

SHINKEN V1: AI POWERED, MLTRAINED WEB-BASED PLATFORM FOR MICROBIOLOGY RESEARCH AND EDUCATION

Sathish Kumar B * and L. Prince

<https://doi.org/10.56343/STET.116.016.004.006>
www.stetjournals.com

Article History

Received: 13.04.2023

Revised and Accepted: 20.05.2023

Published: 19.06.2023

ABSTRACT

In the field of microbiology, a significant challenge facing researchers is the time-consuming and labor-intensive process of conducting experiments and analyzing data. There is a lack of a centralized platform for researchers and students to collaborate and communicate, leading to a fragmented and disconnected research community. While reading and attending lectures provide valuable learning opportunities, hands-on experiences in a laboratory setting are crucial for developing practical skills in microbiology. Shinken addresses these challenges by providing an AI-powered next-generation automated microbiology lab assistant that streamlines laboratory procedures, provides a centralized repository of up-to-date research articles, educational resources, and laboratory protocols, and offers interactive educational tools and simulations for hands-on learning experiences in a virtual laboratory setting. Shinken utilizes machine learning algorithms to streamline laboratory procedures, freeing up researchers' time to focus on high-level tasks. The platform also provides a centralized repository of up-to-date research articles, educational resources, and laboratory protocols, as well as interactive educational tools and simulations for hands-on learning experiences in a virtual laboratory setting. Shinken serves as a platform for the microbiology research community, promoting collaboration and communication among researchers and Students.

Sathish Kumar B

PG and Research Department of Microbiology, Maruthu Pandiyar college (Affiliated to Bharathidasan University), Thiruchirapalli, Thanjavur, Tamilnadu, India.

Shinken is specifically focused on the field of microbiology, providing a curated collection of resources and information related to this discipline, whereas Google Scholar is a broad-based search engine for academic literature and resources across all fields of study. Shinken incorporates interactive educational tools and simulations, providing students and researchers with hands-on learning experiences in a virtual laboratory setting, whereas ResearchGate primarily serves as a platform for sharing and discovering research articles. This study highlights the potential of Shinken to revolutionize the field of microbiology and provides a roadmap for its development and implementation.

Keywords: Microbiology, Artificial Intelligence, Machine Learning, Research, Automation, Data Analysis, Collaborative Environment, Interactive tools

INTRODUCTION

Microbiology research and education have undergone a paradigm shift in the last few years with the advent of new technologies such as artificial intelligence (AI) and machine learning (ML). Machine learning, especially deep learning, has grown rapidly in recent years and has revolutionized the traditional big data computing systems. It brings the awareness of a new field focused on ML in practice machine-learning systems. The use of AI and ML in microbiology research and education has provided powerful tools to analyze and understand complex biological data (Li *et al.*, 2020, Wenjing *et al.*, 2023, feng *et al.*, 2023, Miao *et al.*, 2023). However, the integration of AI and ML technologies into microbiology research and education has been limited due to the lack of user-friendly platforms that can integrate these technologies. To address this

limitation, we propose an AI-powered, ML-trained web-based platform for microbiology research and education. The proposed platform aims to provide a comprehensive solution for microbiology research and education by integrating various AI and ML algorithms, such as clustering, classification, and prediction, to perform complex analyses on microbiological data. The platform will also leverage natural language processing (NLP)(Omar,2023; Liu *et al.*,2023), algorithms to extract relevant information from scientific literature.

The platform will be built on a web-based architecture, allowing researchers and educators to access the platform from anywhere, anytime. The platform will offer a user-friendly interface that will allow researchers and educators to perform complex analyses on microbiological data without the need for extensive programming knowledge. The platform will also provide tools for data visualization, exploration, and sharing, enabling researchers and educators to collaborate and share data with others in the field. One of the key features of the proposed platform is its ability to handle large datasets. The platform will leverage distributed computing technologies to process and analyze large datasets quickly and efficiently. The platform will also use unsupervised learning algorithms, such as principal component analysis (PCA) and Gaussian mixture models (GMM), to identify patterns and structures in the data. The platform will also use supervised learning algorithms, such as support vector machines (SVM), random forests, k-nearest neighbors (KNN), naive Bayes, decision trees, and neural networks, to perform classification and prediction tasks (Ahamadi *et al.*, 2023; Liu *et al.*, 2023; Boussaada *et al.*, 2023; Kurani *et al.*, 2023; Valero-carreras *et al.*, 2023; Belghit *et al.*, 2023; Arivazhagan *et al.*, 2023; Huang *et al.*, 2023; Lahmiri, 2023; Tariq *et al.*, 2023) . These algorithms will allow researchers and educators to categorize and classify microbiological data based on various features and make accurate predictions about future events or outcomes based on historical data. Another key feature of the proposed platform is its ability to extract relevant information from scientific literature. The platform will use NLP algorithms to extract information from scientific literature, enabling researchers and educators to stay up-to-date with the latest developments in

