Neeti Pokhriyal, PhD
AAAS Science and Technology Policy Fellow,
Directorate for Computer and Information Science and Engineering,
National Science Foundation
email: neeti.pokhriyal@gmail.com
web: https://neetip.github.io/

Professional Experience

National Science Foundation September 2022 - present
AAAS Science and Technology Policy Fellow
National Artificial Intelligence Research Institutes Program
Directorate for Computer and Information Science and Engineering

The National Academies of Sciences, Engineering, and Medicine, DC March 2022 - August 2022
Mirzayan Science and Technology Policy Fellow
Committee on National Statistics
Project: Toward a Vision for a New Data Infrastructure for Federal Statistics and Social and Economic Research in the 21st Century

Dartmouth College, NH
Visiting Scholar, Department of Computer Science (CS) Oct 2021 - present
Postdoc, Department of Computer Science Oct 2019 - Sept 2021
jointly with Prof. Soroush Vosoughi (in CS) and Prof. Benjamin Valentino (in Government)
Associate Affiliate, **Irving Institute for Energy and Society** July 2020 - Sept 2021

Inter-American Development Bank, DC Jan 2019 - Dec 2019
Consultant

University at Buffalo, The State University of New York, NY Jan 2014 - Sept 2019
Graduate Research Assistant, Department of Computer Science & Engineering

Oak Ridge National Laboratory, Oak Ridge, TN July 2012 - April 2013
Researcher, Computer Science and Mathematics Division

University of California, Riverside, CA Sept. 2008 - Dec. 2009
Graduate Research and Teaching Assistant, Department of Computer Science

Tata Consultancy Services, Mumbai, India Sept. 2005 - Oct.2007
Assistant Systems Engineer, Nortel Technology Laboratory.

Education

PhD, Computer Science & Engineering Aug 2013 - Sept 2019
University at Buffalo, The State University of New York
Dissertation Title: Learning from Disparate Data: Applications in Biometrics & Sustainability
The thesis proposed novel methods to jointly learn from noisy, uncertain, and high-dimensional data coming from multiple sources, sensors, or modalities for tackling problems in biometric identification and sustainable human development.

Masters, Computer Science April 2008 - Dec. 2009
University of California, Riverside
Thesis Title: Nucleosome Landscape Analysis for Novel Gene Discovery Via Machine Learning
The thesis proposed a computational framework to discover novel genes in the human malaria parasite genome using nucleosome positioning data.

Bachelors in Computer Engineering, *with Honors* July 2001 – May 2005
Aligarh Muslim University, India

Funding secured

1. **Mapping Country-wide Energy Access for the Majority World**
   Awarded by: Irving Institute of Energy and Society, Dartmouth College
   As **Principal Investigator** for **USD 31,000 (100% share)** from July 2020 - Sept 2021.
   Goal: Designing a novel computational model to now-cast energy access for developing countries for policy planning.

2. **Financial Services for the Poor** (OPP1114791)
   Funded by: Bill and Melinda Gates Foundation
   As **Project Lead**, University at Buffalo, for **USD 20,000 (100% share)** from June 2015 - Dec 2016.
   Goal: Building algorithms to map poverty at policy-planning regions using mobile phone and satellite data.

3. **Multi-dimensional poverty mapping from mobile phone data on the OPAL platform**
   Funded by: Overseas Development Institute (ODI), UK
   As **Senior Personnel** for **USD 15,000 (100% share)** from Feb 2019 - August 2019.
   Goal: Building novel algorithms that use mobile phone data in a privacy-preserving manner to map poverty.

Recent Awards

1. **Mirzayan Science and Technology Fellowship**, National Academies of Sciences, March 2022.
   A highly competitive fellowship (<10% acceptance) that exposes fellows to science policy at the federal level.
2. **Chih Foundation Research and Publication Award, University at Buffalo, NY**, May 2019.
   This is a single award of USD 2,500/year given each year for doctoral research related to innovation for the betterment of society at the University at Buffalo, State University of New York.
3. **Doctoral Consortium Scholarship** for AAAI Conference on Artificial Intelligence, Jan 2019.
4. **Winner - National Statistics Prize & USD 2,000 prize**, Data for Development (D4D) Challenge
   International Conference on the Analysis of Mobile Phone Datasets, MIT, 2015.
5. **Finalist**, 3 Minute Thesis (3MT), University at Buffalo, 2019.
6. **Travel Support** to attend International Conference on Computational Sustainability, Cornell 2016.
7. **Dean's Distinguished Fellowship Award** at University of California, Riverside, 2008.

