proximal Development (ZPD), which states that the activities that are the most attractive for students to participate in are ones that are just slightly above their current ability level. According to ZPD's research, the most engaging and interesting exercises for students are ones that are just a little more difficult than what they are currently capable of. Building links between the activities and the varying levels of ability possessed by the users may be avoided by making an estimate of the techniques by which the success rate of each exercise can be increased. This is one approach to avoid building connections between the activities and the users. Estimating the ways by which the success rate of each exercise may be enhanced is one way to go about accomplishing this goal. As a consequence of this, we investigate the unprocessed data to determine how variables such as success rates have changed over the course of the years. This approach does not make any assumptions about the user's level of competence, rather, it utilizes the success rate of the activity to forecast whether or not it should be presented to the user. If the success rate of the activity is high enough, the technique will offer the activity to the user. To put it another way, the procedure in question does not presume anything about the user's degree of experience or ability. If there is a low likelihood of the action being successful, it will not be shown to the user as an option inside the procedure. Even though the developers of Bandit Multi-Arm do not know in advance how these activities will directly affect current skill levels, the algorithm may provide tasks of increasing complexity that are successfully suited to the learner. This is the case even if the developers do not know how these activities will directly impact current skill levels. This is the case regardless of whether or not the system correctly matches. RL is quite comparable to the forthcoming DQN, however, it is simpler to use and requires a lower level of preliminary work upstream

before it can be put into action. The three most important advantages are an increase in one's level of motivation, a reduction in the amount of time spent by the knowledgeable one who designs the ITS learning path, and an increase in the amount of material retained in one's memory.

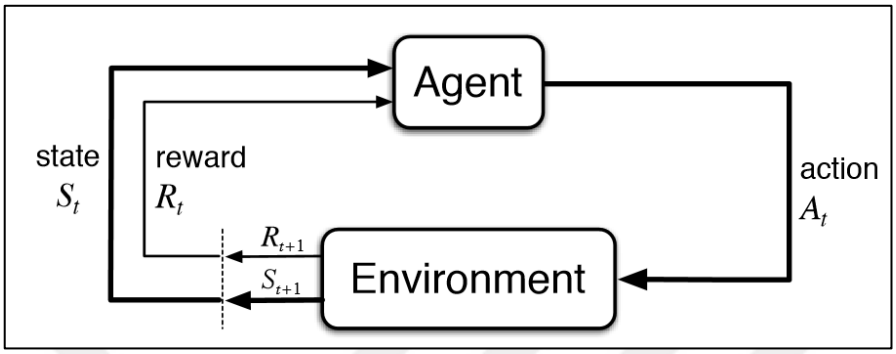

Figure 3.8: RL Model[31].

Engaging in pursuits that lead to the most significant growth in one's knowledge is, as was said before, an efficient tactic for preserving one's motivation over time. By keeping track of how far along a job has proceeded over the course of the most recent time period, this incentive contributes to a measurement of the quality of each individual piece of labor. There are two extreme scenarios in which there is no conceivable reward at all: the first is when the activity has already been gained, and the second is when it is impossible to discover a solution. Both of these scenarios are extreme examples. The notion that activities that can be finished in a shorter period of time are preferable than those that can be finished in a longer period of time is one that is held by many people today.

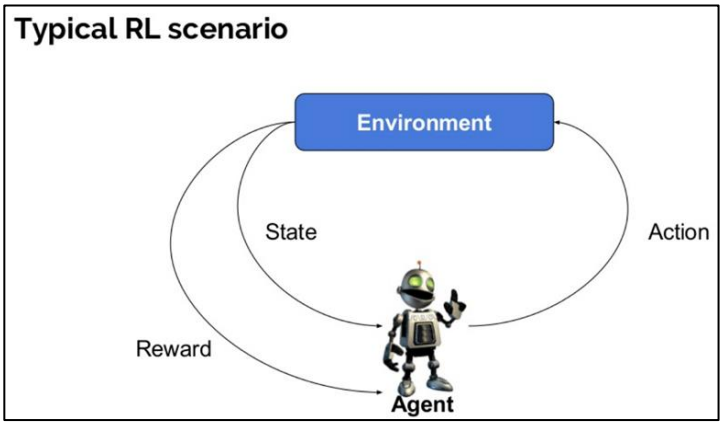

Figure 3.9: Typical RL Scenario[35].

Even with this forward-thinking strategy in place, there are just far too many potential routes to examine, and we do not have any data at the user level on which to base our questions.

Even with this forward-thinking approach in place, there are simply far too many possible paths to investigate. We provide the educational specialist who is in charge of producing the content of the STI the opportunity to describe the ZPD in the form of a graph with prerequisites for progressing from one activity to the next. When we say that we provide them with this opportunity, we are referring to it in this way. This possibility is included into the construction of the STI. In addition, we divide the activities that we provide into two distinct groups: those that do not feature a perceptible increase in difficulty, and those that do have such a gradient. Both the display of prices and the notation of pennies do not match to any order of difficulty, despite the fact that the complexity of the scenario is broken down in a specific sequence. Instead, both of these facets of the story are presented in a way that is entirely arbitrary. In point of fact, before allowing entry to the next zone in the ZPD, we only allow free exploration of activities that have the same level of difficulty as the activities in the zone that was most recently visited. This requirement is in place so that we can ensure a consistent level of challenge throughout the ZPD. When the bandit wa is less than the level of the value of the most difficult parameter, wai wai+1 with 1, we activate the value of the parameter i + I with the rule that wai = 0 and wai+I = wai+I1. When the bandit wa is more than the level of the value of the second most difficult parameter, wai wai+1 with 2, we activate the value of the parameter wai wai+1 with 2. After making the assumption that there is a predetermined order of difficulty for a certain group of activities, such as a1, a2, etc., we activate the value of the parameter known as i + I. By making adjustments to the values of the parameters, and I, one is able to either increase or decrease the amount of variance in the exercise. The most essential principle that lies behind this concept is that the ZPD needs to develop in direct proportion to the degree of difficulty of the action that is being carried out. This is the most significant idea that lies beneath this concept. When there is no evident difficulty order or when the order of difficulty varies from person to person, there has to be the opportunity for more diversified exploration of activities so that these activities may accommodate individual differences. The RL approach is easy to put into practice and needs the user to have very little in the way of previous knowledge on their end. Figure 3.3 provides a good illustration of an exploration graph and, in the opinion of the knowledgeable specialist, shows the information in a suitable manner. Using the core use of learning progress as a reward for each action, it is possible to calculate the level of competence possessed by each bandit. This may be done in a practical manner. This makes it possible to

determine the amount of expertise possessed by each bandit. Program 1 includes a description of the appropriate actions that the program should do.

3.6.3 The DQN Algorithm (Deep Q Learning)

In this piece of work, we provide an alternate method that, in contrast to what is needed by the RL algorithm, takes use of additional data on both the subject matter and the user. This data pertains to both the subject matter and the person using the method. This study is a component of a broader research effort that is being conducted under the working title "Reinventing Learning." As a direct consequence of having access to these additional pieces of information, we are now in a position to make more precise forecasts about the skill levels of our users and to offer them incentives that are more suited to meet their individual needs. The Relationship That Exists Between the Varieties of Educational Experience One Receives and the Variable Quantities of Knowledge Units (Q)The activities might be diverse in a wide variety of various ways and take on a number of different forms (for instance, they could consist of a video lecture followed by questions, an interactive game, or other exercises). Participation in a wide variety of activities will most likely lead to expansion in a number of Qs, although to various degrees, and this is a possibility. This is due to the fact that different kinds of activities serve as vehicles for acquiring different kinds of expertise and information. Despite this, there are circumstances in which the use of more than one Q is essential for the effective completion of a job. Even if there may be certain aspects of this connection that are shared by all users, the particulars of each individual's connection will be one of a kind in their own manner. On the other hand, an information technology service (ITS) may make use of this connection to draw judgments about the user's degree of expertise.

