

# Kaartik Issar

[Personal Website](#) | [Linkedin](#) | [Github](#)

kaartik.issar@alumni.utoronto.ca | +1 437 (984)-8553 | Toronto, Ontario

## EDUCATION AND HONOURS

University of Toronto - *Computer Science Major, Statistics Major*

*Sept. 2020 - Apr.2024*

*Grade : High Distinction (83.3% Avg, cGPA: 3.67/4.0, Final two years GPA: 3.71/4.0 )*

*Dean's List (2020-21, 2021-22, 2022-23, 2023-24)*

Teaching Assistant at Department of Statistics, Research Assistant at Department of Computer Science

Winner - UNESCO - Piaget Hackathon 2022; \$6,800 in cash prize

## TECHNICAL SKILLS

Languages & Tools: Python, SQL, R, Julia, VB, Django, MATLAB, SPSS, Java, JS/HTML/CSS, C++, Power BI, Tableau, Knime Analytics, NumPy, Scikit-learn, Git, GCP, AWS, GraphQL, Spark, REST, Vertex AI

## PROFESSIONAL EXPERIENCE

### **Bell Canada**

Toronto, Canada

*Data Engineering Developer*

*July 2024 - Present*

- Built data pipelines in GCP and SAS to migrate 10M+ customers and 10,000+ agents to Bell's new Contact Centre AI Platform.
- Fine tuned voice transcription models (eg: Chirp2) and built regex pipelines to redact sensitive info in customer agent calls.
- Developed anomaly detection using Isolation Forest and statistical thresholds to flag data discrepancies in consumer data.
- Built a code translation model using Vertex AI and Transformer architecture to convert SAS ETL to GCP workflows.

### **Akiko Sherman Infotech**

Remote, India

*Founding Data Scientist*

*Apr. 2024 - Sept. 2024*

- Led development of an AI driven diagnostic platform that analyses patient reports to assist doctors in early disease detection.
- Designed ML pipelines for feature extraction from structured lab data and unstructured clinical notes using NLP techniques.
- Deployed a clinical insights dashboard using Streamlit and FastAPI, enabling real time predictions for partner hospitals.
- Designed data pipelines with GCP, BigQuery, and Airflow to process large volumes of hospital data securely and efficiently.

### **Kuarts Inc**

Toronto, Canada

*Data Science Intern*

*Jan 2024 – Apr. 2024*

- Developed an NLP powered data pipeline for financial data extraction, converting semi structured reports into machine readable tables with >92% accuracy.
- Leveraged Adobe API and advanced OCR for automated text extraction from diverse document formats & scanned images.
- Integrated machine learning models (decision trees, random forests) to enhance data categorisation and extraction robustness.

### **Bell Canada**

Toronto, Canada

*AI/Data Engineering Intern*

*May 2023 – Aug. 2023*

- Built forecasting models using ARIMA and Prophet to predict SLA adherence and job success rates over time.
- Built Airflow DAGs and support tasks to enable seamless model deployment on GCP by other teams.
- Created and analysed log based metrics in GCP to identify job failure clusters and high resource usage jobs.
- Developed a MicroStrategy dashboard to monitor SLA adherence and job success rates.

### **Akiko Sherman Infotech**

New Delhi, India

*Data Science Intern*

*Apr. 2022 - Aug.2022*

- Implemented bias and calibration checks for supervised risk models (regularised logistic regression) in medical setups.
- Built data quality layers for hospital records using distributional shift tests (Kolmogorov Smirnov and chi-square tests).
- Contributed to a core ML analytics prototype that helped secure a \$120,000 government contract.

## RESEARCH EXPERIENCE

### **Cluster Innovation Centre (University of Delhi)**

*Data Science Researcher*

*Aug 2024 - Oct 2025*

- Built EDA→MI→Linear→GBM→NN pipeline on educational records to model non linear achievement drivers.
- Designed a teacher oriented AI framework translating ML outputs into pedagogically relevant insights - linking learner archetype clustering to targeted tutoring, attendance campaigns, and resource planning.
- Utilised Neural Networks to reveal hidden interactions (e.g., attendance × parental involvement) missed by linear baselines.
- Authored an independent manuscript (under review) on supervised and ensemble models to forecast performance for early interventions.

