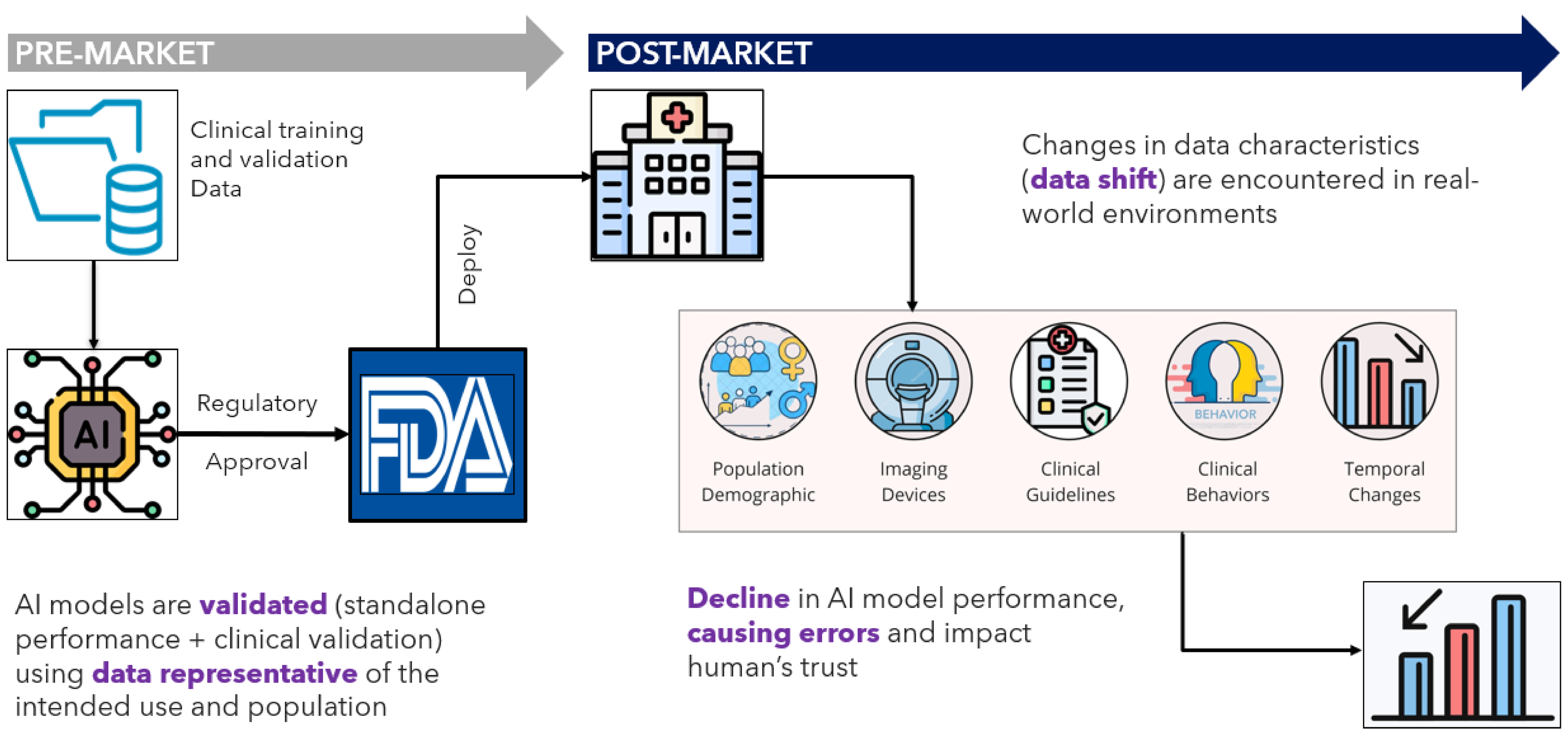
Forecasting Changes in Data Characteristics: A Case Study on Chest Radiograph Images

Elham Konjkav, Ghada Zamzmi, and Jana Delfino

Office of Science and Engineering Laboratories (OSEL), Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration (FDA), Silver Spring, MD, USA

Abstract

We present a framework for forecasting changes in the characteristics of data inputs to artificial intelligence (AI) models, using Chest X-ray (CXR) imaging as the primary modality. Our framework effectively predicts shifts from adult to pediatric CXR, accounting for three types of trends: seasonal, gradual, and incremental changes. Early prediction of such shifts is useful, as data drift—variations in data characteristics—can impact the performance of AI models post-deployment. By anticipating these changes, our framework enables proactive adjustments that help maintain the reliability of AI models in the post-market phase.