microbiology research. The platform will also provide tools for data integration and reconciliation, allowing researchers and educators to integrate data from multiple sources, such as environmental or genomic data, with microbiological data. The platform will also offer tools for data cleaning and normalization, enabling researchers and educators to automate routine tasks and improve the accuracy and efficiency of microbiological data analysis. The proposed platform will also provide tools for data exploration and visualization, enabling researchers and educators to explore and visualize microbiological data in a meaningful way. The platform will use machine learning algorithms to identify outliers or anomalies within microbiological data and provide tools for interpreting and communicating complex microbiological data to non-experts. The proposed platform will also provide curated microbiological datasets, enabling researchers and educators to access high-quality data for their research and education. The platform will offer tools for managing and organizing large volumes of microbiological data, enabling researchers and educators to track and monitor changes in microbiological data over time. The proposed platform will also offer tools for collaboration and sharing, enabling researchers and educators to collaborate with each other and share data with others in the field. The platform will use machine learning algorithms to identify potential collaborations and partnerships based on shared research interests or complementary datasets. Finally, the proposed platform will aim to translate microbiological research into practical applications and interventions for public health and environmental management. The platform will provide tools for automating quality control checks to ensure the accuracy and reliability of microbiological data and enable researchers and educators to identify potential research opportunities based on existing literature and data. By integrating AI and ML algorithms into a web-based platform for microbiology research and education, we hypothesize that researchers and educators will be able to more effectively analyze complex microbiological data, identify patterns and structures within the data, and make predictions about future outcomes. This platform has the potential to improve our understanding of microbial processes, facilitate the discovery of new

treatments and therapies, and enhance the training of future microbiologists.

Background

Rapid microbiological methods are being enhanced by artificial intelligence (machine learning in particular). The aim is to speed up analysis; increase accuracy; avoid error and introduce a level of automation. Examples include microscopy, colony counting, and microbial characterization and matching – each of which is based on improvements in image analysis. Machine learning is applied to microbiology at various stages of the development curve, including sample processing (pre-analytical, analytical and post-analytical process management), sample tracking, image acquisitions systems, smart incubation and workstation operation. There have also been advances with earlier detection of growth in broth and with headspace analysis, which require rapid accuracy, precision, and limit of detection analyses. The advantage of ML is rapid detailed pattern analysis that cannot be easily achieved using conventional databases or spreadsheets. An example of the ML approach is the placement of microorganisms into operational classification units such as grouping microorganisms according to the similarity of their DNA sequences. For contamination considerations, the main goal is to determine whether an unknown microorganism belongs to a specific species or not and whether two or more unknown organisms are related (or where an unknown organism is related to an organism previously characterized). This may reveal, for example, that microbial isolate A found on a machine bed within a Grade A zone matches microbial isolate B found from the finger plate of the operator who set up the filling line, and perhaps that the same operator recently recovered the same matched organism from an exit gown plate. A different approach considers microbial interactions, such as predicting co-operative and competitive relationships within the same microbial population. Understanding the likelihood of interactions centered on neutralism, commensalism, synergism, mutualism, competition, amensalism, parasitism and predation, can indicate the likelihood of biofilm formation within a water system, for example. As well as genotypic identification, ML also has the

