BAU Journal - Creative Sustainable Development

Volume 4 | Issue 2
ISSN: 2789-8334

Article 2

May 2023

A NOVEL SPECIALIZED SEARCH ENGINE FOR AI-MODELS AND THEIR COMPARISON

Imane Haidar
Senior Lecturer, Department of Electrical and Computer Engineering, Faculty of Engineering, Beirut Arab University, Lebanon, i.haidar@bau.edu.lb

Ziad Doughan
Assistant Professor, Department of Electrical and Computer Engineering, Faculty of Engineering, Beirut Arab University, Lebanon, z.doughan@bau.edu.lb

Ali M. Haidar Prof.
Professor, Department of Electrical and Computer Engineering, Faculty of Engineering, Beirut Arab University, Lebanon, ari@bau.edu.lb

Database, artificial intelligence, hybrid models, auto-system generation, performance, search engine, knowledge graph, recommender systems, AI models

A NOVEL SPECIALIZED SEARCH ENGINE FOR AI-MODELS AND THEIR COMPARISON

Abstract

In recent years, the world of AI has tremendously increased in size and depth. Both new and old researchers are facing the problem of fast emerging AI researches, models and services. One needs to continuously read complete papers to understand the idea behind any novel research. This work presents a novel AI service that removes the burdens of long text reading and uncategorized search. It consists of a website that categorizes all the AI researches in a well-designed database. The users just have to select the models they are interested in, and the website will return a table containing the technical data in addition to a graph that shows visual relationships between the AI models, features and datasets. Future work will emphasize on developing the tool by applying NLP in two directions: one on the search box to retrieve the main keyword to search for, and the other on research papers to automatically extract the data into the website categorized database.

Keywords

Database, artificial intelligence, hybrid models, auto-system generation, performance, search engine, knowledge graph, recommender systems, AI models

1. INTRODUCTION

Every five years, the amount of information doubles (Hendrickson2010), for that the need for intelligent search becomes crucial. The literature is full of general search engines that are based on natural language processing and knowledge based reasoning. The browser searches for the data by database indexing then returns a whole bunch of texts. For that, the need for a specialized search engine arises. Instead of returning complete papers and articles, the engine must give back a straightforward technical result to facilitate AI research and learning.

The latest technologies in the 5th-Generation (5G) infrastructure are promising to provide everything-as-a-service (XaaS) especially in the field of Artificial intelligence (AI). Furthermore, with the booming progress in edge computing, edge-AI concept present itself as a main tool for AI services [1]. Thus, building such services requires knowledge aggregation to satisfy both computation and memory requirements. In addition, the success of an AI-as-a-service (AIaaS) platform inquire a wide range of AI and Machine Learning (ML) models to be available [1]. Future works propose to provide multiple agent models to achieve a more robust tool [1].

Selecting the right model in the age of Big Data is becoming exponentially challenging due to enormous number available in this stochastic field of AI. Recommender systems can do a lot to guide users in designing their best models by exploring different aspects of both the application and the dataset [3]. Based on various sources of information, recommenders can present different scenarios of cognitive technologies, and develop the next level AI systems [3]. The progress in Deep Neural Networks (DNN) provided a new direction in recommender systems. This emerged technology proved successful in handling categorical features. Thus, state-of-the-art models for recommendation applications are excelling in simulation and system co-design [2].

2. LITERATURE REVIEW

Parallelism made it more effective in Big Data exploitation and system scaling [2]. Modern tools of recommendation use complex architecture to harness huge amounts of data. For example, applying convolutional (CNN), recurrent (RNN), and generative neural networks (GAN) efficiently may solve hard problems in minimum scales [2]. Such modern systems need access to a vast amount of big data. Although, it is important to focus on models' interconnection and information exchange. In return, this transforms the AI design phase to a fast, accurate and tailored approach where users can easily benefit and fulfill all their needs [3]. In association this parallelism, the effectiveness of AI models in Big Data needs to be strong in model generation. To succeed in advanced recommendation approaches, designer should focus on three factors: cognition, knowledge, and data. Thus, application of knowledge graphs and interaction tools have led to the following substantial recommender types:

