Dimitrios Vamvourellis

Quantitative Researcher, Financial Modeler, NLP Specialist

141 Maujer St, New York, NY 11206

dimvamvourellis@gmail.com | dvamvourellis.github.io | +1 617-229-9771

Education

Harvard University **Cambridge, MA**

MS in Data Science, GPA: 4.0/4.0 *March 2021*

Ranked top school for professional studies in Data Science (acceptance rate ~1%)

Relevant Coursework: Stochastic Methods for Data Analysis, Inference and Optimization, Bayesian Modeling, Statistical Learning Theory, Natural Language Processing.

The University of Edinburgh **Edinburgh, UK**

BS in Computer Science and Mathematics *May 2017*

Ranked in the top 5 Computer Science colleges in the UK

Class Valedictorian, First Class Honors

- Focus in Machine Learning and Statistics
- Thesis: Deep Learning on Low Power GPUs

Prizes

Class Medal for BS Computer Science and Mathematics **Edinburgh, UK | 2017**

- Awarded to the student with the top performance over the entire 4-year degree in a class of 50 students at one of the top Informatics Schools in the world.

Kevin Clarke Memorial Prize **Edinburgh, UK | 2017**

- Awarded to the student with the highest performance in the last year of the BS in Computer Science and Mathematics at the University of Edinburgh.

Google Prize for best performance in System Design Project **Edinburgh, UK | 2016**

- The goal of this group project was to create a robot to play two-a-side football autonomously.
- Our 5-member team was selected among 30 other teams for the Robot with the greatest performance in the Robocup Competition at the University of Edinburgh.

Work Experience

BlackRock **New York, NY**

Vice President - Senior Data Scientist *Nov 2023-Present*

- Largest asset manager in the world managing over 8 trillion dollars.
- Conducting machine learning and NLP research to develop AI-powered analytics and models with broad business impact from portfolio management to trading execution.
- Fully developed and deployed to production ML model which predicts real-time prices for the entire European and US corporate bond market, reducing the pricing error by 24% on average. Used to price transactions exceeding 5 billion dollars per week.
- Built NLP-powered pipelines to extract thematic sentiment signals from unstructured sources (SEC filings, earnings calls etc.), used for reducing equity transaction costs, building ESG alpha generation strategies and improving equity risk decomposition.
- Led a team of 8 data scientists to develop a Python NLP library to centralize and streamline NLP research across business, providing the entire company with a toolkit of the state-of-the-art NLP algorithms.
- Thought leader on Natural Language Processing and core contributor to strategic initiative of adopting cutting-edge Generative AI for financial modeling.
- Fully designed and developed a multi-agent GenAI framework capable of providing answers to complex, multi-step, analytical queries that require LLMs to generate SQL and Python code, as well as reason around data and analytics. The agent will be used to generate insights as well as inform real-time portfolio management and trading decisions.

BlackRock **New York, NY**
Associate - Data Scientist *Mar 2021-Oct 2023*

- Conducting machine learning and NLP research to develop AI-powered analytics and models with broad business impact from portfolio management to trading execution.
- Fully developed and deployed to production ML model which predicts the real-time bid-ask-spread for the entire European and US corporate bond market, used to reduce transaction costs for corporate bond trades.
- Built AI-powered pipelines to extract sentiment analytics from news, used to model trading difficulty for equities, as well as constructing portfolios with optimal ESG characteristics.
- Created NLP models to model the similarity of mutual funds and companies based on textual descriptions.

Intelligencia.ai **Cambridge, MA**
Data Scientist *Mar 2020-Mar 2021*

- Listed in the top 130 startups globally reinventing healthcare with AI.
- Created AI models to assess the probability of success in clinical trials and other analytics to optimize portfolios of the top pharmaceutical companies.
- Built NLP and AI powered models to identify emerging areas of innovation in clinical research.
- Fully developed and deployed product based on dynamic topic models to detect trending topics and influential/upcoming research scientists based on the evolution of medical literature. Product is used by pharmaceutical companies to optimize portfolio by investing in upcoming labs and potential drug development methodologies.

Harvard Data to Actionable Knowledge Lab **Cambridge, MA**
Research Assistant *Jan 2020-Feb 2021*

- Invited to be member of one of the top labs worldwide developing AI methodologies for safety-critical applications.
- Researched techniques for scalable and useful uncertainty estimation of deep learning models with a focus on safety-critical AI applications in healthcare.

BlackRock **Edinburgh, UK**
Analyst *July 2017-June 2019*

- Provided portfolio managers and risk managers with expert advice on using the risk modeling, security analytics and portfolio management workflows of the investment platform.
- Closely worked with top European clients on integrating the Aladdin platform to their investment workflows.

Publications

BlackRock **New York, NY**
A Natural Way of Building Financial Domain Expert Agents *October 2024*

- Developed a framework for building GenAI-powered financial domain expert agents, leveraging zero-shot learning, code generation and an adaptive memory layer.
- Paper published at Workshop on Deploying Real-World AI/ML at the 5th ACM International Conference on AI in Finance, 2024.

BlackRock **New York, NY**
Company Similarity using Large Language Models *October 2024*

- Researched the ability of small and large language models to produce financially meaningful company embeddings based on business descriptions reported in SEC filings.
- First author.
- Paper published at Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr), 2024.

BlackRock **New York, NY**
Enhanced Local Explainability and Trust Scores with Random Forest Proximities *October 2024*

- Developed a novel approach to explain predictions and quantify confidence of Random Forest models.
- Paper published at the 5th ACM International Conference on AI in Finance, 2024.

BlackRock **New York, NY**
Towards enhanced local explainability of random forests: A proximity-based approach *October 2023*
- Developed a novel approach to explain predictions and quantify confidence of Random Forest models.

