



MUHAMMAD SOHAIB


PRINCIPAL ENGINEER (AI, DATA & ML)

 [Portfolio](#)

 +92 323 5083445

 sohaib867@gmail.com

 [Linkedin](#)

 Islamabad, Pakistan

SUMMARY

With over 10 years of experience, I am an accomplished Data and Machine Learning Engineer, specializing in developing advanced Big Data pipelines and implementing scalable ETL processes, data modeling, and warehousing solutions. My technical acumen covers Python, SQL, and key Big Data tools like Apache Spark, Hadoop, and AWS. In computer vision, I possess deep expertise in a broad spectrum of algorithms, including object detection, classification, action recognition, and activity recognition. Recently, I have furthered my skill set with Large Language Models (LLMs), enhancing data insights in vision-based projects. This advancement highlights my commitment to integrating the latest AI and ML technologies for innovative, effective solutions. Renowned for my exceptional problem-solving abilities and collaborative spirit, I consistently deliver high-quality, timely projects.

TECHNICAL SKILLS

PROGRAMMING LANGUAGES

- PYTHON
- SCALA
- C++
- SQL

BIG - DATA

- MAP REDUCE
- SPARK
- KAFKA
- DRILL
- AIRFLOW
- REDIS
- HBASE
- NIFI

MACHINE LEARNING

- PYTORCH
- PADDLE PADDLE
- DEEPSTREAM
- TENSORIT
- TRITON SERVER
- TAO TOOLKIT
- TENSORFLOW
- LANG CHAIN

DATABASES

- POSTGRESQL
- MONGODB
- MAPRDB
- CASSANDRA
- MYSQL
- TIME SCALE
- HBASE
- REDIS

DEVOPS

- DOCKER
- TERFORM
- PACKER
- JENKINS

WORK EXPERIENCE

PRINCIPAL ENGINEER (DATA & ML)

DARVIS

JANUARY 2022 – Present

- Integrated Large Language Models (LLMs) to elevate data analytics capabilities.
- Led the development of an in-house annotation tool, enhancing the efficiency of the data labeling process.
- Engineered and enhanced sophisticated computer vision and machine learning solutions for real-time data processing and analysis, catering specifically to the logistics, retail, and healthcare industries.
- Proficient in a range of technologies including object detection, tracking, classification, re-identification, facial recognition, pose estimation, and action recognition.
- Skilled in Big Data processing, experienced with Apache Spark, Hadoop, and real-time stream processing.
- Managed and guided teams effectively, ensuring the successful execution and delivery of innovative, data-driven solutions.

SENIOR ENGINEER (DATA & ML)

DARVIS

JUNE 2018 – DECEMBER 2021

- Improved system throughput, accuracy, and performance, utilizing tools like Deepstream SDK and TensorRT.
- Advanced computer vision and machine learning systems were developed and optimized for real-time data processing and analysis.
- Proficient in a range of technologies including object detection, tracking, classification
- Created backend services to calculate key performance indicators (KPIs) and combine data from machine learning models for efficient front-end display.
- Created a mono repo architecture that is Dockerized to enable smooth integration of machine learning services with front-end and back-end systems.

SENIOR DATA ENGINEER

PERSONTYLE


JANUARY 2017 – MAY 2018


- The position necessitates consulting and training in Big Data technology and Machine Learning. It entails knowing how to use various open-source Big Data platforms, cloud platforms, NoSQL databases, real-time Big Data analysis, machine learning, and application development.
- Proficient in training and consulting large groups on Hadoop, Spark, and related technologies, while developing plug-and-play modules for diverse data sources using open-source Big Data solutions, fostering collaboration and data accessibility across teams.
- Skilled in fostering Blockchain talent, driving organizational growth through technical leadership, and setting a strategic technological vision while spearheading technological development initiatives.

MUHAMMAD SOHAIB


PRINCIPAL ENGINEER (AI, DATA & ML)

 [Portfolio](#)

 +92 323 5083445

 sohaib867@gmail.com

 [Linkedin](#)

 Islamabad, Pakistan

TECHNICAL SKILLS

CLOUD PLATFORMS

- IBM WATSON
- GOOGLE CLOUD
- AZURE
- AWS

ALGORITHMS EXPERTISE

- DETECTION
- CLASSIFICATION
- TRACKING
- RE-ID
- HOMOGRAPHY
- AR
- OCR
- LLMS

LANGUAGES

- ENGLISH
- URDU

REFERENCES

Paul Warren Chief Technology Officer, **DARVIS**

 pawarren95@gmail.com

Fahim Abbasi, Security Researcher, **CISCO**

 mailto:fahim@gmail.com

HOBBIES

- GAMING
- READING
- TREKKING
- FOOTBALL

WORK EXPERIENCE

- Possessing a deep understanding of real-time data acquisition and processing for Agro Industries, I excel in:
 - Analyzing and managing real-time data streams from various sources including IoT devices, sensors, and machinery.
 - Developing robust data processing pipelines to clean, transform, and aggregate real-time data for efficient storage and analysis.
 - Designing and implementing scalable database solutions to accommodate the high volume and velocity of data generated in the Agro-Industry.

