Kumar Saunack

NLP/AI

GitHub **saunack**
Website **saunack.github.io**
LinkedIn **in/krsaunack**

WORK

MSCI - NLP Developer, Analyst

Aug 2020 - Present
- Working on new data based NLP initiatives for the firm
- Developed 1st cloud and deep learning based services
- Working with Stanford and Imperial University on new areas for the firm

Big Data Institute (BDI), Oxford University

Mar 2020 - July 2020
- Working with collaborators across different teams to develop and maintain toolkits for phylogenetics
- Co-authorship on a paper for the release of the new simulation toolkit

Center for Indian Language Technology(CFILT), IIT Bombay

Aug 2019 - May 2021
- Working on word embeddings for Indian languages
- Released 436 embedding models for public usage
- Analysed monolingual and cross lingual training methods for training in low resource setting

Prodigal Technologies Ltd. - NLP R&D Intern

Mar 2019 - Jun 2019
- Researched novel approaches for call segmentation of transcripts, using HMMs and RL
- Achieved 90% precision on topic identification

PUBLICATIONS

- Efficient ancestry and mutation simulation with msprime 1.0
  Jerome Kelleher et al., (Genetics - under review, 2021)

- How low is too low? A monolingual take on lemmatisation in Indian languages
  Kumar Saunack et al., (NAACL 2021)

- Analysing cross-lingual transfer in lemmatisation for Indian languages
  Kumar Saunack et al., (COLING 2020)

- "A Passage to India": Pre-trained Word Embeddings for Indian Languages
  Kumar Saurav, Kumar Saunack et al., (SLTU-CCURL joint workshop, LREC 2020)

SERVICE

- Developed Android app, MedConnect, for cancer-afflicted families. Now working with an NGO Love-HealsCancer on improving their technical outreach
- Completed workshop on public policy by Global Governance Initiative (GGI)
- Worked with NGO Citizen's Foundation to spread awareness of health issues in development project affected areas
- Contributing to open-source projects and publishing tutorials on GitHub and personal webpage
- Teaching Assistant for undergraduate CS207 course
- Department Academic Mentor for 2nd and 3rd year undergraduate CSE students
- Worked on setting up ranking system for lawn tennis to improve participation for campus residents
- Opened a library for children in my native village

EDUCATION

CSE, IIT Bombay 2016-2020

B.Tech. with Honors. CGPA: 8.82/10
Thesis Word Embeddings and Morphological Tools for Indian Languages

AWARDS

- **Star Award, MSCI** (Sep. 2021): Awarded to top performers in the firm
- **Institute Sports Citation, IIT Bombay** (2020): Awarded to 0.5% of terminal students for contribution to sports community at the institute level
- **NTSE and KVPY scholar**: Awarded by Govt. of India

SPORTS

- Contingent Leader, IIT Bombay sports (2019)
- Captain, Lawn Tennis team (2018-19)
- Sports Head - Lawn Tennis, IITB sports fest (2019)
- Represented state Bihar in Nationals in School Games (SGFI) (2012-13)

HOBBIES

- Books/Reading
- Lawn Tennis
- Violin