The Physician
Journal of The British Association of Physicians of Indian Origin

Introduction to Artificial Intelligence Prediction for Healthcare

An Example of a *Low-Code No-Code* AI Diagnostic *Assistant*

Anindya Chakravorty BTech, MBA
PhD scholar, School of Management, UPES Dehradun, India

theanindyachakravorty@gmail.com

Cite as: Chakravorty, A. (2023) Introduction to artificial intelligence prediction for healthcare. The Physician Vol 8; Issue 2: 1-15 DOI 10.38192/1.8.2.7

Article Information
Submitted 22 Apr 23
Revised 29 Jun 23
Published 22 Jul 23

INTRODUCTION:

Artificial intelligence (AI) aims to mimic human cognitive functions. It is bringing a paradigm shift to healthcare, powered by the increasing availability of healthcare data and rapid progress of analytics techniques.[1] AI can be applied to various types of healthcare data (structured and unstructured). Popular AI techniques include machine learning methods for structured data, such as the classical support vector machine and neural network, modern deep learning, and natural language processing for unstructured data. Major disease areas that use AI tools include cancer, neurology and cardiology.

In recent months, we see an increasing interest in AI and its potential impact on every aspect of human endeavour. In medical diagnostics, there are many AI tools in use, including drug discovery and development, drug repurposing, improving pharmaceutical productivity, and clinical trials, among others; such use reduces the human workload and achieves targets quickly.[2] In protein structure discovery, Deep Minds AI model *AlphaFold* holds a special place in the annals of beneficial outcomes of AI where what used to take professionals months and sometimes years to predict the protein structure now AI does in mins.[3] The neural network-based model, *AlphaFold*, demonstrated accuracy competitive with experimental structures in most cases and significantly outperformed other methods. *AlphaFold* used a novel machine learning approach that incorporates physical and biological knowledge about protein structure, leveraging multi-sequence alignments, into algorithm. Alphabet, the company that owns Deep Minds, has released these algorithms and millions of complex protein structures for free as it can help detect the virus causing the Covid-19 pandemic (SARS-COV-2) behaviour as well.[4]

Doctors need accurate predictions for the outcomes of their patient's diseases. In addition, for accurate predictions, timing is another significant factor influencing treatment decisions. Machine learning aims to mimic the intellectual abilities of humans with machines. Representation and generalisation are used in machine learning. Representations of data instances and functions evaluated on these instances are part of all machine learning methods. Generalisation is the property according to which a machine learning method can provide predictions for previously unobserved data instances. Both supervised and unsupervised methods are used in machine learning.[5]

Supervised Versus Unsupervised Learning

Theoretically, both learning methods vary in structure. In supervised learning, the model describes the outcome that one set of observations (inputs) has on other observations (outputs). Thus, the inputs and outputs that include mediating variables are at opposite ends of the causal chain. Nevertheless, in unsupervised learning, all observations are assumed to be caused by latent variables, which are presumed to be at the end of the causal chain. This approach leaves the probability of the inputs undefined. However, if the inputs are modelled, the missing inputs cause no difficulty because they can be deemed latent variables, as in unsupervised learning.

Natural Language Processing

There is a large volume of natural language text in the connected world, though having a significant content of knowledge. Still, it is becoming increasingly challenging to disseminate it by a human to discover the knowledge or wisdom in it, specifically within any given time limit. This information is usually stored in unstructured and non-standardized formats in electronic healthcare systems, which makes it difficult for the plans to understand the information contents of the narrative information. Thus, NLP techniques can capture unstructured healthcare information, analyse its grammatical structure, determine its meaning, and translate the information so that electronic healthcare systems can easily understand it. Consequently, NLP techniques reduce costs as well as improve the quality of healthcare.[6]

Despite considering various predictions of NLP jobs for individuals or groups of individuals, these assessments could not give the complete result. It is pretty tough to get a precise orientation because of the differences or inconsistencies between the scientific estimates and discrepancies in the methodological evaluations.[7]

This paper aims to do the same for the medical community, where no code AI tool can be used for free and will help healthcare professionals save time, train/ test their juniors, and ultimately help them in saving more lives with the help of an untiring third eye which has a very high level of accuracy.

Using OrangeAI/DM tool

OrangeAI

OrangeAI is an advanced, low-code, no-code AI tool that assists healthcare professionals in detecting COVID-19 and pneumonia through chest x-ray analysis. It leverages the power of OrangeDataMining, a powerful data mining and machine learning platform, to analyse medical images and provide accurate diagnostic results.

