Case Study Deep Learning in Personalized Web Searching_12010285_12010424_12011001.pdf

Deep Learning in Personalized Web Searching

Prof. Rashmi Astangi
Dept. of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
rashmi.astangi@vit.edu

Rohit Jadhav
Dept. of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
rohit.jadhav20@vit.edu

Rahul More
Dept. of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
rahul.more20@vit.edu

Aditya Pachore
Dept. of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
aditya.pachore20@vit.edu

Mahesh Gastgar
Dept. of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
mahesh.gastgar20@vit.edu

***Abstract* - Personalized web searching has become increasingly important in enhancing user satisfaction and engagement. Deep learning techniques have shown promising results in various aspects of personalized information retrieval on the web. This literature review explores the intersection of deep learning and personalized web searching, summarizing key methodologies, applications, challenges, and future directions. We review the fundamentals of personalized web searching, discuss deep learning techniques for personalized search, evaluate metrics and methodologies for system evaluation, present case studies and applications, and highlight future research directions and challenges. The review aims to provide insights into the current state-of-the-art, emerging trends, and potential avenues for future research in this domain.**

***Keywords* - personalized web searching, deep learning, recommender systems, content understanding.**

I. Introduction

In today's digital era, navigating through the vast sea of online information presents a formidable challenge. With an abundance of content flooding the internet, locating relevant and meaningful material can feel like searching for a needle in a haystack. However, the emergence of intelligent personalized web searching techniques offers a promising solution to this problem. Many individuals find themselves grappling with the overwhelming amount of information available online. However, the advent of intelligent personalized web searching techniques presents a potential remedy to this issue. Imagine a web search experience that transcends the conventional method of matching keywords. Picture a system that comprehends users' preferences, interests, and past interactions with online content. This is where neural networks, a cutting-edge technology, come into play, revolutionizing the way people explore the vast expanse of the World Wide Web.

In this case study, the exploration of personalized web searching techniques driven by neural networks will be undertaken. The inner workings of these intelligent systems will be uncovered, delving into the intricate mechanisms that power them. Furthermore, the study will explore the benefits these techniques offer to users and examine the innovative applications that are reshaping the landscape of web search.

II. Literature Review

In recent years, there has been significant progress in leveraging deep learning techniques to enhance personalized web searching. Mitra et al. (2015) introduced the Personalized Ranking Network (PRN), a neural network model that revolutionized the ranking of search results by incorporating user-specific features and query-document features. By learning from user interactions with search results, the PRN model continuously improves ranking accuracy over time, leading to more relevant and personalized search outcomes [5]. Building on this foundation, Wu et al. (2017) conducted a comprehensive survey on deep learning techniques applied to personalized search and recommender systems. Their study covered a wide range of models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep autoencoders, offering insights into how these models address challenges such as data sparsity and the cold-start problem [6]. Goyal et al. (2018) further advanced personalized web searching with the introduction of the Personalized Web Search Model (PWSM). This deep learning architecture integrates user-specific features, search history, and context information to generate highly personalized search results. By training on extensive datasets of user interactions, the PWSM model significantly improves relevance and user satisfaction, setting a new standard for personalized web searching [7]. Additionally, Jin et al. (2019) explored the application of deep reinforcement learning (DRL) to personalized web search. Their framework utilizes a policy network to make ranking adjustments based on user queries and context, while a reward signal guides the updating of the policy through gradient descent. This innovative approach enables direct interaction with users, leading to more adaptive and effective search result rankings. Collectively, these studies demonstrate the transformative potential of deep learning in revolutionizing the personalization and effectiveness of web search algorithms, paving the way for more intelligent and user-centric search experiences [8].

III. Methodology & Case Study

Neural Network working in Web Search:

Artificial neural networks, inspired by the brain's structure, revolutionize personalized web searching. Unlike traditional methods, they adapt to user preferences, improving result relevance. By analyzing data and interactions, they tailor search results, offering a dynamic and personalized browsing experience. With neural networks, web searching evolves beyond keyword matching, promising precise and engaging results for users.

