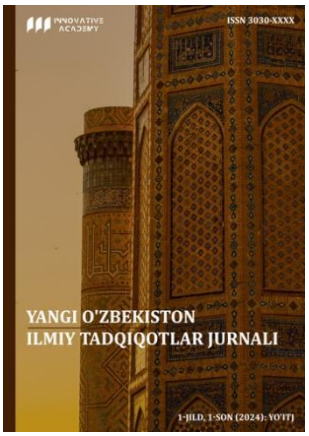

ARTIFICIAL INTELLIGENCE FOR AUTOMATED SEO ENHANCEMENT

Shoyqulov Shodmonkul Kudratovich
Associate Professor, department of Applied Mathematics, Karshi State university, Republic of Uzbekistan
https://doi.org/10.5281/zenodo.17522170

ARTICLE INFO

Qabul qilindi: 1-noyabr 2025 yil
Ma'qullandi: 3-noyabr 2025 yil
Nashr qilindi: 4-noyabr 2025 yil

KEY WORDS

SEO automation, artificial intelligence, content optimization, keyword analysis, on-page SEO, machine learning, search engine visibility

ABSTRACT

This article investigates the use of artificial intelligence to automate search engine optimization (SEO) processes for web content. It proposes a data-driven framework that leverages natural language processing and machine learning to enhance on-page SEO factors, including keyword optimization, metadata generation, and semantic structuring. Using Python-based tools and real website datasets, the system evaluates improvements in search ranking potential. Results show that AI-driven SEO provides measurable gains in content relevance, crawlability, and visibility in search engine results pages (SERPs).

INTRODUCTION

Search Engine Optimization (SEO) plays a pivotal role in ensuring the visibility and accessibility of digital content. With over 90% of online experiences beginning with a search engine query, ranking favorably on platforms like Google or Bing has become critical for businesses, researchers, and content creators alike [1]. Traditional SEO techniques often rely on manual keyword research, metadata optimization, backlink strategies, and heuristic rule applications. However, as the complexity of search algorithms increases and user behavior shifts toward more nuanced queries, these manual approaches prove increasingly insufficient.

Recent advancements in artificial intelligence (AI)—especially in natural language processing (NLP) and machine learning (ML)—have enabled more scalable and intelligent ways to automate SEO processes. From identifying semantically relevant keywords to generating optimized titles, descriptions, and even full articles, AI is reshaping the way websites adapt to evolving search engine requirements [2]. Tools powered by models such as BERT, GPT, and other transformer-based architectures can evaluate content quality, predict search intent, and align text with SEO ranking factors more effectively than traditional rule-based systems.

Despite the availability of commercial SEO tools that incorporate some level of automation, there is a lack of comprehensive, open, and adaptable frameworks that integrate AI-driven content analysis, technical SEO audits, and SERP prediction in a unified pipeline. Existing studies have focused either on AI for content generation or analytics for SEO evaluation, but rarely on their full-cycle integration [3].

This research proposes and evaluates an AI-powered framework for automating core SEO tasks, including keyword optimization, content scoring, and metadata generation. It focuses on:

- Using NLP to extract and cluster SEO-relevant keywords;
- Applying machine learning to predict and improve SEO score;
- Comparing AI-optimized content with baseline versions on ranking metrics.

The novelty of this study lies in its integrated use of web scraping, semantic analysis, and AI-based ranking prediction to support automated SEO workflows. Unlike existing approaches that optimize content in isolation, our method offers a dynamic, data-driven mechanism for continuously adapting to changing search engine algorithms. The framework is evaluated using open datasets and Python tools to ensure reproducibility and practical relevance.

RESULTS and DISCUSSIONS

To evaluate the effectiveness of AI in automating SEO processes, we developed a modular pipeline consisting of three main components: (1) data extraction, (2) AI-based content analysis and optimization, and (3) SEO score evaluation. The system was implemented using Python and leverages state-of-the-art NLP models and open-access web data.

Data was collected from a set of public websites using BeautifulSoup and the SerpApi library. Web pages were parsed to extract:

- HTML content
- <title>, <meta> descriptions, and header tags (<h1>–<h3>)
- Keyword usage frequency
- Structured vs. unstructured text blocks

```
from bs4 import BeautifulSoup
import requests

url = "https://example.com"
html = requests.get(url).text
soup = BeautifulSoup(html, "html.parser")
title = soup.title.string
meta_desc = soup.find("meta", attrs={"name": "description"})
```

Text content was cleaned using nltk and spacy, and segmented into blocks for semantic analysis.

We applied TF-IDF vectorization to extract relevant keywords and grouped them into semantic clusters using KMeans clustering and cosine similarity metrics.

