Josh McGrath

ML Engineer

I'm an engineer with multiple years of experience in startups, and at every level of the stack including machine learning. I have some academic background in using machine learning to solve data quality and systems problems.

EXPERIENCE

Founding Engineer, Distyl AI, San Francisco — April 2023 - Present

- Building RAG based applications with OpenAI models and custom IR systems

Machine Learning Engineer, Snorkel AI, San Francisco — August 2021 - April 2023

- Co-designed and built a monitoring product with a team of four

- Developed new unsupervised error analysis and detection tools deployed in Snorkel Flow

- Worked directly with Snorkel customers as a machine learning engineer in the field

Machine Learning Engineer, Apple, Seattle — April 2020 - August 2021

- Scaled the Inductiv engine to be Apple's knowledge graph platform, adding new workflow orchestration tools

- Trained and deployed entity linking models for graph construction

Founding Engineer, Inductiv (acquired), Waterloo, ON — April 2019 - March 2020

- Built a system using research done at Waterloo and Wisconsin to build a state of the art data quality system for entity linking, data cleaning, and error detection

- Designed and built a suite of statistical tools and pipelines in python for approximate dependency mining that was used to provision generative models over structured data

Academic Research Staff, University of Wisconsin, Madison — Dec 2017 - April 2019

- Researched theoretical and empirical aspects of data cleaning which resulted in two publications in 2019 at UAI (proof editor) and SIGMOD (co-author)

- Built research software for training generative and discriminative models to enable few-shot learning for error detection

- Worked on DARPA ASKE project for automatic knowledge graph construction over scientific documents, building information pipelines to ingest millions of documents using vision and text models

EDUCATION

University Of Wisconsin, Madison — BSc Mathematics and Computer Science, 2018

University of Waterloo, Ontario — PhD Computer Science, DNF

PUBLICATIONS

HoloDetect: Few-Shot Learning for Error Detection — SIGMOD 2019

Fine-Grained Object Detection over Scientific Document Images with Region Embeddings — ArXiv 2019

206-602-4029

joshuawmcgrath@gmail.com