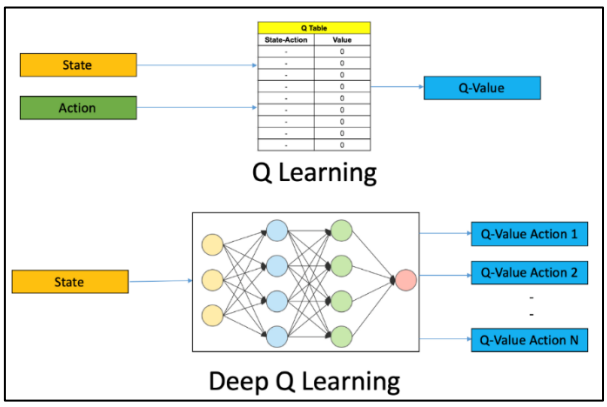

Figure 3.10: DQN vs QL [27].

We describe a user's competence in a specific Q as a continuous number between 0 and 1, somewhat similar to the recent addition of Knowledge Tracing (where 0 signifies no acquired knowledge, 0.6 denotes acquired knowledge at 60%, and 1 suggests totally acquired knowledge). This is achieved by the implementation of a technique that is analogous to the most recent improvement that was made to the Knowledge Tracing system. This level of skill may now be described with the use of the method for measuring knowledge that is known as Qi, and the most recent estimate is Qi. After this, we calculate a value that will be known as qi (a), and we apply it to any action that falls within the Qi category. This number will indicate the level of competence that is required in this KCi in order to realize one's greatest potential in activity a, and it will be shown as a percentage. This worth may be traced back to the activity itself, which acts as its foundation. An inquiry into the manner in which the activities would influence the level of expertise possessed by the users in the many knowledge areas that are the subject of the research. Our approach is built on a forecast of the manner in which each activity will affect the user's degree of competence in the different knowledge domains. This prediction serves as the method's basis. Because of this, it is very necessary to conduct a test to determine the user's current level of expertise in regard to each Qi. We do not expect to adopt regular testing outside of the activities that are designed to assess each user's Qi because we believe that doing so would likely slow down the overall growth of the user. Instead, we plan to stick to the activities that are meant to evaluate each user's Qi. The reason for our choice is that we do not desire to proceed in this manner. Therefore, in order to infer skill levels, discrete testing has to be matched with the q values that were mentioned earlier in this paragraph and utilized in order to evaluate performance in the activities. This is necessary in order to determine how well someone did in the activities. This is essential so that we can evaluate how well each individual participated in the exercises. The method known as "knowledge tracing" is one strategy that may be used to keep track of the ever-expanding toolkits of abilities that individuals possess. A reduced version that takes into consideration the relationship that we have already created between activities and KCs will be used by us going forward. Let's imagine a user believes they've reached the degree of expertise denoted by the code ci with regard to a particular set of information i, and let's say they also have the impression that they've succeeded in doing so. The user's level of achievement in completing the activity will determine the result of action a, which will be referred to as "a." If we take it for granted that this is the situation, then our

estimation of the user's capability with regard to KCi will be off base if ci is lower than qi(a). If the user is unable to complete the job and qi(a) is more than ci, this indicates that the estimated user skill level is too high, and the user should not attempt to complete the task. Depending on the specifics of the situation, the phrase "prize ri" might be interpreted to signify any of the following:

$$a = qi(a) - ci \tag{3.1}$$

There is not much to learn from the other scenarios, thus ri = 0. We factor in this bonus when calculating the user's new projected level of expertise:

$$a = ci + \alpha ri \tag{3.2}$$

Where is a parameter that, by having its value altered, gives us the ability to choose how much weight we give to each new piece of information that we take in. It is conceivable to update the ITS such that it contains specialized data in the form of global constraints. This would make the system more efficient. The knowing person is aware, for instance, that it is a waste of time to give activities on the decomposition of real numbers if the majority of users are unable to do basic mathematical operations. This is because the knowledgeable person is aware that it is a waste of time to offer activities on the decomposition of real numbers. In this scenario, the user's expected skill level may potentially be turned into explicit values, which can then be used to drive the expansion of the RL. The knowledgeable person may thus offer the very basic competence standards for that Qi in order to achieve the goal of allowing the ITS to take part in a certain activity. If the user's current skill level is higher than the minimum required for that activity, then the user will be granted permission to proceed with further exploration of that activity. In addition to this, we provide the expert the chance to choose a threshold value below which a certain action is deleted from consideration for the discovery process. This helps ensure that only relevant information is brought to light throughout the investigation.

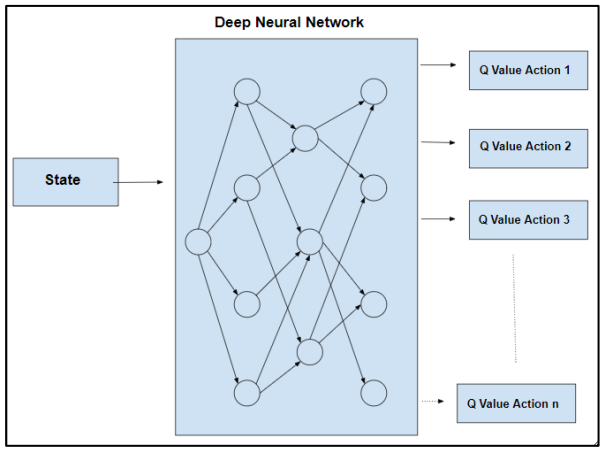

Figure 3.11: DQN State Values (Q1, Q2…Qi)[35].

In the realm of scientific inquiry, the DQN algorithm calls for the accumulation of more pieces of data. Before a new activity can be activated, the expert user must first submit a table that describes the activities' links to the KC, as well as a list of the required competence levels in order to do so. In addition, the expert user must also supply a table that outlines the required competency levels in order to activate the new activity. It is possible to assess the user's level and determine the reward for taking part in this activity by looking at the connection between the user's performance in an exercise, their estimated skill level, and the needed skill level for an exercise. This is done by studying the relationship between the user's performance in an exercise, their estimated skill level, and the required skill level for an exercise. The stages involved in the approach are the same as those involved in method 1, hence the approach and method 1 are interchangeable. Using data that is consistent across all ITS systems, this method gathers information for analysis. When there are a substantial number of KCs, manual methods will not be able to manage the complication of delivering this information to the knowledgeable person who designs the ITS. Methods that are automated will be necessary.

3.7 SIMULATIONS WITH VIRTUAL USERS

In order to study the many components of our algorithms in great depth, we carried out a number of simulations. In order to investigate how the algorithm decides which workouts are appropriate for each user as well as how the effect of user diversity plays a role, we have fabricated two distinct virtual user populations. Imagine that everyone in population Q has access to all they need to learn and grow, but that they all do it at their own pace and at their own degree of skill. This scenario would be called population Q with variable access. An

alternative population, represented by the letter "P," is meant to function as a stand-in for more diverse groups of persons, each of whom may have difficulties comprehending particular sorts of activities. This is because alternate populations are designed to be more flexible than traditional populations. Learning A user is unable to complete an activity because they are unable to hear the instructions for the activity, which prevents them from being able to successfully complete the activity. Since all of the users in population "Q" have access to all of the exercises, we hypothesize that an optimization carried out on this population will not be too sensitive. In order for the algorithm to properly instruct users in the "P" system, it must first build a separate sequence of steps for each user to follow out. This is necessary for the algorithm to achieve its goal. The following formula, derived from Item Response Theory, serves as the conceptual basis for both of these models, which may be used to the problem of determining the probability that a certain endeavor would be accomplished with success:

$$p(\text{successful}) = \frac{\gamma(\text{To})}{1 + e^{-(\beta(c - c(a) + \alpha))}}_{Q} \tag{3.3}$$

Where and allow for a variety of learning rates to be experienced by a population as well as for the form of the probability distribution to be altered. This framework is capable of addressing any and all activities due to the fact that the "Q" model's parameter (a) equals 1. There are variations of the "P" model methods, and some of them do not include 0 a(i) 1, while others do. As a consequence of this, there are some occupations designated with the letter "P" that can never be finished, regardless of how skilled the worker may be. The user's chances of achieving progress are directly proportional to the gap that exists between their current level of ability and the complexity of the task they are attempting to complete.

