

Suhas Bettapalli Nagaraj

Email: hi@bnsuhas.com

Website: bnsuhas.com

Google Scholar

linkedin.com/in/bnsuhas

Phone: 814 862 8156

EDUCATION

	Penn State University	University Park, PA
2021 – 2025	<i>Ph.D. in Informatics, Machine Learning for Health</i> Dissertation: <i>Generative, Human-Centered AI in Mental Health: Supporting Therapeutic Fidelity and Empathy in Interactions</i>. Advisor: Saeed Abdullah. GPA 4.0/4.0.	
2019 – 2021	<i>M.S. in Electrical Engineering, Communications and Signal Processing</i> - Thesis: <i>Privacy-Preserving Assessment of Depression Using Speech Signal Processing</i>. - Graduate minor: Engineering Leadership and Innovation Management.	
	PES University	Bangalore, India
2014 – 2018	<i>B.Tech. in Electronics and Communication Engineering</i>	

EXPERIENCE

Jun 2025 – Present	Amazon, Speed AI Org <i>Applied Scientist</i>	Bellevue, WA
	- Engineered an autonomous LLM-driven agent (CradleWriterMCP) capable of complex multi-step reasoning and environment interaction, enabling zero-shot tool use for large-scale data workflows (catalog search, query synthesis, and automated S3 execution). Now the primary workflow for multiple Speed scientists, who report 5-10x speed up on job runs. - Built and deployed a 0-to-1 ML personalization system reaching 84M customers daily, owning the multi-stage pipeline end-to-end: event-aware candidate generation, filtering (0.93 precision / 0.95 recall), and personalized ranking. Optimized and evaluated model performance at scale for seamless production deployment. - Launched and A/B tested personalized shopping widgets for the FIFA World Cup 2026 with product and business partners. - Filed a U.S. patent application on grocery freshness personalization (August 2026), with two further invention disclosures under IP review.	
Aug – Nov 2024	Amazon Web Services, Health AI <i>Applied Scientist-II Intern</i>	Seattle, WA
	Designed robust, reproducible evaluation frameworks for foundational LLMs, combining automated metrics with structured human-in-the-loop assessments to diagnose hallucinations and measure post-training alignment; first-author paper at INTERSPEECH.	

Apr 2023 – Jun 2025	Pennsylvania State University, NSF TEAMMAIT <i>Graduate Research Assistant</i>	University Park, PA
	- Built LLM-based synthetic-data pipelines producing 3,000+ clinician-evaluated therapy conversations across diverse demographics, released as a NeurIPS 2025 Spotlight paper. - Principal Investigator on two IRB-approved studies on AI-assisted annotation and LLM chatbots for PTSD support (Penn State STUDY00025305 and STUDY00026578); designed human-evaluation protocols with seven clinicians for therapeutic fidelity, empathy, and safety, leading to two first-author EMNLP 2025 papers. - Fine-tuned large audio-language models with LoRA for temporal localization of therapy elements in long-form session audio; first-author ICASSP 2026 paper.	
May – Aug 2023	Samsung Research America <i>AI Research Intern</i>	Mountain View, CA
	Developed multimodal models over ECG and PPG signals for personalized cuff-less blood pressure estimation, reaching a state-of-the-art MAE of 0.29/0.19 mmHg; U.S. patent application and first-author ICASSP 2024 paper.	
Jun – Aug 2020	3M Health Information Systems (M*Modal) <i>Speech Recognition R&D Intern</i>	Pittsburgh, PA
	Improved speaker separation for physician-patient conversations in reverberant, far-field settings by extending ConvTasNet with inter-microphone phase-difference features (SI-SDR 11.96 to 12.69 dB).	
May 2018 – Aug 2019	Indian Institute of Science, SPIRE Lab <i>Research Assistant</i>	Bangalore, India
	- Led multimodal data collection (speech, video, electromagnetic articulography) from 300+ people with ALS and Parkinson's disease, working with neurologists at NIMHANS. - Developed speech-based classification of ALS and Parkinson's disease and their severity, and led deployment of a smartphone diagnostic-support app used by NIMHANS neurologists; papers at INTERSPEECH 2019, ICASSP 2020, and SPCOM 2020.	

PUBLICATIONS

	UNDER REVIEW
ICASSP 2027 submitted	Suhas BN and N. Kudupu. “Unsupervised cross-device speaker diarization and linking through spatiotemporal structure.” arXiv:2610.15978 arXiv
npj Digital Medicine	Suhas BN , et al. “AI Safety Training Can Be Clinically Harmful.” arXiv:2604.23445. arXiv
	CONFERENCE PAPERS
ICASSP 2026	Suhas BN , J. Alaparthi, S. Abdullah, D. Mattioli, R. Arriaga, C. Wiese, and A. Sherrill. “Fine-Tuning Large Audio-Language Models with LoRA for Precise Temporal Localization of Prolonged Exposure Therapy Elements.” <i>IEEE International Conference on Acoustics, Speech and Signal Processing</i> , Barcelona. Paper Link
NeurIPS 2025 Spotlight	Suhas BN , A. Sherrill, R. Arriaga, C. Wiese, and S. Abdullah. “Thousand Voices of Trauma: A Large-Scale Synthetic Dataset for Modeling Prolonged Exposure Therapy Conversations.” <i>Advances in Neural Information Processing Systems</i> . OpenReview

