

Thomas Chang

@thomas.w.chang@gmail.com

 [thomaschangsf.github.io](https://github.com/thomaschangsf)

 linkedin.com/in/thomaswchang

SUMMARY

Machine learning professional who delivers end to end solutions, from data collection, data preparation, training, and deployments in the Search, NLP, and Generative AI domains. Lifetime learner who enjoys working collaboratively within and across teams.

TECHNICAL SKILLS

- ML Frameworks: PyTorch, Spark, Scikit
- Web Service Frameworks: Flask, FastAPI, and Play
- MLOps: Docker, Kubernetes, Helm, AWS, and Google Cloud
- Languages: Python, Scala, and GO

PROFESSIONAL EXPERIENCE

Salesforce - AI Platform Trust

Lead Machine Learning Engineer

Feb 2022 -Current

- Contributed to the Trust Layer on Salesforce AI platforms of Salesforce CoPilot. Trust Layer ensures LLM compliance with Salesforce AI ethics standards. Implemented trust vector detection, including PII identification, prompt defense, and toxicity analysis. Evaluated both internal and external LLMs and models against diverse trust metrics.
- Evangelize Offline Model Analysis in Production (OMAP) for LLM evaluations with real production data, which enables data scientists and safety professionals to investigate, understand, and improve AI agents. Framework provides seamless migration to scalable pipelines which drives dashboards and alerts.
- Led the development of retrieval augmentation for Salesforce Flow GPT, which enables the automatic generation of marketing workflows and code. Enhanced prompt grounding by improving semantic relevance between prompts and meta objects. Managed the project from proof of concept to general availability. Collaborated with research, platform, and product teams to define the architecture and integrate the implementation.
- Responsible for python model service to marketing cloud which powers next generation ML powered use cases such as subject line generation and model bot. Improve and guide python best practices.

EBAY - NLP Core Technology

Member of Technical Staff II

May 2020 - Feb-2022

- Improve trust on the commerce ecosystem by building machine learning models to identify bad actors and remove offensive materials.
- Contribute end to end NLP framework which takes NLP models from exploration to deployments. Features include CLI to load all NLP models, helm chart deployments to Kubernetes clusters, batch model inference in HDFS, and load testing framework.
- Trained Ebert-based Named Entity Recognition (NER) Pytorch models and deploy to production for online inferencing. Federated Kubernetes deployment achieves 4500 TPS. Tune model and services.

EBAY - Core Search

Technical Lead (Search and Ads)

Nov 2018 - May 2020

- Enhance ranking algorithm to maximize revenue between Cost Per Click (CPC) and Cost Per Sale (CPS). To support the new CPC program, I led a team of data scientists and engineers to build a pClick and revenue floor model.
- Build machine learning models to suggest item adrate guidance for Ebay sellers on the Promoted Listings program.
- As part initiative to change the search experience, explore search modules to help customers understand the inventory and make value tradeoffs.

- Improve search ranking algorithm with user segmentation features, which includes hand crafted features and embedding representations.
- Design and architect Spark Data Pipeline used for all search monetization data requirements. Pipeline was scaled to 40+ jobs, adopted by multiple projects, and contributed by data engineers, data scientists, and software engineers.
- Manage a team of 8 machine learning engineers, data scientists, and QA engineers for the search ad projects.

EBAY - HAITO

Member Of Technical Staff I/II (Search and Chatbots)

Jan 2016 - Nov 2018

- Recommendation: Worked on recommendation modules for the item and home page for the Ebay HaiTao Ebay website. Item based recommendation modules utilize a deep learning model to recommend items based on item similarities. Home page recommendation module leverages user view history to make personal recommendations via a collaborative filtering model. Final pipeline includes Airflow orchestration of ETL, ML training, and deployment.
- Recommendation: Worked on commerce Knowledge Graph for aspect recommendation and similar queries. Enrich knowledge graph with general knowledge from Wikidata. Create Spark ETL to ingest 450G of dense data, web service to handle client queries. Implement ontology mapping between wikiData and relevant nodes on the Ebay Knowledge Graph by leveraging Wordnet.
- Search ranking and relevance: Core member in designing and implementing a search engine leveraging structure search and various machine learning models. Components included recall, relevance, and machine learned ranking.
- Search pipeline: Write the near real time data pipeline for search listing index, which achieves a sustained 8,000 Elasticsearch updates/sec and is replicated across multiple international regions. Search index supports the entire 1 billion Ebay items. To ensure data integrity, drove an integration regression framework and various spark jobs used to validate data integrity of the data pipeline.
- Champion best practices so new team members can onboard efficiently and effectively.

SERVICENOW

2012 - 2015

Senior Software Engineer

- Developed cloud infrastructure management software to enable customers to provision, discover, and apply change control to resources on Microsoft Azure and Amazon Web Services cloud.
- Mentored junior developers and helped new team members come up to speed.

LAWRENCE LIVERMORE NATIONAL LAB

2011 - 2012

Computer Scientist

- Wrote simulation jobs for the National Ignition Facility, which is the world's largest laser facility dedicated to achieving fusion. Distributed jobs run on the lab's High Performance Computing Platform using Message Passing Interface (MPI), OpenMP, and MOAB, (a job scheduler).

APPLE

2009 - 2011

Software Engineer

- Designed WiFi calibration models to optimize transmit power over operations modes and frequencies. Algorithm uses machine learning (regression analysis) to select calibration parameters to meet antenna power transmission requirements.
- Explore using models to predict Iphone hardware failures. Build a toolset to collect, analyze, and train models.

CREDENTIALS

MS, UC Berkeley

BS, UC Berkeley

Stanford Advanced Project Management Certificate

Agile ScrumMaster Certificate