Bishal Santra

bsantraigi.github.io, Google Scholar

Email : bsantraigi@gmail.com
Mobile : +91-8670-658-002

EDUCATION

IIT Kharagpur   Kharagpur, India
- **Ph.D. Student (PMRF Fellow)**, *Computer Science & Engineering*   *July 2018 – Present*
  Supervisor: *Prof. Pawan Goyal;* **Research Areas:** *Dialogue Systems, NLP*

IIT Kharagpur   Kharagpur, India
- *Dual Degree in Electronics & Electrical Communication Engineering; CGPA: 9.05* *July 2012 – May 2017*
  Minor in Computer Science

S.M. Arya Public School   New Delhi, India
- *Senior School Certificate Examination (CBSE Board); 90.8%*   *July 2010 - March 2012*

PATENTS

- **Population modeling system based on multiple data sources having missing entries**: *Bishal Santra, Howard Mizes, Kush Motwani,* US Patent 11,256,957 (2022). [Link]

PUBLICATIONS

- **Representation Learning for Conversational Data using Discourse Mutual Information Maximization**: *Bishal Santra, Sumegh Roychowdhury, Aishik Mandal, Vasu Gurram, Atharva Naik, Manish Gupta, Pawan Goyal,* NAACL 2022. [Link]
- **Hierarchical Transformer for Task Oriented Dialog Systems**: *Bishal Santra, Potnuru Anusha and Pawan Goyal,* NAACL 2021. [Link]
- **A Graph Based Framework for Structured Prediction Tasks in Sanskrit**: *Amrith Krishna, Bishal Santra, Ashim Gupta, Pavankumar Satuluri and Pawan Goyal,* Computational Linguistics Vol. 46, No. 4, Dec. 2020. [Link]
- **Incorporating Domain Knowledge into Medical NLI using Knowledge Graphs**: *Soumya Sharma*, Bishal Santra*, Abhik Jana, T Y S S Santosh, Niloy Ganguly and Pawan Goyal,* EMNLP-IJCNLP 2019 [Link]
- **VPDS: An AI-based Automated Vehicle Occupancy and Violation Detection System**: *Abhinav Kumar*, Aishwarya Gupta*, Bishal Santra*, Lalitha Srinivasan*, Manasa Kolla*, Mayank Gupta* and Rishabh Singh*,* IAAI 2019 (application conference of AAAI) [Link]
- **Poetry to Prose Conversion in Sanskrit as a Linearisation Task: A case for Low-Resource Languages**: *Amrith Krishna, Vishnu Dutt Sharma, Bishal Santra, Pavan Kumar Satuluri and Pawan Goyal,* ACL 2019 [Link]
- **Free as in Free Word Order: An Energy Based Model for Word Segmentation and Morphological Tagging in Sanskrit**: *Amrith Krishna, Bishal Santra, Sasi Prasanth Bandaru, Gaurav Sahu, Vishnu Dutt Sharma, Pavankumar Satuluri & Pawan Goyal,* EMNLP 2018 [Link]
- **Word Segmentation in Sanskrit Using Path Constrained Random Walks**: *Amrith Krishna, Bishal Santra, Pavan Kumar Satuluri, Sasi Prasanth Bandaru, Bhumi Faldu, Yajuvendra Singh & Pawan Goyal,* COLING 2016 [Link]

EXPERIENCE

Conduent Labs India (Previously Xerox Research Centre India)   Bangalore, India
Budding Scientist   *July 2017 - July 2018*
  - **Population Health Analytics[PHA]**:
    * Population Microsimulation: Built an RBM based model for learning joint distribution of a population (from survey data) and sampling for simulation of subgroup populations.
    * Cost Analysis Module: Used GLM based cost models built from various survey data to analyze effects of interventions on the health related cost of population.
  - **Vehicle Person Detection System[VPDS]**:
    * Trained a GoogLeNet based person counting model for front and rear window images of a vehicle captured from a toll-booth camera.

