

# Somneel Kumar Saha

[e1129797@u.nus.edu](mailto:e1129797@u.nus.edu) | [LinkedIn](#) | [GitHub: SomneelSaha2004](#) | (+91) 9811199300 | (+65) 96736925

---

## Education

**National University of Singapore (NUS), Singapore**

**Bachelor of Computing** (Computer Science - Specialising in AI, Database Systems) **2023–2027 (Expected)**

Coursework(Yr1, Yr2): Machine Learning, Big Data Systems, Data Structures and Algorithms, Software Engineering, Database Systems ;

Coursework Planned (for Yr3, Yr4): Neural Networks and Deep Learning, Uncertainty Modelling in AI, Knowledge Discovery and Data Mining

**The Shri Ram School - Aravali, Gurgaon, India**

**ISC: 97.5% (2023), ICSE: 98% (2021)**

---

## Work Experience

**MathLogic Consulting, Gurgaon, Machine Learning Intern**

*May 2025 – Present*

- Researching tabular-data applications of Transformer architectures (e.g., TabPFN, TabICL): curate stratified samples of enterprise datasets and optimise inference pipelines for memory efficiency.
- Conduct detailed performance and calibration assessments—such as decile-based analysis—to guide deployment decisions.

**MathLogic Consulting, Gurgaon, India, Data Science Intern**

*Dec 2024 – Jan 2025*

- ML pipelines for farmland segmentation and crop detection using satellite imagery and geospatial data.
- End-to-end data preprocessing of satellite imagery, ensuring high-quality inputs.
- Exploratory data analysis to extract and engineer key spectral and textural features for improved model performance.
- Evaluated machine learning and deep learning architectures (e.g. CNN, UNet) for segmentation accuracy.

**ContractKen Inc., Bangalore, India, Software Engineering Intern**

*May 2024 – Jul 2024*

- Worked at ContractKen, an AI-driven platform providing advanced legal contract analysis and document insights.
- Feature to extract legal definitions with Regex and track their usage within the Microsoft Word plugin.
- Bulk-upload pipeline in React.js, enabling clients to import large contract sets and build robust legal repositories.

**ContractKen Inc., Bangalore, NLP Research Intern**

*May 2022 – Jul 2022*

NLP research and testing, specifically evaluating the Python package LexNLP for Named Entity Recognition tasks on legal contracts.

---

## Technical Skills

**Languages:** Python, Java, SQL

**Frameworks/Libraries:** PyTorch, Pandas, Spark, Hadoop, React.js

**Tools:** Git, PostgreSQL, Linux, Github

---

**Certifications(Coursera):** Deep Learning Models with TensorFlow, End-to-End ML with TensorFlow on GCP, Machine Learning with Python

---

## Projects

**MedGen — Synthetic Data Generation for Healthcare Applications using LLMs**

*Jan 2025–April 2025*

Our team built an LLM-guided framework that generates high-fidelity synthetic diabetes datasets while preserving medical feature dependencies and patient privacy. I led experiment and metric pipelines, aggregated results and co-authored the technical report.

**FitFriends**

*Jan 2025–April 2025*

Our team built a cross-platform fitness-management desktop app, designed to help sporty young adults connect, organize and participate in sports activities together.

---

## Leadership and Honors

- Chairperson : Ad-Hoc National Security Cabinet, Aravali Model UN 2020
- Best Delegate : Le Martiniere Model UN (Nov 2020), Bhavan Vidyalaya Chandigarh Model UN (Jun 2020)

## Extra-curricular Activities

- NUS Sheares Hall Member - Squash, Softball 24/25 and 23/24
  - The Shri Ram School, Gurgaon, India: Member - North West ISC Regional Cricket Team 19/20
-