

# C.S.Bahushruth

✉ bahushruth.179303044@mu.j.manipal.edu | 🏠 bahushruth.in | 📺 PotatoSpudowski | 🐦 @bahushruth

## Education

---

### Manipal University, Jaipur

BTECH - COMPUTER AND COMMUNICATION ENGINEERING

- GPA 8/10

Jaipur

Jul 2017 - Jul 2021

## Research Experience

---

### Lesan.AI

ADVISORS: ASMELASH TEKA HAGDU

Aug, 2021 - Present

- Project: **Efficient Machine Translation for Ethiopian Low Resource Languages**
- Working on techniques to train machine translation models efficiently with far lesser parameters on low resource languages like Amharic, Tigrinya, Afaan Oromo and Somali.
- Performing analysis on existing benchmarks for low resource languages like the FLORES dataset.

### IBM Research - Global Remote Mentoring Program

ADVISORS: SEEMA NAGAR

July, 2021 - Present

- Project: **Multi-modal Misinformation Detection**
- Driving the project "Multi-modal Misinformation Detection" where we introduce a new benchmark for identifying multi-modal type of misinformation and a novel approach to address the drawbacks of existing methods to identify misinformation at scale.

### IBM Research - Global Remote Mentoring Program

ADVISORS: SEEMA NAGAR, DR. KUNTAL DEY

Jun, 2020 - Dec, 2020

- Project: **Modelling Diffusion and Propagation of Affect in Social Networks**
- Implemented Variational Graph Auto-Encoder and Multimodal Graph + Language models using Pytorch and achieved 5-7% improvement over existing techniques.
- Our work was accepted and presented at the "Deep Learning on Graphs" and "Affective Content Analysis Workshop" at AAAI'21 and accepted at the upcoming "Deep Learning on Graphs" workshop at ACM SIGKDD 2021

## Industry Experience

---

### Entrepreneur First

Bangalore, India

FOUNDER IN RESIDENCE

Feb, 2021 - May, 2021

- Built a proof of concept solution using Hugging Face and FastAPI to help the K12 students from the Tier 2, 3 communities practice answering subjective questions.
- Our solution showed improvement of 3-5%, 8-12% and upto 25% in performance of students with a baseline of 90%, 80% and 60% respectively.

### Copper Mobile

Noida, India

SOFTWARE ENGINEERING INTERN, (MACHINE LEARNING)

Aug, 2019 - Apr, 2020

- Architected smart security solutions to enable schools and communities to combat gun violence in Texas using Pytorch and Flask. Performed quantization techniques and achieved 4x reduction in model size and 2x increase in inference speed.

### Topcoder

Remote

FREELANCE DATA SCIENTIST

Sep, 2018 - Jan, 2020

- Gained multi-sector experience by developing solutions for clients from the Petroleum, Finance, Insurance, Market development fund, and the broadband industry with a screening and review success rate of 100% and 85.71% respectively.

## Open Source Contribution

---

## facebookresearch/mmf

[FEAT] ALBEF VISION TRANSFORMER ENCODER #1063 (MERGED)

- Added ALBEF Vision Transformer Encoder + Unit Tests.

## Achievements

---

### Facebook AI Hateful memes challenge 7th Position

Top 1% of the FAIR challenge with the 7th highest AUC score and the 2nd Highest Accuracy score.

### IBM Call for Code Global Semifinalist - 2019

My project titled "malnou (Platform to help diagnose kids for severe acute malnutrition)" was shortlisted for the IBM Call for Code global semifinals in 2019. Out of the 180,000 developers that participated from 165 different nations, we were selected as one of the global semifinalists.

### IISC IBM National Hackathon 1st Runner-up - 2018

I led the team that participated in the national level hackathon conducted by the Indian Institute of Science, sponsored by IBM.

### Topcoder Call for Code Hackathon Winners - 2018

The Call for Code hackathon was also hosted on Topcoder in 2018 and my team won the first place and a cash prize of 5,100 USD.

## Publications

---

Seema Nagar, Sameer Gupta, **C.S Bahushruth**, Ferdous Ahmed Barbhuiya, Kuntal Dey "Hate Speech Detection on Social Media Using Graph Convolutional Networks", Proceedings of the KDD-21 Workshop on Deep Learning on Graphs: Method and Applications, ACM SIGKDD.

