

# Yunlong Jiao

## Machine Learning Scientist

📍 London, UK (No Sponsorship Needed)  
☎ +44 7400 724281  
✉ [yljiao.ustc@gmail.com](mailto:yljiao.ustc@gmail.com)  
🌐 [linkedin.com/in/yunlong-jiao/](https://www.linkedin.com/in/yunlong-jiao/)

### ABOUT ME

I am currently a machine learning scientist at Amazon Alexa, working on building AI assistant that is able to replicate human intelligence. I am passionate about applied AI research that drives tangible impact for products and services. With over 5 years of post-PhD experience in the field of ML/AI, I possess strong skills in model development and performance evaluation, and am proficient in Python and deep learning frameworks. While successfully incubating and delivering multiple projects at Amazon, I am involved in all phases of the development cycle – from building proof of concept to delivering minimum viable products and scaling up production. I have experience in supervising Master's thesis and research internships, as well as coaching and mentoring junior team members.

### EDUCATION

- 2013 – 2017 **Doctor of Philosophy**  
*Centre for Computational Biology*  
*Paris Sciences et Lettres – PSL Research University, Paris, France*
- 2012 – 2013 **Master of Science** (MENTION TRÈS BIEN)  
*Department of Mathematics*  
*University of Paris-Saclay, Orsay, France*
- 2008 – 2012 **Bachelor of Science** (FIRST CLASS HONOURS)  
*Department of Mathematics*  
*University of Science & Technology of China, Hefei, China*

### WORK EXPERIENCE

#### Machine Learning Scientist

*Amazon, London, UK*

CURRENT, FROM NOV 2019

- **Key Skills:** Natural Language Processing, Large Language Models, Deep Generative Models, Neural Text-to-Speech
- At Alexa Shopping Science team (from Mar 2021), my current responsibilities are:
  - 1) Leading research on bias mitigation technologies in NLP while preserving user privacy and fairness;
  - 2) Working closely with cross-functional teams to build and deploy ML solutions and oversee prototype to production;
  - 3) Supervising research internships and coaching and mentoring junior team members.
- At Alexa Text-to-Speech Research team (Nov 2019 – Feb 2021), I had:
  - 1) Proposed and delivered a deep generative model for universal speech synthesis, which generalises well for various out-of-distribution scenarios and can be used regardless of speaker identity, speaking style, or language;
  - 2) Collaborated with engineers to scale up Alexa TTS production and launch new voices in multiple regions and countries.

#### Postdoctoral Research Scientist

*University of Oxford, Oxford, UK*

NOV 2017 – OCT 2019

- **Key Skills:** Gaussian Processes, Time Series Forecasting, Multi-modality
- My responsibilities included:
  - 1) Leading Oxford research in a multi-organizational project (involving organizations in the UK and EU) and developing novel methods for longitudinal modelling of complex disease using multi-modal data integration;
  - 2) Supervising Master's thesis in the Department of Statistics.

#### Doctoral Researcher

*PSL Research University, Paris, France*

SEP 2013 – SEP 2017

- **Key Skills:** Kernel Methods, Representation Learning, Graph Learning, Sparsity Regularisation
- My thesis work contributed novel ML methods and advanced scientific discoveries in biology and cancer research:
  - 1) My research focus was kernel methods and representation learning of non-tabular data, such as rank data and graphs.

- 2) My research outputs significantly improved prediction of breast cancer survival using genetic data, and guided interpretable biomarker discovery with the help of biological networks.
  - 3) I had published multiple first-authored papers in top ML conferences and journals (3 at ICML and 1 at IEEE TPAMI) and developed open-source toolkits.
- A few colleagues and I participated, and finally placed 2nd, in the DREAM Toxicogenetics Challenge, a Kaggle-style community competition in data science that is aimed to advance computational methods in biology.

## Data Scientist Intern

APR 2015 – JUN 2015

Roche Diagnostics GmbH, Penzberg, Germany

■ **Key Skills:** Information Extraction, Feature Engineering, Large-Scale Unstructured Database

■ During the internship, I had:

- 1) Proposed a data pipeline to process large-scale unstructured machinery performance data and built a model to predict failure state for automated immunoassay analyzers;
  - 2) Applied the solution to improve maintenance efficiency for one of Roche's core hardware products.
- My innovative work had a direct impact on strengthening customer trust, and was patented by Roche.

## TECHNICAL SKILLS

---

PROGRAMMING Python (numpy, pandas, scikit-learn), Deep Learning Frameworks (PyTorch, MXNet), R, C/C++  
BIG DATA Accelerated Computing (CUDA), Cloud Computing (AWS, SageMaker), SQL  
DEVOPS Bash, Git, Docker, Open Source, Unit/Integration Testing, CI/CD

## PROFESSIONAL SKILLS

---

COMMUNICATION Experienced speaker at international conferences and workshops  
Confident in presenting project ideas and results to peers, leadership, and stakeholders  
Mentoring and coaching research interns and junior team members  
WRITING Proficient in academic writing and providing guidance to early-stage researchers in their writing  
Experienced in leading R&D proposals and producing technical reports on milestone deliveries  
PROJECT MANAGEMENT Capable tech lead for research projects and trusted in cross-function team collaboration  
Knowledgeable in the principles of Agile development and skilled working in Scrum  
LANGUAGES Mandarin Chinese (native), English (bilingual), French (conversational), Spanish (learning)

## SELECTED PUBLICATIONS AND PATENTS

---

F Liu, **Y Jiao**, J Massiah, E Yilmaz, S Havrylov. "Trans-Encoder: Unsupervised Sentence-Pair Modelling Through Self- and Mutual-Distillations." *ICLR*, 2022. [🔗](#)

A Gabryś, **Y Jiao**, V Klimkov, D Korzekwa, R Barra-Chicote. "Improving the Expressiveness of Neural Vocoding with Non-Affine Normalizing Flows." *Interspeech*, 2021. [🔗](#)

**Y Jiao**, A Gabryś, G Tinchev, B Putrycz, D Korzekwa, V Klimkov. "Universal Neural Vocoding with Parallel WaveNet." *ICASSP*, 2021. [🔗](#)

F Heinemann, S Kobel, S Dahlmanns, JP Vert, **Y Jiao**. "Failure State Prediction for Automated Analyzers for Analyzing a Biological Sample." *US Patent App. 16/416,844*, 2019. [🔗](#)

**Y Jiao**, JP Vert. "The Weighted Kendall and High-order Kernels for Permutations." *ICML*, 2018. [🔗](#)

**Y Jiao**, A Korba, E Sibony. "Controlling the Distance to a Kemeny Consensus without Computing It." *ICML*, 2016. [🔗](#)

**Y Jiao**, JP Vert. "The Kendall and Mallows Kernels for Permutations." *ICML*, 2015. [🔗](#)