- Cognition Recommendations: this type provides accurate and smart recommendations especially in the mechanism of discovering significant explicit and implicit patterns. Also, it may focus on the dimensionality of the models due to various aspects like Text-based, Feature-based, and Action-based applications [3].
- Sequential Recommendations: this type aims to follow a certain behavior on a sequence of inputs. Resent research also focus on a perception of cognitive meaning of a data sequence [3].
- Data-driven Recommendation: this type provide an understanding of Big Data. The main challenge lies in organizing and connecting data from various repositories. This "Data Lake" approach facilitates managing and querying information [3].
- Knowledge Recommendations: this type focuses on experience and knowledge to find patterns of recommendations. This approach uses informative data to guide recommenders in choosing the best system to fit the data [3].

Furthermore, the AI platforms must supply of various dataset types of like random, synthetic, and public data to enhance the training phase. This allows to compare difference models and detect bottlenecks for training and optimization processes [2]. Even though, this type of systems searches for best ML methodologies and prevailing AI techniques to improve cutting-

edge models' design. In addition, it may present the current research topics in practice and optimization. This will directly support users and researchers to carefully choose between multiple ML techniques on a safer basis [4]. Thus, to provide better exploration of the available models, recommenders visualization offers a mapping tool to quickly understand models' comparison. In addition, an interactive interface works a sandbox for testing and evaluation. A lot of recent works discuss the importance of visualization in explaining and selecting models, especially for hybrid solution [4].

Recently, knowledge graphs (KG) are applied side to side with recommenders to provide a semantic explanation of models' intuition and usage. Diverse information sources of AI systems can lead to the construction of large KGs, in a hierarchical integration structure [5]. This leads to a more explainable AI representation. Thus, to achieve the required scope of any application. The platform should provide extracting tools to harvest datasets features, relations, and meaning. Also, illustrate the usage of the AI models as a pre-model, post-model, or in-model [5]. Accordingly, the use of KGs is important to describe the domain of AI from general models to the specific cases. Thus, KGs provide the infrastructure to build search engines or question-answer tools. KGs are usually applied in the following four applications:

- Graph construction to depict different technological concepts and associate related models in groups. This organization and annotation present different types of information in a meaningful distribution.
- Graphs for features extraction to analyze data entities, which are then used to semantically propose AI models and improve the search results.
- Graphs for relations extraction to find interoperability and standardization among different AI models. Although, this approach tries to solve conflicts among unresolved models' integrations.
- Graphs for reasoning to accurately evaluate predictions, compare between models, and find conclusions following a systematic inference.

In conclusion, the available KG techniques support the selection and development of machine learning algorithms. Although, KGs are mostly applied to reason in AI models design, their usage is extended to pre- and post-modelling phases [5]. In this work, graph is used to visualize the result of searching query; the thing that enables accelerated learning and research. In addition, it gives the researcher a clear overview of the models, their details, the relationship with other models and AI techniques, and most importantly gives them the sparks to find novel models and solve AI problems.

3. PROPOSED WORK

Usually, when researchers want to build a new AI model or develop a certain technique, they surf the internet to know what is already done in the literature review. The next step would be putting time and effort to read the research before comparing it to the proposed work.

On the other hand, research papers are not categorized, they just have abstract and keywords that help the reader comprehend the essential idea of the paper, but they are not enough to draw a clear image of the relevance of the work with respect to the field of research.

In this paper, a novel service is provided to remove all the burdens a researcher faces. It consists of a dynamic website that serves as a graphical interface for searching. Users could type what model they are searching for, like CNN, RNN, … and the website will return back all the detailed needed such as number of layers, parameters, the accuracy of the model on different datasets, …

The website also offers the ability to compare many models together in terms of the features the users select. For example, it can compare between the CNN software based model vice versa the CNN hardware based model in terms of the speed feature.

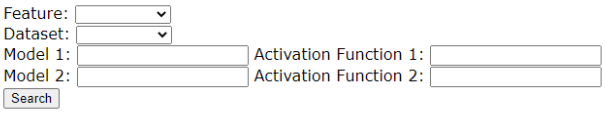
In the back-end, a well-designed database contains all the models, techniques, mathematical formulas, philosophical concepts … of AI researches. When submitting the search query, the website searches for the required data and draw a comparison table in addition to a graph that explicitly shows the relationships between the searched items.

