**SHUMETIE TEFERA**

213-352-9188 | tshumetie5@gmail.com | linkedin.com/in/tshumetie5/

github.com/shum05 | s3shumetie23.com | Los Angeles, CA

**DATA ANALYST | DATA ENGINEER | ML ENGINEER | DATA SCIENTIST**

Data Science professional with 5 years of experience in working to understand business problems, gathering business requirements, identifying relevant datasets, and cleaning and blending data to analyze it to help businesses make informed decisions. Able to leverage technical experience to articulate technical concepts into simpler terms and turn complex technical data into actionable insights that enhance business performance and deliver tangible business value.

**TECHNICAL SKILLS**

**Programming Languages:** Python, R, SQL, JavaScript, HTML, CSS

**Data Visualization:** Power BI, Tableau, Microsoft Excel, Looker Studio, Matplotlib, Seaborn, Microsoft PowerPoint

**Data Tools & Cloud Technologies:** Flask, Docker, CI/CD, Jupyter Notebooks, VSCode, PyTorch, NoSQL, MongoDB, TensorFlow, Keras, Pandas, NumPy, Snowflake, AWS, Microsoft Azure

**CORE COMPETENCIES**

**Data Science:** Statistical Analysis, Machine Learning, Deep Learning, Natural Language Processing (NLP), Mathematics, Quantitative Analysis, Big Data Sets, Time Series Forecasting, Model Tuning, Artificial Intelligence (AI), Regression, Clustering, Classification

**Data Analysis:** Data Collection, Data Cleansing, ETL/ELT, Data Transformation, Data Modeling, Data Mining, Data Warehousing, Data Wrangling, Dashboard, Time Series Analysis, Data Reporting, Metric Design (Key Performance Indicators - KPIs), A/B Testing, Statistics

**Software Engineering**: Software Development Lifecycle(SDLC), Access Control, User Interface (UI), Version Control, Unit Tests

**Professional:** Written & Verbal Communication, Adaptability, Continuous Learning, Growth Mindset, Problem Solving

**CERTIFICATIONS**

**AWS Certified Solutions Architect - Associate,** Amazon Web Services (AWS) March 2020

**MuleSoft Certified Developer (API),** Salesforce July 2022

**Certified Full Stack Developer,** ETech February 2019

**SQL/ETL (MSBI) Developer,** YaYoBE November 2016

**PROFESSIONAL EXPERIENCE**

**Data Scientist (Intern),** Careerera, Herndon, VA June 2023 - November 2023

* Conducted NLP sentiment analysis, leading to scaled up customer feedback and analyzed Time Series data, reducing forecasting errors by 5%
* Crafted visually stunning data visualizations using Power BI, Tableau; facilitated clear communication of complex financial data, enabling strategic decisions resulting in 10% increase in company profitability
* Enriched data quality with 15% fewer errors by building and evaluating ML models for customer churn prediction, achieving a 20% reduction in customer attrition

**Data Analyst,** McKesson Corporation, Chino, CA July 2022 - June 2023

* Developed ETL pipeline to resolved data inconsistencies at 98% data accuracy, developed automated data cleaning scripts using Python, saving 10+ hours per week
* Spearheaded advanced statistical techniques to identify trends, patterns, and correlations to maximize customer retention and a 10% reduction in churn rate
* Leveraged data to drive business growth and efficiency and deliver data-driven results by developing & refining Key Performance Indicators (KPIs) and establishing a feedback loop for continuous improvement

**Data Analyst,** DigiSure Insurance Solutions Inc., Burlingame, CA September 2021 - June 2022

* Executed market research on customer preferences and industry trends for informed product development strategies, leading to a $1.5M revenue growth
* Engineered visual reports, KPI scorecards, filters, parameters using Power BI desktop for claim analysis and leveraged DAX formulas to transform raw data into actionable insights, driving revenue growth by 15%
* Incorporated advanced analytics models and machine learning algorithms along with data visualization to support critical operational and analytical applications
* Activated advanced statistical modeling techniques to conduct exploratory data analysis (EDA), uncovering valuable insights and driving process optimization; identified cost-saving opportunities, leading to $100k in annual savings and significant business growth
* Devised predictive models to accurately forecast sales, introducing a 15% reduction in stockouts
* Presented analytical findings and insights to non-technical audiences in a clear and concise manner with MS PowerPoint, enabling informed decision-making and creative problem-solving

**SQL/ETL Developer,** First General Bank, Los Angeles, CA January 2020 - July 2021

* Exploited Shell Scripting techniques to automate data migrations, prompting a 50% cut down in migration time, 30% reduction in maintenance time and minimal data loss
* Optimized SQL queries, achieving a 15% increase in database efficiency reduced query response times by 40%, leading to an improvement in report generation time by 25%
* Contributed to planning, defining and designing large databases based on business needs for insights & analytics, conducted Joint Requirement Development (JRD) sessions to analyze the database requirements in detail
* Defined Key Performance Indicator (KPI) metrics to create scorecards for optimizing operational efficiency and risk management contributing to the proactive identification and mitigation of potential risks, thus safeguarding the bank's financial integrity

**PROJECTS**

**iCargo Screener Certification Readiness Prediction** October 2023 - November 2023

* Developed a Certification Readiness Prediction model using Python, XGBoost, and scikit-learn, enabling precise assessment of an individual's preparedness, optimizing training efforts, and resource allocation in the logistics and supply chain industry
* Technologies Utilized: Python, XGBoost, scikit-learn

**Question-Answering Tool for Bridge Design Manuals**  August 2023 - October 2023

* Designed a robust Question Answering tool for civil engineering Bridge Design Manuals to extract answers from a dataset, enhancing access to critical information and streamlining the design process
* Technologies Utilized: Python, Haystack, Mistral, ChainLit

**Coffee-Leaf-Rust Classifier -CNN** June 2023 – July 2023

* Designed and implemented a Coffee-Leaf-Rust Classifier using Convolutional Neural Networks with Python, NumPy, TensorFlow, and AWS, achieving accurate disease detection and visualizing results with Matplotlib, contributing to sustainable coffee production and crop management
* Technologies Utilized: Python, NumPy, TensorFlow, AWS, Matplotlib

**Forecasting Economic Evolution - Time Series Analysis** April 2022 - July 2022

* Utilized Python, Facebook Prophet, and Seaborn to conduct a comprehensive time series analysis of currency exchange rates, forecasting future trends and uncovering valuable insights to support informed financial decisions for a 5-year growth and transformation plan
* Technologies Utilized: Python, Facebook Prophet

**EDUCATION**

**Post Graduate Program Certificate in Data Science**, Careerera, VA, USA May 2023

**Master of Science, Applied Mathematics - Numerical Analysis,** Addis Ababa University, Ethiopia July 2012

**Bachelor of Science, Civil Engineering,** Mekelle University, Mekelle, Ethiopia February 2016