OrangeAI utilises machine learning algorithms to identify patterns and features in chest X-ray images that may indicate the presence of COVID-19 or pneumonia. Like how a human mother teaches her child to differentiate between a cat and a dog, the OrangeAI tool is trained on large datasets of labelled chest X-ray images to recognise specific features associated with these diseases. The tool then uses this knowledge to analyse new images and provide accurate diagnostic results.

Advantages of Using OrangeAI in Healthcare

One of the most significant benefits of using OrangeAI is its ability to speed up the diagnostic process. Traditional diagnostic methods, such as manual image analysis by healthcare professionals, can be time-consuming and prone to human error. OrangeAI offers a fast and efficient alternative, enabling healthcare professionals to make quicker decisions and provide timely treatment to patients. OrangeAI's machine learning algorithms have been trained on extensive datasets, allowing the tool to provide highly accurate diagnostic results. OrangeAI can help healthcare professionals make more informed decisions and provide better patient care by reducing the risk of misdiagnosis and false positives. As the COVID-19 pandemic strains healthcare systems worldwide, the need for scalable diagnostic solutions has become increasingly apparent. OrangeAI's low-code, no-code approach allows healthcare professionals to rapidly implement and scale the tool, making it an ideal solution for managing large patient volumes and reducing the burden on healthcare facilities. OrangeAI's low-code, no-code nature makes it accessible to healthcare professionals with limited technical expertise. This user-friendly approach enables medical professionals to harness the power of AI without requiring extensive knowledge of programming or machine learning algorithms.

Applications of OrangeAI in Healthcare

COVID-19 Detection

The ongoing COVID-19 pandemic has highlighted the need for fast, accurate, and scalable diagnostic tools. OrangeAI's ability to analyse chest X-ray images and detect the presence of COVID-19 makes it a valuable tool in the fight against the virus. Early detection and diagnosis can lead to more effective patient management and reduced transmission rates, ultimately helping to control the spread of the virus.

Pneumonia Detection

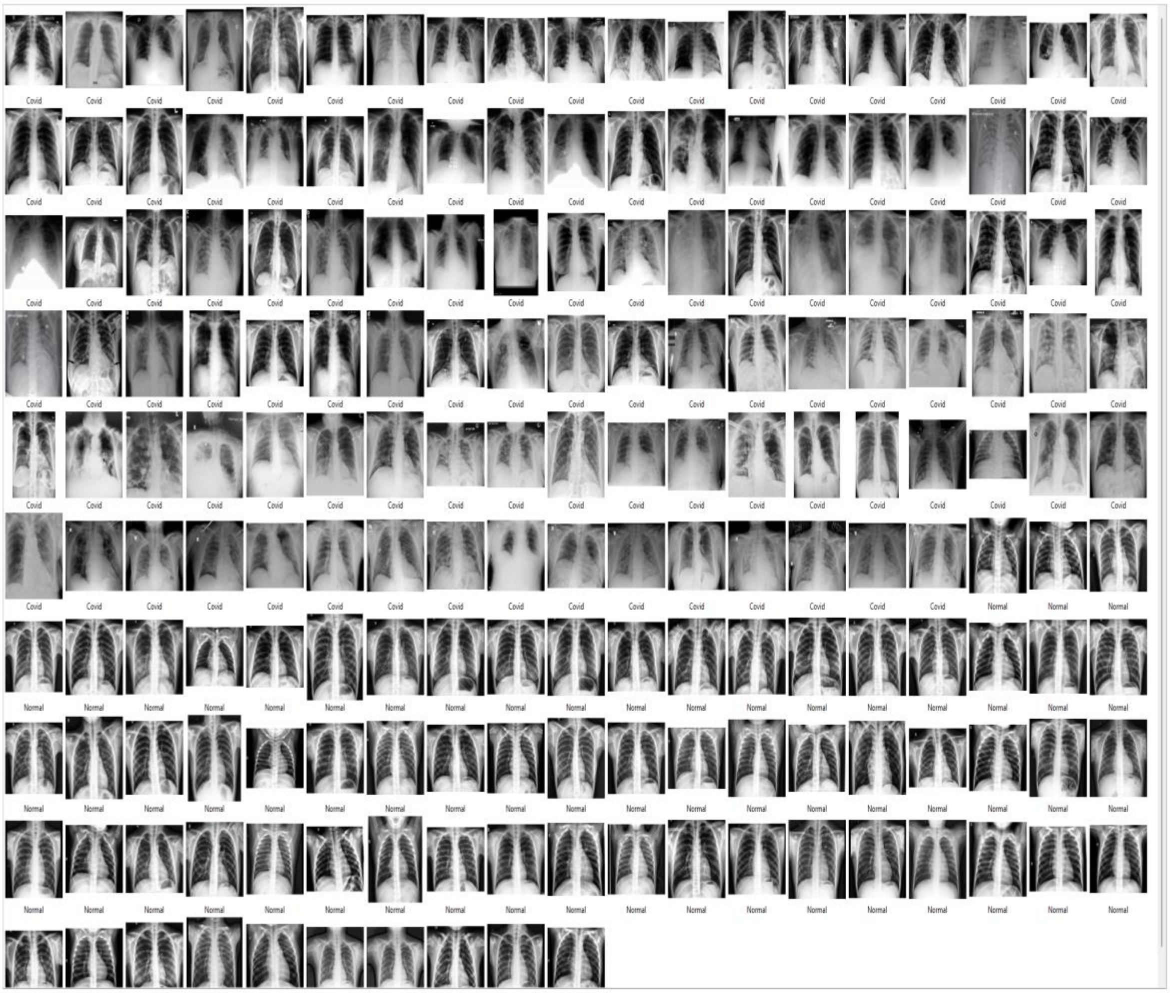
Pneumonia is a common and potentially life-threatening lung infection that requires prompt diagnosis and treatment. OrangeAI's ability to accurately identify pneumonia in chest X-ray images can significantly improve patient outcomes by enabling healthcare professionals to make timely treatment decisions.

