Tony Tong

Cambridge, MA   +1 (617) 710-9767   tongbaojia@gmail.com

github.com/tongbaojia   tongbaojia.github.io   linkedin.com/in/baojiatonytong

Skills

Languages   Python, C++, SQL
Packages   PyTorch, DGL, XGBoost, Sklearn, ROOT
Tools   Git, Bash, LaTeX, SVN

Experience

Suno.ai   Cambridge, MA
Machine Learning Engineer   Aug 2023–present

Meta Platforms, Inc.   Cambridge, MA
Machine Learning Engineer   Aug 2022–Aug 2023

- Tech lead on accuracy workstream. Increased high-impact metric accuracy on novel devices by more than 50% through data-driven algorithm developments, which directly improved user sentiments.
- Collaborate with two engineering XFNs, Data Science and Project Managers, and set directions on novel features and data acquisition methods.
- Planned, executed, validated, and improved data collection experiments.
- Developed data simulation framework for continuous validation and development.

Kensho Technologies   Cambridge, MA
Machine Learning Engineer   Oct 2018–Jul 2022

- Lead the document intelligence team's work and research planning, collaborating with 4 machine learning engineers on OCR, text ordering, and document layout analysis.
- Created and productionized **a novel graph-based deep learning model** for document layout analysis. Achieved state-of-the-art performance in classification and segmentation of documents with less than 10% of the computation cost.
- Combined the graph model with **natural language understanding** to create a key-value extraction tool.
- Invented a novel document layout analysis algorithm and architected it into **a production-level Python package** to **extract tabular and textual** data from PDF documents. The package has been used to extract text and tables from more than **20 million** PDF documents in production.
- Productionized **a financial key-value extraction workflow** for broker research documents. The pipeline automatically extracts $\sim 60\%$ of data with $\sim 98\%$ precision.
- Built and deployed **a general-purpose synonym model** based on Wikipedia open data. The model powers the **text expansion** capability on search platforms daily.

Education

Harvard University, Ph.D. in Physics   Sep 2012–May 2018
California Institute of Technology, B.A. with honors in Physics   Sep 2008–Jun 2012