SOFTWARE ENGINEER

EBRYX

MARCH 2016 – DECEMBER 2016

- The position necessitates evaluating and deriving actionable insights and trends from both offline and live data. Transform the data into aesthetically appealing, real-time dashboards that are easy to interpret.

SOFTWARE ENGINEER

PREDICTIFY ME

DECEMBER 2014 – FEBRUARY 2016

- The role necessitates the creation of a single platform for exploring, analyzing, and strategizing consolidated information from legitimate resources. A one-stop shop for gaining a 360-degree perspective and enhanced expertise in all sectors dealing with human data.

PROJECTS

VISIONARY DATA INSIGHT ENGINE

- Development of an advanced chatbot using LangChain technology and Large Language Models (LLMs) to provide analytical insights from structured data.
- Engineered the system to calculate key performance indicators (KPIs) such as dwell time and zone-specific dwell metrics for strategic analytics.
- Implemented a chatbot interface to enhance user experience, enabling dynamic, conversational interactions for real-time data inquiries.
- Facilitated monitoring of daily activities and object movements within specific areas, providing immediate, data-informed responses.
- This approach significantly improved data accessibility and empowered clients with real-time insights for informed decision-making in various operational scenarios.

INVENTORY TRACKER

- Development of an inference pipeline with object detection, tracking, classification, and QR code recognition using 6 real-time cameras at 7 frames per second on an RTX 3060.
- Training and testing of object detection and classification models regularly to enhance accuracy.
- Increased application's throughput by 30% using Deepstream SDK.
- Homography and georeferencing techniques to accurately align, transform, and enable precise mapping of visual data to floorplan coordinates.
- Custom Algorithms for improving the tracker performance and KPIs.
- Creation of custom gst-plugins.

MUHAMMAD SOHAIB

PRINCIPAL ENGINEER (AI, DATA & ML)



[Portfolio](#)

+92 323 5083445

sohaib867@gmail.com



[Linkedin](#)



Islamabad, Pakistan

PROJECTS

PRODUCTIVITY TRACKER

- Development of an inference pipeline with object detection, tracking, facial recognition, and object reidentification.
- Routinely training, testing, and refining object detection, facial recognition, and object re-identification models to ensure optimal performance.
- Deployment of all models in the deepstream pipeline.
- Homography and georeferencing techniques to accurately align and transform images, facilitating the precise mapping of visual data to floorplan coordinates for advanced indoor navigation and spatial analysis.
- Enhancing accuracy by ensembling color-based and object re-identification models, utilizing complementary features to optimize recognition performance and ensure robust results in diverse environments.

FALL PREVENTION:

- Development and training of action recognition models on both 2D and 3D networks, harnessing the power of spatial and temporal features for enhanced human activity recognition.
- Designing and implementing a high-performance action recognition inference pipeline in C++, optimizing for real-time processing.
- Model optimization using TensorRT and model training using NVIDIA's Transfer Learning Toolkit.

THE MISAAL

- Spearheaded a team of 6 developers in creating an in-house annotation tool, leveraging state-of-the-art AI features to optimize the labeling process and enhance dataset quality.
- Fostered a collaborative environment, driving innovation and efficient development of AI-powered annotation capabilities.
- Coordinated project milestones, ensuring timely delivery of a cutting-edge annotation tool tailored to the team's unique requirements.
- Mentored team members in AI and computer vision techniques, fostering skill development and knowledge sharing.
- Managed resource allocation, effectively balancing team workload and project goals to deliver a high-quality annotation solution.

READY CASH:

- Improving the scalability, stability, accuracy, speed, and efficiency of our existing data systems.
- Designing and developing end-to-end data processing pipelines using technologies such as Apache Spark Apache Drill and Apache Airflow.
- Developing, maintaining, and optimizing the core libraries for batch processing and ingestion of large volumes of data to the big data infrastructure.

EDUCATION

UNIVERSITY OF ENGINEERING & TECHNOLOGY, TAXILA

BS COMPUTER ENGINEERING

SEPTEMBER 2010 – JUNE 2014