SAURABH MARATHE

San Jose, California

Contact: +1-669-265-5167 **LinkedIn:** http://www.linkedin.com/in/saurabhmarathe
Email id: saurabhmarathe1992@gmail.com **GitHub:** http://github.com/SAURABHMARATHE
Personal website: http://saurabhmarathe.me

EDUCATION:

M.S. in Computer Science, San Jose State University, CA (Expected) May 2018
Bachelor of Engineering in Computer Engineering, University of Pune, India June 2014

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

PATENTS: (Provisional)

- System for Brand/Entity Analysis by Performing Sentiment Analysis of Social Media Using Backpropagation Neural Network (Natural Language Processing Text mining project) (Java project) — Indian Patent Appln. No: **201721004567**
- DBMS software to store structured, unstructured and complex datasets using Neural Networks (A project in C language) — 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:

Programming Languages/Frameworks:	Java, C, C++, Python, JavaScript, R(beginner), JEE/J2EE	**ML Libraries:**	Scikit-Learn, TensorFlow, Keras
		Cloud Computing:	AWS, Edge computing, Paxos, Raft
Machine Learning Algorithms:	Neural Networks, PCA, SVM, Decision Trees, HMM, Genetic algo.	**Big Data Analytics:**	Apache Hadoop, HDFS, MapR, Hive HBase, Spark, Kafka, kNN, k-means

PROFESSIONAL EXPERIENCE:

Intern Machine Learning Engineer, Autodesk, San Francisco, US Jun. 2017- Dec. 2017

Work in Agile sprints on machine intelligent product like Autodesk Virtual Agent (AVA), recommender systems etc.

- Implemented a **deep learning RNN/LSTM using Keras** to create word semantic representations of customer chat data
- Developed a **TFIDF-Logistic Regression Active Learning** model to train AVA **1000 times faster** with model-labelled data
- Developed **Word2Vec embeddings** and **DBSCAN clustering** model to find new conversation intents from thousands of chat data
- Researched and designed **Markov Decision process** to optimize dialog flow design for enhanced performance of AVA
- Implemented a **chat based intelligent product recommendation system** for AVA customers

Research Assistant on 2 Machine Learning Software design projects (Ongoing) Feb. 2017- Current

- Decision Support System project: The project involves development of **Decision Support Systems** using **Natural Language** Processing techniques and Neural Network to help companies make better decision. **Python** project. (Ongoing project)
- Graffiti Detection project: Project involved **Development of Deep Learning Convolutional Neural network** for detection of graffiti in street images **Python** project, with TensorFlow. (worked as Research Assistant voluntarily for period of 4 months)

Systems Software Engineer, Infosys Technologies Pune, India Jul. 2014- Jul. 2016

Worked on all phases of Requirement gathering, Design, Development, Integration, Live Deployment and Testing of 3 different software of American Insurance company Aetna Inc., a client of Infosys. Accomplishments are highlighted below

- Reduced insurance processing time by **50%** by designing automated system for classification of problems faced in insurance enrollment. Used Natural Language Processing techniques here. (**A Java/J2EE projec**t)
- Implemented a **SOAP** based service and upgraded project to better framework in very short period of 2 weeks of stringent time. Received **"Employee Award"** for this work. Also implemented **REST** web service for faster execution of light services.
- Implemented 'internationalization' to make Spanish version of Aetna Inc website to help millions of Spanish customers have better experience (**XML internationalization** used)

ACADEMIC PROJECTS: (Machine Learning and Software Development)

DNA Analysis tool for detection of Promoter Regions for Disease Prevention Feb. 2017- Mar. 2017

- Developed a human DNA analysis tool to detect 'promoter regions' in DNA samples for disease protection, using Hidden Markov Model(HMM). Achieved **70% accuracy**.
- Detected principal components by Principal Component Analysis, effectively **reducing** dimensions from **355 to 35**, for protein affinity prediction using Support Vector Machines as part of pre-analysis in same project.

Technology Used: Hidden Markov Model, Support Vector Machines, Principal Component Analysis, Python

Sudoku Solver using Hopfield Neural Network Sept. 2016- Dec. 2016

- Designed and implemented a 729-Neuron Hopfield Neural Network from scratch to solve a full 9x9x9 Sudoku puzzle.
- Neural network solved the Sudoku autonomously with **100% accuracy** through unsupervised learning.

Technology Used: Python

Real time data Brand Analysis website: Trackur.com prototype (patented full stack project) Oct. 2015- May 2016

- Implemented a complete big data analytics website was using JSP, HTML 5, CSS, jQuery and Bootstrap for UI design.
- streamed live Facebook posts were through Kafka. Middleware consisted of Backpropagation Neural Network as learning model

Technology Used: Kafka, Hadoop MapReduce framework, Java/JEE, HTML, CSS, JavaScript, jQuery, Bootstrap