# Manthan Thakar

617-371-7788 | thakar.ma@husky.neu.edu | manthan.io | github.com/Manthan787

## EDUCATION

**Northeastern University, Boston, MA**
College of Computer and Information Science, GPA: 3.7/4.0 — Jan 2016 - present
*Candidate for Master of Science in Computer Science* — Expected graduation: May 2018
Related Courses: Program Design Paradigm, Natural Language Processing, Information Retrieval, Algorithms, Statistics & Data Analysis, Computer Systems, Large Scale Data Processing

**Sardar Vallabhbhai Patel Institute of Technology, Vasad, India**
*Bachelor of Engineering in Computer Engineering,* CGPA: 8.16/10 — Apr 2015

## TECHNICAL KNOWLEDGE

**Languages**: C, Go, Java, Python, R, Scala, Racket
**Platforms**: Linux, Mac OS, Arduino, AWS, Amazon EMR, Leap Motion
**Software**: git, Elasticsearch, Apache Spark, Hadoop, GraphQL, Django, Flask
**Databases:** Postgres, Oracle, SQL Server, MySQL, MongoDB, DynamoDB
**Certifications**: Oracle Database 10g: Introduction to SQL, Oracle Database 10g: PL/SQL Fundamentals

## WORK EXPERIENCE

**Lexumo, Burlington, MA** — Jan 2017 - Aug 2017
Software Engineer Co-op

- Built a crawler to crawl github for 100K open source Python software projects, analyzed programming patterns using ASTs and built a source code similarity model using TF-IDF, Tree Edit Distance and Characteristic Vectors.
- Designed and developed Lexumo's licensing feature for C/C++ and Java source code, responsible for extracting license information from source code using String Matching and Information Retrieval based methods
- Extended Lexumo's analytics pipeline that finds vulnerabilities in source code using CVE databases for Java packages

**Northeastern University, Boston, MA** — May 2016 - Dec 2016
HPC Graduate Assistant – Research Computing, Information Technology Services

- Automated tasks using Python and Bash scripts for monitoring of Northeastern's Discovery HPC cluster
- Assisted researchers in automating and parallelizing jobs by designing SLURM and IBM LSF batch scripts

**Acharya –** *Helping educational institutes create efficient workflows* — Dec 2014 – Dec 2015
Co-founder

- Developed the entire web application consisting of - AngularJS single page application in the frontend and REST API hosted on AWS built on top of Laravel Framework and MySQL database in the backend
- Successfully shipped the product to 10000+ users across 5 institutes

## ACADEMIC PROJECTS

**Spark Join optimization of RDDs** — Oct 2017 – Dec 2017

- Improved Apache Spark's RDD **join operation execution time by 24%** and **reduced network I/O by 89%,** by building a **Scala compiler plugin**
- Modified Spark's join implementation to incorporate Broadcast joins and **improved performance by 50%**

**PyDO: Python Distributed Objects** — Nov 2017 – Dec 2017

- Designed and implemented a distributed object system in Python that allows **distributed object-oriented programming** and keeps the object distribution, caching, placement and migration transparent to programmers

**Distributed Key Value Store** — Nov 2017

- Designed and Built a fault-tolerant distributed key value store with proxy server written in Go and node servers written in Python

**Thread-safe Memory Allocator Library** — Oct 2017

- Built a thread-safe memory allocator library including malloc, free, calloc and realloc functions with per-thread arenas in C using Buddy Memory Allocation technique

**Clustering Music Artists using Apache Spark** — Oct 2017

- Implemented agglomerative hierarchical clustering and k-means clustering from scratch in Scala on top of Apache Spark to cluster music artists and deployed the application on Amazon EMR

**Search Engine, Northeastern University** — May 2016 – Aug. 2016

- Developed a multi-threaded crawler using python with distributed frontier management to crawl 200K web pages, indexed them in a distributed Elasticsearch index and implemented PageRank and HITS to rank websites
- Trained a Linear Regression model to increase the average precision of the IR system by 156%

## INDEPENDENT PROJECTS

**Hitch -** *A cross-platform Airdrop* — Aug 2017 - Sep 2017

- Built a Desktop and Android application for file transfer on the local network using multicast DNS protocol for service discovery, TCP connections for data transfer, electron for Desktop App and Java for Android App