

FAIZA SIDDIQUI

C: 587-969-9970 | Canada | E: faizasiddiqui@uregina.ca | [LinkedIn](#) | [GitHub](#)

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

TECHNICAL COMPETENCIES

Software: Tableau, PowerBI, Hadoop MapReduce, Git/GitHub, **Azure**(Azure Synapse Analytics), **Amazon** (S3, EC2, EMR), Spark, ASP. NET, Microsoft Office Suite including Excel, Word, PowerPoint, Outlook, PyCharm, MATLAB, Computer Vision, Object Detection, Statistics, Databricks, Jupyter Notebook, Docker, Agile, Scrum, Kafka, Visual Studio, VS Code, Hadoop MapReduce, UNIX/Linux, Windows

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

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

DATA SCIENCE EXPERIENCE

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

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. Utilized Python libraries such as Pandas, NumPy in Jupyter notebook
- 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
- Created a Streamlit front-end interface app aligned with PCL's theme for efficient data management in database, enabling non-technical contributors

IQMETRIX – 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, enhancing clarity and future scalability
- Proven track record of strong interpersonal and organizational skills, which allowed to effectively communicate with colleagues
- Successfully worked with little supervision to complete tasks

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
- Troubleshoot & 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 – CURRENT

- 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
- Helped with the development and validation of AI/ML models
- Gave workshops on Tableau teaching how to develop and modify reports and dashboards

KING FAISAL SPECIALIST HOSPITAL & RESEARCH CENTRE (KFSHRC) – SAUDI ARABIA
SOFTWARE QUALITY ASSURANCE ANALYST

APR 2016 – AUG 2021

- Demonstrated excellent problem-solving skills and the ability to think algorithmically, which enabled me to identify and resolve complex issues
- Reported application defects & interoperability flaws to developers using Microsoft Excel & PowerPoint
- Collaborated with teams in Software Development Life Cycle like Requirement Analysis, Design, Implementation, and testing, using Agile Methodology
- Assisted with writing of test scripts, test execution along with business requirements gathering.

Accomplishments:

- ✓ **Awarded thrice** with the Certificate of Appreciation for best employee of the year
- ✓ 50% decrease in the processing time of the requests that were automated

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** **CGPA: 3.87/4**
- 2013 **Training, Android** - Hewlett Packard Global Soft Pvt. Ltd. (HP)

REFERENCES

Dr. Andrei Volodin
Professor of Statistics
Dept of Mathematics & Statistics
Faculty of Grad Studies & Research
University of Regina
Regina, Saskatchewan S4S 0A2
Email: andrei@uregina.ca
Phone: 306-585-4771
Relationship: Evaluated my performance in Statistics & ML

Nova Scheidt, M.Sc.
Lab Instructor
Dept of Computer Science
University of Regina
3737 Wascana Parkway
Regina, Saskatchewan S4S 0A2
Email: nova.scheidt@uregina.ca
Phone: 306-585-4975
Relationship: Supervised my work as Teaching Assistant

Vincent Langevin
Development Lead
Moose Jaw, Regina
Saskatchewan
Email: vincent.langevin@iqmetrix.com
Relationship: Supervised my work at iQMetrix