Jay Lohokare

Email: jlohokare@cs.stonybrook.edu
Phone: +1-631-997-9047
Website: jaylohokare.com
LinkedIn - https://www.linkedin.com/in/jaylohokare
GitHub - https://github.com/JayLohokare

Education

State University of New York	Stony brook	MS in Computer-Information science	2017 – December 2018
College of Engineering Pune	Pune	BTech in Computer Engineering	2013 – June 2017

Research

Working with Professor Fusheng Wang at Biomedical Big data lab (Big data systems) and Professor Fan Ye at MOCA lab (Machine Learning) at Stony Brook university. Published 7 research papers at reputed IEEE/ACM conferences (ACM MobiCom, IEEE TENCON, IEEE TENSYMP, IEEE ICEI).

Skills

Apache Spark/Hadoop ecosystem, Tensorflow-Keras, PyTorch, Cassandra, MongoDB, MySQL, Titan-Gremlin, Python, Scala, R, JAVA, PHP, HTML-CSS, Arduino-RaspberryPi, Kafka, Android-iOS, J-Meter, Unity 3D, Git, Weka, Ethereum, ReactJS, MEAN stack, Django

Work Experience

Data Science Intern, Digital McKinsey (Data engineering team), McKinsey & Co. New York, USA   **Summer 2018**
Worked as Data Scientist in McKinsey's Data engineering group. Worked in a team of 2 at client site(s) to develop datalake strategy, design end-to-end data-pipelines and build machine learning powered data consumption systems. Created automated data processing, validation and ETL engine using Apache Spark. Implemented GIS analytics framework (Plugin) for pySpark. Helped create data consumption API layer (AWS Lambda, NodeJS) to consume data and insights generated. Created recommendation system (Computer Vision) using deep learning algorithms. *[Spark, Python, Tensorflow-Keras, AWS Neptune, AWS (Lambda, SQS, S3, Glue), DynamoDB, datalake strategy]*

Chief data officer at 'DAIX' (http://www.daix.io)   **November 2017 – June 2018**
Lead data scientist at the startup, responsible for data strategy and data platform implementation. Implemented the data architecture for real time sentiment analytics and data capture of news, social media posts using Apache Spark, Apache Nutch, Tensorflow (BLSTM), Cassandra. Built scalable web-sockets server using Go & Redis for consuming data from Crypto-Exchanges and making it available to front-end application with less than 1 second latency. *[Spark, Kafka, Cassandra, Redis, Go, Python, Tensorflow-Keras]*

Technology architect at 'InsureBits' (http://www.insurebits.io)   **November 2017 – June 2018**
Technology lead at the startup, responsible for overall platform design and development. Developed the Ethereum based P2P insurance platform, leading a technology team of 10. Designed and built an MVP – Android application (Interactive Chatbot, system diagnostics, Chat application, Social media integrations, payment integration), NodeJS backend (REST API, Ethereum integration), Data architecture (To manage Insurance 'pools', User to Ethereum interface, Analytics platform), Web-platform. *[Ethereum, ReactJS, NodeJS, Android, MongoDB]*

Founder, CEO and Chief software architect, Skyline Labs, Pune, India   **2015 - 2017**
Founded and led a Tech-startup working on Smart city and IoT solutions. Formed a team of 15, won multiple awards and hackathons. Collaborated with various government and industry organizations in India to develop a varied range of products including an IoT platform, a smart city portal, various smartphone applications and ecommerce systems. Received multiple 'Best startup' awards and got inducted to Facebook's FBStart program. Estimate number of product users - 20,00,000+.

Research intern, 'Macaseinou' Performance engineering team, NTT DATA headquarters, Tokyo, Japan   **Summer 2016**
Worked in RND section of 'Macaseinou' team. Worked in team of 30 to evaluate and improve 'Uniqlo' website performance by stress testing and implementing backend changes to ensure scalability (JAVA). One of the 5 interns selected from all over India.

Projects/Awards

- **Designed and led development of 'ConnectX IoT platform' – Best BTech project award by Cognizant (100+ teams), Finalist at Smart India hackathon (10,000+ teams), Winner at Tech Mahindra IoT hackathon (400+ teams)**
  End to end Industrial IoT platform to automate, deploy, manage, analyze connected devices. Built for scalability and high availability; ConnectX supports 'edge-analytics', data visualization, real-time data-analytics and multiple H/W platforms. Reduces costs and time for IoT deployment by more than 70%. Made open-source contributions to EMQTT (Erlang MQTT Broker) and Kafka to create an open-source scalable data pipeline for IoT data analytics. Implemented Edge analytics feature to allow 'serverless-analytics' using 'NodeRed', Eclipse Paho (MQTT) which can be easily ported to devices through web-dashboard. *[Spark Streaming, Kafka, eMQTT, Android, MongoDB, MEAN stack, NodeRed, Scala, Erlang]*

- **Hadoop GIS / Spark GIS / iSpeed (Graduate Capstone project)**
  Optimizing Hadoop GIS based 'iSpeed' framework to make it compatible with SparkGIS and Redis. iSpeed is a Hadoop based in-memory platform to accelerate 3d spatial data queries on big data involved in 3D mapping, biomedical analytics and imaging. Developing this framework under Professor Fusheng Wang at the Stony Brook Data Management and Bio-Medical data analytics lab. *[Spark, Hadoop, Scala, C++, JAVA, Redis]*

