grazel98@gmail.com
(856) 602-2151

Kevin M. Grazel

graxel.github.io
linkedin.com/in/kgrazel

Education

Rowan University

B.S. in Chemical Engineering

Platforms and Environments

Databricks AWS Azure Google Cloud Linux
MLFlow Docker Jupyter Notebooks PyInstaller

Languages

Python R SQL Bash Markdown

Packages and Frameworks

Spark TensorFlow Keras Pandas NumPy Scikit-learn
PyTorch Plotly XGBoost Matplotlib Seaborn
FastAPI Flask Hadoop Apache NiFi

Experience

Senior Machine Learning Engineer, Cybersecurity — CVSHealth   July 2024 – Present

Toolkit: Databricks, Spark, Python, LLMs, Variational Autoencoders, Transformers, SQL, Deep Learning, Embeddings, MLOps

- Led the design and deployment of robust, production-grade ML systems to detect and block botnet web traffic in real time, training on 2TB of data, demonstrating expertise in building scalable, maintainable ML pipelines.
- Developed and optimized a denoising variational autoencoder for anomaly detection, applying advanced deep learning architectures and rigorous unit and integration testing to ensure production reliability.
- Engineered custom tokenization and transformer-based embedding pipelines to convert sparse web traffic data into dense, model-ready tensors, reflecting hands-on experience with DSP-like data processing.
- Championed MLOps best practices, automating model deployment, monitoring, and retraining workflows for continuous production quality.
- Collaborated cross-functionally with engineering and product teams to integrate ML features into production systems, supporting ongoing feature development and bug triage.

AI Engineer — Parker-Hannifin   February 2024 – July 2024

Toolkit: Python, Azure, LLMs, Autogen, RAG Systems, SQL, MongoDB, Vector Databases, Embeddings, MLOps, NLP

- Led end-to-end development of a retrieval-augmented (RAG) chatbot platform, architecting Python-based backend systems for scalable, production-ready deployment on Azure.
- Integrated ML models and MongoDB vector database systems into enterprise environments.
- Instituted CI/CD automation and code quality standards, mentoring junior engineers and fostering a culture of maintainable, high-quality software development.

Consulting Work & Side Projects   August 2023 – February 2024

Toolkit: Python, AWS, LLMs, SQL, MongoDB, Vector Databases, Embeddings, MLOps, PyInstaller

- Designed and hosted a production ML pipeline for real-time job recommendation, leveraging OpenAI embeddings and vector DBs for fast, relevant search, demonstrating rapid prototyping and scaling from research to production.
- Automated document processing for industrial clients, showcasing experience with data ingestion, transformation, and ML-based decision making.

Machine Learning Engineer — PlanetX, LLC   September 2022 – August 2023

Toolkit: Python, SQL, AWS, FastAPI, Selenium, LLMs, Recommender Systems, MLOps, NLP, Deep Learning, Web Scraping

- Led project strategies as data science lead, ensuring quality, fostering teamwork, and driving accountability for success
- Architected and deployed multi-factor product recommendation engines using Python-based FastAPI on AWS.
- Tripled the number of products available on retail platform using semantic NLP to improve record linkage.
- Built OpenAI ChatGPT integration for environmental consultant chatbot LLM and to enhance product data pipeline.
- Trained deep learning models (CNN BILSTM, Bert) to categorize products and improve user shopping experience.
- Utilized advanced SQL queries to investigate core statistics and overhaul core categorization and peer scoring models.
- Developed web scrapers with Python and Selenium to source alternative data and expand product offering..
- Worked closely with Product and Engineering teams to identify and fix anomalies and to address user UX feedback.

Data Scientist — MUFG Investor Services   November 2019 – September 2022

Toolkit: Python, Pandas, Scikit Learn, SQL, AWS, MLOps, Spark, XGBoost, Forecasting, ETL, dbt, NLP, PowerBI

- Built Python ML training suite to rapidly train and evaluate hundreds of models to forecast investor capital movements.
- Tuned LightGBM model via Bayesian optimization, achieving a 5.0+ lift in targeting users likely to redeem investments.
- Mentored junior analysts in data wrangling and machine learning best practices.
- Led development of backend ETL data architecture (dbt) and frontend PowerBI UI of new ESG reporting product.
- Developed a system to automatically extract information from invoices to reduce invoice processing workload.
- Wrote Python scripts to parse invoice PDF files and trained NLP SpaCy NER models to process text.
- Planned product roadmap toward further business cases given MUFG's unique position as a fund administrator.
- Expedited Ops Team deduplication project by 10x using NLP fuzzy matching techniques to identify duplicate accounts.