Types of neural networks architectures in searching:

1. Graph Convolutional Networks (GCN):

Following are the steps to structure the input data for a GNN-based recommendation system:

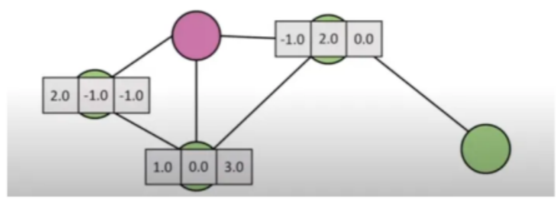
Nodes:

- Blogs: Each blog is a node in the graph. You can represent each blog as a node and include information such as the blog's content, metadata (e.g., title, category, publication date), and any user interactions (e.g., likes, comments).
- Users: If you want to incorporate user behavior into the recommendation system, users can also be nodes. Users may have features such as demographics, past interactions with blogs, and preferences.

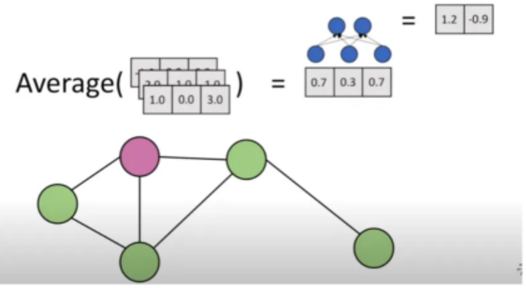
Edges:

- Blog-Blog Edges: Create edges between blogs based on some relationship. This relationship can be determined by content similarity, co-occurrence in user sessions, or shared tags/categories. For example, if two blogs have similar content or are often viewed together, you can create edges between them.
- User-Blog Edges: If you're including user behavior, create edges connecting users to blogs they've interacted with. These interactions might include clicking, liking, sharing, or commenting.

Working of GCN:



Aggregation



Transform

- Aggregate: Calculate the weighted sum of the features of neighboring nodes.
- Transform: Pass feature vector through an MLP to compute the new features for the current node.
- Combine: Concatenate the transformed features.
- Apply Activation, like ReLU [7].

2. Suffix Tree Clustering (STC):

- The transformation of web blog searching, powered by neural networks and deep learning algorithms, goes beyond mere keyword matching. It incorporates cutting-edge techniques like Suffix Tree Clustering (STC) and The Lingo algorithm to refine the search experience.
- STC is a robust clustering technique that aids in organizing web content into meaningful groups. It works by creating a hierarchical structure, like a

tree, that represents the relationships between words and phrases in the documents. This allows for efficient content grouping, making it easier to find blogs that are contextually related.

Example of a suffix tree:

Consider the word "banana":

Then the suffixes are:

banana\0

anana\0

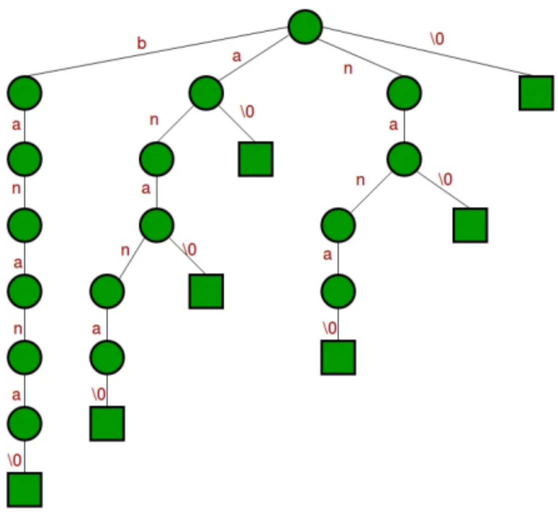
nana\0

ana\0

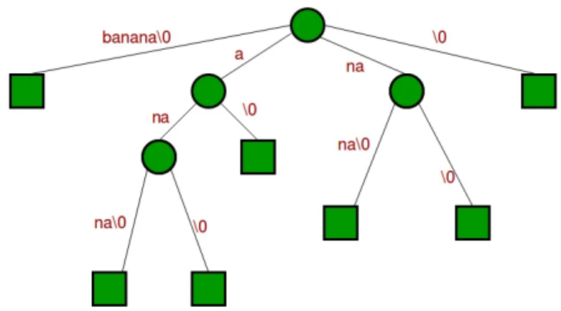
na\0

a\0

\0



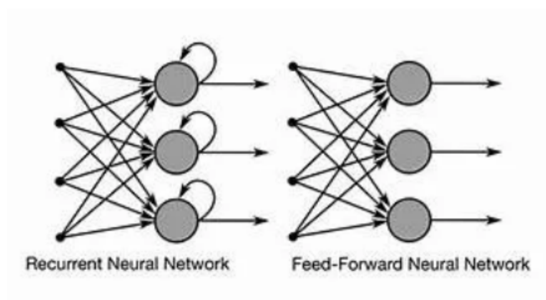
Standard trie for "banana"



Standard trie for "banana"

3. Recurrent neural networks:

Recurrent Neural Networks are deep learning models designed to handle sequential data. They are ideal for understanding patterns in text or other data that have a temporal or sequential nature. In the context of web blog searching, RNNs can be used to capture the context of blog content over time, enabling the system to make more informed recommendations based on the order and flow of information in the blogs.



