Siddharth Sachdeva

DATA SCIENTIST

CONTACT DETAILS
EMAIL: siddsach@gmail.com
WEBSITE: siddsach.github.io
LINKEDIN: linkedin.com/in/siddsach/
GITHUB: github.com/siddsach

Work History

SOFTWARE ENGINEER, MACHINE LEARNING

Roam Analytics | August 2018 - present

PROJECT: MESSAGE ROUTER
Built an NLP system for automatically choosing who should answer a patient message based on the content of the message for a large telemedicine company

- Took implementation leadership in dataset creation, analysis, preprocessing, feature engineering, training, tuning, error analysis, deployment and model monitoring tools in python for Roam's first live-production model, predicting **6k patient messages/week**
- Got **84% Micro-F1** (within **0.1% human performance**),
- Worked with Clinical and Operational stakeholders to identify target metrics, define label schema to suit their needs, and deliver model and performance report leading to **200K client contract conversion**
- Achieved business impact of r**educing message load on doctors by 55%**, equivalent to saving ~million/year
- Designed and Conducted Experiments to validate data/annotation quality

INTERNAL CONTRIBUTIONS
- Trained text classification model for foreign languages, with **89% accuracy** in classifying portuguese abstracts
- Built tool for using embedding spaces to build specialized lexicons, reducing time needed by 10-100x
- Wrote framework for relation extraction experimentation in pytorch, with 3 different starter architectures
- Wrote preprocessing pipelines for text tokenization, normalization in python, used across projects.

SOFTWARE ENGINEER / STUDENT STARTUP FOUNDER

Flipside | December 2016 - June 2018

TECHNICAL
- Trained and Deployed novel Opinion Extraction Algorithm combining Neural Encoder and graph centrality algorithms to achieve **83% test accuracy vs. baseline (55%)**
- Designed, Implemented Data Visualization system that clustered different sides of an issue for **2000 users in <0.1s.** time using Matrix Factorization, Dimension Reduction, Clustering, and Statistical Analysis
- Built REST API, Web Scraping, AWS, Database Management

LEADERSHIP
- Pitched to Chicago Booth alumni and won **33k grant funding** in flagship venture competition
- Led student startup building AI Social Opinion Platform with **750+ users, 3-Person Technical Team**

PROJECTS:
https://siddsach.github.io/portfolio/

Education

UNIVERSITY OF CHICAGO
Statistics + Computer Science
2014-2018
3.7 GPA

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
Computer Science W/ Applications
Applied Regression Analysis

Skills

Python
- PyTorch (Deep Learning)
- Pandas
- NumPy
- Seaborn
- Scikit-learn
- Annoy/FAISS
- Flask

SQL
Docker
Git
Mandarin/Hindi

Accolades

Chicago Booth Social New Venture Challenge Winner $20k
UChicago Polsky Accelerator $10k
Stanford BASES Finalist