I strongly recommend learning about me at my website: https://siddsach.github.io

CURRICULUM VITAE

Siddharth Sachdeva
February 2018
ssachdeva@uchicago.edu

FOCUS

Machine Learning and Natural Language Processing applied to Computational Social Science

EDUCATION

Statistics, CS at University of Chicago — Chicago, IL
GPA**:** 3.7/4.0 — Expected, June 2018

WORK

Flipside — Chicago, IL
Founder/ML Engineer — January 2017~

- Led startup AI Social Opinion Platform with 750+ users, 3-Person Technical Team, $33k raised
- Designed novel Opinion Extraction Algorithm combining Neural Encoder and graph-centrality algorithms to achieve 83% test accuracy vs. baseline (55%)
- Designed, Built Data Visualization system that displayed different sides of an issue for 2000 users in <0.1s. time using Matrix Factorization, Dimension Reduction, Clustering, and Statistical Analysis
- Built Networking, Web Scraping, AWS, Database Management

Content Carnivores — Chicago, IL
Data Science Intern — June 2016-September 2016

- Proposed and implemented machine learning algorithms to speed up lead acquisition 25X
- Designed, Built system that applied graph-based algorithms to Twitter
- Designed, Built Supervised Twitter Profile classifier for investors by engineering a lexicon, improved accuracy over existing system from 64% to 86%

Aavishkaar Venture Capital — Chicago, IL
Impact Investing Intern — June 2015-September 2015

- Successfully pitched Hydroponic Farming to Investment Committee for $2 million investment
- Conducted intensive deal sourcing, finding and calling dozens of startups for preliminary screening for possible investment opportunities.
- Ground-level market research in Mumbai and Bangalore culminating in reports for investment analysts

SELECTED RESEARCH/PROJECTS (code, writing, and results at https://siddsach.github.io/portfolio/)

Using Attention to Interpret Source Domains in Transfer Learning

- Independent Research Project interpreting deep domain transfer using self-attention
- Built Deep LSTM Language Model, LSTM Classifier, 2 Kinds of Attention, and weight sharing
- Advised by Karl Stratos at Toyota Technological Institute of Chicago

Hybrid Recommender System for Movies based on Ratings and Scripts

- Implemented paper from 2011 that jointly learned representations using Probabilistic Topic Modeling and Matrix Factorization to build State of the Art Recommender System
- Able to make high-quality recommendations based solely on text, solving “cold start” problem
- Crawled scripts and linked with IMDB Movie Reviews to create dataset of ratings and texts
- Improved over traditional collaborative filtering approaches by 6 percentage points

Big NLP on Reddit Comments

- Combined Sentiment Analysis Named Entity Recognition, and Knowledge Graph Search to extract (User, Public Figure, Sentiment) tuples and Unsupervisedly extract political groups
- Used MapReduce on Google Cloud Platform to analyze 26 GB dataset of Reddit Comments
- Extracted and averaged Word Embeddings to construct comment similarity matrix for clustering

RELEVANT COURSEWORK

- Deep Learning
- Nonparametric Inference
- Multivariate Data Analysis using Matrix Decomposition
- Machine Learning
- Numerical Linear Algebra
- Statistical Theory and Methods
- Multivariate Calculus
- Cluster Computing and Hadoop
- Applied Regression Analysis

PYTHON DATA SCIENCE STACK

PyTorch (Deep Learning)
Pandas
NumPy
Scikit-learn
NLTK
Spacy
Networkx

ACCOLADES

Chicago Booth Social New Venture Challenge Winner
University of Chicago Polsky Accelerator Winner
Dean's List 2014-2017
Stanford BASES Finalist

SKILLS

Proficient-SQL, Javascript
Familiar-Scala, R, Java, C, Networkx
Mandarin, Hindi