Policy experience

1. **National Science Foundation**:
   (a) I am part of the leadership team for the ExpandAI program, whose aim is to expand AI research and workforce development for underserved communities through capacity building and partnerships with the National AI Research Institutes program. I coordinate efforts to manage the proposals, conduct reviews and coordinate efforts across the different directorates at NSF related to the program.
   (b) I contribute to the NSF's National AI Research Institutes program, which funds long-term, high-reward AI research by engaging in inter-agency partnerships and in developing future solicitations.
   (c) I have assisted the leadership at the Information and Intelligent Systems division at Computer, and Information Science and Engineering directorate, NSF in developing a policy paper in coordination with other agencies and working groups like the Networking and Information Technology Research and Development.
   (d) As an AAAS science and technology policy (STPF) fellow, I participate in professional development activities. Some of the recent ones that I have completed are program management, program evaluation, and policy negotiation across the US government.
2. **National Academies of Sciences, Engineering, and Medicine**:
   (a) I contributed to the **Committee on National Statistics** on the project titled, "Toward a Vision for a New Data Infrastructure for Federal Statistics and Social and Economic Research in the 21st Century". I researched and contributed to aspects of **sustainable models of data sharing, data equity, and privacy-preserving methodologies**. I contributed to another NAS report (*forthcoming*) by working on a chapter on the use of multiple datasets for agricultural statistics.
   (b) I learned skills like consensus-building across diverse stakeholders and how to effectively communicate science to policymakers.

Peer-Reviewed Academic Journals

9. Accurate Intercensal Estimates of Energy Access to Track Sustainable Development Goal 7, **N. Pokhriyal**, Emmanuel Letouzé, Soroush Vosoughi, EPJ Data Science, 2022. (Impact factor: 5.4)
8. An interpretable model for real-time tracking of economic indicators, **N. Pokhriyal**, B. Valentino, S. Vosoughi, Association for Computing Machinery (ACM) Transactions on Data Science, 2021.
7. Combining disparate data sources for improved poverty prediction and mapping, **N. Pokhriyal**, D. Jacques, **Proceedings of the National Academy of Sciences (PNAS)**, 2017. (Impact factor: 12)
6. Estimating and Forecasting Income Poverty and Inequality in Haiti Using Satellite Imagery and Mobile Phone Data, **N. Pokhriyal**, O. Zambrano, J. Linares, H. Hernández *Working Paper*, Inter-American Development Bank, 2020.
5. Learning from disparate data: Applications in Biometrics and Sustainability, **N. Pokhriyal**, PhD thesis, University at Buffalo, State University of New York, 2019.
4. Learning Discriminative Factorized Subspaces with application to Touchscreen Biometrics, **N. Pokhriyal**, V. Govindaraju, IEEE Access, 2020. (Impact factor: 4.6)
3. Cognitive-Biometric Recognition from Language Usage: A Feasibility Study, **N. Pokhriyal**, I. Nwogu, V. Govindaraju, IEEE Transactions in Information Forensics, 2016. (Impact factor: 6.2)
2. Analysis of nucleosome positioning landscapes enables gene discovery in the human malaria parasite Plasmodium falciparum, X. M. Lu, E. M. Bunnik, **N. Pokhriyal**, S. Nasseri, S. Lonardi, K. Le Roch, BMC Genomics, 2015. (Impact factor: 3.5)
1. Nucleosome Landscape Analysis for Novel Gene Discovery Via Machine Learning, **N. Pokhriyal**, Masters thesis, University at California, Riverside, 2009.