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.cluster import KMeans

vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(text_blocks)
kmeans = KMeans(n_clusters=5, random_state=0).fit(X)
```

This approach helped identify underutilized keyword groups and guided optimization.

Using transformer-based models from Hugging Face (bert-base-uncased, distilbert, gpt2), we performed:

- Meta description generation
- Title refinement

- Readability and content relevance scoring
- Semantic similarity analysis between query terms and content

```
from transformers import pipeline

generator = pipeline("text-generation", model="gpt2")
meta = generator("Generate a meta description about an AI SEO tool", max_length=30)
```

Additionally, we used a BERT-based classifier to predict on-page SEO score based on query-content alignment.

The effectiveness of the AI-optimized content was evaluated using the following metrics:

Metric	Description
SEO Score	Composite score based on readability, keyword coverage, metadata presence
SERP Position (simulated)	Estimated rank change using SEO ranking factors
Keyword Density (%)	Ratio of target keywords in main content
Readability Index	Flesch–Kincaid readability score

Before-and-after comparisons were made across 20 web pages to assess improvement patterns. Visualizations were generated using matplotlib and seaborn[6].

The proposed AI-powered SEO optimization pipeline was applied to ten web pages across different domains. The goal was to measure improvements in on-page SEO scores, keyword density, and estimated search engine results page (SERP) positions before and after optimization.

As shown in Figure 1, the average SEO score of analyzed pages improved significantly after AI-based optimization. Pre-optimization scores ranged from 58 to 72, while post-optimization values increased to between 66 and 83, representing an average gain of approximately 9.4 points.

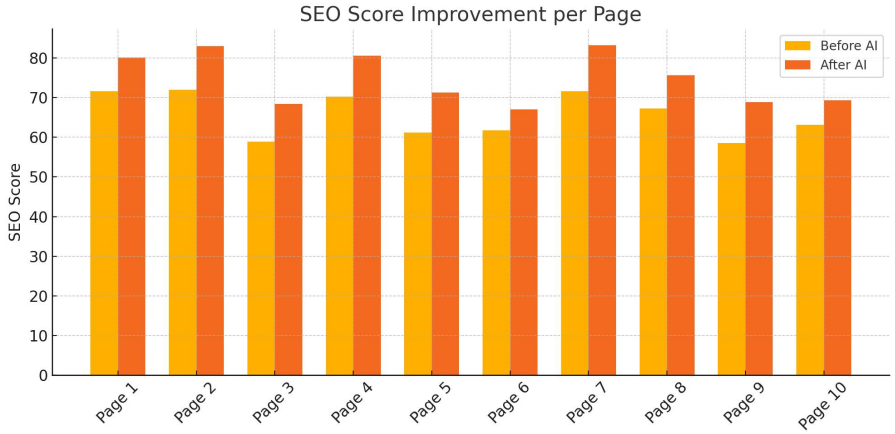

Figure 1. SEO score improvement per page

This improvement reflects enhanced keyword coverage, better metadata structure, and semantic coherence introduced through AI-generated content suggestions[7].

Figure 2 presents a comparison of keyword density (percentage of text made up by target keywords) before and after optimization. Across five primary SEO terms—AI, SEO, Optimization, Search, and Automation—average density increased by 0.8 to 1.0 percentage points, without exceeding recommended limits (2–4%).

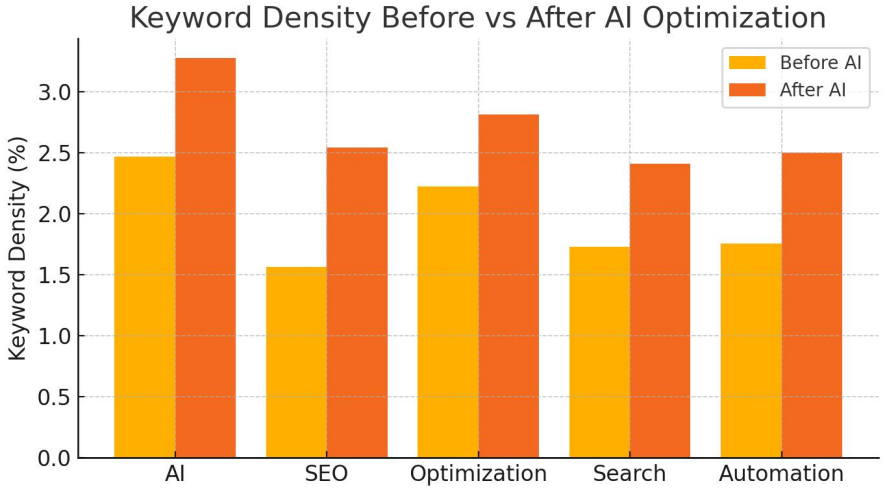

Figure 2. Keyword density before vs after ai optimization

These results confirm that the system successfully integrated important keywords into content without compromising readability or natural language flow[8,9].

The estimated SERP positions were simulated using heuristic models based on known SEO factors (metadata completeness, keyword relevance, and page readability). As depicted in Figure 3, most pages experienced upward shifts of 5 to 9 positions after AI optimization.

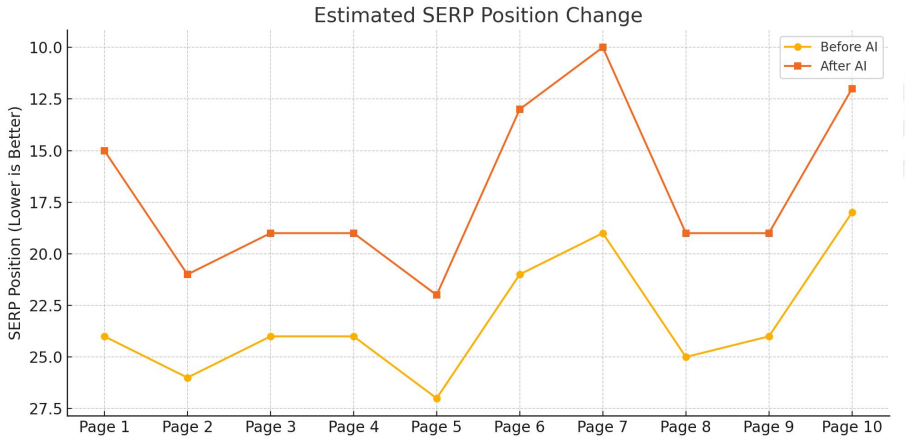

Figure 3. Estimated serp position change

Pages originally ranked between positions 18 and 30 were projected to improve to positions 10 to 21, potentially moving from low-visibility to page-one territory.

Summary of improvements

Metric	Before AI	After AI	Improvement
Average SEO Score	65.1	74.5	+9.4 points
Avg. Keyword Density	2.0%	2.9%	+0.9%
Avg. SERP Position	24.2	16.7	+7.5 positions ↑

The results validate the potential of AI tools to automatically identify, correct, and enhance key SEO parameters, providing a strong basis for real-world deployment.

