

# Venkata Siva Sandeep Yedla

(425) 551-9040 | [yxv180001@utdallas.edu](mailto:yxv180001@utdallas.edu) | <https://www.linkedin.com/in/sandeep-yedla/>

## EDUCATION:

The University of Texas at Dallas; *MS in Information Technology and Management*

Dec 2020

Amrita Vishwa Vidyapeetham University, India; *BS in Electronics and Communication Engineering*

June 2014

**COURSE WORK:** Data Structures & Algorithms, Cloud Computing, Distributed Systems, Database Management, Machine learning, Big Data, Object-Oriented Programming(Python), Programming for Data Science, Agile Project Management, Statistics & Data Analysis.

## CERTIFICATIONS:

• **AWS Certified Solutions Architect – Associate(SAA C02)**

Jan 2021

---

## TECHNICAL SKILLS:

**Databases (SQL & NoSQL):** Oracle, MySQL, MS SQL Server, PostgreSQL, DynamoDB, MongoDB, HBase, Amazon Redshift.

**ETL/Business Intelligence:** Informatica PowerCenter (ETL), Tableau, PowerBI, AWS QuickSight.

**Analytical Techniques:** Data Warehousing, Dimensional Data Modeling, ETL design, Schema Design, Data Mining.

**Big Data/Distributed Systems:** Hadoop, HDFS, Spark, Airflow, Hive, Kafka, Sqoop, Oozie, NiFi.

**Cloud Applications:** AWS Glue, Kinesis, Lambda, AWS RDS, Athena, IAM, EC2, S3, SQS, SNS.

**Programming/Other Tools:** C, SQL, Python, Scala, Linux (Shell Scripting), JSON, XML, Git, JIRA, REST API, Jenkins, Docker, Kubernetes.

---

## PROFESSIONAL EXPERIENCE:

**CBL Solutions, Data Engineer Intern** - Irving, TX (United States)

Aug 2020 - Dec 2020

- Designed and implemented a real-time scalable data pipeline to process semi-structured data by integrating 390 million raw records from 25+ data sources using Kafka and PySpark and store processed data in AWS Redshift.
- Utilized Spark Streaming to consolidate and clean transactional and event data resulting in speed improvements of 20% in the production.

**Neiman Marcus, Software Engineer Intern** - Irving, TX (United States)

Jan 2020 - June 2020

- Developed REST API's for segmentation, personalized recommendations and omnichannel messaging that reduced cart- abandonment rate by 22%, leading to a \$0.4 M increase in online sales.
- Optimized an effective CI/CD pipeline for MuleSoft applications resulting in an 18 percent reduction in time to market of new applications.
- Supported Enterprise Operations team for critical applications with issue triage and resolution and reduced support cases by 15% implementing proactive monitoring.

**Infosys, Business Intelligence Developer** -Hyderabad (India)

July 2014 - Oct 2018

- Automated the data gathering from SQL databases through schedulers and event-based triggers, which reduced team size by 10%.
- Developed, maintained, and enhanced various mappings and workflows in Informatica PowerCenter to perform ETL of data from multiple sources into the data warehouse.
- Saved about 16 hours of manual work each week by generating various Drill-through, parameterized, and linked reports using Tableau.

---

## ACADEMIC PROJECTS:

**NoSQL Project on Yelp Dataset (HBase and MongoDB)**

Oct 2020

- Utilized Non-relational database structures and schemas to store Yelp business attributes and retrieved these data for processing query.
- Substantiated the value of these other ways to store data over using HDFS and joined them with data stored in HDFS in real-time.

**Covid-19 Analysis (NiFi, PySpark, Hive, Kafka, Airflow, AWS QuickSight)**

July 2020

- Simulated a complex real-world data pipeline based on messaging on top of a distributed storage system.
- Created a scalable, reliable, secure data architecture and performed end-to-end Big data pipeline implementation on AWS.

**Rain Prediction in Australia (Regression, Scikit-Learn, NumPy, Pandas)**

May 2020

- Implemented a prediction model using various regression techniques like Linear regression, ridge, lasso, SVM, Decision trees, Random forest, etc., used grid search to tune hyperparameters which led to decision trees as the best model.

**Sales data analysis using Tableau**

June 2019

- Developed interactive dashboards in Tableau using LOD, data blending, drill-down techniques, calculated field, and parameter controls to get insights on the customers and comparison of shopping patterns for millennials and Gen X.

---

## LEADERSHIP & AWARDS:

- Recognized and awarded with the **Star Team Performer** award in a Project team of 35 employees in Infosys for my leadership and overall annual performance.
- Coordinated and managed a Data Analytics fest in the capacity of an Events Chair for the IT Management Leadership Council, UT Dallas.
- Received the prestigious **Golden Comet Award**, awarded for the best organization among 300 organizations in UT Dallas, for Indian Students Association where I served as an Officer.