

Asjad Khan

Data Scientist

+61493883978 @ asjad@outlook.com https://www.asjadkhan.com Sydney

SUMMARY

Data Scientist with a PhD at the intersection of Computer Science and Applied Machine Learning. My experience spans the design and development of large-scale, data-driven software solutions across industry and higher education. I also have experience in teaching assistance, research-led problem solving, developing educational technology (EdTech) solutions, and stakeholder engagement. I stay current with developments in my field through research, publication, and the hands-on delivery of data-driven solutions. I am motivated by using data, technology, and inclusive teaching practices to improve student outcomes and learning experiences.

EDUCATION

Ph.D. (Computer Science and Applied Machine Learning)

Decision Systems Lab, University Of Wollongong (UoW)

2018 - 2022 Wollongong, Australia

M.Sc. (Software Engineering and Intelligent Systems)

University of Wollongong (UoW)

2014 - 2016 Wollongong, Australia

INDUSTRY EXPERIENCE



Data Scientist - Ed. Innovation (Academic Success Monitor)

UNSW

05/2023 - 01/2024 Sydney, Aus

- Played a pivotal role in the design and delivery of the Academic Success Monitor (ASM), an award-winning AI/ML platform that enables early identification of academically at-risk students and timely, evidence-based interventions. ASM achieved 78% student coverage and **67% course adoption** across the university, exceeding project targets.
- Led the re-architecture of the ASM data lake, integrating Moodle, Echo360, ECOS, and SIMS into a unified AI/ML-ready platform. Established scalable MLOps capabilities, eliminated data silos, and implemented secure data governance using **Databricks Unity Catalog**, creating the foundation for institution-wide learning analytics.
- Led the design, development, and deployment of 10+ machine learning models for early academic risk prediction, achieving **82% recall and 75% precision**, exceeding project performance targets. The solution identified students requiring support 3.5× earlier with higher precision and 12.8× more comprehensively than previous manual processes.
- Delivered measurable improvements in student success through AI-driven interventions. Courses adopting ASM demonstrated a statistically significant reduction in failure rates, while 99% of students identified as needing support were automatically matched to UNSW support services and received personalised nudges, with 41% opening these communications.
- Supported stakeholder engagement across **136 academic staff, 1,283 students**, and multiple university divisions, ensuring ASM aligned with academic workflows, institutional policies, governance requirements, and user needs. The platform achieved 70% active adoption among teaching staff, with users reporting an average 2-hour weekly time saving.
- The project received international recognition, winning the **Gartner Eye** on Innovation Award for Education, the **QS Reimagine** Education Award, and being featured as a **UNESCO** Digital Education Impact Case Study, recognising its innovation and measurable impact on student success.



Lead Data Scientist (Predictive Crop Monitoring System)

InvertiGro

05/2023 - 05/2023 Sydney, Aus

- Developed a comprehensive organizational AI strategic roadmap and action plan in collaboration with key project stakeholders
- Collaborated with Engineering team to test and develop a novel crop monitoring system
- Designed and trained machine learning models for predictive maintenance, predicting crop yield, and detecting real-time plant stress using IOT sensor data
- Managed end-to-end ML workflow involving model training and developing cloud infrastructure on databricks platform to support best practices around MLOps
- Worked in collaboration with data engineers to implement CI/CD pipelines for the automatic build, test, and deployment of ML models and ML training pipeline components

INDUSTRY EXPERIENCE



Data Scientist (Semantic Search Engine)

