

Allan Joshua

✉ allanjoshua1984@gmail.com

in [/allan-joshua](https://www.linkedin.com/in/allan-joshua)

🌐 allanjoshua.com

☎ +1-305-301-0643

14+ years of software development experience

5+ years of experience in Machine Learning, Big Data and Apache Spark

Specialties: Machine learning, Computer vision, Apache Spark, Java, J2EE, C, C++

🎓 Education

Stanford University

2010 – 2012

CA, USA

Artificial Intelligence Graduate Certificate - As part of the certificate, Completed CS221 (AI- Principles and Techniques, Sebastian Thrun, 4 units), CS229 (Machine learning, Andrew Ng, 4 units), CS228 (Probabilistic Graphical Models, Daphne Koller, 4 units), CS231A (Computer Vision, Fei-Fei-Li, 4 units) - 16 of the 21 required AI - depth units and CS157 (Logic and Automated Reasoning, Michael Genesereth, 3 units) - 3 of the 6 breadth units (for graduate credit) required for an MS in CS which are transferable to their MS CS program

Anna University

2001 – 2005

Chennai, India

BE Bachelor of Engineering in Computer Science and Engineering

🔧 Experience

Equifax Inc

Aug 2016 – Present

Sr.Big Data Engineer (Machine Learning) - Data Science and Innovation labs

Atlanta, GA

- Worked with researchers to develop the generation 3 of a **proprietary explainable monotonic neural network to predict risk - Issued patent US 10558913** - to provide the ability to extend credit to consumers (particularly) in the sub-prime space who otherwise might not be able to secure credit
- Implemented the algorithm by modifying Apache Spark MLlib library by modifying the objective function using **scala** - backend **HDFS**
- Mine Big Data and other unstructured data - large data-sets (**NLP - 9.6 billion tweets**) to identify signals to help with targeted marketing - using **GloVe, word2vec** vectors, using **Apache spark, Tensor flow**
- Implemented a computer vision POC which would perform Face Recognition - dlib (c++ library ResNET), integrated with IOS to perform Shape prediction and Face matching - Tools/Libraries used (IOS, dlib, opencv) - Willing/Eager to learn new technology and languages to keep myself updated
- Experience working in an **Agile / SCRUM** team to deliver solutions.

Seaview Research

Sep. 2014 – Jul. 2016

Software Development Engineer

Miami, FL

- Built a software to automate the QA and System Testing process which brought down the timelines from **8-weeks** (the industry standard for building a trial) down to **1 week**, thereby bringing down the expense by a considerable amount using **Java**
- Wrote code to export clinical data into a specific format to aid with setting up a new revenue stream.

The Home Depot/(TCS)

Mar. 2009 – July. 2014

Software Technical lead

Atlanta, Ga (Mar 2010 – May 2013), Chennai, India (May 2013 – Jul 2014)

- Worked on POC projects to identify frequently bought products together (Association Rules), Classify if a customer is PRO/CONSUMER based on spending patterns etc.

- > Worked with a Data-Scientist to build a solution to predict a good ordering of the Product lists on the website for the next day. The list could be focused to target seasonal variations, **to maximize margin, to maximize conversion** etc based on the business defined objective. Designed and developed a **MYSQL** database to host data from Omniture, WCS and Endeca and implemented a program to build the features for optimization **Feature engineering**, a User Interface in HTML 5 to provide the ability to modify the Objective function based on business requirements and developed an optimization algorithm to minimize the objective function using the **NDCG metric** to learn the weights from the previous week
- > Lead the offshore team and communicated with several stakeholders at The Home Depot, including Point of Sale, Tool Rental, Finance and EDW to build **RESTful webservices** deployed on **Apache Tomcat** to provide the ability to rent Penske Trucks at The Home Depot locations
- > Involved in Designing and developing an application which spanned across multiple releases to provide rewards to the customer based on the accrued spending. Involved in **Low Level System and Component Design, Sequence diagrams, Class diagrams, Activity diagrams, Component diagrams**
- > Extensively implemented **J2EE design patterns** such as **MVC, Service Locator, Factory Pattern, Session Facade, Singleton, Service Factory**, and **DAO, DTO** and best practices for building enterprise systems.
- > Experience in object-oriented software development and applying good design patterns
- > Provided regular updates to the respective stakeholders of the project on a weekly basis and participated in project management reviews and meeting

Citibank / (TCS)

Nov. 2007 – Mar. 2009

Software Development Engineer

Chennai, India

- > Design, Code, Unit Test, Defect Tracking to closure **Java, J2EE**
- > Developed an online solution for Citibank credit cards which included collections, credit card statements etc.

Polaris Software Labs

Sep. 2005 – Sep. 2007

Software Development Engineer

Chennai, India

- > Design, Code, Unit Test, Defect Tracking to closure, **Java, J2EE**

Awards/Licenses/Certifications

MACHINE-LEARNING TECHNIQUES FOR MONOTONIC NEURAL NETWORKS

- > The generation 3 of the proprietary algorithm used to predict risk - Patent issuer and number US 10558913

DEVELOPER CERTIFICATION FOR APACHE SPARK

- > Completed an Apache Spark Developer certification provided by Databricks

TACKLING THE CHALLENGES OF BIG-DATA

- > A certification provided by MIT professional education

Skills

Languages Java, Scala, Python, JavaScript, SQL, UML, C++

Library/Framework Apache Spark, MLlib, TensorFlow, OpenCV

Machine Learning Discriminative and Generative models, Generalized Linear Models, Logistic Regression, Linear Regression, Support vector Machines, Neural networks, K-means Clustering, RNN

Databases db2, oracle, MySQL

Big Data technologies Apache Hive, Apache Spark

Work Authorization

Visa (H1B) with approved I-140.