

Mithilesh Vaidya

✉ mithilesh.vaidya@gatech.edu • [methi1999.github.io](https://github.com/methi1999) • [LinkedIn](#)

Education

Georgia Institute of Technology, Atlanta, USA [2022 - Present]

- Master's in Computer Science (Specialisation: Machine Learning)

Indian Institute of Technology Bombay, Mumbai, India [2017 - 2022]

- Bachelor and Master of Technology in Electrical Engineering, **CPI 9.51/10**
- **Minor** in Computer Science and Engineering

Honors and Awards

- **Institute Silver Medal** at IIT Bombay for best academic standing in department [2022]
- **Excellence in Academics Award**, for excellent academic performance (Top 3/121) [2022]
- **Undergraduate Research Award (URA03)** for outstanding thesis contributions [2022]
- Bhavesh Gandhi Memorial Prize for standing 1st in the CSP specialization
- **JN Tata Scholarship** for pursuing higher education abroad [2022]
- **AP grade in Control Systems course for exceptional performance** [2020]
- **All India Rank of 388** in JEE Advanced 2017 among 2,00,000 candidates [2017]
- **Kishore Vaigyanik Protsahan Yojana (KVPY) Fellowship** by Govt. of India [2015]
- **National Talent Search Examination (NTSE)** scholarship by Govt. of India [2015]
- **Silver medal** in Homi Bhabha Young Scientist Examination [2011]

Publications and Pre-prints

- **M. Vaidya**, K. Sabu and P. Rao, "Deep Learning for Prominence Detection In Children's Read Speech," ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 8157-8161, doi: 10.1109/ICASSP43922.2022.9747780.
- Sabu, Kamini, **Mithilesh Vaidya**, and Preeti Rao. "CNN Encoding of Acoustic Parameters for Prominence Detection." arXiv e-prints (2021): arXiv-2104.

Research Experience

- **Interpretable latents for neural population responses** Sep'22 - Present
Guide: Prof. Anqi Wu, GeorgiaTech
Research project
 - Studied techniques such as VAEs, CEBRA and LDA for learning **disentangled latent representations** of high-dimensional neural activity
 - Exploring the introduction of supervision in recurrent state space models such as rSLDS for uncovering better underlying dynamics of decision-making
- **Assessing Comprehensibility of Children's Read Speech** May'21 - June'22
Guide: Prof. Preeti Rao, IIT Bombay
Master's Thesis II
 - Implemented a Wav2vec2.0-based end-to-end model for oral reading fluency assessment
 - Outperformed RFC baseline operating on hand-crafted features by 0.06 (absolute Pearson)
 - Probed the internal representations for presence of knowledge-based features using MLP probes
 - Implemented a multi-task learning framework for exploiting hand-crafted features during training
 - Proposed a self-supervised learning framework for utilising an unlabelled dataset
- **Prominence Detection in Children's Read Speech** Jan'21 - Oct'21
Guide: Prof. Preeti Rao, IIT Bombay
Master's Thesis I
 - Replaced a Random Forest Classifier baseline with a **CRNN** framework for predicting the degree of prominence for each word in children's read speech
 - Explored inputs at various hierarchies: raw waveforms, acoustic contours and word-level aggregates
 - Demonstrated an improvement in the acoustic features extracted from word segments using Sinc convolution
 - Exploited phrase boundary labels in various **multi-task learning** paradigms
 - Used part-of-speech tags and various **NLP** embeddings such as **GloVe** and **BERT** for incorporating complementary

lexical information

- **Character Animation from Video for Blender** July'21 - Dec'21
Guide: Prof. Parag Chaudhuri, IIT Bombay Research Project
 - Working on a Blender plugin consisting of an **integrated pipeline** for extracting **3D human pose** from a video using various deep learning backends and **retargeting** it to a rigged character in Blender
 - Added VIBE and MediaPipe to the pose extraction backend of the plugin
 - Explored a **self-supervised** graph neural network framework for dynamic mapping of animations from source to dissimilar target skeletons
- **SIRD Dynamics** Aug'20 - Dec'20
Guide: Prof. Sharayu Moharir, IIT Bombay Research Project
 - Studied the **SIRD** model which is widely used for studying the outbreak of epidemics
 - Simulated the model with various underlying **network topologies** in place of random mixing
 - Formulated multiple mathematical models for calculation of **precise dynamics**