Gray Routes Technology Pvt. Ltd.   Mumbai, India
Android Developer   *Summer 2015*
  - **Product Delivery Statistics**: Worked on a new UI for an Android app, GrayDrop, to show realtime performance of delivery agents & distributions of orders from different places in the last 24 hours. Also made use of various Google Maps APIs (Place API, Heat Maps API, Directions API etc.) to achieve this goal.
  - **Search Algorithm Optimization**: Implemented an online, location based clustering algorithm for optimizing a mysql database which helped lower the bandwidth requirements of the client-side applications.

ExcellenceTech Pvt. Ltd.   Kolkata, India
Embedded System Software Developer   *Summer 2014*
  - Designed a digital power meter (for measuring Vrms, Irms, Power, Power Factor) using ATmega128, for logging power usage of electronic devices (e.g. servers). Built a web server & UI for uploading & displaying logged data.

Achievements & Scholarships

- Our proposal *"Multilingual Dialogue as a Novel Framework for AutoSuggest"* got selected as part of **Microsoft Academic Partnership Grant 2021** program (Dr. Pawan Goyal's Team)
- Our proposal *"Language Understanding via Commonsense Reasoning and Latent Concept Modeling for Effective Dialog Models"* got selected for **Google India AI/ML Research Awards**, 2020 (Dr.Pawan Goyal's Team @IIT-KGP)
- **Prime Minister's Research Fellow**, 2018
- **MCM Scholarship** at IIT Kharagpur, 2012-'16
- All India Rank of **917** in **IIT-JEE**, 2012
- Qualified **KVPY** Scholarship Exam, 2011

Projects

- **Image Super-resolution using Dictionary Learning** IIT KGP
  M.Tech Project *July 2016 - April 2017*
  - Modeled image super-resolution as an *Joint Dictionary Learning* and used Probabilistic Graphical Models(PGM) for solving the dictionary learning problem with Bayesian Inference (Gibbs Sampling).

- **Plagiarism Detection using Tree Kernel Methods** IIT KGP
  NLP Course Project *July 2015 - Dec 2016*
  - **Language Model Based Method**: In this baseline method KL divergence between unigram, bigram and trigram distributions of the two source code files were used to predict plagiarism.
  - **Tree Kernel Based Method**: In this method we trained an SVM for predicting the similarity measure between two source codes. Various approximate Tree Kernel similarities among the abstract syntax trees of the source codes were used as features.

- **Digital Image Watermarking** IIT KGP
  CV Lab Project *Jan 2016 - March 2016*
  - Digital Image Watermarking using Discrete Wavelet Transform and Signular Value Decomposition

- **Imposter Detection using Keystroke dynamics** IIT KGP
  ML Course Project *July 2015 - Dec 2015*
  - **Objective**: The aim of this project was to observe variations in key stroke dynamics (e.g. hold time, delay b/w two key press) from person to person. To understand variations in key stroke dynamics, we used Gaussian Mixture Models to model the data.

- **Mood detection from Facial Images** IIT KGP
  ML Course Project *July 2015 - Dec 2015*
  - **Objective**: Implemented a K-Nearest Neighbor classifier based on facial keypoints' positions for mood detection.

- **Arduino based Quadcopter Controller Design** IIT KGP
  B.Tech Project *July 2015 - April 2016*
  - **Objective**: Objective of the project was to build a Quadcopter with a mounted camera module as part of a *Geospatial Technology based tools Development and Disaster Management project (Sponsored Project by Department of Science and Technology, Govt. of India)*
  - **Improved Firmware**: Modified the source code of the MegapirateNG Autopilot Firmwire to integrate the OV7670 camera module to take images while in flight, by auto-switching into stabilize mode, and store the collected images into a SD Card for further Analysis

- **sugarlabs/JS-Shell** Github
  Open-Source Project *March 2015*
  - Designed and implemented an interactive shell in SugarLabs OS for beginner level Javascript programmers

Computer Skills

Languages:	Python, R, C/C++, Java, C#, Spark/Scala
Toolboxes / Frameworks:	Pytorch, Tensorflow, scikit-learn, Pandas, numpy, nltk, OpenCV, Caffe

Courseworks

- Deep Learning, NLP, Image Processing, Machine Learning, Information Theory, Estimation and Detection

Hobbies

- **Intesets**: Reading Books, Table Tennis

Languages

- **Languages**: English, Bengali, Hindi