Walid Amin Khan

Phone: +12368653348
Email: walid.sec3.034@gmail.com
Address: 7953 St John Crescent, Prince George, BC, V2N 4H9
LinkedIn: https://www.linkedin.com/in/walidaminkhan/

Summary:
Data Engineer with specialization in both data analytics/BI solutions and data engineering. With 5 years of experience across social media, retail, SaaS, healthcare, ride sharing and parcel delivery industries, I have a proven track record of building batch and stream processing ETL/ELT pipelines alongside creating analytics reports and dashboards. Proficient in agile methodologies and experienced in working with cross-functional teams to deliver successful projects on time and within budget.

Technology:

- ETL/ELT, pipeline orchestration: Airflow, PySpark, AWS Glue, Airbyte, SSIS
- Programming languages: Python, Javascript, R, Matlab
- Database, Data warehouse: MySQL, PostgreSQL, MS SQL, Oracle, Neo4j, MongoDB, CosmosDB, Redshift, BigQuery
- BI/Analytics: Tableau, PowerBI, AWS QuickSight, Metabase, Mixpanel, Google Analytics, SSRS
- Others: Git, Bitbucket, Docker, Terraform, Jira, Confluence, PySpark, Jupyter, Flask, Django, AWS S3, Azure blob storage, Kafka, RabbitMQ, SQS, SNS, Lambda, Athena, Azure VM, AWS EC2

Professional Experience:

Database Engineer – HelpUsDefend   **(August 2023 - Present)**
- Data modeling for social media app as well as internal supporting applications.
- Working with backend team for data migrations.
- Tuning performance of cloud data infrastructures – led restructuring of social media app data model that resulted in an average 10% faster response time in several scenarios.
- Documentation of Canonical Data Models and Entity Relation Diagrams.

Data Engineer – Raena   **(February 2022 – June 2023)**
- Building data pipelines for analytics and application use – serving both internal and external customers.
- ETL pipeline orchestration and setting up alerting system for both various metric thresholds and data anomalies.
- Data modeling for both end user applications and data warehousing.
- Cost optimization of data infrastructures – by balancing the data processing load between MWAA and EMR cluster monthly ETL pipeline costs were reduced by at least 12% for similar data load.
- Building web crawlers for competitor and social media data scraping - competitor catalog scraping allowed the business to optimize prices reactively, resulting 17% increased sales in the first month.
- Writing APIs for applications to consume data from data warehouse.

Analytics Engineer – Astha.IT   **(November 2021 - January 2022)**
Worked for a single client - Upsy
- Regular meetings with B2B clients for product specification and requirement documentation.
- Finding and documenting several key business metrics.
- Liaise with backend team for implementing tracking of key data points.
- Built and maintained unified real time data analytics pipeline - enabling customers to monitor impact of the product in real time.
- Building business metric dashboards for client and client's external customers.
- Building models for forecasting critical business metrics.

Analytics Engineer – Maya   **(December 2020 - December 2021)**
- Built ETL pipelines to aggregate data from several sources to one for easier analytical workflows.
- Wrote scripts to measure and record impact of different campaigns/programs.
- Fraud detection and prevention - found solid proof of frauds by 6 doctors in platform and ~10K USD lost in fraudulent transactions.
- Ad-hoc data analysis in various business contexts, e.g. – product acceptance rate, ranking customers, AI acceptance vs cost savings etc.
- Leading and mentoring a team of data analysts.

Data Science Engineer – Shohoz   **(April 2019 - November 2020)**
- Built data pipelines for analytics and application use.
- Fraud detection: Did in-depth visual and heuristic based analyses to detect and reduce fraudulent activities in all business verticals - reduced burn by ~30% in the first month of implementation (compared to monthly median burn).
- Syndicate detection: Network visualization of all users to visually identify potential syndicates - in the first month detected 56 syndicates with 540 customers who amounted to 34% of total burn.
- Scoring system: Multi-dimensional scoring system to rank the customers across ride sharing and food delivery services.
- Oversee data labeling sub-project for building recommendation engine.
- Corona tracer Bangladesh: Worked on the government-approved Corona tracer project - specifically on data pipelines, data archival and retention policies. Created applications for geospatial analyses generation.

Intern – Shohoz   **(January 2019 – March 2019)**
- Identifying zones, areas, and addresses with QGIS for the company's city mapping project.
- Research techniques to identify the correct route and distance on the map with given noisy coordinates.
- Successfully completed unidirectional map-matching project with a Hidden Markov Model based solution.

Education:

• Post Degree Diploma – Data Analytics	• Bachelor of Science – Computer Science and Engineering
Douglas College	BRAC University

References:

Towfiqul Islam	Mohammad Kamruzzaman
Executive Director	Director of Engineering
Lheidli T'enneh Nation	Shohoz
mislam@lheidli.ca	ripbd.94@gmail.com