

# Abhishek Malali

#2645A Harrison Street  
San Francisco, CA 94110  
<https://www.linkedin.com/in/abhishekmalali>

abhishekmalali@gmail.com  
(617) 982-4783

## EDUCATION

### Harvard University

*Master of Engineering in Computational Science and Engineering*, GPA: 3.8/4.0

Cambridge, MA

August 2015 - May 2017

- Coursework: Data Science, Machine Learning, Parallel Computing, Stochastic Optimization, Systems Development for Computing, Extreme Computing, Artificial Intelligence
- Graduate School of Arts and Sciences Commencement Marshal, IACS Scholarship Awardee

### National Institute of Technology - Karnataka

*Bachelor of Technology in Electrical and Electronics Engineering*, GPA : 9.50/10.0, Class Rank: 1

Surathkal, India

May 2014

- Relevant Coursework: Probability Theory, Soft Computing, Random Signal Processing, Linear Algebra
- Awards: Institute Gold Medal 2014, O.P.Jindal Engineering and Management Scholarship 2010, 2012, 2013

## PROFESSIONAL EXPERIENCE

### Tribe Dynamics

Data Scientist

San Francisco, CA

July 2017 - Present

- Applying Natural Language Processing for text classification tasks and building ETL pipelines for handling text data streaming.
- Building internal tools for recommendation tasks and data visualization.
- Spearheading Internal research on trends in social media networking.

### Harvard University/University of Chile/Catolica Data Science Exchange

Graduate Student Researcher

Cambridge, MA

December 2015 - May 2017

- Modified the LSTM module to learn from irregular time series and generate predictions. Added a secondary module for regularizing and removing autocorrelation from the residuals in time series prediction as a part of the thesis research.
- Published a Github package TimeFlow for quick implementation of time series prediction models in TensorFlow in collaboration with H2O.ai.

### H2O.ai

Data Science Intern for leading open-source enterprise level machine learning startup

Mountain View, CA

May - August 2016

- Integrated hyper-parameter optimization with H2O and benchmarked different hyper-parameter optimization methods for H2O AutoML.
- Built custom functions for extending H2O's munging capabilities and replicating pandas functionalities on a big data scale.

## PROJECTS

### Machine Learning in Spark

December 2015

- Implemented random forests and ordinal regression in Apache Spark from scratch. Ordinal Regression functionality did not exist in MLLib which is the standard ML Library in Spark.

### How to ask questions?

May 2016

- Evaluated two different questioning schemes for multiclass labeling problems. Generated data with a confusion matrix for each expert, and an underlying class distribution. Attempted and successfully recovered both parameters and the true labels. Methods used were Expectation Maximization, PyMC and Simulated Annealing.

### Quantitative Analysis of Soccer Player Performance

December 2015

- Studied the current soccer transfer market and perceived the valuation on the subjective measures like the league and the club. In addition we created a metric to assess players on the impact they create during the match.

## SKILLS AND INTERESTS

- Programming : Python, Tensorflow, Theano, SQL, PySpark, H2O, MATLAB, D3, Latex, Unix
- Leadership: Harvard Graduate School Leadership Institute Cohort IX, Teaching Fellow for Data Science (CS209)