Caroll Howard

Machine Learning Engineer

carollhwrd@gmail.com • New Bedford, MA, US

Summary

Highly motivated and results-driven Machine Learning Engineer with 6 years of experience in building and deploying machine learning models across various industries. Skilled in designing and developing custom algorithms and models to solve complex problems using various programming languages such as Python and R.

Expertise in feature engineering, data pre-processing, model selection, hyperparameter tuning, and evaluation techniques. Proficient in various machine learning frameworks such as TensorFlow, PyTorch, and scikit-learn. Experienced in natural language processing (NLP), computer vision, and time-series analysis.

Strong knowledge of statistics, mathematics, and probability theory. Able to efficiently manage large datasets and implement efficient data pipelines. Proficient in SQL and NoSQL databases such as MySQL, MongoDB, and Oracle.

Adept at collaborating with cross-functional teams to integrate machine learning solutions into products and services. Possess excellent communication skills and can effectively communicate complex technical concepts to non-technical stakeholders. Continuously learning and staying up-to-date with the latest advancements in machine learning and artificial intelligence.

EMPLOYEEMENT HISTORY

Innowise-Group

Remote

Machine Learning engineer

04/2021 - 04/2023

- Created and assisted with the development of 5 chatGPT-based custom service chatbots for Business Intelligence(BI) and other commercial purposes and deployment of them into mobile apps and web servers.
- Collaborated with product, engineering, and customer teams to design, build, improve and fine-tune ML models(Large Language Models) tailored to specific enterprise use cases.
- Developed and deployed a DL model for a financial services client, resulting in a 25% reduction in fraudulent transactions and saving the client $50k annually.
- Constructed 3 different deep neural networks and generated, new, realistic faces using GAN(Generative Adversarial Network) on PyTorch environment.

TechCrunch

Remote

Natural Language Processing expert

01/2019 - 04/2021

- Deployed 2 attention-based encoder-decoder models to check grammar error.
- Trained 2 Chinese-English machine translation systems that outperformed Google Translation using Transformers and some manually organizing models.
- Spearheaded transition from firebase to AWS, saving company $27000+ monthly and increasing load speeds by an average of 42%.
- Initiated a character-based transformer spelling correction model using Triton and ONNX; reduced inference latencies below 178ms at 35 req/s.
- Implemented authentication/authorization mechanism using Auth0 and AWS Cogito.
- Developed an NLP system that automatically classified 7.5K emails as spam or advertising mail using 3+ natural language processing methods.

Ample Insight

Machine Learning engineer

07/2017 - 01/2019

- Trained a CNN(Convolutional Neural Network) to analyze images of dogs and correctly identify their breeds and used transfer learning to increase performance by 5% and simplify the model.
- Analyzed and interpreted data for a retail client to identify trends and patterns for 2 years.
- Applied LSTM network to predict the pattern Assisted in creating a DL model that detects anomalies in the building process of a construction robot and machine reducing downtime by 83% & processed action measures to predict the state of machine reducing fix cost by 20%.
- Conceived and discovered a machine learning algorithm that detects deviant behavior in robots using SIFT, HOG, and 20+ other computer vision methods.
- Tracked the health of 15+ robots using React/Redux with NodeJS backend and Python scripts, which collected 1TB of data from its sensors.

Education

York University

Bachelor's Degree in Computer Science

04/2013 - 06/2017

WORK EXPERIENCE

The Replika

09/2022 - 12/2022

Reference: https://replika.ai
Responsibilities:

- prompt engineering for new version of Replika in the field of Science and Tech
- wrote core chatGPT API code
- assisted with data preparation for fine-tuning chatGPT

TECHNICAL SKILLS

Languages: C++ · Java · Python · TypeScript · HTML · JavsScript · Python · Bash Script · GraphQL · SQL

Platform: Amazon Web Services(AWS) · Docker · Kubernetes · Amazon EC2 · Google Cloud Platform(GCP) · Linux · Windows

Libraries:
Numpy · Pandas · sklearn · scipy · Keras · TFLite · XGBoost · LightGBM · onnx · Flask · selenium · Node.js · Bootstrap · PHP · React

Frameworks: Tensorflow · PyTorch · Django · Flask · Angular · Hadoop

Tools:
Microsoft Visual Studio · Semantic Kernel · AutoGPT · LangChain · Visual Studio Code · Git · AWS SageMaker · Comprehend · Azure · Hugging Face · Slack · GitHub · GitLab · Jenkins · Jupyter · NGINX · Jira · Prometheus · AWS CloudFormation · Amazon EKS · GPU · CUDA · dbt

Paradigms:
Deep Learning · Math · algebra · Probability · Statistics · Natural Language Processing · LLM(Transformer-based LM/GPT) · chatGPT · chatGPT API · Prompt Engineering · Computer Vision · GAN · Reinforcement Learning · Model Optimization · Continuous Integration(CI) · Continuous Delivery(CD) · Algorithms · Data Science · Data Structure · Data Visualization · Web Services · User Interface(UI) · Agile Software Development · DevOps · MLOps · automated testing · performance testing · ETL · data warehouse

Storage: MySQL · PostgreSQL · MongoDB · Redis · SQLite · AWS DynamoDB