Foretheta

📅 01/2017 - 12/2017 📍 Foretheta Spokane, Washington, USA

- Designed and implemented a scalable patent retrieval system trainindex.io by leading a team of remote engineers.
- Latent Semantic Indexing (LSI) and Latent Dirichlet Allocation (LDA) for improving search accuracy beyond keyword-based methods and captures latent relationships in patent text, addressing challenges like synonymy and polysemy.
- Worked with data engineers to design an end-to-end data processing pipeline with Apache Spark's MLlib and Gensim, optimizing patent text preprocessing, tokenization, and feature extraction, reducing processing time from 9 hours to ~30 minutes.
- Optimized search retrieval performance through benchmarking with TREC, CLEF, and NTCIR datasets, improving precision-recall trade-offs and ensuring robust search relevance for legal domain.
- Deployed an efficient query processing pipeline, integrating indexing and distributed computing to enhance search scalability for large-scale patent databases (1TB of text data).
- Replaced the legacy Java based Search System to reduce cloud training costs by around 30%



Data Scientist (Recommender System)

Patari

📅 07/2016 - 12/2016 📍 Patari - Lahore, PK

- Designed and implemented an LSTM-based sequential recommender system in TensorFlow, modeling user listening patterns for personalized multi-lingual and multi-genre music recommendations.
- Worked with Data Engineering to develop a scalable data pipeline for preprocessing and encoding implicit user interactions, optimizing session-based recommendations using embedding-based similarity models.
- Addressed cold-start and sparsity challenges by leveraging deep learning techniques, improving content-agnostic recommendations without explicit user ratings.
- Led A/B testing experiments on live production traffic, evaluating retention, diversity, and novelty metrics to refine the recommendation engine.
- Benchmarked model performance against traditional collaborative filtering approaches, demonstrating superior user engagement and recommendation accuracy.
- Defined project roadmap, and facilitated agile methodologies to enhance collaboration with the Engineering team



Software Engineer (BlockChain dev.)

Redbrick Systems

📅 01/2012 - 01/2014 📍 Islamabad, PK

- Contributed to the development of the Onename identity system, enabling users to create secure, blockchain-tied profiles for decentralized and trustless identity verification.
- Designed and implemented a decentralized domain registration system for the .id namespace on Blockstack, allowing users to register usernames and manage blockchain-powered profiles.
- Developed a high-performance user profile search system for Onename using Python, Flask, Elasticsearch, Memcached, and MongoDB, improving search efficiency and retrieval speed.
- Designed and built scalable web crawlers for an E-Commerce Search Engine, leveraging Python to extract, index, and analyze product data across multiple online marketplaces.

TEACHING EXPERIENCE



Lead Instructor, Code Camp

📅 2017

- Taught basic computer programming to primary and secondary school students through Sydney Code Camp, supporting learners from diverse backgrounds and building confidence in foundational coding skills.
- Delivered engaging, hands-on coding activities using age-appropriate programming tools, helping students develop problem-solving, computational thinking, and digital literacy skills through project-based learning.
- Guided students in designing and building their own games and interactive digital projects, encouraging creativity, collaboration, and confidence in using technology.
- Adapted teaching approaches to meet the needs of students with different learning styles and prior coding experience, fostering an inclusive and supportive learning environment



Teaching Assistant, University of Wollongong (UOW)

📅 2016

- Tutored **50+ undergraduate and graduate students** in Service-Oriented Software Engineering through practical laboratories covering **SOA, RESTful APIs, microservices, distributed systems**, and enterprise software architecture.
- Mentored students in designing, implementing, testing, and debugging service-oriented applications using **BPMN, TOGAF**, and industry-standard software engineering methodologies.
- Delivered hands-on instruction in service modelling, API development, business process modelling, software architecture, and collaborative software engineering practices.
- Facilitated technical discussions, code reviews, and problem-solving sessions that promoted best practices in software design, system integration, and maintainable application development.

TEACHING EXPERIENCE



Teaching Assistant, GIKI

2012

- Taught hands-on lab courses in Unix, Algorithm Design, C Programming, Python, and Compiler Construction.
- Supervised a team of Teaching Assistants, ensuring effective delivery of undergraduate Computer Science practical classes.
- Delivered 10+ programming practicals per teaching term for groups of up to 30 students, providing hands-on instruction and academic support.
- Collaborated with the Course Coordinator to improve teaching materials, practical exercises, and laboratory delivery.
- The Compiler Construction course was highly appreciated during a routine Engineering Council (PEC) visit.