Peer-reviewed Academic Conferences and Workshop Proceedings

11. AI-assisted diplomatic crisis management, **N. Pokhriyal**, Till Koebe, *forthcoming* in AI and Diplomacy workshop at AAAI Conference on Artificial Intelligence, Washington DC, 2023.
10. Social media data reveals signal for public consumer perceptions, **N. Pokhriyal**, A. Dara, B. Valentino, and S. Vosoughi, ACM International Conference on AI in Finance 2020.
9. Assessing countrywide socio-economic deprivations using auxiliary data sets, **N. Pokhriyal** and S. Vosoughi, AI for Africa for Sustainable Economic Development Workshop, ACM International Conference on AI in Finance 2020.
8. Multi-view learning from disparate sources for Poverty Mapping, **N. Pokhriyal**, AAAI Conference on Artificial Intelligence, 2019.
7. A Computational Approach to Poverty Mapping, **N. Pokhriyal**, V. Govindaraju, International Conference on Computational Sustainability, Cornell, 2016.
6. Virtual Network and Poverty Analysis in Senegal, **N. Pokhriyal**, W. Dong, V. Govindaraju, International Conference on the Analysis of Mobile Phone Datasets, MIT, 2015
5. A Large-scale Study of Language Usage as a Cognitive Biometric Trait *Invited*, **N. Pokhriyal**, I. Nwogu, V. Govindaraju, Elsevier Handbook on Big Data Analytics, 2015 .
4. Use of Language as a Cognitive Biometric Trait, **N. Pokhriyal**, I. Nwogu, V. Govindaraju, IEEE International Joint Conference on Biometrics, 2014.
3. Novel Gene Discovery in the Human Malaria Parasite using Nucleosome Positioning Data, **N. Pokhriyal**, N. Ponts, E. Harris, K. Le Roch & S. Lonardi, Intl Conf. on Computational Systems Bioinformatics, 2010.
2. Anomaly Detection for High Fidelity Core Simulators, **N. Pokhriyal**, U. Mertyurek, A. Godfrey, J. J. Billings, In Proc. of the American Nuclear Society Annual Meeting, 2013.
1. Knowledge Discovery from Nuclear Reactor Simulation Data, **Neeti Pokhriyal**, Ugur Mertyurek, Andrew Godfrey, Jay Jay Billings, Workshop on Analytics for Cyber-Physical Systems, SIAM International Data Mining Conference, 2013.

Journal articles under review

2. Quantifying participation biases on social media, **N. Pokhriyal**, B. Valentino, S. Vosoughi, 2022.
1. Fairness in learning human development indicators from mobile and satellite imagery, **N. Pokhriyal**, 2023, *In preparation*,

Selected Talks and Presentations

12. Understanding existential societal problems using a computational lens,
2023 AAAS Annual Meeting, *to be held March 2-5*, 2023 in Washington, DC
11. Novel data and methods for predicting and mapping multi-dimensional poverty index,
Invited talk at the **Oxford Poverty and Human Development Seminar Series**, Human Development Report Office at the United Nations Development Program and the Institute of International Economic Policy at George Washington University, Nov 2021.
10. Estimating poverty, inequality and social deprivations in Haiti via machine learning techniques,
National Statistics Office of Haiti, Port-au-Prince and Inter-American Development Bank, Washington DC, 2020.
9. Social media data reveals signal for public consumer perceptions, ACM International Conference on AI in Finance (ICAIF '20), 2020.
8. Assessing countrywide socio-economic deprivations using auxiliary data sets, AI for Africa for Sustainable Economic Development Workshop, ACM International Conference on AI in Finance 2020.
7. Multiple talks on Combining disparate data sources for improved poverty prediction and mapping at National Statistics Office of Senegal, United Nations Development Program (UNDP), UNICEF, Sonatel Telecom, Dakar, Senegal, 2019.
6. Multi-view learning from disparate sources for Poverty Mapping, AAAI Doctoral Consortium, 2019.
5. A Computational Approach to Poverty Mapping, Intl Conf on Computational Sustainability, Cornell, 2016.
4. Virtual Networks and Poverty Analysis, National Statistics Office, Sonatel, Senegal, June and November 2015.
3. Virtual Networks and Poverty Analysis in Senegal, NetMob, MIT, April 2015.
2. Computational Framework for Novel Gene Discovery via Machine Learning, Oak Ridge National Laboratory, Computer Science Research Seminar, February 2012. (Invited)
1. Knowledge Discovery from Nuclear Reactor Simulation Data, International Workshop on Analytics for Cyber-Physical Systems, SIAM International Data Mining Conference, 2013.

