

Neel Shah

688B Highpoint Ave, Waterloo, ON

Email : neelshah181993@gmail.com

+1-807-356-5415

TECHNICAL SKILLS

- **Technologies:** AWS, Linux server, Git, Rest API, Flask, ETL, Agile, Docker, MongoDB, Elasticsearch, Sql, Hadoop
- **Programming Languages and Tools:** Python, R, C, Go, Tensorflow, Scipy, Keras, Qgis, NLTK, scikit-learn, Stanford-NLP, Tableau, SAS Visual Analytics, Power BI, AWS Data pipeline
- **Machine/Deep Learning:** Regression Models, Tree Models, Clustering Models, SVM, CNN, RNN, Gradient models, LSTM

EXPERIENCE

• Data Analysis Research Assistant

Lakehead University

Thunder bay, ON

November 2017 – May 2019

- Led Research Assistants Team of 12 members for real-time analytical system development.
- Documented the complete process flow to describe program development, logic, testing, and implementation, application integration, coding.
- Implemented discretization and binning, data wrangling: cleaning, transforming, merging and reshaping data frames with NoSQL databases.
- Integrated applications with designing database architecture and server scripting, studying & establishing.
- Worked closely with different stakeholders of the system to understand their needs for this system integration.
- Performed troubleshooting, fixed and deployed many Python bug fixes of the two main applications that were a main source of data for both Canada health department and internal department service team.
- Created RESTFUL API's for several of our Intranet applications using open source software.
- 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 suing python.
- Created Technical specifications documents based on the functional design document for the ETL coding to build the data mart.
- Automated MySQL container deployment in Docker using Python and monitoring of these containers.
- Deployed and tested different modules in Docker containers and GIT.
- Usage of advance features like pickle/unpickle in python for sharing the information across the applications.

• Data Analyst

Datalog.ai

San Diego, CA

January 2017 – August 2017

- Involved in various phases of Software Development Life Cycle (SDLC) such as requirements gathering, modeling, analysis, design and development.
- Created Python and Bash tools to increase efficiency of call center application system and operations, data conversion scripts, REST, JSON, and CRUD scripts for API Integration.
- Automation of unit testing using Python.
- Used Unix/Linux OS based cloud for the deployment. Wrote scripts in Python for extracting data from HTML file using Python library BeautifulSoup.
- 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.
- Extensively involved in Data Extraction, Transformation and Loading (ETL process) from Source to target systems using MongoDB and Oracle database.
- Worked with Technical Experts to Improve Report Performance and performed System testing.
- Providing End-user Training and documentation for customer reporting services.
- Excellent experience with Python development under Linux OS (Debian, Ubuntu, CentOS SUSE Linux, Red Hat Linux, Fedora)
- Co-Project Managed the successful upgrade of NER and analytical module of chatbot systems.

• **Data Analyst/ Python Developer**

Panchamrut Dairy

Godhra, GJ

July 2014 – December 2016

- Working real time data analysis system for raw product cost and product transportation using SAP.
- Developed Python based API (RESTful Web Service) to track sales and perform sales analysis using Flask and SQLAlchemy.
- Developed a fully automated continuous integration system using Git, Gerrit, MySQL and custom tools developed in Python and Bash.
- Played a key role in a department wide transition from Subversion to Git, which resulted in an increase in efficiency for the development community.
- Worked on daily basis with lead Data Warehouse developers to evaluate impact on current implementation, re-design of all ETL logic.
- Worked on exporting reports in multiple formats including MS Word, Excel, CSV and PDF.
- Responsible for making the report available for scheduling or viewing on demand.
- Forecasted sales pattern of ice-cream using time-series with 71% efficiency.
- Familiarity with development best practices such as code reviews and unit testing.
- Experience with Version Control, ideally SVN, CVS and GIT.
- Wrapper developed in Python to run this application along with other applications.
- Ported the application from Linux to Windows.

EDUCATION

• **Master of Science in Computer Science**

Lakehead University; GPA: 3.9/4.00

Thunder Bay, ON

September 2017 – May 2019

• **Master of Engineering in Information Technology**

Gujarat Technology University; CGPA: 8.78/10

Vadodara, India

August 2012 – June 2014

• **Bachelor of Engineering in Computer Science**

Gujarat Technology University; CGPA: 7.31/10

Vadodara, India

August 2008 – July 2012

OPEN SOURCE LIBRARIES

- **Emot:** Developed and Maintainer of widely use open source "emot" python library(11k+ Users) 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.
- **Kaggle:** Open source two database "Arxiv analysis data" and "Indian's AI growth" on kaggle with 25k+ Downloads.
- **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

- Shah, Neel, Darryl L. Willick and Vijay Mago. "A framework for social media data analytics using Elasticsearch and Kibana." *Wireless Networks* (2018): 1-9.
- Neel Shah and Vijay Mago. Canadian's health analysis through social media. *IEEE Machine Learning* 2019: (submitted)