SHULAV KARKI

Kathmandu, Nepal, +977 9815028713
✉ shulavkarki@gmail.com | Linkedin | GitHub | Portfolio

Highly motivated and detail-oriented Machine Learning Engineer with over 2 years of experience in implementing and deploying machine learning models (RecSys, CV, Fraud Detection., etc) for various applications. Proficient in Python, PyTorch, PySpark, Pandas and AWS cloud services. Skilled in building pipelines for data pre-processing, feature engineering, model building and model evaluation. Strong communicator and team player with a passion for learning and innovation.

EDUCATION

FUSEMACHINES — **Kathmandu**
Artificial Intelligence Fellowship — March 2022 - Present

- Full Merit **Scholarship** recipient
- Studied ML algorithms from scratch, engaged in weekly paper Reading Sessions and their implementation
- **Projects**: ML (SVM, Naive Bayes, Logistic Regression) for Parkinson's Disease Classification, Image Caption Generator and Multi language translator(English-German)

INSTITUTE OF ENGINEERING, IOE — **Dharan, Nepal**
Bachelor of Computer Engineering — May 2017 - Aug 2022

- Full Merit **Scholarship** recipient; maintained **First Division** grades throughout
- Teaching Assistant for AI, DSA and Big data course
- Final Year Project: "**Landslide Detection From Remote Sensing Imagery Using Autoencoder And Attention boosted CNN**" under Pravin Sangraula (Head of Department)
- **Relevant Coursework**: AI, Data Mining, Operating Systems, Big Data, DBMS, DSA, Probability and Statistics, ToC, Microprocessor, Distributed Systems, Network and Security, Embedded Systems, DSP

PROFESSIONAL EXPERIENCE

Fusemachines — **Kathmandu, Nepal**
Data Scientist / Machine Learning Engineer — Sept 2022 – Present

- Conducted large-scale data analysis tasks using PySpark for clients to extract insights
- Developed CTR prediction model and time-series approach for improving click through rates
- Designed and implemented landing page and next item recommendation, boosting sales by 20%
- Collaborated on research-grade multi-task recommendation systems under Dr. Pradip Mainali, implementing algorithms like ESM2 for personalized fashion recommendations
- Developed fraud detection models for single and multi-transactions, identifying 10% more fraudulent patterns than rule-based approach
- Worked on multi channel speech enhancement and its optimization(quantization, knowledge distillation) with a focus on edge computing
- Led technical interviews and mentored interns

Chulo Solutions — **Kathmandu, Nepal**
Machine Learning Intern — May 2022 – Aug 2022

- Image Classifier, Object Detection and Research paper reading session
- Worked on a project to count people and make a log of each profile in a video using YOLOv3
- Teamwork, Communication and Idea sharing session

SELECTED PROJECTS

Healthcare AI: RAG-based Chatbot
- Implemented Mistral(7B) quantized model with Selenium web scraping Cleveland Clinic
- Streamlit for chatbot interaction

Contextual and Sequential Food Recommendation | Fusemachines
- Developed context-based (DeepFM, FibiNet) and sequential (DSIN, DIN, SRGNN) recommenders
- Implemented Multi-task learning algorithm like ESM2

Single Channel Speech Enhancement
- The model reduces background noise, improves clarity, and increases intelligibility
- Tech stack used: Python, Numpy, Librosa, PyTorch

Interactive Personalized Product Recommendations
- Applied Thompson Sampling for exploration and developed a visual demo of distribution updates

Landslide Detection From Remote Sensing Imagery Using Autoencoder And Attention boosted CNN
- Achieved 20% accuracy boost over pre-trained CNN by fine-tuning encoder and CNN
- Implemented Squeeze and Excitation attention, further improving accuracy by 5%

TECHNICAL SKILLS

Programming Language: Python, SQL, C, C++
Libraries: Tensorflow, Keras, Scikit-Learn, Pandas, Dask cuDF, NumPy, Matplotlib, Seaborn
Framework: PyTorch, PySpark, Flask, FastAPI
Database: MySQL, PostgreSQL
Deep Learning Models: MLP, CNN, RNN, LSTM, Transformer, BERT
NLP: SpaCy, NLTK, Gensim, LLM, Chatbot, Prompt Engineering
Version Control System: Git, BitBucket
Ops: MlFlow, Weights & Biases, Evidently
Model Optimization: Quantization, Pruning, Low-rank Matrix Factorization, Knowledge Distillation
Cloud Deployment: AWS(EC2, Redshift, S3, Athena, EMR, Lambda), GCP(Big Query, Cloud Storage), IAM
Domain: RecSys, Computer Vision, NLP, Audio

LEADERSHIP AND EXTRACURRICULAR ACTIVITIES

- Workshop on Git and Intermediate Python for Freshmen
- Workshop on Machine Learning with Python for Sophomores
- Volunteering and Organizing Internal Hackathons Within Campus

AWARDS AND HONORS

AI Hackathon(Winner), Fusemachines (2023)
Appreciation, Fusemachines (2023)
Full Scholarship, AI Fellowship, Fusemachines (2022)
Delta Hackathon(Runner-up), Institute of Engineering, Nepal (2019)
Full Merit Scholarship, Institute of Engineering (2017)