Div Dasani

281.701.8058 · divdasani@gmail.com · divdasani.com

SKILLS

Languages: Python (PySpark, NumPy, Pandas, Flask), Scala, Java, SQL

Tools: AWS (Kinesis, EMR, Redis, EC2, S3), Databricks (Apache Spark, Delta Lake), Elasticsearch

Domain Expertise: Recommendations, Ranking, Personalization, ML Infrastructure

WORK EXPERIENCE

Discovery Bellevue, WA
Machine Learning Engineer- Recommendations April 2021 – Present

- Building (micro-batch) event ingestion system for our ML platform which materializes features from interaction data and publishes to online and offline feature stores (Kinesis, Spark Streaming, DynamoDB, Redis, Feast)
- Engineered A/B testing platform to power online experimentation for recommendations rails (Scala, EMR)
- Refactored data pipeline code to be configuration driven, allowing for logic to be ported to international regions

Scribd San Francisco, CA
Machine Learning Engineer- Recommendations March 2020 – April 2021

- Built online embedding-based retrieval and reranking platform with Elasticsearch to serve recs in real-time:
    - The system serves all homepage traffic, handling ~50rps with p95 latency <60ms
    - This work was part of the Personalization project and A/B test, which saw statistically significant ($p<0.01$) increases in CTR and read time for Scribd subscribers
- Worked on a faceted-search EBR system to power multiple recs services and scale for large (>10^8) corpora (Vespa)

Facebook Seattle, WA
Data Science Intern- Infrastructure September 2019 – November 2019

- Created model to estimate CPU utilization of Facebook live video transcoding requests using encoded parameters
- Developed dynamic slotting dispatcher to allocate compute resources to requests, increasing server efficiency by 35%

Roku Los Gatos, CA
Software Engineering Intern- Recommendations June 2019 – August 2019

- Built *More Like This* feature for The Roku Channel, which uses matrix factorization to personalize recommendations
- Employed Spark and AWS Redis to build data pipelines, capable of handling user requests with <10ms latency

TECHNICAL PROJECTS

Cross-Corpus Recommendations (Discovery Hack Week) January 2022 – January 2022

- Leveraged Gracenote metadata to prototype ML system capable of recommending D+ shows from a non-D+ query
- Achieved finalist status and presented demo to CTO and other executive members (TensorFlow, Docker, EC2)

Netflix Workshop on Personalization, Recommendation, and Search June 2021 – June 2021

- Presented a poster about building an embedding-based retrieval system to serve recommendations with Elasticsearch
- Answered questions from and networked with machine learning practitioners across various industries

EDUCATION

Northwestern University, McCormick School of Engineering Evanston, IL
Master of Science in Electrical Engineering September 2016 – March 2020
Bachelor of Arts in Statistics, Mathematical Methods in the Social Sciences September 2016 – March 2020