**SASHIDHAR DONTHIRI**

San Francisco, CA, 94102 | +1-469-703-5424

[Shashi.donthiri9@gmail.com](mailto:Shashi.donthiri9@gmail.com) | <https://www.linkedin.com/in/sashidhardonthiri/> | <https://github.com/shashidonthiri9/>

**Objective and Summary :** Seeking internship for Summer 2018**.** 2+ years of industry experience with strong fundamental knowledge of software development that includes developing web services and web architecture from scratch

**EDUCATION:**

**The University of Texas at Dallas, TX, USA** **Aug 2017- May 2019**

Master of Science in Computer Science, GPA: 3.33

**Indian Institute of Technology, Patna, INDIA** **July 2012- May 2016**

Bachelor of Science in Computer Science, GPA: 3.52

**Coursework:** Web Programming Languages, Database Design, Big Data and Analysis, Advanced Machine Learning, Artificial intelligence, Algorithms and Data Structures, Formal Language and Automata

**SKILLS:**

**Languages**: Java, Python scripting (Keras, Numpy, TensorFlow, SciKit Learn, OpenCV), C#, Scala, R programming

**DBMS**: MySQL, NoSQL, MongoDB

**Data Sciences:** Machine Learning, Data engineering, Map-reduce, cleanup and transformation

**Technologies:** Hadoop, Apache Spark, Azure COSMOS, PySpark, kafka

**Web Design**: JavaScript, JQuery, Bootstrap, NodeJS, SOAP, RESTFUL APIs, AngularJS, XML, JSON, HTML5, CSS

**Frameworks**: Selenium automation framework, Play for Scala/Java, Django, Flask for python

**Packages**: Visual Studio, Selenium, Android SDK, Anaconda, Jupyter, R (rpart, prune tree for regression analysis)

**Statistics:** Time-space series analysis, model selection, ggplot

**NLP**: NLTK, Stanford Core-NLP, Elastic Search-Kibana, Sentimental Analysis, tweepy

**WORK EXPERIENCE:**

**Software Engineer, WIPRO Technologies: July 2016- Aug 2017**

* Developed automation suite for Azure COSMOS dashboards using selenium on CSS, JavaScript, JQuery based pages
* Designed scripts in SCOPE language based on C# and SQL for Microsoft Big data project for data integration
* Deployed Scripts on Azure COSMOS to design fact tables that process, filter, organize and reflect data on dashboards
* Achieved 20% more efficiency in speed implementing scripts for stored procedures using PL/SQL on clusters

**Data Science Research Intern at Institute of High performance Computing Singapore: May 2015-Aug 2015**

* Performed Spatial and Temporal analysis of Marine Accidents: Implemented classification and regression trees on marine accidents that improved accident prevention rate by 40% and accuracy up to 70%
* Analyzed unusual Space-time Patterns by Moran’s I, Local Geary and knox ratio test

**Gift Registry E-commerce Website: Oct 2017- Dec 2017**

* Designed a scalable and responsive client server architecture; e-commerce website using webservices and microservices in python using django framework, AngularJS, SQLite database and Nginix as server
* Implemented Token based Authentication of users for security
* Implemented shared key authentication for micro-services
* Designed and developed REST APIs for backend web, microservices using Django framework and AngularJS, Framework for the UI for the web app

**Library Management and Recommendation system: Sep 2017- Nov 2017**

* Data Engineering of large dataset with huge noise and inserted into SQL database in a structured format
* Designed client-server architecture based multiuser Environment system that handles large amount of user data
* Implemented Search, check in, check out, fines and update system at API level using python and flask framework in RESTFUL webservices that included nested queries in SQL

**Command line SQLite Database Engine: Sep 2017- Nov 2017**

* Developed a basic SQL database engine from scratch that supports data manipulation commands like SELECT, DELETE, CREATE, DROP and UPDATE using B Tree Data Structure with configurable page size in JAVA
* Implemented Memory management using file-per table approach to physical storage where each table will be physically stored as separate file and each table is subdivided into logical sections of fixed equal size call pages

**Sentimental Analysis and Visualization of Twitter data: Feb 2018- Mar 2018**

* Extracted and Scraped tweets in real-time using tweepy, spark streaming by removing noise such as emojis, special characters and save the metadata like location and timestamp that is sent to kibana later used for data visualization
* Analyzed relative strength of tweet if its positive, neutral or negative using Stanford CoreNLP sentiment analyzer
* Visualized metadata using heat maps based on geolocation distribution of tweets through Elastic Search and Kibana

**Kaggle Puerto Seguro Safe-driver Prediction: Sep 2017- Oct 2017**

* Implemented machine learning models like Logistic Regression, Decision Trees, Random Forests, ANN, KNN, Gradient Boosting and achieved Gradient Boosting as best classifier with 0.26 Gini Co-efficient while best was 0.29 on Kaggle