Marios Kokkodis, *PhD*
Applied Scientist @ Thumbtack | ex Meta | ex Academic
marios.kokkodis@gmail.com — (951) 941-4357
linkedin.com/in/kokkodis/

Machine learning IC leader with 10+ post-PhD years of experience driving the design, deployment, and optimization of ML systems in production. Particularly focused on search and ranking problems, recommender systems, and personalization. Experience with designing and deploying solutions that orchestrate multi-task learning, fine-tuned LLMs, and multi-shot prompting to deliver a personalized search and ranking experience. Publications in top applied ML outlets (KDD, WWW, WSDM, Management Science, ISR, POMS, MISQ, Plos 1). Experienced in cross-functional leadership, collaborating with product, design, content, research, marketing, and data science to shape ML strategy. Driving roadmap and execution while mentoring 4 applied scientists and several engineers.

2024 - Current ***Staff (L7) Applied Scientist*** at **Thumbtack** (*Search and Ranking*)

- ◇ Machine learning
  - ❑ Designed, aligned, and currently leading the deployment of the first **multi-task ranker** algorithm at Thumbtack. The new system focuses on increasing supply by recommending broadly relevant professionals, while optimizing both engagement (contacts) and satisfaction (hires, lead declines, refunds) metrics. An early version of the ranker is currently deployed in a single market (major US city), showing a **3% increase in engagement** (contacts). (Multi-task learning, fine-tuning, LLM summarization and feature generation – Pytorch, MLFlow, Fiddler)
  - ❑ Designed, aligned, and currently leading the deployment of the new **multimodal** (image, voice, natural language) **search** feature for Thumbtack. The new search ML system orchestrates several classification models along with conversational prompting to deliver a personalized search user experience (**link to relevant paper** of an early version of this system). Experiments of several early versions of the system showed that the it is already revenue neutral with the baseline search system, while it outperforms the current system by 28% in terms of customer trust and satisfaction (OpenClip, Finetuned classifiers, LLMs for feature engineering, LLMs for Prompting, LLMs as a judge – Scikit-learn, Pytorch, MLFlow, Fiddler | **patent filed** | received **innovation award** from Thumbtack)
  - ❑ Designed, aligned, and currently leading the deployment of the new prompt-based match explanations system that summarizes why a pro is a good match for a customer and vice versa. The system is currently deployed for the first time on the pro side. Survey results expected in Q3. (prompting, LLMs, LLMs as a judge)
  - ❑ Aligned and currently designing the new dynamic pro-preferences system that will infer the types of tasks that a pro is willing to take, recommends new tasks, and increases transparency and pro control. The first iteration of this system is expected to be deployed by the end of Q3. (**provisional patent filed**)

2022 - 2024 ***Staff (IC6) Applied Scientist Tech Lead*** at **Meta** (*Ads Infra and Ranking*)

- ◇ Machine learning
  - ❑ Designed and deployed an end-to-end machine learning framework (supervised learning, unsupervised learning, anomaly detection) for predicting launch risk. By accelerating the launch of low-risk Ad models while flagging risky models for additional review, this framework has **generated $94M in 2023H2**. (Python–Scikit-learn)
  - ❑ Fine-tuned (sampling adjustments, explained and measured counterfactual impact) a system that rejects model snapshots that are not within certain performance boundaries. Precision improvements prevented **monthly losses of $\sim$ $10M**. (Python, SQL)
  - ❑ Designed and deployed an end-to-end ranking framework—bootstrapping, probabilistic *language models*, and ranking models—to personalize the app experience for billions of users. (Python, R, Dataswarm–PyTorch)
  - ❑ Created an end-to-end machine learning predictive framework (PCA, regression analysis) for outlier app experience detection and monitoring. (Python, R, Dataswarm, Unidash–Scikit-learn)
  - ❑ Established end-to-end pipelines for analyzing observational data (regression analysis, fixed effects, propensity score matching) to robustly measure wins and provide trade-off analysis that informs VP-level decision-making. (Python, R, SQL, Dataswarm, Unidash–Scikit-learn)
- ◇ Leading, Mentoring, and Cross-Functional Collaboration
  - ❑ Designed a framework for understanding and measuring the impact of 25+ reliability systems (e.g., systems that stress-test and monitor ML models, feature serving, calibration, A/B testing, signal growth, and interpretability) within Meta's Ads ML infra. The framework provides input to ROI calculations and guides

project prioritization and headcount investment; it is projected to generate $\sim$ **$400M**. (Metrics, Strategic alignment with 20+ XFNs–statsmodels)

- Tech-led and mentored junior colleagues by (1) providing weekly guidance and actionable feedback on their work, (2) scoping new opportunities for them, and (3) helping them roadmap their work.
- Organized and led several org-wide study groups on recommender systems, Ad auctions, and experimentation.

