1713 Crest Road, Raleigh, NC 27606
443-917-8052
oachary@ncsu.edu

Omkar Acharya

https://linkedin.com/in/omkaracharya
https://github.com/omkaracharya
https://omkaracharya.github.io

EDUCATION

NORTH CAROLINA STATE UNIVERSITY — Raleigh, NC
MASTER OF COMPUTER SCIENCE — Aug 2016 – Dec 2017
Coursework: Design and Analysis of Algorithms, Business Intelligence, Foundations of Data Science, Databases, Artificial Intelligence, Graph Theory, Automated Learning and Data Analysis, Computer Networks — GPA: 4.00 / 4.00

PUNE INSTITUTE OF COMPUTER TECHNOLOGY — Pune, India
BACHELOR OF ENGINEERING, COMPUTER ENGINEERING — Aug 2012 – Jun 2016
Coursework: Data Structures, Software Engineering, Data Mining, Natural Language Processing — GPA: 3.78 / 4.00

EXPERIENCE

QUANTWORKS INC. | Graduate Intern – Backend — May 2017 – Aug 2017
Technologies: C++, Python, NTLK, Pandas, REDCap database, Microsoft Visual Studio, Adobe C++ plug-ins — Raleigh, NC

- Built a Smart Abstractor Assistant (SAA) to extracting the structured data from medical records in REDCap database
- Wrote Adobe Acrobat plug-ins 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 Operation Research system for a Fortune 500 company

iKNOWLATION RESEARCH LABS PVT. LTD. | Machine Learning Intern — Aug 2015 – Apr 2016
Technologies: Python, Nvidia Caffe, Numpy, Pandas, Scikit-Learn, Python Bottle, Chainer, HTML5, CSS3 — Pune, India

- Led the backend team of a project to describe videos in English using Convolutional and Recurrent Neural Nets
- 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%**

TECHNICAL SKILLS

LANGUAGES / DATABASES	C++, C, Python, R, Java, Shell Scripting, Q, Markdown, 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 - 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 exploring 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