ARUN JAGANATHAN

http://arunjaganathan.me

ajagana@ncsu.edu - +1-(415)-969-0602 - Raleigh, North Carolina

EDUCATION

Master of Computer Science ***Aug 16 – May 18***

North Carolina State University, Raleigh ***GPA 3.7***

Coursework: Algorithms, Operating Systems (OS), Software Engineering, DevOps, Database Management Systems (DBMS), Software Security, Data Intensive Computing, Foundations of Data Science, Algorithms for Data Guided Business Intelligence, Automated Learning and Data Analysis, Data-Driven Decision Making

Bachelor of Technology in Electronics and Communication Engineering ***Jun 10 - May 14***

SASTRA University, India ***GPA 8.0***

SKILLS

- **Languages/OS** C++, Java, JavaScript, Python, R, NodeJS, Linux
- **Technologies/Databases/Packages** Ansible, Spark, Kafka, Git, JIRA, RESTful APIs, SQL, NoSQL, MongoDB, Selenium, Express, pySpark, NLTK, scikit-learn
- **Application** SAP BO Design Studio, SAP HANA, SAP Bex Query, SQL Developer

PROJECTS

- Web Application: OObot - Object-Oriented Bot (Python, Github APIs, Slack, Git, Selenium)
  Developed a slack chat-bot that provides starter template code for a design pattern, based on user's request. Implemented three use cases – *Provide* to the user, *Learn* from the user, *Search* GitHub public repository.
- Web Application: Todo API (NodeJS, MongooseJS, Rest APIs, MongoDB, Heroku, Git, Robomongo)
  Developed a simple Todo application that enables user to maintain a list of things to-do, using REST APIs, nodeJS and mongoDB as the database. Used mongooseJS to connect with the db, and postman to test the APIs.
- Web Application: Work Force Pressure Calculator (NodeJS, ReactJS, Postgresql)
  Engineered an application to analyze the workforce pressure for medical staffs in Uganda, using two open sourced data sources – DHIS2 and iHRIS, for IntraHealth International.

- Operating Systems: Key-value pseudo device (C, Linux)
  Developed loadable kernel module in C under Ubuntu Linux, that allows user to add, delete and maintain a key-value store allowing different processes to share data through accessing this device.
- Operating Systems: Performance Measurement of my system (C, MacOS Sierra)
  Developed a set of programs that evaluates the performance metrics of a computer system like RAM access time, RAM Bandwidth, Page fault service time, time taken for thread and process creation, etc.

- Database Application: Course Registration System (Java, SQL, JDBC, Oracle DB, Git)
  Developed a University Course Registration System with minimal application constraints, involving complex triggers and procedures. Gained experience about data modeling and frontend-backend interaction.
- Database Application: Buffer and File Management in SimpleDB (Java, SimpleDB, Git)
  Extended the buffer with a map data structure to keep track of its contents and modified the buffer replacement policy to Most Recently Modified (MRM). Also, extended the file management system to display file statistics.

- Data Analysis: Twitter Sentiment Analysis (Python, Apache Spark, Apache Kafka, NLTK, pyspark)
  Engineered twitter live stream processor to gauge the text polarity. The system is aimed to analyze the impact of events (concerts, games) using the tweets, to improve audience satisfaction on-the-go.
- Data Analysis: Sentiment Analysis using Logistic Regression, Naïve Bayes Classifier (Python, scikit, gensim, NLTK)
  Developed a python program to generate feature vectors using traditional NLP technique and Doc2vec for use in Logistic Regression and Naïve Bayes classifiers and compared their accuracy using confusion matrix.
- Data Analysis: Yelp User Rating prediction (Python, R, NumPy, SciPy, sklearn, pandas, e1071, Git)
  Analyzed the yelp dataset (3 GB in size) extensively to determine the top attributes influencing the performance of restaurants, thereby predicting the restaurant ratings using Artificial Neural Networks, SVM and regressions.
- Data Analysis: Music Recommendation System (Python, Apache Spark, pyspark)
  Created a recommender system using spark and collaborative filtering techniques (ALS), that will recommend top five musical artists to a user based on their listening history. Used python library for spark for data loading and transformation.

LinkedIn: https://www.linkedin.com/in/arun-jaganathan **GitHub:** https://github.com/arun5493

- Data Analysis: Adwords Placing using Bipartite Graph Matching (Python)
  Implemented the following algorithms: Greedy, MSVV, Balance; to solve the Adwords Placement problem. It was concluded that MSVV algorithm achieves optimal solution with 0.92 competitive ratio.

- Data Engineering: Log Management and Analysis (Python, Apache Kafka, LogStash, Elasticsearch, Kibana)
  Developed an application integrating Kafka and LogStash to parse and store the real-time log data. Used Elasticsearch for querying through the log and analyzing the data using the Visualisation tool Kibana.

- Software Security: Risk Analysis of OpenMRS 2.6.0 (Java, SQLMap, OpenZAP)
  Tested the OpenMRS application extensively – and analyzed the vulnerabilities in it, referring to OWASP Top 10 2017 List. Submitted a pull request fixing some of the vulnerabilities.

PROFESSIONAL EXPERIENCE

Research Assistant - Software Engineer, Professor Michael Kowolenko ***Oct 17 – Present***

Technologies: Python, django, NLTK, Scikit-learn, matplotlib

- Extracted useful medical information from WHO's articles using Scikit-learn package from Python. Built Bag of Words model and TF-IDF model to extract features from the texts and applied Linear SVM and Latent Semantic Analysis (LSA) for classification of the texts into useful domain. Visualized the findings and classifications using matplotlib package from python.
- Working on developing a web application- Yurika, that crawls the web and filters data for further text analytics.

Intern - Office of Information Technology, NCSU ***Jul 17 – Oct 17***

Business Intelligence Analyst - Visual BI Solutions, India ***Oct 14 - Jun 16***

Technologies: HTML5, CSS3, JavaScript, SAP BO Design Studio, JIRA, Agile.

- Addressed client requirements and their Business Intelligence needs throughout the SDLC
- Developed applications using SAP BO Design Studio, representing the client's financial data which includes data modeling in the back-end, optimization, and dynamic representation of data. Made sure applications are mobile compatible.
- Developed advanced table component/extension, one of the utilities for Design Studio, using HTML5, CSS3 and JavaScript, with features such as searching, sorting, conditional formatting etc.

ACADEMIC ACHIEVEMENTS

- Secured the Top 2% (among 3000+ students) in freshmen year of Undergraduate study
- Secured top 0.1% (among 7 lakh students) in Grade 12 Examination.