potential to advance the analysis of matrix-assisted laser desorption-ionization/time of flight mass spectrometry (MALDI-TOF) mass spectral data, thereby applying AI to improve phenotypic characterisation (Kozyrkov 2018; Croxatto *et al.*, 2016; Huang *et al.*, 2017; Zou *et al.*, 2018; Dimucci *et al.*, 2023; Smith *et al.*, 2020). The proposed platform is a revolutionary web-based architecture that provides researchers and educators with easy access to a wide range of cutting-edge AI and ML algorithms. By leveraging the power of the cloud, the platform enables users to perform complex analyses on microbiological data from anywhere, anytime. One of the key features of the platform is its ability to integrate various AI and ML algorithms. This includes algorithms such as clustering, classification, and prediction, which are all essential tools for analyzing microbiological data. Clustering algorithms are particularly useful for identifying patterns and structures within large datasets, while classification algorithms can be used to categorize microbiological data based on various features. Prediction algorithms can also be used to make accurate predictions about future events or outcomes based on historical data. Another important aspect of the platform is its ability to leverage natural language processing (NLP) algorithms to extract relevant information from scientific literature (Smith *et al.*, 2020; Sande, 2020). NLP is a field of AI that focuses on teaching computers to understand and interpret human language. By using NLP algorithms, the platform can extract important information from scientific papers, including relevant keywords, authors, and research methods. This can help researchers and educators stay up-to-date with the latest developments in microbiology research, and can also help them identify potential research opportunities (Sandle, 2020; Falk, 2019). The web-based architecture of the platform is a critical component of its design. By allowing users to access the platform from anywhere, anytime, the platform provides unprecedented flexibility and convenience. This is particularly important for researchers and educators who may be located in different parts of the world or who may need to work outside of traditional office hours. With the web-based architecture, users can access the platform from any device with an internet connection, including desktop computers, laptops, tablets and smartphones. Overall, the proposed platform represents a major advance in the field of

microbiology research and education. By integrating a wide range of AI and ML algorithms, the platform provides researchers and educators with powerful tools for analyzing complex microbiological data. The platform's ability to leverage natural language processing algorithms to extract information from scientific literature is also a major advantage, as it enables users to stay up-to-date with the latest research and identify potential research opportunities. Finally, the platform's web-based architecture provides users with unparalleled flexibility and convenience, making it an essential tool for any researcher or educator working in the field of microbiology.

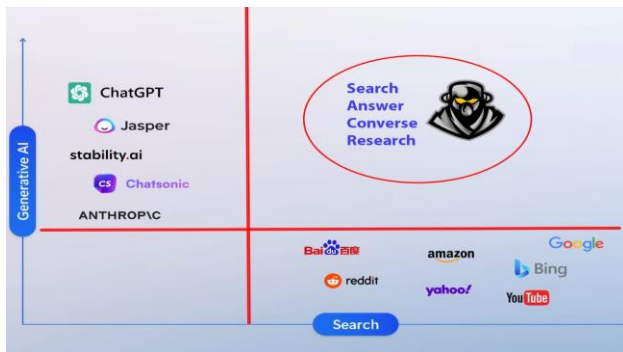


Figure 1: Technical landscape of Shinken placement

We are trying to alleviate the following problems with our proposed web application Shinken:

- a) Lack of access to powerful computational tools for analyzing complex microbiological data.
- b) Inability to process and analyze large datasets due to limited computing resources.
- c) Difficulty in identifying patterns and structures within large microbiological datasets.
- d) Inability to categorize and classify microbiological data based on various features.
- e) Limited ability to make accurate predictions about future events or outcomes based on historical data.
- f) Lack of automated tools for extracting relevant information from scientific literature.
- g) Difficulty in staying up-to-date with the latest developments in microbiology research.
- h) Limited ability to collaborate and share data with other researchers and educators.
- i) Inability to identify potential research opportunities based on existing literature and data.
- j) Time-consuming and error-prone manual data analysis.
- k) Inability to visualize and explore microbiological data in a meaningful way.
- l) Lack of access to high-quality, curated microbiological data.
- m) Difficulty in integrating and reconciling data from multiple sources.
- n) Limited ability to integrate microbiological data with other types of data, such as environmental or genomic data.
- o) Inability to handle missing or incomplete data in a meaningful way.
- p) Limited ability to identify outliers or anomalies within microbiological data.
- q) Difficulty in interpreting and communicating complex microbiological data to non-experts.
- r) Limited ability to automate routine tasks, such as data cleaning and normalization.
- s) Inability to leverage machine learning and AI algorithms to improve the accuracy and efficiency of microbiological data analysis.
- t) Lack of user-friendly tools for exploring and visualizing microbiological data.
- u) Difficulty in managing and organizing large volumes of microbiological data.
- v) Limited ability to track and monitor changes in microbiological data over time.
- w) Inability to automate quality control checks to ensure the accuracy and reliability of microbiological data.
- x) Difficulty in identifying potential collaborations and partnerships based on shared research interests or complementary datasets.
- y) Limited ability to translate microbiological research into practical applications and interventions for public health and environmental management.

