Atif Tahir

509 Nagle St. Apt# 202, College Station, TX 77840 | (979) 985-7132 | atif.tahir@tamu.edu

OBJECTIVE -

Looking for full time opportunities in the field of data warehousing and business intelligence. Experienced in building data pipelines by extracting, processing and transforming operational data.

EDUCATION

Texas A&M University	**College Station, Texas**
Master of Science in Information Systems / Department Scholarship	GPA: 3.33	May 2017

Chhattisgarh Swami Vivekananda Technical University (CSVTU), India	**Bhilai, India**
Bachelor of Engineering in Information Technology / GPA: 3.25	May 2011

EXPERIENCE

Texas A&M Health Science Center (TAMHSC)	**College Station, TX**
Biomedical Informatics Research Assistant – Data Engineer	Jan 2016 – Present

- Engineered an ontology by leveraging the electronic health records(EHR) for testing the possibility of a clinical decision support system (DSS) enabled feedback loop for outpatient encounters.
- Implemented data warehouses with a biomedical star design by doing extract, transform and load (ETL) process, for two hospitals based on a knowledge-engineered ontology.
- Designed an unsupervised clustering algorithm using python's sklearn for vectorization of 617,177 patient visits from EHRs to enable a machine to distinguish between the episodes of care.
- Enabled clinicians to narrow down DSS' knowledge base for mapping patient problems to diagnosis codes in EHRs by implementing principal component analysis (PCA) using sklearn.

Accenture	**Pune, India**
Software Engineering Senior Analyst	Jan 2015 – June 2015

- Orchestrated a service for transporting big data files from SAP seamlessly to third-party warehouse, reducing transaction time by 10% for large files.
- Innovated file compression for big data files using Java and shell scripting. Packaged compression solution as a reusable library and was used by multiple projects throughout Accenture.

Tata Consultancy Services (TCS)	**Mumbai, India**
Systems Engineer – Decision Support Systems and Process Automation	Oct 2011 – Dec 2014

- Lead developer for Rules Management solution used by 20,000 helpdesk employees for decision-making, resulting in increase of accuracy by 25% and eliminating the need for training.
- Improved efficiency by 15% through a solution using Hibernate and web services to automate shipment process.
- Mentored and trained teams in various industry groups on implementing customized solutions by utilizing open source packages. Conducted 5 open source knowledge-training sessions for 250 employees.

ACADEMIC PROJECTS

Data Warehouse – Center of Veterinary Medicine (CVM)	March 2016 - May 2016

- Built warehouse for CVM by redacting 298,000 protected health information (PHI) records using natural language processing (NLP). Used Tableau for building reports, which addressed CVM's research needs.

Redaction utility using natural language processing (NLP)	March 2016 – July 2016

- Created a library to redact protected health information (PHI) according to safe harbor method

Hadoop Project: Resume Filter	Oct 2015- Dec 2015

- Parallel filtering in MapReduce and MongoDB, received a throughput time of 90 seconds for 10,000 resumes.

TECHNOLOGY AND PLATFORM KNOWLEDGE

Programming Languages: Java, Python

Database: Microsoft SQL Server, SSIS, SSAS, SSRS ,MySQL ,Hibernate, MongoDB

APIs and Tools: Hadoop, Camel, Tableau, sklearn , Apache Spark , Agile methodology

ACHIEVEMENTS AND INTERESTS

- TCS "Classroom Faculty Award" for conducting training on open source BPM packages in TCS
- Rotary Club at Texas A&M - "Rotaractor of the Month" award for committed service in College station

LinkedIn: https://www.linkedin.com/in/tahiratif   Github: http://www.github.com/mars137