

RAHUL TANGSALI

tangsali@usc.edu ◇ 213-675-5916 ◇ Google Scholar ◇ LinkedIn ◇ Portfolio ◇ GitHub

EDUCATION

University of Southern California

Los Angeles, CA

Masters in Computer Science (GPA: 3.85/4)

08/2023 - 05/2025

Courses taken: Analysis of Algorithms, Database Systems, Web Technologies, Machine Learning for Data Science

Pune Institute of Computer Technology

Pune, India

Bachelors in Computer Engineering (GPA: 3.73/4)

08/2019 - 06/2023

Courses taken: Data Structures and Algorithms, Object Oriented Programming, Discrete Mathematics, Data Science and Big Data Analytics, Artificial Intelligence, Machine Learning, Natural Language Processing

TECHNICAL SKILLS

Programming Languages: Python, C++, JavaScript, SQL

ML/DL libraries: PyTorch, Tensorflow, scikit-learn, nltk

Development frameworks: HTML, CSS, Bootstrap, Flask, Streamlit, Linux, AWS, Azure LUIS, Github

WORK EXPERIENCE

L3Cube Pune

Pune, India

NLP Research Intern

06/2022 - 06/2023

- Led creation of Marathi's largest sentiment analysis dataset, L3Cube-MahaSentMD of 60000 training examples, with research published at DMLR workshop, ICML 2023.
- Innovated data augmentation methods for sentiment analysis using named entity-based masking, back-translation and GPT-based methods.
- Led research in news article summarization by fine-tuning on top of SOTA architectures such as PEGASUS, BRIO, T5 and IndicBART, with top 3 ranks in workshop leaderboard and research published in FIRE 2022.

Pune Institute of Computer Technology

Pune, India

Undergraduate Research Assistant

10/2021 - 05/2023

- 1st place at DravidianLangTech, ACL 2022 for research work in neural machine translation, which included training Seq2Seq models such as LSTMs, CNNs and transformers from scratch.
- Achieved accuracy of 92% on abstractive document summarization for medical literature using models such as BART-large, DistilBART and T5-base, with research at COLING 2022.
- Implemented unsupervised methods for NMT on German and Sorbian, with research at EMNLP 2022.

ResoluteAI.in

Bengaluru, India

ML Intern

10/2021 - 01/2022

- Engineered a staff productivity tracking system using YOLO-v4; used PowerBI, Streamlit for displaying insights.
- Created a ML pipeline featuring labeling for data annotation, and fine-tuning ResNet50 architecture for towel defect detection, achieving 96% accuracy.

PROJECTS

Strategic Talent Retention using ML

scikit-learn, pandas, Flask, Python, nginx, AWS

- Trained a XGBoost classifier to predict employee retention; Achieved 99.23% accuracy with GridSearchCV.
- Developed REST APIs using Flask for model training and prediction.
- Utilized HTML, CSS, JS and Bootstrap for developing the front-end, with application hosted on AWS EC2 and nginx.

Image Captioning Systems for Dravidian Languages

Python, PyTorch, nltk

- Developed attention-based merge architecture models for image captioning in 4 Dravidian languages.
- Designed an encoder-decoder system with ResNet50, VGG16, and InceptionV3 as image feature extractors (encoders) and LSTM, Simple RNNs, and GRUs as caption generators (decoders).
- Trained a total of 36 models to optimize BLEU scores, with research published at CGIP 2023.

EXTRA-CURRICULARS

PICT ACM Student Chapter (PASC)

Pune, India

Volunteer and Technical Contributor

01/2020 - 04/2023

- Managed organization of ML events as part of PASC's annual event, drawing 500+ students across Pune.
- Recognized with an award for outstanding research enthusiasm in 2022.