Matt **Krzus**

Applied Machine Learning Scientist · Professional Goofball · Solver of Problems

1253 W Melrose St. Chicago, Illinois

(+1) 630-258-8799 | mkrzus3@gmail.com | www.mattkrzus.com | KayneWest | mattkrzus

"'You miss 100p of the shots you dont take - Wayne Gretzky' -Michael Scott."

Summary

Machine Learning Scientist with over 10+ years experience in the field. In the past 6 years, I've worked with over 60+ vastly different customers on every imaginable type of Machine Learning problem from anomaly detection on oil rigs, text summarization for Big Pharma, "Amazon-Go" style people tracking, and even synthetic data generation using 3D scanning. My main focus over the past few years has been edge-based applied ML/Computer Vision and with tons and tons of embeddings schemes (I've never met something I couldn't embed). Given the recent rise in popularity of Large Language Models and Generative AI, I have an absolute ton of experince in the deployment of NLP/Language Models and synthetic data generation using real 3D based models (along with some fun Stable Diffusion stuff, I have a few blogs posted in the writing section on NLP topics/exploration). For fun, I dable in Statistical Arbitrage and Quant Finance execution strategies. Outside of all things code, I'm a grappling/wrestling/jiu-jitsu freak and absolutely love spending time on the mats when I'm not hanging out with my family.

Work Experience

Amazon Web Services

Chicago, Illinois

Sr Data Scientist — *Nov. 2018 - Present*

- IC on a machine learning team in the IoT/AWS-Industrials division. Hiring and mentoring a team of 15 Applied Scientists. Leading the Engineering Standards proserv initiative
- Scalable C++ based multi-camera 3D tracking system using ROS for Amazon Fresh
- Player tracking using CenterTrack, camera pose/homography reconstruction using SLAM methods, ground plane detection for the NFL using MeTRAbs pose estimator, and PoEM (Human POse EMbedding) View Invariant Probalistic Embeddings for Human Pose for injury query/detection for the NFL
- Synthetic data creation utilizing 3D scanning and Blender that helped propel SageMaker Ground Truth Synthetics -blog post here
- TrueLies: Novel motif/time-series detection algorithms utilizing MASS/Matrix Profiles and Dynamic Time Warping for 'period of time' similarity/embeddings
- Inspectron: Rules-based visual QA framework utilizing local features, SuperGlue, semantic segmentation, and a custom rules-engine for manufacturing deployed at a major car manufacturer
- Pizza-Pizza: multi-model-serving framework for massive image-similarity queries using FAISS-cpu: detecting objects using ShuffleNet-YOLOv3, embedding objects using ResNet, and running similarity queries via FAISS for a major beverage producer
- Adverse text event detection system for a major Pharma company utilizing BERT embeddings;
- Vile seal detector using segmentation for trailer transportable Vaccine Pod meant for deployment to conflict zones for a major pharma company;
- Deploying TrueLies similarity system to a major US Energy provider;
- Hand-Written tattoo recognition system with a customer-specific probabilistic post-processor using the pizza-pizza framework for a major pork producer. Work published in the 2021 Amazon Machine Learning Conference
- Return prediction via pytorch deep-ar, avellaneda stoikov market-making bot, and portmanteau criterion based statistical-arbitrage bot for an aws trading-club
- With over 60+ projects during my time, only my favorites are listed as there are simply too many projects to include, however, the bullet below will list some of the skills used:
- Smaller/other projects: clustering, customer segmentation, object classification, object detection; instance and semantic segmentation; recommenders via embeddings for all kinds of features; keypoints/local features for SLAM; boosted models for text; mlops; tvm/sagemaker neo;

Amazon Web Services

Chicago, Illinois

Data Scientist — *Apr. 2017 - Nov 2018*

- Stack-Detector: poker chip detection and segmentation for a major casino
- Customer Response classification for a major car manufacturer - architecture blog here
- Device failure detection using SMOTE and XGBoost for an IoT partner
- As one of the first 4 data scientists in my org, I was responsible for conducting interviews for building a team of 100 new scientists (thus have more experience than I wish in the art of team building and interviewing.)
- Master Chef: computer vision state-machine involving object/employee tracking/detection, and inventory management of the back-of-kitchen events for a major fast-food restaurant. Tracking employee actions, monitoring major state changes in inventory and cooking positions. Much of this was done through the generous collaboration from our AmazonGo colleagues
- Master Chef cont... due to my initial computer-vision state-machine, our internal aws team won the invite-only competition and won the business of the fast-foot provider

mattkrzus.com

Chicago, Illinois

Consulting — *Jun. 2013 - Present*

- Most recently a beer-recognition system for a client-side app: Building the api, scaling compute, rolling deployments, developing and training the models, etc. Bunch of random client work ranging from hiring to scoping to deploying

Shoppertrak

Chicago, Illinois

Lead Machine Learning Scientist — *Jul. 2015 - Apr. 2017*

- Originally I was brought into this role to help build out ShopperTrack's customer insights tool. I transitioned into a pure CV role working with the hardware team to build methods to collect relevant visual data.
- In-house stereo cams to track blobs coming in and out of locations using handcrafted local features.
- Worked towards building a new camera capable of using modern methods (for the time) to track customers inside stores. However, we were acquired by Tyco – and then Tyco was acquired by Johnson Controls – and our team became a consulting team within Johnson Controls.
- Computer Vision using OpenCV C++, Stereo/camera homography, blob detection, deployment of TensorFace, customer entry/exit modeling using XGBoost/RandomForests/LSTMs

Kaplan, Inc.