Triage and Resource Allocation

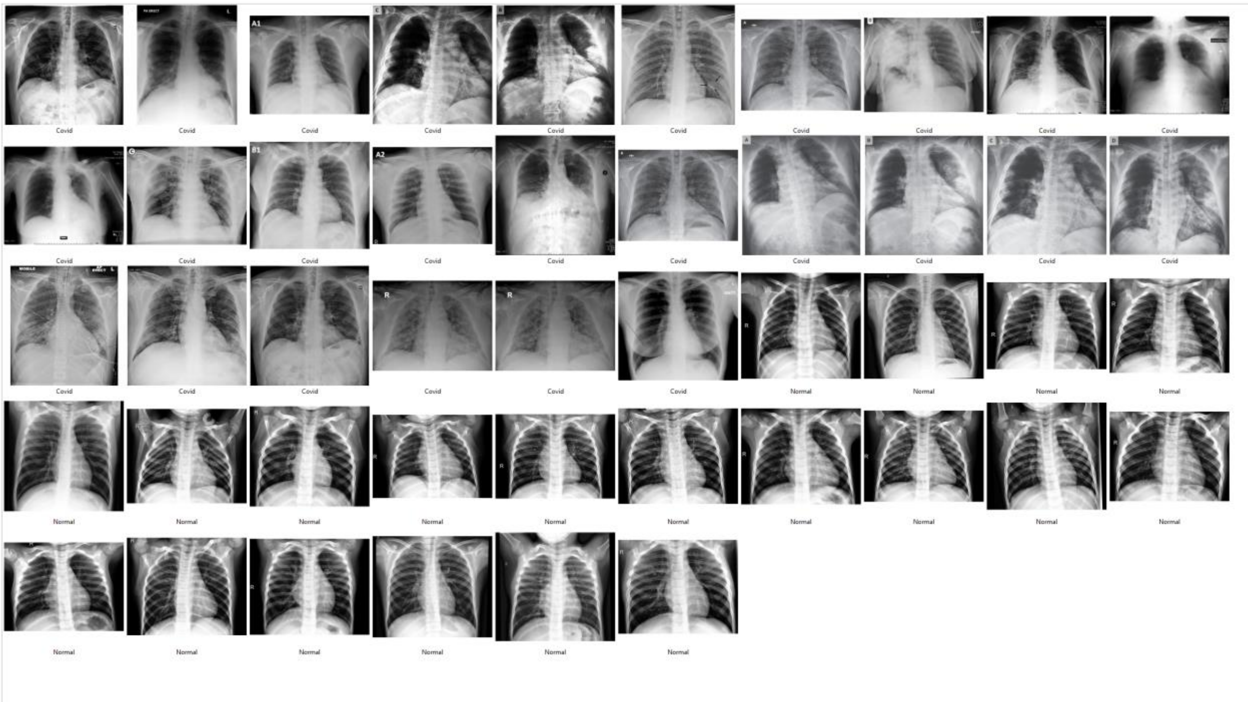
In addition to its diagnostic capabilities, OrangeAI can assist healthcare professionals in triaging patients and allocating resources more effectively. By quickly analysing chest X-ray images and providing accurate diagnostic results, the tool can help healthcare professionals prioritise patients based on their needs and ensure that resources are allocated to those who require them most urgently.