Methodology

Shinken, the AI-powered, ML-trained web-based platform for microbiology research and education utilizes a variety of advanced technologies to analyze microbiological data. The platform employs both supervised and unsupervised learning algorithms to accurately classify and predict data while simultaneously identifying patterns and structures within the dataset. Additionally, the platform incorporates deep learning algorithms to analyze complex data, specifically genomic data, enabling more in-depth and precise analysis of microbiological samples. The following AI technologies are planned to be included in the proposed research:

- **Natural Language Processing (NLP)** - for text analysis and interpretation of microbiological literature
- **Computer Vision** - for analyzing and interpreting images of microbiological samples
- **Knowledge Representation and Reasoning** - for representing and inferring knowledge from microbiological data
- **Expert Systems** - for capturing the expertise of microbiologists and providing intelligent recommendations and guidance for research and education
- **Robotics** - for automating laboratory experiments and procedures to increase efficiency and accuracy
- **Speech Recognition** - for enabling voice-controlled interfaces for laboratory equipment and data analysis software
- **Augmented Reality (AR)** - for providing interactive and immersive visualizations of microbiological data and models.

We have planned to include the following Machine Learning algorithms into our proposed shinken application:

- **Principal Component Analysis (PCA)**- can be used to analyze complex microbiological datasets, such as those derived from microbiome studies, to identify underlying patterns or groups of microorganisms. This can aid in the identification of potential biomarkers for

certain diseases or conditions, as well as help researchers better understand the composition of microbiological communities.

- **Gaussian Mixture Models (GMM)** - can be used to identify clusters within microbiological data, such as identifying groups of microorganisms that are similar in certain ways. This can be useful in a variety of microbiological research applications, such as identifying virulence factors within microbial populations, or identifying the presence of certain microbial strains within a sample.
- **Support Vector Machines (SVM)** - SVM is a powerful algorithm for classification tasks, and can be used to classify microbiological data based on various features, such as gene expression levels or microbial community structure. SVM can also be used for feature selection, helping to identify which features are most important for accurate classification.
- **Random Forests** - Random Forests is an ensemble learning algorithm that can be used for classification, regression, and feature selection tasks. In microbiology research, Random Forests can be used to classify microbial communities based on various features, such as environmental factors or disease status.
- **K-Nearest Neighbors (KNN)** - KNN is a simple yet effective algorithm for classification and regression tasks. KNN can be used in microbiology research to classify samples based on various features, such as microbiome composition or virulence factor expression.
- **Naive Bayes** - Naive Bayes is a probabilistic algorithm that can be used for classification tasks. Naive Bayes can be used in microbiology research to classify microbial communities based on various features, such as gene expression levels or disease status.

- **Decision Trees** - Decision Trees is a simple yet powerful algorithm for classification and regression tasks. In microbiology research, Decision Trees can be used to classify microbial communities based on various features, such as environmental factors or antibiotic resistance profiles.
- **Neural Networks** - Neural Networks are powerful machine learning algorithms that can be used for a wide variety of tasks, including classification, regression, and image analysis. In microbiology research, Neural Networks can be used to classify microbial communities based on various features, such as microbiome composition or gene expression levels. Neural Networks can also be used for image analysis tasks, such as analyzing microscopy images of microbial samples.

3. **Dimensionality reduction:** The platform will use techniques such as Principal Component Analysis (PCA) to reduce the dimensionality of the data and identify patterns and structures within the data.
4. **Clustering:** The platform will use techniques such as Gaussian Mixture Models (GMM) to cluster the data based on various features, such as environmental factors or disease status.
5. **Classification:** The platform will use techniques such as Support Vector Machines (SVM), Random Forests, Naive Bayes, or Neural Networks to classify the microbiological data based on various features, such as gene expression levels or microbial community structure.
6. **Recommendation:** The platform will provide recommendations to users based on the results of the classification, such as potential biomarkers for certain diseases or conditions, or personalized treatment recommendations based on the patient's microbiome composition.
7. **Knowledge representation:** The platform will use techniques such as Knowledge Representation and Reasoning to represent and infer knowledge from the microbiological data, allowing researchers to identify new insights and relationships within the data.