Collaborative grant writing experience

4. Facebook Research's proposal for statistics for improving insights, models, and decisions (2021) (with Prof. Soroush Vosoughi, Dartmouth) (**Finalist**).
3. Assisted in a proposal for computational models of narrative understanding (with Prof. Soroush Vosoughi, Dartmouth).
2. Neukom Institute for Computational Science, Dartmouth, compX grants with Prof. Soroush Vosoughi, Computer Science and Prof. Benjamin Valentino, Government Department on risk assessment of violence and mass atrocities for countries across the globe (2020).
1. NSF Center for Identification Technology Research proposal on biometric identification on social media and mobile networks (2016) (**Finalist** with collaborators from Clarkson University).

Recent leadership and initiative skills

1. Irving Institute seed grant program, Dartmouth College (2020) (**Awarded** as Principal Investigator) with academic-industry collaboration - I conceived the idea, managed, lead, and successfully completed the work.
2. Mentoring middle school Science Olympiad team at Fairfax County Public schools, VA (2022-2023).
3. Meeting with Quisqueya University, Haiti, and Inter-American Development Bank in Feb 2020 for a participatory exercise to build technical capacity for mapping poverty and inequality using the environment and mobile phone data.
4. Workshop on using mobile data for poverty projections at National Statistics Office, Senegal, 2019.
5. Invited to blog on poverty mapping at **the Brookings Institution** at `https://www.brookings.edu/blog/africa-in-focus/2015/06/02/big-data-for-improved-diagnosis-of-poverty-a-case-study-of-senegal/`.
6. Did a TV Interview for encouraging women to join STEM fields in Buffalo, NY in Nov 2016.
7. Participated in panel and biometrics STEM outreach event at Niagara Falls High School, NY in 2016.
8. Protégé in the Women in Computing Mentorship program, Oak Ridge National Laboratory, TN in 2013.

Mentoring

Middle school science olympiad team (2022-23 at Longfellow middle school, VA) Kshitij Tayal (for a year as visiting Master's student at UB in 2015, now a Ph.D. student at the University of Minnesota); Saumya Tripathi (for a summer as a visiting undergraduate student, UB, 2016); several MS and Ph.D. students in CS Dept, UB (for semester-long engagements).

Reviewer

Journals: Proceedings of National Academy of Sciences (PNAS), Nature Human Behavior, Nature Humanities and Social Sciences Communications, Sociological Methods and Research (SMR), SAGE Journals, Information Technology for Development (Taylor & Francis).

Conferences: International Conference on Biometrics (ICB), Biometrics: Theory, Applications, and Systems (BTAS).

Teaching

Lectures for graduate machine learning seminar, 2014.
Intermediate Data Structures and Algorithms, Fall 2009.

Professional Membership

ACM USTPC (US Technology Policy Committee)
ACM (Association for Computing Machinery)
AAAS (American Association for the Advancement of Science)

Dear Members of the Hiring Committee,

I am excited to apply for the position of Assistant professor at the School of Data Science at the University of Virginia. I am a computer scientist with research background at the intersection of machine learning and policy. Currently, I work in the Directorate for Computer and Information Science and Engineering at National Science Foundation (NSF) as an American Association for the Advancement of Science (AAAS) Science and Technology Policy fellow (STPF)[1]. At NSF, I am part of the leadership team of the ExpandAI program[2], whose aim is to start AI research and workforce development for underserved communities through capacity building and partnerships with the AI Research Institutes. I am also part of the NSF's National AI Research Institute program, and as a policy fellow, I gain an understanding of the policies and priorities related to the development and use of AI technologies across NSF and, more broadly, across various federal agencies.

My research interests and technical expertise are in developing statistical machine learning and data science solutions for tackling pressing public policy problems. I distill complex societal problems, like poverty, inequality into computational tasks that are characterized by jointly learning from disparate data sources, often in limited training data settings, and involve quantifying model uncertainty. I focus on building interpretable ML models and quantifying the different types of biases in digital data. I have always been a collaborative researcher having worked in interdisciplinary settings with economists, political scientists, physicists, and statisticians. I have collaborated with developmental banks, and governmental offices, and have worked in building consensus with experts from diverse disciplines.

I was the 2022 Mirzayan Policy fellow at the National Academies of Sciences, Engineering, and Medicine in DC where I worked with the Committee on National Statistics in the Division of Behavioral and Social Science and Education on issues of visioning a new data infrastructure for the federal statistical agencies. My special focus was on policy aspects related to data sharing of private data, data governance, and data equity.