A worked example

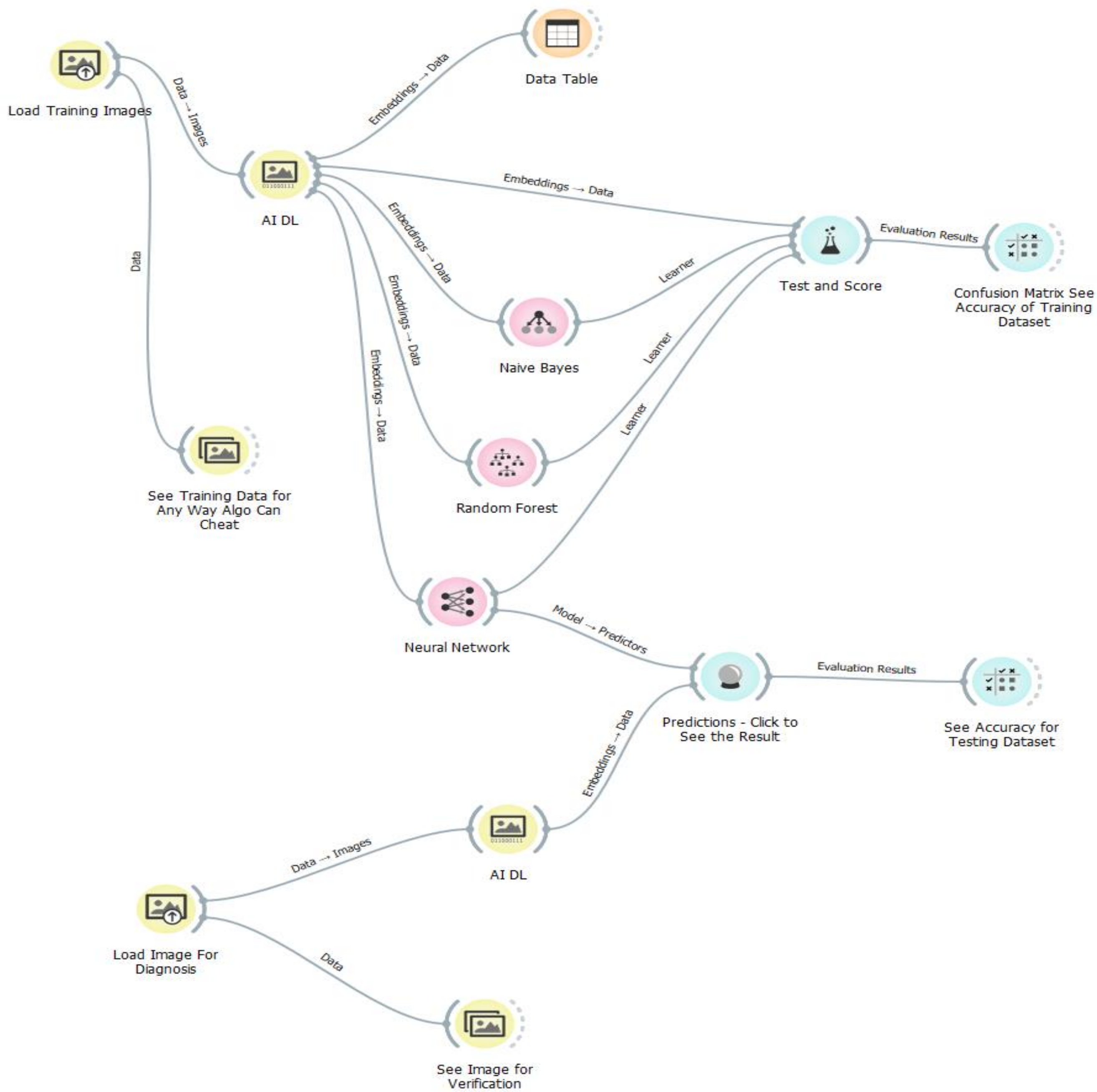
This is how a data set looks like- (Training)



This is how a dataset looks like- (Testing):



This is the AI Tool OrangeAI and what the Algorithm workflow (AI Model) looks like.



