RAHUL PANDEY

2900 1st Ave APT N203, Seattle, WA 98121 · rpandey4@gmu.edu · +1 (571) 421-3101
Website: https://mason.gmu.edu/~rpandey4
Linkedin: @RahulPandeyGMU · **Twitter:** @RahulPandeyGMU

INTERESTS

Natural Language Understanding, Applied Machine Learning, Speech Recognition, Human-AI Collaboration

EDUCATION

George Mason University   Fairfax, VA

Ph.D. Information Technology (Information Sciences & Technology)   Jan 2018 - Jun 2023
Advisor: *Prof. Hemant Purohit*
Thesis: *Jointly Improving Performance of Human Annotators and Models in Human-in-the-loop Machine Learning Systems*

M.S. Applied Information Technology (Cyber-Human Systems)   Jan 2018 - May 2020

The LNM Institute of Information Technology   Jaipur, India

B.Tech. Computer Science and Engineering   Aug 2013 - May 2017

WORK EXPERIENCE

Amazon.com   Seattle, WA
Applied Scientist II   Dec 2023 - Present

- Building robust widgets for product advertising on Amazon detail page using machine learning.

Applied Scientist II Intern   Jul 2023 - Dec 2023

- *Advised by Phani Nidadavolu, Milind Rao, Aparna Khare, Gautam Tewari, and Shalini Ghosh*
- Developing a robust speech recognition model using reinforcement learning from human feedback.

Applied Scientist II Intern   May 2022 - Aug 2022

- *Advised by Roger Ren, Jing Liu, Qi Luo, and Ariya Rastrow*
- Developed a framework that integrates textual and pronunciational lexicon knowledge into an RNN-T model to enhance the recognition of rare words in End-to-End (E2E) automatic speech recognition (ASR) systems.
- Achieved a significant 44% reduction in Word Error Rate (WER) for personalized entities while increasing the model size for training by only 1%.
- Published the research at ICASSP 2023 [16], and submitted a patent application to USPTO.

Applied Scientist II Intern   May 2021 - Aug 2021

- *Advised by Linda Liu, Qi Luo, Aparna Khare, and Brian King*
- Constructed an end-to-end framework for speaker-aware training of the Automatic Speech Recognition model.
- Introduced low-latency (one-hundredth of a second) frame-level training and high-latency utterance-level training, effectively minimizing the performance gap between the two approaches.
- Improved the performance of the ASR model by up to 7% in Word Error Rate (WER) compared to the baseline, without requiring speaker adaptation for utterances with multiple speakers and background speech.

Comcast Applied AI Lab   Washington, DC (Remote)
Research Contractor - Machine Learning   Jun 2020 - Aug 2020

- *Advised by Wenyan Li, Vikrant Sagar, and Ferhan Ture*

- Significantly reduced the manual effort involved in extracting erroneous voice queries from Xfinity remote for annotations by 80%. Developed an end-to-end pipeline that automates the process and further improves the performance of the Automatic Speech Recognition (ASR) algorithm.
- Utilized weak supervision to generate a dataset for training a feedback-driven multi-task learning model.
- Demonstrated the performance and live usage of the proposed model using Streamlit.io hosted in AWS.

Philips Research North America Cambridge, MA
Research Intern - AI May 2019 - Aug 2019

- *Advised by Sadid Hasan and Md Shamsuzzaman*
- Enhanced the performance of clinical concept identification for clinical trial matching by developing a model based on an intelligently fine-tuned BERT-based language model [14].

Humanitarian Informatics Lab, George Mason University Fairfax, VA
Graduate Research Assistant Jan 2018 - Jun 2023

- *Advised by Prof. Hemant Purohit*
- Addressing the challenges of Human-AI collaboration in designing online active learning-based systems by incorporating human factors into AI systems [2,7,8,10,12,13,15,17,19,21].
- Modeling the intent and behavior of online users using ML, NLP, and Semantic Computing methods to tackle complex human-centered problems in the humanitarian domain [1,3,4,5,9,11,20].

Culture Machine Media Pune, India
Software Engineer - Data Science Jun 2017 - Dec 2017

- *Advised by Meghana Negi and Arvind Hulgeri*
- Enhanced the social presence of clients by implementing supervised classification for brand entity recognition at scale using 20M+ YouTube metadata records using CNN layers as feature representation.