Chicago, Illinois

Lead Machine Learning Scientist — *Jan. 2013 - Jun. 2015*

- Kaplan decided to build a proprietary OCR software with the goal of trying to sell it to compete with available products like Adobe, etc. This was great for me because it was my first introduction into serious Computer Vision, so I learned an absolute boatload
- Improvement on the ocropus/ocropy libraries for deployment. Basically the deployment of the seminal blog by danvk.org

Kaplan, Inc

Chicago, Illinois

Machine Learning Scientist — *Oct. 2013 - Jan. 2013*

- NLP on Hive/Pig/Hadoop. Constructing a knowledge graph in Neo4j
- Constructing a transcript summarizer using word2vec/doc2vec

Edo Interactive

Chicago, Illinois

Data Scientist — *Oct. 2012 - Oct. 2013*

- In retrospect, we ran unnecessarily complicated hadoop jobs on nodes that were too slow for what we needed them for, but this was a great place to learn deep NLP techniques. This job was mainly focused on using bayes methods for customer segmentation and nlp methods for extracting location of purchase data.

Advito Consulting

Chicago, Illinois

Data Analyst — *Aug. 2011 - Oct. 2012*

- At this company I taught myself python so I wouldn't have to use excel. Oddly, this job probably was the most pivotal job in my career.

Schneider Logistics

Chicago, Illinois

Account Executive — *Feb. 2011 - Oct. 2011*

- Post-GFC, no one was hiring so I did what I had to do. All Cold calling and sales. Oddly, the sales shtick has utility.

Village of Orland Park

Orland Park, Illinois

Sewer Cleaner Extraordinare — *Summer. 2006 - Summer. 2010*

- Quite literally cleaned sewers with the Streets and Sanitation Crew. Basically if The Office meets Parks and Rec meets Dirty Jobs. Disgustingly gross, but absolutely loved every minute of it

Skills

Machine Learning	Computer Vision, Language Models, Time-Series, Anomaly Detection, Embedding Systems, Synth Data, Edge
Programming	Python, C++, SQL
Major Libraries	Pytorch, OpenCV, XGBoost, FAISS, Numpy/Scipy, Cython, TVM, TensorFlow, MXNet, Sklearn, Pandas, SkOpt, AWS, CVXPy, Blender, Pytorch-C++, Jit, Numba, Eigen, ROS-py/c++, OpenPTrack, PCL
MLOps	SageMaker, Hydra, AirFlow
AWS	SageMaker, Ground Truth, LookoutForVision, Monitron, Panorama, EC2, CloudFormation, S3, Route53, IoT Core, GreenGrass, CodeCommit, DynamoDB, Redshift, ECR/S, Batch, Athena, Glue, Kinesis

Certificates

2023 **AWS Machine Learning - Specialty,** Amazon Web Services (AWS)

2020 **AWS Certified Solutions Architect – Associate,** Amazon Web Services (AWS)

2020 **Brazilian Jiu Jitsu Brown Belt,** Chicago Mixed Martial Arts — *cringiest cert ever*

2019 **Certified Good Dad,** My Wife

Presentation

Amazon Machine Learning Conference

Virtual

Presenter — *Dec. 20221*

- Presented a paper called An Edge Computing Workflow for Tattoo Recognition
- Discussed the creation of a novel deep network and deplopyment infrastructure
- dropbox download link

Writing

Create synthetic data for computer vision pipelines on AWS

AWS ML Blog *2022*

https://aws.amazon.com/blogs/machine-learning/create-synthetic-data-for-computer-vision-pipelines-on-aws/

Text Classification with Gluon on Amazon SageMaker and AWS Batch

AWS ML Blog *2018*

https://aws.amazon.com/blogs/machine-learning/text-classification-with-gluon-on-amazon-sagemaker-and-aws-batch/

Whiskey Embeddings

Personal Blog *2015*

https://www.mattkrzus.com/posts/whiskeyembedding/

Facial Recognition on a TX1

Personal Blog *2016*

https://www.mattkrzus.com/posts/face/

Bourbon Secondary

Personal Blog *2018*

https://www.mattkrzus.com/posts/bourbon_secondary/

What I Think is Wrong with Whiskey Reviews

Personal Blog *2019*

https://www.mattkrzus.com/posts/whiskey_reviews/

An Edge Computing Workflow for Tattoo Recognition

Amazon Machine Learning Conference *2021*

dropbox download link

Education

UIUC(Univerity of Illinois Chicago)

Chicago, Illinois

B.A. in Applied Psychology *Aug. 2006 - May. 2010*

- Became a phenomal pool player. Shout out to Rose's Lounge