William Duggan

Morris Plains, NJ | wmduggan@gmail.com
(973) 303-3354 | https://www.linkedin.com/in/wmduggan/

EXECUTIVE SUMMARY

A visionary and performance-driven **Data Scientist / Business Scientist** and key team member, with a successful track record of aligning business and data strategy to deliver initiatives that improve processes, systems, and profitability.

- **Communication / Relationship Building:** Exceptional communication and presentation skills. An influencer with the ability to build robust and mutually beneficial business relationships at all levels.
- **Experience Includes**: Data strategy, complex data analysis, data-driven advisement, predictive modeling, machine learning, statistical analysis, data preparation, model evaluation/design, and process improvement.
- **Demonstrated Expertise:** Using extremely large, complex datasets to draw conclusions and effectively communicate findings with both technical and non-technical team members and stakeholders.
- **Data Lead:** Championing data-driven outcomes and business problem solutions via data analysis and building clear visualizations, models, and dashboards.
- **Strategic Business Partner:** Interacting with key stakeholders including executive leadership, providing strategic advisement on data insights, solutions, business impacts, and improvement opportunities.
- **Additional Credentials:** M.S. in Data Science, Certified AWS Cloud Practitioner.

CORE COMPETENCIES

- Data Science / Business Science
- Big Data Modeling
- Business Insights / Impacts
- Deep Learning
- Machine Learning
- Data Mining
- Data Analytics
- Data Visualizations / Dashboards
- Model Development / Deployment
- Statistical Programming
- Scalable Improvements
- Performance Optimization / KPIs
- Data Integration
- Data Cleaning
- Project Management
- Documentation / Testing
- Trends / Pattern Evaluation
- Cross-Functional Collaboration

PROFESSIONAL EXPERIENCE

INDEPENDENT CONSULTANT - Remote 2023 - Present

Data Scientist

Contracted by blockchain company Mask.io to optimize and streamline application transactions. Serving as a mentor for CloudWolf LLC community, providing expertise in the well-architected framework for implementing cost-effective resources on demand in AWS.

- Collaborated with Mask.io, a blockchain company, to optimize and streamline application transactions on-chain. Leveraged data science expertise to analyze blockchain performance metrics and implement data-driven improvements, increasing transaction speed and reducing processing costs.
- Mentored CloudWolf LLC community members, sharing best practices for building cost-effective, on-demand AWS resources within the Well-Architected Framework. Guided members in optimizing resource utilization, implementing automation, and managing serverless functions, resulting in significant cost savings and enhanced operational efficiency.
- Designed and automated data pipelines to seamlessly integrate blockchain data with traditional data sources. Implemented continuous integration/continuous delivery (CI/CD) practices to ensure data consistency and reliability, enabling real-time monitoring and insightful reporting for application performance and user behavior.
- Conducted in-depth exploratory data analyses on diverse data sets, including on-chain and off-chain sources. Developed and evaluated predictive models to identify opportunities for further blockchain optimization and inform data-driven decision-making for long-term business goals.

WILMINGTON PAPER CORPORATION - Pine Brook, NJ 2015 - 2023

Operations Manager, Data Lead

Provided governance and vision for supply chain operations and data strategy. Developed solutions and processes for addressing complex business and technical problems. Managed supply chain operations across 30+ accounts, optimizing workflows for suppliers, customers, and carriers. Served as a Power BI creator for the organization.

- Architected and executed the strategic data vision. Partnered with senior and Lines of Business (LOB) leadership, providing strategic advisement on data-driven insights, business impacts, and uncovered opportunities.
- Implemented standardized operating procedures across all accounts, enhancing operational consistency and scalability.
- Collaborated with cross-functional departments and conducted data cleansing, management, troubleshooting client source connection, admin control and maintenance, and reporting dashboard and visualization for a private company database infrastructure.
- Built and implemented machine learning and statistical models for data strategy using R, Git, and Power BI. Developed and delivered the data infrastructure required for the needs of predictive modeling and analytics.
- Mined and translated data into meaningful insights and built dashboards and visualizations. Presented results and advisement to senior leadership and key stakeholders.
- Researched, designed, built, optimized, and maintained reliable, efficient, and accessible data systems, data pipelines, and models.
- Created and delivered training sessions for employees on database navigation, queries, and report generation.
- Leveraged SQL for data extraction from internal databases, facilitating data-informed decision-making and enhancing customer understanding.
- Employed ETL methodologies to derive and categorize key features matching regional vendor costs and business requirements, thereby supporting trend analysis.
- Provided back-end support for the IT group in customer web portal development, including DevOps and administrative credentials.
- Led multiple fully-lifecycle data projects, including a project that pinpointed and addressed supply chain inefficiencies, achieving a 6% net cost reduction.

H&M INTERNATIONAL - Jersey City, NJ 2012 - 2015

Assistant Terminal Manager (2013 - 2015)

Strategically coordinated Operation Manager shifts with 10 daily inbound/outbound cargo train schedules. Assisted in negotiating local ILA labor contracts, ensuring 24/7 lift operation coverage for Norfolk Southern and CSX rail yards.

