

ADITYA RAJKAMAL GUDISEVA

**Contact No:** +1 (469) 258-2406 E-Mail: aditya.rajkamal@gmail.com

**Professional Summary:**

***Data Engineer*** with Hands-on and up to date skills on the state-of-the art automation technologies in Data Management and integration including ETL, DataOps, Cloud, Warehousing and Big Data. Strong background in Development Methodologies, Solution Delivery and mission critical Enterprise Applications Operational Support. Deep technical expertise and Wide business exposure to understand and solve business problems with technology.

***Data Ingestion & Processing***: Proven background of Designing, Developing and Deploying End to End Data Pipelines both on-premises and in cloud environments. Experienced in handling high volume and high velocity loads using parallel processing tools & methodologies. Expertise in designing BI Solutions using diverse ETL tools and databases.

***Reporting and Analytics***: Experienced in enabling down-stream reporting and analytics applications both in batch and real-time scenarios. Designed Data models such that the proper load balancing is ensured between the Data and BI layers.

***Solution Design and Delivery*:** Experience and capability to deal with complex and large environments across various business verticals. Assisted teams in designing highly scalable, highly concurrent, and low latency data applications. Ability to adapt cutting edge technologies to enterprise requirements including high availability and disaster recovery. Experience working with and evaluating open source technologies and demonstrated ability to make objective choices.

**SKILLS SUMMARY**: -

* Having 12+ years of Software Development Experience in Health Care Domain.
* Experienced in all phases of software development life cycle, Requirement gathering, Planning, Designing, Development and testing.
* Expertise in Azure Services – Azure Data Factory, Data Lake, Azure Functions, Azure Synapse,Azure Databricks (SPARK), Hadoop, Data Warehousing, Linux, and Scala/Python.
* Experience in Developing Spark jobs using Scala for faster real-time analytics and used Spark SQL for querying.
* Strong knowledge of core Spark components including - RDDs, Data frame and Dataset APIs, Data Streaming, in memory capabilities, DAG scheduling, data partitioning and tuning
* Performed various optimizations like using distributed cache for small datasets, partition and bucketing in hive
* Expertise in Developing Spark application using PySpark and Spark Streaming API's in Python, deploying in yarn cluster in client, cluster mode. Used Spark Data frame APIs to ingest data from HDFS to Azure Blob.
* Experience developing iterative Algorithms using Spark Streaming in Scala and Python to builds near real-time dashboards.
* I also have experience with CICD (continuous integrating and continuous deployment) pipelines using Jenkins.
* Experience with Data flow diagrams, Data dictionary, Database normalization theory techniques, Entity relation modeling and design techniques.
* Experience with the Azure logic apps with different triggers and worked on ISE environment in logic apps.
* Expert in Data Management, Transformation & Quality Services.
* Extensive experience in Data Modeling, Design, Development and ETL Development.
* Experienced in Data Warehousing, and Business Intelligence (BI) applications.
* Experienced in all phases of the Data warehouse project life cycle including Requirements Analysis, Dimensional Modeling, ETL and Business Intelligence.
* Highly adept at Information Gathering, Requirement Analysis, ETL Design, Development, and deployments & on time Delivery.
* Have good experience in implementing data pipeline solutions using Azure, Snowflake and Python.
* Experience in managing and securing the Custom AMI's, AWS account access using IAM.
* Experience in handling large volumes of storage in AWS using Elastic Block Storage, S3.
* Hands of experience in GCP, Big Query, GCS bucket, G - cloud function, cloud dataflow, Pub/sub cloud shell, GSUTIL, BQ command line utilities, Data Proc, Stack driver.
* Experience in SQL and PL/SQL programming.
* Experience on Hadoop Ecosystem like HDFS, Hive, Sqoop, Spark and Python.
* Experience in cloud Datawarehouse snowflake.
* Experienced in creation of reports using Power BI.
* Working knowledge of Agile development methodologies.
* Experienced to write queries in Splunk and creation of Dashboards in Splunk.
* Experienced to configuration of Jenkins in the Windows servers.
* Experienced to creation of Jobs to run the PowerShell scripts in Jenkins.
* Good Experience in Waterfall and Agile (Scrum) software development methodologies.
* Develops software that is solid, extensible, reliable, well-documented and user-friendly.
* Excellent Technical and Analytical skills and ability to work independently with minimal supervision and can also perform as part of a team.
* Strongly initiative, creative and able to learn new technologies and systems.
* Self-confidence, total commitment, patience and positive attitude towards situations in life.
* Strong troubleshooting and analytic skills.
* Has strong ability to learn new technologies in short span and implement independently.

