

TU VU

 [tuvllms.github.io](https://github.com/tuvllms) |  tuvu@vt.edu |  [Google Scholar](#) |  [@tuvllms](#)

APPOINTMENTS

Virginia Tech

Fall 2024 — *present*

Assistant Professor, Computer Science

Research Interests: *natural language processing & machine learning*

Google

Faculty Researcher

Spring 2025 — *present*

Google DeepMind Research Scientist

2023 — 2024

EDUCATION

University of Massachusetts, Amherst

2016 — 2023

M.S/PH.D. in Computer Science

Advisor: [Mohit Iyyer](#)

2018 — 2023

Thesis committee: [Mohit Iyyer](#), [Subhransu Maji](#), [Hamed Zamani](#), [Thang Luong](#), [Colin Raffel](#)

Vietnam National University, Hanoi

2009 — 2013

B.S. Honors Program in Computer Science

Highest distinction (class rank: 1/100)

PROFESSIONAL EXPERIENCE

Google DeepMind

Fall 2022 — Spring 2023

Student Researcher *with* [Thang Luong](#) & [Quoc Le](#)

Google DeepMind

Summer 2022

Research Intern *with* [Thibault Sellam](#) & [Elizabeth Clark](#)

Google DeepMind

Winter 2021 — Spring 2022

Student Researcher *with* [Noah Constant](#)

Google DeepMind

Summer 2021 — Fall 2021

Research Intern & Student Researcher *with* [Daniel Cer](#) & [Noah Constant](#)

Google DeepMind

Winter 2020 — Spring 2021

Student Researcher *with* [Thang Luong](#) & [Quoc Le](#)

Google DeepMind

Summer 2020

Research Intern *with* [Grady Simon](#) & [Zi Yang](#) & [Nan Hua](#)

Microsoft Research

Summer 2019

Research Intern *with* [Tong Wang](#) & [Tsendsuren Munkhdalai](#) & [Adam Trischler](#)

*: equal contribution

For an up-to-date list of my research papers, please see my [Google Scholar](#) profile.

Foundational Autoraters: Taming Large Language Models for Better Automatic Evaluation

Tu Vu*, Kalpesh Krishna*, Salaheddin Alzubi, Chris Tar, Manaal Faruqui, Yun-Hsuan Sung
EMNLP 2024

// The top-performing generative model on RewardBench as of July 15, 2024, trained only on publicly available data

What Matters for Model Merging at Scale?

Prateek Yadav, **Tu Vu**, Jonathan Lai, Alexandra Chronopoulou, Manaal Faruqui, Mohit Bansal, Tsendsuren Munkhdalai

Under review @ ICLR 2025

Gemini: A Family of Highly Capable Multimodal Models

Google Gemini Team: Rohan Anil, Rohan Anil, Sebastian Borgeaud, Yonghui Wu, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew Dai, Anja Hauth, and others including **Tu Vu**

arXiv preprint 2023

// Google AI Blog

FreshLLMs: Refreshing Large Language Models with Search Engine Augmentation

Tu Vu, Mohit Iyyer, Xuezhi Wang, Noah Constant, Jerry Wei, Jason Wei, Chris Tar, Yun-Hsuan Sung, Denny Zhou, Quoc Le, and Thang Luong

ACL 2024 Findings

// Our dataset and method have inspired or been used for the development of Google's Gemini, Perplexity.AI's Online LLMs, You.com, and Contextual AI's RAG 2.0

The Flan Collection: Designing Data and Methods for Effective Instruction Tuning

Shayne Longpre, Le Hou, **Tu Vu**, Albert Webson, Hyung Won Chung, Yi Tay, Denny Zhou, Quoc Le, Barret Zoph, Jason Wei, and Adam Roberts

ICML 2023

// Google Research Blog

Mixture-of-experts meets instruction tuning: A winning combination for large language models

Sheng Shen, Le Hou, Yanqi Zhou, Nan Du, Shayne Longpre, Jason Wei, Hyung Won Chung, Barret Zoph, William Fedus, Xinyun Chen, **Tu Vu**, Yuexin Wu, Wuyang Chen, Albert Webson, Yunxuan Li, Vincent Zhao, Hongkun Yu, Kurt Keutzer, Trevor Darrell, and Denny Zhou

