

# Pranjali Basmatkar

NLP Master's student with an unbridled appetite for artificial intelligence backed with 3+ years of industry experience designing, experimenting with, and developing innovative AI solutions for clients across the globe.



pranja.li



pranjali97



pranjali-basmatkar



pranb.mail@gmail.com



+1 (669)-877-4344

## EDUCATION

**M.S. Natural Language Processing** | UNIVERSITY OF CALIFORNIA, SANTA CRUZ  
CGPA : 3.9/4.0 | California, USA

SEP 2022 – PRESENT

**B.Tech. Computer Engineering** | NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES  
CGPA : 3.19/4.00 | Mumbai, India

JUL 2015 - MAY 2019

## SKILLS

Python • Pytorch • Keras • Fairseq • Transformers • LangChain • Python • Machine Learning • NLP • Large Language Models • AWS • HuggingFace • Streamlit • GCP • Docker

## EXPERIENCE

**Research Engineering Intern** | ALLEN INSTITUTE FOR ARTIFICIAL INTELLIGENCE  
Seattle, WA

JUN 2023-PRESENT

→ Research Engineering Intern with the AllenNLP team for the Summer of 2023.

**Senior Machine Learning Engineer** | QUANTIPHI ANALYTICS SOLUTIONS PVT. LTD.  
Mumbai, India

APR 2020 – AUG 2022

- **Spoken language identification** - Implemented a transformer-based model to identify distinct language segments on language-mixed audio files to improve transcription quality by 40 %.
- **Intent identification on call log data** - Developed a call log insight generation tool using BERT and LDA embeddings for topic modeling in the conversational AI domain.

**Machine Learning Engineer** | QUANTIPHI ANALYTICS SOLUTIONS PVT. LTD.  
Mumbai, India

MAY 2019 – MAR 2020

- **Credit risk analysis** - Feature engineered 300+ banking related features to build a boosting-based classification model
- **Anomaly detection on emails** - Solutioned an 84% accurate, ML-based data loss prevention software for a Canadian insurance organization

**Intern** | CRIMSON INTERACTIVE  
Mumbai, India

JULY 2018 – FEB 2019

- **Grammatical Error Correction** - Incorporated a grammatical error correction model based on Seq2Seq with Attention mechanism achieving an F1 score of 0.81 to replace their existing rule-based system using OpenNMT and FairSeq libraries.

## PROJECTS AND PUBLICATIONS

**Formality Controlled Neural Machine Translation** | ACL

FEB 2023 - APR 2023

- Formality-controlled NMT using the style-embedding intervention. Accepted in IWSLT, 2023.

**Multilingual text intimacy detection** | UCSC

OCT 2022 - DEC 2022

- Tuned Multilingual XLM Roberta to achieve Pearson's R of 0.82 using few-shot and zero-shot learning for SemEval 2023.

**Contextual Topic Modelling using BERT** | SPRINGER

JAN 2022 - MAR 2022

Proceedings of the Third ICCES, Springer, Singapore DOI: 10.1007/978-981-16-8862-1\_32

- Proposed a hybrid context-based approach to improve the overall quality of topic modeling tasks by performing qualitative analysis between existing related work.

**Neural Machine Translation for Multilingual Translation System** | IEEE

JAN 2019 - MAR 2019

Proceedings of Third ICCMC, IEEE. DOI: 10.1109/ICCMC.2019.8819788

- Proposed a Seq-to-Seq neural machine translation system tool between English and Hindi using attention.