

NIKHIL RANJAN

SENIOR NLP ENGINEER, MBZUAI, ABU DHABI

✉ nikhilranjan7@gmail.com | [nikhilranjan7.github.io](https://github.com/nikhilranjan7) | [nikhilranjan7](https://www.linkedin.com/in/nikhilranjan7) | [in nikhilranjan7](https://www.linkedin.com/in/nikhilranjan7)

PUBLICATIONS AND PATENTS

DEC 2023 Published	LLM360: TOWARDS FULLY TRANSPARENT OPEN-SOURCE LLMs (MBZUAI) <i>A framework for open-source LLMs to foster transparency, trust, and collaborative research</i>
DEC 2020 Submitted	DISRUPTION DISCOVERY TOOL (ELISA R&D) <i>Patent submitted at Elisa for automated topic mining related to latest disruptions</i>
DEC 2020 FI20206334	NATURAL LANGUAGE PROCESSING MODEL ADAPTATION (ELISA R&D) <i>Patent filed in Finland for novel chat intent recognition technique for agglutinative languages</i>

EXPERIENCE (NLP, ASR, RECOMMENDATION)

JUL 2022 - Ongoing	MBZUAI <i>Senior NLP Engineer, Abu Dhabi</i> <ul style="list-style-type: none">Working with CMU CASL project in Forte and AdaptDLPretraining and deployment of LLMKey Technologies: Python, C++, Pytorch
APR 2021 - Jul 2022	BYTEDANCE TIKTOK <i>Machine Learning Engineer, Singapore</i> <ul style="list-style-type: none">Worked on Short and Live Video recommendation systemsKey Technologies: Python, Java, C++, Tensorflow
Nov 2020 - Mar 2021	INTEL LABS EUROPE PRIVATEAI COLLABORATIVE RESEARCH INSTITUTE <i>Private AI Researcher, Germany</i> <ul style="list-style-type: none">Worked on privacy-preserving machine learning with PrivateAI Collaborative Research Institute LinkDeveloped federated learning and watermarking tools to increase robustness of current deployed systemsSuccessfully converted workflow of a major network anomaly detection ML system to utilize decentralized architecture without performance degradation
JUN 2019 - Mar 2021	ELISA OYJ FCAI - FINNISH CENTER FOR ARTIFICIAL INTELLIGENCE <i>Senior Research Engineer, Helsinki</i> <ul style="list-style-type: none">Involved in Private AI Research in collaboration with Intel LabsInvolved in Speech Recognition and Conversational AI team in collaboration with Aalto UniversityDeveloped product pipelines to automate customer support based on latest research in the field of ASR and NLPAchieved best result for Finnish language on multiple open-source and internal dataset and attained significantly lower WER compared to Google and Microsoft products Link-1 Link-2Key Technologies: Kaldi, Transformers LM (BERT, GPT, GPT-2), Pytorch, Docker, Kubeflow, Streamlit
AUG 2020	HUAWEI R&D <i>Speech Algorithm Engineer, Helsinki</i> <ul style="list-style-type: none">Worked on Huawei Petal Search TTS and ASR for Finnish languageKey Technologies: Tacotron-2, ESPnet, Pytorch Kaldi, Attention & CTC based models, Docker
DEC 2018 MAY-JUL 2018	SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN (SUTD) <i>Machine Learning Research Internship, Singapore</i> Supervisor: DR. DORIEN HERREMANS github

Product

- Worked on a model for hit song prediction based on early adopters data and audio features.
- Created large open source dataset scraped from Billboard, Last.FM and Echo Nest (Spotify). OpenSMILE was used for audio feature extraction.
- Prototype deployed as web interface constituting 8 different models.
- **89% accuracy** was achieved from the best model after stratified 10-fold cross validation with **ROC AUC score of 0.95**.

DEC 2017 | NANYANG TECHNOLOGICAL UNIVERSITY (NTU)

NLP Research Internship, Singapore

Supervisor: DR. ERIK CAMBRIA

- Worked on a Named Entity Recognizer used to extract retail and consumer product details from any text and categorize them based on given set of labels.
- Created using Named Entity Recognizer by Stanford and spaCy. This NER is being used by SenticNet for their corporate products.
- Checked consistency of the established concepts of parts-of-speech synonyms filtration, thus improved 1000 synsets from NLTK WordNet

MAY-JUL 2017 | DIGISOL SYSTEMS LTD.



Software Developer Internship, Goa

- Led web development team and implemented backend in Django to automate the process of scrap management which used to be done manually.
- The portal is used by all the three branches of the company located in Bangalore, Mumbai and Goa.

ACHIEVEMENTS

- 2020 **Award:** Nominated for Elisalaisuus Award Q2 for fast supply of high quality results.
- 2019 **Hackathon:** Winner at Hack Junction (Largest Hackathon in Europe) in Spam Email Categorization track.
- 2018 **Hackathon:** Runner-up at Hack Moscow Russia (Russia biggest hackathon) in Voice Analysis track.
- 2015 **Scholarship:** Awarded INSPIRE scholarship for being in top 1 percentile at national level in board examination.


EDUCATION


- 2020 M.Sc. (Hons.) Biological Sciences & B.E. (Hons.) Computer Science, **BITS Pilani, India**
 - Research Assistant *AI Lab*: Worked on speech synthesis models based on Deep Voice architectures for various dialects of Hindi
 - Teaching Assistant *Machine Learning*: Involved in creating assignments and conducting doubt sessions
 - Projectwork *Cancer Prediction*: Formulated unsupervised methods to explore relationship between miRNA and cancer cells
 - Projectwork *Pose Detection*: Created architecture for fitness evaluation using CMU OpenPose and Microsoft Kinect


TECHNICAL SKILLS

Proficient	Python, C++, Bash
Comfortable	Java, C#
Tools	Pytorch, Kaldi, Tensorflow, Scikit-Learn, Docker, Kubernetes
Cloud	Azure, AWS, Heroku
Familiarity	Django, Flask, R, SQL, PySpark, Scala, Prolog

MISCELLANEOUS

CNN SCALA: Two layered CNN written from scratch in Scala functional programming paradigm.  [CODE](#)

CONSTRAINT PROPAGATION: AI that will solve sudoku based on search and constraint propagation techniques made as a part Udacity AI Nanodegree.  [CODE](#)

FUN SCRIPTS: Mini projects including crawling and scraping scripts. Example includes selenium based Typeracer.com player  [CODE1](#) [CODE2](#)

STUDENT GUIDE, AI LAB, BITS PILANI: Mentored freshmen and sophomores through their first AI projects.