**Smit Ambalia**  +**91 6351387041 smit.ambalia@gmail.com**

**SUMMARY**

* Experienced IT Professional in Data Analytics.
* Oracle Database Administration Ver. 12C/11G on a variety of OS like LINUX, Windows, Installation, Configuration, Upgrades from 11g to 12c,10g to 12c, Space Management, System maintenance, patching, Sql\*Plus, documenting daily practices and troubleshooting steps.
* Worked on developing MapReduce jobs in java, configured, installed, and used eco system components like Hadoop, MapReduce, HiveQL, Pig, Sqoop in several Big Data projects.
* Installed Oracle Database Server on WINDOW and migration of oracle database server from 11g to 12c and applied oracle patches.

**Machine Learning Engineering**

**Applied AI Course, Online Dec 2018 -Current**

Trained in Machine Learning (**ML**) using Python, and Deep Learning with TensorFlow and Keras.

* Developed K-NN ML program, used W2Vec,BoW,TFIDF to convert text to vector data with tuning hyperparameter (K) to predict target label.
* Worked on Classification problem using Logistic Regression analysis on datasets, and plotted AUC graph on train and cross validation data, and tuned hyperparameter to prevent model to overfit.
* Developed classification algorithms like Naïve Bayes, Support Vector Machine and Decision tree, Random Forest, XGBoost and obtained desired AUC score in python.
* Worked on different Multiple Layer Perceptron models on MNIST datasets in Keras with TensorFlow as backend.
* Developed a model that predicts whether an order for a given product can go on backorder or not, exposed APIs using Flask which can predict backorder of the product for a bulk and a test data.
* Developed CNN models on MNIST datasets to predict handwriting images correctly, and obtained batter accuracy less.
* Developed DenseNet model on CIFRA-10 datasets to classify 10 classes, used image augmentation, and achieved 90 up accuracy on train and test datasets.
* Deployed Sentiment Analysis Prediction on AWS Cloud with EC2 instance.

**Environment:** GIT, Python, NumPy, Pandas, Flask (APIs), Visual Studio Code (IDE)

**Software Engineer Intern**

**Improvised Tech Solutions, Rajkot, Gujarat Nov 2019 - April 2020**

* Hands-on experience with relational databases and NoSQL databases (MongoDB, MySQL).
* Learned best practices of version control tool **GIT**.
* Developed APIs in Node.js that interact with backend database (MongoDB).
* Developed Front-end with Nuxt.js that interacts with backend APIs
* Knowledge in unit and integration testing with Mocha.
* Used best development practices for quality code.

**Environment:** GIT, Node.js, Vue.js, Nuxt.js, MongoDB, MySQL.

**Hadoop Administrator**

**GRASS Solutions Feb 2018 - Nov 2018**

Trained in Hadoop Administration.

* Manageinstall up to 4 nodes Hadoop cluster, added and removed nodes from the cluster, configured the NameNode high availability.
* Performance tuning of Hadoop clusters.
* Installed Hadoop echo systems, for example, HBase, MongoDB, Cassandra, Pig, and Hive, Sqoop, Zookeeper, Oozie.
* Learned python basics packages like NumPy, Pandas, scikit-learn.
* Worked on Data ingestion mini-project from twitter to HDFC using Flume and Sqoop.

**Hadoop Developer**

**Edureka Training Center Jan 2017 – July 2017**

Trained in Big Data Hadoop Ecosystem which includes **Sqoop, Hive, Pig Latin** (Data Flow Languages) and Custom **MapReduce** programs in java.

* Worked on Hive data warehousing tool for creating tables, and optimizing Hive Query.
* Developed MapReduce jobs in java on unstructured files for YouTube data analysis project to YouTube video trending predictions, removal and detection of adult content in Linux Environment.
* Worked on Airlines Analysis project by injecting unstructured data into HDFS, created tables in HIVE, load that data from HDFS to perform query in HIVE for Airline data analysis which includes flight delays, fuel consumption.
* Worked on StackOverflow data analysis project, used Pig Latin ecosystem for data analysis.
* Worked on Movie Data Analytics project, loaded unstructured data into HIVE tables, analyzing the data by performing Hive queries.
* Used Sqoop to load data into HDFS from MySQL and extracted the data from MySQL into HDFS.