### **National Council of Educational Research & Training (Ministry of Education, Govt. of India)**

*ML Research Assistant*

*Jan 2023 - July 2023*

- Modelled literacy progression using nationwide demographic and assessment datasets.
- Built scalable preprocessing (missing-data imputation, feature engineering, PCA) and unsupervised segmentation (K-Means, DBSCAN, hierarchical clustering).
- Trained predictive models (Gradient Boosted Trees, Logistic Regression, Random Forest) with hyper-parameter tuning.
- Conducted survival analysis (Cox PH, Kaplan–Meier) to estimate time-to-dropout; applied Bayesian inference & hypothesis testing to assess interventions (digital aids, peer groups, phonetic modules).

### **CHAI Labs (Department of Computer Science, University of Toronto)**

*Research Assistant*

*Jan 2022 - Apr 2022*

- Built a secure ETL pipeline using Dexcom and Fitbit APIs; unified schemas, data quality checks.
- Automated participant operations with a discord bot for IRB approved reminders and adherence tracking.
- Set up a bias minimised survey (validated scales, pilot, branching/randomisation, attention checks) with a codebook and power plan; also managed participant device distribution and inventory.

## Cluster Innovation Centre (University of Delhi)

Research Assistant

Apr 2020 - Sept 2020

- Co-authored a Springer published study on geometric tilings and their complex number formulations.
- Implemented complex number transformations and matrix driven geometric algorithms in MATLAB to model pentagonal tilings.
- Developed simulation scripts to verify tiling constraints, generate visualisations, and support reproducible computational experiments.

## PUBLICATIONS

### Supervised and Ensemble Methods for Early Academic Adaptive Intervention (Under peer review at Journal Revista de Psicología)

Abstract: [Link](#), Peer Review: [Link](#)

This study demonstrates how supervised learning frameworks can be deployed by educational institutions to proactively forecast student performance and trigger early interventions. Leveraging a realistic, temporally stratified synthetic dataset, the work benchmarks regularised linear models against gradient boosted ensembles under rolling origin evaluation.

### [Application of Complex Numbers and Matrix Transformations in Pentagonal Tiling](#), Springer Vol 25, Pages 1369-1384

The article delves into the realm of application of complex numbers and matrix transformations in pentagonal tiling. A basic understanding of the concepts in these two branches helps to plot the tiles in any computer algebra system (CAS). The application of the concepts has been demonstrated for the special case of Type 13 pentagonal tiling.

## TEACHING EXPERIENCE

### University of Toronto

Toronto, Canada

Teaching Assistant (STA313 - Data Visualisation)

Jan 2024 – Apr. 2024

- Led tutorials on EDA and data visualisation using ggplot2, R Shiny, Altair, and Tableau for 100+ students.
- Mentored student projects on visual storytelling, design practices, and data ethics.
- Developed grading rubrics and evaluated assignments on data wrangling, chart grammar, and reproducibility.
- Provided 1:1 support via office hours on improving visualisation and code optimisation.

## PROJECTS

### [MajorMatch.ai](#) | NumPy, Scikit-learn, XGBoost, Git

- Built and deployed a Balanced Random Forest based recommender with class balancing and standardised features to map student personality profiles to suitable undergraduate majors.
- Engineered a pipeline for data preprocessing, probability scoring, and top 3 major ranking with Big Five interpretability.
- Delivered 500+ personalised reports for rural Delhi students, reducing mismatches in stream selection.

### [Semantic Climate - CEVOpen](#) | Docker, Git, spaCy, Pytesseract

- Built pipelines to extract and parse phytochemical literature from EuropePMC using automated tools.
- Developed NER and dictionary matching modules to link compounds, plants, and activities to Wikidata.
- Implemented classification models to categorise annotations and resolve annotator disagreements.
- Dockerized the workflow for reproducible end to end data processing and analysis.

### [Intraday Stock Breakout Classifier](#) | Python, Yfinance API, AsyncIO

- Built a real time intraday feature pipeline (volatility ratios, intrabar expansion, range compression) for breakout signals.
- Engineered a low float universe using fundamental and liquidity constraints, structured for downstream ML models (GBDTs, logistic, contextual bandits).
- Implemented an asynchronous classifier like decision engine over streaming 1 minute microstructure data using the IB API.