

Milan Gautam

Ratnanagar-10, Tandi
Chitwan
Nepal

Website: <https://gautammilan.com.np>
E-mail: milangautam193@gmail.com
LinkedIn: <https://www.linkedin.com/in/milan-gautam/>
Github: <https://github.com/gautammilan>
Phone: +977 9865138476

EXPERIENCE

AI Engineer
Palua.AI Ltd, London
06/2022- Present

- Developed a Nepali language model, NepBERTa, and its benchmark, Nep-Glue, and evaluated the model's performance in it.
- Working on active learning research in NLP and CV to make the annotation task faster and increase label efficiency.

EDUCATION

Bachelor in Computer Engineering
Pashchimanchal Campus, Institute of Engineering,
Tribhuvan University, Nepal

Nov 2017 - 2022

PROJECTS

Transformer-based Model for Nepali Language Model

- By scrapping 14.5 GB of text data (0.8 Billion words), a BERT base model (110 M parameters) was pre-trained from scratch on TPU v3-128.
- Introduced Nepali Language Understanding Evaluation (Nep-gLUE) benchmark which contains four tasks NER, Sentence Pair Similarity, Content Classification and POS Tagging.
- Finetuned the model on Nep-gLUE tasks.

Single Image Super-Resolution

- Implemented SR-GAN from scratch on Tensorflow.
- Used Wasserstein Distance with Gradient Penalty to enforce smooth training of the generator-discriminator network.

Nepali Speaker Recognition

- Scapped audio of 34 Nepali politician (male and female) from Youtube.
- Trained Siamese network with contractive loss to identify the individual from there short audio recordings.

Transaction Entity Extraction

- Trained a bidirectional LSTM model to extract entity from the transaction using cross entropy loss and dice loss
-

PUBLICATION

- Gautam, M., Timilsina, S., & Bhattarai, B. (2022). “*NepBERTa: Nepali Language Model Trained in a Large Corpus*”.
The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (ACL-IJCNLP)
-

TECHNICAL SKILLS

- Python with libraries used for machine learning and deep learning like *Pandas, Numpy, Scikit-Learn, Tensorflow, huggingface*.
 - Distributed training on *GPU* and *TPU*.
 - Cloud Computing with *Google Cloud Platform*.
 - Deployment platform like *streamlit* and *kubernetes*.
 - Version Controlling with *Git*.
 - Linux skills like shell scripting and familiarity with different Linux distributions.
-

LANGUAGES

- Nepali : Native/ Bilingual Proficiency.
- English : Fluent/ Full Working Proficiency.
- Hindi : Fluent/ Full Working Proficiency.