Aman Sinha

Email : aman.sinha{@atilf.fr, @univ-lorraine.fr}

https://amansinha09.github.io/

Address : Bureau 107, 44 Avenue de la Libration, 54000 Nancy

RESEARCH INTEREST

Natural Language Processing, Graph Representation Learning.

EDUCATION

- **Université de Lorraine - Institut de Cancérologie Strasbourg** — Nancy, France
  PhD Candidate (Advisor: Marianne Clausel, Mathieu Constant, David Cox) — *October 2021 –*
- **Indian Institute of Technology(Indian School of Mines), Dhanbad** — Jharkhand, India
  Integrated Master of Technology in Mathematics and Computing; GPA: 8/10 — *2015 – 2020*

PUBLICATIONS

CONFERENCE PAPERS

1. A Sinha, P Patekar, R Mamidi. (2019). Unsupervised Approach for Monitoring Satire on Social Media. In *Proceedings of the 11th Forum for Information Retrieval Evaluation, 36-41* [Link]

WORKSHOP PAPERS

1. A Sinha, Sandrine Ollinger, Matheiu Constant. (2022). Word Sense Disambiguation of French Lexicographical Examples Using Lexical Networks. To appear in *TextGraphs-16 Workshop, COLING 2022*
2. A Sinha*, CG Holgado*, Marianne Clausel, Mathieu Constant. (2022). Catch me if you can! Capturing complex disease mentions in tweets. To appear in *Social Media Mining for Health 2022, COLING 2022*
3. CG Holgado*, A Sinha*. (2022). HalBERT at PoliticEs 2022: Are Machine Learning Algorithms better for Author Profiling?. Development, 5, 4-496. [Link]
4. AK Jena, A Sinha*, R Agarwal*. (2020). C-net: Contextual network for sarcasm detection. In *Proceedings of the Second Workshop on Figurative Language Processing*, 61-66 [Link]
5. P Gupta, H Gupta, A Sinha. (2020). DSC IIT-ISM at SemEval-2020 Task 8: Bi-Fusion Techniques for Deep Meme Emotion Analysis. In *arXiv preprint* arXiv:2008.00825 [Link]
6. A Singh, P Kumar, A Sinha. (2020). DSC IIT-ISM at SemEval-2020 Task 6: Boosting BERT with Dependencies for Definition Extraction. In *arXiv e-prints*, arXiv: 2009.08180 [Link]
7. SD Laxmi, R Agarwal*, A Sinha*. (2020). DSC-IIT ISM at WNUT-2020 Task 2: Detection of COVID-19 informative tweets using RoBERTa. In *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, 409-413 [Link]

ORAL PRESENTATIONS

1. A Sinha, R Mamidi, Evaluation of Machine Translation System for Sarcasm Transfer, Accepted at *16th Annual Workshop of The Australasian Language Technology Association 2018.*
2. Sinha Aman, Bairolia Utkarsh, Sorav Vishal, Competitive Balance in Football Leagues: Domestic vs International, Accepted at *MathSport Asia Conference 2018 at XLRI Jamshedpur, India* [Link]