Examples of Searching:

- Google Search: Google Search uses a variety of neural network techniques to improve the relevance and accuracy of its search results. For example, Google uses neural networks to understand the context of user queries, identify synonyms and related terms, and rank search results based on their relevance to the user's query.

- Bing Search: Bing Search also uses neural networks to improve the relevance and accuracy of its search results. For example, Bing uses neural networks to understand the meaning of user queries, identify the intent of the user, and rank search results based on their relevance to the user's intent.

- Quora: Quora uses neural networks to personalize the content that users see on their feeds. For example, Quora uses neural networks to understand the topics that users are interested in, the types of questions they ask, and the types of answers they like to read. Quora then uses this information to recommend new questions and answers to users.

- Reddit: Reddit uses neural networks to personalize the content that users see on their front pages. For example, Reddit uses neural networks to understand the subreddits that users subscribe to, the types of posts they upvote and downvote, and the types of comments they make. Reddit then uses this information to recommend new posts and comments to users.

- Amazon: Amazon uses neural networks to personalize the product recommendations that users see on its website. For example, Amazon uses neural networks to understand the products that users have purchased, the products they have viewed, and the products they have reviewed. Amazon then uses this information to recommend new products to users.

IV. Results and Discussions

Some of the use cases and the real-time implementation of algorithms are mentioned in the following techniques:

- The Google Algorithm

  When discussing Google's "algorithm," individuals often refer to how Google delivers search results. Many organizations closely monitor "algorithm changes" that impact their search traffic. However, Google's "algorithm" comprises a sophisticated network of AI-powered algorithms that dictate the appearance and ranking of search results. While outsiders lack a comprehensive understanding of these algorithms, they endeavor to grasp certain elements to enhance SEO practices. Google frequently provides insights into how its AI system processes search results, emphasizing the optimization for delivering the best possible results tailored to user queries. Presently, there is no definitive way to "beat the algorithm" aside from creating exceptionally high-quality content tailored to human users.

- RankBrain

  RankBrain serves as a fundamental component of AI within Google search, contributing significantly to enhancing search result accuracy. This system aids Google in comprehending the relevance of topics to users' search queries, thereby refining the search experience. By discerning the user's intent more accurately, RankBrain minimizes the likelihood of confusing search terms with similarly sounding topics that possess distinct meanings.

- BERT

  Google uses a sort of artificial intelligence called BERT, or Bidirectional Encoder Representations from Transformers, to better understand the meaning and intent of searches.BERT processes each word in a search query in relation to every other word in a sentence using artificial intelligence (AI) techniques such as sentiment analysis, natural language understanding (NLU), and natural language processing (NLP). The way Google search used to operate is not the same as this. Google's AI used to process words sequentially, one after the other. The outcomes were literal but correct. For instance, according to Google, a search query such as "2019 brazil traveller to usa need a visa" would have previously been read as coming from a US citizen who needed a visa to enter Brazil. In such case, prepositions alter the context of queries, which Google failed to take into consideration. The word "to" would not have been taken into consideration, which radically alters the search aim. BERT is not like that. Considering the entire text, BERT determines that the searcher is a Brazilian who needs a US visa. This leads to more precise searches that ultimately accomplish your goals for each given search.

- MUM

  MUM, short for Multitask Unified Model, represents a more advanced iteration of AI utilized by Google, surpassing the capabilities of BERT. Employing sophisticated AI techniques, MUM enhances its understanding of search context, intent, and queries in diverse languages. In a presentation, Google illustrated MUM's functionality with an example: Suppose a user inputs the search query, "I've hiked Mt. Adams and now want to hike Mt. Fuji next fall, what should I do differently to prepare?" MUM possesses the ability to comprehend the underlying context of this query, a feat that poses challenges for conventional search engines. MUM can provide insights into the similarities and differences between the two mountains and suggest appropriate equipment for the new hike.

