

MARCUS JOHNSON

Senior Data Scientist & Machine Learning Engineer

✉ marcus.johnson@email.com | ☎ (555) 987-6543 | 💼 linkedin.com/in/marcusjohnson
| 📍 New York, NY | 🐙 github.com/marcusj

EXECUTIVE SUMMARY

Data-driven professional with 8+ years of experience in machine learning, statistical analysis, and big data processing. Expert in developing predictive models that drive business decisions and optimize operations. Led cross-functional teams to deliver AI solutions resulting in \$10M+ cost savings and revenue generation.

CORE COMPETENCIES

Machine Learning: Deep Learning, Neural Networks, Random Forest, SVM, Clustering, NLP, Computer Vision **Programming:** Python, R, SQL, Scala, Java, MATLAB **ML Frameworks:** TensorFlow, PyTorch, Scikit-learn, Keras, XGBoost, LightGBM **Big Data Technologies:** Apache Spark, Hadoop, Kafka, Airflow, Databricks **Cloud Platforms:** AWS (SageMaker, EMR, Redshift), Google Cloud (BigQuery, Vertex AI), Azure ML **Visualization:** Tableau, PowerBI, Matplotlib, Seaborn, Plotly, D3.js **Databases:** PostgreSQL, MySQL, MongoDB, Cassandra, Snowflake

PROFESSIONAL EXPERIENCE

Senior Data Scientist | *FinanceFirst Bank* | Apr 2020 - Present

- Lead a team of 6 data scientists developing ML models for fraud detection and risk assessment

- Built deep learning model reducing false positive fraud alerts by 45%
- Developed customer churn prediction model with 92% accuracy, preventing \$3M annual revenue loss
- Created automated feature engineering pipeline processing 100GB+ daily transaction data
- Collaborated with business stakeholders to translate requirements into technical solutions

Major Projects:

- **Credit Risk Model:** Developed ensemble model improving loan approval accuracy by 25%
- **Real-time Fraud Detection:** Implemented streaming ML pipeline detecting fraudulent transactions within 100ms
- **Customer Segmentation:** Built unsupervised learning system identifying 12 distinct customer personas

Data Scientist | *RetailMax Corporation* | Sep 2018 - Mar 2020

- Designed recommendation engine increasing online sales conversion by 28%
- Built demand forecasting models optimizing inventory management across 500+ stores
- Developed pricing optimization algorithms resulting in 15% margin improvement
- Implemented A/B testing framework for data-driven product decisions
- Created automated reporting dashboards for executive leadership

Key Achievements:

- **Recommendation System:** Collaborative filtering model serving 50M+ daily recommendations
- **Supply Chain Optimization:** ML model reducing inventory costs by \$5M annually
- **Dynamic Pricing:** Real-time pricing algorithm increasing revenue by 12%

Machine Learning Engineer | *TechStartup Analytics* | Jan 2017 - Aug 2018

- Developed computer vision models for automated quality control in manufacturing
- Built NLP pipeline for sentiment analysis of customer feedback across social media

- Implemented MLOps practices including model versioning, monitoring, and deployment
- Created data pipelines processing structured and unstructured data from multiple sources

Data Analyst | *ConsultingPro* | Jun 2016 - Dec 2016

- Performed statistical analysis and created visualizations for client presentations
- Built predictive models for marketing campaign optimization
- Conducted market research and competitive analysis using data science techniques

EDUCATION

Master of Science in Data Science | *Columbia University* | 2014 - 2016

- **GPA:** 3.9/4.0
- **Thesis:** "Deep Learning Approaches for Time Series Forecasting in Financial Markets"
- **Relevant Coursework:** Machine Learning, Statistical Learning, Big Data Analytics, Natural Language Processing

Bachelor of Science in Mathematics | *MIT* | 2010 - 2014

- **GPA:** 3.7/4.0, Magna Cum Laude
- **Minor:** Computer Science
- **Honors:** Phi Beta Kappa Honor Society

CERTIFICATIONS & TRAINING

- **AWS Certified Machine Learning - Specialty** (2023)
- **Google Cloud Professional Machine Learning Engineer** (2022)
- **TensorFlow Developer Certificate** (2021)
- **Databricks Certified Associate Developer for Apache Spark** (2020)
- **Microsoft Azure AI Engineer Associate** (2019)

PUBLICATIONS & RESEARCH

Peer-Reviewed Papers:

1. "Advanced Ensemble Methods for Financial Fraud Detection" - *Journal of Machine Learning Research* (2023)
2. "Deep Learning for Real-time Customer Behavior Prediction" - *ACM Conference on Data Mining* (2022)
3. "Optimizing Supply Chain Networks using Reinforcement Learning" - *Operations Research Letters* (2021)

Conference Presentations:

- **PyData NYC 2023:** "Scaling Machine Learning Models in Production"
- **Strata Data Conference 2022:** "Building Robust ML Pipelines for Financial Services"
- **KDD 2021:** "Feature Engineering Automation in Large-scale Systems"

NOTABLE PROJECTS

AI-Powered Investment Portfolio Optimizer | 2023

- Developed reinforcement learning model for automated portfolio rebalancing
- Integrated multiple data sources including market data, news sentiment, and economic indicators
- Achieved 18% better risk-adjusted returns compared to traditional methods
- **Tech Stack:** Python, TensorFlow, Apache Kafka, AWS SageMaker

Multi-modal Sentiment Analysis Platform | 2022

- Built NLP system analyzing text, images, and video content for brand sentiment
- Processed 1M+ social media posts daily with 94% accuracy
- Implemented real-time alerting system for reputation management
- **Tech Stack:** PyTorch, Transformers, OpenCV, Kubernetes, Redis

Predictive Maintenance System | 2021

- Created IoT-based ML system predicting equipment failures 30 days in advance
- Reduced unplanned downtime by 60% across manufacturing facilities
- Integrated sensor data, maintenance logs, and environmental factors
- **Tech Stack:** Apache Spark, TensorFlow, InfluxDB, Grafana

AWARDS & RECOGNITION

- **Data Scientist of the Year** - FinanceFirst Bank (2023)
- **Innovation Excellence Award** - RetailMax Corporation (2019)
- **Best Paper Award** - International Conference on Machine Learning Applications (2022)
- **Outstanding Graduate Student** - Columbia University (2016)

PROFESSIONAL MEMBERSHIPS

- **Association for Computing Machinery (ACM)** - Member since 2016
- **IEEE Computer Society** - Senior Member since 2019
- **American Statistical Association (ASA)** - Member since 2017

TECHNICAL BLOG & THOUGHT LEADERSHIP

- **Medium Blog:** 50+ articles on machine learning with 100K+ views
- **Kaggle:** Competitions Master with 15 gold medals
- **Open Source:** Contributor to TensorFlow, Scikit-learn, and Apache Spark
- **Speaking:** Regular speaker at data science conferences and meetups

LANGUAGES

- **English:** Native
- **Spanish:** Professional proficiency

- **French:** Conversational

INTERESTS & HOBBIES

- Competitive chess (USCF rating: 2100)
- Marathon running and triathlon
- Mentoring students in STEM programs
- Contributing to open-source ML projects