

MACHINE LEARNING DATA SCIENTIST

Over 20 years of high technology experience developing Enterprise quality software and database systems. Includes 6 years as a Data and Machine Learning Scientist in the Microsoft Cloud organization, 15+ years of data-driven system development and analysis, 10+ years as a Development Manager and Program Manager.

EXPERIENCE

Machine Learning Data Scientist

Microsoft Cloud 2013 – 2019: Senior Data Scientist, Statistics and Analytics Lead

Machine Learning Projects

- Designed, trained, and operationalized prediction models in support of the Microsoft (MS) Cloud Supply Chain. Performed all phases of the standard Machine Learning process flow: exploratory data analysis (EDA), data cleaning, feature engineering, model training/testing, accuracy evaluation, release to production, and visualization of results. Performed continuous monitoring, iterative improvement, and retraining of production models.
- Iteratively developed and evaluated numerous models to answer key business questions related to process cycle times, delivery dates, anomalies, and risks to on-time delivery. Improved overall date prediction accuracy by 300% from initial version over 2-year period.
- Used both supervised and unsupervised algorithms including Regression, Logistic Regression, Boosted Decision Trees, Random Forests, Fast Tree Regression and other ensemble methods, Naïve Bayes, Support Vector Machines (SVM), k-Nearest Neighbor (KNN), k-Means Clustering, Principal Component Analysis (PCA), and Time Series Analysis.
- Prototyped and evaluated text/Natural Language Processing (NLP) models for system logs, deep learning Neural Networks for image processing.
- Simulated end-to-end Cloud deployment process by various methods including Monte Carlo simulations, Queueing Systems, and the ProModel simulator tool. Used for throughput estimation, process optimization, and to drive improvements into upstream systems to create more accurate capacity plans.
- Drove efficiencies by automating model validation and best model selection. Automated hypothesis testing, and determination of best fit data distributions for over 20k scenarios.
- Developed using R, Python, MS SQL, MS Azure Machine Learning Studio (Azure ML), and other technologies described below.
- Conducted regular Executive Level project reviews, showing continuous improvement, roadmaps, and project plans for executive approval. Presented prediction algorithms, accuracy results, and impact analysis.
- Mentored junior Data Scientists and was mentored by more senior Data Scientists. Collaborated with Data and Software Engineers across the organization. Defined analytics requirements for data warehouses and transactional software systems.

Dashboards and Metrics

- Designed and produced the Quarterly Analyst Report used by Executive Level Management. Created manually initially, then automated as an interactive web site using R, Azure ML, C#, SQL. Created other dashboards and reports using MS Power BI.
- Established best statistics and metrics for our organization and helped inform Leadership Team on how to interpret and use them. Produced aggregated metrics with drill-down for different levels in the organization.

Database

- Extensive experience developing SQL database solutions. Designed and built databases, stored procedures, views, and jobs. Used in Data Science analysis and applications.
- Performed complex automated and ad-hoc quantitative analyses. Combined data from disparate sources, authored pipelines using SQL and Python based ETL frameworks.

- Improved organizational data quality by identifying issues and driving improvements into source data collection systems and ETL processes.

Leadership and Training

- Set statistics standards and conducted organization-wide statistics training series, raising the level of understanding and proficiency across the organization. Trained multiple teams in statistical analysis, machine learning algorithms, R programming, and Azure ML.
- Conducted technical data science presentations for audiences with over 1k attendees, including Executives, Data Scientists, and Developers. Conducted full day demos of ML projects and results at technology fair.
- Wrote White Papers on improving the Cloud Supply Chain using process simulation and Machine Learning.

Development Manager

Microsoft 2002 – 2008: Software Development Manager

- Managed teams of 5 – 30 direct reports in support of Microsoft Windows and IT. Teams included software, database, user interface, and report developers.
- Lead teams to successful on-schedule project completion. Lead through major technology transformations, migrations, and organizational transformations. Established team standards and best practices. Recruited, conducted performance reviews, managed career growth, and managed team budgets.

Program Manager

StanCorp Financial Group 2009 – 2012: Senior Program Manager

Microsoft 2008 – 2009: Principal Program Manager

- Successfully delivered multi-year, multi-phase portfolio of technically complex IT programs with budgets over \$15 million and teams of 30+ people. Leveraged Global Model and managed remote teams in India and China. Delivered significant business value incrementally with each phase using a combination of Agile and Waterfall methodologies.
- Drove design and delivery of enterprise-wide strategic Data Warehouse (DW), Master Data Management (MDM) and Business Intelligence (BI) platforms for B2B customer data using the Microsoft BI stack: SSIS, SSAS, SSRS, and SharePoint.

EDUCATION AND CERTIFICATIONS

Recent Certifications: [DataCamp](#)

[Machine Learning Scientist](#) 3/2020

[Data Scientist](#) 1/2020

Degrees: University of Texas at San Antonio

Master of Science, Mathematics with Statistics Concentration

Bachelor of Science, Mathematics

SKILLS, TOOLS, AND TECHNOLOGIES

- *Data Science Languages/Tools:* Python, R, Microsoft Azure Machine Learning Studio, Minitab
- *Python Packages:* Numpy, Pandas, Matplotlib, Seaborn, SciPy, Scikit-Learn (sklearn), TensorFlow, Keras, PyTorch, NLTK, spaCy, Scikit-Image (skimage), statsmodels, PySpark, Bokeh, Jupyter Notebooks/Lab
- *Databases:* MS SQL Server, MS Azure SQL, PostgreSQL, MySQL, SQL Server Analysis Services (SSAS, OLAP), SQL Server Integration Services (SSIS)
- *Visualization and Reporting Software:* Power BI, MS Excel, Tableau, ASP.NET, SQL Server Reporting Services (SSRS)
- *Other Software Development Languages/Tools:* ASP.NET Framework, Visual Studio, TFS, GIT, HTML/XML/XHTML, CSS, C#, VBScript, Java, JavaScript
- *Proficient in Microsoft Business Tools:* Visio, MS Project, MS Office suite
- *Some Experience With:* Azure Data Lake, Data Factory, Data Bricks, NoSQL

Previous professional experience available upon request.