Athiya Deviyani

CONTACT
INFORMATION

EMAIL: adeviyan@cs.cmu.edu
WEBSITE: athiyadeviyani.github.io

RESEARCH
INTERESTS

Trustworthy machine learning and natural language processing; AI for social good; societal and cultural impacts of recommender systems

EDUCATION

Carnegie Mellon University, Pittsburgh, PA, United States
Ph.D. Language and Information Technologies, GPA: 4.00   Aug 2023 - May 2027 (expected)

- Advisor: Dr. Fernando Diaz

Carnegie Mellon University, Pittsburgh, PA, United States
M.S. Artificial Intelligence and Innovation, GPA: 4.02 (transcript)   Aug 2021 - May 2023

- Advisor(s): Dr. Michael Shamos, Dr. Alan W. Black, Dr. Maarten Sap
- Relevant courses: Advanced Deep Learning, Advanced NLP, Machine Learning with Large Datasets, Computational Ethics for NLP, Deep Reinforcement Learning

University of Edinburgh, Edinburgh, United Kingdom
B.Sc. Artificial Intelligence and Computer Science, GPA: 4.00 (transcript)   Sep 2017 - May 2021

- Academic award: 2021 Class Prize for Artificial Intelligence and Computer Science (pdf)
- Honors thesis: Assessing Dataset Bias in Computer Vision (received outstanding classification)
- Research advisor(s): Dr. Ajitha Rajan, Dr. Hakan Bilen, Dr. Steven R. Wilson
- Relevant courses: Machine Learning, Foundations of NLP, Automatic Speech Recognition

PUBLICATIONS

Men Have Feelings Too: Debiasing Sentiment Analyzers using Sequence GANs
Athiya Deviyani*, Haris Widjaja*, Mehak Malik*, Daniel Hoskins*
Published in the *AI for Social Good Workshop* at AAAI 2023 (oral presentation)

Text Normalization for Speech Systems for All Languages
Athiya Deviyani*, Alan W. Black
Published in the *Speech for Social Good Workshop* at INTERSPEECH 2022 (oral presentation)

Assessing Dataset Bias in Computer Vision
Athiya Deviyani*
Outstanding Informatics Undergraduate Dissertations, The University of Edinburgh 2021

ACADEMIC
RESEARCH
EXPERIENCE

Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, United States
Graduate Research Assistant   Dec 2021 - present

- Conducted research under Dr. Alan W. Black to build a tool for generalizable text normalization for speech synthesis using the Flite synthesizer and a Python backend, published in the Speech for Social Good workshop at INTERSPEECH 2022
- Audited text-to-image diffusion models (DALLE2, StableDiffusion) for gender, race, and age-related social biases under Dr. Maarten Sap, via controlled prompt generation and qualitative and quantitative analysis of generated images using a zero-shot classification method using CLIP (in preparation for FAccT 2023); this work is part of a collaboration with AI2 to document the LAION-5B dataset used to train StableDiffusion
- Investigate the short- and long-term consequences for multiple stakeholders (consumers, providers, and platform) of the continual experimentation in cultural curation, distribution, and experience powered by large-scale recommendation systems

SMASH Lab, The University of Edinburgh, Edinburgh, United Kingdom
Research Assistant   Jun 2020 - Mar 2021

- Worked with the Social Media Analysis and Support for Humanity (SMASH) research group on a project using NLP tools (GenSim, FastText, SpaCy) and a variety of ML classification techniques (SciKit Learn) to expand the LIWC psycholinguistic lexicons for improved coverage of online slang in social media data using the UrbanDictionary corpus

PROFESSIONAL EXPERIENCE	**Apple**, Seattle, WA, United States	
	Machine Learning Research Intern	May 2023 - Aug 2023

- Implemented and evaluated the benefits and drawbacks of data augmentation, parameter-efficient large language model (LLM) fine-tuning techniques such as LoRA, prefix-tuning, prompt-tuning, and SentenceBERT fine-tuning, for neural query deduplication for applications such as Siri search and web answers on the Query Resolution team at Siri and Information Intelligence

Apple, Seattle, WA, United States
Machine Learning Research Intern   May 2022 - Aug 2022

- Implemented and evaluated a GraphSAGE Graph Neural Network (GNN) with a recall at 3 rate (R@3) of 0.147 using the Deep Graph Library (DGL) framework and PyTorch to suggest related questions to be used in various applications such as Siri search and web answers
- Conducted experiments on the effect of implementing diversity penalties on beam search to diversity and perplexity of query suggestions using distributed training paradigms

Goldman Sachs, London, United Kingdom
Summer Analyst   Jun 2021 - Aug 2021

- Designed and implemented a data processing pipeline for capital reporting and analysis

Google, London, United Kingdom
Software Engineering Intern   Jun 2020 - Sep 2020

- Implemented an algorithm that selects a user's most interesting visits within a selected time frame for the Timeline feature on Google Maps

Google, Zurich, Switzerland
Software Engineering Intern   Jul 2019 - Sep 2019

- Worked on an end-to-end project within the YouTube Creator Commerce team specializing on Growth and User Engagement to create a Channel Shop Tab for creators

HONORS AND AWARDS		
Class Prize for Artificial Intelligence and Computer Science (highest overall GPA)		2021
Outstanding Informatics Undergraduate Thesis		2021
Best Health Hack at Big Red Hacks 2021, Cornell University		2021
1st place Transferwise challenge at HackCambridge4D		2019
2nd place overall and 3rd place Amazon AWS challenge at OxfordHack		2018

TEACHING EXPERIENCE	**Carnegie Mellon University**, Pittsburgh, PA, United States	
	Graduate Teaching Assistant	Aug 2021 - present

- Provided office hours, co-delivered lectures and recitations (video), designed and graded assignments for the following graduate courses: 11-631 Data Science Seminar (Prof. Eric Nyberg, Fall 2021), 11-737 Multilingual Natural Language Processing (Prof. Graham Neubig, Spring 2022), 10-617 Intermediate Deep Learning (Prof. Ruslan Salakhutdinov, Fall 2022), 11-830 Computational Ethics for NLP (Prof. Maarten Sap, Spring 2023), 10-707 Advanced Deep Learning (Prof. Ruslan Salakhutdinov, Spring 2023), 10-703 Deep Reinforcement Learning (Prof. Katerina Fragkiadaki, Fall 2023)

The University of Edinburgh, Edinburgh, United Kingdom
Graduate Teaching Assistant   Sep 2020 - Jun 2020

- Lab demonstrator for Foundations of Data Science (INFR08030), Introductory Applied Machine Learning (INFR10069, INFR11152), and Cognitive Science (INFR08020) at the University of Edinburgh, and tutor and marker for Introductory Applied Machine Learning (INFR10069, INFR11152), Software Testing (INFR10057)

VOLUNTEERING AND SERVICE		
Grace Hopper Celebration representative from Carnegie Mellon SCS		2022
Machine Learning Instructor at Bangkit Academy		2020-2021
Data Science Instructor and Curriculum Organizer at Generation Girl		2020-2021
Python and Web Development Instructor at CodeFirstGirls		2020-2021

SKILLS AND INTERESTS	
	TECHNICAL: Python (Tensorflow, PyTorch), Java, Haskell, SQL, JavaScript, C++, Isabelle, MATLAB
	LANGUAGES: Indonesian, English, Malay, Mandarin
	INTERESTS: Piano and Violin (ABRSM Grade 5), Singing, Dance, Guitar, Muay Thai, Kickboxing