**Marios Kokkodis, *PhD***
Applied Machine Learning Scientist @ Meta
marios.kokkodis@gmail.com — (951) 941-4357
linkedin.com/in/kokkodis/ — kokkodis.com — Google Scholar

Machine learning generalist with 10+ years of experience in predictive (supervised and unsupervised) modeling, recommender systems, personalization, deep learning, reinforcement learning, NLP, anomaly detection, causal inference, and A/B testing. Experience with Transformers and LLM fine-tuning for supervised learning. Deployed several end-to-end customer and internal-facing machine learning solutions. Tech lead working with diverse cross-functional stakeholders and mentoring several junior data and machine learning scientists. Publications in top applied machine learning science outlets (KDD, WWW, WSDM, Management Science, ISR, POMS, MISQ, Plos One).

---

2022 - Current ***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)
  - ❑ (WIP) Studied and fine-tuned (defined boundary conditions, explained and measured counterfactual impact, improved precision) a system that rejects model snapshots that are not within certain performance boundaries. The system prevents losses of roughly $250M every half. Precision improvements **are expected to prevent additional 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. (Python, SQL, 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
- 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).
- Dataswarm (Airflow equivalent), Presto, Hive, Manifold (HDFS equivalent), MySQL, Unidash, Deltoid3, FBlearner, MongoDB, portable batch system, Hadoop, MapReduce. Meta internal devservers; AWS, Sagemaker Studio for personal projects.
- Jupyter, Tableau, VSCode, Rstudio, PyCharm, Vim.

---

Education

- **New York University**. ***PhD*** in Applied Machine Learning (Information Systems)
- University of California Riverside. MS in Computer Science
- 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**. 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)