BlackRock **New York, NY**
Learning Mutual Fund Categorization using Natural Language Processing *June 2022*
- Developed BERT-based models to learn mutual fund embeddings based on textual descriptions and explored business applications related to fund similarity learning.
- First author.
- Paper published at International Conference for AI in Finance, 2022, the most reputable conference for AI applications in the financial field.
- Patent application had been filed by BlackRock for the methodology described in this paper.

Harvard University **Cambridge, MA**
Bayesian Classifiers with Out-of-Distribution Uncertainty *June 2020*
- Developed a novel Bayesian framework for estimating and decomposing the predictive, in and out-of-distribution uncertainty of Bayesian Neural Networks for classification tasks.
- First author.
- Paper published at ICML Workshop on Uncertainty and Robustness in Deep Learning, 2020, the top machine learning conference globally.

Judging

5th ACM International Conference for AI in Finance **New York, NY | 2024**
- Invited to be member of the Program Committee at the most significant global conference for AI in Finance.
- Served as a judge conducting official peer reviews for papers submitted at the conference.

4th ACM International Conference for AI in Finance **New York, NY | 2023**
- Invited to be member of the Program Committee at the most significant global conference for AI in Finance.
- Served as a judge conducting official peer reviews for papers submitted at the conference.

Blackrock, Machine Learning Model Approval Committee **New York, NY | 2023**
- Appointed member on a committee of responsible for reviewing Machine Learning models used in production across business.
- Serving as subject matter expert on machine learning and Natural Language Processing, judging models and methodologies brought for approval at the committee.
- Committee is responsible for safeguarding the company in terms of using safe and transparent machine learning models that comply with SEC regulations.

Patents

BlackRock **New York, NY**
1st Named Inventor *July 2023*
- Patent application filed by BlackRock with title Learning Categorization with Natural Language Processing Networks.

Teaching

Harvard University **Cambridge, MA**
Teaching Fellow *Aug - Dec 2020*
- Taught the graduate class on Advanced Scientific Computing: Stochastic Methods for Data Analysis, Inference and Optimization.
- Reviewed weekly assignments and was responsible for helping students during office hours.

Columbia University **New York, NY**
Teaching Fellow *January 2024*
- Invited to teach parts of a class on applications of Machine Learning and NLP in Finance.

Conference Presentations

IEEE Symposium on Computational Intelligence for Financial Engineering and Economic **Hoboken, NJ | 2024**

- Invited to present work on *Company Similarity using Large Language Models.*
- Presentation attended globally by hundreds of finance practitioners and academics in field of AI in Finance.

5th ACM International Conference for AI in Finance **New York, NY | 2024**

- Invited to present work on *A Natural Way of Building Financial Domain Expert Agents* and *Enhanced Local Explainability and Trust Scores with Random Forest Proximities*.
- Presentation attended globally by hundreds of finance practitioners and academics in field of AI in Finance.

4th ACM International Conference for AI in Finance **New York, NY | 2023**

- Invited to present work on *Company Similarity using Large Language Models.*
- Presentation attended globally by hundreds of finance practitioners and academics in field of AI in Finance.

3rd ACM International Conference for AI in Finance **New York, NY | 2022**

- Invited to present work *Learning Mutual Fund Categorization using Natural Language Processing.*
- Presentation attended globally by hundreds of finance practitioners and academics in field of AI in Finance.

37th International Conference in Machine Learning **Remote | 2020**

- Invited to present work *Bayesian Classifiers with Out-of-Distribution uncertainty.*
- Top Machine learning conference worldwide attended by the best academics in the field of AI.

Other Presentations and Panels

MLOps Community **New York, NY | 2023**

- Global community organizing large scale events focusing on sharing real-world Machine Learning Operations best practices from engineers in the field of machine learning and AI.
- Over 17k members globally.
- Invited to participate in panel discussion with theme “Best Practices for MLOps in Financial Services”.
- Invited to present on Applications of Large Language Models for Financial Modeling and Quantitative Research.

BlackRock Innovation Network **New York, NY | 2023**

- Largest forum for innovation within BlackRock inviting both internal and external speakers to give talks on applications of Machine Learning in Finance.
- Invited to present work on Company Similarity using LLMs
- Presentation attended globally by over 300 people.

Risk and Quantitative Analysis Global Staff Meeting **New York, NY | 2023**

- Invited by the Chief Risk Officer to give presentation on recent advances of Natural Language Processing and its applications within BlackRock and in finance.
- Presentation attended globally by over 100 people.

BlackRock Investment Institute Forum **New York, NY | 2023**

- Offsite bi-annual forum for top BlackRock leaders across asset classes with focused presentations and sessions on current investment themes, geopolitical and innovation trends.
- Presented work on current applications of Large Language Models for investment management.

BlackRock Innovation Network **New York, NY | 2022**

- Largest forum for innovation within BlackRock inviting both internal and external speakers to give talks on applications of Machine Learning in Finance.
- Invited to present work on Mutual Fund Categorization using Natural Language Processing.
- Presentation attended globally by over 300 people.

Books

Introduction to Computer Science **Athens, Greece**

Assistant Author *Jun - Sep 2014*

- Invited by the main author, Professor Christos Douligeris to contribute in the writing of a textbook for Computer Science.
- Wrote chapters on Artificial Intelligence, Information Systems, Operating Systems.
- After a national competition, book was selected by the Greek Ministry of Education as the most suitable to be used nationally for teaching Introduction to Computer Science at the 5th grade of Greek high schools.

Languages

- Greek
- English

Skills

Programming: Python (Pandas, PySpark, Pytorch, Scikit-learn, Keras, Numpy, PyMC3, Gensim, Langchain), SQL.