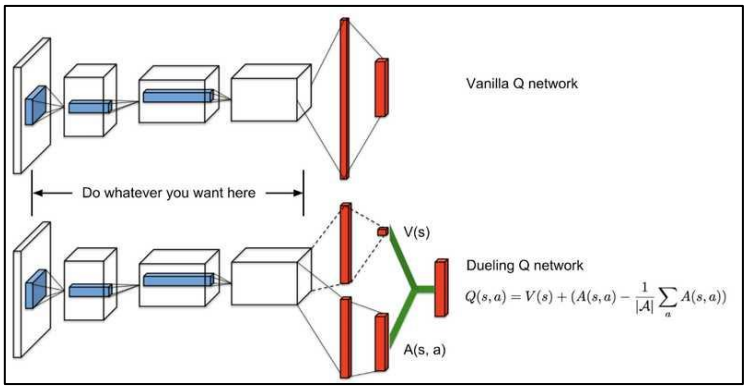

Figure 3.12: Simulation Parameters for DQN [37].

3.8 CONCLUSIONS AND PERSPECTIVES

We have shown that an effective algorithm that keeps track of a student's academic advancement and makes recommendations for suitable assignments bearing that growth in mind has the potential to provide highly favorable results. We presented two algorithms, DQN and RL, with very distinct information requirements: the first one is reliant on prior knowledge about the complexity and qualities of the jobs that are being suggested, whereas the second one does not require any knowledge whatsoever about the jobs' level of difficulty or characteristics. It was to be anticipated that DQN, which makes use of more information, would outperform RL, which, despite its lack of data, manages to do well in the majority of cases. However, it was to be expected that DQN would come out on top. On the other hand, given that DQN makes use of more information, it was reasonable to anticipate that it would outperform RL. On the other hand, considering that DQN makes use of a greater quantity of data, it was logical to assume that it would have higher performance. Our team has devised and put into practice a one-of-a-kind approach for intelligent systems that takes use of multi-armed bandits. This method was created by our team. We are the ones who came up with it, but the results it provides are on par with those of the most successful training sequences. The fact that these algorithms need just a moderate amount of specialized subject matter knowledge and that their performance may be improved in real time based on the outcomes generated by individual users makes them an especially useful resource. We introduce RL, a very simple algorithm that creates a one-of-a-kind lesson plan only based on the outcomes of the exercises and a more or less preset exploration network. This plan is then used to teach the student something new. The results of the exercises served as the basis for the development of this strategy. It is possible to tap into extra domain knowledge by making

use of an additional algorithm that goes by the name of DQN. This allows for the evaluation of the skill levels of individual students and the successful customization of instruction to meet the requirements of the students. Even if RL performs better in real-world user trials, we discover that DQN, which has more data, performs better in simulation. This is because DQN has more data to work with. This is because DQN has a greater quantity of data to work with. This is due to the fact that DQN has a more accurate representation of the data, which provides an explanation as to why this is the case. It is unlikely that this discovery will come as a shock to anyone, and it lends credence to the findings of previous study, which shown that it may not be in everyone's best interest to aim to customize the learning experience using criteria that apply to the whole population. This finding lends weight to the findings of previous research.

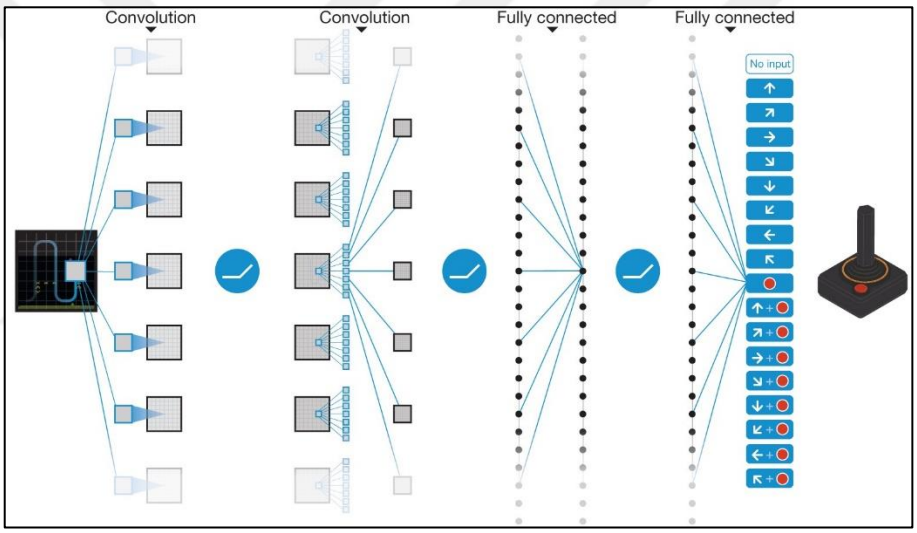

Figure 3.13: DQN for MAB [47].

Developing a system that can automatically give users with exercises that are tailored to their present level of skill is a goal of ours, but it ranks lower on our list of priorities compared to the production of learning sequences that are superior to those designed by professionals. The findings of our study, on the other hand, suggest that our algorithms could be able to attain a level of learning performance that is equivalent to that of a sequence that is supplied by an experienced user. It is possible to achieve this goal without requiring a considerable quantity of user data or a significant number of instructional components from the knowledgeable person. We have shown that we are capable of getting the same outcomes when working with user groups that are homogenous. On the other hand, the learning advantage is significantly amplified when it is applied to user populations that are

characterized by a larger variation in terms of both the breadth and depth of the challenges they confront. We make the point that a formulation of the issue on the basis of the KC is already useful, even if there is no decrease in learning time, since it allows for better identification of the obstacles that each specific user is encountering. This is why we argue that a formulation of the problem on the basis of the KC is already beneficial. User tests show a considerable improvement in learning speed across a wide range of abilities, as well as an expanded ability to customize the order in which lessons are given by using algorithms. This improvement may be attributed to the fact that there are more lessons available. RL has the most potential for real-world applications because to the little quantity of information it needs (even less than DQN) and the high degree to which it can be changed. This gives it the greatest potential for usage in the real world. We are now doing research into a number of educational settings in order to determine the optimal conditions that would enable us to get the most potential benefit from our tactics while also making their use as straightforward as is humanly feasible. Even though it is one of the strengths of our device that it requires a great deal less assumptions on the interactions between cognitive models and users, this advantage cannot be fully recognized until first an empirical evaluation of the learning gain connected with each activity has been carried out. This advantage cannot be fully recognized until after our device has been fully tested. We believe that this will be effective in settings where there are a large number of activities and interactions with systems that are relatively easy. It is possible to read the definitions that are provided in (KR KOEDINGER et al., 2013). It will operate more efficiently with the inside loop of an ITS, which is also referred to as the internal activities, as opposed to the outside loop of an ITS, which is also referred to as the transversal activities. Because each methodology operates under a unique set of assumptions, comparing the various techniques may be a challenging task. Given that we have access to a user model that is sufficiently characterized, we have the ability to anticipate that techniques that are based on POMDP will function much better for user groups that have considerable disparities. This is a requirement that has to be satisfied. In the event that this is not the case, however, our strategy will be better equipped to handle known issues and differences within the user community. However, there is always room for improvement in the assessment of precise approaches for learning models, in particular those that make use of variables like those that we warn might be detrimental when personalization is the major purpose of the endeavor. It is also necessary to do research on emerging MAB techniques.

