

Neel Shah

<https://neelshah18.github.io>

Email : neelshah181993@gmail.com

Mobile : +1-807-356-5415

SUMMARY

Experienced interdisciplinary researcher in health and customer analysis. My academic and extracurricular pursuits have developed a solid knowledge base in data analysis, machine learning, programming languages and scientific programming applications for large scale system design. Recent graduate with my second Master's degree and second publication.

TECHNICAL SKILLS

- **Programming Languages:** Python, R, Java, C, Go
- **Technologies:** AWS, Linux server, Rest API, Flask, ETL, Agile
- **Database:** MongoDB, Elasticsearch, Sql, Spark
- **Machine/Deep Learning:** Regression Models, Tree Models, Clustering Models, SVM, CNN, RNN, Gradient models, LSTM
- **Libraries:** Tensorflow, Scipy, Keras, Qgis, NLTK, scikit-learn, Stanford-NLP

EXPERIENCE

- **Data Analysis Research Assistant**

datalab.science

Thunder bay, ON

November 2017 – May 2019

- Spatial analysis of Social media data to predict health problems in Canada with 94% accuracy.
- Successfully implemented real-time health analysis system using social media data with 93% multiclass classification using Random forest.
- Design, develop and test Machine learning model for analysing and predicting diabetic trend in Canada through social media data on distributed NVIDIA Titan-X CUDA server using tensorflow.
- Created largest open source food database “Food in one” with 333K data points for statistical analysis of country level food buying and consuming pattern.
- Designed and developed “Elasticsearch” on distributed server with 20 nodes which allow searching from 330M tweets/second.
- Performed quantitative analysis and visualization of social media data for health analyzation.
- Created data pipelines for data analysis and machine learning models.

- **Data Analyst**

Datalog.ai

San Diego, CA

January 2017 – August 2017

- Developed backend NER functionality for NLP algorithm used in chatbots.
- Design and developed clustering algorithm for decision making based on chat.
- Successfully implemented 5+ visualization dashboard for backend analysis of chatbot data: semantic analysis, topic extractor & other.
- Implement asynchronous API for chatbot analysis system, which is capable of handling thousands request per second.

- **Data Analysis Intern**

Panchamrut Dairy

Godhra, GJ

July 2016 – December 2016

- Designed and developed a real time data analysis system for raw product cost and product transportation using python, pandas and numpy.
- Forecasted sales pattern of ice-cream using time-series with 71% efficiency.
- Performed cost analysis and report generation to reduce production cost by 1.2%.

EDUCATION

- **Lakehead University**

Master of Science in Computer Science; GPA: 3.9/4.00

Thunder Bay, ON

September 2017 – May 2019

- **Gujarat Technology University**

Master of Engineering in Information Technology; CGPA: 8.78/10

Vadodara, India

August 2014 – June 2016

- **Gujarat Technology University**

Bachelor of Engineering in Electrical and Electronics; CGPA: 7.31/10

Vadodara, India

August 2010 – July 2014

OPEN SOURCE & PROJECTS

- **Emot:** Developed and Maintainer of widely use open source "emot" python library(9k+ Downloads) to extract the emojis and emoticons from a text with their respective meaning. Lead library maintainer team for bug fixing and new features.
- **Sentencelab:** Developed and Maintainer open source python library for text data cleaning, sentiment analysis and topic extraction.
- **Bug solving and test cases:** Solved more than 13 bugs and 10 test model for different open source libraries.

FUNDING

- **Research grant:** 7000\$/Yearly NSERC Discovery Grant.

PUBLICATION

- Neel Shah, Darryl Willick, and Vijay Mago. A framework for social media data analytics using elasticsearch and kibana. Wireless Networks, pages 1–9
- Neel Shah and Vijay Mago. Canadian's health analysis through social media. Data Science, IEEE(submitted)

CERTIFICATION

- Optimize Python code: Lynda.com
- Web API development: Lynda.com

References will be provided upon request.