

1236 East Lemon Street,
Apt #102, Tempe, AZ 85281
+1 (480) 758 2202

RAMAN AHUJA

rvahuja@asu.edu

[linkedin.com/in/ahujaraman](https://www.linkedin.com/in/ahujaraman)
[https://ahujaraman.github.io](https://github.com/ahujaraman)
github.com/ahujaraman

EDUCATION

Arizona State University – Master of Science in Computer Science

August 2018 – May 2020

• GPA : 4.0/4.0

• Courses: Foundations of Algorithms [CSE 551], Fundamental of Stat. Learning [CSE576], Natural Language Processing [CSE569]

Visvesvaraya National Institute of Technology – Bachelor of Technology in Computer Science & Engineering July 2012 – April 2016

• GPA : 7.80/10

• Electives: Analysis of Algorithms, Data Mining-Warehousing ,Neuro-Fuzzy Techniques, Information Retrieval, Artificial Intelligence

SKILLS AND TECHNOLOGIES

• Programming & Databases: Java, Python, C, MySQL, Amazon Redshift, HTML, CSS, PHP, Oracle SQL 11g

• Frameworks & Tools: Apache Spark, Jenkins, AWS, Git , Numpy, Scikit-learn, MATLAB, OCTAVE, IntelliJ, Jupyter Notebook

WORK EXPERIENCE

Associate Software Engineer | Numerify, India

August 2016 – June 2018

Performance Optimizing Engine for Petabyte Scale Real Time Analytics:

• Developed a highly scalable distribution and sorting framework for millions of records on columnar databases to boost analytical queries, system was **deployed for 30 + tenants** and **reduced daily ETL batch processing time from ~10 hrs to ~6hrs**

• Implemented compression encoding to reduce disk I/O & storage which increased Real Time Analytics performance by 60%

Enterprise Level Feedback Driven Decision Making Platform:

• Built an extractor pipeline for Microsoft-Yammer cloud source to pull scheduled data by continuous streaming

• Trained ML models on features like feedback texts, groups, incidents to detect the bottlenecks from unstructured data

• Built intuitive dashboards on MicoStrategy cubes which led to **20% faster resolution of incident tickets**

Healthcare Infrastructure Outage Analytics Engine:

• Built a predictive model which progressively gets trained on historic incidents, to build RCA trends on critical client outages

• Architected the data warehouse model and developed an ETL pipeline which runs daily on ~20 million records

• Developed interactive dashboards to find actionable insights on outage trends that **elevated outage prevention by ~2x**

Software Development Intern | DoboZ

May 2015 – July 2015

• Designed and developed an android application for digital wallet as a service. Implemented loyalty points program, rewards and incentives, and gift cards purchase-sale on platform

PROJECTS

Live Tweet Analysis of continuous streaming data using Apache Spark

Dec 2018

• Continuous feed to Spark and analysed the tweets every 10 seconds based on Popular Hash tags, Popular Mentions using spark RDD, Data frames and used various spark operations like Map, Reduce. Dashboards reflect live changes in the trends

Conversational Question Answering System (CoQA)

Oct 2018 – Nov 2018

• Developed a novel architecture for machines to answer series of interconnected questions in a conversation, seq to seq modelling

De-noising and Stacked Auto Encoders for Fashion MNIST

Oct 2018 – Nov 2018

• Demonstrated the effect of greedy layer wise pre-training criterion and how it helps to achieve significant results on supervised classification task for deep neural architectures, for very less labelled data

Intruder Detection through Webpage Session Tracking

Aug 2018

• Predicted whether a webpage session belongs to legitimate user or someone else, by training ML models on sequence of historical visits, used sequential pattern mining techniques

Enhanced Naïve Bayes Model for Text Categorization

July 2015 – May 2016

• Developed an novel approach for text based categorization on linear Naïve Bayes classifier at par to state-of-art complex models

• Achieved accuracy of 85.1% on IMDB dataset. Model incorporates N-grams, negation handling , feature selection

Information Retrieval – Indexing Engine

November 2015

• Developed a distributed system on Hadoop Map Reduce framework for construction of inverted index on large scale documents, to improve search query performance. Implemented pre-processing techniques using NLP constructs to improving ranking.

INITIATIVES AND RESPONSIBILITIES

Numerify | Training Across the teams on AWS project

February 2018

Visvesvaraya National Institute of Technology | Student Council – Training and Placement Secretary

June 2015 - May 2016

ACHIEVEMENTS

• Received **“Star of Month – April 2018”** at Numerify, for Impeccable performance on developing optimization engine which reduced AWS infrastructure cost of organisation by 30%