

SAI VIVEK KANAPARTHY

Sunnyvale, CA-94086; (210)-712-9966; sai.vivek15@gmail.com; [linkedin.com/in/saivivek15](https://www.linkedin.com/in/saivivek15); [saivivek15.github.io](https://github.com/saivivek15)

Professional Experience of 3+ years in developing Big Data and ML applications

EDUCATION

The University of Texas at Dallas

Aug. 2016 – Dec. 2018

Master's Degree; Computer Science (Spl: Data Science); Jonsson School **Graduate Scholarship Recipient**

3.82/4.00

(Coursework: Machine Learning, Natural Language Processing, Big Data Analytics, Database Design, Web Programming Languages, Implementation of Advanced Data Structures and Algorithms, Statistical Methods for Data Science, Information Retrieval)

Indian Institute of Information Technology & Management Gwalior

Jul. 2011 – May 2016

Integrated Post Graduate (B. Tech + M. Tech); Information Technology

7.64/10 (Cumulative)

10/10 (M. Tech Thesis)

PROFESSIONAL EXPERIENCE

Realtor.com (Move Inc)

Santa Clara, CA

Senior Machine Learning Engineer

Jan. 2019 – Present

- Productionalized and scaled personalized recommended systems in the real estate domain for 16M users on a daily basis and improved the user engagement by 6% and reduced the performance time by 20% using PySpark, AWS Athena and Elastic MapReduce
- Lead, designed and deployed the architecture of large scale machine learning data pipelines based on terabytes of user behavior data integrated with AWS Glue for ETL transformations, Sagemaker for hosting ML Inference endpoints orchestrated with Step functions
- Developed a platform to track the performance of the ML models to analyze the business impact user engagement, and retention
- Developed the personalized notifications and email campaigns for new listings in the market, recommended listings based on the user's activity and A/B tested with multiple model variants for different audience segments
- Research, design and implemented ML modules on the website to improve buyer experience like Similar Homes to the home the user is viewing, Affordability Calculator which recommends affordable neighborhoods
- Developed Real profile, which gives the portfolio of the user based on his activity in the past and deployed it as a candidate generation model. Developed a robust model based on LightGBM, Match Score, which gives the relative importance of any listing for a user and deployed it as a candidate ranking model for multiple use cases

Mavenir Systems

Dallas, TX

Software Engineer Intern

Jan. 2018 – Aug. 2018

- Designed a framework to create state chart XMLs. Developed the front-end Graphical User Interface using Angular 4
- Developed the back-end of the framework using Python, which builds the SCXMLs from the flow charts on the GUI

VOMO

Dallas, TX

Jr. Developer Intern

Sep. 2017 – Dec. 2017

- Developed the accounts application for sign up and sign in of a user and registered organization on the VOMO platform using Node.js and deployed it on AWS EC2
- Analyzed data to recommend volunteer projects for users using machine learning models, content based and collaborative recommender systems

Securonix

Dallas, TX

Big Data Developer Intern

May. 2017 – Aug. 2017

- Developed integrated framework using SOLR, Hadoop and Spark for improving the quality of big data solutions on large datasets
- Optimized the indexing and search performance of Apache SOLR Cloud using Java, Spark and Web Services built on Spring MVC in Cloudera environment (SAAS) which uses Hadoop Distributed File Systems

University of Texas at Dallas

Dallas, TX

Web Developer

Nov. 2016 – Jan. 2017

- Designed web pages of Prof. Yvo Desmedt using HTML, CSS, JavaScript and exposure to tools like Latex

PUBLICATIONS & AWARDS

- "Stylometric Analysis of E-mail Content for Author Identification.", *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*. Oct. 2017.

- "Hack Day Winner", Realtor.com, Mar. 2020

CERTIFICATIONS & TECHNICAL SKILLS

Certifications: **Coursera:** Machine Learning, Neural Networks and Deep Learning, Object Oriented Programming in Java, Java Programming: Solving Problems with Software, Advanced Data Structures in Java, Algorithms Design and Analysis;

IBM: Big Data Foundations, Hadoop Foundations; **MongoDB University:** MongoDB for Developers (Python);

Programming Languages: Java, Python (NLTK; Flask; NumPy; Pandas; Matplotlib; Seaborn, Scikit-learn), R

Big Data Technologies: Hadoop, Spark, Kafka, SOLR, Elastic Search, Kibana, Scala, TensorFlow, AWS EMR, Glue, Sagemaker, SQS

Databases: SQL (AWS Athena, MySQL), NoSQL (MongoDB, DynamoDB) **Misc:** Git, UNIX/Linux, Agile

ACADEMIC PROJECTS

- Search Engine Soccer (Crawler4j, Java, Lucene, Spring MVC, Bootstrap, JSP) ([GitHub](#))

- Near Real Time Event Detection (Python, Spark, Kafka, Core NLP, SEMAFOR, MongoDB, scikit-learn) ([GitHub](#))