BISHAL SANTRA

Ph.D. Student, CSE Dept.
IIT Kharagpur, India
(+91) 8670 658 002 / 8250 715 615

bsantraigi (at) gmail.com
http://studiobytestorm.com/home
https://github.com/studiobytestorm

EDUCATION

July 2018 - Current   **Ph.D. Student, Computer Science & Engineering**
IIT Kharagpur, Kharagpur
Research Area: Dialog Systems, NLP
Advisor: Dr. Pawan Goyal

July 2012 - May 2017   **Dual Degree (M.Tech and B.Tech)**
IIT Kharagpur, Kharagpur
Major: Electronics and Electrical Communication Engineering
Minor : Computer Science
M.Tech Specialization: Visual Information Processing and Embedded Systems
CGPA: 9.05/10

2010 - 2012   **Senior Secondary School Exam (CBSE Board)**
S.M. Arya Public School, New Delhi
Percentage: 90.8 %

PUBLICATIONS

- *Soumya Sharma*, Bishal Santra*, Abhik Jana, T Y S S Santosh, Niloy Ganguly and Pawan Goyal.* **Incorporating Domain Knowledge into Medical NLI using Knowledge Graphs. EMNLP-IJCNLP 2019,** November 3-7, Hong Kong (short paper, Link)

- *Abhinav Kumar*, Aishwarya Gupta*, Bishal Santra*, Lalitha Srinivasan*, Manasa Kolla*, Mayank Gupta* and Rishabh Singh*.* **VPDS: An AI-based Automated Vehicle Occupancy and Violation Detection System. IAAI 2019 (application conference of AAAI)**, (Link)

- *Amrith Krishna, Vishnu Dutt Sharma, Bishal Santra, Pavan Kumar Satuluri and Pawan Goyal (2019).* **Poetry to Prose Conversion in Sanskrit as a Linearisation Task: A case for Low-Resource Languages. ACL 2019**, July 28th - August 3rd, Florence, Italy (short paper, Link)

- *Amrith Krishna, Bishal Santra, Sasi Prasanth Bandaru, Gaurav Sahu, Vishnu Dutt Sharma, Pavankumar Satuluri & Pawan Goyal,* **Free as in Free Word Order: An Energy Based Model for Word Segmentation and Morphological Tagging in Sanskrit, EMNLP 2018** (Link)

- *Amrith Krishna, Bishal Santra, Pavan Kumar Satuluri, Sasi Prasanth Bandaru, Bhumi Faldu, Yajuvendra Singh & Pawan Goyal,* **Word Segmentation in Sanskrit Using Path Constrained Random Walks, COLING 2016**, the 26th International Coference on Computational Linguistics (Link)

WORK EXPERIENCE

JULY, '17 - JULY, '18

Population Health Analytics

Conduent Labs India (Previously Xerox Research Centre India)
Budding Scientist

- **(Patent Filing in Progress)** Built an RBM based Probabilistic Graphical Model for sample population generation and Diabetes Prevalence Prediction in U.S. counties by learning from multiple datasets - BRFSS / NHANES
- Working on a novel algorithm for data imputation in survey datasets (eg. BRFSS, NHANES) with help of generative models
- Built GLM from NHANES dataset for predicting health cost (distribution) of a person given his demographics and health conditions. Analyzed effect of educational interventions on health cost.

VPDS

- Trained a GoogLeNet based model for person detection and counting in Vehicles for tolling on high occupancy lanes

MAY. - JULY, 2015

Gray Routes Technology Pvt. Ltd.
Android Developer
- Worked on Google Maps (Place API, Heat Maps API, Directions API etc.) Android plugins to build a new UI which can show real time performance of delivery persons & distributions of orders from different places in the last 24 hours.
- Implemented a progressive location based clustering algorithm into the mysql database to lower the bandwidth requirements and thus the loading time of the web/mobile applications.

