FAIZA SIDDIQUI

C: 587-969-9970 │ Canada │ E: faizasiddiqui@uregina.ca │ LinkedIn │ GitHub │ Portfolio

Professional Summary: Highly motivated and outstanding Data Scientist recently graduated from University of Regina with 7 years of proven experience in Software & Data Quality Assurance and in Data Science

TECHNICAL COMPETENCIES

Technologies and Tools: Tableau, PowerBI, Hadoop, Git/GitHub, **Azure** (Azure Synapse Analytics), **AWS** (S3, EC2, EMR), Spark, PySpark, .NET, Microsoft Office Suite including Excel, Word, PowerPoint, Outlook, Databricks, Docker, Jupyter, Agile, Scrum, Kanban, Kafka, Visual Studio, VS Code, UNIX/Linux, Windows, Jira, Confluence

Languages: **SQL**, Java, **Python** (TensorFlow, NumPy, Pandas, Streamlit, Panel, Scikit-learn, Keras, PySpark, Matplotlib, PyTorch, Dash, Dask)

Modeling: Machine Learning, Deep Learning, NLP, Predictive Analytics, Regression, Clustering, Classification, Time Series Analysis, Generative Adversarial Network, Normalization, Object Detection, Computer Vision, Data Visualization, Big Data Analytics, Text Analytics, Unsupervised learning, Supervised learning, Anomaly Detection, Convolutional Neural Network, Exploratory Data Analysis, Forecasting, Data Modeling, Statistics

DATA SCIENCE PROJECTS

Big Data Analytics & Cloud Computing Project   JUL 2022 - AUG 2022

- Effectively managed 3.2M records of Amazon's Reviews on Electronics purchases with AWS EMR & **PySpark** displaying strong problem solving and analytical skills for data upload and storage in **AWS**
- Developed using Python and various data science libraries such as Scikit-learn, Pandas, NumPy, PyTorch, Keras
- Loaded, processed, explored, analyzed big data efficiently using cloud computing platform & performed natural language processing using Vader Model, visualized data with **Tableau**

Time Series Analysis - Forecasting Prices of Gold Project   MAY 2022 – JUN 2022

- Cleaned & analyzed time series data by developing various models namely, Persist, Exponential Smoothing, Holt's Linear Trend method, Holt-Winters' Seasonal method, ARIMA & SARIMA
- Built predictive models in Python to forecast monthly average gold prices of data

"Approximation of Losses on a Lending Project" for Farm Credit Canada (FCC)   FEB 2022 – APR 2022

- Independently developed linear, lasso, ridge regression models by looking at objectives & data & applying the right data science technique to help FCC achieve great value by faster loss calculation
- Collaborated with team members to develop mL models with hyperparameter tuning, namely decision tree, random forest, deep neural networks using Python. Conducted exploratory data analysis & visualized data

Accomplishments: **Awarded by FCC** for the best recommendations and findings throughout the entire class. Predicted loss calculation on a loan with 92.7% accuracy

"Housing Price Prediction": 24-Hour Solo Data Science Project   DEC 2022

- Independently completed housing price prediction project within 24-hours using LSTM, multiple regression, SARIMAX & ARIMA models, showcasing efficient time management & solo project execution in data science

"PCL Worker Return Prediction"   SEP 2023 – DEC 2023

- **Achieved Exceptional Workforce Insights:** Leveraged cutting-edge mL models, including Random Forest, Logistic Regression, Gradient Boosting, Neural Networks and achieved a remarkable accuracy rate of 76%, empowering PCL recruiters with actionable insights which significantly optimized workforce management, resulting in substantial operational efficiency gains and cost savings for PCL

RECENT PROFESSIONAL EXPERIENCE

S.I. SYSTEMS & AVENUE LIVING ASSET MANAGEMENT *(Data Science & Analytics Team)* – CANADA
DATA QUALITY ANALYST   FEB 2024 – PRESENT

- Led the development of robust testing plans for data assets, reports, and dashboards, ensuring successful implementation and accuracy.
- Successfully facilitated data corrections through collaboration with business groups, ensuring accurate and reliable data by implementing efficient resolution processes.
- Accomplished the establishment and maintenance of data quality rules and validation checks to ensure ongoing data accuracy, by implementing a systematic approach to rule creation and maintenance.
- Achieved proactive monitoring of data quality metrics to identify trends and areas requiring attention, ensuring a high standard of data integrity through consistent monitoring and analysis.
- Successfully created and maintained documentation on data quality procedures, contributing to a standardized and efficient data governance framework.
- Generated regular reports on data quality and compliance, providing stakeholders with actionable insights for strategic decision-making. Successfully managed the Data Quality Issues log, ensuring a systematic and organized approach to issue resolution by implementing effective log management and resolution processes.