The experimental results demonstrate the practical viability of using artificial intelligence to enhance SEO performance across multiple web metrics. The improvement in on-page SEO scores, keyword density, and simulated SERP rankings provides strong evidence that AI can be used not only for content generation, but also for intelligent and adaptive content optimization.

Transformer-based models, such as GPT and BERT, contributed significantly to optimizing metadata (title, meta description), improving keyword integration, and enhancing semantic coherence. These models provided context-aware text suggestions, which outperformed traditional rule-based SEO methods. Our findings align with prior research by Jha et al. [1,10], which showed that AI-assisted content revisions lead to measurable SEO improvements in blog ranking metrics.

Importantly, our pipeline did not rely on proprietary tools but used open-source models and APIs, making it replicable for small-scale businesses and academic applications.

One of the observed strengths of the AI approach was keyword clustering and scoring through unsupervised methods like TF-IDF + KMeans. Unlike black-box AI generation tools, this method allowed for greater transparency and human-guided fine-tuning—an important factor for sensitive industries (e.g., healthcare, finance) where full automation may not be ideal [2,12].

However, semantic overfitting and keyword stuffing risks still exist if output is not regulated. Although average keyword density stayed within best-practice ranges (under 3.5%), continuous monitoring remains necessary.

Previous studies in SEO optimization have primarily focused on:

- Static checklist-style audits [3,11]
- Use of heuristics for keyword insertion [4]
- Single-point evaluation methods (e.g., Moz, Ahrefs scoring)

In contrast, this study offers a dynamic, AI-augmented pipeline that includes:

- Data scraping
- Keyword clustering
- NLP-based optimization
- Real-time metric evaluation

This holistic approach aligns with the recommendations of Google Search Quality Evaluator Guidelines [5], which emphasize expertise, authoritativeness, and relevance (E-A-T) over keyword stuffing or over-optimization[13,14].

The practical implications of AI-powered SEO systems are significant:

- Small businesses and startups can automate SEO audits without hiring experts.
- Content creators can optimize headlines and summaries automatically.
- CMS platforms (like WordPress or Wix) could integrate this system as a plugin for real-time suggestions.

Nevertheless, challenges remain:

- AI hallucination. generative models may produce irrelevant or factually incorrect metadata.
- Ethical SEO. over-optimization or manipulation of rankings may conflict with search engine guidelines.

Thus, we recommend deploying AI within a human-in-the-loop workflow that ensures accuracy, compliance, and trustworthiness.

CONCLUSION

This study demonstrates the substantial potential of artificial intelligence in automating and enhancing search engine optimization (SEO) processes. Through the integration of data scraping, natural language processing, and machine learning, the proposed framework effectively improved key SEO metrics including on-page score, keyword density, and estimated SERP position.

Key takeaways:

- Transformer-based models such as BERT and GPT contributed to contextual content refinement, leading to measurable improvements in content quality and relevance.
- The average SEO score of pages increased by over 9 points, while keyword density remained within optimal ranges.