From 2019-2021, I was a postdoc at the Department of Computer Science (CS), Dartmouth College for two years, under Prof. Soroush Vosoughi (CS), and Prof. Benjamin Valentino, Department of Government. I completed my Ph.D. from the Department of Computer Science at the State University of New York, Buffalo in 2019. My doctoral work was awarded the university-wide Chih Foundation Research Award in 2019.

I have years of experience in independently writing grants and effectively communicating the project progress with the funders. I was awarded a competitive grant as a Principal Investigator from the Irving Institute, Dartmouth for 2020-21 on energy poverty. During my Ph.D., I was awarded a grant to contribute to a Gates Foundation-funded project. I have also consulted and successfully finished projects with various development banks. While I have sought opportunities and written several collaborative grants during my academic training, I mention two proposals that were chosen as finalists: one for Facebook research (on a computational framework for understanding bias in large-scale language models) in 2021 and a second for the NSF program (on biometric authentication on mobile devices) in 2016.

I have several first-author publications in high-impact multidisciplinary journals, like the Proceedings of the National Academy of Sciences (PNAS), ACM Transactions of Data Science, EPJ Data Science, and IEEE Transactions on Information Forensics. I have presented my work not only at CS conferences but at multidisciplinary venues, like the Oxford Poverty and Human Development Seminar Series and governmental offices. I am also a reviewer for interdisciplinary journals, like PNAS, Nature Human Behavior, etc. I have knowledge of NSF's review processes and, also, emerging funding opportunities in AI across federal agencies.

Being a female with a circuitous path to earning my Ph.D., I am personally aware of some of the challenges faced by underrepresented groups in the computing field and am involved in NSF-wide efforts on diversifying the workforce.

Based on my collaborative and multidisciplinary approach to problem-solving, knowledge of science policy, commitment to broadening participation in computing and scholarly achievements on socially meaningful problems, I am excited to apply to the School of Data Science at UVA. I thank you for your time and consideration and look forward to hearing from you.

Sincerely,
Neeti Pokhriyal, Ph.D.
AAAS Science and Technology Policy Fellow,
Directorate for Computer and Information Science and Engineering,
National Science Foundation

[1] http://www.stpf-aaas.org/

[2] https://beta.nsf.gov/funding/opportunities/expanding-ai-innovation-through-capacity-building

Research Statement

Neeti Pokhriyal

My primary research interests lie in the area of data science and statistical machine learning to develop interpretable computational models for scenarios characterized by noisy, uncertain, and high-dimensional data coming from multiple sources. My aim is to **jointly learn** and reason under **uncertainty** from such **heterogeneous** data. I am passionate about problems that have **broader social impact**, are best tackled in interdisciplinary settings, and have direct policy implications. I am also interested in researching the ethical and social-technical implications of today's data-driven systems.

1. Philosophy: Given the intricate linkages between technology and society, I seek to develop solutions, where the notions of fairness, inclusion, and explainability are as important as robustness, technical rigor, and performance accuracy. I believe that transformative solutions that affect people's lives and livelihoods can only be developed by equal ownership and participation from multiple scientific domains and stakeholders who are involved.

I will, now, describe how some of these ideas have been the guiding philosophy of my research journey till now; and later, propose my vision for the future.

2. Doctoral Thesis: Many learning problems involve data coming from multiple sources, sensors, modalities, or feature spaces, corresponding to different views of some underlying phenomenon and providing both unique and complementary information. The varied data sources are termed as views, and the task of learning from such multi-view data is known as multi-view learning [14]. For example, learning from audio and video, text and images, web pages and click-through data, spatially and temporally displaced sensors, texts in different languages, etc.

My thesis, **Learning from Disparate Data: Applications in Biometrics and Sustainability**, was supervised by Prof. Venu Govindaraju, Department of Computer Science, University at Buffalo in September 2019. I was on maternity leave from 2017 - 2018. My thesis focused on developing two classes of models for learning from multiple views in a supervised setting. The first class of modeling framework is based on Gaussian Processes which employs Bayesian uncertainty to combine the predictions from multiple views. The second class of methods dealt with factorized subspace learning, where the idea is to learn representation from multiple views such that it captures both the shared as well as the per-view components [2].

