KHIZAR HUSSAIN

 Khizar-Hussian ◇ khizarhussain

 khizarhussain.dev ◇ ✉ khizar@vt.edu ◇ ✉ m.hussainkhizar@gmail.com ◇ ▮ +1-540-824-0048

RESEARCH INTERESTS

High Performance Computing, Distributed Computing, Computer Architecture, Systems and Networking, Network Security, Machine Learning, Deep Learning, Computer Vision.

EDUCATION

Virginia Polytechnic Institute and State University - Blacksburg   2024 - Present
PhD Computer Science
Relevant Courses: Advanced Machine Learning, Advanced Topics in Language and Translation

National University of Computer and Emerging Sciences, Lahore   2017 - 2021
BS Computer Science   Cumulative GPA: **3.73**
Relevant Courses: Computer Architecture, Operating Systems, Compiler Construction, Assembly Language, Computer Networks, Artificial Intelligence, Computer Vision, Data Science, Algorithms, Probability and Statistics, Databases, Advanced Databases

RESEARCH PROJECTS

Ransomware Analysis using SPADE provenance graphs and Machine Learning

- Developed a method for automated creation of a virtual machine loaded with a ransomware provided by the user
- Trained KNN, Decision Trees, Linear Regressors and Neural Networks on data acquired from the SPADE kernel for malicious activity pattern recognition.
- Proposed methods for classifying processes as malicious and/or ransomware using Graph Neural Networks and transformer-based networks.

Uncovering the Provenance of Lies: Empirical Study on the Exploitation of Data Voids

- Developed methods for the detection of data voids that propagate from non-mainstream to mainstream media.
- Trained and deployed SOTA language models, for encoding and decoding, such as BERT for extracting language semantics from n-grams.
- Proposed methods for classifying data voids utilizing measuring semantic shift from word and phrase temporal embeddings.

Soccer Player Re-Identification Through Broadcast Video Streams

- Created a solution for real-time re-identification of subjects using a combination of real-time tracking algorithms and deep learning models.
- Prepared a custom, MOT compliant, soccer player dataset for jersey number detection and recognition, and player tracking.
- Deployed an end-to-end real-time re-identification pipeline compatible with live broadcasts of soccer matches.
- Demoed the solution to FIFA, who later bought out the product from OMNO-AI.

Performance analysis Of MPI and MPJ for scaling up DL4J on Commodity HPC Clusters

- Benchmarked MPI libraries in C against MPJ Express for Java and Apache Spark and Hadoop for distributed deep learning.
- Explored the open-source DL4J library and added MPJ Express compatibility to distributed deep learning loads over commodity clusters.
- Proposed options for integrating MPI in Python 3 for potentially faster machine learning using distributed computing.

WORK EXPERIENCE

Abyss Solutions
L6 Software Engineer - Team Lead Scalers   Feb. 2023 - Dec. 2023

- Delivered critical initiatives and brought them to production by collaborating with the ML teams, such as Graph Assisted Tagging, and 2D ML Asset Labeling.
- Automated deployments to our V2 product through a migration tool reducing the time it takes for deployment from 3 days to 1-4 hours.
- Wrote a suite of CLI tools called fabric-deploy-v2 to manage V2 deployments, to purge, export, import, transfer and transform data using MongoDB Compass compatible bson files, and over the network.
- Implemented tools for automating ops tasks such as user management, data extraction, links migration on spreadsheets between two different deployments, copying over work packs, and manipulating point cloud data.
- Developed and used tools to automate the upload, download, and metrics collection of our datasets on the V7 Darwin platform using the Darwin-py SDK.
- Implemented data ETL and processing pipelines for critical processing using Python and Prefect stack, and scaled it up using Ray clusters.