Data Science Fellow — Metis   July 2018 – December 2018

Toolkit: Python, Pandas, Scikit Learn, SQL, AWS, MLOps, Spark, XGBoost, Forecasting, ETL, dbt, NLP, PowerBI

- Developed multiple end-to-end ML projects, including time series forecasting and NLP topic modeling, with a focus on reproducibility and code quality.

Senior Risk Consultant — Process Risk, LLC   April 2016 – November 2019

- Automated critical business processes with custom software, reducing project cycle times by 50% and demonstrating commitment to engineering efficiency and quality.
- Led teams of three to ten engineers and operators, serving as project consultant and client point of contact.

Validation Engineer — Bristol Myers Squibb   September 2015 – April 2016

- Expedited validation procedures by developing automation methods in Visual Basic.
- Completed validation of cryogenic freezers in half the budgeted time.
- Became validation lead after training new team members on validation equipment.
- Validated and certified calibration of critical climate-control units; focus on pharma, biotech, and life sciences.

Project Engineer — Grünenthal   June 2014 – September 2015

- Designed, implemented, and validated EHS (Environment, Health, and Safety) projects in pharmaceutical plant.
- Oversaw design, installation, and integration of a climatic test chamber to assess pharmaceutical product stability.
- Projects include plant ventilation and filter upgrade, explosion-proofing production and storage areas, water purity monitor installation, pneumatic system upgrade.

Field Service Engineer — Analitica Weisser   June 2013 – June 2014

- Provided in-field service and support for high precision chemical instruments.
- Consulted with laboratory directors on strategy, optimization, and analytical insights.
- Coached and trained new engineers for solo field service.

Continuous Improvement Engineer — SQM   June 2011 – February 2013

- Used SQL to query historical ion concentration data and developed a simulation model of vat leaching iodine extraction.
- Conducted studies in R with survey data to characterize behavior of caliche ore heap leaching process.
- Developed gPROMS simulations to model precipitation of salts and ion concentrations in evaporation ponds.

Consulting Work and Research Projects

Refinery Maintenance Dashboard — A consulting job for an oil refinery I'd worked with previously. Small accelerometer internet-of-things (IoT) sensors were used to monitor vibration levels of pumps and compressors at various refinery units. The data was logged to a Snowflake SQL database hosted on AWS. I developed a dashboard in PowerBI that would connect to Snowflake and prioritize maintenance based on planned unit downtime schedules and vibration level history.

CampOut — Displays campsites available for reservation on an interactive map. Behind the scenes, a web scraper running on an AWS EC2 server continuously gathers the latest campsite availability data.

WhaleWatch — A personal research project that monitors the SEC's EDGAR database for declared purchases and sales of stock by company board members. I use time series analysis to determine if a company's stock price might warrant a closer look by watching for board members buying or selling large amounts of corporate stock.

Some buzzwords and keywords for the ATS scanners: Big Data Machine Learning Artificial Intelligence (AI) Deep Learning Data Mining Predictive Analytics Natural Language Processing (NLP) Neural Networks Algorithm Data Visualization Cloud Computing Internet of Things (IoT) Business Intelligence (BI) Data Engineering Data Lake Data Warehouse Feature Engineering Regression Classification Clustering Dimensionality Reduction Anomaly Detection Reinforcement Learning Supervised Learning Unsupervised Learning Semi-supervised Learning Ensemble Learning Decision Trees Random Forest Gradient Boosting Support Vector Machines (SVM) K-Nearest Neighbors (KNN) Recommender Systems Time Series Analysis Sentiment Analysis Text Mining Data Wrangling Data Cleansing Data Preprocessing Data Normalization Data Augmentation Overfitting Underfitting Cross-Validation Hyperparameter Tuning Bias-Variance Tradeoff Feature Selection A/B Testing Ensemble Methods Bayesian Methods