<div align="center">

CURRICULUM VITAE

Siddharth Sachdeva
https://siddsach.github.io
December 2017

</div>

FOCUS

Machine Learning and Natural Language Processing (code at Github, projects and writing at my website)

EDUCATION

University of Chicago   Chicago, IL

B.A. in Statistics, Minor in Computer Science   Expected, June 2018

GPA: 3.7/4.0

WORK EXPERIENCE

Flipside   Chicago, IL

Founder/CEO/ML Engineer   January 2017~

- Led startup AI Social Opinion Platform with 750+ users, 3-Person Technical Team, $33k raised

- Designed novel Main Opinion Extraction Algorithm combining Neural encoder, TextRank, and MPQA Dataset to achieve 83% test accuracy, a 60% improvement over standard NB-SVM (55%)

- Designed, Built Automatic Opinion Clustering Interface that determined different opinion groups 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

- Designed, Built system that applied network theory algorithms to Twitter to speed up lead acquisition 25X

- Designed, Built Supervised Twitter Profile classifier for investors by engineering a lexicon, improved accuracy over existing system from 64% to 86%

SELECTED RESEARCH/PROJECTS

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

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

Hybrid Recommender System for Movies based on Ratings and Scripts

- Jointly learned representations using Probabilistic Topic Modeling and Matrix Factorization to build SOA Recommender System
- Able to make high-quality recommendations on totally new content, 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

RELEVANT COURSEWORK

- Multivariate Data Analysis using Matrix Decomposition
- Machine Learning
- Numerical Linear Algebra
- Statistical Theory and Methods
- Multivariate Calculus (MATH 200)
- Intro to Probability
- Computer Science w/ Applications (CS121)
- Data Management (CS122)
- Cluster Computing and Hadoop (CS 123)
- Applied Regression Analysis
- Applied Price Theory
- Honors Calculus
- Deep Learning
- Nonparametric Inference

ACCOLADES

Social New Venture Challenge Winner
Polsky Accelerator Winner
Dean's List 2014-2017
Stanford BASES Finalist
Advanced 1st Place at Deerfield Debate tournament, largest tournament in New England Boarding School League
Music Director's Award at NMH School

SKILLS

Python Data Science Stack – PyTorch (Deep Learning), scikit-learn, gensim, NLTK, Spacy, Networkx, Pandas, NumPy (proficient)
SQL, Javacript; (familiar) Scala, R, Java, C
Certified Proficiency in Mandarin, Hindi
Plays Indian, African, Korean Latin Drums