Mahtab Khan

Calgary, AB · (403)714-9366 · mahtabukhan16@gmail.com · linkedin.com/in/mmahtabkhan · github.com/Mmahtab10

Education

University of Calgary   **Sep 2019 – Apr 2024**
Bachelor of Science, Software Engineering   *Calgary, Alberta*
cGPA: 3.8/4.0 (Dean's Honours List 2019-2022)

Relevant Coursework: Data Structures and Algorithms, Database Management Systems, Data Science for Engineers, Web Development, Distributed Systems, Scalable Data Analysis, and Statistics

Work Experience

Pason Systems   **May 2022 – Sep 2023**
Data Science Intern   *Calgary, Alberta*
- Managed the development and maintenance of high-scale data analysis software developed in **Python**, serving over **2,000** weekly users and providing critical functionalities for daily business operations
- Spearheaded the creation and deployment of section detection algorithms for industrial process sensors, improving data processing **speed by 17%** and **accuracy by 15%**
- Implemented K-means clustering algorithms to analyze patterns in drilling surfaces
- Enhanced performance analysis by 20% and boosted critical event detection by 22% through sensor data analytics
- Created documentation of processes and actively collaborated with product managers, designers, and sales representatives to gather client requirements and generate client reports

University of Calgary   **Jun 2021 – Aug 2021**
Software Engineering Research Intern   *Remote*
- Researched video data analytics for human biometrics, focusing on advanced action detection and accurate breathing rate estimation
- Led the development of user interfaces for agent-based modeling, enhancing accessibility for interdisciplinary teams

Technical Projects

Sign4Language   **Python | ML Libraries**
- Developed a gesture recognition ML model for Bay Area Hacks 2020, recording hand movements as .mp4 with custom hand tracking software
- Used Keras, OpenCV, TensorFlow, and Python 3 to feed gesture data into a neural network, translating it into plain text
- Successfully trained the machine learning model on a comprehensive dataset, achieving a 94% accuracy rate

Traffic Crash Prediction   **PySpark | ML Libraries | Databricks**
- Led a Traffic Crash Prediction project using the **Databricks** platform, handling data extraction, cleaning, and analysis on datasets comprising over **700,000 records** retrieved through its API
- Developed a predictive random forest model using PySpark, securing high validation scores in forecasting accidents
- Implemented Matplotlib and Seaborn for advanced data visualizations, boosting data-driven decision-making with concise and actionable insights

TrapEase (capstone project)   **AWS | Python**
- Deployed portals using AWS and docker, increasing threat detection and compromised credentials identification by **40%**
- Optimized operational efficiency by integrating AWS services, including **EC2** for scalable hosting, **RDS** for database management, and **Power BI** for dynamic data visualization, achieving a 30% reduction in data processing times
- Implemented **agile** methodologies, accelerating feature deployment by 20% and boosting productivity by 25% with iterative development

Technical Skills and Interests

Languages: Python, C, C++, **SQL**, JavaScript, Java, Php
Developer Tools: AWS (VPC, RDS, Lightsail, Secrets Manager, Lambda), Postman, Docker, Git, Databricks, Power BI
Libraries/Frameworks: Spark, **JSON**, Scikit-learn, Seaborn, Pandas, boto3, Keras, Matplotlib, XGBoost, AWS SDK
Interests: Soccer, Ping pong, Basketball, E-sports, Golf, Hockey, Cycling