

Facilitating AI/ML Translation in Healthcare



Dr. Sagnik Dakshit Assistant Professor, Computer Science

Al/ML/Data Science Consultant. OSP Research Design & Data Analysis Lab

The University of Texas at Tyler



01 Envision the Future

02 Challenges

03 Examples

04 LLMs and VR



https://www.aidoc.com/learn/blog/ai-in-healthcare/



https://onlinelibrary.wiley.com/doi/10.1155/2022/8421434





https://www.researchgate.net/figure/oT-devices-in-healthcare_fig1_340371190

https://onlinelibrary.wiley.com/doi/full/10.1002/eom2.12448





Digital Twin Technology Relies on Data









Components of AloT

- Devices and Sensors
- Gateway Services
- Cloud Services
- Communication Infrastructure



https://www.researchgate.net/figure/oT-devices-in-healthcare_fig1_340371190



https://gloriumtech.com/10-benefits-of-ai-in-healthcare/



https://arxiv.org/pdf/2202.02868







Challenges in Development





Challenges in Adoption



- Data security
- Privacy crucial

Ethical and					
Legal					
Concerns					

- Ethical dilemmas
- Inequitable care.
- Liability and accountability

Resistance to Change

- Job displacement
- Lack of trust in technology
- Complexity of integrating



Challenges in Deployment

	 High computational power
Scalability	 Secure data storage Reliable network access
Compliance with Regulations	 FDA (U.S.) and EMA (Europe) require evidence of safety and efficacy for Al-driven applications
Monitoring and Maintenan <u>ce</u>	 Evolving knowledge, protocols Continuous retraining



Challenges in Implementation

Integration with Existing Systems

Legacy Systems

Interoperability

- Diverse and fragmented data sources
- EHR systems, medical devices and imaging platforms

Training and Usability

- Training to use AI tools
- Lack of user-friendly interfaces

DATA CHALLENGES



- Insufficient Data Quantity
- Imbalanced Data
- Missing Data
- Lack of Diversity in Data
- Label Noise and Mislabeling
- Data Distribution
- Data Bias
- Data Privacy

CORE-SET / SUBSET SELECTION



- Coreset selection is a method for selecting a small, representative subset of an entire dataset.
- Most informative samples
- Quality of samples

• Dakshit, Sagnik, et al. "Core-set selection using metrics-based explanations (CSUME) for multiclass ECG." 2022 IEEE 10th International Conference on Healthcare Informatics (ICHI). IEEE, 2022.

• Dolatabadi, H.M., Erfani, S.M. & Leckie, C. Adversarial Coreset Selection for Efficient Robust Training. Int J Comput Vis 131, 3307–3331 (2023). https://doi.org/10.1007/s11263-023-01860-4

GENERATIVE AI



- Missing Data
- Synthetic Data
- Quality and Quantity
- Lack of Control

• Dakshit, Sagnik, et al. "Core-set selection using metrics-based explanations (CSUME) for multiclass ECG." 2022 IEEE 10th International Conference on Healthcare Informatics (ICHI). IEEE, 2022.

• Dolatabadi, H.M., Erfani, S.M. & Leckie, C. Adversarial Coreset Selection for Efficient Robust Training. Int J Comput Vis 131, 3307–3331 (2023). https://doi.org/10.1007/s11263-023-01860-4

DISTRIBUTED LEARNING



- Improved Generalization
- Data Heterogeneity
- Data Privacy
- Computational Overheads

BIAS CHALLENGES

DATA BIAS

Training data is not representative of the realworld population or task the model is expected to perform on.

ALGORITHMIC BIAS

Arise from how the algorithm weighs certain features, optimizes objectives, or handles fairness constraints.

MODEL BIAS

Errors or limitations from assumptions of the learning algorithm leading to underfitting.



https://medium.com/@cloudhacks_/federated-learning-a-paradigm-shift-in-data-privacy-and-model-training-a41519c5fd7e



Chen, Xing & Li, Jingtao & Chakrabarti, Chaitali. (2022). Energy and Loss-aware Selective Updating for SplitFed Learning with Energy Harvesting-Powered Devices. Journal of Signal Processing Systems. 94. 1-15. 10.1007/s11265-022-01781-4.



- Explanations vs Interpretability
 - Are the explanations interpretable?

• Evaluating Explanations

- Reliability of Explanations
- Insufficiency of human judgment to validate XAI



XAI

Das, Arun and Paul Rad. "Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey." *ArXiv* abs/2006.11371 (2020): n. pag.

Types of Explanations



Sujatha Ravindran, Akshay & Contreras-Vidal, Jose. (2023). An empirical comparison of deep learning explainability approaches for EEG using simulated ground truth. Scientific Reports. 13. 10.1038/s41598-023-43871-8.

The Challenge in XAI

• XAI methods do not address well-defined problems and are not evaluated against objective criteria of explanation correctness.

• Human Bias: In mammographic data analysis, a radiologist would likely trust a cancer diagnosis made by an AI if told that the decision was based on a patch of tissue they themselves identify as cancerous.

Conversely, if the XAI method assigns high "importance" to features that are known not to be associated with cancer, this might lead to the dismissal of the model itself as being wrong (Saporta et al., 2022).