**Environment:** MapReduce, HDFS, Sqoop, Hive, Pig Latin**, CentOS, Hadoop, YARN**.

**PROFESSIONAL EXPERIENCE**

**Oracle DBA**

**UM-KECC, Ann Arbor, MI Feb 2016 – Sep 2016**

The **University of Michigan Kidney Epidemiology and Cost Center (UM-KECC)** is an interdisciplinary research group drawing from the Departments of Biostatistics, Health Management and Policy, Surgery and Nephrology.

As an Oracle DBA, I was responsible for Installing, Configuring, managing 10g,11g,12c database, Configuring OEM, migrating 10g and 11g Oracle databases to 12c database in Production environment, backup and recovery, performance tuning, training junior Oracle DBAs.

* Installation and creation of 10g, 11g, and 12c Databases for Development, Quality and Production environments and upgraded from 10g to 11g, 11g to 12c.
* Performance Monitoring and capacity planning, Security Compliance (FISMA), audit trail monitoring.
* Compile and validate data; reinforce and maintain compliance with corporate standards.
* Wrote and fixed numerous SQL queries, varying from simple to complex, multi-page queries.
* Defined, design and configured a security policy to support FISMA compliance which gets adopted by all the departments in the university.
* Configured RADIUS for two-factor Authentication in Windows and Linux environment
* Created 12.1.0.2 databases for Testing, and Production on Linux and Windows.
* Used to export and import of different databases using Oracle Exp / Imp and data pump utilities
* Performed all daily Maintenance tasks like monitoring job failures, scheduling backup jobs, and troubleshooting RMAN backup job failures and interacting with Oracle Team when required.
* Installed and configured Oracle 12c database.
* Design and implement database backup and recovery procedures using RMAN
* Having experience in Migrating data to other databases and Database storage management
* Monitoring database performance and database operation from OEM
* Support developer by tuning SQL query and PL/SQL codes
* Created table spaces and managed table spaces adding/deleting/ resizing data files
* Performed logical backup & recovery using export/import utilities and Datapump
* Hands on experience in Performance tuning of the database - SQL Tuning, Used Tuning utilities like EXPLAIN PLAN, AWR, ADDM and Tuning of SGA, Distribution of disk I/O, Sizing of tables and indexes.
* Performed SQL and PL/SQL tuning and Application tuning using various tools like AWR, SQL\*TRACE, TKPROF and AUTOTRACE.
* Hands on experience in documenting Database management practices.

**Environment**: Oracle 10g/11g/12c, RMAN, DBCA, DBUA, Export/Import (EXP/IMP), Data Pump, AWR.

**TECHNICAL SKILLS**

**Big Data:** Hadoop, MapReduce, HDFS, Hive, Pig and Sqoop

**Database/Platform**: Oracle 11g/12c, MySQL, MongDB

**Script languages:** Python, TensorFlow, Keras

**Oracle DB:** Database up gradation 11g/12C, Disaster Recovery Solutions, Recovery Manager (**RMAN**), Oracle Enterprise Manager (**OEM),** Performance Tuning (DB).

**Environment:** CentOS Linux, Manjaro Linux

**EDUCATION**

* Master in Computer and Information Science, Cleveland State University, USA
* Bachelor of Technology, Computer Engineering, Atmiya Institute of Technology, GTU India
* Diploma in Computer Engineering, AVPTI Rajkot, Technical Examination Board , GTU India

**CERTIFICATION**

* Accessing Hadoop Data using Hive at Big Data University
* MapReduce and Yarn Certificate at Big Data University
* Hadoop 101 Certificate at Big Data University
* Big Data and Hadoop Developer Certificate at Edureka Training center
* Training in Oracle Certified Professional (OCP) in 10g ,Oracle Certified Association (OCA) in 10g, Oracle SQL Expert Certified in 10g