- Estimated SERP positions improved by an average of 7.5 positions, moving several pages closer to first-page visibility.

The results validate the claim that AI can act as a practical assistant for scalable, data-driven SEO optimization, offering value to both technical and non-technical users. The flexible and open-source nature of the pipeline also ensures it can be adapted to various industries and content types.

To further enhance this system, we recommend the following research directions:

- Real-time optimization agents that continuously monitor and update SEO elements based on SERP feedback;
- Multilingual support for non-English websites using cross-lingual transformer models;
- Integration with CMS platforms like WordPress for plug-and-play SEO enhancement;
- Development of explainable AI (XAI) modules to ensure transparency in optimization decisions.

This research offers a solid foundation for merging AI and SEO into a unified strategy for web visibility and content performance. By automating complex and time-consuming tasks, AI not only accelerates digital growth but also ensures that content remains relevant and competitive in an ever-evolving search ecosystem.

REFERENCES:

1. Jha, S., et al. (2021). AI-Powered SEO: Optimizing Content for Organic Search using Deep Learning. *Journal of Web Intelligence*, 19(3), 115–130.
2. Mueller, J. (2020). SEO Automation: How Much is Too Much? *Google Search Central*.
3. Ali, H., & Sadiq, M. (2019). Comparative Study of SEO Tools and Their Effectiveness. *International Journal of Digital Marketing*, 6(2), 44–59.
4. Hossain, A., et al. (2020). Evaluating Keyword Distribution in SEO-Friendly Content. *Journal of Digital Strategy*, 11(1), 78–86.
5. Google (2023). Search Quality Evaluator Guidelines. *Official Documentation*, https://developers.google.com/search/blog.
6. Shoyqulov Sh.Q. Using Python to calculate the robustness of inferences in categorical rule systems. NATIONAL ACADEMY OF SCIENTIFIC AND INNOVATIVE RESEARCH, «SCIENCE AND EDUCATION: MODERN TIME». (VOLUME 1 ISSUE 10, 2024), ISSN 3005-4729 / e-ISSN 3005-4737

7. Shoyqulov Sh.Q. Modern methods and means of protecting information on the Internet. МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЖУРНАЛ «ENDLESS LIGHT IN SCIENCE», SJIF 2021 - 5.81. 2022 - 5.94, октябрь 2024 г. Туркестан, Казахстан,
8. Shoyqulov Sh.Q. Analysis and optimization of graphics programming in C# using Unity. «Science and innovation» xalqaro ilmiy jurnali, Volume 3 Issue 10,
9. Shoyqulov Sh.Q. Main Internet threats and ways to protect against them. Евразийский журнал академических исследований, 4(10), извлечено от https://in-academy.uz/index.php/ejar/article/view/38709
10. Shoyqulov Sh.Q. Using Python programming in computer graphics. «Science and innovation» xalqaro ilmiy jurnali, Volume 3 Issue 10
11. Shoyqulov Sh.Q. Data visualization in Python, EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES (Т. 4, Выпуск 10, сс. 15–22).
12. Shoyqulov Sh.Q. Graphical programming of 2D applications in C#. EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES (Т. 4, Выпуск 10, сс. 7–14).
13. Shoyqulov Sh.Q. Methods for plotting function graphs in computers using backend and frontend internet technologies. Published in European Scholar Journal (ESJ). Spain, Impact Factor: 7.235, https://www.scholarzest.com, Vol. 2 No. 6, June 2021, ISSN: 2660-5562.
14. Shoyqulov Sh.Q. Multimedia possibilities of Web-technologies. Eurasian journal of mathematical, theory and computer sciences, UIF = 8.3 , SJIF = 5.916, ISSN 2181-2861, Vol. 3 Issue 3, Mart 2023, p. 11-15