

## EDUCATION

<b>Bachelor of Technology in Mathematics and Computing</b> , Indian Institute of Technology, Delhi, <i>CGPA: 8.15/10</i>	2022
<b>All India Senior School Certificate Examination</b> , Chennai Public School, Chennai, <i>Percentage: 96.4%</i>	2018
<b>All India Secondary School Examination</b> , Chennai Public School, Chennai, <i>CGPA: 10/10</i>	2016

## EXPERIENCE

<b>KnowDis Data Science LLP, Delhi</b> Data Science Intern: <i>Product Search and Recommendation Engine</i>	Jan 2022 – Present MARS Team
<ul style="list-style-type: none"><li>Developed multiclass classification baselines using <i>BERT</i>, <i>CANINE</i> &amp; <i>GPT-2</i> and unconstrained text generation baselines using <i>BART</i> for <b>product tagging</b></li><li>Improving inference time &amp; performance of Generative Entity Retrieval (<b>GENRE</b>) model by rescoring on smaller candidates set and using LSTM decoder</li></ul>	

<b>Samsung R&amp;D Institute, Delhi</b> Software Engineering Intern: <i>Sound Source Localization, Tracking and Separation</i>	Jun 2021 – Jul 2021 Intelligence Software Team
<ul style="list-style-type: none"><li>Developed <b>sound source direction estimation</b> module, using time delay of arrival between pairs of microphones in an array, for real-time audio input</li><li>Added modules for mapping &amp; tracking active sources and extracting signals of individual sources for downstream object identification pipeline</li><li>Integrated stationary noise estimation module for ambient noise removal and reduced maximum direction of arrival error to 7°</li></ul>	

<b>MateRate Education Pvt. Ltd., Delhi</b> Machine Learning R&D and Backend Web Development Intern: <i>Students' Latent Knowledge Space Modelling and Results Portal Development</i>	May 2020 – Jul 2020
<ul style="list-style-type: none"><li>Developed Item Response Theory-based probabilistic models to estimate and analyze the ability of 5000+ students &amp; difficulty of 200+ questions</li><li>Designed database schema &amp; workflows, devised <b>system design</b> for production, built Web APIs using <b>Django REST framework</b> and tested them</li><li>Deployed <i>Django</i>-based backend to EC2 with <i>MySQL</i> database layer using <b>RDS</b>, and <i>React</i>-based frontend to <b>S3</b> with <b>CloudFront</b> CDN integration</li><li>Set up <b>Auto Scaling</b> group with scaling policies (triggered by <b>CloudWatch</b> alarms) and attached <b>Load Balancer</b> for traffic-based horizontal scaling</li></ul>	

## SKILLS

**Development:** Python, Java, C++, C, Bash, **Django**, **AWS**, SQL, PHP, MATLAB, HPC Cluster, Docker, Git, CSS, jQuery  
**NLP/ML Tools:** PyTorch, Transformers, PyTorch-Lightning, TensorFlow, Keras, NLTK, Selenium

## PROJECTS

<b>Tracking State Changes for Entities in Technical Procedural Text</b> Prof. Srikanta Bedathur & Prof. Maya Ramanath   In collaboration with IBM AI Horizons Network	Research Project, Data Group, Feb 2021 – Present [arXiv preprint]
<ul style="list-style-type: none"><li>Prepared a dataset by <b>scraping WikiHow pages</b> from <i>Computers and Electronics</i> category and developed <b>knowledge-injected BERT-based baselines</b></li><li>Deployed NLP scheme to predict the actions taken &amp; changes in properties of entities involved at each step of a process and performed error analysis</li><li>Currently, surveying the literature and building models to predict the <b>next most-likely action</b> to be performed given a sequence of performed actions</li></ul>	

<b>chaii - Hindi and Tamil Question Answering</b> <ul style="list-style-type: none"><li>Fine-tuned <i>XLM-RoBERTa</i> for <b>multilingual Q/A</b> using <i>chaii-1</i> dataset augmented with <i>MLQA</i>, <i>XQuAD</i> &amp; <i>SQuAD</i> and attained test <b>Jaccard score of 68.72%</b></li></ul>	Prof. Mausam, NLP Course, Oct 2021 – Nov 2021
--	---

<b>Rule-based Written-to-Spoken Text Converter</b> <ul style="list-style-type: none"><li>Built a <b>regex</b>-based system that accounts for chunks with abbreviations, dates, numerical quantities and inflections; obtained test <b>F1-score of 97.94%</b></li></ul>	Prof. Mausam, NLP Course, Aug 2021 – Sep 2021
---	---

<b>Corporate Bankruptcy Prediction</b> <ul style="list-style-type: none"><li>Reviewed state-of-the-art bankruptcy prediction models &amp; observed poor recall; hypothesized <b>class imbalance</b> &amp; <b>missing values</b> to be the reasons</li><li>Trained an ensemble model with <b>Mean Imputation</b> &amp; <b>SMOTE</b> transformations on <i>Polish companies</i> dataset and <b>gained 10% improvement in recall</b></li></ul>	Prof. Niladri Chatterjee, Data Mining Course, Feb 2021 – Apr 2021
--	---

<b>Extended Vector Space Model for News Articles Retrieval</b> <ul style="list-style-type: none"><li>Implemented an end-to-end retrieval system indexed using <b>TF-IDF weights</b>, with support for prefix search &amp; named-entity based filters at query time</li><li>Reduced index size by half with <b>gap encoding</b>; applied <b>pseudo-relevance feedback</b> based probabilistic <b>query expansion</b> to rerank retrieved articles</li></ul>	Prof. Srikanta Bedathur, IR Course, Oct 2020 – Nov 2020
---	---

### More projects:

- Bachelor's Thesis: Identification of Hate Spreaders on Social Media:** Extracting features from data, analyzing them and building baselines (*Ongoing*)
- Context-Sensitive Word Sense Disambiguation:** Compared non-contextual and contextual embeddings (GloVe+BiLSTM vs BERT) using WiC dataset
- Tweet Sentiment Classifier:** Processed tweets with tweet normalization, internet slang dictionary, stemming, etc.; vectorized with TF-IDF; fed into LR
- Adaptive Neuro-Fuzzy Inference System for Diabetes Prediction:** Trained a *Takagi-Sugeno* type fuzzy neural network with a test accuracy of 81.3%

## ACTIVITIES & POSITIONS HELD

<b>Teaching Assistant</b> , Information Retrieval and Web Search, Prof. Srikanta Bedathur	Aug 2021 – Dec 2021
<b>General Secretary</b> , Mathematics Society, IIT Delhi	Aug 2021 – Present
<b>Web Development Executive</b> , Entrepreneurship Development Cell, IIT Delhi	Sep 2019 – Jun 2020
<b>Volunteer in Teaching projects</b> , National Service Scheme (NSS), IIT Delhi	