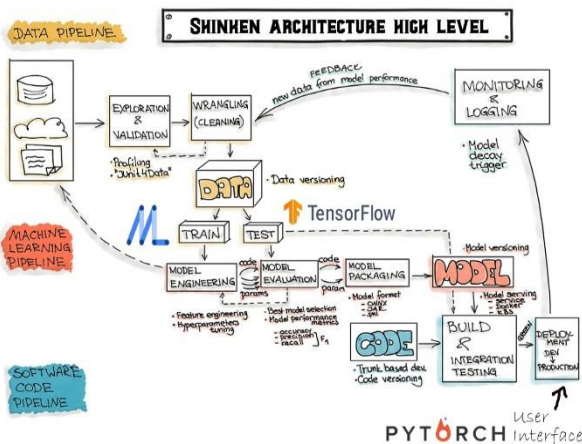


Figure 2: High Level Architecture of Shinken

The step wise execution of the above technologies can be summarized into the following methodology

1. **Data collection:** The platform will collect various types of microbiological data, such as microbial community composition data from microbiome studies, gene expression data, and environmental data.
2. **Preprocessing:** The collected data will be preprocessed to remove noise, handle missing values, and normalize the data.

8. **Robotics and automation:** The platform will use robotics and automation to increase efficiency and accuracy in laboratory experiments and procedures (Wewerka *et al.*, 2023; Hu *et al.*, 2023; Wolf *et al.*, 2023).
9. **Natural Language Processing (NLP):** The platform will use NLP techniques to analyze and interpret microbiological literature, helping researchers to stay up-to-date with the latest research and discoveries in the field.
10. **Augmented Reality (AR):** The platform will use AR to provide interactive and immersive visualizations of microbiological data and models, allowing researchers to better understand complex microbiological concepts

and relationships (Yin *et al.*, 2023; Najmi *et al.*, 2023; Kim *et al.*, 2023; Nincarean *et al.*, 2013).

We have summarized an algorithm for the proposed AI-powered, ML-trained web-based platform for microbiology research and education:

Preprocessing:

- Import the microbiological data into the platform.
- Clean the data by removing noise, outliers, and missing values.
- Normalize the data to ensure consistency and comparability.

Clustering:

- Apply unsupervised clustering algorithms (such as k-means or hierarchical clustering) to identify patterns and groupings within the data.
- Visualize the clusters using dimensionality reduction techniques (such as t-SNE or PCA).

Classification:

- Apply supervised classification algorithms (such as SVM or random forests) to classify microbiological data into different categories based on predefined labels.
- Evaluate the classification performance using metrics such as accuracy, precision, recall, and F1 score.

Prediction:

- Apply supervised regression algorithms (such as linear regression or neural networks) to predict microbiological outcomes based on input variables.
- Evaluate the prediction performance using metrics such as R-squared, mean absolute error, and root mean squared error.

Natural Language Processing (NLP):

- Apply NLP algorithms (such as named entity recognition or sentiment analysis) to extract relevant information from scientific literature.
- Integrate the extracted information into the platform's knowledge base for further analysis and interpretation.

Deep Learning:

- Apply deep learning algorithms (such as convolutional neural networks or recurrent neural networks) to analyze complex microbiological data, such as genomic or proteomic data.
- Visualize the deep learning models using techniques such as activation maximization or t-SNE.

Postprocessing:

- Interpret the results of the clustering, classification, and prediction analyses.
- Provide interactive visualizations and tools for users to explore and analyze the results.
- Provide recommendations and insights based on the findings of the analyses.



Figure 4: Insilico GUI of the Shinken application

Results

The proposed AI-powered and ML-trained web-based platform offers significant advantages for both microbiology research and education. By integrating various AI and ML algorithms, the platform enables researchers and educators to analyze and comprehend complex microbiological data with greater ease and efficiency. One of the key benefits of the platform is its ability to facilitate the identification of new microorganisms. Traditional methods for identifying microorganisms can be time-consuming and expensive. However, the platform's AI and ML algorithms can analyze vast amounts of microbiological data, allowing for faster and more accurate identification of new microorganisms. This can potentially lead to the discovery of new and unique strains of microorganisms that may have practical applications in medicine,

