

Graduate student at the Dept. of Computer Science, University of Toronto enthusiastic about making a fundamental contribution in the field of machine learning

EDUCATION	<ul style="list-style-type: none">UNIVERSITY OF TORONTO 2019-Present MSc. Applied Computing GPA: 4/4R. V. COLLEGE OF ENGINEERING, BANGALORE 2014-2018 B.Eng., Computer Science and Engineering GPA: 9.72/10
COURSES	<ul style="list-style-type: none">Machine Learning and Data MiningIntelligent Adaptive InterventionsReinforcement LearningNeural Networks and Deep Learning
SKILLS	<ul style="list-style-type: none">Languages: Python, MATLAB, C, C++Machine Learning: Unsupervised Learning (mixture modelling, expectation maximization, variational autoencoders), Generative Modelling (gaussian discriminant analysis), PyTorch, TensorflowCI/CD: git, Docker, Jenkins, SonarQubeDevelopment: Unit Testing, Functional Testing, SCAVA Vulnerability scans
RESEARCH	<p>Learning Robust Latent Representation for Controllable Speech Synthesis May'20-Present</p> <ul style="list-style-type: none">Learning disentangled latent features for controlled text to speech synthesis.Proposed a Transformer based VAE through identity map reordering and gating.Discovering improved clusters of latent variables robust against collapse in case of limited data compared to LSTM based state-of-the-art for speech generative models.Forcing discovery of different features in supervised and unsupervised latent variables. <p>Adaptive Transformers in RL Jan'20-April'20</p> <ul style="list-style-type: none">Using a learnable context length for attention in partially observable and memory intensive tasks in RL.Showing performance improvements compared to current state-of-the-art Transformer XL and XL-1 architectures in these Partially Observable MDP environments.To our knowledge this was the first work to stabilize Adaptive Attention Span in RL.Project Homepage– https://github.com/jerrodparker20/adaptive-transformers-in-rl <p>Adaptive Attention based Kernels for Image Classification Jan'20-April'20</p> <ul style="list-style-type: none">Learning self-attention kernel sizes for image classification.Compare its performance to fixed-size local attention and convolution kernels.Discussion on whether adaptive attention can be helpful in correlating global features and yield any reduced FLOP count over CNN and attention based architectures.Project Homepage– https://github.com/JoeRoussy/adaptive-attention-in-cv <p>Cooperative Learning for Anomaly Correction Sept'19-Jan'20</p> <ul style="list-style-type: none">Modelling of an anomaly-correction system via inverse reinforcement learningLearning to execute corrective actions in different conditions by working alongside humans in a Cooperative Inverse Reinforcement Learning settingProject Homepage– https://github.com/shaktikshri/adaptiveSystems <p>Approximate Cross Validation in Incremental Learning Jan'17-Sept'18</p> <ul style="list-style-type: none">Invented a method for functional estimation of cross validation in incremental learning models for features extracted via caffe2 architecture; Pending patent status.Functionally estimating the hyperparameters at each incremental step instead of a conventional hyperparameter search on previously learnt inputsRemoves the need to store previous data points in online learning systems for cross validation purpose

PAPERS	<ul style="list-style-type: none"> • [In Preparation] Shakti Kumar, Jithin Pradeep, Hussain Zaidi. <i>Learning Latent Features through Disentangled Sequential Autoencoders</i>. • Shakti Kumar, Jerrod Parker, Panteha Naderian. <i>Adaptive Transformers in RL</i>. 2020. arXiv: 2004.03761 [cs.LG]. • Shakti Kumar, Jerrod Parker, Joe Roussy. <i>Adaptive Attention Span in Computer Vision</i>. 2020. arXiv: 2004.08708 [cs.CV]. • Method, System and Apparatus for Providing Efficient and Secured Authentication using Biometric Credentials, Indian Patent Application Number 201841036854, filed September 2018. Patent Pending. Available online at https://ipindiaservices.gov.in/PublicSearch/PublicationSearch/ApplicationStatus • M. K. Giluka, T. Priyadarshi, S. Kumar, A. A. Franklin and B. R. Tamma, "An enhanced EAB algorithm to reduce RACH congestion due to IoT traffic in LTE-A networks," 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 2018, pp. 395-400. https://ieeexplore.ieee.org/document/8355156
EXPERIENCE	<p>Vanguard CAI, Research Intern, Toronto May 2019-Present</p> <ul style="list-style-type: none"> • Working in unsupervised learning of latent features in speech • Architectural changes in Transformers for robust latent representation learning <p>Cisco Systems, Inc. Software Developer, Bangalore 2018-2019</p> <ul style="list-style-type: none"> • Learning of new malwares with unknown profile by observing their activity on exclusively set Honeypots • Used tf-idf vector conversion of system calls of these malwares for incremental threat model updates • Developed Long Short Term Memory (LSTM) bots using RASA and Cisco's MindMeld libraries
TEACHING	<p>CSC384 Introduction to Artificial Intelligence Winter 2020</p> <ul style="list-style-type: none"> • Designing tests and assignments on probability, bayesian networks, markov models and monte carlo tree search methods. The course was supervised by Prof. Bahar Aameri and Prof. Sonya Allin. • Course Homepage– https://www.teach.cs.toronto.edu/~csc384h/winter/
AWARDS	<ul style="list-style-type: none"> • Student Presentation Award, Applied Research in Action (ARIA), MScAC Graduating Batch 2020 • Vector Scholarship in Artificial Intelligence (VSAI), CDN 17,500 2019 • Best Final Year Project, Dept. of CSE, R. V. College of Engineering 2018 • Summer Research Fellow of the Indian Academy of Sciences. INR 20,000 2016 • Youngest to become member of Network Mobility Systems Group, IIT Hyderabad 2016