Andrew Ingalls

Kansas City, MO | LinkedIn | 508-517-0139 | ingallsandrew0@gmail.com

WORK EXPERIENCE

A Place for Mom   **Kansas City, MO**
Machine Learning Engineer   *Nov 2022 - Present*

- Developed and deployed a multi-layered perceptron neural network for recommendation using Tensorflow, resulting in a 56% improvement in Mean Rank Error, and a 10% increase in conversion.
- Co-engineered the implementation of a recommendation engine using LightGBM Ranker, achieving an 8% increase in conversion rate.
- Contributed to the development of transcript summarization and information extraction techniques, leveraging large language models such as ChatGPT API to improve data comprehension and utility.
- Designed and implemented Spark SQL ETL pipeline from 20+ semi-structured sources, improving data quality for ML model ingestion and boosting accuracy by 15%.
- Conducted A/B testing power analysis for all algorithm experiments, enforcing proper methodology.
- Spearheaded adoption of a Bayesian A/B testing paradigm, significantly reducing decision timeframe.
- Enhanced a FastAPI Python application for machine learning serving, improving performance.
- Integrated New Relic monitoring to optimize application performance and proactively address issues.
- Automated ML pipelines using Pyspark and SQL, reducing technical debt and improving efficiency.
- Implemented CI/CD pipelines with GitHub Actions for machine learning API, leveraging AWS ECS to efficiently spin up new Docker containers for deployment.
- Conducted and presented ad-hoc analyses in Databricks to C-suite officers for $20M decisions.
- Managed ML clusters on Databricks for various projects, optimizing resource allocation.

Stowers Institute for Medical Research   **Kansas City, MO**
Associate Scientist II / Data Scientist   *July 2018 - Nov 2022*

- Developed an automated time series ETL pipeline and reporting tool using PowerBi, Python, Linux, PostgreSQL, Docker, and Airflow, resulting in a workload reduction of over 100 hours per year.
- Decreased company report crafting costs by 60+ hours per year by automating processes with SQL.
- Utilized polynomial regression modeling with scikit-learn to optimize lab specimen feeding costs, achieving an average cost reduction of 40%.
- Developed and deployed a XGBoost classifier for lab specimen feed size to optimize growth and breeding for age, sex, and subspecies.
- Engineered and implemented a time series forecasting model with XGBoost, effectively predicting salt dosing pump clogs through analysis of daily pump activation frequency. This model significantly reduced after-hours emergencies by 20%.
- Used Deep Learning to conduct behavioral tracking experiments with the package DeepLabCut.
- Improved water quality procedures through descriptive analysis of historic data, reducing employee workload by 300+ hours.
- Spearheaded an overhaul of the training system and created over 50 new training documents, which resulted in a direct increase in training opportunities and mentorship for 7 employees.
- Co-authored 3 scientific books focused on model organism research.

Trifecta Ecosystems, Inc.   **Meriden, CT**
Chief Operations & Research Officer   *June 2016 -2018*

- Managed finances, inventory, and projections including $250,000 in annual operating costs.
- Developed an in-situ sensor suite and Angular web app for real-time analysis of water quality metrics.
- Engineered 25+ recirculating aquaponics systems including a 3,000 sqft. facility worth over $150,000.

PROJECT EXPERIENCE

Real-Time Greenhouse Weather Monitor and Forecaster *Nov 2023*
- Developed an ESP-32S microcontroller system for transmitting environmental data to AWS IoT Core, focusing on efficient data handling and system stability.
- Implemented a real-time AWS data streaming pipeline using Kinesis, Lambda, and API Gateway, achieving 1-second update intervals for an S3 static web app.
- Utilized AWS SageMaker's DeepAR for real-time weather forecasting, predicting maximum temperature 30 minutes ahead to enable proactive greenhouse alerts.
- *Technologies Used*: AWS (IOT Core, Kinesis Data Streams, Firehose, Managed Service for Apache Flink, Gateway API, Sagemaker, S3, Glue, Lambda, DynamoDB), Python, Arduino (C++), HTML, JavaScript

Dungeons & Dragons Monster Stat Block Generator *May 2022*
- Scraped over 300 monster data entries using Selenium; wrangled complex data features
- Built, trained, and deployed an MLP neural network using Keras to generate a balanced monster stat block for inexperienced DMs with 85% accuracy.
- Automatically deployed to Heroku through Docker and GithubActions CI/CD pipeline
- *Technologies Used*: Python, Selenium, Pandas, Numpy, Tensorflow, Keras, Dash, Docker, AWS, Heroku, Docker, GitHubActions

EDUCATION

Kansas State University **Manhattan, KS**
Master of Science: Geology *2014*
Northeastern University **Boston, MA**
Bachelor of Science: Environmental Science *2012*

SKILLS & INTERESTS

Skills: Python (Numpy, Pandas, Scikit-learn, Tensorflow, Keras, Dash, Flask, Matplotlib, Seaborn, Plotly, BeautifulSoup, FastAPI, Selenium), Databricks, SQL (Spark, MySQL, PostgreSQL), PySpark, Linux, Git, PowerBI, AWS (Kinesis, EC2, S3, Glue, Sagemaker, Gateway API, CloudWatch, DynamoDB) Apache Flink, Docker, Airflow, Statistics, Bayesian Statistics, A/B Testing

Interests: Deep Learning, Traveling, Analytics, Python, Dungeons and Dragons, Carpentry, Gardening