**Certifications:**

DP – 203 [**Microsoft Azure Data Engineer Associate**]

**Academic Profile:**

Bachelor of Technology (**B.Tech)** in **Computer Science & Engineering**

**Technical Skills:**

Databases : SQL Server 2008R2 /2012/2014, Oracle 9i/10g, MySQL, MS Access , MongoDB, Azure Cosmos DB, Azure Synapse

Data Warehousing : Teradata, Snowflake, Kimball, Data Vault

BIG Data : Spark, Hive

Languages : Python, Scala, T-SQL, PL/SQL

Cloud : Microsoft Azure, SQL Azure, AWS, EC2, Redshift, S3, RDS, EMR

Modelling : Kimball, Inman, Hybrid

ETL/Data-Integration: SSIS, QLIKVIEW, INFORMATICA

Reporting Tools : SSRS, D3, Dundas, Fusion Charts, Power BI

Scripting : Shell Scripting

Tools & Utilities : Microsoft visual Studio, VSS, TFS, SVN, ACCUREV, Eclipse, Toad

DevOps : GitHub, Azure Git

**1. PROJECT PROFILE:**

**Project**  : **BITS 2.0**

**Client** : AMERI HEALTH CARITAS, USA

**Period** : June 2022 to Till Date

**Environment** : Azure Cloud, Azure Data Factory, Data Lake and Azure Databricks, Azure SQL Database, Logic Apps, Power BI, Python, Azure Functions, Azure Synapse.

**Role** : DATA ENGINEER LEAD

**Description:**

BITS Project is to migrate Hadoop system to Azure cloud, it has existing dashboards where all needs to be migrated to Azure cloud environment.

Product stack for this assignment is ADF, ADLS, Databricks, Python and Spark SQL.

**Roles & Responsibilities:**

#### 

* Analyzing requirements of SAS code for each Dataset and identified the sources required for dataset.
* Conversion of SAS code manually to spark SQL based on the requirement.
* Reading the data from ADLS through Databricks notebook and performed required transformations as per the SAS code.
* Write back the transformed data to dedicated sqlpool.
* Orchestrated data extraction pipeline in ADF to get the data from ADLS to dedicated sqlpool.
* Creating pipelines, data flows and complex data transformations and manipulations using Azure Data Factory (ADF) and PySpark with Databricks.
* Requirements gathering and followed scrum methodology.
* Load the data into Azure synapse using pipeline.
* Designed the workflow engine for the entire audit lifecycle.
* Importing the data using ADF pipelines.
* Updated the existing Azure functions and re-deploy.
* Creating Logic Apps with different triggers, connectors for integrating the data from workday to different destinations.
* Explore on Integrated Service Environment (ISE) of the Logic apps with storage account integration.
* Learnt and implemented new technologies as and when needed.
* Grooming the team members on any technical clarifications.
* Extensively worked on Power Query to extract data from various sources such as Relational (SQL Server), Flat Files (Excel, CSV), SharePoint Files/Folders etc.
* Created and used Date Dimension, DAX tables, Columns and measures in Power BI.
* Extensively used Disconnected Tables to switch Measures/Dimensions in visuals using Slicers and Bookmarks.
* Create spark data frames to read the data from source files
* Created widgets in Databricks to dynamic files and dynamic tables.
* Loaded Data Azure Synopsis manage Database.
* Implemented ADF pipelines for data-driven workflows in cloud for orchestrating and automating data movement and data transformation.
* Implemented pipelines according to business rules by using different ADF activities like IF Condition, For Each, Lookup, Get Metadata, Set Variable, Filter and Switch etc.
* Extract Transform and load data from sources systems to Azure Data Storage services using a combination of ADF and Data Lake Analytics.
* Data ingestion to one or more Azure services (Azure Data Lake, Azure Storage, Azure SQL) and processing the data in Azure Databricks.
* Develop Spark applications using pyspark and spark SQL for data extraction, transformation and aggregation from multiple file formats for analyzing and transforming the data uncover insight into the customer usage patterns.
* Responsible for estimating the cluster size, monitoring and troubleshooting of the spark data bricks cluster.
* Ability to apply the spark Data Frame API to complete Data manipulation within spark session.
* Implemented PySpark data bricks notebooks to ingest and transform the data from source(raw) to target(output) container by applying business logics.