4. EXPERIMENT

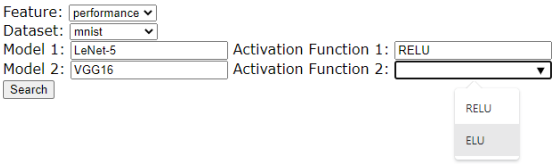
This work presents a novel AI service that let the users search for AI models and get their details. The service also has the ability to compare between models using tables and graphs.

A prototype is done to illustrate the idea. The website will be under continuous development to integrate new services needed by researchers and learners.

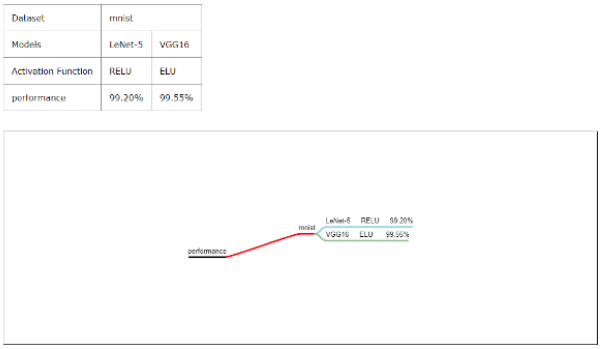
Each service contains a search box or a form for the user to search queries or select options.



The user has the ability to write the model name or select it from the list.



When submitted, the website shows a table containing the data plus a graph illustrating the tree of models details.



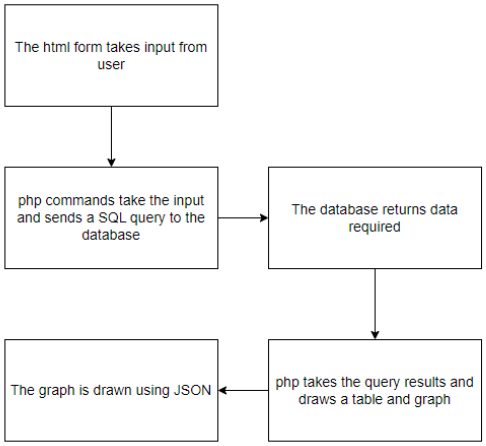
The table details the models' criteria and values. In the example illustrated, the performance feature is chosen, in addition to the MNIST dataset. The table then compares the model's performance based on the activation function.

Dataset	mnist	
Models	LeNet-5	VGG16
Activation Function	RELU	ELU
performance	99.20%	99.55%

The graph illustrated the aforementioned details. This helps the researchers view the relationship between all details to learn it, and conclude advanced model. In future work, this will be extended to larger and more complicated graphs.

The software is implemented using html and css for webpage building. Javascript is used for internal functions and modifications. The form input is sent as SQL queries to the database which returns them as data. Php commands are responsible for website-database connection. When received, the data is stored in a table then drawn as a graph using JSON.

In future work, AI models will be used to enhance the functionality of the website. NLP and reasoning are examples of the concepts to be employed for a better user experience.



The tools and languages used:

- WAMP Server
- Phpmyadmin database
- HTML, CSS, Javascript, PHP
- JSON
- Python will be added for AI analysis and NLP

5. ALTERNATIVE DESIGNS

The graph drawn to represent the results could be created using alternative graph designs.

A simple one is implemented in this work while a more advanced one is to be created in future works.

- Knowledge Graph [6]
- Graph Neural Network [7]
- Property graph [8]
- Petri nets [9]
- Automata [10]

6. CONTRIBUTION

- Create a platform that offers a graphical interface between researchers and technical data without the need to read complete articles and papers.
- Design a database that gives the ability to search for the details of any model or techniques
- Develop an algorithm that compares AI models and techniques
- Implement a graphical tree of the results.
- Offer a teaching tool for AI instructors to illustrate the importance of mathematical formulas and AI algorithms in every AI model.

