

---

# Shahil Karki

Phone: 0424718577 | E-mail: [shahil.karki@students.mq.edu.au.com](mailto:shahil.karki@students.mq.edu.au.com) | Address: Sydney, Australia |  
LinkedIn: <https://www.linkedin.com/in/shahil-karki-8660592b9> | Website: shahilkarki.com

---

## SUMMARY

Results-driven **Data Scientist / Data Engineer** with strong analytical skills and a deep passion for extracting insights from complex data. Currently completing a Master's in Data Science at Macquarie University, with hands-on experience in machine learning, big data technologies, and data pipeline development. Skilled in designing scalable data solutions, cleaning and transforming raw data into actionable insights, and deploying models in production environments. Proficient in using statistical methods, programming, and cloud platforms to drive decision-making. Experienced in working with cross-functional teams and clearly presenting data findings to help guide decisions.

## TECHNICAL SKILLS

<b>Programming &amp; Data Analysis:</b>	Python (NumPy, Pandas, Scikit-learn, PySpark), R, SQL, Bash, Jupyter Notebooks
<b>Machine Learning &amp; AI:</b>	Supervised & Unsupervised Learning, Model Evaluating & Tuning, Scikit-learn, TensorFlow, Keras, NLP, Time Series Forecasting
<b>Big Data &amp; Cloud Platforms:</b>	Apache Spark, Hadoop, AWS (S3, EC2, Lambda, Redshift), Microsoft Azure, Google Cloud Platforms (BigQuery, Dataflow)
<b>Data Engineering &amp; ETL:</b>	Airflow, Apache Kafka, Spark Streaming, Data Wrangling, Data Modeling, ETL Pipelines, Data Warehousing
<b>Databases:</b>	MySQL, PostgreSQL, MongoDB, Snowflake
<b>Data Visualization &amp; Reporting:</b>	Power BI, Tableau, Matplotlib, Seaborn, Plotly
<b>DevOps &amp; Version Control:</b>	Docker, Git, GitHub, CI/CD, MLflow
<b>Tools &amp; Platforms:</b>	Jupyter, VS Code, Linux, Anaconda, Postman

## PROFESSIONAL EXPERIENCE

**HURDEC Pvt. Ltd**  
**Associate Data Analyst**

Kathmandu, Nepal  
February 2021 – April 2023

- Contributed technical expertise in data cleaning, transformation, and validation to ensure the accuracy and consistency of datasets used in development research and policy studies.
  - Designed and developed data tables, summary statistics, and visual reports (charts, graphs dashboards) to support evidence-based decision making in social and institutional development projects.
  - Collaborated with researchers and field teams to analyze survey data and extract key indicators relevant to health, education, governance, and rural development.
  - Prepared structured datasets and metadata documentation for project reporting, publication, and compliance with donor and government data standards.
  - Utilized tools such as Excel, SPSS, and Python to process large-scale survey and administrative datasets across multiple project assignments.
  - Assisted in drafting data-driven insights and visuals for presentations, proposals, and final reports submitted to clients, including government agencies and international development partners.
-

## Rooster Logic Internship

Kathmandu, Nepal  
July 2020 – December 2020

- Contributed to the development of a data visualization tool to display solar and lunar events (e.g., Dawn, Sunrise, Sunset, Dusk) for any geographic location and specific time.
- Utilized the Python module Astral to compute astronomical timings based on geographic coordinates and time zones.
- Assisted in building the backend logic to process user input and dynamically generate solar cycle data, with the user interface developed using the Flask web framework.
- Supported the development and testing of REMO (Research and Monitoring System), a mobile-based research and monitoring system, by assisting in form design, data validation workflows, and location/image-based input functionalities.
- Worked with the technical team to connect the mobile app with the backend system, helping enable real-time data collection and reporting.

## Deerwalk Institute of Technology Final Year Project – Music Genre Detection using MFCC and FFT

Kathmandu, Nepal  
July 2019 – December 2019

- Developed a music genre classification system using Mel-Frequency Cepstral Coefficients (MFCC) and Fast Fourier Transform (FFT) for feature extraction.
- Implemented Support Vector Machine (SVM) algorithm to classify audio signals into distinct music genres with high accuracy.
- Designed the application using Python and deployed the user interface via the Flask web framework.
- Preprocessed audio datasets, extracted key spectral features, and visualized data for analysis and model tuning.
- Evaluated model performance using precision, recall, and F1-score to optimize classification results.

## EDUCATION & CERTIFICATION

**Master's in Data Science** / Macquarie University, North Ryde, Sydney | 2024 - running  
**Bachelor of Science in Computer Science and Information Technology** /  
Deerwalk Institute of Technology, Nepal | 2020

## PERSONAL ATTRIBUTES

- |                             |                              |                                    |
|-----------------------------|------------------------------|------------------------------------|
| • Analytical Thinking       | • Technical Curiosity        | • Adaptability                     |
| • Problem-Solving Abilities | • Communication Skills       | • Time Management                  |
| • Data Storytelling         | • Teamwork and Collaboration | • Ethical and Responsible Data Use |

## LANGUAGE

English: Native Proficiency

---