[Link to: Paper \(Stay Tuned\), Workshop](#)

Seema Nagar, Sameer Gupta, **C.S Bahushruth**, Ferdous Ahmed Barbhuiya, Kuntal Dey "Homophily - A Driving Factor for Hate Speech on Twitter", Proceedings of the AAAI-21 Workshop on Deep Learning on Graphs: Method and Applications, New York, USA, AAAI.

[Link to: Paper, Workshop](#)

Seema Nagar, Sameer Gupta, **C.S Bahushruth**, Ferdous Ahmed Barbhuiya, Kuntal Dey "Empirical Assessment and Characterization of Homophily in Classes of Hate Speeches", Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI.

[Link to: Paper, Workshop](#)

## Presentations

---

### CONTRIBUTED PRESENTATIONS

Victor Callejas Fuentes, **C.S.Bahushruth** "Addressing the lack of pragmatic knowledge in multimodal models", The Hateful Memes Workshop at NeurIPS 2020.

[Link to: Presentation, Challenge Website](#)

## Skills

---

**Programming Languages:** Python, C/C++, Java, Bash, Javascript, Go

**Machine Learning:** Scikit-learn, Keras, Pytorch, Tensorflow, Fast.ai, Hugging Face, Weights & Biases

**Web Development:** MongoDB, SQL, Redis, Node, Flask, Fastapi, AWS, Azure, Kubernetes, Docker, Solr, Kafka

## Projects

---

### **FACEBOOK AI CHALLENGE - HATEFUL MEMES**

- Designed a multi-modal ensemble architecture to predict multi-modal hate using pytorch and achieved an AUROC score of 0.788 placing myteam "MemeLords" in the 7th position on the leader board out of 3000 participants (Top 1%).
- Our final submission had the 7th highest AUC score and the 2nd highest Accuracy score.
- Experimented around with different techniques to reduce bias in the model by performing adversarial training and adding custom input tags.

### **MALNOU (PLATFORM TO HELP DIAGNOSE KIDS FOR SEVERE ACUTE MALNUTRITION)**

- Project was shortlisted for the IBM Call for Code global semifinals in 2019. Out of the 180,000 developers that participated from 165 different nations, we were selected as one of the global semifinalists.
- Built a low cost affordable IoT device that calculates the health metrics and verifies the identity of the child and stores the health data on Cloudant database. Used IBM's Watson IoT platform to manage IoT devices and the data collected by them.

### **GPT3 STARTUP IDEA GENERATOR**

- Web app that uses Open AI's GPT3 to generate startup ideas based on Input Vertical/Sector and Tech Stack.
- Input prompt to the GPT3 model was extracted from a list of series A startups funded from Nasdaq in 2021.
- Built a web app using Flask and HTML to interact with the Open AI's API.

### **RIBITTER (TECHNIQUES TO COMBAT MISINFORMATION AND HATE SPEECH ON TWITTER)**

- For the Twitter Codechella hackathon, My team worked on a project to fighting against misinformation and hate speech on Twitter using scalable semantic matching and user-specific topic filtering systems based on personal triggers.
- Our project consisted of 3 modules that I worked on:
  1. Frontend app - An incredible Twitter clone web to showcase our advancements.
  2. Muting Service - Service to filter tweets based on user preferences and topic subtraction. We use zero-shot classification techniques to identify tweets based on personal triggers of the users and remove them from their feed.
  3. Misinformation Service - Service to fact verify a tweet using information retrieval and sequence classification technique.

### **CONVERTING LEGAL TEXTS TO PLAIN TEXT USING GPT2**

- Fine tuned GPT2-Medium to convert unilateral contracts (such as terms of services) to plain simple english texts. Very few people read these contracts as they are too long and complicated. Hence I worked on this project.
- Dataset used was from the "Plain English Summarization of Contracts" paper submitted to "Proceedings of the Natural Language Processing Workshop 2019"

## Leadership & Volunteer Work \_\_\_\_\_

### **PEER REVIEW**

**EMNLP 2021**, The 2021 Conference on Empirical Methods in Natural Language Processing, Upcoming conference

**NAACL-HLT 2021**, The 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, [Link to the proceedings](#)

### **VOLUNTEER**

**ACL Mentorship 2021**, The Association for Computational Linguistics year-round Mentorship Program.

## Research Interests \_\_\_\_\_

AI/Technology For Social Good, Natural Language Processing, Multi-modal Research, Graph ML, Computer Vision, Social Network Analysis