

BASANT KHANAL

AI & Machine Learning Engineer | Data Science

+977 9800749714 • me.basantkhanal07@gmail.com • Kathmandu, Nepal

[LinkedIn](#) • [GitHub](#) • basantkhanal.com.np

PROFESSIONAL SUMMARY

Final-year Bachelor of Information Technology student specializing in Artificial Intelligence and Machine Learning. Experienced in building LLM-powered applications, Retrieval-Augmented Generation (RAG) systems, and end-to-end ML pipelines. Hands-on with LangChain, Google Gemini, Hugging Face, vector databases, and FastAPI. Driven by a deep interest in applied AI — from intelligent chatbots to predictive ML APIs.

TECHNICAL SKILLS

Languages & Tools: Python, FastAPI, Git, GitHub

AI & LLM Stack: LangChain, LLM APIs, Hugging Face, Pinecone, RAG, Prompt Engineering

Machine Learning: Scikit-learn, Random Forest, Logistic Regression, KMeans, PCA, KNN, Decision Tree

Deep Learning & NLP: Neural Network Fundamentals, CNN, NLTK, Sentiment Analysis

Data Science: NumPy, Pandas, Matplotlib, Seaborn, Feature Engineering, EDA, Model Training

Databases: SQLite, Redis, Pinecone, FAISS, ChromaDB

PROJECTS

RAG System with Interview Booking Chatbot

Tools: Python · FastAPI · LangChain · Google Gemini · Pinecone · Hugging Face · Redis · SQLite · Streamlit

Link: [GitHub →](#)

- ▶ Architected a production-ready RAG system using LangChain and Google Gemini, capable of ingesting PDF/TXT documents and answering context-aware questions via a Streamlit frontend.
- ▶ Engineered a custom document pipeline with fixed and semantic chunking strategies, Google Gemini embeddings, and Pinecone vector storage for high-accuracy similarity search.
- ▶ Built a multi-turn LLM conversational engine with Redis-backed session memory, maintaining context-awareness across extended chat interactions.
- ▶ Designed an AI-driven intent detection module that extracts structured booking data (name, email, date, time) from natural language and persists records to SQLite via FastAPI.

Insurance Premium Prediction API

Tools: Python · Scikit-learn · FastAPI · Pandas · Pydantic

Link: [GitHub →](#)

- ▶ Developed an end-to-end ML classification system predicting insurance premium tiers (Low / Medium / High) with a custom-engineered synthetic dataset.
- ▶ Implemented advanced feature engineering: BMI calculation, age grouping, lifestyle risk scoring, and city tier classification to maximize model signal.
- ▶ Built a robust Scikit-learn pipeline (OneHotEncoder → ColumnTransformer → RandomForestClassifier) and deployed it as a production-ready FastAPI REST API.
- ▶ Added Pydantic-based input validation, health-check endpoints, and real-time inference — making the API integration-ready from day one.

Customer Behavior Insights & Churn Prediction

Tools: Python · Pandas · NumPy · Scikit-learn · NLTK · Matplotlib · Seaborn · Streamlit

Link: [GitHub](#) →

- ▶ Analyzed synthetic dataset of 300+ customer records with in-depth EDA to surface patterns in demographics, spending behavior, and purchase history.
- ▶ Benchmarked four ML algorithms (Random Forest, Logistic Regression, Decision Tree, KNN) for churn prediction, achieving 72% accuracy; evaluated with precision, recall, F1-score, and confusion matrices.
- ▶ Applied KMeans clustering for customer segmentation to support targeted marketing and revenue optimization strategies.
- ▶ Conducted NLP-based sentiment analysis on customer feedback using NLTK; delivered findings via an interactive Streamlit real-time prediction app.

TRAINING & CERTIFICATION

Data Science & Machine Learning Trainee

2025

[Deerwalk Institute of Technology](#) | Kathmandu, Nepal

- ▶ Trained on core ML concepts including supervised learning, model evaluation, and Scikit-learn pipeline construction.
- ▶ Gained hands-on experience with NumPy, Pandas, Matplotlib, and Seaborn for data wrangling and visualization.
- ▶ Explored deep learning fundamentals and NLP processing pipelines.
- ▶ Completed applied mini-projects involving real-world datasets, feature engineering, and insight extraction.

EDUCATION

Bachelor of Information Technology (BIT) — 4th Year

2022 – Present

[Patan Multiple Campus](#) | Patandhoka, Lalitpur

+2 Science Faculty

2020 – 2022

[New Horizon School](#) | Devi Nagar, Butwal

GPA: 3.01

Secondary Education Examination (SEE)

2010 – 2020

[Moon Light Secondary Boarding School](#) | Siddharthanagar

GPA: 3.75

REFERENCES

Available upon request.