AppRocket
Full-stack Engineer   Sept. 2022 - Dec. 2022

- Lead the development for the entire backend API for the 3i Capital investment platform.
- Added integrations for Slack, Stripe Payments, MailChimp Emails, and Docusign ClickWrap Agreements.
- Implemented an alert and monitoring system by incorporating E2E tests written in AWS Canaries with AWS Lambda functions which brought down our live-site turnaround time by 60%.
- Added rotating logs with auto compression for each of our instances, for reduced space request logging saving upwards of 22% in space costs.
- Actualized the plans for creating 3 in-house products and worked with the CEO to create bespoke teams for those products.
- Lead the development of one of the products *(under an NDA)* from scratch to a fully functioning MVP in under a month.
- Implemented an SMTP server in NodeJS using Web-Sockets and data streaming to ensure high scalability.

Lahore University of Management Sciences
Technical Specialist - Technology for People Initiative Lab   Jan. 2022 - Oct. 2022

- Lead the development of Pakistan's first Child Protection IMS funded by Unicef Pakistan.
- Oversaw a team of 6 developers from Septem System and Xprolabs for the development of the application and deployment to all 5 provinces of Pakistan.
- Implemented Ansible Playbooks for seamless remote deployments to servers in inaccessible flood-stricken areas like Gilgit-Baltistan, and politically unstable areas like Balochistan.

- Coordinated with PITB and NITB for provisioning and placement of the latest server equipment and data facilities to fast-track the deployment and usage of the CPIMS by the 60 million youth in Pakistan.

Research Assistant - Technology for People Initiative Lab   Mar. 2021 - Oct. 2022
NLP, Internet Measurement and Privacy, Systems and Security
Advisor: Dr. Fareed Zaffar

OMNO AI
Computer Vision Research Engineer   Jun. 2020 - Jul. 2021
As a research engineer at OMNO-AI, I worked on my BS(CS) capstone project to develop a novel deep-learning solution for the real-time identification of football players in a live broadcast.

- Prepared an in house, MOT compliant, soccer player dataset from raw videos from Premiere League, La Liga, and Bundesliga.
- Created a real-time player tracking pipeline in PyTorch using a combination DeepSort, YOLOv5 and various other CV algorithms.
- Trained models for player jersey number detection using YOLOv5, fine-tuned on our dataset.
- Trained models for jersey number recognition using ResNet trained on SVHN dataset and fine-tuned on our soccer player dataset.
- Deployed an end-to-end real-time re-identification pipeline compatible with live broadcasts of soccer matches.

FAST - NUCES
Teaching Assistant - Compiler Construction   Sept. 2021 - Mar. 2022
Professor: Aamir Raheem

Teaching Assistant - Comp. Arch. and Assembly Language   Sept. 2020 - Jan. 2021
Professor: Samin Iftikhar

Research Assistant - FAST   Jun. 2019 - Apr. 2020
HPC and Deep Learning
Advisor: Dr. Aamir Shafi

Lab Instructor - BRAIA Summer School - NUST   16th July 2019

- Singlehandedly managed the creation and delivery of a 3 hour lab for international students
- Introduced HPC and parallelizing physics and engineering simulations using Java and MPJ-Express
- *Accompanied by: Dr. Aamir Shafi*

TECHNICAL STRENGTHS

Programming Languages: Python, JavaScript, TypeScript, Java, C/C++, C#, SQL, Bash, Go
Tools & Frameworks: Keras, Pytorch, Tensorflow, Pytesseract, PySpark, OpenCV, OpenCV-Python, Jadx, Android Studio, DL4J MongoDB, NodeJS, ExpressJS, WebSockets, Docker, Kubernetes, AWS, GCP, Linux, Django, Flask, Fast API, Spring Boot, .Net, LaTeX

HONORS AND AWARDS

Rector's Silver Medal for 2nd Highest SGPA	Spring 2021
Rector's Silver Medal for 2nd Highest SGPA	Fall 2017
Dean's Honour List	2017 - 2021
President Voice Debating Society	2020 - 2021
Vice-President Voice Debating Society	2019 - 2020
General Secretary Voice Debating Society	2018 - 2019