In contrast to linear bandits, which may study more in-depth relationships between parameters, contextual bandits are able to take into consideration not just the present state of the user but also the parameters that are made accessible to them. After some consideration, we reached to the realization that each parameter needed its very own special bandit. That is the same thing as proposing that the less common characteristics not be taken into consideration. In order to have a proper understanding of the components of such a system, more research of a much more in-depth sort is essential. Exploiting the clustering that are formed as a natural result of our algorithms is another possible topic of investigation that might be pursued. We could opt to share data between two different users if we find out that they have a lot of information in common with one another. It is essential to do investigation into the newly created strategies for generating profits from the sale of multi-armed bandits. Another approach may be to choose tasks to complete based on how far we have progressed in our (recent) learning, and then to move on to something else that has a degree of difficulty that is roughly equivalent to the first activity. If these facts are accurate, then MAB could be able to provide more complex workouts without forcing users to make educated guesses about the benefits of participating in activities of this kind. This technique also generates a flow of training that is more organic since it reduces the number of times that "jumps" are performed between sessions.

4. PROPOSED METHODOLOGY

This chapter provides a detailed description of the proposed methodology for implementing a Reinforcement Learning (RL) approach using Multi-Armed Bandits (MABs) for dynamic content optimization of web performance.

4.1 PROBLEM FORMULATION

The problem at hand revolves around the optimization of dynamic content on websites to enhance their performance. In essence, the objective is to select the best content that promotes efficient web performance, while considering the dynamic and volatile nature of internet users' preferences. We propose to model this problem as a Multi-Armed Bandit problem, where each bandit arm represents a different piece of web content, and the reward is a measure of user engagement or satisfaction (such as click-through rate, session duration, bounce rate, or conversion rate). The objective is to develop a policy that can adaptively select the arm (content) that yields the highest reward (user engagement or satisfaction). The task at hand involves the optimization of dynamic content on websites to amplify their performance. More specifically, the primary objective is to choose the most suitable content that promotes the most effective web performance. This selection process must acknowledge the volatile and unpredictable nature of online user behavior and preferences, which can fluctuate based on a variety of factors such as time, location, cultural trends, market developments, and user demographics. A website's content, in this context, can include a variety of elements such as text, images, videos, animations, layout designs, and more. All these components contribute to the overall user experience on a website and thus, can heavily influence the performance metrics of a website. A piece of content that resonates with a user could lead to increased engagement, higher satisfaction, and ultimately better conversion rates. Conversely, inappropriate or less engaging content could result in decreased user satisfaction, higher bounce rates, and a decrease in return visitors. The challenge lies in the fact that user preferences are not static. What works today may not work tomorrow. This dynamism needs to be captured when optimizing web content. Traditionally, website owners or content managers might rely on intuition, experience, or even rudimentary A/B testing to decide which content to put forth. However, this is not efficient or scalable in the long run, especially considering the rapid pace at which internet users' preferences evolve. Thus, the problem can be formulated as a decision-making task: given a set of content options and a

user at a given point in time, which piece of content should be chosen to maximize user engagement or satisfaction? This question is made more complex by the fact that the reward (user engagement or satisfaction) associated with each content choice is uncertain and can vary over time.

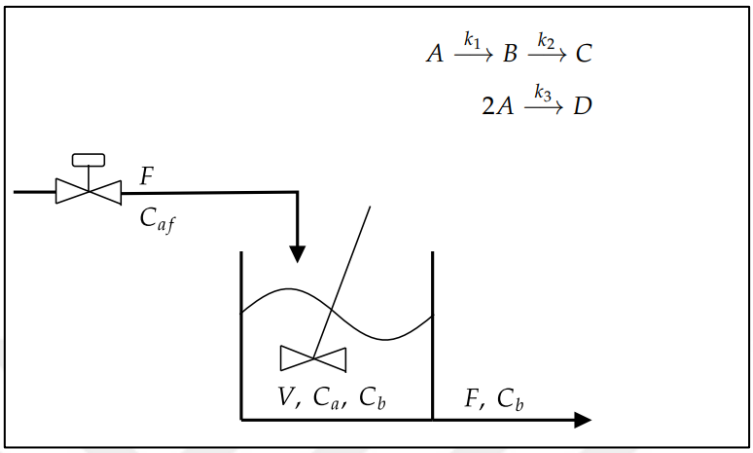

Figure 4.1: Problem Formulation [42].

We propose to model this problem using the framework of the Multi-Armed Bandit (MAB) problem, which is a classical problem in probability theory and machine learning. In this model, each piece of web content is analogous to a slot machine (or "arm") in a casino. The website owner is the gambler who must decide which machine (content) to play (display) at each time step. The reward is then a measure of user engagement or satisfaction such as click-through rate, session duration, bounce rate, or conversion rate. The goal is to develop a policy that can adaptively select the arm (content) that yields the highest expected reward (user engagement or satisfaction). This policy must balance the trade-off between exploiting arms with high known rewards and exploring arms with uncertain rewards to gain more information. The optimal policy will dynamically adapt to changes in user behavior and preferences to continually optimize web content and performance. This proposed problem formulation allows us to leverage powerful techniques from reinforcement learning and multi-armed bandits to develop a data-driven, adaptive, and scalable approach to dynamic web content optimization. The approach would be capable of continuously learning from user interactions and adjusting the content accordingly to maximize user engagement and satisfaction. Thus, it can lead to improved web performance, higher user satisfaction, and increased conversion rates, all of which are key objectives for any website owner or content manager.

Arm	Reward
1	0
2	0
3	1
4	1
5	0
3	1
3	1
2	0
1	1
4	0
2	0

Figure 4.2: Arm and Reward Functions for the RL[51].

4.2 REINFORCEMENT LEARNING WITH MULTI-ARMED BANDITS

Our proposed methodology leverages Reinforcement Learning and the concept of Multi-Armed Bandits. RL is a branch of machine learning that allows an agent to learn from its environment by interacting with it and receiving rewards or penalties. On the other hand, MABs is a problem in probability theory that portrays a gambler's dilemma of choosing between multiple slot machines (bandits), each with an unknown payout probability.

4.2.1 Reinforcement Learning Agent

We will build an RL agent that learns the optimal policy over time. The agent will interact with the environment (the website and its users), and through these interactions, it will learn to choose the most rewarding content. The agent's actions are the different pieces of content it can choose to display, and the feedback (reward) is based on user interactions with the content.

4.2.2 Exploration vs Exploitation Dilemma

A critical aspect of the MAB problem is the trade-off between exploration (trying out all content to understand their rewards) and exploitation (selecting the content with the highest known reward). To manage this, we propose to use ε-greedy policy. The agent will exploit the current knowledge and choose the best content with probability 1- ε and explore the less rewarding content with probability ε. The value of ε will decrease over time, meaning that as the agent learns more about the environment, it will exploit its knowledge more and rely less on exploration.

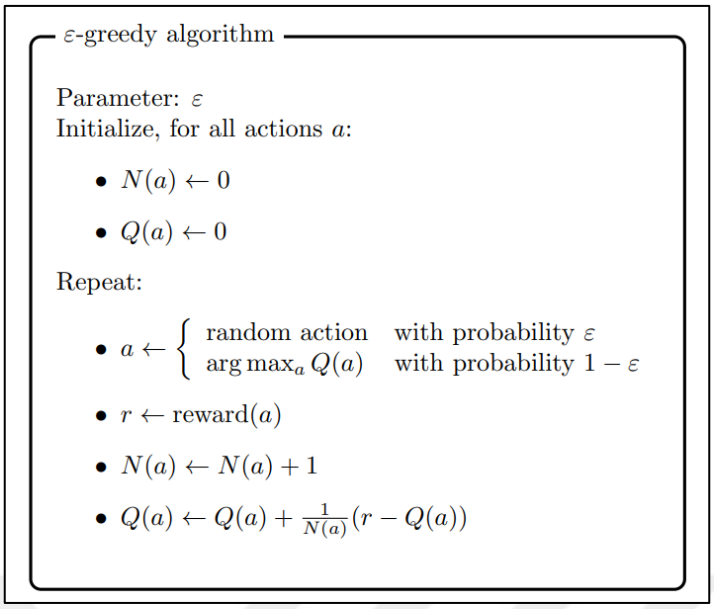

Figure 4.3: E-greedy Algorithm or RL [46].