ICLR 2024

SPoT: Better Frozen Model Adaptation through Soft Prompt Transfer

Tu Vu, Brian Lester, Noah Constant, Rami Al-Rfou, and Daniel Cer

ACL 2022

Overcoming Catastrophic Forgetting in Zero-Shot Cross-Lingual Generation

Tu Vu, Aditya Barua, Brian Lester, Daniel Cer, Mohit Iyyer, and Noah Constant

EMNLP 2022

STraTA: Self-Training with Task Augmentation for Better Few-shot Learning

Tu Vu, Thang Luong, Quoc Le, Grady Simon, and Mohit Iyyer
EMNLP 2021

Exploring and Predicting Transferability across NLP Tasks

Tu Vu, Tong Wang, Tsendsuren Munkhdalai, Alessandro Sordoni, Adam Trischler, Andrew Mattarella-Micke, Subhransu Maji, and Mohit Iyyer
EMNLP 2020

FUNDING

Adobe Research Gift (\$5,000) PI: Tu Vu	2024
--	------

ADVISING

PHD ADVISEES:

Quyet Do, 1 st year PhD student @ Virginia Tech	Fall 2024 — <i>present</i>
Thinh Pham, 1 st year PhD student @ Virginia Tech	Fall 2024 — <i>present</i>
Rishab Balasubramanian, 1 st year PhD student @ Virginia Tech	Fall 2024 — <i>present</i>
Pin-Jie Lin, 1 st year PhD student @ Virginia Tech	Fall 2024 — <i>present</i>

OTHERS:

Prateek Yadav, Research Intern @ Google Gemini	Summer 2024 — Spring 2025
Simeng Han, Student Researcher @ Google DeepMind	Summer 2024 — Spring 2025
Salaheddin Alzubi, Masters student @ UMass Amherst	Fall 2022 — Spring 2023
Dheeraj Mekala, PhD student @ UCSD	Spring — Summer 2022

TEACHING

CS-5624: Natural Language Processing	Spring 2025
--------------------------------------	-------------

RECENT INVITED TALKS

Efficient Model Development in the Era of Large Language Models VinAI	November 2024
Efficient Model Development in the Era of Large Language Models Mila / McGill NLP Seminar	October 2024
Efficient Adaptation of Large Language Models Graph Neural Networks Reading Group, Google	November 2023
Effective and Efficient Transfer Learning in the Era of Large Language Models Faculty job talk	Spring 2023
Overcoming Catastrophic Forgetting in Zero-Shot Cross-Lingual Generation Parameter Efficient Tuning Methods Sync, Google	October 2022

Transfer Learning with Large-scale Language Models
Lecture @ **The New Turing Institute**

August 2022

The Appeal of Parameter-efficient Transfer Learning
Natural Language Accelerated Team, **Google**

June 2022

SPoT: Better Frozen Model Adaptation through Soft Prompt Transfer
Parameter Efficient Tuning Methods Sync, **Google**

December 2021

ACADEMIC SERVICE

Area Chair for NAACL 2025, COLING 2025, ACL 2024, EMNLP 2024

Session Chair (Machine Learning for NLP) @ EMNLP 2024

Program Committee/Reviewer for ICML 2025; NEURIPS 2024; TL4NLP@NEURIPS 2022; COLM 2024; AAAI 2023; ACL 2022, 2021, 2020, 2019; EMNLP 2022, 2021; NAACL 2022, 2021; COLING 2020; CoNLL 2019; INLG 2020, 2019

SELECTED MEDIA

GEMINI: [Google AI Blog](#) 2023

FRESHLLMs: [ZDNET](#) 2023

THE FLAN COLLECTION: [Google Research Blog](#) 2023

SPoT: [Headlines of Google AI's Natural Language Accelerated Newsletter](#) Q1, 2022

SELECTED AWARDS & HONORS

Google Student Researchships 2020 — 2023

UMass Amherst Graduate Assistantships 2016 — 2023

Honda Y-E-S Award for young engineers and scientists, Vietnam 2013
// in the top 10 nationally

Outstanding Academic and Co-curricular Achievements, Vietnam National University 2013

Prominent Young Figure Award, Vietnam National University 2010 & 2012

First Runner-up Prize, International Programming Contest, Japan 2011
// ranked 2nd among 64 teams internationally

Outstanding Young Talent of the Capital City, Vietnam 2010
// in the top 100 most outstanding young talents selected from a wide range of fields

Champion Prize, National Mathematical Olympiad, Vietnam 2010
// ranked 1st among more than 600 contestants nationally

A number of prizes in National/International Olympiads (in both Mathematics and Informatics) 2009 — 2013

A number of academic scholarships for undergraduate students 2009 — 2013

*: original inventor

Frozen Model Adaptation Through Soft Prompt Transfer

Tu Vu*, Daniel Cer, Noah Constant, Brian Lester, Rami Al-Rfou

U.S. Patent Application, 17/863,840

Task Augmentation and Self-training for Improved Few-shot Learning

Thang Luong, **Tu Vu***, Quoc Le, Grady Simon

U.S. Patent Application, 17/826,690