PUBLICATIONS

Cross-Silo Process Mining with Federated Learning

Service-Oriented Computing: 19th International Conference (ICSOC)

Asjad Khan, Aditya Ghose, and Hoa Dam.

2021 https://link.springer.com/chapter/10.1007/978-3-030-91431-8_38

DeepProcess: Supporting Business Process Execution Using a MANN-Based Recommender System.

Service-Oriented Computing: 19th International Conference (ICSOC)

Asjad Khan, Hung Le, Kien Do, Truyen Tran, Aditya Ghose, Hoa Dam, and Renuka Sindhgatta

2021 https://link.springer.com/chapter/10.1007/978-3-030-91431-8_2

Decision support for knowledge intensive processes using RL based recommendations

International Conference on Business Process Management

Asjad Khan, Aditya Ghose and Hoa Dam,

2021 https://link.springer.com/chapter/10.1007/978-3-030-85440-9_15

Mining goal refinement patterns: distilling know-how from data

International Conference on Conceptual Modeling 2017

Santiputri M, Deb N, Asjad Khan, A. Ghose, H. Dam, Chaki N

2017 https://link.springer.com/chapter/10.1007/978-3-319-69904-2_6

HIGHLIGHTED RESEARCH PROJECTS

Predictive and Prescriptive Process Monitoring with Deep Learning

2020 DSL, University of Wollongong

- Led a research project on Memory Augmented Neural Networks (MANNs) for real-time operational support in business process execution, in collaboration with Applied Artificial Intelligence Institute (A²I²), Deakin University, and IBM Research India.
- Developed a novel process-aware suffix recommender system using MANN-based architectures to enhance prescriptive monitoring and improve decision-making for optimizing real-time business recommendations for process efficiency and operational intelligence.
- Conducted exploratory data analysis (EDA) using data visualizations and dimensionality reduction techniques to uncover insights and validate modeling approaches.
- Built and maintained multiple datasets, performing data collection, cleaning, and preprocessing for effective model training and evaluation.

Clinical Decision Support System using Deep Reinforcement Learning

2021 DSL, University of Wollongong

- Designed and developed a Clinical Decision Support System using Deep Reinforcement Learning to enhance decision-making in knowledge-intensive medical processes.
- Led the conception, planning, and prioritization of key project tasks, ensuring structured research execution and development.
- Preprocessed the large-scale MIMIC critical care dataset using AWS compute services, to prepare data for model training.
- Conducted extensive experimentation and modeling, developing a proof-of-concept Deep RL-based system that generates policy-driven recommendations for emergency and critical care scenarios.

SKILLS

ML Stack: Pandas,
Scikit-learn,
NumPy, Matplotlib
Proficient



Python
Programming
Proficient



Cloud: Apache
Spark, Databricks
Notebooks,
MLFlow, Delta
Lake, Delta Tables,
Azure
Intermediate



Deep Learning
Keras, Tensorflow, Intel
Coach



SQL, Flask Web
API, GIT, JIRA for
agile development
Intermediate



ADDITIONAL EXPERIENCE

Consulting

Catalysr

📅 2017 - Present

Technical consulting: Catalysr is a Sydney-based incubator for early-stage startups. My consulting work with Catalysr involves mentorship and working directly with startup founders, where we formulate product execution strategies and brainstorm solutions to any technical challenges which the founders may face during product development.

AI Meetup

CafeDSL: University of Wollongong

📅 2018 - 2022

I have organized and given several talks on AI, Big Data Tools and Large-Scale Machine Learning at UOW AI Meetup group.

Organizing Chair

Service Science Society

📅 2016

As a member of the Service Science Society Australia, I helped organize the Annual 2016 Service Science Innovation Awards Event