



Twin Cities ASA Chapter FALL VIRTUAL SHORT COURSE Nov 6-7, 2020, -4.30PM

Empowering the Statistician with Spark, Machine Learning and Deep Learning

Data can be a valuable asset, especially when there's a lot of it. Exploratory data analysis, business intelligence, and machine learning can benefit tremendously if such big data can be wrangled and modelled at scale. Apache Spark is an open-source distributed engine for querying, processing and modeling big data. In this one-day workshop, you will learn how to leverage Spark and R/Python to process and model big data with a common machine learning algorithm. By the end of this workshop, you will have a solid understanding of how to process big data using Spark and how to build common machine learning models in the cloud environment. You will also learn the motivation and use cases of deep learning through hands-on exercises.

This workshop is designed for audiences with a statistics education background. This course bridges the gap between traditional statisticians and data scientists. No software download or installation is needed; everything is done through a laptop internet browser (Chrome or Firefox) with Databricks free cloud environment.

Instructors:

Dr. Ming Li is currently a Research Scientist at Amazon. He organized and presented 2018 JSM Introductory Overview Lecture: Leading Data Science: Talent, Strategy, and Impact. He was the Chair of Quality & Productivity Section of ASA. He was a Data Scientist at Walmart and a Statistical Leader at General Electric Global Research Center before joining Amazon. He obtained his Ph.D. in Statistics from Iowa State University in 2010. With deep statistics background and a few years experience in data science and machine learning, he has trained and mentored numerous junior data scientists with different backgrounds such as statistician, programmer, software developer, database administrator and business analyst. He is also an Instructor of Amazon's internal Machine Learning University and was one of the key founding members of Walmart's Analytics Rotational Program.

Hui Lin is the head of data science at Netflix where she is leading and building the data science department. Before Netflix, she was a Data Scientist at DuPont. She provided data science leadership for a broad range of predictive analytics and market research analysis from 2013 to 2018. She is the co-founder of Central Iowa R User Group, blogger of <https://scientistcafe.com/>, and 2018 Program Chair of ASA Statistics in Marketing Section. She enjoys making analytics accessible to a broad audience and teaches tutorials and workshops for practitioners on data science (<https://course2019.scientistcafe.com/>). She holds MS and Ph.D. in statistics from Iowa State University.

REGISTRATION on Eventbrite Page

FREE for chapter members, \$10 membership fee for non-members

REGISTRATION CLOSING NOVEMBER 4, 2020

See amstatmn.org and/or contact TwinCitiesChapterASA@gmail.com for more information