Professional Experience

- **LLM-powered Natural Language Interface** May'23 - Aug'23
Qualcomm, San Diego | Guide: Mr. Vasudev Nayak, Principal Engineer Machine Learning Internship
 - Designed an end-to-end hierarchical modular pipeline for structured parsing using LLMs such as MPT-7B
 - Developed techniques for reducing hallucinations and compared performance across various open-source LLMs
 - Evaluated performance of NVIDIA Nemo and OpenAI Whisper for STT and machine translation
- **Verification of FPGA-based High Frequency Trading Platform** Apr'20 - June'20
APT Portfolio Pvt. Ltd. | Guide: Mr. Vivek Pannikar, Senior Verification Engineer Internship
 - Implemented **DPI**, a protocol for exchanging data between SystemVerilog and other languages, for **speeding up verification** of testbenches using Cocotb, Quartus and Riviera by **3x**
 - Used Python **metaclasses** for automatically generating Python, SystemVerilog and C DPI header and implementation files from high-level JSON inputs
- **Autonise AI** Sep'18 - May'19
Machine Learning Startup Co-Founder
 - Implemented **PixelLink** and a GRU for word-level text detection, **invariant** to font size, colour, background, orientation, etc. and demonstrated an accuracy of **74%** on a proprietary dataset of documents like Aadhar Card, Passport, Driving Licence, etc.
 - Implemented a robust model with a **UNet** backbone for **segmenting** out spots, patches and wrinkles in selfies and exposed it through **AWS for demonstration**

Key Projects

- **Compiler Optimizations for speeding up Capsules** Mar'23 - May'23
Guide: Prof. Vivek Sarkar, GeorgiaTech Course Project
 - Implemented compiler optimization techniques such as unroll and jam, scalar replacement and loop reordering in CUDA for reducing execution time of the core capsule operation
 - Explained the **4x speedup** in CUDA execution time by examining the profiler output
- **Legendre Memory Unit** | Course Project: Advanced Machine Learning Jan'21 - May'21
 - Implemented and analysed the performance of Legendre Memory Units (LMU), **an improved sequential model**, on various tasks and datasets such as JSB Chorales, Mackey-Glass dynamics, etc.
 - Suggested modifications to the **core equations** by studying various **basis** functions
- **Audio Steganography** | Course Project: Automatic Speech Recognition Jan'21 - May'21
 - Exploited **adversarial attacks** on ASR systems for hiding any given sequence of tokens in any audio file
 - Analysed performance as a function of various token sequence properties such as **length and perplexity**
 - Demonstrated **high PESQ scores** which indicate low perceptibility of deviation from original audio
- **Video Toonification** | Course Project: Digital Image Processing Aug'20 - Dec'20
 - Used **Mean Shift Segmentation** across both time and spatial dimensions for toonification of videos