7. CONCLUSION

This work presents a novel AI service that facilitates AI learning and research. It consists of a website which takes user query and returns back straightforward and technical results without the need to read complete papers and articles. The results are given in the form of well categorized table that shows the comparison between models in terms of the feature chosen by user. In addition, a graph is drawn showing all the details of selected AI models. Using this AI search engine service will accelerate the process of learning and research since it helps the researcher visualize the relationships between models. Also it helps them in inferring new remarks and models based on the general and detailed overview offered by the service. The future work will emphasize on developing the tool by integrating NLP to analyze queries and retrieve technical data from research papers.

REFERENCES

- M. S. Munir, S. F. Abedin and C. S. Hong, "Artificial Intelligence-based Service Aggregation for Mobile-Agent in Edge Computing," *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, 2019, pp. 1-6, doi: 10.23919/APNOMS.2019.8892984.
- *Naumov, Maxim & Mudigere, Dheevatsa & Shi, Hao-Jun & Huang, Jianyu & Sundaraman, Narayanan & Park, Jongsoo & Wang, Xiaodong & Gupta, Udit & Wu, Carole-Jean & Azzolini, Alisson & Dzhulgakov, Dmytro & Mallevich, Andrey & Cherniavskii, Ilia & Lu, Yinghai & Krishnamoorthi, Raghuraman & Yu, Ansha & Kondratenko, Volodymyr & Pereira, Stephanie & Chen, Xianjie & Smelyanskiy, Misha. "Deep Learning Recommendation Model for Personalization and Recommendation Systems",2019.*
- *Elahi, Mehdi & Beheshti, Amin & Goluguri, Srinivasa. (2021). Recommender Systems: Challenges and Opportunities in the Age of Big Data and Artificial Intelligence, 2021, 10.1201/9781003102380-2.*
- *Zhang, Qian & Lu, Jie & Jin, Yaochu., Artificial intelligence in recommender systems. Complex & Intelligent System, 2020, 7. 10.1007/s40747-020-00212-w.*
- *Rajabi, Enayat & Etminani, Kobra. Knowledge-graph-based explainable AI: A systematic review. Journal of Information Science, 2022, 016555152211128. 10.1177/01655515221112844.*
- *Hogan, Aidan & Blomqvist, Eva & Cochez, Michael & d'Amato, Claudia & de Melo, Gerard & Gutierrez, Claudio & Kirrane, Sabrina & Labra Gayo, Jose & Navigli, Roberto & Neumaier, Sebastian & Ngonga Ngomo, Axel-Cyrille & Polleres, Axel & Rashid, Sabbir & Rula, Anisa & Schmelzeisen, Lukas & Sequeda, Juan & Staab, Steffen & Zimmermann, Antoine. (2021). "Knowledge Graphs". ACM Computing Surveys, 2021, 54. 1-37. 10.1145/3447772.*
- *Chai, Ziwei & You, Siqi & Yang, Yang & Pu, Shiliang & Jiarong, Xu & Cai, Haoyang & Jiang, Weihao. "Can Abnormality be Detected by Graph Neural Networks?". In the proceedings of Thirty-First International Joint Conference on Artificial Intelligence {IJCAI-22}, 2022, 1920-1926. 10.24963/ijcai.2022/267.*
- *Bonifati, Angela & Dumbrava, Stefania & Martinez, Emile & Ghasemi, Fatemeh & Jaffré, Malo & Luton, Pacôme & Pickles, Thomas. "DiscoPG: Property Graph Schema Discovery and Exploration. Proceedings of the VLDB Endowment". Proceedings of the VLDB Endowment (PVLDB) At: Sydney, Australia, 2022, 15. 10.14778/3554821.3554867.*
- *Blaga, Florin & Pop, Alin & Hule, Voichița & Indre, Claudiu.. "The efficiency of modeling and simulation of manufacturing systems using Petri nets". IOP Conference Series: Materials Science and Engineering, 2021, 1169. 012005. 10.1088/1757-899X/1169/1/012005.*

- *C. A. Tecson and M. M. T. Rodrigo, "Tutoring Environment for Automata and the Users' Achievement Goal Orientations," 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), 2018, pp. 526-533, doi: 10.1109/TALE.2018.8615234.*