2.1. Interpretable computational models using novel datasets for better measurement of multi-dimensional poverty and inequality at scale

A major component of my thesis dealt with developing computational techniques for improved measurement of country-wide poverty estimates; and was published in the **Proceedings of National Academy of Sciences (PNAS)**, in Oct'17 [6]. Poverty mapping is the task of estimating and mapping the spatial distribution of poverty and associated deprivations for a country. These maps are important to get a baseline depiction of poverty and to assess the impact of interventions by governments and developmental agencies.

Current ways to estimate poverty are costly and time-consuming, and as a result, these estimates are grossly delayed for poor economies. I showed the efficacy of using anonymized mobile data for high-resolution country-wide poverty mapping for Senegal [4]. This work won the National Statistics prize and a monetary award of 2.5K in the **Data for Development Challenge** held Netmob, MIT, 2015 [4].

I received a grant from **Bill and Melinda Gates Foundation** (OPP1114791)[1] to develop a robust methodology for poverty mapping as well as to travel to Senegal and work with the National Statistics Office of Senegal (NSO), a governmental body responsible for policy planning. It was a collaborative grant with different universities and public-private partnerships in Senegal. My share was **USD 20K** and I was the **project lead** from University at Buffalo, and this grant partly supported by graduate studies.

I developed a novel computational framework that incorporates data from multiple sources, like environmental data (related to food security, economic activity, and accessibility to facilities) and anonymized mobile phone data. I formulated the task of poverty mapping as a regression problem and build a framework that is based on Gaussian Process regression (GPR), which provides uncertainties associated with the predictions and allowed us to combine them from multiple sources. This work showed that combining disparate data sources,

[1] https://www.gatesfoundation.org/How-We-Work/Quick-Links/Grants-Database/Grants/2014/10/OPP1114791

like environmental data, and mobile data provides more accurate predictions of poverty and its individual dimensions, and using this approach poverty can, now, be estimated more frequently, accurately, and for spatially finest micro-regions [6].

A major challenge occurs due to the existence of varying spatial and temporal granularity of different datasets that are used to predict poverty. We employ an aggregation mechanism that brings the different data sets to a common spatial granularity. Via the use of semi-parametric learning methods we provide insights into the important factors from these datasets that our model deems important as well as uncertainties as a measure of trust. This work also received the **Chih Foundation Research and Publication Award**, May 2019, which is a single award of USD 2.5K given each year for doctoral research related to innovation for the betterment of society at the University at Buffalo.

This work emphasized that *big data* sources like mobile phone data and satellite imagery can provide intercensal estimates of poverty, which is especially important in view of the **Sustainable Development Goals** (SDGs).

2.2. From Research to Practice - issues on data privacy and data governance: As a part of this collaborative effort, I presented my work with multiple stakeholders (Sonatel, Senegal) and local governmental agencies like the National Statistics Office, UN Development Program, UNICEF, etc in 2019, so that the benefit of these technologies can reach to the most vulnerable of human populations.

As a follow-on grant for **USD 15K**, I collaborated with the Oversees Development Institute (ODI), London and Datapop Alliance [2] to implement the poverty prediction module using mobile phone data into a privacy-preserving platform (called Open Algorithms [1]).

During 2019-2020, I started a new collaboration (as an independent consultant) with the **Inter-American Development Bank**[3] (IDB), Washington DC, to develop a computational framework to map inequality and poverty for Haiti, the poorest country in the Western hemisphere that is constantly battered by natural disasters and political instability. I showed how publicly available data can be used to *nowcast* the inequality and social indicators of poverty. Some of the detailed inferences for Haiti are given in the working paper [13].

As a follow-on work, I am engaged with the bank to do a participatory exercise with Quisqueya University in Port-au-Prince during 2021 to transfer the algorithms for poverty mapping in Haiti. I am excited about this work, as it involves building the capacity of the Haitian people and enabling change at grassroots level.

2.3. Explainable models for learning from multiple modalities for building secure and private biometric authentication systems

The second part of my thesis focused on biometrics. I worked in the area of **behavioral biometrics**, which attempts to learn a usage pattern for a person's behavioral and psychological attributes, like swipe statistics, gait analytics, language usage, etc; and can be combined with physical biometrics, like a fingerprint to provide better authentication capabilities. I developed a novel cognitive biometric modality [8, 9], that can be used to build more secure and seamless continuous authentication systems via understanding the language usage of an individual from social media data.

