# Sunipa Dev

Email: sunipad@cs.utah.edu
Mobile: +1 (214) 566-4559
Website : sunipa.github.io

## Research Interests

My research interests primarily pertain to text representations and language modeling, and various modules of Machine Learning, NLP and Algorithmic Fairness. My current work actively checks for social biases in AI with the goal of fairer and more interpretable AI in practice. In particular, I work with the geometry of distributed embeddings such as word embeddings to identify and decouple concepts captured by the different subspaces, leading to more coherent usage in text applications and tasks.

## Education

- **University of Utah** — Aug 2016 – present
  *PhD Computer Science; Advisor : Jeff M Phillips* — *Salt Lake City, UT*
  - **Coursework:** Data Mining, Machine Learning, Advanced Algorithms, Data Visualization, NLP, Databases
- **Indian Institute of Science Education and Research, Kolkata** — Jul 2011 – Jun 2016
  *Integrated Bachelors and Masters in Mathematics and Statistics* — *Kolkata, India*
  - **Coursework:** Differential Geometry, Topology, Statistical Inference, Probability, Stochastic Processes

## Skills

Python, PyTorch, Tensorflow, MATLAB, R, C, Perl, Data Science, Linear Algebra, Algorithms, NLP, Big Data

## Recent Projects

- **Closed Form Word Embedding Alignment** — University of Utah
  - Formulated linear transformation between two high dimensional embeddings based on single iteration rotation, translation and scaling to align embeddings generated from different mechanisms (GloVe, Word2Vec) for boosting word embedding performance and multilingual translations with >80 % top 10 accuracy.
- **On Measuring and Mitigating Biased Inferences of Word Embeddings** — University of Utah
  - Developed a probe using the extrinsic measure of textual inference to measuring bias in word embeddings
  - Debiased contextual word embeddings ELMo and BERT to contain approximately 30% less bias
- **Orthogonalize Concept Subspaces and Rectify Biases in Word Embeddings** — University of Utah
  - Minimized the meaningful information by instead orthogonalizing concept subspaces which should not have interdependence such as gender and professions. Preserved $> 15\%$ more correctly gendered information than other debiasing techniques, while debiasing to the same extent.

## Internships

- **Building Interest Specific Embeddings for a User** — Microsoft Bing Ads
  *Manager: Yi Zhang; Mentor: Haijing Wang* — *June 2020 - Aug 2020*
  - Currently conducting research on ways to build distinct embeddings for a given user based on their preferences, analogous to the vector representations of polysemous words
- **Disentangling Features from Transaction Embedding Space** — VISA Research
  *Manager: Wei Zhang; Mentor: Yan Zheng* — *May 2019 - Aug 2019*
  - Built merchant and user embeddings from transaction data for improved reccommendation systems
  - Reduced interdependence of subspaces of cuisine and location from the embedding from 99% to 68% for better translation of user behavior patterns **(patent filed)**

## Publications

- **S. Dev**, T. Li, J.M. Phillips and V. Srikumar; "On Measuring and Mitigating Biased Inferences of Word Embeddings"; AAAI 2020, New York City, USA
- **S. Dev**, S. Hassan, J.M. Phillips; "Closed Form Word Embedding Alignment"; ICDM 2019, Beijing, China; **Invited to Special Issue of KAIS for Best Papers**
- **S. Dev**, J.M. Phillips; "Attenuating Bias in Word Vectors"; AISTATS 2019, Okinawa, Japan;
- **S. Dev**, T. Li, J.M. Phillips and V. Srikumar; "OSCAR: Orthogonal Subspace Correction and Rectification of Biases in Word Embeddings"; under review;

## Teaching and Awards

- Teaching Mentee for Data Mining (Spring 2018, 103 students) and Foundations of Data Analysis (Fall 2017, 43 students) : conducted lectures, tutorials, project evaluations and office hours
- CI Fellows Award Recipient 2020 (by CRA and NSF); Grace Hopper Scholar 2019;