"Multimodal Machine Learning at Scale: Democratizing AI for Academic Research"

Monday, April 11, 2022
12:15pm – 1:45pm EST
Zoom registration

PRESENTED BY:
Sanjiv Das, William and Janice Terry Professor of Finance and Data Science, Santa Clara University
Leavey School of Business; Amazon Scholar at AWS

ABSTRACT
Data analytics is mostly geared towards tabular data (numerical and categorical). Humans form decisions using this data but also make judgments based on text they read, such as news, reports, etc. Econometrics and Machine Learning have been used successfully on tabular data and also on text and images, but the combination of text and tabular data is much more powerful.

This presentation of recent tools in AWS SageMaker JumpStart will demonstrate that combining natural language processing of text with tabular data brings better results and also brings machine cognition closer to that of humans. This is especially useful in finance, where humans have been using multimodal data to make decisions, and the seminar will present and demo some use cases and tools for multimodal ML, showcasing how ML comes closer to being human.

BIO
Sanjiv Das is the William and Janice Terry Professor of Finance and Data Science at Santa Clara University's Leavey School of Business, and an Amazon Scholar at AWS.

He previously held faculty appointments at Harvard Business School and UC Berkeley. He holds postgraduate degrees in Finance (M.Phil and Ph.D. from New York University), Computer Science (M.S. from UC Berkeley), an MBA from the Indian Institute of Management, Ahmedabad, B.Com in Accounting and Economics (University of Bombay, Sydenham College), and is also a qualified Cost and Works Accountant (AICWA). He is a senior editor of The Journal of Investment Management, Associate Editor of Management Science and other academic journals, and is on the Advisory Board of the Journal of Financial Data Science.

Prior to being an academic, he worked in the derivatives business in the Asia-Pacific region as a Vice-President at Citibank. His current research interests include: portfolio theory and wealth management, machine learning, financial networks, derivatives pricing models, the modeling of default risk, systemic risk, and venture capital. He has published over a hundred and twenty articles in...
academic journals, and has won numerous awards for research and teaching. His recent book "Derivatives: Principles and Practice" was published in May 2010 (second edition 2016). Sanjiv's research may be accessed at https://srdas.github.io/research.htm

LINKS/RESOURCES

Blogs:


Documentation for more reading:
3. Smjsindustry SDK: https://pypi.org/project/smjsindustry/
5. Github Repo: https://github.com/aws/sagemaker-jumpstart-industry-pack/
Get started by going to JumpStart using SageMaker Studio: https://aws.amazon.com/sagemaker/getting-started/

For Model Bias and Explanations, see AWS Clarify: https://aws.amazon.com/sagemaker/clarify/

Request a free account on SageMaker Studio Lab here: https://studiolab.sagemaker.aws/

Research papers
1. Context, Language Modeling, and Multimodal Data in Finance: https://jfds.pm-research.com/content/3/3/52
3. On the Lack of Robust Interpretability of Neural Text

CONTACTS
Melina Denebeim ’12
Co-Director, Program for Financial Studies (PFS) and Master of Science in Financial Economics (MSFE)
md2953@gsb.columbia.edu

Harry Mamaysky
Professor of Professional Studies in the Faculty of Business; Faculty Director, Program for Financial Studies (PFS) and Master of Science in Financial Economics (MSFE)
hm2646@columbia.edu