FDA / CDRH ORISE Staff Fellow Research Position: Performance metrics for Multi-Class Decision Support Systems

Application Deadline: **09/30/2022**

The U.S. Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) has a research position available through The Oak Ridge Institute for Science and Education (ORISE) Research Participation Program. Eligibility includes PhD candidates in Mathematics and Statistics, and in Computer, Information, and Data Sciences.

The ORISE Research Participation Program at the FDA is an educational and training program designed to provide students and recent graduates, opportunities to participate in project-specific research and developmental activities at CDRH. Participants will have an opportunity to gain a hands-on research experience on a variety of regulatory research projects related to CDRH's mission. The program is designed for participants to engage with an expert mentor or mentors to examine a question of interest related to those projects within the placement office. The participant will receive a monthly stipend commensurate with educational level and experience.

Project Title, Description, and Aims are given in the description below.
Interested? Please contact lead investigator Kenny Cha at Kenny.cha@fda.hhs.gov

Qualification: You need to have lived in the United States of 3 of the past 5 years. J1 visas are accepted.
Salary: GS series equivalent commensurate with education level and experience
Benefits: Health insurance is included at no cost to employee.
Appointment Length: Research project is funded for up to 2 years.
Questions? Please contact lead investigator Kenny Cha (Kenny.cha@fda.hhs.gov)

For these positions, other positions, and other information, please see the twitter page Aldo Badano (@aldo_badano) / Twitter for Aldo Badano, Director, Division of Imaging Diagnostics and Software Reliability (DIDSR). DIDSR has several research charters, in particular for Medical Devices and Diagnostics and in Artificial Intelligence and Machine Learning (AI/ML)

Project Title: Performance metrics for multi-class decision support systems
Lead Investigator: Kenny Cha, PhD, Kenny.cha@fda.hhs.gov

Project Description: This project will develop methods for evaluation of devices that performs complex classification tasks such as multi-outcome diagnoses. The growth in artificial intelligence (AI) and machine learning (ML) has led to increased development of decision support systems for medical classification and prediction. Binary classification tasks include discriminating between presence/absence of disease, e.g., cancer/non-cancer or malignant/benign. Methods for evaluating the performance of binary classification and prediction are well-established. However, AI/ML-based systems are increasingly being developed for non-binary, polychotomous classification tasks, for which performance evaluation is less understood. For example, a device may classify a skin lesion as a mole, rash, melanoma, or basal cell carcinoma. While the evaluation of these multi-class classification tasks is common in the pattern recognition field – the field for enabling computers to identify and process patterns in data such as images, a consensus is lacking on the study design to determine the safety and effectiveness of devices that perform such tasks. In addition, agreement in the field is lacking regarding which performance metrics are appropriate for medical applications attempting to solve these types of problems. This is important as the device outputs may influence patient diagnosis and treatment, and improper evaluation of these devices can lead to harm to the patients who are affected by these devices, either directly or indirectly.

This research aims to develop principles and methodology for evaluating devices that perform multi-class classification. A team of FDA scientists will study the state of the field, and provide information

and data to help industry, as well as the FDA reviewers, in designing a performance evaluation study to show the safety and the effectiveness of a device performing non-binary classification. Our research will provide guidance on how to select clinically meaningful metrics based on the task being performed, while addressing common issues such as study design, prevalence, and operating point. This project will assist in developing assessment metrics for multi-class decision support systems, helping to bring these devices to the US market in a safe and effective manner, giving access to cutting-edge technologies for patients, and providing least-burdensome approaches for device manufacturers.

Aim 1: Survey analytical performance metrics in multi-class classification including their meaning, underlying assumptions, and computation methods. Review the metrics being used in non-medical disciplines, such as neural network-based classification in computer vision and multi-spectral analysis, and metrics cited in statistical literature for evaluating nominal and ordinal classification variables.

Aim 2: Assess the feasibility of using existing metrics in medical device evaluation, by understanding each metric's strengths, weaknesses, and key properties relevant to various clinical tasks. Study new and underutilized metrics and methods to estimate confidence intervals to quantify sampling variability for tasks where currently existing metrics or methodology are inadequate. Explore how different study designs affects the metrics.

Aim 3: Study the fit the metrics identified from Aim 2 to different diagnostic areas with different study design condensations. Apply the metrics to simulated data and real-world data to evaluate the metric's fit to the given clinical task. Provide recommendations for metrics that should be utilized in demonstrating effectiveness in evaluation of a multi-class classification task. If no single metric is found to be acceptable, provide a list of characteristics that a set of metrics should have to be acceptable for a medical application. The metrics recommended may depend on study design.

Thank you,

Gene Pennello, PhD, *Mathematical Statistician*

U.S. Food and Drug Administration
Center for Devices and Radiological Health
Office of Science & Engineering Laboratories

Division of Imaging, Diagnostics, & Software Reliability
10903 New Hampshire Avenue, White Oak 64, Room 3020
Silver Spring MD 20993-0002 gene.pennello@fda.hhs.gov