Abhishek Niranjan

+91 (851) 390 2048
abhishek.niranjan.iitkgp@gmail.com
abhishek-niranjan.github.io
github.com/abhishek-niranjan

Education

2013 – 2018 **Master of Technology + Bachelor of Technology(Hons.) in Computer Science & Engineering**, *Indian Institute of Technology*, Kharagpur, India.
GPA: 7.96/10.00

2011 – 2012 **AISSCE (Grade XII)**, *Kendriya Vidyalaya*, Jodhpur, India.
Percentage : 93.6%

Experience

Industry

Jun 2018 – Present **Senior Software Engineer**, *Voice Intelligence, Samsung Research*, Bangalore, India.

Grapheme-to-Phoneme (G2P) : Bixby is a voice assistant indigenous to Samsung smart devices. As a part of *Automated Speech Recognition(ASR)* team, I built Copy-Augmented Encoder-Decoder Bi-LSTM based architecture to achieve state-of-the-art results. Research paper accepted at ASRU 2019.

Grammatical Error Correction (GEC) : Approached the GEC problem as a sequence-2-sequence task with (hypothesis, reference) as the (source, target) sentence pair. Modified Transformer architecture to handle attention from multi Encoders in an hierarchical fashion for on-device textual processing module in smartphones to assist voice and keyboard enabled services.

Speech-to-Speech Translation : Presently working on translation problem involving conversion of Korean audio signal features to English audio signal features. Using Transformer model as the fundamental sequence-to-sequence architecture with addition of auxiliary decoders to train on parallel tri-phone aligned data.

May-Jul 2017 **Data Science Intern**, *Amplus Solar*, Gurgaon, India.

Soiling Rate : Photovoltaic plants experience soiling phenomena which results in decrement of power generation. Developed an algorithm to compute the soiling rate from limited data to devise a *cost-optimized* cleaning schedule of solar plants.

Power-Generation Forecasts : Engineered a forecasting module to predict hourly active power generation by a PV plant using *gradient boosted trees*. achieving a correlation of 0.97; Augmented the feature set by utilizing OpenWeatherMap API.

Power-BI Reports : Automated the generation of daily visualization reports of Solar Plants portfolio in *Power BI* by connecting MySQL server hosted on AWS EC2 instance. Built a GUI application using *PyQt*4 to facilitate the data downloading from multiple dashboards.

May-Jul 2016 **Data Science Intern**, *Bidgely Technologies*, Bangalore, India.

Vacation Detection : Bidgely's energy dis-aggregation technology helps consumers and utilities to adopt eco and pocket friendly power consumption. Devised a sliding window algorithm to predict the vacation instances of the residents in *MATLAB with precision >95% and recall >70%,* which got incorporated into production.

Refrigeration Consumption : Extended the sliding-window algorithm to compute refrigerator consumption from low resolution energy data, which was pushed into dis-aggregation pipeline.

May-Jul 2015 **Application Developer Intern**, *Outsy Inc.*, Mumbai, India.

Information Retrieval : Outsy is a lifestyle and event recommendations android application. Extracted artists' names from 15,000 Facebook textual posts using Stanford POSTagger after NLTK assisted pre-processing; Generated artists' profile database using Wikipedia API, Youtube API, and SoundCloud API.

Academic

Jul 2017 – May 2018 **Researcher**, *Complex Networks Research Group*, Indian Institute of Technology, Kharagpur, India.
Advised by Prof. Pawan Goyal and Prof. Mayank Singh
Master's Thesis: Document Clustering - Aim of the project was to cluster articles from ACL Anthology on the basis of NLP tasks and the methodology (algorithm, deep architecture, etc.) proposed to solve that task. Labelled the research articles with one of the NLP tasks (Machine Translation, NER, etc.) using rule-based pattern search in bibliographic text (title, abstract and citations received and given) to formulate a semi-supervised learning problem of tagging research papers with the NLP task. Created the feature-set comprising of learned vector embeddings of research articles' bibliographic text(doc2vec) and citation network features for each of the rule-based tagged research article. The algorithm achieved *recall of 91.52%* and *precision of 76.31%*.