**2. PROJECT PROFILE:**

**Project**  : **ClaimShark**

**Client** : iSpace, USA

**Period** : June 2020 to June 2022

**Environment** : Azure Cloud, Azure Data Factory, Data Lake and Azure Databricks, Azure SQL Database, Logic Apps, Power BI, Python

**Role** : SENIOR DATA ENGINEER

**Description:**

ClaimShark tracks the entire lifecycle of a claim – starting from the time the claim file is sent to the vendor (internal audit team and external vendors), to when a claim is tagged for savings - as it goes through the validation process, and later through the collection effort, receipt of funds, and appeals. While tracking the lifecycle, it provides robust reporting on all the activities of the audit teams, including concepts leading to savings, vendors tagging the claims, and the recovery. It provides real-time information on savings and helps with managing payment integrity departments.

**Roles & Responsibilities:**

#### 

* Requirements gathering and followed scrum methodology.
* Designed the workflow engine for the entire audit lifecycle.
* Importing the data using ADF pipelines.
* Creating Logic Apps with different triggers, connectors for integrating the data from workday to different destinations.
* Explore on Integrated Service Environment (ISE) of the Logic apps with storage account integration.
* Learnt and implemented new technologies as and when needed.
* Grooming the team members on any technical clarifications.
* Extensively worked on Power Query to extract data from various sources such as Relational (SQL Server), Flat Files (Excel, CSV), SharePoint Files/Folders etc.
* Created and used Date Dimension, DAX tables, Columns and measures in Power BI.
* Extensively used Disconnected Tables to switch Measures/Dimensions in visuals using Slicers and Bookmarks.
* Create spark data frames to read the data from source files
* Created widgets in Databricks to dynamic files and dynamic tables.
* Loaded Data Azure Synopsis manage Database.
* Implemented ADF pipelines for data-driven workflows in cloud for orchestrating and automating data movement and data transformation.
* Implemented pipelines according to business rules by using different ADF activities like IF Condition, For Each, Lookup, Get Metadata, Set Variable, Filter and Switch etc.
* Extract Transform and load data from sources systems to Azure Data Storage services using a combination of ADF and Data Lake Analytics.
* Data ingestion to one or more Azure services (Azure Data Lake, Azure Storage, Azure SQL) and processing the data in Azure Databricks.
* Implemented PySpark data bricks notebooks to ingest and transform the data from source(raw) to target(output) container by applying business logics.

**3. PROJECT PROFILE:**

**Project**  : **QANI Archive Report**

**Client** : CaseNet, USA

**Period** : Apr 2018 to May 2020

**Environment** : Azure, ADF, Data Bricks, Logic Apps, Power BI, Python,HTML5, Ajax, SQL Server, Powershell

**Role** : DATA ENGINEER

**Description:**

#### This application is used to create the Archive report to members in system.

#### The application is basically will create .json files and PDF files of a member.

#### Json files will display the member data in HTML5 application. The PDF files

Contains member authorizations, Assessments, Notes, Care plans and cases

Summary.