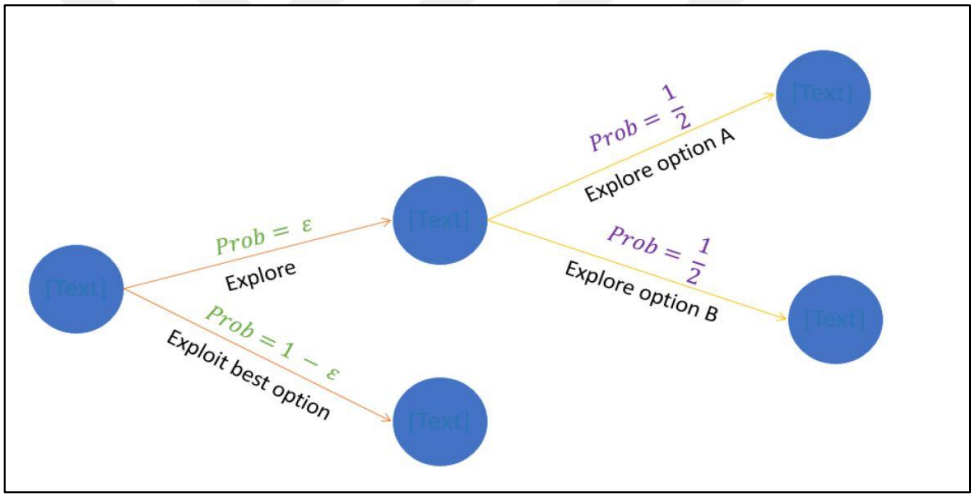

Figure 4.4: Exploration and Exploitation for RL [18].

4.3 EXPERIMENTAL SETUP

To evaluate our proposed methodology, we will need to set up a controlled experimental environment. We propose to use A/B testing for this purpose. In this setup, a fraction of users will be subjected to the standard (A) content selection method, while the rest will be subjected to our RL-based (B) method. We will then compare the performance of the two methods in terms of key performance indicators such as click-through rate, session duration, bounce rate, and conversion rate.

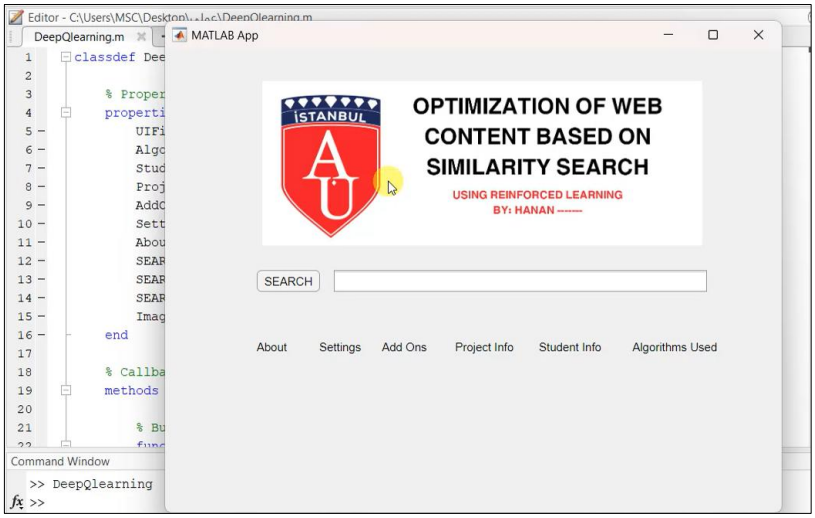

Figure 4.5: UI of the Proposed SEO.

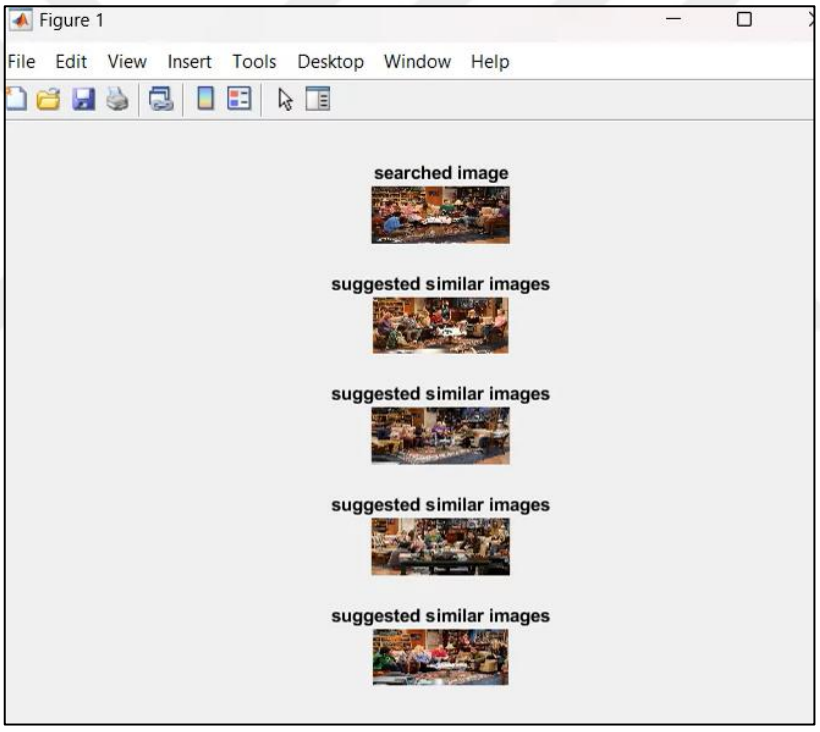

Figure 4.6: Similarity Suggestion for Optimized Search in the Proposed Model.

4.4 PERFORMANCE METRICS

To evaluate the success of our RL-based dynamic content optimization strategy, it is essential to identify and define a set of key performance indicators (KPIs). These KPIs, which reflect different aspects of user engagement and behavior, will provide insight into the efficacy of our reinforcement learning algorithm. Click-through rate (CTR): The click-through rate is a measure of how many users, upon being presented with a specific piece of

content, click on it to interact further. It can be calculated as the number of clicks that a piece of content receives divided by the number of times that content is shown, often expressed as a percentage. A higher CTR generally indicates that a piece of content is more engaging and resonates more with users. Using CTR as a KPI, we can determine whether our RL-based strategy is effectively choosing content that encourages users to interact more with the website. Session duration: Session duration refers to the length of time a user spends on the website during a single visit. This KPI can provide insights into user engagement and satisfaction.

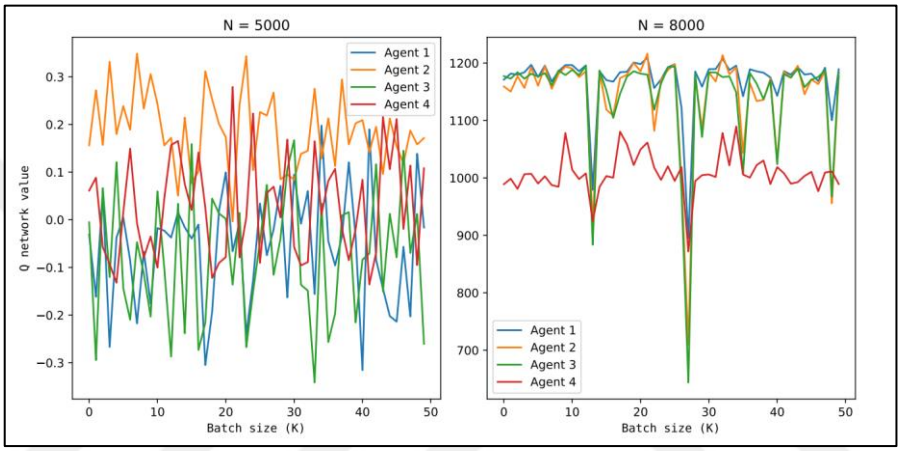

Figure 4.7: Q Values for each Armed Bandit.