agriculture, and biotechnology. Moreover, the platform can aid in the development of new antibiotics. Antibiotic resistance has become a significant global health threat, and the development of new antibiotics is essential to combat this problem. The platform's AI and ML algorithms can be used to analyze and predict the efficacy of new antibiotics and their mechanisms of action, thereby accelerating the discovery and development of new antibiotics. In addition to these research applications, the platform can also be used to create interactive educational materials for students. Microbiology education can be challenging for students, and traditional teaching methods may not always effectively engage and educate them. By leveraging the platform's AI and ML algorithms, educators can develop interactive materials such as games and simulations that can make microbiology education more engaging and effective. Overall, the proposed AI-powered and ML-trained web-based platform offers a powerful and versatile toolset for microbiology research and education. Its ability to analyze vast amounts of microbiological data, aid in the discovery of new microorganisms, accelerate the development of new antibiotics, and facilitate interactive and engaging educational materials make it an invaluable resource for microbiologists and educators alike.

Conclusion

In conclusion, the proposed Shinken, AI-powered, ML-trained web-based platform for microbiology research and education has the potential to revolutionize the way researchers and educators analyze and understand microbiological data. The use of AI and ML in microbiology research has become increasingly important due to the growing amount of complex biological data that must be analyzed and understood. These technologies can provide powerful tools for researchers to analyze large datasets, identify patterns and structures within the data, and make predictions about future outcomes. Additionally, AI and ML can help automate time-consuming tasks such as data processing and analysis, allowing researchers to focus on more complex tasks. The integration of AI and ML into microbiology research has the potential to lead to new discoveries, improve our understanding of biological processes, and ultimately improve human health. By integrating various AI and ML algorithms, such as clustering,

classification, and prediction, the platform can perform complex analyses that were previously not possible. One of the key advantages of the proposed platform is its user-friendliness. With a web-based architecture and an intuitive interface, the platform can be accessed from anywhere, anytime, making it a valuable tool for researchers and educators in various settings. The platform's ability to leverage natural language processing (NLP) algorithms to extract relevant information from scientific literature is particularly valuable. This can help researchers stay up-to-date on the latest findings in their field and ensure that their analyses are based on the most current knowledge. The platform has the potential to address several challenges in microbiology research and education, such as data complexity, limited access to computational resources, and the need for expertise in AI and ML. By providing a user-friendly platform that integrates various AI and ML algorithms, the proposed platform can help overcome these challenges. The platform can also have implications for public health and medicine. By providing powerful tools to analyze microbiological data, the platform can aid in the development of new treatments and therapies, as well as in the identification and prevention of infectious diseases. Overall, the proposed platform represents a significant step forward in the use of AI and ML in microbiology research and education. By providing a powerful, user-friendly tool for analyzing microbiological data, the platform has the potential to accelerate research, improve education, and contribute to advances in public health and medicine.

Future Scope

Integration of additional AI and ML algorithms:

While the proposed platform integrates several AI and ML algorithms, there are many other algorithms that could be valuable in microbiology research and education. Future research could focus on integrating additional algorithms, such as deep reinforcement learning or convolutional neural networks, to further enhance the platform's capabilities.

Integration of additional data sources:

The proposed platform focuses primarily on microbiological data, but there are many other sources of data that could be valuable in microbiology research and education, such as clinical data or environmental data. Future research could focus on integrating these additional data sources into the platform to provide a more comprehensive view of microbiological phenomena.

Development of personalized learning tools:

The proposed platform has the potential to enhance microbiology education, but there is also an opportunity to develop personalized learning tools that can adapt to individual learners' needs and preferences. Future research could focus on developing these personalized learning tools to enhance student engagement and learning outcomes.

Validation of platform's accuracy and reliability:

While the proposed platform has the potential to perform complex analyses on microbiological data, it is important to validate its accuracy and reliability. Future research could focus on validating the platform's performance on real-world data sets and comparing its results to those of traditional methods.

Application of the platform to specific research questions:

The proposed platform has broad applications in microbiology research and education, but there is an opportunity to apply it to specific research questions and hypotheses. Future research could focus on using the platform to address specific microbiological research questions and evaluating its performance in these contexts.