**Roles & Responsibilities:**

#### 

* Performance Management - Monitor and tune system to achieve optimum performance levels, and provide reports to support metrics collection conducts code reviews. Provides guidance to developers for continuous improvement.
* Lead Technical Development team with development tasks, follow-up and monitoring.
* Extract Transform and load data from sources systems to Azure Data Storage services using a combination of ADF and Data Lake Analytics.
* Data ingestion to one or more Azure services (Azure Data Lake, Azure Storage, Azure SQL) and processing the data in Azure Databricks.
* Creating Logic Apps with different triggers, connectors for integrating the data from workday to different destinations.
* Develop Spark applications using pyspark and spark SQL for data extraction, transformation and aggregation from multiple file formats for analyzing and transforming the data uncover insight into the customer usage patterns.
* Responsible for estimating the cluster size, monitoring and troubleshooting of the spark data bricks cluster.
* Ability to apply the spark Data Frame API to complete Data manipulation within spark session.
* Clearly communicate task estimates, ETAs and work breakdown structure to the management and client side.
* Ensure development artifacts are in alignment with patterns/ frameworks designed by software engineering teams
* Full lifecycle application development using standard frameworks and coding standards/patterns.
* Created Power BI model by extracting and transformed data using Power Query. Defined relationships and created custom measures, tables and columns using DAX.
* Created Drill through reports, Drill down reports, sub-reports, cascading parameters, linked reports and cross tab reports.
* Swapped visuals in Power BI by extensively using Bookmarks to provide app like experience to users.
* Lead, coach, task manage, technical plan, and mentor team members and more junior engineers.
* Automates testing processes by developing scripts as part of coding.
* Ensuring smooth delivery of the product.
* Understanding the business requirement for database and application design and development.

**4. PROJECT PROFILE:**

**Project**  : **App Management Automation**

**Client** : CaseNet, USA

**Period** : Oct 2017 to March 2018

**Environment** : Azure Cloud, Azure Data Factory, Data Lake and Azure Databricks, Azure SQL Database, Logic App, Power BI

**Role** : DATA ENGINEER

**Description:**

#### This application is used to take the backup of servers, moving the IMEX rules from

#### Golden Configuration server to Dev, Train, Stage and Production, moving the BRS

#### rules from Golden Configuration server to Dev, Train, Stage and Production,

#### moving the reports from one server to another server and auto start windows

#### services if they are stopped.

**Roles & Responsibilities:**

#### 

* Ensuring that development tasks are completed within the timeline provided and that issues are fully tested with minimal defects.
* Create spark data frames to read the data from source files
* Created widgets in Databricks to dynamic files and dynamic tables.
* Loaded Data Azure Synopsis manage Database.
* Implemented ADF pipelines for data-driven workflows in cloud for orchestrating and automating data movement and data transformation.
* Implemented pipelines according to business rules by using different ADF activities like IF Condition, For Each, Lookup, Get Metadata, Set Variable, Filter and Switch etc.
* Implemented PySpark data bricks notebooks to ingest and transform the data from source(raw) to target(output) container by applying business logics.
* Created hive tables with partitions on target Parquet and Delta file format folders.
* Implemented code reusability technique's by creating separate notebooks for variable initialization and Pyspark function, calling them in each dimension and fact table notebooks execution.
* Created Azure Logic App resource to post the request using web activity in ADF to send automated emails to business/project resources to acknowledge Pipeline Success/Failures
* Created DB scripts which suits for both Oracle & MSSQL database as our product is supporting on both databases.
* Created Transformations & Jobs using Pentaho ETL tool.
* Implemented object level restart ability for Initial & Refresh Data to overcome the manual intervention between loads.
* Implemented batch process mechanism to load Initial & Refresh loads.
* Prepared data extraction scripts required for ML project.
* Created data pipelines and cleansed data using python pandas and Spacy libraries.
* Used Spacy module to generate named entities from one of the data frame columns.

**5. PROJECT PROFILE:**

**Project**  : **Agile Care and Case Management**

**Client** : CaseNet, USA

**Period** : Mar 2015 to Sept 2017

**Environment** : iReports/Jasper reports, Python, Java, spring, Sql Server 2008

**Server** : JBoss

**Role** : SOFTWARE ENGINEER

**Description:**

CaseNet applications are client-server software platforms supporting the entire spectrum of care management. The software is tailor able to rapidly fit to any care management workflow of the industry. The system has even been designed in a very generic way to accommodate future moves into different industries. It has modules like Case Management, Disease Management, Utilization Management, Sponsors Management, Appointments, Third party integration, Authorizations and Referrals.