Deep Learning breaks the images into hundreds of elements and gives weight to each. The rapid evolution of artificial intelligence (AI) and machine learning (ML) technologies has revolutionised various sectors, including healthcare. One such development is the OrangeAI low code no code tool, which is helping healthcare professionals speed up the diagnostic process, particularly in detecting COVID-19 and pneumonia through chest x-ray analysis. This article explores the potential of OrangeAI as a valuable tool for medical professionals and discusses its various features, benefits, and applications in the healthcare sector.

Challenges and Limitations of OrangeAI

As with any AI tool that processes sensitive medical data, data privacy and security are critical concerns when using OrangeAI. Ensuring that patient data is stored and processed securely and complies with relevant regulations is essential to maintaining patient trust and safeguarding sensitive information. Machine learning algorithms, including those used by OrangeAI, are only as accurate as the data they are trained on. The tool's diagnostic accuracy may be compromised if the training data is biased or unrepresentative. Ensuring that OrangeAI is trained on diverse and representative datasets is crucial to mitigating algorithm bias and ensuring accurate diagnostic results. Integrating OrangeAI with existing healthcare systems and workflows can be challenging, particularly in facilities with limited technical resources.

Seamless integration allows healthcare professionals to utilise the tool effectively without disrupting existing processes. Despite the promising results, several challenges and limitations are associated with using AI in medical imaging, including

- Heterogeneous study methodologies: Differences in study design, patient selection, and AI model development make comparing and generalising results across studies difficult.
- Risk of bias: Inconsistencies in reporting, potential selection bias, and lack of transparency in AI methodologies may limit the validity and reliability of AI models.
- Black box problem: The complex nature of AI models can make it difficult to understand and explain the rationale behind their decisions, which could impact trust and adoption among healthcare professionals.

Future Developments and Potential of OrangeAI

Expanding Applications & the Scope of AI in Medical Imaging

As AI and machine learning technologies continue to evolve, the potential applications of OrangeAI in healthcare are vast. In addition to COVID-19 and pneumonia detection, the tool could be adapted to detect other diseases and conditions using medical imaging data, further enhancing its value to healthcare professionals. As AI technology advances, it will likely play an increasingly significant role in medical imaging. Potential applications include predicting disease progression, evaluating treatment response, and identifying high-risk patients.

Improving Algorithm Accuracy

Continued research and development in machine learning algorithms will likely lead to even more accurate diagnostic results from tools like OrangeAI. As these algorithms improve, healthcare professionals can expect to benefit from increased diagnostic accuracy and better patient outcomes.

Integration with Telehealth Platforms

Integrating tools like OrangeAI with telehealth platforms could offer significant benefits as telehealth becomes increasingly popular. By providing accurate diagnostic results remotely, healthcare professionals can offer patients faster, more efficient care regardless of location.

In conclusion, OrangeAI is a robust low-code, no-code tool with significant potential for improving the speed and accuracy of the diagnostic process in healthcare. By harnessing the power of AI and machine learning, OrangeAI can help healthcare professionals detect COVID-19 and pneumonia more effectively, ultimately leading to better patient outcomes and more efficient use of healthcare resources. While challenges and limitations do exist, the future of OrangeAI and similar tools in healthcare is promising, with potential applications and improvements likely to enhance their value even further.

Integrating AI into Clinical Practice

To fully realise the potential benefits of AI in medical imaging, it is essential to integrate AI models into clinical workflows seamlessly. This may involve developing user-friendly interfaces, providing adequate training and support for healthcare professionals, and establishing effective communication between AI models and human readers.

OrangeAI low code no code is crucial in helping healthcare professionals speed up the diagnostic process for detecting COVID-19 and pneumonia in chest X-rays and CT scans. By providing an accessible and efficient platform for developing AI models, OrangeAI enables healthcare professionals to harness the power of AI for improved patient care. As AI technology advances, it is vital to address the challenges and limitations associated with its use in medical imaging to ensure its safe and effective integration into clinical practice.

Prospects of AI in Medical Imaging – Future Directions?

Expanding the Scope of AI in Medical Imaging

As AI technology advances, it will likely play an increasingly significant role in medical imaging. Potential applications include predicting disease progression, evaluating treatment response, and identifying high-risk patients. Many more diseases like cataracts, breast cancer, acne and others are places where AI can be helpful.

