Sunipa Dev

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

Research Interests

My research interests primarily pertain to high dimensional spaces and their properties, and various modules of Machine Learning, Data Mining and NLP. I currently work with the geometry of word embeddings, the bias and concepts captured by subspaces and ways to disentangle such subspaces.

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.
- **Attenuating Bias in Word Vectors** — University of Utah
  - Devised different novel ways that perform better than state-of-the-art to detect and remove bias from word embeddings; Used names as an effective way to detect inherent gender, racial or age based bias in textual data.
  - Successfully reduced gender bias in GloVe embeddings by 25% and extended to racial and age based bias.
- **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

Internships

- **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
  - Developed novel metrics to evaluate the quality of embeddings generated
- **Simulation of Stimuli Processing in the Human Visual Cortex** — OIST Japan
  Advisor: Dr. Robert Sinclair — *May 2015 - Aug 2015*
  - Generated stimulus data using a Gaussian distribution and simulated flow of stimuli in the visual and auditory cortex to understand Synathesia using mathematical modeling

Publications

- **S. Dev**, J.M. Phillips; "Attenuating Bias in Word Vectors"; AISTATS 2019, Okinawa, Japan;
- **S. Dev**, S. Hassan, J.M. Phillips; "Closed Form Word Embedding Alignment"; ICDM 2019, Beijing, China;
- **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

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
- Received Graduate Fellowship at School of Computing, University of Utah
- Secured the DST Inspire awarded to the top 1% students in science in India