**Roles & Responsibilities:**

* Maintaining, tuning and repairing applications to keep them performing according to technical and functional specifications providing user support.
* Working effectively with the Test & Verification team to resolve issues.
* Hands on JAVA, Python and Reports development including design, development, testing production rollout.
* Being involved in Development, Continuous Integration and time-boxed iterative development
* Performing code reviews and reviews of technical documentation
* Ensuring smooth delivery of the product.
* Understanding the business requirement for database and application design and development.
* Estimation, design and development of various modules.
* Involved in unit testing of the code.
* Involved in probing the application for memory leaking.

**6. PROJECT PROFILE:**

**Project**  : **TC-XML Generator**

**Client** : CaseNet, USA

**Period** : Jan 2014 to Feb 2015

**Software** : Java, Spring MVC, Hibernate, XML, JQuery

**Server** : Tomcat 6/JBoss 5.0

**Role** : Application Developer

**Description:**

TC-XML Generator is a web-based application. Can be used to provide an alternate out of box solution to create and run data feeds of different kind ex: member, provider, product etc. Otherwise user has to login to TC application for all these things. This application will allow to do operations in a bulk manner and is portable can be deployed against different TC versions. By changing, the properties file only.

Application will take different xsd -schemas as input related to different data-feeds. All the entities created using this application will behave all most similar way that they are created using TC application

**Roles & Responsibilities:**

* Analyzed business needs and documented functional and technical specifications based upon user requirements with extensive interactions with business users.
* Requirement gathering, Analysis, Effort Estimation and design, Prepare documentation on all aspects of ETL processes.
* Working effectively with the Test & Verification team to resolve issues.
* Hands on JAVA development including design, development, testing production rollout.
* Used the Spring Framework for data transfers between Business layer and the Data Layer.
* Involved in deployment of application in different TC servers.
* Used Hibernate for database interactions.
* Ensuring smooth delivery of the product.

**7. PROJECT PROFILE:**

**Project**  : **Authorization management**

**Client** : CaseNet, USA

**Duration** : May 2012 to Nov 2013

**Software** : Java, Spring MVC, XML, JQuery

**Server** : Tomcat /JBoss

**Role**  : Application Developer

**Description:**

AM is designed with two major modules named Auth loader and Auth extractor. Auth loader deals with creation of bulk authorizations into TC, we are working with xml format. The Auth Extract Utility will support ONLY Inpatient (IP) and Service/Procedure (SP) authorizations. The following defines service lines that will be extracted for each auth that has been changed (determined/discharged)

i. Determined lines only.

ii. Only lines that have been marked "sent to claims".

**Roles & Responsibilities:**

* Being involved in the requirement analysis, design, and development as well as in the testing of this module.
* Being involved in development of both client and server-side programming.
* Used the Spring Framework for data transfers between Business layer and the Data Layer.
* Used Hibernate for database interactions.
* Worked with design patterns like Business delegate, service locator, Data Access Object (DAO), Factory Patterns.
* Involved in design and coding and unit testing.
* Involved in deployment of application in different TC servers.
* Involved in probing the application for memory leaking.

**8. PROJECT PROFILE:**

**Project**  : **Real Time Resource Allocation System**

**Client** : INTERNAL

**Duration** : Jun 2010 to Apr 2012

**Software** : Java, XML, JQuery, HTML

**Server** : Tomcat /JBoss

**Role**  : Jr.Developer

**Description:**

Resource Allocation System (RAS) is an Enterprise Resource Portal, designed, developed and used to keep track of resources within the organization. This application manages resource allocation to projects. It consists of a welcome screen, which displays the status of all the resources available for deployment and the project list displays client wise information. RMS application provides a search functionality, which can be used for searching a resource or project by id or name. Reports generated by the application can be viewed in html.

**Roles & Responsibilities:**

* Being involved in the requirement analysis, design, and development as well as in the testing of this module.
* Involved in design and coding and unit testing.
* Have created objects, custom fields, tabs, different Page layouts based on the Record Types.
* Developed SQL Queries for Insert & update the Fields of Custom Objects