SARAH CHEN

Senior Data Scientist & Machine Learning Engineer

Email: sarah.chen@email.com | Phone: +1 (555) 987-6543
LinkedIn: linkedin.com/in/sarahchen | GitHub: github.com/sarahchen
Portfolio: https://sarahchen.dev

PROFESSIONAL SUMMARY

Passionate data scientist with 6+ years of experience in machine learning, deep learning, and statistical modeling. Expertise in Python, TensorFlow, PyTorch, and cloud-based ML solutions. Proven track record of developing and deploying production-level AI systems that drive business value. Strong communicator with experience leading cross-functional teams and mentoring junior data scientists.

EDUCATION

Ph.D. in Computer Science - Machine Learning
Stanford University | Stanford, CA | 2016 - 2020
GPA: 3.92/4.0 | Graduated
Dissertation: 'Deep Reinforcement Learning for Autonomous Systems' Advisor: Dr. Andrew Ng

Master of Science in Data Science
University of California, Berkeley | Berkeley, CA | 2014 - 2016
GPA: 3.88/4.0 | Graduated
Concentration: Statistical Learning and Big Data Analytics

Bachelor of Science in Mathematics and Computer Science
Massachusetts Institute of Technology | Cambridge, MA | 2010 - 2014
GPA: 3.85/4.0 | Graduated with Honors
Minor: Cognitive Science

PROFESSIONAL EXPERIENCE

Senior Machine Learning Engineer | Tech Innovations Inc. | San Francisco, CA | Jan 2021 - Present
• Lead ML team of 8 engineers developing next-generation recommendation systems serving 50M+ users
• Architected and deployed deep learning models that improved click-through rates by 35% and revenue by $12M annually
• Implemented MLOps pipelines using Kubernetes, Docker, and Airflow for automated model training and deployment
• Designed A/B testing framework that reduced experiment cycle time from 4 weeks to 3 days
• Mentored 5 junior data scientists and conducted weekly technical workshops on advanced ML techniques

Data Scientist | CloudTech Solutions | Seattle, WA | Jun 2018 - Dec 2020
• Developed predictive models for customer churn reduction, achieving 92% accuracy and saving $8M in retention costs
• Built real-time fraud detection system using ensemble methods and neural networks, reducing fraud by 45%
• Created automated data pipelines processing 10TB+ daily using Apache Spark and AWS EMR

• Collaborated with product teams to integrate ML models into production applications
• Published 3 research papers on deep learning applications in industry

Research Assistant | Stanford AI Lab | Stanford, CA | Sep 2016 - May 2018
• Conducted research on reinforcement learning algorithms for robotics and autonomous systems
• Co-authored 5 papers published in top-tier conferences (NeurIPS, ICML, CVPR)
• Developed novel deep Q-learning algorithm that improved sample efficiency by 60%
• Collaborated with industry partners on applied research projects

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, Java, C++, JavaScript, Scala, Julia
Machine Learning: TensorFlow, PyTorch, Scikit-learn, XGBoost, LightGBM, Keras, JAX
Deep Learning: CNNs, RNNs, LSTMs, Transformers, GANs, Reinforcement Learning, Transfer Learning
Big Data & Cloud: Apache Spark, Hadoop, AWS (SageMaker, EMR, S3, Lambda), Google Cloud Platform, Azure
MLOps & Tools: Docker, Kubernetes, Airflow, MLflow, Git, CI/CD, DVC
Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, D3.js
Databases: PostgreSQL, MySQL, MongoDB, Redis, Cassandra, Elasticsearch
NLP: BERT, GPT, Word2Vec, spaCy, NLTK, Hugging Face Transformers
Computer Vision: OpenCV, YOLO, ResNet, VGG, Image Segmentation, Object Detection

KEY PROJECTS

AI-Powered Content Recommendation Engine
Developed a hybrid recommendation system combining collaborative filtering and deep learning to personalize content for 50M+ users. Implemented using PyTorch, deployed on AWS with auto-scaling capabilities. Achieved 40% improvement in user engagement and 25% increase in time-on-platform. Technologies: Python, PyTorch, Redis, AWS SageMaker, Apache Kafka

Real-Time Fraud Detection System
Built end-to-end fraud detection pipeline processing millions of transactions per day. Utilized ensemble methods (Random Forest, XGBoost) and LSTM networks for sequence modeling. Reduced false positives by 60% while maintaining 99.5% fraud detection rate. Technologies: Python, TensorFlow, Apache Spark, Kafka, PostgreSQL, Docker

Natural Language Processing Platform
Created NLP platform for sentiment analysis, named entity recognition, and text classification. Fine-tuned BERT and GPT models for domain-specific tasks. Deployed as RESTful API serving 10M+ requests daily. Technologies: Python, Transformers, FastAPI, Docker, Kubernetes, MongoDB

Computer Vision for Medical Image Analysis
Developed deep learning models for automated diagnosis from medical images (X-rays, MRIs). Achieved 94% accuracy in disease classification, assisting radiologists in faster diagnosis. Collaborated with Stanford Medical School for data collection and validation. Technologies: Python, TensorFlow, OpenCV, ResNet, U-Net, DICOM

LEADERSHIP & EXTRACURRICULAR ACTIVITIES

• **President, Women in Data Science (WiDS) Chapter** - Organized monthly workshops and networking events for 200+ members (2019-2021)
• **Lead Mentor, Code for Good Initiative** - Mentored 15 underrepresented students in data science

and ML (2020-Present)
• **Speaker, PyData Conference** - Delivered talks on production ML systems and best practices (2019, 2021, 2022)
• **Volunteer, DataKind** - Applied data science skills to solve social impact problems (2018-Present)
• **Organizer, Bay Area Machine Learning Meetup** - Host monthly technical discussions with 500+ members (2020-Present)
• **Kaggle Competitions Master** - Top 1% globally, 3 gold medals, 5 silver medals
• **Open Source Contributor** - Active contributor to Scikit-learn, TensorFlow, and PyTorch projects

PUBLICATIONS & AWARDS

• Chen, S. et al. (2020). 'Deep Reinforcement Learning for Autonomous Navigation.' NeurIPS.
• Chen, S. et al. (2019). 'Efficient Transfer Learning in Computer Vision.' CVPR.
• Chen, S. et al. (2018). 'Scalable Machine Learning on Big Data.' ICML.
• **Best Paper Award**, International Conference on Machine Learning (ICML 2019)
• **Outstanding Graduate Researcher Award**, Stanford University (2019)
• **Grace Hopper Scholarship** for Women in Computing (2017)

CERTIFICATIONS

• AWS Certified Machine Learning - Specialty
• Google Cloud Professional Machine Learning Engineer
• Deep Learning Specialization - Coursera (Andrew Ng)
• TensorFlow Developer Certificate

CAREER ASPIRATIONS

Seeking opportunities to lead innovative AI/ML initiatives that create meaningful impact at scale. Passionate about developing cutting-edge machine learning systems, mentoring the next generation of data scientists, and pushing the boundaries of what's possible with artificial intelligence. Particularly interested in applications of deep learning in healthcare, autonomous systems, and ethical AI. Goal is to become a Chief AI Officer or VP of Machine Learning at a technology company driving positive change.