I also developed an algorithm to fuse information from multiple views called **Discriminative Factorized Subspace** that learns factorized subspaces from the different views in a supervised setting, by mapping them to a low dimensional single shared subspace (that captures the common information among the different views) and view-specific subspaces (that captures the information private to each view), and showed its efficacy for touchscreen biometrics [5].

3. Postdoctoral Research

In Oct 2019, I began my postdoctoral work funded by the **Institute for Security, Technology, and Society**, Dartmouth College, NH, and jointly done under the supervision of Prof. Soroush Vosoughi, Department of Computer Science and Prof. Benjamin Valentino, Department of Government. Some of the exciting avenues of my research are as follows:

3.1. Novel probabilistic models to understand demographic biases in social media data and mitigate the ethical and privacy concerns of existing works

While numerous studies have shown the efficacy of using digital data to derive measures of social, political and economic realities of the population, but the question remains can inferences derived from such data be used to inform policy, which is currently being done using sample surveys. Many researchers, however, have raised important concerns about the biases inherent to these data resulting from the self-selection of individuals onto social media platforms and among those users who choose to participate in discussions. If we hope to use digital data to draw conclusions about the broader public, therefore, we need to understand and account

[2] https://datapopalliance.org/

[3] https://www.iadb.org/en

Teaching Statement

Neeti Pokhriyal

I believe teaching and mentoring are central to academia. My teaching philosophy involves **instilling creativity, critical thinking, research ethics, discipline, and hard work** in students. I believe that it is extremely important for students to have deeper conversations about ethics that guide them not only in their research and collaboration but also help them understand the **societal implications** of the technology they create. I believe that developing critical reflective thinking about technology is as important as cultivating mathematical rigor and technical finesse.

Some of the pillars of my teaching philosophy are described below:

1. **Instilling ethics training and reflective thinking about the societal implications of technology**: To cultivate discussions about ethics and broader societal implications of the technologies that we create, I will develop training and teaching modules by collaborating with experts in ethics and fairness areas and making them either mandatory readings or as good practices in projects.

2. **Fostering creativity and collaborative thinking**: Some of the examples via creativity and critical thought can be implemented in traditional classroom teaching by providing students with various media of expression, either through coding, open-ended projects, video games and augmented reality. Besides weaving creativity into instruction, I will pay special emphasis to bringing it into research design by fostering out-of-the-box thinking.

   Based on my personal experiences with collaborative work, I think it is important to realize the importance of listening and an open dialogue while participating in an interdisciplinary discussion. I propose to promote good interpersonal skills and a synergistic work environment by leading by example. I will also craft training modules to impart such skills and provide ample opportunities for students to participate and contribute to discussions.

3. **Bringing innovative classroom teaching tools**: I plan to bring in a discussion and interaction in the classroom, by employing various networking tools like *Piazza*. This increases student engagement and assists in long-term retention. I want to develop effective teaching methodologies that increase student engagement, attendance, retention, and learning in the classroom. Some of my key pedagogical foundations include active learning in the classroom, interactive engagement between faculty and students and between students and their peers, and a synergistic environment on the confluence of teaching (imparting necessary knowledge), application in real life scenarios, and research (seeking for avenues of growth with the current understanding).

4. **Designing new collaborative courses and mentoring**: At the graduate level, I can teach data mining and related courses. At the advanced graduate level, I can propose a project-based collaborative course, whose content includes the myriad ways in which auxiliary data sources are used to measure important outcomes of human development, like poverty, education, electricity access, etc. via the use of machine learning techniques.

   In the past, I have been involved with mentoring high school and undergraduate students, who have all gone to pursue higher education. I plan to continue mentoring especially high school students with a special focus on increasing the participation of women, persons with disabilities, LGBTQ communities, and other historically under-represented groups in CS via many programs like the Computing Research Association's (CRA) Collaborative Research Experience for Undergraduates (CREU) and Distributed Research Experience for Undergraduates (DREU) [1].

