

MOHAMED MOSAED

Senior Data Analyst | Data Scientist | AI & Analytics Solutions

mohamedmosaed405@gmail.com | 732-630-7495 | New Jersey, NJ | Portfolio: <https://mohamedmosaed.github.io/>

PROFESSIONAL SUMMARY

Results-driven Senior Data Analyst with 7+ years of experience leveraging advanced analytics, machine learning, and Generative AI to solve complex business challenges across insurance, healthcare, and financial services. Expertise in predictive modeling, statistical analysis, data engineering, and cloud-based analytics solutions using Python, SQL, PySpark, Databricks, and Azure. Proven success in building scalable data pipelines, deploying machine learning models, and developing AI-powered applications that improve efficiency, reduce operational costs, and accelerate decision-making. Adept at translating complex data into strategic insights and partnering with cross-functional stakeholders to drive measurable business outcomes and enterprise-wide data initiatives. Portfolio showcasing AI, machine learning, and analytics projects available at [Mohamed_Mosaed_Portfolio](#).

CORE SKILLS

Languages	Python, R, SQL, PySpark
ML Libraries	Scikit-learn, TensorFlow, PyTorch, LightGBM, XGBoost, NLTK, SpaCy
Data & ETL	Pandas, NumPy, SciPy, PySpark, Azure Data Lake, Databricks
Visualization	Tableau, Power BI, R Shiny, Matplotlib, Seaborn, Plotly
Cloud & DevOps	AWS (S3, CodePipeline, CodeBuild), Microsoft Azure, GitHub, JIRA
Databases	MySQL, MS SQL Server, NoSQL
Methodologies	Agile / Scrum, A/B Testing, Statistical Modeling, NLP

PROFESSIONAL EXPERIENCE

Senior Data Analyst | AXA XL, New York, NY - Sep 2023 – Present

- Designed and deployed end-to-end predictive models (classification, regression, survival analysis) using Python and Scikit-learn, improving risk scoring accuracy by 15%.
- Built scalable PySpark distributed pipelines on Databricks integrated with Azure Data Lake, cutting ETL run time by 30% for high-volume financial and claims datasets (50M+ records).
- Collaborated with actuaries and underwriters to deliver data-driven insights that reduced manual review time by 25%.
- Led migration of legacy reporting systems to Azure SQL and Databricks, reducing report generation time by 60% and improving query performance for 200+ business users.
- Developed complex SQL pipelines integrating 50M+ records from multiple internal and external sources, reducing data preparation time by 40% and improving reporting reliability.
- Created interactive Power BI and R Shiny dashboards used by 100+ users, reducing manual reporting effort by 35% and improving visibility into underwriting KPIs.
- Developed an anomaly detection platform using BERT embeddings and autoencoders to identify underwriting policy exceptions and high-risk transactions, reducing manual review effort and enabling proactive risk monitoring.
- Developed a Retrieval-Augmented Generation (RAG) solution using Python, OpenAI APIs, vector embeddings, and semantic search to analyze 500+ annual reports and business documents, reducing information retrieval time by 70% and enabling users to access critical insights in under 30 seconds.
- Championed best practices in data quality, reproducibility, and governance using Git, JIRA, and automated testing frameworks.
- Automated end-to-end Excel reporting by developing VBA macros that trigger a Python script to extract and load data from Databricks into scheduled reports, improving data accuracy, eliminating manual refresh cycles, saving 15+ hours per week, and eliminating manual refresh errors.

Data Analyst | UMass Chan Medical School, Worcester, MA - Jun 2023 – Sep 2023

- Cleaned, transformed, and analyzed clinical datasets containing 500K+ patient records, improving data quality and supporting multiple translational medicine studies.
- Applied ML models (classification, survival analysis) with Scikit-learn and inferential statistics (logistic regression, multivariate modeling) to evaluate treatment outcomes and population risk factors.
- Built Power BI and Plotly dashboards used by research teams to monitor 1,000+ patients across multiple clinical studies.
- Developed visualizations and summary tables for peer-reviewed publications and conference posters.

- Designed custom data visualizations using Python (Matplotlib, Seaborn, and Plotly) to present complex clinical findings in clear, publication-ready formats tailored to research and medical audiences.

Data Analyst | KPMG, New York, NY - Jan 2022 – Jun 2023

- Delivered automated Tableau and Power BI solutions for 10+ client engagements, reducing reporting effort by 30% and improving executive visibility into key performance metrics.
- Built and deployed machine learning models (classification, regression, clustering) that improved prediction accuracy by up to 20% across client-specific business use cases.
- Processed and analyzed 1M+ text records using NLP techniques like tokenization, lemmatization, and stemming using NLTK and SpaCy, enabling automated categorization and insight generation.
- Designed and deployed machine learning and deep learning solutions for NLP, computer vision, and forecasting use cases, improving prediction accuracy and automating business processes for client engagements.
- Improved model performance by 15–25% through feature engineering, hyperparameter tuning, and cross-validation.

Data Analyst | McKesson, New York, NY Mar 2020 – Dec 2021

- Performed data profiling, cleansing, mapping, and quality assessment on complex clinical and operational datasets using SQL (MS SQL Server) and Python.
- Developed SQL and ETL pipelines processing millions of healthcare records, reducing data preparation time by 35%.
- Built predictive ML models using Python and Scikit-learn; optimized algorithms for distributed computing environments.
- Developed interactive visualizations using Seaborn, Plotly, and Matplotlib to uncover trends in multi-dimensional datasets.

EDUCATION

Master of Science in Applied Data Science - (Expected 2027) In Progress

University of Michigan, Ann Arbor, MI

Master of Science in Applied Artificial Intelligence - Dec 2022

University of San Diego, San Diego, CA

Bachelor of Science in Mechanical Engineering - Jun 2019

City College of New York, New York, NY

CERTIFICATIONS

- IBM AI Engineering Professional Certificate
- IBM Data Science Professional Certificate
- Python for Everybody Specialization
- Fundamentals of Engineering Exam (FE Exam)