CONCLUSION

When we have a hospital dataset of COVID-19 Pneumonia and routine chest X-rays, we can train an AI model to understand and classify them. Here, one needs to understand what the Deep Learning algorithm is doing to find the pattern and how it understands what COVID-19 and pneumonia are, and what normal human beings will not know how exactly it is doing. There are three common AI problems

- Lazy algorithm

The lazy algorithm is one such phenomenon where the dataset has been taken from varied sources; in artificial intelligence (AI), a lazy algorithm refers to an approach where most computation is deferred until the algorithm is required to make a prediction or provide an output. Unlike eager algorithms that pre-process and generate a model or summary of the data upfront, lazy algorithms postpone most of their computation until the last moment. This characteristic allows lazy algorithms to be more flexible and adaptable to changing data but can also lead to increased computational costs.

- Coded Bias

When the dataset contains only one species/race, coded bias in AI refers to the inherent biases and discriminatory outcomes that can emerge from developing and implementing artificial intelligence systems. While humans create AI algorithms, they can inadvertently reflect and perpetuate biases present in the data used to train them or in the design choices made during their development. This phenomenon raises significant concerns about fairness, accountability, and the potential for reinforcing societal inequalities.

One aspect of coded bias in AI is the reliance on biased datasets. AI algorithms learn from historical data, which can often contain tendencies present in society. For example, suppose historical data used to train a hiring algorithm primarily consists of resumes from male applicants. In that case, the algorithm may learn to favour male candidates, unintentionally perpetuating gender bias in hiring practices. Similarly, biases related to race, age, or other protected attributes can also be encoded into AI systems if not carefully addressed.

The impact of coded bias becomes particularly concerning when AI systems are deployed in critical domains such as criminal justice, healthcare, or lending. Biased AI algorithms can lead to unfair outcomes, perpetuating discrimination against marginalised groups. For instance, biased predictive policing algorithms may disproportionately target minority communities, leading to over-policing and unjust treatment. Biased healthcare algorithms may result in differential access to medical treatments or misdiagnosis for specific demographics.

Coded bias can also emerge from the design choices made during the development of AI systems. For example, if a facial recognition system is predominantly trained on lighter-skinned individuals, it may struggle to accurately recognise or misidentify darker-skinned individuals, perpetuating racial bias. Similarly, language processing algorithms may exhibit biases in their interpretation or translation of specific languages or dialects, reinforcing cultural or linguistic prejudices.

Addressing coded bias in AI requires proactive measures. One approach is to ensure diverse and inclusive representation during the development and decision-making processes. By involving individuals from different backgrounds and perspectives, biases can be more effectively identified and mitigated. Additionally, data collection practices must be critically examined to ensure that training datasets are diverse, representative, and carefully curated to minimise bias.

- Cheating algorithms

When data has been taken and used from particular sources with metadata giving away the secrets. Cheating algorithms, also known as cheating detection algorithms or cheating detection systems, refer to computational methods designed to identify instances of cheating or fraudulent behaviour in various contexts. These algorithms aim to detect and prevent dishonest practices in academic, gaming, or online platforms.

In academia, cheating algorithms are used to identify instances of plagiarism, where students may submit work that is not their own or that contains substantial portions of copied content without proper attribution. These algorithms compare submitted assignments or papers against a database of existing sources to detect similarities and potential instances of plagiarism.

Challenge definition

How can complex Image analytics CT scan reports be read by AI by healthcare professionals so that with no outside help and maintaining medical data privacy rules which are divergent and strict, they can still use the wonder of AI?

Limitations

A larger dataset is needed—an actual user dataset of the required community. Remove algo cheating opportunities. It is ensuring transparency and reproducibility in AI model development and evaluation. In conclusion, the utilisation of the OrangeAI tool for COVID-19 versus pneumonia prediction has the potential to save the time and effort of healthcare professionals significantly. The COVID-19 pandemic has placed an immense burden on healthcare systems worldwide, with an overwhelming number of patients requiring prompt and accurate diagnosis. Differentiating between COVID-19 and pneumonia early is crucial for effective treatment planning and resource allocation. OrangeAI, as a powerful machine learning tool, offers a promising solution to expedite the diagnostic process and improve patient outcomes.

The findings presented in this paper demonstrate that the OrangeAI tool exhibits a high level of accuracy and efficiency in distinguishing between COVID-19 and pneumonia cases. By training the model on a large dataset of clinical data and imaging studies, OrangeAI can recognise patterns and features that might not be immediately apparent to human observers. This enables it to provide reliable predictions and support healthcare professionals in making informed decisions.

One of the critical advantages of OrangeAI is its ability to rapidly analyse large volumes of data, allowing for quick and efficient processing. In the context of COVID-19, where time is of the essence, this tool can significantly reduce the workload on healthcare professionals. By automating the initial screening and triage process, OrangeAI can swiftly identify urgent cases, streamlining the workflow and ensuring that critical patients receive immediate care.

