KARTIK ANIL REDDY

Email: kartiklaw@gmail.com | Phone: +1(313) 413-5874 | Website: http://www.kartik-reddy.com

---

EMPLOYMENT:

Comcast Corporation, *Big Data/ Software Engineer* **Nov '17 to present**

Building a wholesale, highly scalable and automated platform called "Timeline" to facilitate real-time ingestion, analytics and prediction for large scale data. Part of multiple teams like Elements (real-time ingestion) and Clairvoyance (analytics).

- Helped develop a wholesale, automated and scalable platform called "Timeline" that will, and has helped improve customer service in Comcast.
- Built an ETL tool (using Mapreduce and Data pipeline) to migrate all legacy data (3 billion records) from on-premise hardware onto AWS. This ETL tool facilitated migration from Hbase to S3, S3 to DynamoDB and DynamoDB to DynamoDB.
- Developed a real-time ingestion ETL tool (using Apache Flink, Kafka, NiFi, Kinesis, EMR) for consuming data (16,000 records per second) from Kafka servers and and pushing them into Kinesis Streams for further analytics and consumption onto the Timeline UI.
- Helped develop a REST api for producers to push their data onto the timeline platform for analysis. Designed and implemented an infrastructure orchestration tool to facilitate this entire process (RDS, Elastic Beanstalk, Kinesis, Lambda, EMR, S3).
- Built a DateTime conversion library integrated with Jolt Transformations to allow conversions between all 5 major US timezones and across 40 formats used across various Comcast products. This library can be used independently or can be invoked using Jolt Specifications.
- Build a real-time ingestion hook using Apache Flink to consume data from Apache Kafka, apply transformations (deserialization, normalization, encryptions, etc) and put into Kinesis Streams for further processing by downstream analytics components (Flink, EMR, Kafka, Kinesis, KMS).
- Built an analytics platform called Clairvoyance that consumes 10 million records/minute from Kinesis streams and transforms it for ingestion into business analytics tools like Interana and Pointillist (using tools like AWS Lambda, Kinesis, DynamoDB, S3, Firehose, Jolt Transform, KMS and SNS).

HeavyWater Solutions Inc, *Software Engineer, Machine Learning* **Jan '17 to Nov '17**

Build a product to liberate human talent from drudgery by automating various tasks in the Mortgage/Insurance industry.

- Built a scalable platform entirely on AWS called AIVA (Artificial Intelligence Virtual Assistant) which can perform multiple skills. Worked on the skills for reading and comprehending documents which includes File Intake, OCR, Document Classification and Entity Extraction/Classification orchestrated through multi-level workflows (SWF).
- Designed and built a continuous integration pipeline that enables flow and deployment of artifacts through the various stages and environments of AGILE Development (Development, Test, Exploratory and Production) using GitHub, S3, Cloudformation, API Gateway and Jenkins.
- Built an Intake/Output and Packaging Component that can perform zip/unzip operations and allows dual communication between S3 and a secure FTP server, via EC2 instances.
- Created an Elasticsearch domain with dynamically adjustable configurations to index over a million documents. Also built interactive webservices that allow interactions (Create Index, Delete Index, Index and Search documents) with this domain using API Gateway and Lambda functions.
- Used the Tesseract open source OCR engine to transform Images to XML and eventually JSON so that the data can be consumed and processed easily for other applications such as Entity Extraction and classification. The webservice was designed to run on ECS and save the
- Built REST applications using the apache and spring frameworks to train, extract and classify Addresses, Dates, Currencies, Names, Phone Numbers and other document information from an input JSON file that contains all words and their coordinates present in that document. Followed the microservices architecture by building independent webservices for each functionality using Lambda, EC2 and API Gateway.
- Configure UI to access and display documents to Operators based on their access levels (Using DynamoDB, SQS and S3). The operators can review and modify the classified data and submit the changes. This feedback is used for training the classification algorithms.

DataMetica Solution Pvt. Ltd, Pune, India, *Software Engineer (Intern)* **June'14 to Aug '14**

- Effectively delivered assigned NLP module for POS tagging and topic modeling, for a call center client as part of a project team, implementing a speech to text conversion to facilitate reduced storage costs & data search ease.

PROJECTS:

DDI Analysis: Analyzing drug-drug interactions using Python and C++. **FL/ Sept '16**

- Obtaining a database of drug-interactions and drugs from Medline. Extracting mesh terms and calculating z-score and probability of interaction between drugs. Parallelizing the program using threads.

ZIKA ANALYSIS: Mining tweets about the Zika virus in Florida using Python. **FL/ July '16**

- Creatively collected Zika virus tweets restricted geologically for data mining, sentiment analysis and centrality leveraging on Natural Language Processing & Python twitter APIs.

EMOTION AND HAND GESTURE RECOGNITION (*Sponsored by Persistent Systems*) **Pune/May'15**

- Developed software to count number of fingers for gestures using Canny Edge Detection and used Principal Component Analysis and Support Vector Machine (OpenCV library) to classify emotions as one of 8 (angry, sad, disgust, etc.).

EDUCATION:

University Of South Florida		**Tampa, Florida, USA**
Master of Science in Computer Science	GPA: 3.88 / 4.0	Graduated Dec' 2016
Pune Institute of Computer Technology (PICT)		**Pune, India**
Bachelor's Degree in Computer Engineering	GPA: 3.1 / 4.0	Graduated May' 2015

TECHNICAL SKILLS:

Programming Languages: Java, Python, C, C++

AWS: EMR, Elasticsearch, Elastic Beanstalk, EC2, ECS, API Gateway, S3, DynamoDb, SQS, Cloudformation, Lambda, SNS, Route53, SWF, Workspaces, Data Pipeline, Kinesis, Cloudwatch.

Apache: Flink, Kafka, Hadoop, Maven, Tomcat, NiFi.

Database: PostgreSQL, MySQL.

Others: Git, Android SDK, Linux, MacOSX, Jenkins, REST, Spring, LucidChart, Weka, Tesseract, Jolt, Multi-mechanize, Datadog.

PERSONAL(Github link: https://github.com/Pajkouisn**):**

- **Scrum master** of an Agile development team at HeavyWater Solutions Inc Mar '17- Nov '17
- **AWS Certified Solutions Architect Associate** Mar '18 - Mar '20
- **AWS Certified Solutions Architect Professional** Feb '19 - Feb '21