

**Pankesh Bamotra**  
Mob : (412) 636 1222  
Mail : pbamotra@andrew.cmu.edu  
Github : pbamotra  
Blog : <http://sciontist.com>

## Education

**MCDS** **CARNEGIE MELLON UNIVERSITY, SCHOOL OF COMPUTER SCIENCE.**  
**MASTER OF COMPUTATIONAL DATA SCIENCE - DEC. 2016 - GPA 3.63.**  
▷ *Intro to Computer Systems.* ▷ *Machine Learning.* ▷ *Search Engines.* ▷ *Probability and Mathematical Statistics I.* ▷ *Natural Language Processing.* ▷ *Probabilistic Graphical Models.* ▷ *Multimedia Databases and Data Mining.*

**B.Tech CSE** **VELLORE INSTITUTE OF TECHNOLOGY.** BACHELOR OF TECHNOLOGY IN  
COMPUTER SCIENCE AND ENGINEERING - MAY 2013 - GPA 9.24.  
▷ *Data structures and algorithms.* ▷ *Operating systems.* ▷ *Linear algebra.* ▷ *Computer programming and problem solving.* ▷ *Digital Logic.* ▷ *Algorithm design and analysis.*

## Experience

**Autodesk, Inc.** **SOFTWARE INTERN - BIG DATA ANALYTICS.** MAY '16 - AUG '16.  
Worked with cloud platform team to migrate data from legacy system and ingest it into new Spark based platform  
▷ *Worked with AWS, Spark, Oozie, Hadoop.* ▷ *New platform aimed to tune data ingestion, transport, and fast compute layers.*

**PayPal India Private Limited** **SOFTWARE ENGINEER.** JUN '13 - JUL '15.  
Worked with Open Analytics Platform group to develop self-service database provisioning tool for analysts and developers  
▷ *Refactored legacy code into a maintainable and testable stack.* ▷ *Developed plug and play solution for Qlikview and Teradata.* ▷ *Delivered SQL based solution to analyse clickstream data.* ▷ *Worked with one of the first agile teams at PayPal.*

**PayPal India Private Limited** **SOFTWARE DEVELOPMENT INTERN.** JAN '13 - JUN '13.  
Worked with seller risk management to create early warning dashboards  
▷ *Developed a SQL based solution to analyse buyer account habits.* ▷ *Early warning indicator reports used by teams across NA and EMEA for fraud analysis.* ▷ *Modeled post transaction risk mitigation.*

## Projects

**Info Retrieval** **IMPLEMENTED A SEARCH ENGINE ON LUCENE.**  
▷ *Supports exact match (Ranked Boolean).* ▷ *Best match using Okapi BM25 and Language model.* ▷ *Sequential dependency model.* ▷ *Query expansion.* ▷ *LeToR enabled.*

**ML** **BRAIN fMRI IMAGE DATA ANALYSIS.**  
▷ *Classification of fMRI images using C-SVM.* ▷ *Prediction of fMRI image voxels using ElasticNet Ridge Regression.*

**NLP** **QUESTION GENERATION AND ANSWERING SYSTEM - BEST PROJECT AWARD.**  
▷ *Wikipedia dataset.* ▷ *CLI interface to generate and answer questions.* ▷ *Generate semantically, grammatically, fluent questions and answers.*

**Prob. Graphical models** **CONTEXT GENERATIVE RECURRENT NEURAL NETWORK LANGUAGE MODEL.**  
▷ *Penn Treebank dataset.* ▷ *Optimized by Gensim, faster-rnnlm.* ▷ *Deep learning.*

**Data mining** **DATA MINING ON LARGE SCALE GRAPHS.**  
▷ *Social graph datasets - Epinions, Gnutella.* ▷ *SQL based algorithms - Pagerank, K-core decomposition, SVD, triangle counting.* ▷ *Pattern analysis on static and temporal graphs using power law analysis.*

**Capstone** **FINANCIAL PRICE PREDICTION FROM EARNING CALL TRANSCRIPTS.**  
▷ *In-progress.* ▷ *Predicts directional movement of stock on next trading day.* ▷ *Predicts change in stock price on next trading day.* ▷ *Analysis purely using NLP and historical stock prices.*

## Technical

### Advanced Intermediate

Java 7/8, Spring framework  
Python 2.7, Scikit-learn, Stanford Core-NLP, R, MATLAB, C, Postgres, Apache Hadoop, Apache Spark, Apache Oozie, Apache Hive, Apache Zeppelin, AWS, Splunk, Oracle 10g

## Publications

### IOSR Journal (Sep '12)

**SECURING GROUP COMMUNICATION IN PARTIALLY DISTRIBUTED SYSTEMS.** SYMMETRIC KEY CRYPTOGRAPHY, KEY DISTRIBUTION.

### IJSCE (Nov '12)

**SECURE TRANSMISSION OF GRAYSCALE IMAGES USING DISCRETE FOURIER TRANSFORM.** IMAGE PROCESSING, SECURITY.

### IJARCSSE (Dec '12)

**IMAGE ENCRYPTION USING PIXEL SHUFFLING.** ENCRYPTION.

### IJAIS, FCS (Jan '13)

**LRUD SHIFT BASED ENCRYPTION METHOD USING MATRICES.** ENCRYPTION, SECURITY, LINEAR ALGEBRA.

### IJAIS, FCS (May '13)

**HADOOP AND RISK ANALYTICS.** ANALYTICS, FINANCE, BIG DATA.

## Certifications

### Coursera

Computing for Data Analysis, The Data Scientist's Toolbox, R Programming, Getting and Cleaning Data, Exploratory Data Analysis

## Extra-Curricular Activities

### Public speaking

Vice President for three terms at PayPal IDC Toastmasters and awarded for Competent Communicator milestone, area level international speech contest runners-up in 2014

### Sports

Cricket, Badminton, High-altitude trekking