

THAPA SANUS

☎ (+33) 07 53 83 83 32 | ✉ sanuzh.thapa@gmail.com | 🔗 linkedin.com/in/sanuz-thapa | 🌐 sanuzthapa.github.io

EDUCATION

Université Marie-et-Louis-Pasteur

Behavioral and Digital Economics for Effective Management

Besançon, 25, FR

Sep. 2024 – Present

Rungta College of Engineering and Technology

Computer Science and Engineering (CSE)

Bhilai, India

July 2015 – Sep 2019

SKILLS & COMPETENCIES

Programming & Data Science: Python, R, SQL, pandas, NumPy, scikit-learn, TensorFlow, tidyverse (R)

Machine Learning & NLP: Sentiment Analysis, Text Classification, Multimodal Analysis (Text + Image), OCR, Image Captioning, LLM Integration, NLP, Data Annotation & Labeling

Automation & Data Engineering: Pipedream (Workflow Automation), Web Scraping (Selenium), Data Cleaning & Preprocessing, ETL Pipelines, API Integration, Data Storage & Management

Analytics & Visualization: Power BI (Modeling & DAX), Advanced Excel, R Shiny, Data Storytelling, Dashboard Design

Tools & Platforms: Visibrain, Linux, Git/GitHub, Google Sheets, MySQL, PostgreSQL

Core Competencies: Communication, Presentation Skills, Time Management, Strategic Thinking, Teamwork

EXPERIENCE

Intern – Opinion Monitoring & Data Science

Meteco SAS

April 2025 – July 2025

Besançon (25), France

- Developed an automated data mining and monitoring pipeline in Pipedream, enabling real-time collection of social media data from over 5 platforms including Instagram, LinkedIn, YouTube, Bluesky, and CloudMedia, increasing data retrieval efficiency.
- Engineered and optimized data pipelines by filtering, cleaning, and structuring over 28k Instagram Visibrain datasets, enhancing machine learning model accuracy.
- Implemented LLM-based pipelines (ChatGPT API) for sentiment analysis, text extraction, and visual content interpretation with OCR & image captioning (GPT-4v Vision API / JinaAI SceneXplain).
- Conducted the “Migration Project”, leveraging NLP and multimodal data analysis to study online narratives around immigration policy.
- Integrated textual and visual embeddings to enhance sentiment and narrative detection, supporting model-driven insight generation.

Key Skills: Python · Data Mining · Pipedream Automation · NLP · LLMs (ChatGPT API) · Sentiment Analysis · OCR · Multimodal Learning · Social Media Analytics

Management Information System (MIS) Manager

Sunwarshi Municipality (Government of Nepal)

Dec. 2020 – Dec. 2021

Morang, Nepal

- Managed data workflows including cleaning, validation, preprocessing, modeling, and importing.
- Performed SQL-based data extraction, updates, and database operations for efficient management.
- Developed advanced Excel reporting tools, including pivot tables, queries, and macros.
- Designed and maintained Power BI dashboards to support data-driven decision-making.
- Implemented RPA tools to automate repetitive tasks and streamline operational processes.

CERTIFICATIONS

Complete A.I. & Machine Learning, Data Science Bootcamp — by Andrei Neagoie, Daniel Bourke on Udemy (Link)

PROJECTS

Impact of Scholarship on Academic Performance & Outcomes | *R, Econometrics* (Link)

- Conducted an econometric analysis to evaluate how scholarship allocation affects academic outcomes in the Master's BDEEM program.
- Investigated differences in academic performance (first- and second-year grades) between scholarship recipients and non-recipients.

Predicting Heart Disease using sklearn (RandomForestClassifier) | *Python, ML* (Link)

- Predicted Heart Disease using RandomForestClassifier with hyperparameter tuning (RandomizedSearchCV/GridSearchCV) to achieve 90% accuracy.

Predicting the Sale Price of Bulldozers using sklearn (RandomForestRegressor) | *Python, ML* (Link)

- Predicted the sale price of bulldozers using RandomForestRegressor with hyperparameter tuning and pipeline modeling.

PowerBI Interactive dashboard for Uniqlo | *PowerBI, DAX Modeling, Forecasting* (Link)

- Developed an interactive PowerBI dashboard for Uniqlo with KPIs, slicers, and visual analytics for sales, inventory, and performance tracking

Web Scraper and Summarizer with FastAPI | *Python, FastAPI* (Link)

- Developed a FastAPI-based web scraper that fetches the main content from a URL, summarizes it using OpenAI API, and stores results in Google Sheets for newsletter.