

# Arun kumar Jegarkal

arunjeg@gmail.com

(806) 620-8017

<https://arunjegarkal.github.io/arunjegarkal/>

---

**Summary:** Developer with 4+ years of extensive experience in Java, Python, Amazon Web Services, Microsoft Azure, Power BI, Tableau, Data Analytics and ETL technologies; worked for various clients and projects in every phase of Software Development Life Cycle, from business requirement gathering to project delivery. Excellent design and integration problem solving skills.

## Professional Experience

---

### Data Engineer, Peterbilt Motors Company (Paccar Inc.), Denton, TX (2019 Nov - Present)

#### Responsibilities

- As a Data Engineer I focus on making possible fast, accurate, and reliable access to data.
- Build data pipelines, manage a data warehouse, and support the production use of our data.
- I interact with business users and data scientists to gather the requirements and develop the code pipeline using AWS services based on the need and deploy into production environment.
- Also, I will involve in data migration process to move data from OnPrem Data sources to cloud platform.
- I also build a continuous integration and continuous development pipelines for machine learning models.
- Use the knowledge of Agile / Scrum methodology as a development process.

### Business Systems Analyst, Netrush, Vancouver, WA (2018 Oct - 2019 Nov)

#### Primary Duties

- Developing Azure ETL Pipelines to extract data from Azure storage and transform into dimensional data and load into analytical database.
- Scheduling ETL pipelines using Azure Automation.
- Developing Power BI reports and dashboards to transfer insights in data to visuals.
- Writing Runbook scripts in Azure Portal to Auto scale Power BI premium Tier and trigger reports dataset refresh.
- Dimensional Data Modeling using star schema. And writing SQL Queries and Procedures to pull reports.
- Peer review reports and dashboards before publishing to production.
- Interacting with Stakeholders and End users to collect User Requirements.
- Translating User Stories into development features.
- Testing the Netrush internal developed tools, Reports and Dashboards.
- Writing release, user documents as a guidance of the tools for End Users and Stakeholders.
- Experience in agile methodology.

### Graduate Assistant at Athletics Department, Texas Tech University, Lubbock, TX (2016 Dec - 2018 May)

#### Responsibilities

- I developed tableau dashboards and stories to get insights and trends of expenditures and athletes' performance in each game and in their academics.
- I created charts, reports and motion charts using tableau.
- Deployed dashboards to the Tableau public and used tableau bridge to keep the data UpToDate at the server side.
- Scheduled Tableau server to send reports via email.
- Involved in writing SQL queries, PL/SQL procedures and functions to get data from various data sources.
- Also performed logical and physical Data Modeling using normalization Techniques.
- Performed pivoting, clustering, created calculated and parameter fields to model the data according to the business needs.

### Software Engineer S2-Data Warehouse Analyst at EVRY India Pvt Ltd, Bangalore, India (2014 Aug - 2016 Aug)

#### Primary Duties

- I was responsible for the Extraction Transformation and Loading (ETL) data from multiple sources into Data warehouse.
- Extensively worked on designer to designing and developing mappings from varied transformation logic using Expression, Source Qualifier, Filter, Router, Joiner, Lookup, Sorter, Aggregator and Update Strategy etc.
- Designed ETL logic for Dimension Tables using SCD - Type1 and Type2.
- Worked with heterogeneous data sources like DB2, Flat Files, and Mainframe Files.
- Experienced in identifying and solving performance bottlenecks in various levels like source, target, mappings and session levels.
- Experienced in SQL and Informatica Cloud services and working experience on PL/SQL and UNIX commands.

## Projects

---

- **Netrush Portal** - This is a Netrush Analytical and reporting project, focused on developing reports and dashboards to get insights of sales, inventory, compliance, violations, pricing etc.,

- **Netrush Platform** – This is a Netrush inhouse retail domain project, focused on developing ecommerce platform. I am responsible for gathering user requirement and translating into development features. Also involve in testing tools.
- **National Eligibility Database (NEDB Carrier Loads)** – This is an Insurance domain project for the client HMS. The goal is to use proprietary logics to avoid cost from the government payers and recovery from the commercial or third-party insurance vendors. For which we received files on timely from all the vendors and performed ETL operations to uniform the data collected, before applying match processing.
- **Geospatial Visualization of Twitter Data** – This is a National Science Foundation funded research, it is focused on building a model for researchers to analyze trends in Twitter using machine learning algorithms for detecting the topics discussed and obtaining people attitude towards the discussions. Moreover, to find co-occurring topics and its evolution over a period.
- **A High Performance, Scalable Event-Bus System for Distributed Data Synchronization** – Our primary goal is to provide the support for fully transactional communication between the clients and the event bus thereby making the communication more effective. Achieved in Java technology using Web sockets, multithreading and Avro.
- **Movie Scope** – Analyzed the movie database from IMDB database, to compare the box office collection towards the user review for a movie, for which machine learning algorithms used to calculate the sentiment of the review and used created tableau dashboard to visualize the trends and insights based on releasing years, genres and movie name.

## Technical Skills

---

<b>Programming languages</b>	:	Java, Python, C, SQL, PL/SQL, Shell scripting, Data Analytic Expressions (DAX)
<b>Web Designing Skills</b>	:	HTML5, HTML, CSS, JavaScript, PHP, XML, JSON
<b>Database</b>	:	Oracle 9i, SQL Server
<b>Cloud Experience</b>	:	<b>Microsoft Azure</b> : Data Factory, Azure Storage, Azure Automation <b>Amazon Web Services</b> : S3, Lambda, SQS, SES, SNS, EC2, Step function. Snowflake
<b>BI Tools</b>	:	Power BI Pro, Power BI Desktop, Tableau, Tableau Desktop, Tableau Bridge, Tableau Server and Tableau Public, Google Analytics, Looker, QuickBase, D3.js
<b>Tools &amp; Utilities</b>	:	Informatica PowerCenter Client 9.5.1, Informatica Data Explorer (IDE), Informatica Cloud, SQL Developer and Toad, Power Exchange, Eclipse, Net Beans.

## Education

<b><i>Master of Science in Computer Science</i></b> Texas Tech University, Lubbock, TX	<b>Graduation: May 2018</b>
<b><i>Bachelor of Engineering in Computer Science</i></b> Visvesvaraya Technological University, Belgaum, India	<b>Graduation: June 2014</b>
<b><i>Diploma in Computer Science</i></b> R N Shetty Polytechnic, Belgaum, India	<b>Graduation: June 2011</b>

## Honors, Awards and Achievements

---

- Geospatial Visualization of Twitter Data - Poster presentation at national workshop for NSF grant support.
- Awarded competitive scholarship: 75% fee waiver by Texas Tech University.
- WebMaster at Texas Tech India student Association 2017-18.