Summer Intern May 2016 - Aug 2016

- *Advised by Amit Garde*
- Streamlined the video generation process of trending news articles by developing an end-to-end AI-based platform utilizing text summarization and entity recognition, resulting in an 80% reduction in processing time.

INRIA - French Institute for Research in Computer Science and Automation Nice, France
Research Intern - STARS Lab Dec 2016 - May 2017

- *Advised by François Brémond*
- Designed and validated an improved deep learning architecture for people detection on depth images, outperforming AlexNet and VGG16 on a private dataset from Nice Hospital.

DRDO - Defence Research and Development Organisation Delhi, India
Winter Intern - INMAS Lab Dec 2015 - Jan 2016

- Generated cognitive visualizations using d3.js on big data to derive meaningful insights.

Kritikal Solutions Noida, India
Summer Intern May 2015 - Jul 2015

- Achieved 95.7% accuracy on pedestrian detection & optimized vanishing-point detection in the project of Advanced driver-assistance systems (ADAS) for smart cars.

AWARDS

Tier 1 - Doctoral Research Scholars	George Mason University (Aug 2022 - May 2023)
Best Paper Award	IEEE/WIC/ACM International Conference on Web Intelligence (WI 2018)
Regional Mathematics Olympiad	Homi Bhabha Centre for Science Education (Bihar State, 2011)

Teaching Experience

Graduate Teaching Assistant AIT 664: Information: Represent, Process, & Viz   George Mason University
with *Dr. Charles Lynch, Dr. Ebrima N. Ceesay, & Dr. Hemant Purohit*   Aug 2021 - May 2022
with *Dr. Charles Lynch, Dr. Vivian G. Motti, & Dr. Hemant Purohit*   Aug 2020 - May 2021

Services

Organizing Committee   WI 2019 (Workshop Co-chair)

Program Committee   ISCRAM 2023, HCOMP 2023/22, WI 2023/21/20/19
CSCW 2022, WebSci 2021, VLIOT 2021/20, AAAI FSS 2020
SBP-BRiMS 2020, ISCRAM 2020, SML (in ICSC) 2020

Journal Reviewer   IEEE Access, Elsevier FGCS

Sub-reviewer   TheWebConf 2023/22/21, ASONAM 2022/20, ICWSM 2021
NeurIPS 2019, EMNLP 2019, CoNLL 2019, TheWebConf 2019

General Secretary Science & Technology   The LNM Institute of Information Technology
Apr 2015 - Mar 2016

Publications [Google Scholar]

Book Chapter

[21] **Citizen-Helper System for Human-Centered AI Use in Disaster Management** *in International Handbook of Disaster Research, Springer*   2023
Yasas Senarath, **Rahul Pandey**, Steve Peterson, and Hemant Purohit

[20] **Intent Mining for the Good, Bad & Ugly Use of Social Web: Concepts, Methods, and Challenges** *in Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining, Springer*   2019
Hemant Purohit and **Rahul Pandey**

Journal Paper

[19] **Modeling and Mitigating Human Annotation Errors to Design Efficient Stream Processing Systems with Human-in-the-loop Machine Learning** *in International Journal of Human-Computer Studies*   2022
Rahul Pandey, Hemant Purohit, Carlos Castillo, Valerie L. Shalin

[18] **VUER: A Model for Rating Videos to Curate Content for Learning** *in Education and Information Technologies*   2022
Rahul Pandey, Hemant Purohit, Aditya Johri

[17] **Ranking and grouping social media requests for emergency services using serviceability model.** *in Social Network Analysis and Mining, Springer*   2020
Hemant Purohit, Carlos Castillo, **Rahul Pandey**

Conference Paper

[16] **PROCTER: Pronunciation-aware contextual adapter for personalized speech recognition in neural transducers** *in 48th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2023)*   2023
Rahul Pandey et al.

[15] **CitizenHelper-training: AI-infused System for Multimodal Analytics to assist Training Exercise Debriefs at Emergency Services** *in 17th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2020) (WiP Paper)*   2020
Rahul Pandey, Brenda Bannan, Hemant Purohit

[14] **BoostER: A Performance Boosting Module for Biomedical Entity Recognition** *in AIBH Workshop, IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 2019
Rahul Pandey et al.