Longer session durations can indicate that users are finding the content more engaging, compelling them to stay on the website for longer periods. By monitoring changes in session duration, we can assess whether the RL-based strategy is improving user engagement and satisfaction by selecting more compelling content. Bounce rate: The bounce rate is defined as the percentage of users who navigate away from the site after viewing only one page. This metric provides insights into user dissatisfaction or lack of engagement. A high bounce rate can indicate that the content on the landing page did not meet the user's expectations or needs, prompting them to exit the website without further interaction. Therefore, by using bounce rate as a KPI, we can evaluate whether our RL strategy is effectively choosing content that meets user needs and encourages further exploration of the website. Conversion rate: Conversion rate is one of the most critical performance metrics for many websites.

Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Ad 9	Ad 10
1	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0

Figure 4.8: Where (0) is the Incorrect Values of each Bandit and the (1) Represents the Correct Similarity Action for each Bandit.

It refers to the percentage of users who complete a desired action on the website, such as making a purchase, signing up for a newsletter, or filling out a form. An increase in conversion rate signifies that more users are performing the desired actions, which is typically the ultimate goal of optimizing web content. By tracking the conversion rate, we can assess whether the RL-based content selection is not only improving engagement metrics like CDF and session duration but also driving users to complete the desired actions.

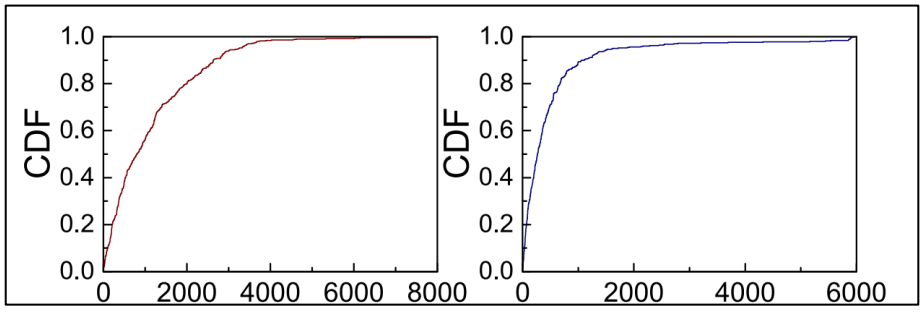

Figure 4.9: Accuracy for both the Training and Testing CDF.

These KPIs collectively provide a comprehensive picture of how well the RL-based content optimization strategy is working. Each of them addresses a different aspect of user behavior and engagement, providing multifaceted insights into the strategy's effectiveness. However, it is important to remember that these metrics do not operate in isolation. They are

interconnected, and changes in one may impact others. For instance, improvements in session duration or CDF could lead to lower bounce rates and higher conversion rates. Therefore, when evaluating the performance of our RL strategy, we need to consider all these metrics together, analyze their interplay, and take a holistic view of web performance. Moreover, the ideal values for these metrics can vary based on the type of website, the nature of the content, and the specific goals of the website owner. Therefore, it is essential to establish benchmarks or goals for these KPIs based on historical data, industry standards, or specific business objectives. The success of the RL strategy can then be measured against these benchmarks or goals. By utilizing these performance metrics in a thoughtful and integrated manner, we can effectively evaluate our RL-based dynamic content optimization strategy. This, in turn, will help us continually refine and improve the strategy to maximize web performance and user engagement.

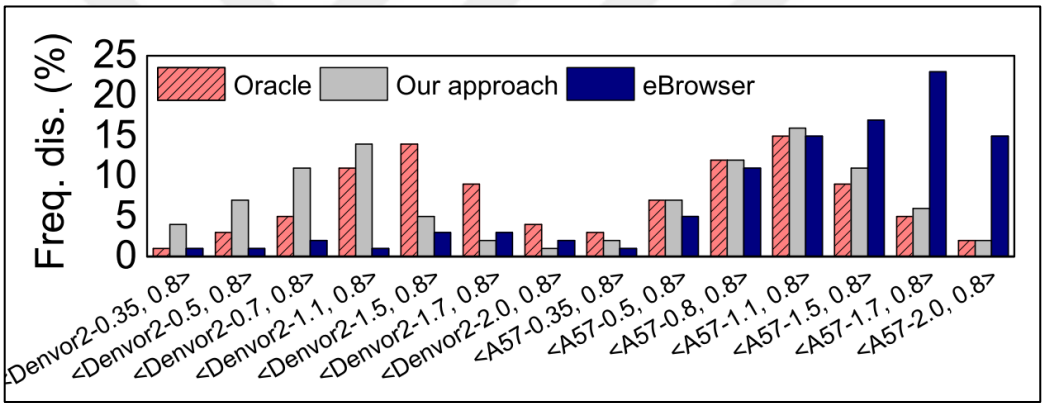

Figure 4.10: Comparing the Accuracy of Our Model to the Oracle Optimizer and Browser Optimizer Across Multiple Datasets.

5. CONCLUSIONS AND FUTURE WORK

5.1 CONCLUSIONS

This chapter presents the conclusions drawn from the research conducted in this thesis on the application of reinforcement learning with multi-armed bandits for dynamic content optimization of web performance. The primary objective was to explore the effectiveness of this approach in improving user engagement, enhancing web performance metrics, and delivering personalized experiences. The following conclusions can be derived from the findings:

a. Firstly, the integration of reinforcement learning techniques with multi-armed bandits offers a promising approach for dynamic content optimization. By continuously exploring and exploiting different content variants based on real-time user interactions, the proposed framework demonstrates the ability to adapt to changing user preferences and traffic patterns. This adaptability leads to improved web performance and enhanced user experiences.
b. Secondly, the comparative analysis with traditional approaches, such as static A/B testing or rule-based algorithms, reveals the superiority of the reinforcement learning framework. The adaptive nature of the framework enables it to outperform static approaches in terms of adaptability, scalability, and performance improvement. The findings support the notion that data-driven decision-making through reinforcement learning can significantly enhance content optimization strategies.
c. Moreover, the evaluation of performance metrics demonstrates the positive impact of the reinforcement learning framework on web performance. Increased user engagement, higher retention rates, improved conversion rates, and overall user satisfaction are observed as a result of dynamic content optimization. These findings validate the effectiveness of the proposed approach and its ability to deliver tangible benefits to organizations seeking to optimize their web performance.
d. Furthermore, the practical recommendations and guidelines derived from the research provide valuable insights for practitioners. The best practices, implementation considerations, and potential challenges outlined in this thesis can serve as a roadmap for organizations looking to implement dynamic content optimization strategies using reinforcement learning with multi-armed bandits. The guidelines aim to facilitate

successful adoption and maximize the benefits of the proposed approach in real-world scenarios.

In conclusion, the research conducted in this thesis demonstrates the effectiveness and potential of reinforcement learning with multi-armed bandits for dynamic content optimization of web performance. The findings support the notion that data-driven decision-making, adaptive strategies, and personalized experiences are crucial factors in achieving optimal web performance and user satisfaction.

5.2 FUTURE WORK

While this thesis contributes significantly to the understanding and implementation of reinforcement learning with multi-armed bandits for dynamic content optimization, there are several avenues for future research and improvement. The following areas can be explored to further advance the field.

Advanced reinforcement learning algorithms: The current research primarily focuses on basic reinforcement learning techniques with multi-armed bandits. Future work can involve investigating more advanced algorithms, such as deep reinforcement learning or contextual bandits, to improve the efficiency and effectiveness of dynamic content optimization. These advanced algorithms can leverage more complex data structures and capture more nuanced relationships between content variants and user responses.