Introduction

AI models are typically developed in controlled, static environments; however, real-world clinical settings are highly dynamic. Model performance can degrade during post-market deployment due to shifts in data characteristics, influenced by factors such as changes in patient demographics, scanner technology, clinical guidelines, or practices.

Given the unavoidable reality of post-market changes in data characteristics, there is a critical need to develop methods that can forecast these changes. By accurately predicting data shifts, healthcare providers can implement preventive measures or, at the very least, approach patient care with enhanced vigilance.

Our framework represents a novel application of forecasting methods for forecasting changes in the characteristics of medical imaging data.

Materials and Methods

We evaluated our method using two datasets: the NIH Adult Chest X-ray (CXR) dataset [1] and a pediatric CXR dataset [2]. These datasets were specifically chosen to enable the simulation of changes in data characteristics that occur due to demographic shifts from adult to pediatric populations.

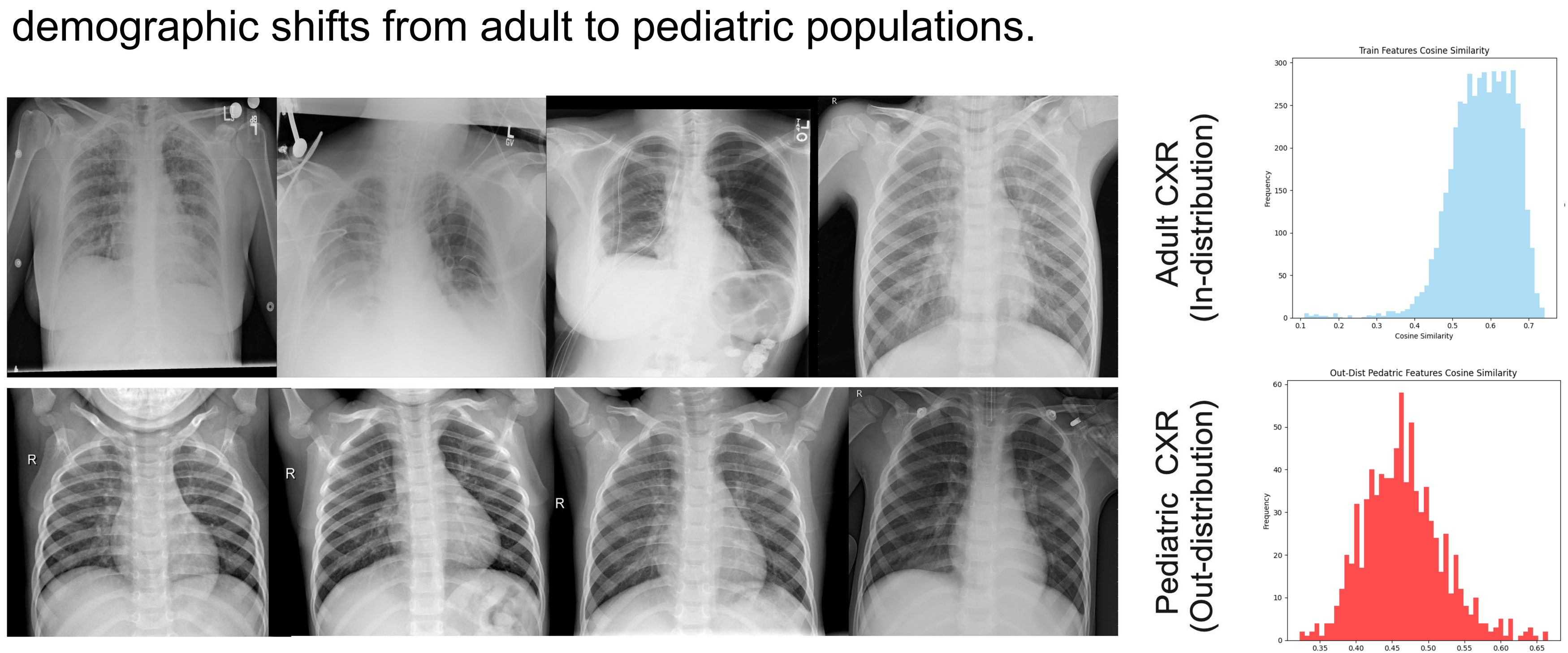
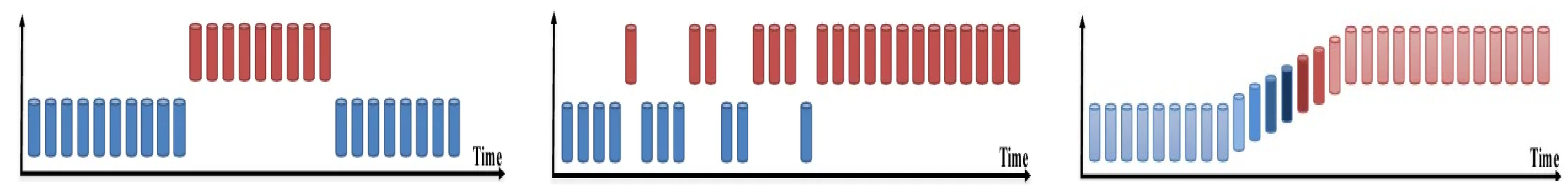


