

Shubham Bagi

www.linkedin.com/in/shubhambagi
Website : shubham12595.github.io/

Email: shubham.bagi@stonybrook.edu

Mobile : +1-217-518-7469

Kaggle : kaggle.com/shubham1205

EDUCATION

- **Stony Brook University** Stony Brook, New York
Master of Science in Computer Science; GPA: 3.7 *Aug. 2019 - Present*
 - Courses taken - Machine Learning, Data Science, Artificial Intelligence, Big Data Analytics, Theory of Databases;
- **B.V.Bhoomaraddi College of Engineering** Hubli, India
Bachelor of Engineering in Electronics and Communication; GPA: 4 (9.15/10.0) *Aug. 2013 - July. 2017*

PROGRAMMING SKILLS

- **Programming Languages:** Python, C, SQL, Javascript, Java, C++
- **Skills:** Machine Learning, Data Analysis, Software Development, NLP, Big Data, Deep Learning, Statistics.
- **Libraries:** Pandas, Numpy, SciKit-Learn, matplotlib, Keras/TensorFlow, Pytorch, Mlib, BeautifulSoup, NLTK, Scipy, Seaborn, Plotly.
- **Databases/BigData:** Map-Reduce, Hadoop, Spark, DB2, MySQL, SQL Server, MongoDB, Hive
- **Others** AWS, REST, HTML5, CSS, Unix Scripting, OpenCV, ETL, Tableau, Oracle, XML, JIRA, Confluence, Git, MS-Office, Selenium, Splunk, Postman.

EXPERIENCE

- **Remote Roofing** SouthLake, Texas
Machine Learning Engineer-Data Scientist Intern *May 2020 - Present*
 - Fabricating web applications in **Vanilla Javascript, JQuery, Ajax, HTML** in use by 30 employees, resulting in 10+ new features, reduction of 50% in save/load time and 25% operation time.
 - Implementation of **Deep Learning** using **Mask R-CNN, MS COCO, ResNet101** to remotely inspect and assess roof damage; identified the roof damage with an accuracy of **96.58%**.
- **Stony Brook University** Stony Brook, New York
Teaching Assistant *Aug. 2019 - Dec 2019*
 - Orchestrated the **TA** duties of **250 under-graduate students** under Prof. Michael Tashbook.
- **Robert Bosch Engineering and Business Solutions** Bangalore, India
Associate Software Engineer *Oct 2017 - June 2019*
 - Supervised a team of 4 in **Germany** as **Functional Safety Developer**, to achieve **80%** reduction in testing time by developing autonomous testing software for fault management of CAN Frames.
 - Executed responsibilities on various requirements like **Tester Diagnostic Services, CAN frames, Cruise Control functionality, Driver Demand Requests**. Remodelled the safety software for complex systems like Engine Control, Remote Parking Assistant, Handling Transmission and Vehicle Stability interventions; increased the throughput by an average of **20%** and reduced customer complaints by **50%**.

PROJECTS

- **Twitter Sentiment Analysis** - Drafted a data pipeline which extracts 6000 real time tweets using Twitter Stream API. Used **AWS Kinesis FireHose, S3 bucket and Amazon Elastic MapReduce(EMR)** to stream, store and analyse the Tweet Data. Utilized **PySpark** jobs to load tweets from S3. Achieved an accuracy of 83%.
- **Political Polarization and Marriage** - Collaborated to analyse the 20 million census data of **Florida and New York state** in order to obtain insights on the impact of political party affiliation of 5 million+ couples on their marriages and its correlation to the divorces using **Pandas, Numpy and Sklearn libraries**.
- **MapReduce Library** - Designed a **Hadoop-like MapReduce** facility, with master and worker nodes to process vast amounts of data in parallel on large clusters with 100% fault-tolerance.
- **Image and Video Classification using CNN and RNN techniques** - Classified Human actions in images and video using **CNN and RNN techniques** and **Pytorch** with an accuracy of 89.93%.
- **Recommended System** - Engineered a Recommended System using **Item-Item Collaborative filtering** method and **minhashing method(Similarity Method)** using **Apache Spark** and **HDFS** to achieve an RMSE of 0.765.
- **Spatial Queries Analysis on Covid-19(DB2)** - Explored the New York Health Information and US Census to find closest healthcare facility with ER Room for **NY Counties** using **DB2 Spatial Databases**.