

Ethiraj Srinivasan

Big Data and Machine Learning Engineer

EXPERIENCE

Lomotif, Singapore— Data Engineer

FEBRUARY 2019 - PRESENT

- Built a data lake by dumping data from various sources of event tracker application into S3 using lambda and EMR
- Built dashboards on Quicksight using Athena as the query engine
- Developed jobs on EMR(pyspark) to score users based on their urban factors
- Built an inhouse event tracker application using vertx, Kinesis streams and lambda
- Clustering of video tags using Fasttext
- Building an inhouse recommendation engine for the feeds

Knorex, Singapore— Big Data and Analytics Engineer

MAY 2018 - JANUARY 2019

- Migrated ad events tracker to Vertx and optimized the throughput using Kinesis and Lambda architecture
- Built Reporting Analytics tool on Spark with Scala to determine attributions and generated aggregated data for visualisation

Trakomatic, Singapore— Freelancer

DECEMBER 2017 - MARCH 2018

- Developed Data Management and Annotation feature for Smart Digital Signage with **facial recognition** prototype with Django.

National University of Singapore, Singapore— Graduate Researcher

AUGUST 2017 - MARCH 2018

- Built an **avatar based open domain conversational chatbot** with facial recognition feature to engage the visitors of NUS under the guidance of **Dr. Matthew Chua** and **Prof Chuk Mun Lee (Java Evangelist)**. The application is built using **Java**, **Python** and supporting **Machine Learning** libraries (**dlib**, **keras** with **TensorFlow** as backend)

Pramati Technologies, Chennai— Senior Software Development Engineer

AUGUST 2015 - DECEMBER 2016

- Involved in the development of web based portal (**Ruby on Rails**, **Memcached**, **Postgres**) for patents and built a content management application where content was manipulated, authorized and accessed based on roles. Responsible for conducting **Open Source Meetup**. Handled terabytes of data using **Solr** search engine and hold responsibility as an **author** of Sunspot Solr gem.

Tata Consultancy Services, Chennai— Assistant System Engineer

JANUARY 2014 - JULY 2015

- **Lead** the development of a collaborative platform (**Ruby on Rails**, **Memcached**, **Redis**, **Resque**, **Solr**, **Postgres**) used to share the content, idea and process among group of companies. Developed a platform that provides **gamification**, **webinar management** and **personalized media directories**.

ethirajsrinivasan@gmail.com

ethigeek.com

github.com/ethirajsrinivasan

(+65) 9811-4927

EDUCATION

National University of Singapore

M.Tech Software Engineering

JANUARY 2017 – APRIL 2018

Anna University, India

B.E Electronics and

Communication

SEPTEMBER 2009 - MAY 2013

SKILLS

Language

Python, Scala, Ruby, Java

Web Development

HTML, CSS, JavaScript, jQuery, CoffeeScript

Database Management

Postgres, SQLite, MySQL, MongoDB, RDS, Athena

Framework

Spark, Vertx, Ruby on Rails, Django, Flask, Bootstrap, Foundation

Technologies

Solr, Redis, Memcached, Resque, Sidekiq, Kinesis

Tools

AWS, Google Cloud, Quicksight, Heroku, Digital Ocean, Travis CI, Capistrano, Eclipse, Git, Nginx, Rspec, Jasmine

AWARDS

Winner - Hackathon

DNA Capitals, Singapore

Star of the Learners group

Tata Consultancy Services.

On the Spot Award

Tata Consultancy Services

RESEARCH PROJECTS

- **Deep Learning Models for Medical Image Diagnostics** - Developing models capable of diagnosing lesions suspicious of malignancy in mammogram screening. **Computer vision techniques** are used to preprocess Radiological images and **Keras** with **tensorflow** as backend to build deep learning model.

PROJECTS

Machine Learning

- **Brand Analysis of Social Network feeds** - Data Ingestion using **Flafka** and batch processing on **Spark** were performed, brands were recognized using deep learning model and online dashboard was created using **NodeJS**
- **Relay Activation using Facial Recognition system** - **Raspberry pi** controlled relay is activated by **flask** based **facial recognition** system.
- **Resource constrained project scheduling** - A combinatorial optimization problem where a set of activities and a set of resources of known characteristics are optimized using **genetic algorithm**.
- **uAsk** - A utilitarian **android** mobile app to get answers to questions about student life. The app is also integrated with **chatbot** implemented using **Amazon Lex** to answer frequently asked questions.
- **Movie Recommendation System** - Web log analysis of movie ratings with demographic data (gender, age and occupation) to build a **recommendation engine** to recommend movies that would suit the user using **apriori** algorithm
- **Fatality & Catastrophe Investigation** - Analyzed textual dataset of common workplace injuries to find patterns that might be useful to prevent frequently occurring accidents using **SVM classifier**, **NER** and **CFG**.
- **Sentiment Mining of Customer Reviews** - Extracted Restaurant Review for various cuisines in Singapore from **Yelp's** API. The text data was preprocessed and models were built to classify the review sentiments using **Naïve Bayes**, **SVM**, **K-NN classifier**
- **Social Community Expansion** - The project aims to expand the social community of Facebook users. The three major steps performed were **community detection**, identifying **influencers** in the community and **expanding** the community with a **seed set** of influencers. The community expansion was implemented using **LEMON** algorithm.
- **Spatio -Temporal Analysis of Students' Travel** - The objectives of this project were to help optimize travel times for NUS students and provide helpful information about facilities near their home or travel routes such as nearby restaurants/Hawker centers, Shopping Mall using **Geofencing**, **Hot Spot Analysis**, **Cluster** and **Outlier Analysis**. The application was built using **R Shiny** and **leaflet.js**

Web Applications

- **Tablizer** - This application built on **Sinatra** scraps tables from either external **URL** or **HTML** input and displays it as **CSV**. It also allows users to download the output as **CSV** or **Excel** file.
- **CSS To Inliner** - This single page application is built on **Ruby on Rails** to combine **HTML** and **CSS** to give **HTML** with inline styles. This is extremely useful for mailer templates.
- **Url Shortner** - Url Shortner application gives a quick statistics on the short and provides both web and API based services. The application is built on **Ruby on Rails**.

CERTIFICATIONS

Scala and Spark for Big Data and Machine Learning

The ultimate hands on hadoop tame your big data

Data Engineering on Google Cloud Platform Specialization

Machine Learning by Stanford University on Coursera

Complete Guide to TensorFlow for Deep Learning with Python

Oracle Certified Java Professional

AMCAT Certified Data Processing Specialist

OPEN SOURCE CONTRIBUTIONS

Sunspot - Sunspot is a **Ruby** library for expressive, powerful interaction with the **Solr** search engine.

Supercache - A totally unobtrusive add-on that runs along your **Rails** application rapidly improving your development time by caching **Activerecord Queries** and **HTTP** requests.

Rails Fort - A wrapper on JavaScript plugin that comes with modern progress bar

Rails Protip - Configurable tooltip gem written on top of a next-generation **jQuery** tooltip plugin

Technology Icons - An easy to use rails web font gem for technology icons.

Devise Foundation Views - Creates custom views for **devise** on top of **foundation** framework.

Collection2csv - Download **Activerecord** collection to **CSV** with minimal syntax and provides support across associations.