Jul 2017 – Apr 2018 **Teaching Assistant**, *Indian Institute of Technology*, Kharagpur, India.
Programming and Data Structures for Prof. Rogers Mathew.
Tutored freshmen in data structures concepts and algorithmic programming exercises.

Jan 2017 – May 2017 **Undergraduate Researcher**, *Complex Networks Research Group*, Indian Institute of Technology, Kharagpur, India.

Bachelor's Thesis:Copying Citation Contexts - Objective of the project was to analyze massive dataset comprising of nearly 1.5 million computer science articles and more than 26 million citation contexts to understand the behaviour of "Copying Citation Contexts" amongst the researchers. Conducted experiments to reveal the copying patterns across 24 fields of computer science as well as the prominence of self-citation (authors citing their previous work) in the research community. Research Paper published in *Joint Conference on Digital Libraries(JCDL), 2017*

Publications

Abhishek Niranjan, M A B Shaik. (2019). *Improving grapheme-to-phoneme conversion by investigating copying mechanism in recurrent architectures*. ASRU 2019

Ankan Mullick, Anindya Bhandari, **Abhishek Niranjan**, Nitesh Sekhar, Shrey Garg, Riya Bubna, Mayank Roy. (2018). *Drift in Online Social Media.* IEMCON 2018

Mayank Singh, **Abhishek Niranjan**, Divyansh Gupta, Nikhil Angad Bakshi, Animesh Mukherjee, Pawan Goyal. (2017). *Citation sentence reuse behavior of scientists: A case study on massive bibliographic text dataset of computer science.* JCDL 2017

Projects

Jan 2017 – Mar 2017 **Competitive Strength Prediction of ATM Vendors in California : Data Analytics**.
Assessed the competitive strength of 3 major ATM vendors using spatial ATM network data and the demographics information of California.
Visualised feature importance using Tableau and clustered the ATM locations by utilising k-means approach.
Studied the features significance by applying Random Forest Classifier on the clustered data for each county.
Built a county specific linear regressor to model the revenue of each ATM to capture spatial locality information.

Aug 2016 – Nov 2016 **Sign Language Translation Through Sensory Gloves : Machine Learning**.
Worked in a team to translate American sign language gestures to text using flex sensing gloves.
Compared several classifiers on input data from sensory gloves to train the alpha-numerical character recognition algorithm.
Built a statistical language model based on stochastic grammar to recognise meaningful words from a stream of characters.

Mar 2016 – Apr 2016 **Data Extractor from 2D Plots : Software Development** .
Worked in a team of 15 members in Inter-Hall Software Development Competition to build a graph extractor that detects plots in any PDF and digitises the graphs.
Built a module which digitises the plot lines using the pixel values from the graphical(BW) image by aptly scaling them to the given range of the axes imported from the OCR module using Python Imaging Library.

Honors and Achievements

2018 **First Place**, *Data Science Competition*, Inter IIT Technical Meet.
Selected as the captain, from a pool of 400+ applicants, of the gold- winning data science team to represent IIT Kharagpur in the technical tournament attended by 19 IITs.

2017 **Runner-Up**, *Analyze This*, American Express.
Secured 2nd position in the leaderboard out of 1400+ teams participated from top-tier colleges in India.

2016 **Runner-Up**, *Data Analytics*, General Championship, IIT Kharagpur.
Captained LBS Hall of Residence's Data Analytics team to secure 2nd position in the annual event attended by a total of 20 teams.

2013-2016 **Recipient**, *MCM Scholarship*, Indian Institute of Technology Kharagpur.

2010-2012 **Recipient**, *IAF AFWWA Scholarship for High School Students*.

Skills

Languages: Python, C++, C, SQL

Frameworks & Libraries: TensorFlow, PyTorch, Keras, Scikit-learn

Softwares & Tools: MATLAB, Tableau, Power-BI, Git

Documentation: LaTeX, UML