2015 - 2022 ***Research Professor of Applied Machine Learning*** at **Boston College**

- Machine learning
  - Designed deep *reinforcement learning* frameworks (DQN, Dueling DQN, Double DQN, NLP) to provide career path recommendations that could result in up to a 6% increase in market revenue and a 22% increase in worker wages. (Python, R, Keras, PyTorch, published paper)
  - Designed *sequence-aware recommender systems* (HMM, LSTM) that match workers with potential employers; predictive performance (better quality matches) up to 40% better than state-of-the-art baselines. (Python, R, SQL, Keras, published paper)
  - Proposed machine learning frameworks that predict user engagement; feature-engineered new engagement metrics that increase predictive performance by up to 40%. (Python, R, SQL, Keras, published paper)
  - Designed dynamic expertise assessment systems (HMM, W2V, NLP) that yielded 20%$-$60% better outcomes than state-of-the-art baselines. (Python, R, SQL, Keras, published paper)
  - Explained (panel data, instrumental variables, W2V, NLP) the trade-offs of skillset diversification. (Stata, Python, R, SQL, published paper)
  - Explained (panel data, matching, topic modeling, NLP) biases that affect the reputation systems of online platforms. (Stata, Python, R, SQL, published paper)
  - Explained (natural experiment, difference-in-differences, panel data, subsample analysis) the effect of purchase verification on a market's reputation system. (Python, R, Stata, published paper)
- Leading, teaching, and mentoring
  - Organized and ran several international conferences and workshops. Designed programming, machine learning, and database classes. Negotiated for hundreds of students to get free access to R Studio Cloud.
  - Taught core programming, machine learning, and database classes for undergraduate and graduate students. Repeatedly awarded for influential teaching.
  - Mentored 10+ students by supervising undergraduate and graduate theses and providing career advice.

2015 ***Machine Learning Scientist*** at **Upwork**

- Prototyped, deployed, and fine-tuned (feature engineering) the platform's first employer-worker matching algorithm, which led to a 4% increase in revenue and an 8% increase in outcomes. (Java, SQL, published paper)

Prior to 2015: ***Research Scientist*** (intern) at **Microsoft Research**, ***Machine Learning Scientist (intern)*** at **Upwork**, and ***Software Engineer*** at the **National Technical University of Athens**.

Topics and Methodologies

- Predictive modeling, supervised and unsupervised learning, feature engineering, recommender systems, reinforcement learning, deep learning, neural networks, NLP, time-series forecasting, anomaly detection, LLMs, prompting, Finetuning, RAG
- A/B testing, difference-in-differences, panel data, instrumental variables, Heckman selection models, propensity score matching, parametric and non-parametric survival analysis

Programming Languages and Tools

- Python (Pandas, Numpy, Scikit-learn, Keras, PyTorch, Transformers (Hugging Face), Statsmodels), R, SQL, Stata, Shell Scripting (prior to 2016: Java, C#, C++, C).
- Git, Gerrit, Dataswarm (Airflow equivalent), Presto, Hive, Manifold (HDFS equivalent), MySQL, Unidash, Deltoid3, FBlearner, MongoDB, portable batch system, Hadoop, MapReduce. AWS, Databricks, MLflow, Big query, GCP, Sagemaker Studio, Fiddler, Docker.
- Jupyter, Tableau, VSCode, Rstudio, PyCharm, Vim.

Education

- **New York University**. ***PhD*** in Applied Machine Learning & Causal Inference (Information Systems)
- National Technical University of Athens. BS in Computer Engineering (5-year program)

Selected Honors and Awards

- INFORMS ISS Gordon B. Davis **Young Scholar Award** and ISS Nunamaker-Chen **Dissertation Award**
- INFORMS Data Mining **Best Paper Award**
- Represented Cyprus in the **International Physics Olympiad**
- Represented Cyprus in the **International Biology Olympiad**
- Several awards and honorary mentions in **Annual Mathematics Cyprus Olympiads**

Incomplete selection of Solo Applied Machine Learning Publications – (Google Scholar profile.)

- **Kokkodis Marios**. 2025. Beyond the Hype: Embeddings vs. Prompting for Multiclass Classification Tasks. ***Under review***. (link to paper)
- **Kokkodis Marios**. 2022. Adjusting Skillset Cohesion in Online Labor Markets: Reputation Gains and Opportunity Losses. ***Information Systems Research***. (link to paper)
- **Kokkodis Marios**. 2021. Dynamic, Multi-dimensional, and Skillset-specific Reputation Systems for Online Work. ***Information Systems Research***. (link to paper)
- **Kokkodis Marios**. 2019. Reputation Deflation Through Dynamic Expertise Assessment in Online Labor Markets. ACM World Wide Web Conference (***WWW***). (link to paper)
- **Kokkodis Marios**. 2018. Dynamic Recommendations for Sequential Hiring Decisions in Online Labor Markets. ACM International Conference on Knowledge Discovery and Data Mining (***KDD***). (link to paper)