ARIHANT PATAWARI

apatawari.hit@gmail.com +91-7014322581
www.linkedin.com/pub/arihant-patawari/1b/43/884

Education

University of Southern California (USC), Los Angeles, CA — GPA: 3.31/4.0
MS in Engineering Management, (Minor Data Informatics) — May 2016

Vellore Institute of Technology, Vellore, India — GPA: 3.80/4.0
Masters in Software Engineering. — May 2014
Special Achiever's award for Academic Excellence by VIT University.

Courses Taken: Information Integration on Web, Data Mining, Data Management, Database Systems, Applied Machine Learning, Natural Language Processing, Data Informatics, Project Management, Management of Engineering Teams. Software Design and Development, Software Process Management.

Tools and Technical Skills

Tools: Eclipse, WEKA, R-Studio, H2O, JIRA, Docker, Confluence, Putty, Scikit-learn, OpenRefine, Data Wrangler
Programming skills: Core Java, R, Python, Drools, Clinical Quality Language (CQL), JS, JSP, JSF, SQL, PHP

Summary

- More than 4 years experience in software design and development.
- More than 2 years of hands on experience in data mining and applications of machine learning and NLP.
- Extensive experience in developing various web based applications.
- Goal oriented and effective problem solver.
- Strong interpersonal and communication skills with an ability to lead a team.

Work Experience

Software Engineer, Motive Medical Intelligence, San Francisco, CA — **July 2016 - June 2018**

Motive Medical Intelligence is involved in improving quality of care by intelligent-evidence based systems. Worked as an integral part of team to build and enhance a Clinical Decision Support System (CDSS) as a service.

- Experienced in FHIR - Fast Healthcare Interoperability Resources, customised clinical data models to better serve system and organisational needs.
- Developed a multi-model auxiliary tool to read patient information from text file to generate data interchange structure.
- Provided support and maintenance for CDSS legacy software.
- Developed and maintained Drools rule execution engine, for analysing patient facts against clinical rules.
- Maintained an automated CQL to Drools translator for converting expressions to drools supported rules.
- Involved in designing product requirements and user stories to developing well-designed software components
- Involved in documenting all the software components and also creating UML diagrams as part of software design.
- Developed and maintained a REST based spring web application to serve CQL content followed by translation to drools and publishing of the content with backend support.
- Deployed web applications to different environments and tested them as and when required .
- Developed micro services as a part of revamping the application to create independent services.
- Coordinated across projects as well as teams to efficiently implement the data modelling changes.

NLP Applications Engineer Intern, City of Hope, Duarte, CA — **Feb 2016 - May 2016**

City of Hope is a cancer research centre. Worked for informatics division in Research Informatics Applications group and was involved in developing a semantic based search application for clinical trials.

- Analysed existing approaches used for searching clinical trials in Clinical Trials Online System (CTOL).
- Installed and analysed various existing semantics based applications like MetaMap, cTAKES and UMLS thesaurus provided by National Library of Medicine.
- Analysed performance of above tools and also utility for search application.
- Designed a process flow to receive a clinical document and extract keywords and phrases with locational annotations.
- With experimentation, designed a work flow to use pass generated keywords and phrases to cTAKES to attach semantic annotations, making it a rich resource.
- Designed a mechanism to search the key terms and synonyms using Solr.

Graduate Researcher, Institute of Neuro-Imaging and Informatics, USC **Jan 2015 - Jan 2016**

Part of the Global Alzheimer's Association Interactive Network(GAAIN). Worked on data integration and data distribution technology to improve effective data access for Alzheimer's research community.

- Reduced manual effort of data extraction from the PDF's with 80% accuracy, employing text mining and machine learning CRF based classification, leading to automated input pipeline for the GEM system.
- Improved accuracy of data integration system by 10%, by formulating a new approach for identifying and classifying composite name match pairs and leveraging into GAAIN Entity Mapper (GEM) system.
- Conceptualised and installed a cross platform working prototype for automating secure scientific work-flow analysis over globally distributed medical data, by integrating 3 technologies.
- Developed Web-Service for GEM system for Name Matching using Rest service API.

Software Engineer Intern and Volunteer, Meliorix Incorporated, San Diego, CA **Jun 2015 – Dec 2015**

Meliorix is a **h**ealth informatics based company specialised in clinical decision support service (CDSS). Worked closely in a team to identify data extraction and translation requirements. Designed a model to manage rules for data extraction.

- Converted standard CCDA XML schema's into java objects and vice versa.
- Wrote rules code for data extraction; mapped as per data specifications.
- Built an internal REST service API to perform CRUD operations for managing operations in database.

Volunteer, Institute of Neuro-Imaging and Informatics, USC **Sep 2014 - Dec 2014**

- Performed schema mapping accuracy evaluations for tools such as GEM, Harmony and COMA++, using f-measure and compared to determine best tool for medical datasets mapping based on entities extracted.
- Performed test results verification and validation for GEM system using WEKA and python scripting.

Project Experience

Integration of online content: Record Linkage of performed on three websites MedicineNet, HomeRemedies and Rxlist for compacting necessary information regarding illnesses.

Twitter Sentiment Analysis : Analysed millions of tweets to summarise moods of people across several hot topics using map-reduce framework.

Comprehensive Data Centric Cancer Patient's lifestyle analysis : Involved in extraction of tweets/posts to analyse lifestyle users parameters and to conclude preventive measures which can be effective and/or in-effective during cancer treatment.

Publications

- Patawari, A. (2015, Jul). GAAIN Virtual Appliances: Virtual Machine Technology for Scientific DataAnalysis. In Data Integration in the Life Sciences (pp. 259264). Springer International Publishing.
- Ashish, N., Patawari, A., Chhabra, S. S., & Toga, A. W. (2016). Name Similarity for Composite Element Name Matching. Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics - BCB '16. doi:10.1145/2975167.2975203
- Ashish, N. and Patawari, A. (2018). Machine Reading of Biomedical Data Dictionaries. *Journal of Data and Information Quality*, [online] 9(4), pp.1-20. Available at: https://dl.acm.org/citation.cfm?doid=3177874.
- Scott C. Neu, Naveen Ashish, Judy Pa, Arihant Patawari, Arthur W. Toga,(In process) An Analysis Framework for the Global Alzheimer's Association Interactive Network