- Enhanced lift per man hour (LPMH) by 10%, leading to improved productivity and efficiency.
- Ensured punctual outbound train releases and departures for NS and CSX railroads.
- Implemented effective manpower scheduling to increase terminal efficiency and reduce delays.
- Established and monitored key performance indicators (KPIs) for railroad operations.
- Reduced vendor dwell time by 25%, increasing operational efficiency and reducing downtime.

Operations Manager (2012 - 2013)

Liaised with Norfolk Southern (NS) managers to facilitate timely cargo train activities and resolve delays while managing 16+ company assets to load and unload 6 daily cargo trains.

- Analyzed daily train manifests to ensure receiving and delivery of hazardous and non-hazardous cargo.
- Conducted daily safety briefings, calculated daily lift per man hour, and reported mechanical downtimes.

APM TERMINALS - Port Elizabeth, NJ 2009 - 2011

Operations Manager, Rail Department (2010 - 2011)

Oversaw daily rail loading and unloading activities for a team of 20+ ILA workers. Promoted a productive working environment for the Millennium Marine Rail integration project.

Operations Manager, Marine Department (2009 - 2010)

Ensured safe work environments, managing daily labor and overseeing marine and vessel operations. Collaborated with management, ILA labor, and the marine foreman to achieve deadlines and maximize productivity to 50+ moves per hour.

Full CV Available Upon Request

EDUCATION / CERTIFICATIONS / PROFESSIONAL DEVELOPMENT

M.S. in Data Science - New Jersey Institute of Technology

- Built, deployed, and hosted an interactive web application using a public dataset that predicts future sales data by regional product categories.
- Utilized ETL, data cleansing, mining, visualization, and machine learning to highlight key variables for areas of improvement or concern (i.e., K-means clustering | Linear regression.)

B.S. in Business Management - York College of Pennsylvania

AWS Cloud Practitioner Certification

Data Mining - New Jersey Institute of Technology

Additional Certifications: Microsoft Power BI - LearnPowerBI.com, Machine Learning with Python - Coursera, Python for Applied Data Science - Coursera, Data Analysis with Python - Coursera, Data Visualization with Python - Coursera, Open-Source Tools for Data Science - Coursera, Data Science Methodology - Coursera, Business Science - Shiny Web Applications.

TECHNICAL SKILLS

Microsoft Office, Cloud Computing, CSS, HTML, JavaScript, Python, R, SQL, Power BI, AWS, Hadoop, MapReduce, Java, Linux, Hugging Face, PyTorch, NumPy, Pandas, Sci-kit Learn, MatPlotLib.

PROJECTS

1. Sales Demand Forecast Application:
   - *Functionality/Purpose*: This application serves as a critical tool for Marketing and Sales teams, enabling them to accurately predict and analyze trends in product demand. By leveraging advanced analytics, it aids in strategic planning and inventory management, ensuring optimal resource allocation and efficiency in meeting market demands.
   - *Technologies Used*: The core of this application is built using RStudio's Shiny framework, known for its interactivity and user-friendly interfaces in web applications. The predictive capabilities are powered by machine learning algorithms, which analyze historical sales data to forecast future trends. The integration of ShinyR allows for dynamic and responsive dashboard displays, making the analysis visually accessible and actionable. Utilized tools such as R Programming, Git for version control, and machine learning libraries like parsnip and xgboost.
   - *Key Achievements*: Successfully deployed predictive models that improved forecasts up to 2 years in the future, which could lead to a significant reduction in inventory costs. Enhanced decision-making processes for the marketing and sales teams by providing real-time, data-driven insights.

2. Knowledge Distillation (NLP Project):
   - *Functionality/Purpose*: This project focuses on optimizing Natural Language Processing (NLP) tasks using knowledge distillation techniques. It involves fine-tuning a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model from HuggingFace on a specific text classification task, enhancing the model's accuracy and efficiency. Our model performed with 95% accuracy to classify the intent of "i like you" with "i love you."
   - *Technologies Used*: Implemented using the HuggingFace's Transformers library, known for its comprehensive collection of pre-trained models and ease of use. The project leverages a BERT-based-uncased model, suitable for understanding the context in text data without the need for case sensitivity. The transformer pipelines are used for tokenization, forward pass, and backward pass processes, ensuring efficient training and evaluation of the model.
   - *Innovative Approach*: Developed an efficient training pipeline that reduced model training time by 30% without compromising accuracy by leveraging CUDA. Applied advanced techniques in transformer architectures and fine-tuning, leading to a 15% improvement in model performance on benchmark datasets.
   - *Collaboration and Integration*: Worked closely with two other NJIT master students to integrate the model into an existing text analytics platform. Utilized HuggingFace, Python and PyTorch for model development and deployment, along with Google Collab for experimentation and Git for source code management.