Ramesh Oswal

ramesho@andrew.cmu.edu **linkedin.com/in/ramesh-oswal** **rameshoswal.tumblr.com** **Phone - (412) 587-3598**

Objective – Seeking internship in the field of Data Science specializing in Machine Learning and Big Data Analytics

Education

- **Master of Science - Language Technologies Institute (School of Computer Science)** **Aug'16 - May'18 (Expected)**
  Carnegie Mellon University, PA
- **Bachelors in Computer Engineering** **Aug'10 - May'14**
  University of Pune,IN. Division: First Class with Distinction

Relevant Coursework **(*In Progress)**

- Design & Analysis of Algorithm
- Machine Learning*
- Advanced Databases
- Theory of Computation
- Coding & Algorithm Bootcamp
- Big Data in Practice*

Technical Skills

	Proficient	**Familiar**
Core Languages:	C, Core Java, Visual Basic, Python	C++, R, Javascript, SAS, MATLAB
Databases:	Oracle 9i and 10g, MySQL	MS Access
Development/Productivity Tools:	Turbo C, Informatica Siperian, Excel, Anaconda, Eclipse, PyCharm	MS Visual Studio 2010 & 14, Jupyter Notebook

Academic Projects

Project Intern **Talencea Inc, Pittsburgh** **Oct '16 - Present**

- Working with a Pittsburgh based startup founded by LTI Director Dr. Jaime Carbonell.
- 1st phase of project involves working on Big Data from different external sources like client and social media platforms and building Skill Repository.
- 2nd Phase includes building a cognitive model which matches candidates with appropriate job openings.
- Data Munging activities include Data Cleanup, Indexing, Classification, Redundancy Removal & etc.

Image classification to classify proteins into subcellular localization patterns (CMU) **Aug'16 - Dec'16**

- Built an Active Learning Framework containing Pool Based Data Access Model, Uncertainty based Querying Strategy and different base learners like SVM, Gaussian NB, KNN and Logistic Regression
- Accuracy score of 0.97 was achieved on test data using SVM as base learner.
- **Tool Used: Spyder, Python (sklearn, NumPy, matplotlib, SciPy)**

Using Probabilistic Graphical Model to forecast Stock Prices for Time Series data (CMU) **Aug'16 - Nov'16**

- Transformed into stationary time series by using log space for removing unequal variances and difference to handle trend component.
- Checked staionarity using Dickey-Fuller Test (Apple, MS, Hecla, NEM Mining, GM, Ford)
- Created precision matrix using transformed features and marginalized Precision Matrix for missing data.
- Conclusively was able to predict with minimal error rate the stock prices for Apple by using only 3 days worth of data and stock prices for companies MS, Hecla, NEM.
- **Tool Used: Spyder, Python (NumPy, SciPy)**

Linear and Forward Stagewise Regression on unknown Dataset (Carnegie Mellon University) **Aug'16 - Nov'16**

- Implemented Linear and Forward Stagewise Regression from scratch.
- Implemented Feature transformation like One Hot Encoding and Polynomial Feature transformation.
- Cross validated different train and test dataset combination to get optimized weights for each features.
- **Tools Used: Jupyter, Python (NumPy, SciPy)**

Load Balancer for OpenFlow compliant SDN architecture (Sponsored by GS Labs Pvt. Ltd) **July '13 - Jun '14**

- Aimed at enhancing s/w load-balancer in distributing traffic based on server capacity by adding generic flows.
- Based on paper "OpenFlow-based server load balancing gone wild" published in ACM Hot ICE'11 conference.
- R. Oswal et. al. "A Survey of Past, Present and Future of Software Defined Networking"
- **Tools Used - Mininet with POX controller, OpenVswitch and OpenFlow protocol**

Paper Presentation "A Cloud Framework for Parameter Sweeping Data Mining Application" **Jan '13 – Feb '13**

- Explained the system framework i.e. its architecture and execution mechanism of how parameter sweeping could be achieved in data mining application
- Finally, concluded by showing a performance evaluation w.r.t clustering & classification algorithms

Professional Experience

Business Operation Associate **ZS Associates Inc.** **Sept'14 – June'16**

- **Master Data Management - Role (Data Steward)**
  - Automated processes like loading client data and QCing client deliverable and performed Ad-hoc analysis.
  - Automation of Processes to reduce response time for file processing by over 80%
  - **Technologies Used – Python, MS Excel, VBA, Informatica Siperian, PL/SQL**
- **Smart Data Quality Management – (Quest '15)** **24hr – Hackathon (Oct'15)**
  - Participated & won in Quest'15 organized by ZS Associates which had 44 participating teams.
  - Created the architecture of product detailing communication between different modules.
  - Implemented an algorithm for "Thomson Tau Method of Outlier detection" to detect outliers
  - **Technologies Used – R, MS Excel, VBA and MS Access**

Summer Intern **Softkoash Solutions Pvt. Ltd** **May '12 – July '12**

- Implemented Microsoft's NerdDinner project as a POC
- Fixed bugs and made changes to proprietary ERP Solution used by customers in production.
- **Technologies Used – C#, Microsoft's .NET Framework, HTML, CSS and JavaScript**

Co-curricular Activities

- Won in Quest '15 (Hackathon at ZS Associates India Office)
- 2nd Best Project - PICT's "Impetus & Conceptus'14"
- 2nd Prize in College TechFest Event 'Network Raptors'
- Best Project in Operations Excellence (ZS Associate Global Offices)