- **Designed and led development of 'FindX smart-city platform' – Winner at Avaya smart city hackathon (100+ teams), Winner at IEEE PICT Credentz 2017 (Inter-college), Winner Digital Pune city hackathon (500+ teams), Winner at Honeywell Mobility hackathon (200+ teams)**
  Social media platform POC to connect citizen of Pune city with government officials in real time. Revolutionized emergency services by drastically reducing response time through the 'Smart SOS' system. Built on a real-time location tracking framework supporting integration with smart-phone and web-applications. Got incubation support from Persistent systems for building the location tracking service for Public buses in Pune city (Uber for public transport) with distributed-scalable backend. *[MQTT, JAVA, Android, LAMP Stack, Spark]*

- **Designed and led development of 'Kym' decentralized credit score platform – Winner at HackD hackathon by Envestnet (300+ teams)**
  Alternative credit score calculation system. Uses secure data mining, social media/internet crawling and machine learning to extract data from previously untouched data points – Social media, Emails and SMS. Reinforcement learning helps accurately predict credit score thereby bringing more users into the scope of credit scoring. Built multiple applications over this framework – Social lending platform (YHacks 2017), GoPay credit-based e-wallet (GoJek hackathon finalist), Personal finance manager chatbot (Barclays open-minds hackathon finalist). *[R programming, MongoDB, nltk, Android, D3JS-MEAN stack analytics dashboard]*

- **Machine learning projects**
  Phenotypic prediction of Transcriptomic features (Multi-label classification using sklearn SVM, Decision tree, Random forests – 87.3% accuracy), MNSIT dataset Generative adversarial network (PyTorch), Actions classification CNN (PyTorch – 61% accuracy), SVM implementation (Quadratic programing), Multiclass SVM implementation (Stochastic gradient descent), Ridge regression Lasso coordinate descent implementation, GAN MNSIT digits generation, RNN for Kinect human action recognition (84% accuracy)
- **Super sensor Adhoc network**
  Adhoc network of 'super sensors' that can detect any events happening in the room using Machine learning and active listening of different sensor signals. Eliminates need to have sensors in every appliance; provides extended range and mobility using Adhoc networks over BLE. Implemented CNN for event classification achieving 86% test accuracy for 6 events classifications using 4 sensors (LDR, Magnetometer, Mic, Temperature). Implemented Pseudo-AODV protocol for Android using JSONs based data transfer on BLE advertisements *[TensorFlow, RaspberryPi, Python, Android Things, AODV protocol, BLE beacons]*
- **Brain Access**
  A plug and play Brain smartphone interface system – Marks all possible interactions on a display with numbers and listens for numerical data coming from BCI kits to perform the corresponding action. Implemented the system using OpenBCI Ganglion kit. Used Android's accessibility service to overlay all 'onClickListeners' with numbers and implemented BLE interface for BCI-Smartphone communication. Used CNN for creating an EEG feature extraction model and used SVM to classify using these features (84% accuracy for 5 class classification). *[Tensorflow, Keras, Python, Android, BLE]*
- **'Echo Print' Smartphone face recognition using Sound**
  Non-invasive face recognition system using high frequency sound waves. The application classifies faces based on patterns in echo received due reflection of sound waves the system creates. Used Keras based CNN model for classification of echo data for face classification (93.5% cross validation accuracy on equally distributed data). Proved the potential and viability of the system by thoroughly studying learning model and system performance. *[Android, Keras, Tensorflow]*
- **E-commerce framework – Finalist at Rajasthan state hackathon (1000+ teams)**
  A dynamic framework for food ordering. Contains admin app with options to customize content, menu items, advertisements and a customer facing app. Includes backend for discount coupons, targeted push notifications, payments integrations, real time order tracking. This platform is currently used by over 10 restaurants in India *[Android, PHP REST APIs, MySQL, LAMP Stack]*
- **'Ema' Intelligent Home automation system – Winer at Tech-Mahindra/ATT hackathon (500+ teams), Paper published at IEEE TENCON 2017**
  Voice controlled Home automation based on ConnectX IoT Platform. Android application based personal assistant and RaspberryPi based active listening application detects natural language commands and triggers actions on ConnectX IoT platform, to control home appliances. *[Android, MQTT, RaspberryPi, Python, CMU Sphinx based NLP]*
- **Mixed reality interactions platform – Winner at FC Bayern Hackdays 2018 Siemens challenge, Germany (1500 participants)**
  A computer vision and AI powered interactive platform redefining chatbots. Renders life-like avatars of football legends who are capable of interactions (With life-like expressions, voice and motion). AI is powered by various data APIs, BLE Beacon & camera feeds to have a personalized interaction with users. *[ReactJS, OpenCV-JS, NodeJS, IBM Watson APIs, ArtyomJS – active listening for chatbot, TrackerJS, WebgazerJS]*

Publications

- **EchoPrint: Two-Factor Authentication using Acoustics and Vision on Smartphones**
  - Accepted at ACM MobiCom 2018
- **Classification of Cardiotocography signals using Machine Learning**
  - Accepted at IEEE IntelliSys 2018
- **Emergency services platform for smart cities**
  - Published at IEEE TENSYMP 2017
- **An IoT Ecosystem for the Implementation of Scalable Wireless Home Automation Systems at Smart City Level**
  - Published at IEEE TENCON 2017
- **Diagnosis of liver diseases using machine learning**
  - Published at IEEE ICEI 2017
- **Automated data collection for credit score calculation based on financial transactions and social media**
  - Published at IEEE ICEI 2017
- **Scalable tracking system for public buses using IoT technologies**
  - Best paper award at IEEE ICEI 2017