5. **Building diversity of thought and experiences and promoting inclusive spaces where ideas can be openly shared**: Diversity brings innovation; and I believe in bringing people of diverse thoughts, experiences, and skill set so that the research products that we create are reflective and inclusive of all. I will actively seek diversity and inclusion in my teaching, and make an explicit effort to point out unconscious bias while forming teams for research and reviewing papers/proposals. [2]

[1] https://cra.org/cra-wp/

[2] https://royalsociety.org/ /media/policy/Publications/2015/unconscious-bias-briefing-2015.pdf

6. **Nurturing scientific writing, presentation skills**: It is very important for graduate and undergraduate students to learn how to effectively communicate their ideas via scientific writing and presentation skills. I plan to build these skills by providing students with an early exposure to writing and reviewing manuscripts, and encouraging them to submit their work to conferences and workshops early on.

Diversity Statement

Neeti Pokhriyal

I believe that diversity leads to innovation and by bringing people of diverse thoughts, experiences, and skill sets together, we can create a future that is reflective and inclusive of all. I believe in encouraging and retaining a diverse group of individuals by providing them with a conducive environment that lowers their barriers to study and increases their chances of success and satisfaction in the computing discipline.

I am currently in a leadership role in the NSF's ExpandAI program, whose aim is to expand AI research access and workforce development activities for underserved communities, and work towards the goal of diversifying the AI-ready workforce.

As a minority female in Computer Science, who had a circuitous journey navigating her Ph.D., I am personally aware of the special challenges faced not only by minorities but also by caretakers, who have to intercept their research/career path because of a major change like maternity or taking care of sick parent/sibling. Owing to the lack of policies regarding providing financial and professional support during and after returning from these breaks adversely affects the retention and growth of minorities, especially women.

I am excited to see the growth of activities or opportunities regarding widening participation[1] of people across different groups, that are traditionally underrepresented in the computing discipline. I have participated in various activities and in multiple roles in such programs at different stages of my career. I was a part of the mentor-protege program in 2012-2013 while I worked as a research associate in the Oak Ridge National Laboratory (ORNL), where I learned first-hand some of the challenges faced by women in different roles in the industry as well as some of the policies at national labs that help mitigate those biases. I was also part of the Women in Science initiative in 2012-2013 at ORNL.

While my previous roles were mostly participative, I took a leadership role in a STEM Outreach Event at Niagara Falls High School in conjunction with the NSF Center for Identification Technology Research (CITeR) in 2016. I was part of a panel that encourage high-school students, most of whom were from under-served communities about the exciting careers in STEM. I partook in hands-on activities related to forensics as a part of outreach. I did a television interview encouraging girls with a personal narrative on STEM education.

I have first-hand seen the exodus of young girls in my high school from science and technical classes, mostly due to a lack of mentoring and a lack of strong female role models. This trend resulted in very few women in undergraduate and graduate programs as well as in workplaces. At a personal level, this makes the research journey very isolating. I have been part of numerous teams while working in industry and academia. I have personally noticed how my creativity and satisfaction have flourished when working in diverse teams.

While there is lots of impetus in increasing the entry of people from under-served communities, especially women, I believe there is another area that needs work if we want to see a more diverse workforce, which is policies that encourage their retention in the research field. I believe we need more inclusive family friendly policies for all genders, and researchers across all roles from graduate students to postdocs to early career scientists that can support them when they return from periods of extended absence.

I believe that for building transformative technology solutions that are free of biases, are embedded in ethics, and are inclusive of all, one needs to assertively convey that computing and computer science are for everyone and anyone can do it. I plan to realize this via my research questions, technical solutions, teaching paradigms, and outreach activities.

[1] https://cra.org/cra-wp/

Names of references for Neeti Pokhriyal

1) Prof. Venu Govindaraju
   Vice President for Research and Economic Development, University at Buffalo
   SUNY Distinguished Professor, Department Computer Science and Engineering
   Director, Center for Unified Biometrics and Sensors
   University at Buffalo, State University of New York
   Email: govind@buffalo.edu

2) Prof. Benjamin Valentino
   Professor of Government
   Chair, Department of Government
   Dartmouth College, NH
   Email: Benjamin.A.Valentino@dartmouth.edu

3) Prof. Soroush Vosoughi
   Assistant Professor of Computer Science
   Dartmouth College, NH
   Email: Soroush.Vosoughi@dartmouth.edu