Sushil Khyalia

Google Scholar | sushil-khyalia.github.io | skhyalia@andrew.cmu.edu | +(1) 4124502585

Education

Carnegie Mellon University — Pittsburgh, PA
Master of Science in Machine Learning, GPA: 4.11/4.00 — December 2024

Indian Institute of Technology Bombay — Mumbai, India
Bachelor of Technology in Computer Science with Honors, GPA: 9.00/10.00 — June 2020

Publications

(⋆ and † indicate equal contribution, in order of decreasing contribution)

Transformers Get Stable: An End-to-End Signal Propagation Theory for Language Models
Akhil Kedia⋆, Mohd Abbas Zaidi⋆, **Sushil Khyalia**⋆, Jungho Jung, Harshith Goka, Haejun Lee
Accepted at the Forty-first International Conference on Machine Learning link

Meta-Learning for Effective Multi-task and Multilingual Modelling
Ishan Tarunesh, **Sushil Khyalia**, Vishwajeet Kumar, Ganesh Ramakrishnan, Preethi Jyothi
Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics link

Upper Bounds for All and Max-gain Policy Iteration Algorithms on Deterministic MDPs
Ritesh Goenka, Eashan Gupta⋆, **Sushil Khyalia**⋆, Pratyush Agarwal†, Mulinti Shaik Wajid†, Shivaram Kalyanakrishnan
ArXiv, abs/2211.15602 link

Data Driven Phoneme Representations for a Lexicon Free Text to Speech of Low-Resource Languages
Abhinav Garg, Jiyeon Kim, **Sushil Khyalia**, Chanwoo Kim, Dhananjaya Gowda
Accepted at IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2024 link

STING: Self-attention based Time-series Imputation Networks using GAN
Eunkyu Oh, Taehun Kim, Yunhu Ji, **Sushil Khyalia**
Proceedings of the 2021 IEEE International Conference on Data Mining (ICDM) link

SR-GCL: Session-Based Recommendation with Global Context Enhanced Augmentation in Contrastive Learning
Eunkyu Oh, Taehun Kim, Minsoo Kim, Yunhu Ji, **Sushil Khyalia**
Accepted at Deep Learning on Graphs: Methods and Applications (DLG-AAAI 22) link

PAC Mode Estimation using PPR Martingale Confidence Sequences
Shubham Anand Jain⋆, Rohan Shah⋆, Sanit Gupta†, Denil Mehta†, Inderjeet J. Nair†, Jian Vora†, **Sushil Khyalia**, Sourav Das, Vinay J. Riberio, Shivaram Kalyanakrishnan
Proceedings of The 25th International Conference on Artificial Intelligence and Statistics link

Work Experience

Samsung Research | Engineer — Seoul, South Korea
Data Analysis Lab — October 2020 - December 2021

- Achieved improved performance on session-based recommendation systems using Graph Neural Networks
- Used contrastive learning along with maximum likelihood loss to get more general graph representations

Language and Voice Team — January 2022 - March 2023

- Identified problems with current large scale transformers and proposed a new initialization and output scaling scheme for transformer models enabling us to train transformers with 100s of layers and improved performance
- Designed a new mechanism to train Grapheme to Phoneme models without any need of lexicon and instead using speech representations generated by HuBERT model

- Worked on improving performance of Open-Domain Question Answering systems by adding a semi-supervised loss which tries to generate query back from the retrieved passages

Summer Intern May 2019 - July 2019

- Worked on automating the process of hyperparameter optimisation in an AutoML pipeline
- Formulated the problem of hyperparameter optimisation problem as a case of infinite armed stochastic bandit and used policy gradient methods to find the best configuration in limited budget
- The proposed method outperformed the SMBO (Bayesian Optimisation) used in auto-sklearn

Academic Projects

Structure of the policy space of 2-action MDPs: Using Acyclic Unique Sink Orientations (AUSOs) drew insights on the policy space of 2-action MDPs and proved that Howard's Policy Iteration is an optimal deterministic algorithm for 3,4 state 2-action MDPs

Adversarial Examples for Keyword spotting: Developed a GAN which generated adversarial examples for keyword spotting systems which were then augmented with the training data resulting in 2% improvement in classification accuracy

Top-k Tournament Ranking From Pairwise Preferences: Explored schemes for fully-sequential sampling by modelling tournament ranking as a stochastic multi-armed bandit and reducing the problem to PAC subset selection in stochastic multi-armed bandits

Traffic Flow Prediction: Used Message Passing Neural Network on graph created by traffic sensor data on road network and combined it with Diffusion Convolutional Recurrent Neural Networks to incorporate both spatial and temporal dependency in traffic flow

Generating Super Resolution Images using GANs: Implemented a GAN to perform super-resolution with the generator generating images with 4x upscaling factor

Neural Image Captioning: Developed an encoder-decoder model employing the use of CNN and LSTM to caption images with Soft-Attention mechanism for visualisation and dynamic representation of salient features

Technical Skills

Languages: Python (Proficient), C++, C, Java, MATLAB, Prolog, Scheme, Bash

Machine Learning: PyTorch, Keras, TensorFlow, NumPy, auto-sklearn, OpenCV

Relevant Coursework: Intermediate Statistics, Multimodal Machine Learning, Machine Learning in Practice, Automated Speech Recognition, Information Retrieval & Mining for Hypertext & the Web, Advanced Machine Learning, Advances in Intelligent and Learning Agents, Computer Vision