[13] **Sensor-Based Adaptive Instructional Systems in Live Simulation Training** *in International Conference on Human-Computer Interaction (HCI International 2020)* 2020
Brenda Bannan, Elisa M. Torres, Hermant Purohit, **Rahul Pandey**, Jody L. Cockroft

[12] **Modeling Human Annotation Errors to Design Bias-Aware Systems for Social Stream Processing** *in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2019) (Short Paper)* 2019
Rahul Pandey, Carlos Castillo, Hemant Purohit

[11] **Relevancy Classification of Multimodal Social Media Streams for Emergency Services** *in IEEE International Conference on Smart Computing (SMARTCOMP 2019) (Short Paper)* 2019
Ganesh Nalluru, **Rahul Pandey**, Hemant Purohit

[10] **EMAssistant: A Learning Analytics System for Social and Web Data Filtering to Assist Trainees and Volunteers of Emergency Services** *in 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)* 2019
Rahul Pandey, Gaurav Bahl, Hemant Purohit

[9] **Distributional Semantics Approach to Detect Intent in Twitter Conversations on Sexual Assaults** *in IEEE/WIC/ACM International Conference on Web Intelligence (WI '18)* 2018
Rahul Pandey, Hemant Purohit, Bonnie Stabile, Aubrey Grant

[8] **Ranking of Social Media Alerts with Workload Bounds in Emergency Operation Centers** *in IEEE/WIC/ACM International Conference on Web Intelligence (WI '18)* 2018
Hemant Purohit, Carlos Castillo, Muhammad Imran, **Rahul Pandey**

[7] **Social-EOC: Serviceability Model to Rank Social Media Requests for Emergency Operation Centers** *in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2018)* 2018
Hemant Purohit, Carlos Castillo, Muhammad Imran, **Rahul Pandey**

Workshop/Symposium Paper

[6] **CrowdRL: A Reinforcement Learning Framework for Human Error Mitigation in Crowdsourcing-based Stream Processing** *in CSCW 2021 Workshop - Investigating and Mitigating Biases in Crowdsourced Data* 2021
Rahul Pandey, Hemant Purohit

[5] **Classifying Relevant Social Media Posts During Disasters Using Ensemble of Domain-agnostic and Domain-specific Word Embeddings** *in AAAI 2019 Fall Symposium Series* 2019
Ganesh Nalluru, **Rahul Pandey**, Hemant Purohit

[4] **AI for Trustworthiness! Credible User Identification on Social Web for Disaster Response Agencies** *in AAAI 2018 Fall Symposium Series* 2018
Rahul Pandey, Hemant Purohit, Jennifer Chan, Aditya Johri

[3] **Generic architecture of a social media-driven intervention support system for smart cities** *in Smart and Connected Communities, ICDCN 2018* 2018
Rahul Pandey and Hemant Purohit

Demo Paper

[2] **CitizenHelper-Adaptive: Expert-augmented Streaming Analytics System for Emergency Services and Humanitarian Organizations** *in IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2018)* 2018
Rahul Pandey and Hemant Purohit

Extended Abstract

[1] **Modeling Policy-relevant Intent related to Gender Violence Myths on Social Media using Social Construction Theory** *in Annual International Conference on Computational Social Science (IC^{2}S^2 2018)*   2018
Hemant Purohit, Bonnie Stabile, Aubrey Grant, **Rahul Pandey**

Skills and Coursework

Programmatical	Python(pytorch, pandas, scikit, numpy, nltk, tensorflow, flask, boto3), Java, Elasticsearch, Kibana, Tableau, Streamlit, Bash, draw.io, LaTeX, Keras, SQL, NoSQL
Analytical	Online Active Learning, Text Mining, Applied Machine Learning, Deep Learning, NLU, NLP, Intent Mining, Learning to Rank
Relevant Coursework	Approximate Dynamic Programming (OR 774), Deep Learning for Predictive Analytics (OR 750), Applied Machine Learning (AIT 736), Data Analytics in Social Media (AIT 724), Natural Language Processing (AIT 690 & 726), Research in Information Science (AIT 602), Algorithms and Data Structures (AIT 512)

Extracurricular Activities

Runner up, International Student Photo Competition   George Mason University
                                                                                              Apr 2022

1st Prize, International Student Photo Competition   George Mason University
                                                                                              Apr 2018

1st Prize, Chess - Inter College Sports Fest   RIET, Jaipur, India
                                                                                              Feb 2014