Image to Drift Metric: To predict changes in image characteristics, we first need to convert image data into quantifiable features for prediction. We extracted deep learning features from the images using a convolutional neural network (VGG16). Once the features are extracted, we use cosine similarity to measure the similarity between the feature vectors of adult CXR and pediatric CXR data. This metric calculates the cosine of the angle between two vectors, which provides an indication of how closely related the feature representations of the two data sets are.

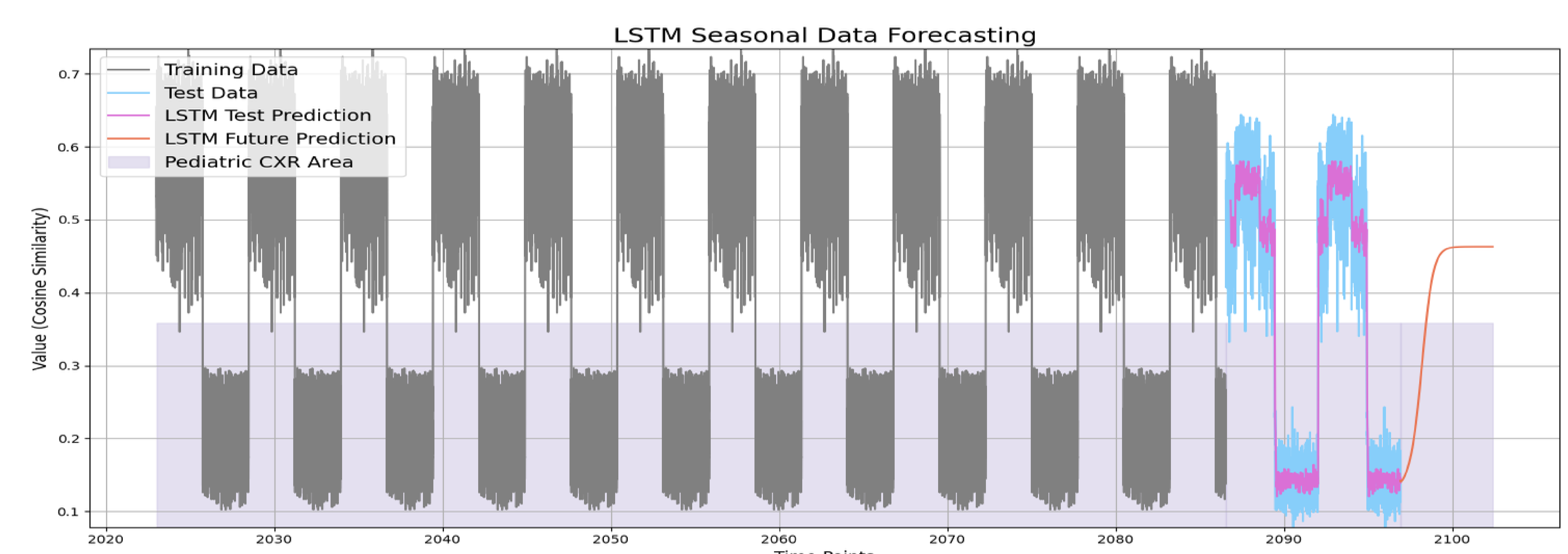
Forecasting Data Trend: The similarity metric was used as input for training forecasting methods such as Long Short-Term Memory (LSTM), SARIMA (Seasonal Auto-Regressive Integrated Moving Average), and Holt-Winters. These models were trained to identify and learn different types of trends, including seasonal, gradual, and incremental.