Incorporating contextual information: The proposed framework primarily focuses on user interactions and feedback as the basis for content optimization. Future research can explore the integration of contextual information, such as user demographics, browsing history, or location, to personalize content further. By considering contextual factors, the reinforcement learning framework can deliver even more tailored and relevant experiences to users.

Hybrid approaches: Another avenue for future work is exploring hybrid approaches that combine the strengths of reinforcement learning with other optimization techniques. For instance, integrating reinforcement learning with genetic algorithms or particle swarm optimization can potentially enhance the exploration and exploitation capabilities of the framework. Such hybrid approaches can lead to more robust and efficient content optimization strategies.

Real-time adaptation: While the proposed framework operates in real-time, there is scope for further research on enhancing the speed and responsiveness of the optimization process.

Investigating techniques to enable faster decision-making and content allocation based on real-time user interactions can result in more immediate and seamless optimization.

Generalization across domains: The research in this thesis focuses on web performance optimization in a general context. However, future work can explore domain-specific optimizations, such as e-commerce, news, or social media platforms. Investigating how the reinforcement learning framework can be customized and adapted to specific domains can lead to more tailored and effective optimization strategies.

Ethical considerations: As reinforcement learning techniques become more prevalent in content optimization, it is essential to address ethical considerations. Future research can delve into the ethical implications of dynamic content optimization, such as privacy concerns, algorithmic fairness, and transparency. Developing frameworks and guidelines to ensure responsible and ethical implementation of reinforcement learning in web content optimization is a crucial area for further exploration.

Long-term impact analysis: While the research in this thesis focuses on short-term web performance metrics, future work can investigate the long-term impact of dynamic content optimization on user behavior, customer loyalty, and business outcomes. Understanding the lasting effects and potential trade-offs of continuous content optimization can provide valuable insights for organizations in strategic decision-making.

In summary, there are several opportunities for future research to further advance the application of reinforcement learning with multi-armed bandits for dynamic content optimization. By exploring advanced algorithms, incorporating contextual information, investigating hybrid approaches, focusing on real-time adaptation, considering domain-specific optimizations, addressing ethical considerations, and analyzing long-term impacts, researchers can continue to push the boundaries of web performance optimization and contribute to the development of more effective and responsible content optimization strategies.

REFERENCES

[1] S. Gupta, N. Rakesh, A. Thakral, and D. K. Chaudhary, "Search engine optimization: Success Factors," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2016. doi:10.1109/pdgc.2016.7913146.

[2] C.-J. Luh, S.-A. Yang, and T.-L. D. Huang, "Estimating google's search engine ranking function from a search engine optimization perspective," Online Information Review, vol. 40, no. 2, pp. 239–255, 2016.

[3] C. Zhu, G.Wu, " Research And Analysis Of Search Engine Optimization Factors Based On Reverse Engineering," In Proceedings of the 2011 Third International Conference on Multimedia Information Networking and Security (MINES), Shanghai, China, 4–6 November 2011, pp. 225–228.

[4] J. Zhang, A. Dimitroff, "The Impact Of Metadata Implementation On Webpage Visibility In Search Engine Results (Part II)." Inf. Process. Manag. 2005, 41, 691–715.

[5] J. Zhang, A. Dimitroff, "The Impact Of Webpage Content Characteristics On Webpage Visibility In Search Engine Results (Part I)." Inf. Process. Manag. 2005, 41, 665–690.

[6] A.S. "Hussien, Factors Affect Search Engine Optimization." Int. J. Comput. Sci. Netw. Secur. 2014, 14, 28–33

[7] R. Hernandez, (2023, May 5). "The 2022 Guide to Content Optimization." The HOTH. https://www.thehoth.com/blog/content-optimization/ In-Text Citation: (Hernandez, 2023).

[8] Ecarter, "What is an SEO specialist," SEO.com, https://www.seo.com/blog/what-is-an-seo-specialist (accessed Sep. 24, 2023).

[9] S. Zhang, N. Cabage, "Search Engine Optimization: Comparison Of Link Building And Social Sharing." J. Comput. Inf. Syst. 2017, 57, 148–159.

[10] S. Brin, P. Lawrence, "The anatomy of a large-scale hypertextual web search engine." Comput. Netw. ISDN Syst. 1998, 30, 107–117.

[11] J.M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, “A. Tomkins, The web as a Graph: Measurements, Models, and Methods.” In International Computing and Combinatorics Conference, Springer: Berlin, Germany, 1999, Volume 1627, pp. 1–17.

[12] R.W. White, M. Richardson, W.-T. Yih, “questions vs. Queries in informational search tasks.” In proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18 May 2015, pp. 135–136.

[13] Y. Li, X. Nie, R. Huang, “Web Spam Classification Method Based On Deep Belief Networks.” Expert Syst. Appl. 2018, 96, 261–270.

[14] M.A. Adebowale, K. Lwin, E. Sanchez, M.A. Hossain, “Intelligent Web-Phishing Detection And Protection Scheme Using Integrated Features Of Images, Frames And Text.” Expert Syst. Appl. 2019, 115, 300–313.

[15] P. Meel, D.K. Vishwakarma, “Fake News, Rumor, Information Pollution In Social Media And Web: A Contemporary Survey Of State-Of-The-Arts, Challenges And Opportunities.” Expert Syst. Appl. 2020, 153, 112986.

[16] R.W. Bello, F.N. Otobo, “Conversion of Website Users to Customers-The Black Hat SEO Technique.” Int. J. Adv. Res. Comput. Sci. Softw. Eng. 2018, 8, 29.

[17] S. Duari, V. Bhatnagar, “Complex Network Based Supervised Keyword Extractor.” Expert Syst. Appl. 2020, 140, 112876.

[18] S.A. Özel, “A Web Page Classification System Based On A Genetic Algorithm Using Tagged-Terms As Features.” Expert Syst Appl. 2011, 38, 3407–3415.

[19] L. Moreno, P. Martinez, “Overlapping Factors In Search Engine Optimization And Web Accessibility.” Online Inf. Rev. 2013, 37, 564–580.

[20] A.-J. Su, Y.C. Hu, A. Kuzmanovic, C.-K. Koh, “How to Improve Your Search Engine Ranking: Myths And Reality.”ACM Trans. Web 2014, 8, 1–25.

[21] S. Sagot, E. Ostrosi, A.-J. Fougères, “A Multi-Agent Approach for Building A Fuzzy Decision Support System To Assist The SEO Process.” In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016, pp. 4001–4006.

[22] D. Giomelakis, A. Veglis, “Investigating Search Engine Optimization Factors In Media Websites: The Case Of Greece.” Digit. J. 2016, 4, 379–400.

[23] M.N.A. Khan, A. Mahmood, “A Distinctive Approach To Obtain Higher Page Rank Through Search Engine Optimization.” Sādhanā 2018, 43, 43.

[24] F. Buddenbrock, “Search Engine Optimization: Getting to Google’s First Page.” In Google It, Lee, N., Ed., Springer: New York, NY, USA, 2016, pp. 195–204.

[25] G. Matošević, “Measuring The Utilization Of On-Page Search Engine Optimization In Selected Domain.” J. Inf. Organ. Sci. 2015, 39, 199–207.

[26] C. Ziakis, M. Vlachopoulou, T. Kyrkoudis, M. Karagkiozidou, “Important Factors For Improving Google Search Rank.” Future Internet 2019, 11, 32.

[27] S. Cirovic, “comparative analysis of seo factors across and within distinct industries—ecommerce, hospitality, telecommunications.” Master’s thesis, school of journalism and mass communications, faculty of economic and political sciences, Thessaloniki, Greece, January 2020.

[28] H.-J. Tsuei, W.-H. Tsai, F.-T. Pan, G.-H. Tzeng, “Improving Search Engine Optimization (SEO) By Using Hybrid Modified MCDM Models.” Artif. Intell. Rev. 2020, 53, 1–16.

