SAURABH MARATHE
San Jose, California

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

EDUCATION:

M.S. in Computer Science, San Jose State University, CA, GPA: 3.8/4.0 Aug. 2016- (Expected) May 2018
Bachelor of Engineering in Computer Engineering, University of Pune, India, GPA: 3.3/4.0 June 2014

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

PATENTS:

1. 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**
2. A 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) (Java project) Oct. 20, 2014
Publication URL: http://troindia.in/journal/ijcesr/ijcesr/paper14.pdf (International Journal)

TECHNICAL SKILLS:

Programming Languages/Frameworks:	Java, C, C++, Python, JavaScript, R(beginner), JEE/J2EE, Spring	**UI /Web/Full stack Technologies:**	Bootstrap, HTML 5, JSP, CSS, AngularJS, MEAN stack, jQuery
Machine Learning Algorithms:	Neural Networks, PCA, SVM, Decision Trees, HMM, Genetic algo.	**Databases:**	Hive, HBase, Oracle DB, MySQL
		Big Data Tech:	Hadoop, Spark, Kafka

PROFESSIONAL EXPERIENCE:

Intern Machine Learning Engineer, Autodesk, San Francisco, US (Ongoing) Jun. 2017- Current
Work on machine intelligent product like Autodesk Virtual Agent (AVA), recommender systems, intelligent routing etc.

- Implemented a **deep learning RNN/LSTM using TensorFlow** to create word semantic representations of customer chat data
- Developed a **TFIDF-Logistic Regression** machine learning model to train AVA 1000 times faster and with increased accuracy
- Used the **LSTM word embeddings** and **k-means clustering** model to find new conversation intents from millions of chat data
- Currently implementing **Markov Decision process** to optimize dialog flow design for enhanced performance of AVA

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.
- Graffiti Detection project: Project involves **Development of Deep Learning Convolutional Neural network** for detection of graffiti in street images as part of preliminary research work for Smart City San Jose project. **Python** project, with TensorFlow. (pilot proj.)

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: (Software development and full-stack development included)

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.
- Also detected principal components by reducing dimensions from 355 to 35 for protein affinity prediction using Support Vector Machines as part of pre-analysis of same project.
  Technology Used: Hidden Markov Model, Python

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

- A 729-Neuron Hopfield Neural Network was designed and implemented from scratch to solve a full 9x9x9 Sudoku puzzle.
- The Neural network solved the Sudoku with 100% accuracy and solved the Sudoku autonomously through learning.
  Technology Used: Python

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

- A complete data analytics website was implemented using JSP, HTML 5, CSS, jQuery and Bootstrap for UI design.
- Middleware consisted of Java code. Live Facebook posts were fetched using Facebook4J API. Posts streamed through Kafka.
- The Middleware consisted of Backpropagation Neural Network that identified entities in the posts, identified features of the brand/product discussed in the posts and did sentiment analysis on data. RitaWordNet library was used for NLP pre-process.
  Technology Used: Java/JEE, HTML, CSS, JavaScript, jQuery, Bootstrap, Kafka