In advance of the first 24-hour virtual Women in Data Science (WiDS) Worldwide Conference March 8, 2021, we invite you to build a team, hone your data science skills, and join us for the 4th Annual WiDS Datathon focused on social impact. Register now!

The WiDS Datathon 2021 focuses on patient health, with an emphasis on the chronic condition of diabetes, through data from MIT's GOSSIS (Global Open Source Severity of Illness Score) initiative. Brought to you by the WiDS Worldwide team at Stanford, the West Big Data Innovation Hub, and the WiDS Datathon Committee, this year's datathon is open until March 1, 2021. Winners will be announced at the WiDS Conference via livestream, reaching a community of 100,000+ data enthusiasts across more than 85 countries.

Background

Getting a rapid understanding of the context of a patient's overall health has been particularly important during the COVID-19 pandemic as healthcare workers around the world struggle with hospitals overloaded by patients in critical condition. Intensive Care Units (ICUs) often lack verified medical histories for incoming patients. A patient in distress or a patient who is brought in confused or unresponsive may not be able to provide information about chronic conditions such as heart disease, injuries, or diabetes. Medical records may take days to transfer, especially for a patient from another medical provider or system.

Knowledge about chronic conditions such as diabetes can inform clinical decisions about patient care and ultimately improve patient outcomes. Learn more about the global scale of diabetes on our Datathon News page, and make sure you are subscribed to the WiDS Datathon Mailing List to receive the latest updates, tutorials, and articles.

Overview

This year's challenge will focus on models to determine whether a patient admitted to an ICU has been diagnosed with a particular type of diabetes, Diabetes Mellitus. Using data from the first 24 hours of intensive care, individuals and teams will explore labeled training data for model development. Participants will then upload predictions for unlabeled data to Kaggle and these predictions will be used to determine the public leaderboard rankings, as well as the final leaderboard revealed at the close of the competition.

Data analysis can be completed using your preferred tools. Tutorials, sample code, and other resources will be posted throughout the competition at widsconference.org/datathon-resources and on the Kaggle Discussion Forum.

The WiDS Datathon 2021 dataset is similar to the WiDS Datathon 2020 dataset, but prior experience is not needed. This year's WiDS Datathon will also feature a Phase 2 hosted by the National Science Foundation Big Data Innovation Hubs, to encourage research papers and further collaboration. Special thanks to Kaggle for supporting the suite of WiDS Datathon cash awards this year, totaling $10,000 USD.

Who can participate

We invite anyone from those new to data science to veterans of the field to participate. For those who have never tried machine learning or worked with health data before, check out the series of tutorials and webinars from the WiDS Datathon Committee, workshop organizers, and community of participants to help you get started.

The WiDS Datathon aims to inspire women worldwide to learn more about data science, and to create a supportive environment for women to connect with others in their community who share their interests. Toward these ends, we open the datathon to individuals or teams of up to 4; at least half of each team must be women (individuals identifying as female participants). Participants can include students, faculty, and individuals with various roles in non-profit, academic, government, and industry organizations.

Acknowledgements

The WiDS Datathon 2021 is a collaboration led by the WiDS Worldwide team at Stanford University, the West Big Data Innovation Hub, and the WiDS Datathon Committee. WiDS Datathon 2021 cash prizes are provided by Kaggle and the Excellence in Research Award is supported by the National Science Foundation under Grants 1916573,