Challenges in searching:

- Cold-Start Problem: For a new user, user interest data is very less items therefore it becomes difficult for recommendation systems to give good results.
- Data sparseness: Even for established users, data can be sparse. Many users may interact with only a few web searches, making it challenging to understand their preferences accurately.
- Privacy and Trust: Users might object to the collection of data regarding their personal preferences.
- Content Diversity: Web search can cover a wide range of topics and styles. Recommending diverse and relevant content that matches the user's interests can be difficult.
- Scalability: As the number of users grows, the computational complexity of recommendation systems increases. Scalability is a significant challenge, especially for real-time or large-scale platforms.

V. Conclusion

In today's digital landscape, the challenge of managing information overload and optimizing the search experience is widely acknowledged. To address this challenge, intelligent personalized web searching techniques driven by neural networks have emerged as a promising solution. These techniques have the potential to reshape the way individuals interact with online content, leveraging advanced technologies such as Graph Convolutional Networks (GCN), Suffix Tree Clustering (STC), the recurrent neural network algorithm. By employing these technologies, content organization is refined, and search accuracy is enhanced, promising a more tailored and efficient browsing experience.

Real-time recommendations, facilitated by neural networks, ensure that users receive relevant content suggestions promptly, aligning with their preferences and interests. Additionally, semantic search capabilities, empowered by neural networks' natural language understanding, enable the delivery of contextually relevant results, even when user queries exhibit variations in wording.

The ongoing evolution of personalized web searching methods represents a pivotal advancement in online information exploration. With AI-driven algorithms driving this transformation, the future promises a more user-centric, efficient, and rewarding journey through the vast expanse of digital content.

References

[1]. https://medium.com/@ashmi_banerjee/drawbacks-of-recommender-systems-e6a596fc937e

[2]. https://www.mdpi.com/2306-5729/4/1/7

[3]. https://www.algolia.com/blog/ai/neural-networks-or-ai-whats-the-difference-when-it-comes-to-search/

[4]. https://www.marketingaiinstitute.com/blog/how-search-engines-use-artificial-intelligence

[5]. Learning to Rank Personalized Search Results in Web Search" (2015) by Bhaskar Mitra et al.

[6]. Deep Learning for Personalized Search and Recommender Systems" (2017) by Liang Wu et al.

[7]. Personalized Web Search Using Deep Learning" (2018) by Navin Goyal et al.

[8]. Deep Reinforcement Learning for Personalized Web Search" (2019) by Lianhong Jin et al.

[9]. https://www.researchgate.net/publication/340594113_An_intelligent_personalized_web_blog_searching_technique_using_fuzzy-based_feedback_recurrent_neural_network

[10]. https://www.youtube.com/watch?v=2KRAOZIULzw

[11]. Felix Wu, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, Kilian Weinberger Proceedings of the 36th International Conference on Machine Learning, PMLR 97:6861–6871, 2019

[12]. https://www.geeksforgeeks.org/pattern-searching-using-suffix-tree/

[13]. Ji-Rong Wen, Zhicheng Dou and Ruihua Song "Personalized Web Search" ResearchGate 2015.

[14]. Sunny Sharma "A Systematic Literature Review Of Web Search Personalization" Recent Advances in Computer Science and Communications 2020.

Case Study Deep Learning in Personalized Web Searching_12010285_12010424_12011001.pdf

ORIGINALITY REPORT

4%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.coursehero.com Internet	32 words — 1%
2	Sunil N. Nikhate, Archana K. Chaudhari, Jayant V. Kulkarni. "Determination of Extent of Similarity between Mimic and Genuine Voice Signals Using MFCC Features", 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), 2019 Crossref	16 words — 1%
3	Rakhi Bharadwaj, Soham Ratnaparkhi, Rajendrasingh Rajpurohit, Kashish Rahate, Rahul Pandita, Samarth Thosar. "Deepfake detection for preventing Audio and Video frauds using Advanced Deep Learning Techniques", 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS), 2023 Crossref	14 words — 1%
4	Satheesh Abimannan, El-Sayed M. El-Alfy, Yue-Shan Chang, Shahid Hussain, Saurabh Shukla, Dhivyadharsini Satheesh. "Ensemble Multifeatured Deep Learning Models and Applications: A Survey", IEEE Access, 2023 Crossref	14 words — 1%
5	idus.us.es Internet	13 words — 1%