

Raghav Gupta



rwg5289@psu.edu | +1(814) 384-5075 | New York City, NY

EDUCATION

The Pennsylvania State University

Bachelor of Science in Computer Engineering

Graduated: *May 2019*

- Relevant Coursework: Algorithms, Data Structures, Big Data Programming Models, Data Mining, Database Management, Operating Systems, Discrete Mathematics, Probability & Statistics

TECHNICAL SKILLS

Programming Languages: Python, C/ C++, Scala, MATLAB **Web/Databases:** Flask, Django, MySQL, NoSQL, PostgreSQL, HDFS

Data Science Ecosystem: Pandas, MapReduce, Spark, Pig, Hive, Flume, Numpy, Bokeh, SKLearn, MongoDB, Tableau

WORK EXPERIENCE

Hashtag Loyalty, Mumbai

Data Engineer/Scientist

Aug 2017 - Aug 2018

- Segmented existing customers using RFM & LTV scores and built a churn prediction model using the SKLearn framework, to identify customers at risk or lapsed.
 - Targeted these clusters to decrease customer attrition rate by at least 10%.
- Implemented a marketing automation model using K-means for email/sms communications.
 - Increased open rates by 6% using Send Time Optimization, to predict the best times to reach customers.
 - Used customer behaviors and demographics to personalize communications, and tailor better offers.
- Created visualizations using Python and Bokeh framework, to identify top performance statistics per brand/outlets for the sales team to use in business decisions.
 - Architected an in-memory datastore using Redis, to store most queried data for faster retrievals.
 - Designed and Implemented APIs using Flask to push data for the internal dashboard.

1Decisionsuite, Menlo Park, CA

Big Data Engineer

Aug 2016 - Nov 2016

- Built a Big Data automation pipeline for migrating incoming batches of Sales Data(~10 TB) onto HDFS stores.
 - Structured, cleaned and performed analysis on this data using Apache Spark to be used in machine learning applications.
 - Aggregated resultant dataset onto a NoSQL database (MongoDB) for further visualizations.
 - Projected sales trends using Tableau, in order to aid client's decision making.
- Built a real-time processing stream using Spark Streaming to push high velocity data.
 - Created an Apache Flume agent to aggregate log data from various web channels.
 - Integrated Spark to restructure incoming logs to further store on MongoDB.

PROJECT EXPERIENCE

Wireless Ultrasonic Acquisition System with Cloud Interface

Jan 2019 - May 2019

- Built a Flask API to receive ultrasonic signal data routinely, from a remote & wireless hardware monitor installed on a metal pipeline.
- Wrote Backend application to process & store incoming signals and identify possible structural integrity faults in the pipeline.
- Extended the API to push information to a .NET based WebApp for visualizations for the end user.

PeerSkills(HackPSU winner)

Oct 2018 - May 2019

- Extracted valuable data from resumes of various formats using spaCy Framework and string matching.
- Created a dynamically updating YAML configuration file to parse incoming data into respective domains(skills, hobbies, experience).
- Visualized and compared any new resume to an existing database of resumes within the same major of study.

Semantic Segmentation

Mar 2017 - Apr 2017

- Segmented and classified image into semantically meaningful parts(such as background, human, animals, object, etc) using Caffe's Deep Learning Network.
- Created Flask APIs to run the network onto an HTML front end.

User-Space Device Driver

Feb 2017 - Apr 2017

- Implemented a file system Device Driver in C, built on top of an Object Storage Device.
- Translated file system commands (mount, unmount, format, open, read, write, and close) into storage array commands.