PCL CONSTRUCTION – CANADA
DATA SCIENCE STUDENT   MAY 2023 – DEC 23
- Conducted groundbreaking workforce data analysis using Python and SQL in Azure Databricks. Validated results with PowerBI and Excel, presenting key insights to stakeholders
- Integrated data from diverse sources into Azure SQL database using MS SQL & Azure Data Studio, implementing data cleaning, transformation with SQL. Created Streamlit front-end app for efficient data management in database
- Implemented a variety of mL algorithms with Scikit-learn to enhance recruitment processes for effective future workforce planning. Also, automated data quality checks through unit tests for data validation
- Organized team-building activities as Data Science Intern Lead, supporting team of interns, and supported Agile processes for task execution as a Scrum Leader
- Dockerized and deployed data-driven Panel app & Dash app on Azure for PCL employees, enabling exclusive access. Established continuous integration on GitHub for seamless maintenance of the apps

IQMETRIX (Data & Analytics Team) – CANADA
DATA ENGINEER CO-OP   JAN 2023 – APR 2023
- Implemented Data Flow design and data definition workflows
- Conducted verification processes of data flow processes, with strong attention-to-detail and analytical skills
- Developed & merged multiple data streams in Kafka Confluent Cloud to facilitate real-time data analysis
- Efficiently produced detailed documentation for data flow designs and Kafka Confluent Cloud configurations,

UNIVERSITY OF REGINA – CANADA
OBJECT-ORIENTED DESIGN LAB INSTRUCTOR   JUL 2022 – DEC 2022
- Assisted students in developing problem-solving techniques in C++ by displaying strong attention-to-detail
- Monitored student progress and provided ongoing feedback using multiple data sources such as lab reports, test scores, and attendance records with effective time management and organizational skills
- Troubleshot & resolved technical issues in the lab to ensure the accuracy and reliability of the systems
- Trained students of all levels of expertise with effective and excellent communication both verbal and written

GOOGLE DEVELOPER STUDENT CLUB of University of Regina (URGDSC) – CANADA
ARTIFICIAL INTELLIGENCE / MACHINE LEARNING LEAD   MAR 2022 – DEC 2023
- Established relationships with for 150 students and ensured all students felt attended with personalized feedback, class discussions and engagement demonstrating excellent presentation skills
- Delivered presentations to students & community groups with a proven ability to adapt and collaborate
- Collaborated with students to plan AI/ML workshops for students on campus with communicative approach
- Gave workshops on Tableau teaching how to develop and modify reports and dashboards

---

EDUCATION

2023	**Master's Degree**, **Data Science** – University of Regina, Saskatchewan, Canada	
	FGSR Excellence Award Nominee	**CGPA: 4.0/4.0**
	Coursework: Software Development Fundamentals, Python & Data Fundamentals, Data Science, Statistics & Machine Learning, Applied Machine Learning, Big Data Analytics & Cloud Computing	
2014	Bachelor's Degree, Information Technology – Software Engineering	**CGPA: 3.87/4**

REFERENCES

Dr. Andrei Volodin	**Calvin Feng**, PMP, M.Eng.	**Vincent Langevin**
Professor of Statistics	Senior Data Scientist	Development Lead
Dept of Mathematics & Statistics	Data Science Team	Moose Jaw, Saskatchewan
Faculty of Grad Studies & Research	PCL Construction	iQmetrix
University of Regina	10003 56 Ave NW,	*Email:* vincent.langevin@iqmetrix.com
Regina, Saskatchewan S4S 0A2	Edmonton, AB T6E 5L7	*Relationship:* Supervised my work at
Email: andrei@uregina.ca	*Email:* calvinfeng66@gmail.com	iQMetrix
Phone: 306-585-4771	*Phone:* 587-337-4134	
Relationship: Evaluated my performance in Statistics & ML	*Relationship:* Senior Data Scientist in the same team	

Notice Period: Available after 15 days

Relocation: Preferably willing to work in Calgary. May be open to other locations.

Work Preference: Open to remote, in-person, or hybrid

Work Authorization: Legally eligible to work in Canada without any restrictions or sponsorship