APURBA MAHANTA
apurba5964@gmail.com |+1(716)578-4804 | linkedin.com/in/apurba-mahanta/ | https://apurba5964.github.io/

EDUCATION

Master of Science, Computer Science: **Expected Dec 2019**
University at Buffalo, The State University of New York; GPA: 3.46/4
Bachelor of Technology, Computer Science & Engineering: **July 2015**
VIT University, Vellore, India; CGPA: 8.48/10

TECHNICAL SKILLS

- Programming Languages: Java, Scala, Python, Perl, C++, Javascript, PHP, R
- Database: DB2, Greenplum, Oracle, MySQL
- Frameworks and Tools: Keras, Tensorflow, OpenCV, Numpy, Sklearn, Scipy, Spring Framework, Spring Integration, Spring Boot, Hibernate, Junit, Restful Webservices, AngularJS, Ext JS , Git.
- Big Data Stack: Hadoop, Apache Spark, Hive, Impala,Docker,AWS
- Operating System: Linux, Windows, Android

COURSE WORK

- Graduate Level Focus: Machine Learning,Deep Learning,Algorithm design analysis, Distributed Systems, Data Intensive Computing, Computer Security, Software Engineering.

PROFESSIONAL EXPERIENCE

Publicis Sapient – Senior Associate Technology L1 **July 2015 - June 2018**
(Technology Consultant at Morgan Stanley)

- Developed the data ingestion framework with a very performant / scalable code running on top of Big Data platforms (Spark/Hive/Impala/Hadoop) that ingests **terabytes** of files every day into hdfs. Data loading timings improved massively loading **400 million rows under 4 minutes.** Impala query timings was also **2.9x faster** than Greenplum for analytical queries**.**

- **Flatfile Loader** : Revamped legacy data aggregation framework used in Risk Tech in order to make it scalable, maintainable and improve overall throughput. Apart from rewriting the entire code, performed analysis of GC logs to tweak JVM Args, leveraging JProfiler/VisualVM to identify hotspots in code. Final build was able to provide **30% more throughput consuming 50% less memory**.
- **Subscription Manager** : A publisher-subscriber based model in Java and Spring MVC for notifying via MQ Channel and allowing users to run **dynamic queries on Greenplum database** using load balancer to handle multiple requests

ACADEMIC AND RESEARCH PROJECTS

- **Distributed key value storage on Android(Amazon Dynamo) -** Implemented a failure resilient distributed key-value storage inspired by Amazon Dynamo that provides Linearizability and availability. *Java, Android, Socket Programming, Linearizability.*
- **Distributed Hashtable on Android -** Implemented a distributed hash-table based on Chord that provides node joins, ID space partitioning and ring based routing. *Java, Android, Socket Programming, Chord*
- **Group Messenger Application-** Implemented a group messenger system using ISIS algorithm while maintaining FIFO and Total ordering with a local persistent key value storage .Features include scalability, reliability and fault tolerance.
- **Data Analytics Pipeline using Apache Spark Python-** Implemented data acquisition & analytics pipeline to fetch articles using NY-Times API. Used Hadoop MapReduce and Apache Spark to build article classification model.
- **Exploratory Data Analysis - R**: Analyzed Influenza Outbreak by performing Exploratory data analysis by extracting tweets by using Twitter REST APIs and comparing the data with Official Influenza Statistics
- **Handwritten data recognition in forensics-** Developed an Explainable AI system based on Bayesian Inference model and Deep learning models (Autoencoder and Siamese network) to distinguish between same and different writers.
- **Hierarchical Variational Autoencoder for Music-** Developed a generative novel model on Hierarchical Variational Autoencoder based on Google Magneta's work for music interpolation between music melodies .
- **Reinforcement Learning using Deep Q Network-**Combined reinforcement learning and deep learning techniques to teach the agent to navigate in the grid-world environment.

PROFESSIONAL ACCOMPLISHMENTS

- Awarded with Best Debutante of the year 2015 award, Client Focused Delivery award for Q1 2016 and Core Value Relationships award for Q1 2017.
- Awarded with "**Star of the Quarter**" by Morgan Stanley RiskTech team for continuous 2 quarters, for my dedication to complete an impact-full project at crucial time.