# Arpit Kapoor

PhD Candidate | Data Scientist
Phone: +61 412 982 005 | Email: kapoor.arpit97@gmail.com
Website: www.arpit-kapoor.com | GitHub (arpit-kapoor) | Google Scholar | LinkedIn

## Professional Summary

Results-driven Data Scientist with experience in building and deploying **AI/ML solutions across environmental science, logistics, and e-commerce sectors**. Proven track record in **reducing operational costs, improving prediction accuracy**, and **enabling data-informed business decisions**. Adept at working in cross-functional teams and communicating insights to stakeholders; currently completing a **PhD in Data Science**. Actively seeking data science roles where impact, innovation, and collaboration matter.

## Technical Skills

**Programming Languages**:

Python, C/C++, SQL

**Frameworks & Tools:**

Deep Learning: TensorFlow, PyTorch, Flax

ML & Numerical Computing: Jax, Scikit-learn, Spark

MLOps: MLflow, Docker, Kubernetes

**Machine Learning & AI:**

Supervised & Unsupervised Learning: Random Forest, Gradient Boosting, SVM, Bayesian Statistics

Deep Learning: Computer Vision (CNN), Autoencoders, Seq2Seq models

Reinforcement Learning: Deep Q-Learning, Policy Gradient Methods

Retrieval Augmented Generation (RAG)

**Cloud & DevOps:**

AWS, Azure and GCP

High-Performance Computing (HPC)

Big Data Processing: Dask, PySpark, DataBricks

## Professional Experience

**Bureau of Meteorology, Australia** Feb 2023 – Apr 2025, Part-Time
Research Support Scientist

- Implement multivariate bias correction for climate models for the **Australian Climate Service**.
- Developed a scalable **Python interface** for legacy FORTRAN-based models
- Integrated a **Dask-based** distributed computing approach for improved efficiency.
- Reduced simulation runtime by **80%** on the National Computing Infrastructure (NCI)

**Quince, India** Mar 2022 – Aug 2022, Full-Time
Data Scientist

- Developed customer churn models leading to a **15% lift in repeat user engagement**.
- Designed an **ML-based logistics cost forecasting** model that estimated a shipping **cost reduction of 10%**
- Led stakeholder engagement to align modelling outcomes with operations teams

**3Qi Labs, India** Nov 2019 – Nov 2021, Full-Time
Data Scientist

- Designed and deployed ML-based **anomaly detection** for data pipelines, **boosting detection accuracy by 40%**.
- Integrated automated ML pipelines for data quality monitoring, saving **50+ analyst hours/month**.
- Automated ML workflows on the DataBricks platform using PySpark and related technologies.
- Introduced and deployed **MLOps infrastructure (MLflow & Docker)** for better model lifecycle management.

**Bomotix, India** Jan 2019 – Nov 2019, Full-Time
Machine Learning Engineer

- Developed a real-time **computer vision pipeline** for efficient tracking of player performance in sports
- Developed, trained and maintained various deep learning based computer vision models.
- Improved model efficiency using mixed-precision and GPU optimisations, leading to a **15% reduction in model training costs**
- Cleaned, processed, and curated video data for model training and validation
- **Maintained CI/CD pipelines for deep learning** model development and deployment

---

## Education

***Doctor of Philosophy (PhD), Mathematics and Statistics*** Aug 2022 – Expected Feb 2026
University of New South Wales, Sydney
Thesis: ***Enhancing process-based hydrological modelling with deep learning***

***Bachelor of Technology (B.Tech)*** Jul 2015 – May 2019
SRM Institute of Science and Technology, India
Major in **Computer Science & Engineering**
Extra Curricular: Team Leader of SRM Team Humanoid (student-led humanoid robotics team)

---

## Other Achievements

- Selected Participant, **Information Resilience PhD School**, University of Queensland, Brisbane (2024)
- **Challenge Facilitator,** Data Study Group - Theyr Challenge, **Alan Turing Institute** (UK, 2024)
- PhD Scholar, Australian Research Council Training Centre in Data Analytics for Resources and Environment (DARE) Centre & UNSW **University International Postgraduate Award (UIPA) Grant**
- Finalist, IEEE/RSJ IROS 2017 Humanoid Application Challenge (Canada)
- **Team Lead, SRM Team Humanoid** (Robogames 2017 – 1 Gold, 2 Silver, 1 Bronze)
- Published in several tier one machine learning and engineering journals - more details on personal website and Google Scholar profile

---