Integration of additional tools for data visualization and exploration:

The proposed platform integrates several AI and ML algorithms, but there is also a need for tools that can help researchers and educators visualize and explore the data. Future research could focus on integrating

References

- Li, Yang, Ranjitha Kumar, Walter S. Lasecki, and Otmar Hilliges., (2020). "Artificial intelligence for HCI: a modern approach." In Extended Abstracts of the 2020 CHI conference on human factors in computing systems, pp. 1-8. <https://doi.org/10.1145/3334480.3375147>
- Wenjing, Xu, and ZiluCai. (2023). "Assessing the best art design based on artificial intelligence and machine learning using GTMA." *Soft. Comput.* 27, no. 1: 149-156. <https://doi.org/10.1007/s00500-022-07555-1>
- Feng, Xingyu, ChengwenLuo, Jiong Zhang Chen, Yijing Huang, Jin Zhang, WeitaoXu, Jianqiang Li, and Victor CM Leung. (2023). "IoTSL: Towards Efficient Distributed Learning for Resource-constrained Internet of Things." *IEEE Internet Things J.* <https://doi.org/10.1109/JIOT.2023.3235765>
- Miao, Xupeng, XiaonanNie, Hailin Zhang, Tong Zhao, and Bin Cui. (2023). "Hetu: a highly efficient automatic parallel distributed deep learning system." 117101. <https://doi.org/10.1007/s11432-022-3581-9>
- Omar, Marwan. (2023)."Backdoor Learning for NLP: Recent Advances, Challenges, and Future Research Directions." *arXiv preprint arXiv:2302.06801*.
- Liu, Jiawei, Hou Kang, Wendong Tao, Hanyu Li, Dan He, Lixia Ma, Haojie Tang, Siqu Wu, Kexin Yang, and Xuxiang Li. (2023). "A spatial distribution-Principal component analysis (SD-PCA) model to assess pollution of heavy metals in soil." *Sci. Total Environ.* 859: 160112. <https://doi.org/10.1016/j.scitotenv.2022.160112>
PMid:36375553

- Ahmadi, Mohsen, Abbas Sharifi, MahtaJafarianFard, and NastaranSoleimani. (2023). "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA." *Int. J. Neurosci.* 133, no. 1: 55-66. <https://doi.org/10.1080/00207454.2021.1883602> PMID:33517817
- Liu, Bo, Changfu Liu, Yang Zhou, Daohai Wang, and Yichao Dun. (2023). "An unsupervised chatter detection method based on AE and merging GMM and K-means." *Mech. Syst. Signal Process.* 186: 109861. <https://doi.org/10.1016/j.ymsp.2022.109861>
- Boussaada, Rim, AbdelazizHakimi, and MajdiKarmani. (2023). "Non-performing loans and bank performance: what role does corporate social responsibility play? A system GMM analysis for European banks." *J. Appl. Account. Res.* <https://doi.org/10.1108/JAAR-10-2021-0283>
- Kurani, Akshit, PavanDoshi, AaryaVakharia, and Manan Shah. (2023). "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting." *Ann. Data Sci.*, 10(1): 183-208. <https://doi.org/10.1007/s40745-021-00344-x>
- Valero-Carreras, Daniel, Javier Alcaraz, and Mercedes Landete. (2023). "Comparing two SVM models through different metrics based on the confusion matrix." *Comput. Oper. Res.* 152: 106131. <https://doi.org/10.1016/j.cor.2022.106131>
- Belghit, Amar, MouradLazri, FethiOuallouche, KarimLabadi, and SoltaneAmeur. (2023). "Optimization of One versus All-SVM using AdaBoost algorithm for rainfall classification and estimation from multispectral MSG data." *Adv. Space Res.*, 71(1): 946-963. <https://doi.org/10.1016/j.asr.2022.08.075>
- Arivazhagan, Naveen, and Tielman T. Van Vl (2023). "Natural language processing basi *Cli. J. Am. Soc.Nephrol.* 18(3): 400- <https://doi.org/10.2215/CJN.0000000000081> PMID:36763809
- Huang, Anzhong, RuiXu, Yu Chen, and MeiwenC (2023). "Research on multi-label 1 classification of social media based on 1 KNN algorithm." *Technol. Forecast. Soc. Chu* 188: 122. <https://doi.org/10.1016/j.techfore.2022.171>
- Lahmiri, Salim. (2023). "Integrating convolutic neural networks, kNN, and Bayes optimization for efficient diagnosis Alzheimer's disease in magnetic resona images." *Biomed. Signal Process. Control.* 104375. <https://doi.org/10.1016/j.bspc.2022.104375>
- Tariq, Aqil, Yan Jiango, Qingting Li, JianweiC Linlin Lu, WalidSoufan, Khalid F. Almut and Muhammad Habib-ur-Rahman. (2023). "Modelling, mapping and monitoring of fo cover changes, using support vector mach kernel logistic regression and naive bayes models with optical remote sensing da *Heliyon* 9: 2122. <https://doi.org/10.1016/j.heliyon.2023.e12122> PMID:36785833 PMCID:PMC9918775
- Kozyrkov C. (2018) Explaining supervised learnin a kid (or your boss), Towards Data Science <https://towardsdatascience.com/explaini-supervised-learning-to-a-kid-c2236f423e0f>
- Croxatto A, Prod'hom G, Faverjon F, et al. (2015) Laboratory automation in clin bacteriology: what system to choose? *Clin Microbiol Infect.* 22(3):217-235 PMID:26806 <https://doi.org/10.1016/j.cmi.2015.09.030>
- Huang YA, You ZH, Chen X, et al. (2017). Prediction of microbe-disease association from the integration of neighbor and graph with collaborative recommendation model. *Transl. Med.* 15:209. doi: 10.1186/s12967-017-1304-7 <https://doi.org/10.1186/s12967-017-1304-7> PMID:29037244 PMCID:PMC5644104