Facilitating AI / ML

- Experts-in-the-loop
- Formal definitions
 - Problem
 - Explanations
 - TASK
 - DATA MODALITY
 - STAKEHOLDER



Example 1: Explaining Bias

BIAS = = DISCRIMINATION !!

WHAT IS CONSIDERED BIAS-FREE ? 80% or 4/5th rule of Participation

The 80% rule states that if for a protected group, the rate of selection is less than 80% of the group with the largest number of samples, there can be adverse impact.





Dakshit, Sagnik, et al. "Bias Analysis in Healthcare Time Series (BAHT) Decision Support Systems from Meta Data." *Journal of Healthcare Informatics Research* 7.2 (2023): 225-253.











Example 2: Explaining Augmentation

- Augmentation is a common preprocessing step to match the class with the highest number of samples.
 - Data Balancing
 - Increasing Dataset Size

- Magnitude Warping
- Time Warping
- Window Slicing
- Window Warping
- Jittering
- Scaling
- Flipping
- Permutation

Balasubramanian, Nikil Sharan Prabahar, and Sagnik Dakshit. "Explanations of Augmentation Methods for Deep Learning ECG Classification." International Conference on Artificial Intelligence in Medicine. Cham: Springer Nature Switzerland, 2024.



Balasubramanian, Nikil Sharan Prabahar, and Sagnik Dakshit. "Explanations of Augmentation Methods for Deep Learning ECG Classification." *International Conference on Artificial Intelligence in Medicine*. Cham: Springer Nature Switzerland, 2024.

Example 3: Explaining Outlier Filtering Method

- Various outlier filtering methods are used prevalently as pre-processing step
- Traditionally the methods are selected blindly based on performance improvement

OF Method	Training Sample Count	Test Accuracy	Precision	Recall	AUC
LOF	1311	0.9819	0.984	0.996	0.99
One-Class SVM	1311	0.9819	0.9738	0.995	0.99
Elliptical Envelope	1311	0.9819	0.9738	0.995	0.98

Dakshit, Sagnik. "Multiview Outlier Filtered Pediatric Heart Sound Classification." *International Journal of Advanced Computer Science & Applications* 15.5 (2024).





(C) UMAP Visualization before Outlier Filtering (D) UMAP Visualization after Outlier Fil



(A) UMAP Signal Visualization after Outlier Filtering

One Class SVM



Elliptical Envelope



Local Outlier Filtering

Example 4: Reducing FP and FN



Abstaining classifier: Mitigates risk in critical multiclass tasks through explainable prototypical spaces consequently improving reliability.

Dakshit, Sagnik. "Abstaining ECG Classifiers Through Explainable Prototypical Spaces." 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI). IEEE, 2024.



Example 5: Lack of Control in Generation



Dakshit, Sagnik, and Balakrishnan Prabhakaran. "CVAE-based Generator for Variable Length Synthetic ECG." 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI). IEEE, 2023.

• LLMs cannot Reason, Plan

- Inability to plan
- LLMs have their Bias
- LLMs hallucinate





Large Language Model (LLM)

['lärj 'laŋ-gwij 'mä-dəl]

A deep learning algorithm that's equipped to summarize, translate, predict, and generate human-sounding text to convey ideas and concepts.

2 Investopedia

Retrieval Augmented Generation



- RAG reduces hallucination
- RAG improves output formatting
- RAG improves consistency

How many r's are there in strawberry?

There are **2 r**'s in the word "strawberry."

() ቆ ቆ ₽ ℃~

B

2

What happens if I hold a piece of paper horizontally with two hands and I let go of one hand?

\$

If you were holding a piece of paper horizontally with two hands and let go of one hand, the paper would likely tilt or rotate in the direction of the hand that you let go of. This is because the paper is no longer being held evenly on both sides, and the weight distribution is now uneven. If the paper is a light weight, it will drift to the side where you let go of and fall to the ground. If the paper has a little more weight or if you have a strong grip, it will stay in air for a moment before falling down.

• GPT-63% for self-diagnosis

• Gemini 24% for self-diagnosis

Balasubramanian, Nikil Sharan Prabahar, and Sagnik Dakshit. "Can Public LLMs be used for Self-Diagnosis of Medical Conditions?." *arXiv* preprint arXiv:2405.11407 (2024).

Gemini

- User Role : 18 year old male

Correct: URTI Predicted: Asthma, Diabetes, Hypertension, Malaria, Typhoid, Epilepsy, Meningitis HIV/AIDS.

- User Role : 90 year old female

Correct: URTI Predicted: Trigeminal neuralgia, Cluster headache.

GPT - 4.0

- User Role : 18 year old male

Correct: URTI Predicted: Sinusitis, Tension headache, Migraine, Upper respiratory infection

- User Role : 90 year old female

Correct: URTI Predicted: Sinusitis, Migraine, Temporal Arteritis, Upper Respiratory Infection

Balasubramanian, Nikil Sharan Prabahar, and Sagnik Dakshit. "Can Public LLMs be used for Self-Diagnosis of Medical Conditions?." arXiv preprint arXiv:2405.11407 (2024).



AI /ML in Virtual Reality

- Smaller Al models
- Improved Cloud
- Improved Connectivity
- Lower Costs



Thank you

Dr. Sagnik Dakshit sdakshit@uttyler.edu www.sagnikdakshit.com

