

Arushi Sharma

Boston, MA 914-484-8552
<https://github.com/Arushi04>

sharma.aru@northeastern.edu
<https://www.linkedin.com/in/arushi04/>

SUMMARY

A passionate data analytics professional, with 5+ years of industry experience in financial, telecom and technology domains, with an aim to drive optimal decision making for real-world business problems

EDUCATION

- **MS (Data Science)**; Northeastern University, Boston, MA August, 2021
Relevant Courses: Algorithms, Linear Algebra and Probability, Information Retrieval, Machine Learning
- **B.Tech (Computer Science)**; Rajasthan Technical University, India, **Honours** May, 2012
Relevant Courses: Data mining and Warehousing, Distributed Systems, Statistics and Probability Theory

TECHNICAL SKILLS

- **Programming:** Python, Shell, SQL, R, Terraform, Numpy, Pandas, Scikit-Learn
- **Technologies:** Hadoop, Spark, Kafka, Hive, Cloud (AWS), Docker, Elastic Search, PyTorch
- **Machine Learning:** Regression (Linear, Logistic), Regularization, Decision Trees, Boosting, Bagging (Random Forest), SVM, Naïve Bayes, PCA, CNN, RNN(LSTM), NLP

ACADEMIC PROJECTS

- Document ranking using vector-space, language and machine learning models (Python, Elastic Search)
- Spam classification of emails (Decision Tree, Logistic Regression, Naïve-Bayes)
- Prediction of fatality and progress of COVID-19 (SIR model, SVM, Random Forest, One-class classification)
- Image Classification using Neural Networks (PyTorch, Logistic Regression, CNN)
- Classifying names with a Character Level RNN (PyTorch, LSTM)

PROFESSIONAL EXPERIENCE

Bridge2i Analytics Solutions, India - Senior Analytics Consultant April 2017 - May 2019
Bridge Funnel (Docker, Terraform, AWS : S3, EC2, Elastic Beanstalk, ElastiCache, Lambda, Jenkins)

- Optimized the creation of AI sales product cloud infrastructure, reducing the setup time by 60% using automation
- Containerized algorithms and configured Jenkins for continuous builds, reducing man power by 50 hours/week

Image Analytics (Kafka, Spark Streaming, Cassandra, Tensorflow, Docker, AWS)

- Developed an object detection and classification framework using Kafka and Spark, ingesting live feed from sensors
- Integrated algorithms with AWS, increasing scalability and cost efficiency and reducing execution time by 30%

Anomaly Detection (Spark Streaming, Cassandra, Kafka)

- Designed and developed data pipelines for anomaly detection by collecting data from various sensors
- Performed data cleaning and manipulation on sensor data, using it for real time monitoring of the anomalies

Lead Engine (Intent Marketing) (Pyspark, AWS, DataBricks)

- Optimized the code in PySpark to run on Spark 2.x version and improved the performance of Lead Engine, lowering the real-time execution by 75%
- Designed a level control framework in the application, controlling multiple execution points from a single program.

Accenture, India - Application Development Analyst Feb 2016 - April 2017

Data Architecture and Warehousing (Hortonworks, Diyotta, HDFS, Hive, Hue, Sqoop, Flume)

- Developed data ingestion pipeline, consuming historical and incremental data from multiple sources to Hadoop data lake, helping generate business insights
- Reduced the operation cost by 35% by building automated pipelines and automating hive scripts using Diyotta

IBM, India - Software Developer

Jan 2013 - Feb 2016

DCR Analysis (Cloudera, Hive, Sqoop, Oozie, Pig)

- Developed hive scripts to generate insights from call center data, improving dropped call rate by 25%
- Migrated data to HDFS using sqoop and optimized queries performance using hive partitioning and bucketing

Content Management System (Java, HTML, CSS, Javascript, Teamsite)

- Developed and supported language translation module in Java, increasing revenue and client satisfaction
- Performed root cause analysis of user's functional and technical issues, improving customer satisfaction

ACHIEVEMENTS

- **Certificate of Recognition, Bi2i** for mentoring analytics professionals in Big data, Pyspark and Cloud 2018
- **Managers Choice Award, IBM** for excelling in the practice of "Put the Client First" 2015
- **Orion Award Winner, IBM** for exemplary performance and exceptional dedication to IBM 2014
- **Outstanding Contributor Award, IBM** for excellent performance for the year 2013-2014 2014