Samrat Thapa

tsamrath40@gmail.com | Barrie, Ontario

Personal Website | LinkedIn | Medium | GitHub

Dedicated **Data Scientist and Data Analyst** with a keen interest in leveraging advanced technologies to drive impactful solutions. With over **2+ years of experience in ML, AI, Python, and SQL**, I have successfully helped companies enhance their operations and achieve their goals by harnessing the power of data and AI.

Skills

- **Programming Languages:** Python, R, SAS Hadoop, PySpark
- **Cloud Computing:** Azure, AWS
- **Data Visualization:** Power BI, Tableau, Plotly, Matplotlib, Seaborn
- **Database:** SQL, MongoDB, Cassandra
- **ML and AI**: sklearn, Keras, TensorFLow, PyTorch
- **Deployment:** Flask, Azure Machine Learning Studio, Streamlit
- **Version Control**: Git and Github
- **Front End**: HTML, CSS, JS
- **Data Handling and Scraping**: Pandas, Excel, Selenium, BeautifulSoup

Education

Artificial Intelligence – AIDI
Georgian College • Dec 2022 • 89.94%
Big Data Analytics (BDAT)
Georgian College • Dec 2021 • 90.58%
B. Tech in Biotechnology
Kathmandu University • Nov 2019 • 81%

Course Work

- AI Algorithms and Mathematics
- Vision System
- Neural Networks
- Data Visualization
- Data Manipulation Techniques
- Data System Architecture

Certification

- Data Science Math | Duke University
- Machine Learning | Stanford University
- AWS Academy Cloud Foundation
- AWS Academy Machine Learning Foundations
- SAS Enterprise Miner

Work Experience

Skinopathy | AI Engineer + Team Lead

Sept 2022 to Dec 2022 | Internship | Toronto, Canada |
- Achieved 95% accuracy with 0 false positives by leading a team of 4 to develop a lesion detection system for a CNN model in collaboration with the Skinopathy team
- Improved CNN classifier accuracy from 81% to 89% by optimizing hyperparameters and model design, reducing the Skin Cancer risk.

Beenova AI | Data Analyst

June 2021 to Dec 2021 | Internship | Toronto, Canada |
- Developed a 90% accurate ML model detecting elephant poaching in Southern Africa, aiding conservation and reducing poaching by 10X.
- Created training data for the model by consolidating data from 15+ sources.

Cygnus Tech Nepal | Jr. Data Analyst

July 2020 to Dec 2020 | Kathmandu, Nepal |
- Successfully collected data through web scraping for research and client needs.
- Revamped client's dashboards, resulting in a 20% increase in satisfaction and a 15% improvement in issue resolution.

Projects

Eye Classification | GitHub | Medium | App |
- Designed an 88% accurate End-to-End CNN Gender Classifier from eye images, honing deployment skills for large models.
- Technology Used: Streamlit, OpenCV, TensorFlow, Keras, scikit-learn

EDA on Seattle Airbnb Data | GitHub | Medium |
- Helped a friend in Seattle vacation planning via trend analysis, pricing research, availability and Airbnb reviews saving him about 500$ in that vacation.
- Technology Used: Matplotlib, NLP, TextBlob, Pandas, Seaborn, NLTK

Sentiment Analysis of Amazon App using Azure | GitHub | Medium |
- Developed an Azure Based End–to-End Sentiment Analysis pipeline starting from data scraping to data visualization using Azure Resources.
- Helped solidify understanding of Cloud Services
- Technology Used: Python, Azure Resources like Blob Storage, Data Factory, Azure SQL Database, PowerBI

German Credit Card Assessment App | GitHub | Medium | App |
- Built and launched a German Credit Card Assessment app with an 80% accurate Random Forest classifier, enabling users to determine their credit eligibility by answering questions.
- Technology Used: Streamlit, scikit-learn, Pickle, Pandas