[29] L.L. Carvalho, “Search Engine Ranking Factors Analysis: Moz Digital Marketing Company Survey Study.” Master’s Thesis, Universidad Nova de Lisboa, Lisbon, Portugal, 2018.

[30] V. Aul, “Harnessing Search Engine Optimization Experience to Enhance the Visibility of Websites.” Ph.D. Thesis, University of West London, London, UK, 2018. [Google Scholar]

[31] M. Hashemi, “Web page classification: A Survey of Perspectives, Gaps, And Future Directions.” Multimed. Tools Appl. 2020, 79, 11921–11945.

[32] C. Balim, K. Özkan, “Functional Classification of Web Pages with Deep Learning.” In Proceedings of the 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24–26 April 2019, pp. 1–4.

[33] J. Salminen, J. Corporan, R. Marttila, T. Salenius, B.J. Jansen, “Using Machine Learning to Predict Ranking of Webpages in the Gift Industry: Factors for Search-Engine Optimization.” In Proceedings of the 9th International Conference on Information Systems and Technologies, Caro, Egypt, 24–26 March 2019, pp. 1–8.

[34] C.-M. Chen, H.-M. Lee, Y.-J. Chang, “Two Novel Feature Selection Approaches For Web Page Classification.” Expert Syst. Appl. 2009, 36, 260–272.

[35] C. Rovira, L. Codina, F. Guerrero-Sole, C. Lopezosa, “Ranking by Relevance and Citation Counts, a Comparative Study: Google Scholar, Microsoft Academic, WoS and Scopus.” Future Internet 2019, 11, 202.

[36] A. Giannakoulopoulos, N. Konstantinou, D. Koutsompolis, M. Pergantis, I. Varlamis, “Academic Excellence, Website Quality, SEO Performance: Is there a Correlation,” Future Internet 2019, 11, 242.

[37] I.H. Witten, E. Frank, A.M. Hall, C.J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann: Burlington, MA, USA, 2016.

[38] X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J McLachlan., A. Ng, B. Liu, P.S. Yu, et al. “Top 10 algorithms in data mining.” Knowl. Inf. Syst. 2008, 14, 1–37.

[39] J.R. Saura, “Using Data Sciences in Digital Marketing: Framework, Methods, And Performance Metrics.” J. Innov. Knowl. 2020, in press.

[40] D.T. Larose, C.D. Larose, Data Mining and Predictive Analytics. John Wiley & Sons, Inc, NJ, USA, 2015.

[41] T. Cover, P. Hart, “Nearest Neighbor Pattern Classification.” IEEE Trans. Inf. Theory 1967, 13, 21–27.

[42] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning, Springer: New York, NY, USA, 2013.

[43] L. Gaudette, N. Japkowicz, “Evaluation Methods For Ordinal Classification.” In proceedings of the 22nd Canadian Conference on Artificial Intelligence, Canadian AI 2009, Kelowna, BC, Canada, 25–27 May 2009, pp. 207–211.

[44] J.S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, “Algorithms For Hyper-Parameter Optimization.” Adv. Neural Inf. Process. Syst. 2011, 24, 2546–2554.

[45] R.G. Mantovani, T. Horváth, R. Cerri, J. Vanschoren, A. Carvalho, “Hyper-Parameter Tuning Of A Decision Tree Induction Algorithm.” In Proceedings of the 5th Brazilian Conference on Intelligent Systems (BRACIS), Recife, Brazil, 9–12 October 2016, pp. 37–42.

[46] A. Ben-Hur, J. Weston, J. “A User’s Guide To Support Vector Machines.” In Data Mining Techniques for the Life Sciences, Carugo, O., Eisenhaber, F., Eds., Humana Press: Totowa, NJ, USA, 2010, Volume 609, pp. 223–239.

[47] S. Aliakbary, H. Abolhassani, H. Rahmani, B. Nobakht, “Web Page Classification Using Social Tags.” In Proceedings of the International Conference on Computational Science and Engineering, Vancouver, BC, Canada, 29–31 August 2009, pp. 588–593.

[48] J.-H. Lee, W.-C. Yeh, M.C. Chuang, “Web Page Classification Based On A Simplified Swarm Optimization.” Appl. Math. Comput. 2015, 270, 13–24.

[49] S.T. Marath, M. Shepherd, E. Milios, J. Duffy, Large-Scale “Web Page Classification.” in proceedings of the 47th Hawaii International Conference on System Sciences (HICSS), Waikoloa, HI, USA, 6–9 January 2014, pp. 1813–1822.

[50] A.L. Berger, V.O. Mittal, OCELOT: “A System for Summarizing Web Pages.” In Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Athens, Greece, 24–28 July 2000, pp. 144–151.

[51] J.-T. Sun, D. Shen, H.-J. Zeng, Q. Yang, Y. Lu, Z. Chen, “Web-Page Summarization Using Click Through Data.” In Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil, 15–19 August 2005, pp. 194–201.

[52] Y. Zhang, N. Zincir-Heywood, E. Milios, “Term-Based Clustering And Summarization Of Web Page Collections.” In Advances in Artificial Intelligence, Tawfik, A.Y., Goodwin, S.D., Eds., Springer: Berlin/Heidelberg, Germany, 2004, Volume 3060, pp. 60–74.

[53] L. Mostafa, “Webpage Keyword Extraction Using Term Frequency.” Int. J. Comput. Theory Eng. 2013, 5, 174.

[54] J.R. Landi, G.G. Koch, “The Measurement Of Observer Agreement For Categorical Data.” Biometrics 1977, 33, 159–174. [Google Scholar] [CrossRef][Green Version]

[55] M.G. Kendall, A New Measure of Rank Correlation.” Biometrika 1938, 30, 81–93.

[56] K.D. Abdullah, “Search Engine Optimization Techniques By Google’s Top Ranking Factors: Website Ranking Signals, 2017.” Available online: https://dl.acm.org/doi/book/10.5555/3169359

[57] V. Andersson, D. Lindgren, “Ranking Factors to Increase Your Position on the Search Engine Result Page: Theoretical and Practical Examples.” Faculty of Computing, Blekinge Institute of Technology: Karlskrona, Sweden, 2017.

[58] T. Mavridis, A.L. Symeonidis, “Identifying Valid Search Engine Ranking Factors In A Web 2.0 And Web 3.0 Context For Building Efficient SEO Mechanisms.” Eng. Appl. Artif. Intell. 2015, 41, 75–91.

[59] J. Sujata, S. Noopur, N. Neethi, P. Jubin, P. Udit, “On-Page Search Engine Optimization: Study of Factors Affecting Online Purchase Decisions of Consumers.” Indian J. Sci. Technol. 2016, 9, 1–10.

[60] A. Jović, K. Brkić, N. Bogunović, “A Review Of Feature Selection Methods With Applications. In Proceedings of the 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO).” Opatija, Croatia, 25–29 May 2015, pp. 1200–1205.

[61] R. Mcgill, J. W. Tukey, and W. A. Larsen, “Variations of box plots,” The American Statistician, vol. 32, no. 1, pp. 12–16, 1978. doi:10.1080/00031305.1978.10479236.N.

[62] Japkowicz, M. Shah, “Statistical Significance Testing.” In Evaluating Learning Algorithms: A Classification Perspective, Japkowicz, N., Shah, M., Eds., Cambridge University Press: Washington, DC, USA, 2011, pp. 206–291. [Google Scholar]

[63] T.G. Dietterich, “Approximate Statistical Tests For Comparing Supervised Classification Learning Algorithms” Neural Comput. 1998, 10, 1895–1923.

[64] F. A Milton, “Comparison Of Alternative Tests Of Significance For The Problem Of M Rankings.” Ann. Math. Stat. 1940, 11, 86–92. [Google Scholar] [CrossRef]

[65] P. Nemenyi, “distribution-free multiple comparisons.” Biometrics 1962, 18, 263.