- EMNLP 2025 **Suhaz BN**, Y. Mahajan, S. Abdullah, D. Mattioli, R. Arriaga, C. Wiese, and A. Sherrill. “The Pursuit of Empathy: Evaluating Small Language Models for PTSD Dialogue Support.” *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 30888–30910, Suzhou. [Paper Link](#)
- EMNLP 2025 Findings **Suhaz BN**, S. Abdullah, D. Mattioli, R. Arriaga, C. Wiese, and A. Sherrill. “How Real Are Synthetic Therapy Conversations? Evaluating Fidelity in Prolonged Exposure Dialogues.” *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 20986–20995, Suzhou. [Paper Link](#)
- INTERSPEECH 2025 **Suhaz BN**, J. P. Cohen, H. C. Shing, J. C. Moriarty, L. Xu, M. Strong, J. Burnsky, J. Ofor, J. R. Mason, S. Chen, C. Shivade, and S. Srinivasan. “Fact-Controlled Diagnosis of Hallucinations in Medical Text Summarization.” *Proc. INTERSPEECH*, pp. 3070–3074, Rotterdam. [Paper Link](#)
- INTERSPEECH 2024 **Suhaz BN**, A. Rebar, and S. Abdullah. “Speaking of Health: Leveraging Large Language Models to Assess Exercise Motivation and Behavior of Rehabilitation Patients.” *Proc. INTERSPEECH*, pp. 3155–3159, Kos. [Paper Link](#)
- ICASSP 2024 **Suhaz BN**, R. S. Srinivasa, Y. M. Saidutta, J. Cho, C. H. Lee, C. Yang, Y. Shen, and H. Jin. “End-to-End Personalized Cuff-less Blood Pressure Monitoring Using ECG and PPG Signals.” *IEEE ICASSP*, pp. 2101–2105, Seoul. [Paper Link](#)
- INTERSPEECH 2023 **Suhaz BN**, S. Rajtmajer, and S. Abdullah. “Differential Privacy Enabled Dementia Classification: An Exploration of the Privacy-Accuracy Trade-off in Speech Signal Data.” *Proc. INTERSPEECH*, pp. 346–350, Dublin. [Paper Link](#)
- ICASSP 2022 **Suhaz BN** and S. Abdullah. “Privacy Sensitive Speech Analysis Using Federated Learning to Assess Depression.” *IEEE ICASSP*, pp. 6272–6276, Singapore. [Paper Link](#)
- CogMI 2020 **Suhaz BN**. “Automatic Bird Sound Detection in Long Range Field Recordings Using Wavelets and Mel Filter Bank Features.” *IEEE International Conference on Cognitive Machine Intelligence*, pp. 218–226. [Paper Link](#) & [Code](#)
- SPCOM 2020 **Suhaz BN**, J. Mallela, A. Illa, B. K. Yamini, N. Atchayaram, R. Yadav, D. Gope, and P. K. Ghosh. “Speech Task Based Automatic Classification of ALS and Parkinson’s Disease and Their Severity Using Log Mel Spectrograms.” *International Conference on Signal Processing and Communications*, pp. 1–5. [Paper Link](#) & [Code](#)
- ICASSP 2020 J. Mallela, A. Illa, **Suhaz BN**, S. Udupa, Y. Belur, N. Atchayaram, R. Yadav, P. Reddy, D. Gope, and P. K. Ghosh. “Voice-Based Classification of Patients with Amyotrophic Lateral Sclerosis, Parkinson’s Disease and Healthy Controls with CNN-LSTM Using Transfer Learning.” *IEEE ICASSP*, pp. 6784–6788, Barcelona. [Paper Link](#)
- INTERSPEECH 2019 **Suhaz BN**, D. Patel, N. R. Koluguri, Y. Belur, P. Reddy, A. Nalini, R. Yadav, D. Gope, and P. K. Ghosh. “Comparison of Speech Tasks and Recording Devices for Voice Based Automatic Classification of Healthy Subjects and Patients with Amyotrophic Lateral Sclerosis.” *Proc. INTERSPEECH*, pp. 4564–4568, Graz. [Paper Link](#)
- IC4 2018 **Suhaz BN**, S. Bhagavat, V. Vimalanand, and P. Suresh. “Wireless Sensor Networks Based Monitoring of Railway Tracks.” *International CET Conference on Control, Communication, and Computing*, pp. 187–192. [Paper Link](#)

- Springer 2023 H. J. Han, **Suhas BN**, L. Qiu, and S. Abdullah. “Automatic Classification of Dementia Using Text and Speech Data.” In *Multimodal AI in Healthcare: A Paradigm Shift in Health Intelligence*, pp. 399–407. Springer, Cham. [Chapter Link](#)
- Collaborators at Emory University, Georgia Tech, Penn State, IISc Bangalore, and NIMHANS. Full list on [Google Scholar](#).