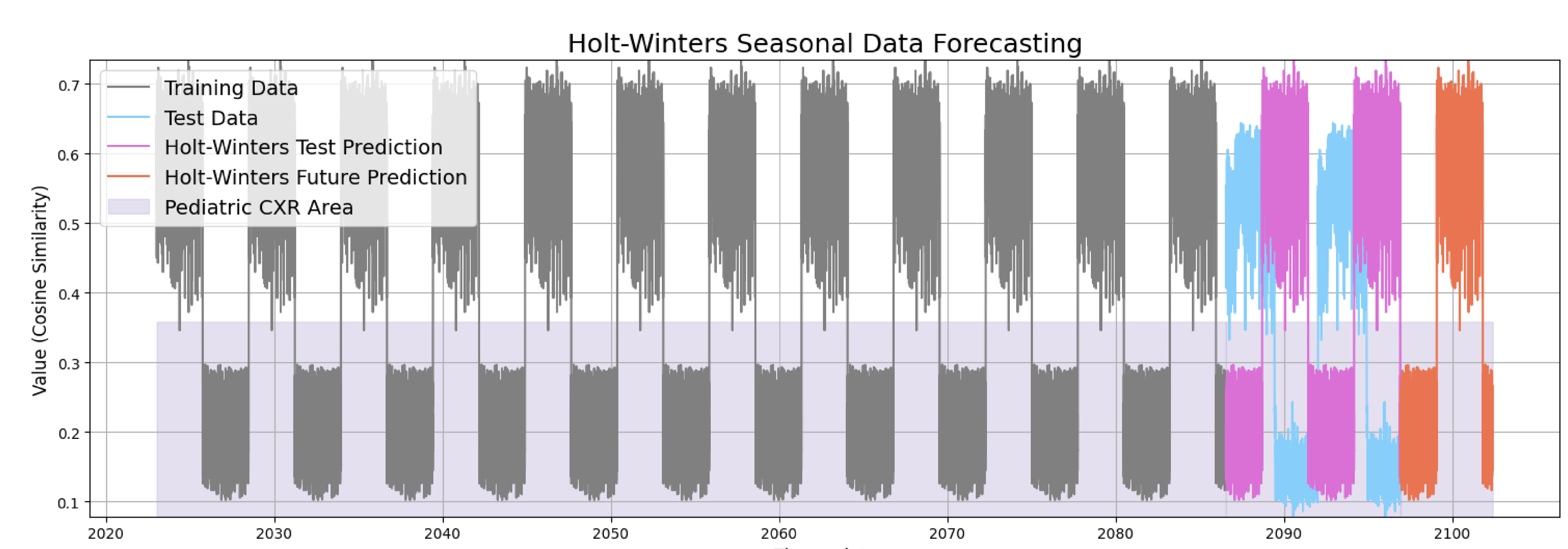


Seasonal Trend — Gradual Trend — Incremental Trend

Results and Discussion

The figures below demonstrate that both the Holt-Winters and LSTM models effectively captured seasonal trends. However, the Holt-Winters model showed a delay in its predictions compared to the LSTM model. The highlighted purple area in the figures indicates that all data within that region are from pediatric CXR, demonstrating that the forecasting models can successfully predict changes in data from adult to pediatric.





Model	Drift Type	MAE	RMSE	MAPE%
LSTM	Seasonal	**0.02885**	**0.04079**	**10.27**
	Incremental	**0.00270**	**0.00445**	**0.87**
	Gradual	**0.01280**	**0.01495**	**7.35**
SARIMA	Seasonal	0.20410	0.27556	126.84
	Incremental	0.14737	0.18704	40.60
	Gradual	0.13554	0.20139	31.56
Holt-Winters	Seasonal	0.33098	0.36065	156.01
	Incremental	0.14934	0.19648	36.31
	Gradual	0.15328	0.19738	46.59

The table shows that SARIMA performed the worst amongst the three models. While LSTM is the overall best model, it struggles to capture seasonality in future predictions, a task the Holt-Winters model accomplishes effectively as shown in the figures.

Conclusion

We propose a framework for forecasting changes in data characteristics, focusing on seasonal, incremental, and gradual drift from adult CXR to pediatric CXR data.

This framework is valuable in clinical practice, as it predicts changes that could affect AI model performance, which might enable proactive measures for cautious deployment. By serving as an alerting system for potential shifts in data characteristics—such as forecasting shifts from adult to pediatric populations—it supports the safer deployment of AI models.

ACKNOWLEDGEMENT: Elham Konjkav acknowledges funding by appointment to the Research Participation Program at the Center for Devices and Radiological Health administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the US Food and Drug Administration.

References:

[1] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2097-2106). doi:10.1109/CVPR.2017.36

[2] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22), 2402-2410. doi:10.1001/jama.2016.17216