

## SAURABH MARATHE

San Jose, California

**Personal Email id:** [saurabhmarathe1992@gmail.com](mailto:saurabhmarathe1992@gmail.com)  
**LinkedIn URL:** [www.linkedin.com/in/saurabhmarathe](http://www.linkedin.com/in/saurabhmarathe)  
**GitHub URL:** <https://github.com/SAURABHMARATHE>

**Contact Number:** +1 (669) 265-5167

### EDUCATION:

M.S. in Computer science, San Jose State University, CA GPA: 3.8/4.0  
Bachelor of Engineering in Computer engineering, University of Pune, India

Aug 2016- (Expected)May,2018  
Jun, 2014

**COURSEWORK:** Advanced Topics in AI, Topics in Mobile Networking & Cloud Computing, Big Data Analytics, BioInformatics

### PATENTS:

1. System For Brand/Entity Analysis By Performing Sentiment Analysis Of Social Media using Neural Networks Indian Patent Appln No: **201721004567**
2. A DBMS software to store structured, unstructured and complex datasets using Neural Networks Indian Patent Appln No: **201721006423**

**PUBLICATION:** Product Analysis Using Customer Reviews (using Naïve Bayes and Rule Based algorithm)

Oct 20, 2014

Publication URL: <http://troindia.in/journal/ijcesr/ijcesr/paper14.pdf>

### TECHNICAL SKILLS:

<b>Programming Lang:</b>	Java, C, C++, Python, JavaScript, PHP	<b>Development Frameworks:</b>	JEE/J2EE,JPA, JSF, Spring
<b>UI/Web/Full stack</b>	Bootstrap, HTML 5, CSS, AngularJS,	<b>Databases:</b>	Oracle, MySQL, SQL Server, HBase,
<b>Technologies:</b>	MEAN stack, jQuery		Hive
<b>Wireless Nw/Cloud:</b>	SDN, NFV, AWS, Google App Engine	<b>Web Services:</b>	SOAP-based services, REST
<b>Machine Learning</b>	Neural Networks, PCA, SVM, Decision	<b>Data Clustering techniques:</b>	k-Means algo, k-Nearest neighbors
<b>Algo:</b>	Trees, HMM, Genetic algo		algo
<b>ML Libraries:</b>	ScikitLearn, NumPy, SparkML, PyKE,	<b>Big Data Analytics:</b>	Apache Hadoop, HDFS, MapR, Hive,
	Weka, NLTK, TensorFlow		HBase, Spark, Kafka

**CURRENT WORK:** Research Assistant under Prof Leonard Wesley. Building a deep learning Convolutional Neural Network to analyze images of micro array chips

**INDUSTRY EXPERIENCE:** Systems Engineer, Infosys Technologies (2 years)

Jul 2014- Jul 2016

**Infosys Projects:** (All projects are proprietary software of American insurance company Aetna Inc.)

1. **Plan Sponsor Application:**
  - Reduced insurance processing time by 50% by designing automated system for classification of problems faced in enrollment. Used Natural Language Processing here
2. **Aetna Voice Advantage:**
  - Implemented a SOAP based service and upgraded project to better framework in very short period in 2 weeks of stringent time. Received "Employee Award" for this work
3. **DocFind Search Engine:**
  - Implemented 'internationalization' to make Spanish version of Aetna Inc website to help millions of Spanish customers have better experience

### FULL STACK DEVELOPMENT: (MEAN stack and JSP-Java MVC stack)

- Designed and implemented a dating website using MongoDB, Express, AngularJS, NodeJS
- Developed a Brand Analysis website is a full stack website (HTML/JQuery, JSP, JavaScript, Java, Plain text files as DB. The project was hosted on AWS, and is offline because of expiry of free tier hours

### PROJECTS/ASSIGNMENTS: (Cloud Computing/Wireless Networking):

- Study of Distributed systems: coordination, concurrency, consensus, Proxos algorithm (labs by Nutanix company)
- Cloud computing paradigms: Zookeeper, Map-Reduce, Curator, Leader Election using ML techniques (on going project)
- Study of SDN, NFV, LTE and 5G, Internet of Things and use of tools like Mininet, OpenDaylight

### RESEARCH EXPERIENCE: (Machine learning/AI) (2 years)

2014-Present

- Brand Analysis using real time tweets and Facebook posts by implementing a neural network from scratch (patented)
- Solving a Sudoku using single layer 799 neuron Hopfield Neural Network
- Analysis of UCI datasets and building prediction models from ScikitLearn provided algorithms (like SVM, PCA, D-tree)
- Bio-informatics- DNA nucleotides alignment prediction using HMM(Ongoing project)
- Use of TensorFlow library to build CNN to analyze images of chips for classification of faulty chips

### PROJECTS: (Data Science) (1 year) (Few important proj highlighted below)

- A online live sentiment analyzer was built using IBM Bluemix Alchemy. Kafka was used to store tweets
- Data Analytics of CARS dataset using all algorithms from SparkML
- Implemented Parallel processing of novel sentences for sentiment using Hadoop MapReduce paradigm