PATENTS AND INVENTIONS

- Filed Aug 2026 **Suhas BN**, et al. “Preference-Based Classification and Selection of Objects Using Sensor-Derived Data.” U.S. patent application, Amazon. Pending.
- Published 2025 **Suhas BN**, Y. M. Saidutta, R. S. Srinivasa, J. Cho, C. H. Lee, C. Yang, Y. Shen, and H. Jin. “Method for End-to-End Cuff-less Blood Pressure Monitoring Using ECG and PPG Signals.” U.S. patent application 18/883,902, Samsung Electronics; published as US 2025/0090033 A1. [Google Patents](#)
- Under IP review **Suhas BN**, et al. “Intent-Based Pre-Departure Flash Fulfillment with Speculative Pre-staging.” Invention disclosure, Amazon.
- Under IP review **Suhas BN**, et al. “System and Method for Geographic Event-Based Shopping Interface with Location-Embedded Recommendations.” Invention disclosure, Amazon.

HONORS AND AWARDS

- 2026 Amazon Fellow. One of 50 selected company-wide for a six-month engagement applying ML to a nonprofit’s use case, with the Southeast Seattle Education Coalition (SESEC).
- 2025 Spotlight, NeurIPS 2025, for *Thousand Voices of Trauma*.
- 2025 Best Poster Award, Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL), first of 60+ posters.
- 2025 Third Place, Penn State AI Week poster session, of 80+ posters. [Link](#)
- 2024 Graduate Student Travel Grant, College of IST, Penn State.
- 2023 Finalist, Nittany AI Challenge, for SeniorSync, a system for monitoring older adults. [Link](#)
- 2022 – 2023 NSF NRT LinDiv Fellowship, Penn State. [Link](#)
- 2022 Google CS Research Mentorship Program Scholar (mentor: Jo Schaeffer).
- 2022 Google Developers Machine Learning Bootcamp Scholar.
- 2022 Winner, AAAI 2022 Hackallenge, multimodal analysis of dementia.

SERVICE

- 2023 – Present Paper Reviewer: NeurIPS (2026 – Present), JMIR (2026 – Present), BMJ Digital Health & AI (2026 – Present), ACM Health (2026 – Present), ACM CHI (2024 – Present), ACM CUI (2023 – Present), IEEE Signal Processing Letters (2023 – Present).
- 2024 Program Committee, IEEE International Conference on Healthcare Informatics (ICHI).
- 2024 Judge, Three Minute Thesis Competition, Penn State Graduate School.
- 2024 Judge, Undergraduate Exhibition, Penn State.
- 2024 Hiring Committee, Millennium Scholars Program, Penn State.

IN THE MEDIA

- 2025 Penn State News. Featured as an AI Week 2025 winner for research on generative AI for empathetic mental health conversations and audio-based therapy evaluation. [Link](#)
- 2024 ISCA-SAC Interspeech Podcast. Interviewed on the official conference podcast about my research and impressions from Interspeech 2024. [Link](#)

TEACHING

- Graduate Teaching Assistant, Penn State. 500+ students across six semesters.
- Fall 2023 IST 504: Theories and Methods of Information Science Research (Dr. Luke Zhang).
- Fall 2023 SRA 450: Cyber Crime and Cyber Warfare (John G. Hodgson).
- Spring 2022 Data Visualization (Dr. Yasser Elmanzalawi).
- Fall 2021 Data Integration and Fusion (Dr. Yasser Elmanzalawi).
- Spring 2021 Data Analytics for Healthcare (Dr. Fenglong Ma).
- Spring and Fall 2020 Data Analytics for Machine Learning (Dr. Fenglong Ma).

LEADERSHIP

- 2022 – 2023 President, Indian Graduate Student Association, Penn State. Led an organization of 600+ graduate students and managed a \$40,000 annual budget.
- 2020 – 2021 Editor, CERS engineering research journal, Engineering Graduate Student Council, Penn State. Supervised 200+ student contributors.

SKILLS

- Methods LLM & Agentic AI: autonomous agent design (tool-use, planning), alignment (DPO, RLHF), Parameter-Efficient Fine-Tuning (LoRA), recommender systems and ranking, synthetic data generation, differential privacy and federated learning, speech and physiological signal processing.
- Evaluation LLM and human evaluation design, clinician studies and IRB protocols, factuality and hallucination diagnosis, A/B testing, controlled ablations.
- Tooling PyTorch, Hugging Face Transformers, vLLM, Whisper, wav2vec 2.0, ModernBERT, scikit-learn, PySpark, MLflow, MCP servers for agent workflows.
- Infrastructure Python, C++, SQL, CUDA, AWS, Google Cloud, Git.