Moreover, OrangeAI can be seamlessly integrated into existing healthcare systems, making it accessible to many medical professionals. Its user-friendly interface and intuitive design enable healthcare practitioners to interact with the tool and interpret the predictions easily. This eliminates the need for extensive training or expertise in machine learning, making it a practical solution for healthcare facilities of varying sizes and resource availability.

The implementation of OrangeAI as a diagnostic aid has the potential to enhance resource allocation within healthcare systems. By accurately identifying COVID-19 cases, healthcare professionals can proactively allocate the necessary resources, such as isolation facilities, ventilators, and personal protective equipment (PPE). This optimisation of resources improves patient care and reduces the strain on healthcare infrastructure during times of crisis.

While OrangeAI shows great promise, it is essential to acknowledge its limitations. The tool heavily relies on the quality and representativeness of the training data. Therefore, continuous updates and refinements to the dataset are crucial to maintaining its accuracy in an ever-evolving healthcare landscape. Additionally, OrangeAI should be viewed as a diagnostic aid rather than a substitute for clinical judgment. The predictions provided by the tool should always be interpreted in conjunction with other clinical information and expertise.

In conclusion, implementing the OrangeAI tool for COVID-19 versus pneumonia prediction offers a valuable solution to expedite the diagnostic process and save time for healthcare professionals. By leveraging the power of machine learning and data analysis, OrangeAI can accurately differentiate between these two conditions, allowing for prompt decision-making and resource allocation. However, ongoing research and development efforts are necessary to enhance the tool's performance and ensure its integration within existing healthcare systems. With the potential to revolutionise the diagnostic process, OrangeAI represents a significant advancement in the fight against COVID-19 and pneumonia.

References

1. Jiang, F. *et al.* Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* **2**, 230–243 (2017).

2. Paul, D. *et al.* Artificial intelligence in drug discovery and development. *Drug Discov Today* **26**, 80–93 (2021).
3. Highly accurate protein structure prediction with AlphaFold | Nature. https://www.nature.com/articles/s41586-021-03819-2.
4. Higgins, M. K. Can We AlphaFold Our Way Out of the Next Pandemic? *Journal of Molecular Biology* **433**, 167093 (2021).
5. Alanazi, H. O., Abdullah, A. H. & Qureshi, K. N. A Critical Review for Developing Accurate and Dynamic Predictive Models Using Machine Learning Methods in Medicine and Health Care. *J Med Syst* **41**, 69 (2017).
6. Chowdhary, K. R. Natural Language Processing. in *Fundamentals of Artificial Intelligence* (ed. Chowdhary, K. R.) 603–649 (Springer India, 2020). doi:10.1007/978-81-322-3972-7_19.
7. Roy, K. *et al.* Application of Natural Language Processing in Healthcare. in *Computational Intelligence and Healthcare Informatics* 393–407 (John Wiley & Sons, Ltd, 2021). doi:10.1002/9781119818717.ch21.

Resources:

There are many datasets of Covid19 & normal and pneumonia on medical data-sharing websites, hospitals, and GitHub forums. The best ones are in GitHub. They adhere to the norms of keeping the training set diverse and from the same X-ray centres.

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8594127/
- https://imgcreator.zmo.ai/ai-generator
- Joseph Paul Cohen PhD, Director Institute of Reproducible Research, AWS Heath AI, Covid dataset, GitHub, Butterfly Networks
- The RadImageNet Database
- COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University: This repository provides comprehensive global data on COVID-19 cases, including confirmed, recovered, and deceased cases. (Website: https://github.com/CSSEGISandData/COVID-19)
- Kaggle COVID-19 Open Research Dataset (CORD-19): Kaggle, a popular data science community, provides a dataset that comprises thousands of scholarly articles related to COVID-19, SARS-CoV-2, and related coronaviruses. (Website: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge)
- COVID-19 Image Data Collection: This dataset contains chest X-ray and CT images of COVID-19 cases, pneumonia cases, and normal cases. It can be useful for developing machine learning models for image-based diagnosis. (Website: https://github.com/ieee8023/covid-chestxray-dataset)
- Open COVID-19 Dataset by Google Cloud: This dataset includes various COVID-19-related data, such as cases, deaths, and testing, from multiple sources worldwide. (https://console.cloud.google.com/marketplace/product/bigquery-public-datasets/covid19-dataset-list)
- COVID-19 Data from the World Health Organization (WHO): WHO provides a comprehensive dataset containing COVID-19 cases, deaths, testing, hospitalizations, and other related information from various countries. (Website: https://covid19.who.int/)
- MIMIC-CXR: This dataset consists of chest X-ray images from over 65,000 patients, including cases of pneumonia, normal cases, and various other conditions. It is available through PhysioNet, a research resource supported by the National Institutes of Health (NIH). (https://physionet.org/content/mimic-cxr/)
- COVID-19 Data Hub: The COVID-19 Data Hub offers a wide range of COVID-19-related datasets from various sources, including global cases, vaccination data, and more. (Website: https://covid19datahub.io/)
- Radiological Society of North America (RSNA) COVID-19 Imaging Data Repository: RSNA provides a dataset of chest X-ray and CT images of COVID-19 cases, pneumonia cases, and normal cases. It is intended to facilitate the development of AI models for diagnosing COVID-

