

Sumit Kumar

+91 8102176022 | sumit.atlancey@gmail.com | [linkedin](#) | [Github](#) | [Google Scholar](#)

EDUCATION

Birla Institute of Technology

Bachelor of Engineering in Computer Science

Mesra, India

Aug 2017 – May 2021

EXPERIENCE

Undergraduate Research Assistant[1]

June 2020 – Present

Carnegie Mellon University - Advisor: Prof. Kathleen M. Carley

Remotely

- Worked on a fine-grained classification and inference task for Tweets misinformation erupted during the Covid-19 pandemic.
- Crawled the web and collected data for misinformation from Twitter using Tweepy library
- Developed and fine-tuned several deep learning language models for misinformation prediction task
- Published the work as the first author in a special issue of springer named “Computational and Mathematical Organization Theory on disinformation”

Undergraduate Research Assistant

May 2020 – Present

University of California, Berkeley - Advisor: Prof. Kurt Keutzer

Remotely

- Worked on low resource language Sanskrit for word segmentation task.
- Collected and pre-processed sentence-pair data in SLP1/IAST format to create an embedding vector.

CNRS Research Intern

May 2019 – July 2019

University of Montpellier, LIRMM - Advisor: Prof. Vincent Berry

Montpellier, France

- Developed an inference program for biological data set allowing users to input their data into the program for visualization.
- Scripted file format transformation using data mining methodology and dealt with new types of databases.
- Implemented various interfacing programs with the Galaxy web platform.
- Wrote inference analysis programs to summarise the programs finding by statistics and graphics to predict the likelihood of ancestors in a phylogenetic tree.

PUBLICATIONS

- [1] **Sumit Kumar**, Raj Ratn Pranesh, Kathleen M. Carley “A Fine-Grained Analysis of Misinformation in COVID-19 Tweets” In proceedings with Computational and Mathematical Organization on Disinformation. [PAPER](#) [WEBSITE](#)
- [2] Raj Ratn Pranesh, Ambesh Shekhar, **Sumit Kumar** “Automated Medical Assistance: Attention Based Consultation System” In proceedings with NeurIPS 2020 MLPH Workshop [PAPER](#) [WEBSITE](#)
- [3] **Sumit Kumar**, R.R Pranesh, A. Shekhar “Biomedical Network Link Prediction using Neural Network Graph Embedding” In Proceedings of CoDS-COMAD 21 [PAPER](#) [ACM DL](#)
- [4] Smita Pallavi, Raj Ratn Pranesh*, **Sumit Kumar*** “A Conglomerate of Multiple OCR Table Detection and Extraction”. In proceedings of ICDAR-2020 [PAPER](#) [*Equal contribution]
- [5] Raj Ratn Pranesh, **Sumit Kumar**, Ambesh Shekhar “CLPLM: Character Level Pretrained Language Model for Extracting Support Phrases for Sentiment Labels” Accepted at ICON 2020 [PAPER](#)

EXTRACURRICULAR PROJECTS

Black Lives Matter PAPER | *Python, Tensorflow, NLTK, Regex*

July 2020 – August 2020

- Developed a framework for hate speech detection for Tweets mitigation
- Collected and annotated dataset related to black lives matter using Tweepy, a python library for accessing twitter API.
- Presented at the institute and published the dataset on Zenodo for opensource contribution and further development

Handwritten Digit Classification | *TensorFlow, Sklearn, Git, Terminal, Matplotlib*

May 2018 – June 2018

- Developed a python program for predicting images of handwritten digits on handwritten MNIST dataset
- Implemented python frameworks such as sklearn, Tensorflow, and Matplotlib for predictive learning and visualisation
Deep learning models achieved significant improvement from classical ML algorithms like random forest and SVM. The task could be extended for the alphabet and character prediction given the dataset

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL, JavaScript

Scientific Libraries: Tensorflow, Keras, Sklearn, Pandas, Numpy, NLTK

Developer Tools: Git, LATEX, Docker, Google Cloud Platform, VS Code, Google Colab, Terminal, Android Studio

TEACHING ASSISTANTSHIP

- Assisted Prof. S.C Pandey in coursework CS 5105 at Birla Institute of Technology, Mesra, Patna campus for the 7th semester CERTIFICATE (2020)
- Assisted Prof. K.Lal in coursework CS 8102 at Birla Institute of Technology, Mesra, Patna campus for the 5th semester CERTIFICATE (2019)
- Assisted Prof. Partha Paul in coursework CS 4206 at Birla Institute of Technology, Mesra, Patna campus for the 4th semester CERTIFICATE (2019)

ACHIEVEMENTS AND AWARDS

- Selected for the title "InSc Young Researcher Award" LIST (2020)
- Mentioned in the annual college magazine's achievement section [page 15] MAGAZINE (2020)
- Received NeurIPS 2020 registration grant support for presentation in MLPH workshop RECEIPT (2020)
- Entitled scholarship grant for visiting scholar at Jean Monnet University, Lyon, France DOCUMENT (2020)
- Recognized in the achievement section of the institute website for research WEBSITE (2019)
- Accepted idea for a presentation at the annual start-up summit "National Seminar on Pragmatic Role of Technological Innovation in Startup NPTIS-2020" organized by the government of India ABSTRACT (2019)

EXTRA CURRICULAR ACTIVITIES

- Attended a 3-day workshop at IIT Patna titled "Interplay of IoT and Machine Learning", explored a wide range of research and development in the interdisciplinary fields of ML and AI CERTIFICATE (2019)
- Developed and deployed Google assistant service for Google console used for promoting E-Learning WEBSITE (2019)
- Member of the Robotics club responsible for organizing workshops for students and teach them the complexity of the algorithms and data structures (2017-2019)