- Benchmarked results with standard techniques such as Bilateral Filtering
- **Auction Theory** | Course Project: Game Theory Aug'20 - Dec'20
 - Studied various models in Auction Theory such as **Vickrey Auction** and **First Price Sealed Bid Auctions**
 - Discussed equilibrium and optimal auction design analysis
- **FMX Rendering and Animation** | Course Project: Computer Graphics Nov'20 - Dec'20
 - Designed and rendered an FMX track with obstacles of varying shapes such as **cylinders and ramps**
 - Designed, rendered and animated a rider and a motorbike on the track using **keyframing**
 - Employed **Phong Shading, Texture Mapping** and used a **Skybox** for a realistic look
- **Pipelined RISC Processor** | Course Project: Microprocessors Oct'19 - Nov'19
 - Designed a 16-bit, 8-register, 6-stage **Pipelined RISC** processor in VHDL
 - Employed **Branch Prediction** and **Hazard Mitigation** techniques for optimizing the performance
- **FindIt** | Self Project: Audio Fingerprinting May'19 - June'19
 - FindIt is a Python program for identifying a song given a **short noisy segment**, similar to Shazam
 - An **audio fingerprint** consisting of constellations of major time-frequency peaks is stored in a hash table
- **Handwriting Recognition Pen** | Summer Project May'18 - June'18
 - Built a pen which can instantly convert **handwriting strokes** on ordinary paper into text
 - Designed the pen from scratch in AutoCAD and **3D printed** it
 - Generated own training data for each letter using a custom OpenCV script
- **Furniture Classification** | Kaggle Competition May'18 - June'18
 - Participated in the iMaterialist Challenge organised by Malong Technologies and CVPR 2018
 - Implemented **ResNet in PyTorch** for classification of furniture images into **128 classes**, each class containing around 1500 training images with low inter-class variation
 - Achieved an accuracy of 87.4% and ranked **30** among 428 teams across the globe

Technical Skills

- **Programming Languages:** Python, C++, C, Bash, Verilog, VHDL, OpenGL, SQLite
- **Softwares:** Matlab, Arduino, L^AT_EX, Blender, VHDL, AutoCad, Solidworks, Android Studio
- **Data Science:** PyTorch, Pandas, Numpy, OpenCV, TensorFlow, MATLAB

Key Coursework

- **Electrical Engineering:** Digital Communication, Digital Systems, Digital Signal Processing, Data Analysis and Interpretation, Control Systems, Information Theory and Coding, Markov Chains, Microprocessors, Microelectronics, Foundations of VLSI CAD, Power Systems
- **Computer Science:** Automatic Speech Recognition, Advanced Machine Learning, Computer Graphics, Foundations of Intelligent and Learning Agents, Data Structures and Algorithms, Digital Image Processing, Network Security, Computer Networks, Operating Systems
- **Miscellaneous:** Calculus, Complex Analysis, Linear Algebra, Differential Equations, Biology, Chemistry, Economics, Psychology, Engineering Drawing, Environmental Studies

Positions of Responsibility

- **Editorial Board Member, Insight** Apr'21 - Mar'22
Insight is IIT Bombay's student media body with over 10,000+ readers
 - Surveyed the **effectiveness** of the Faculty Advisor program by taking inputs from both students and faculty and **suggested various reforms**
 - Initiated a series on **startups from research labs** at IIT Bombay as part of the **LinkedIn** team
 - Interviewed authorities and current international students for understanding the causes behind **poor international representation** at IITB and suggested **remedies** for the same
 - Awarded **Special Mention for Journalism** by IIT Bombay for work carried out as part of Insight
- **Department Academic Mentor** Apr'21 - June'22

- Selected as part of a 35-member team on the basis of **ethics, peer-reviews** and an **interview**
- **Mentoring** 6 sophomores in academic and co-curricular activities

- **Teaching Assistant**

July'21 - Apr'22

- TA for EE679, a graduate-level course on Speech Processing which covers speech production, analysis techniques and applications such as ASR, speech synthesis, etc. (Autumn 2021-22)
- TA for EE352, a lab course on Digital Signal Processing which covers practical aspects of signal processing algorithms such as FFT on digital signal processors (Spring 2021-22)

Extra-Curricular Activities

- National-level quarter-finalist at Bournvita Quiz Contest; appeared on **National TV** for the same
- **Won 2nd prize** in **Android app development** competition organised by Web and Coding Club
- Successfully completed a 12-month **Lawn Tennis** course under National Sports Organisation and represented Hostel 4 in inter-hostel tournaments
- Attended the **Vijyoshi** camp conducted by **IISC**, Bangalore which served as a platform for interaction between bright young students and leading researchers in the field of science
- Awarded **Best Outgoing Student** of the year 2014-15 by Nirmala Convent High School