RESEARCH EXPERIENCE

- **ATILF, University of Lorraine** — Nancy, France
  Research Engineer — *November 2020 – June 2021*
  - **Word Sense Disambiguation**: Working on qualitative experiments for word disambiguation on text resources by incorporating lexical networks.
- **University of Lorraine** — Nancy, France
  Research Intern (Mentor: Prof. Marianne Clausel and Prof. Mathieu Constant) — *February 2020 – October 2020*
  - **Information retrieval in linguistic networks**: Develop graph matching techniques well adapted to lexical data.
- **Indian Institute of Technology Kanpur** — Kanpur, India
  Master Thesis (Mentor: Dr. Ashutosh Modi) — *February 2020 – July 2020*
  - **Multimodal Emotion Recognition**: Worked on developing a multimodal emotion recognition system.[Link]
- **IBM India Research Labs** — Bangalore, India
  Research Intern (Mentor: Parag Jain and Priyanka Agarwal) — *May 2019 – July 2019*
  - **Multi-task learning under heterogeneous supervision for semantic parsing**: Leveraging heterogeneous supervision in multitask learning to boost weak supervised learning based semantic parsing.
- **Language Technologies Research Centre, IIIT Hyderabad** — Hyderabad, India
  Summer Research Intern (Advisor: Dr. Radhika Mamidi) — *May 2018 – July 2018*
  - **Evaluation of Machine Translation Systems for Sarcasm Transfer**: Worked on evaluation of translation systems for transfer of sarcasm from source to translated language. Built a Sarcasm Detection System for Hindi language using various machine and deep learning models.
- **Oriserve** — Mumbai, India
  Data Science Intern — *May 2017 – July 2017*
  - **Pre-processing Method for Text Correction**: Worked on character-based RNN model to pre-process user responses of DishTV chatbot conversation logs to change the corrupted text to English language.
  - **Intent Classification for User Responses**: Implemented a Multi-class Naive Bayes Classifier based on Bag-of-Words algorithm to classify the client responses based on the intent of the conversation.

PROJECTS

- **Flipkart Object Localization Challenge**: Implemented a Convolutional Neural Network model for localization of bounding box for the object in images.
- **Ad-Click Prediction Model**: Implemented a Support Vector Machine Model with linear kernel to predict Ad click-through rates with training accuracy of 97.4% and test accuracy of 88.7%.
- **Language Model for Text Generation**: Implemented Recurrent Neural Network based Word Generation Model from scratch using with Novel 'Warpeace' to generate new sentences.
- **OpenAI Gym CartPole Game Simulation**: Implemented a Deep neural network model based on reinforcement learning to solve OpenAI gym CartPole balancing environment using pygame platform, the model attained an average score of 195.

HACKATHONS

- **Multilingual Customer support Agent for Indian Languages [IBM IRL Hackathon 2019]**:
  An end to end speech to speech customer support agent using IBM Watson APIs for supporting Indian languages.
- **Extraer-language translator tool [HACKFEST '17]**:
  A live translator used to capture image and translate the text using image segmentation through Open-CV for pre-processing, and pyteserract library to read the text and translate extracted text offline.
- **Smart Home Automation [HACKFEST '16]**:
  Implemented an electronic prototype to minimize the amount of power consumption using motion sensors, humidity checkers in room using Arduino Programming.

KEY COURSES

- **MATHEMATICS**: Probability and Statistics, Linear Algebra, Statistical Inference, Numerical Methods, Discrete Mathematics, Operational Research, Soft Computing Techniques, Topology, Information and Coding Theory
- **COMPUTER SCIENCE**: Object Oriented Programming, Data Structure, Computer Organization, Design and Analysis of Algorithm, Computer Networks, Operating System, Data Base Management System, GPU Computing.
- **MOOCs**: Neural Networks and Deep Learning(deeplearning.ai), CS224D, Machine Learning (Stanford University).

ACHIEVEMENTS

- Received Student Grant for the FIRE 2019 conference.
- Secured 2nd prize in IBM IRL Hackathon 2019.
- Mentor at IIIT-Hyderabad Advanced Summer School on Natural Language Processing 2018
- Secured 10th in Code Marathon-2015 conducted by CSE Society, IIT(ISM) Dhanbad among 900 students.
- Secured 1801th rank in IIT-JEE ADVANCE-2015 from among 124741 students

POSITION OF RESPONSIBILITY

- **Organizer**: CaféTAL, NLP Reading Group, ATILF, Université de Lorraine
- **Senior Mentor**: Data Science Club, IIT ISM Dhanbad
- **Event Coordination Member**: Society for Industrial and Applied Mathematics (SIAM), IIT ISM Student Chapter
- **Teacher**: Taught underprivileged children at Kartavya, Student Run NGO