- Zou Q, Lin G, Jiang X, et al. (2018). Sequence clustering in bioinformatics: an empirical study. *Bioinform.* bby090. <https://doi.org/10.1093/bib/bby090> PMID:30239587 PMCid:PMC6954402
- DiMucci D, Kon M, Segre D. (2018). Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem networks. *Msystems* <https://doi.org/10.1128/mSystems.00181-18> PMID:30417106 PMCid:PMC6208640
- Smith K, Wang A, Durant T, et al. (2020) Applications of Artificial Intelligence in Clinical Microbiology Diagnostic Testing, *Clin. Microbiol. Newsl.*, 42 (8): 61-70 <https://doi.org/10.1016/j.clinmicnews.20.03.006>
- Smith K, Kang A, Kirby J. (2017) Automated Interpretation of Blood Culture Gram Stains using a Deep Convolutional Neural Network. *J.Clin. Microbiol.* <https://doi.org/10.1128/JCM.01521-17> PMID:29187563 PMCid:PMC5824030
- Sandle T. (2018) Automated, Digital Colony Counting: Qualification and Data Integrity, *J. GxP Compliance*, 22 (2): <https://www.ivtnetwork.com/article/automated-digital-colony-counting-qualification-and-data-integrity-0>
- Wewerka, Judith, and Manfred Reichert. (2023). "Robotic process automation-a systematic mapping study and classification framework." *Enterp. Inf. Syst.* 17(2): 1986862. <https://doi.org/10.1080/17517575.2021.1986862>
- Wolf, Ádám, Stefan Romeder-Finger, KárolySzéll, and PéterGalambos. (2023). "Towards Robotic Laboratory Automation Plug & Play: Survey and Concept Proposal on Teaching-free Robot Integration with the LAPP Digital Twin." *SLAS Technol.* <https://doi.org/10.1016/j.slasd.2023.12.004> PMID:38101573
- Yin, Yue, PaiZheng, Chengxi Li, and Lihui Wang. (2023). "A state-of-the-art survey on Augmented Reality-assisted Digital Twin for futuristic human-centric industry transformation." *Robot. Comput.-Integr. Manuf.* 81: 102515. <https://doi.org/10.1016/j.rcim.2022.102515>
- Najmi, Ali Hassan, WaleedSalimAlhalafawy, and MarwaZakiTawfiqZaki. (2023). "Developing a Sustainable Environment based on Augmented Reality to Educate Adolescents about the Dangers of Electronic Gaming Addiction." *Sustainability* 15(4): 3185. <https://doi.org/10.3390/su15043185>
- Kim, Jung-Hwan, Minjeong Kim, Minjung Park, and JungminYoo. (2023). "Immersive interactive technologies and virtual shopping experiences: Differences in consumer perceptions between augmented reality (AR) and virtual reality (VR)." *Telemat.* 77: 101936. <https://doi.org/10.1016/j.tele.2022.101936>
- Nincarean, Danakorn, Mohamad Bilal Alia, Noor Dayana Abdul Halim, and MohdHishamuddin Abdul Rahman. (2013). "Mobile augmented reality: The potential for education." *Procedia-soc. Behave. Sci.* 103: 657-664. <https://doi.org/10.1016/j.sbspro.2013.10.385>