19. (Website: https://www.rsna.org/COVID-19/COVID-19-Imaging-Data-Repository)

- UCI Machine Learning Repository: The UCI repository hosts various datasets related to health and medical domains, including pneumonia and respiratory-related data that can be used for analysis and model development. (https://archive.ics.uci.edu/ml/index.php)
- COVID-19 Data Lake: The COVID-19 Data Lake is a comprehensive collection of datasets related to COVID-19, including cases, vaccination data, mobility data, and more. It provides a centralized repository for accessing various COVID-19 datasets. (Website: https://covid19datalake.org/)
- IH Chest X-ray Dataset: This dataset contains chest X-ray images, including cases of pneumonia, from the National Institutes of Health (NIH) Clinical Center. It is a widely used dataset for pneumonia classification tasks. (Website: https://nihcc.app.box.com/v/ChestXray-NIHCC)
- ChestX-ray8: The ChestX-ray8 dataset consists of 108,948 frontal-view chest X-ray images from 32,717 patients, including cases of pneumonia. It can be used for various medical imaging tasks, including pneumonia detection. (Website: https://www.kaggle.com/nih-chest-xrays/data)
- RSNA Pneumonia Detection Challenge: This dataset is part of a Kaggle challenge and contains chest X-ray images labeled for pneumonia detection. It includes both bacterial and viral pneumonia cases. (Website: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data)
- MIMIC-CXR: As mentioned earlier, this dataset from PhysioNet includes chest X-ray images from over 65,000 patients, including cases of pneumonia, normal cases, and other conditions. (Website: https://physionet.org/content/mimic-cxr/)
- Montgomery County X-ray Set: This dataset consists of chest X-ray images collected from the Department of Radiology, Montgomery County, USA. It contains 138 images, including cases of normal and tuberculosis pneumonia. (Website: https://ceb.nlm.nih.gov/repositories/montgomery-county-xray-set/)
- Shenzhen Hospital Dataset: This dataset includes chest X-ray images obtained from the Shenzhen Hospital in China. It comprises 662 images, with 326 images labeled as pneumonia and 336 as normal. (Website: https://ceb.nlm.nih.gov/repositories/tuberculosis-chest-x-ray-image-data-sets/)
- Chest Radiograph Dataset of Pneumonia: This dataset contains 5,856 chest radiographs, including cases of pneumonia, collected from various sources. It is commonly used for training and evaluating pneumonia detection models. (Website: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia)
- Mendeley Pneumonia Dataset: This dataset provides chest X-ray images of pneumonia cases obtained from Mendeley Data. It consists of 3,500 images, including bacterial, viral, and COVID-19 pneumonia cases. (Website: https://data.mendeley.com/datasets/rscbjbr9sj/2)
- Guangzhou Women and Children's Medical Center Dataset: This dataset contains chest X-ray images of pediatric pneumonia cases collected from the Guangzhou Women and Children's Medical Center in China. (Website: https://github.com/ieee8023/covid-chestxray-dataset)
- Stanford CheXpert: The CheXpert dataset includes chest radiographs labeled for various pathologies, including pneumonia. It consists of over 200,000 images from more than 65,000 patients. (Website: https://stanfordmlgroup.github.io/competitions/chexpert/)
- RSNA Pneumonia Detection Challenge Stage 2: This dataset is an extension of the RSNA Pneumonia Detection Challenge dataset, containing additional chest X-ray images labeled for pneumonia. It was used for the second stage of the challenge. (Website: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data)
- Pneumonia Detection on Chest X-rays Dataset: This dataset, available on Kaggle, contains chest X-ray images labelled for pneumonia. It includes images from pediatric patients,