DHRUV MADEKA https://github.com/dmadeka ◆ www.linkedin.com/in/dhruvmadeka
222 East 39th Street, Apt. 11C, New York, NY - 10016 ◆ (734)-845-4252 ◆ dm4087@nyu.edu

EDUCATION

New York University, Machine Learning and Data Science, New York, NY September 2016-
• Courses in Foundations of Machine Learning, Deep Learning, Big Data
• GPA: 4.00 • Faculty Award for Best First-Year Project

Boston University, M.S. Mathematical Finance, Boston, MA January 2014
• GPA: 4.00 • Special Recognition Award: Best Student by GPA

University of Michigan, B.S.E. Industrial & Operations Engineering, Ann Arbor, MI 2012
• Graduated Cum Laude • Dean's List, University Honors

EXPERIENCE

Amazon Corporate LLC 2017-
Senior Machine Learning Scientist, Consumer Division New York, NY

• Research Scientist primarily focused on inventing and prototyping Deep Learning models for the consumer division at Amazon

Bloomberg LP 2014-2017
Quantitative Researcher, Quantitative Research Department New York, NY

Mathematical Research:
• Developed new models for the US Presidential Elections, using techniques from Statistics and Machine Learning, that outperformed established players such as FiveThirtyEight and the New York Times
• Application of deep networks and advanced stacking models to predicting bond default
• Research in the application of Deep Learning (classification using ConvNets, object detection via TensorBox) to data retrieval from charts
• Research in new methods for scoring model predictions and their applications to online learning - which was subsequently incorporated in multiple Bloomberg functions
• Research in the use of functional Ito calculus as a more efficient way to estimate conditional expectations
• Developed novel models related to pricing, risk management, statistical estimation and portfolio allocation under the direct supervision of Bruno Dupire

Software Development:
• Founding developer and primary maintainer of the leading open-source python interactive visualization library - bqplot
• Work in prototyping different research projects using PyTorch and Keras
• Development of an object oriented framework used for creating interactive dashboards in the Jupyter Notebook to visualize training of deep networks – including the development of innovative new quantities to visualize for gauging the generalization capability of the network
• Assisted in the development of a Python based quantitative library for an advanced research platform using numpy, pandas, scikit-learn
• Gained experience in high-performance Python development, including Cython (primarily) and JIT compilers such as numba (for certain use cases)
• Developed advanced functionality for the generic simulation of Stochastic Differential Equations through higher order numerical schemes

Citizens Financial Group 2013
Quantitative Research Intern, Model Risk Validation Department Boston, MA
• Performed sensitivity analysis and Bayesian model selection statistics on New Generation OpsRisk Models
• Generated interest rate scenario simulations and modeled corporate bond spreads to determine discount factors for establishing the funding status of the RBS Americas Pension Plan

SKILLS Python – *development on the numpy, pandas, ipywidgets and bqplot libraries*, Deep Learning, PyTorch, Keras, JavaScript, Machine Learning, Data Science, Data Visualization, Stochastic Calculus, Malliavin Calculus, Machine Learning, Probability Theory, Advanced Statistics, R, Econometrics, Derivatives

PUBLICATIONS Madeka, D., Paripovic, N., Shah, S., Steigleder, K. "UM Bus Route Optimization: Diag to Diag Express." Winter Simulation Conference. Phoenix, Arizona

INVITED PRESENTATIONS

- Morgan Stanley, Internal Presentation, "An economic understanding of the Recovery Theorem"
- University of Michigan, Practitioners Seminar, "Dynamic Factor Model Asset Allocation"
- Bloomberg Quant Seminar, April 2015, "Dynamic Factor Model Asset Allocation"
- Bloomberg Quant Seminar, August 2015, "Malliavin Calculus made simple"
- Bloomberg Quant Seminar, November 2016, "Election Methodologies"
- Columbia University, Invited Guest Lecturer, MATH 4079, "McKean Stochastic Differential Equations"
- Startup ML Conference, Invited Speaker, "Machine Learning for Delinquency and Default Risk"
- NYC d3-js Meetup, Invited Speaker, "Interactive Visualization in the Jupyter Notebook"
- PyData Ann Arbor, Invited Speaker, "Interactive Visualization in the Jupyter Notebook"
- TD Ameritrade Data Science Innovation Center, Invited Guest, "Machine Learning in Finance"
- University of Michigan, Mathematics Seminar, "Machine Learning in Finance"
- PyData Seattle, July 2017, Invited Speaker, "bqplot – Seamless Interactive Visualizations in the Jupyter Notebook"
- SciPy, July 2017, Invited Speaker, "bqplot – Seamless Interactive Visualizations in the Jupyter Notebook"

CONFERENCE PUBLICATIONS

- Cliché, M., Rosenberg D., Madeka, D., Yee, C., "Scatteract: Automated extraction of data from scatter plots", European Conference on Machine Learning (EMCL-PKDD) 2017

  Charts are an excellent way to convey patterns and trends in data, but they do not facilitate further modeling of the data or close inspection of individual data points. We present a fully automated system for extracting the numerical values of data points from images of scatter plots. We use deep learning techniques to identify the key components of the chart, and optical character recognition together with robust regression to map from pixels to the coordinate system of the chart. We focus on scatter plots with linear scales, which already have several interesting challenges. Previous work has done fully automatic extraction for other types of charts, but to our knowledge this is the first approach that is fully automatic for scatter plots. Our method performs well, achieving successful data extraction on 89% of the plots in our test set.

WORKS IN PROGRESS

- Madeka, D., Nilsen, W., "Optimal Asset Allocation in a Dynamic Factor Model framework"

  Factor Models are the primary basis for tactical asset allocation amongst funds and institutional investors. They have largely been used in a static sense, with allocations based on myopically optimal weightings along risk factors. However, factor returns are rarely static, exhibiting diffusive behavior across time. As a result, dynamic strategies for hedging these factors are required for adequate return and risk management. Here we suggest a method for optimal portfolios in a dynamic sense, with a factor structure for the spanned risks in the market. The model produces an optimal portfolio in the sense of Merton (1971), with an intertemporal hedge towards the market factors, which is also studied.

- Madeka, D., "Accurate Prediction of Electoral Outcomes"

  We present novel methods for predicting the outcome of large elections. Our first algorithm uses a diffusion process to model the time uncertainty inherent in polls taken with substantial calendar time left to the election. Our second model uses the Option Market. Our third model uses Online Learning along with a novel ex-ante scoring function to combine different forecasters along with our first two models. We also consider Bayesian Regressions and a Graphical Model. We also present new density based scoring functions that can be used to better evaluate the efficacy of forecasters.