MAY. - JUNE, 2014

ExcellenceTech Pvt. Ltd.
Embedded System Software Developer
Designed a digital power meter (for measuring Vrms, Irms, Power, Power Factor) using ATmega128

PROJECTS

BAYESIAN LEARNING
STATISTICAL SIGNAL
PROCESSING

Image Super-resolution using Dictionary Learning(M.TECH PROJECT)

- Modeled image super-resolution as an *Joint Dictionary Learning* and used Probabilistic Graphical Models(PGM) for solving the dictionary learning problem with Bayesian Learning (Gibbs Sampling).

NLP

Plagiarism Detection using Tree Kernel Methods

- In this baseline method KL divergence between unigram, bigram and trigram distributions of the two source code files were used to predict plagiarism.
- 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.

MACHINE LEARNING

Imposter Detection using Keystroke dynamics
- 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.

MACHINE LEARNING

Mood detection from Facial Images
- Implemented a K-Nearest Neighbor classifier based on facial keypoints' positions for mood detection.

EMBEDDED SYSTEMS	**Arduino based Quadcopter Controller Design** (B.TECH PROJECT) - Helped the team for *Geospatial Technology based tools Development and Disaster Management project (Sponsored Project by Department of Science and Technology, Govt. of India)* to develop a Quadcopter with mounted camera module - Modified the source code of the MegapirateNG Autopilot Firmwire to integrate the OV7670 camera module - Programmed to take images while in flight by switching into stabilize mode of Autopilot control and store the collected images into a SD Card for further Analysis
COMPUTER VISION LAB	Digital Image Watermarking using Discrete Wavelet Transform and Signular Value Decomposition
SUGARLABS/JS-SHELL	Designed and Implemented an Interactive Shell in SugarLabs Software-pack for Beginner Level Javascript programmers
IP PORTAL	Designed and programmed a secure online portal for moderation of abstracts of research poject done by students/proferessors within our institute.
TWO ARCHERS (GAME)	- Created this Arcade style web-based game (on Unity Game Engine) during a online Game-JAM. - Later redesigned and launched this game for Android on Google Play Store. (LINK)
DECYCLEDJSON	A Javascript module for converting objects with cyclic structures into a JSON string format for easier portability. (LINK)

ACHEIVEMENTS AND SCHOLARSHIPS

JULY. 2018 - CURRENT	Awarded **Prime Minister Research Fellowship (PMRF)** towards pursuing PhD in Computer Science and Engineering
JULY. 2016 - APRIL 2017	**Teaching Assistantship** for students with a CGPA Higher than 8/10 in 5th Year at IIT Kharagpur
JULY. 2012 - 2016	**Means Cum Merit(MCM)** Scholarship for Undergraduate students with CGPA higher than 7.5/10 at IIT Kharagpur
SEPT. 2012	Secured **All India Rank of 917** in IIT-JEE
FEB. 2012	Qualified in both the rounds(written test and final interview) of **KVPY 2011**

LANGUAGES

BENGALI:	Mother Tongue
HINDI:	Native Tongue
ENGLISH:	Fluent

COMPUTER SKILLS

Languages:	C/C++, Python, Java, C#,
Scripts:	Python, R, MATLAB, JavaScript
IDEs:	Visual Studio, Android Studio
Toolboxes / Frameworks:	Tensorflow, scikit-learn, Pandas, numpy, nltk, gensim, Spacy, OpenCV, Caffe, Hadoop

COURSEWORK

Deep Learning	NLP	Information Theory and Coding
Pattern Recognition	Estimation and Detection	Machine Learning
Object Oriented System Design	Digital Image Processing	Algorithms

INTERESTS AND EXTRA CURRICULAR ACTIVITIES

Hobbies: Table Tennis, Reading Non-Fiction, Gaming

Extra Curricular: National Service Scheme(NSS) Volunteer, IIT Kharagpur (2012 - 2014)
Member of Open IIT Bengali Dramatics, 2014 team of Patel Hall of Residence
Represented Patel Hall of Residence at Open IIT Bengali Elocution in 2nd Year