

EDUCATION

North Carolina State University, Raleigh, NC

Aug 2016 – May 2018

Master of Computer Science

GPA: 4.00 / 4.00

Coursework: Design and Analysis of Algorithms, Algorithms for Data Guided Business Intelligence, Computer Networks, Databases Management Concepts and Systems, Artificial Intelligence, Automated Learning and Data Analysis, Foundations of Data Science

Pune Institute of Computer Technology, Pune, India

Aug 2012 – Jun 2016

Bachelor of Engineering, Computer Engineering

GPA: 3.78 / 4.00

Coursework: Data Structures, Software Engineering, Data Mining, Natural Language Processing

EXPERIENCE

Backend Developer (Co-op)

Quantworks Inc., Raleigh, NC

Sept 2017 – Present

- Extending the backend functionalities of Lenovo's automation system to calculate the optimized target inventory
- Continuing the work on Smart Abstractor Assistant from last summer and adding NLP features for context based searching
- Automating the process of loading clinical data from Medical Records into REDCap database using RED-I APIs and Python

Backend Developer Intern

Quantworks Inc., Raleigh, NC

May 2017 – Aug 2017

- Built a Smart Abstractor Assistant in Python to extracting the structured data from Electronic Medical Records (EMR)
- Wrote Adobe Acrobat plug-ins in C++ for Optical Character Recognition (OCR) and RegEx for the vital information extraction
- Decreased the single PDF processing time by ~30% by automatic generation of highlights, bookmarks, and indexing
- Performed code review and bug fixing of a live, production Operations Research (OR) system for a Fortune 500 company

Machine Learning Intern

iKnowlation Research Labs, Pune

Aug 2015 – April 2016

- Led the backend team of a project to describe videos in English using Convolutional and Recurrent Neural Nets in Python
- Finetuned two of the Nvidia's CNN caffemodels (Google-Net and VGG-Net) on ImageNet dataset of 1.2M images
- Trained Long Short-Term Memory (LSTM) models on MSCOCO dataset of 120K images with Chainer framework
- Tested both the models on cross-validation dataset and improved the system's accuracy from 63% to 77%

SKILLS

Languages / Databases

C++, Python, C, R, Java, Node.js, Shell Scripting, Q, MySQL, Oracle 11g, MongoDB

Machine Learning

Scikit-Learn, Caffe, Numpy, Pandas, GNU Octave, Weka, Doc2Vec, NLTK, Apache Spark

Web Technologies

HTML5, CSS3, PHP, Bootstrap, phpMyAdmin, Python - Bottle, WordPress, REST API

Tools / Frameworks

Tableau, JDBC, Eclipse, Microsoft Visual Studio, Adobe SDK Plug-ins, IntelliJ, GitHub

PROJECTS

Kaggle Competition, Yelp Restaurant Photo Classification

(Python, Nvidia Caffe, Numpy, Pandas, Scikit-Learn)

- Implemented a three-stage transfer learning pipeline with Convolutional Neural Nets and Support Vector Machine
- Extracted features from with ~250K training images and ~1.2M testing images using BAIR Reference Caffemodel
- Built an SVM classifier and performed cross validation to get **134th** rank on Kaggle with the F1 score of **0.765**

Unsupervised Learning, Settlement Mapping

(Python, R, Numpy, Pandas, Scikit-Learn, Weka, QGIS, GNU Octave)

- Developed an unsupervised learning system to explore the settlement activities for disaster recovery
- Implemented Gaussian Mixture Models using Expectation Maximization algorithm and 1-Holt rule based classifier
- Achieved **80%** accuracy compared to with Weka's RIPPER rule-based classifier and Scikit-Learn's GMM

Supervised Learning, Sentiment Analysis

(Python, Scikit-Learn, Apache spark, Apache Kafka, Streaming API)

- Built an application using the dataset containing the real-time tweets and IMDB reviews from the internet
- Trained Logistic Regression with important words as features to get **60.3%** accuracy on Twitter and **84.8%** on IMDB
- Improved the system using Doc2Vec and Artificial Neural Nets with **63.56%** accuracy on Twitter and **83.49%** on IMDB

Database Application, Personal Health Manager

(Java, JDBC, Oracle 11g, IntelliJ)

- Developed an application to manage patient's health information, recommendations, and emergency alerts
- Used Java Database Connectivity (JDBC) to interfacing with the Oracle SQL 11g containing the tables for patients, health history, their health supporters, and diseases. Implemented triggers in Oracle for the unusual events

Web Application, Leave Manager for College Faculty

(HTML5, CSS3, PHP, MySQL, phpMyAdmin)

- Built a web application to make the college's paper-based leave management system obsolete
- Led the back-end team to create a database schema in MySQL and wrote PHP scripts to handle database queries
- Helped the front-end team in designing the leave application form widget using HTML5 and CSS3