Ajinkya Khamkar

812-606-2086 | ajinkya.khamkar93@gmail.com | linkedin | github

Education

Indiana University — Bloomington, In
Master of Science in Data Science — *Aug. 2016 – May 2018*

University of Mumbai — Mumbai, India
Bachelor of Science in Computer Engineering — *Aug. 2011 – May 2015*

Experience

Senior Machine Learning Engineer — September 2021 – Present
Criteo Corporation — *Palo Alto, CA*

- Build machine learning algorithm to **estimate a Brand's presence and performance** on multiple e-commerce platforms. The algorithm uses Brand's product placements for popular search keywords and categories the Brand is associated with
- Integrate new marketplaces, scale existing systems to account for increased product traffic using airflow and kubernetes
- Maintain existing code base, identify performance bottlenecks and rewrite code to tackle them, troubleshoot system failures
- Built **Recommender System** to help Brands identify material changes they should make to their ad placements on e-commerce websites to increase Revenue and presence

Data Scientist - Search and Recommendations — April 2019 – September 2021
Shipt, Inc — *San Francisco, CA*

- Designed a **sequential prediction** model for "Peripheral vision" system, which determines next category member will hop to given current composition of basket
- Python Tensorflow Sequence-to-Sequence Model trains and predicts parallelly on a 4 GPU AWS instance
- System increased browse conversion rate by **4 %** on search pages and increased conversion rate on PDP shelf recommender by **3%**
- Built **personalized recommender system**, a wide and deep network architecture which jointly models member signals such as dietary preferences and frequency of orders, product signals such as brand, entity and nutritional information as a supervised learning problem
- Model was deployed on homepage shelf and achieves a conversion rate of **11 %**
- Deployed model to train and predict parallelly on 16 GPU's. Train / test / validation features are generated in parallel on 96 core CPU instance on AWS
- Built **weakly supervised model** to disambiguate products into entities (Milk, Bread, Soap) and modifiers (Skim, Rye, Body vs Hand). Model improves search precision and reduces recall, significantly reduces the number of secondary searches members need to perform to find correct product
- Model uses historically converted search queries to determine the entity and modifiers for products

Data Analyst — June 2018 – April 2019
Shipt, Inc — *Birmingham, AL*

- Built **anomaly detection system** tracking ratio of order price to cost. Margins are dynamic and cyclic, anomalous margin indicates incorrect product pricing
- Decomposed aggregated series using local regression, Used GESD to identify anomalies
- Built secondary system to detect failures in ETL pipeline responsible to update product prices on platform
- Both systems ran bi-hourly and generated failure reports accessed and consumed by the **finance team and data engineering team**
- Caught over **30+ pricing failures** across retailers in a period of 8 months
- For non-partner retailers built **price identification system**
- Correctly matched products were determined using deviation of scraped price from historical price in catalog

Technical Skills

Languages: Python, SQL (Postgres/Snowflake), R
Deployment Frameworks: GCP, Airflow, Kubernetes, AWS batch, EC2, MLflow, Docker, Kubernetes
Databases and Warehouse: Snowflake, Postgres, Redis
Libraries: pandas, numpy, tensorflow, scikit-learn, pytorch, huggingface, caret, ggplot, dplyr.