CHICHILI VINEETHA

• +91-9949659936 • vineetha.chichili@gmail.com • <https://in.linkedin.com/in/chichili-vineetha-9ba7b3ab>

• 929, AECS Layout-C Block, BANGALORE, INDIA - 560037

**EDUCATION**

NAtional institute of technology WarAngal, India May 2017

*Bachelor of Technology* in **Electrical and Electronics Engineering** (GPA: 7.62/10)

## EXPERIENCE

CAPGEMINI TECHNOLOGY SERVICES LIMITED, Bangalore, India July 2017 – Present

Associated with Insights and Data business Unit as an Associate consultant in Telecommunication and media industry.

TECHNICAL SKILLS

* **Languages:** Scala, Java, Shell Scripting, SQL , Python
* **Big Data Tools**: Hadoop ,Hive ,Spark, Spark Streaming, Oozie, Sqoop, Pig, Kafka, Phoenix, Apache NiFi, Ambari, YARN
* **Cloud Technology :** AWS
* **Databases:** Oracle, MySQL, Hbase
* **Versioning and CI/CD Tools:** Git, Tortoise SVN, Jenkins, Bit bucket
* **Hands-On Tools:** Elastic Search, Informatica Power Center
* **Proficiencies:** Algorithms, Data Structures, OOPS

PROJECTS

 Data Masking Automation, Capgemini Feb 2021 – Present

* + - Implementing a framework which will mask (Right to Forgot) PII columns in a nested AVRO data identified by business using Spark and Scala.

 Linear Service Addressability, Capgemini Oct 2020 – Jan 2021

* + - Consumed real time Linear Advertisement impression data using Spark Streaming from Kafka. Parsed the Semi-Structured XML and JSON Messages and generated reports which helps the business to calculate the ads revenue.

Linear and VOD Reports, Capgemini Apr 2020 – Sept 2020

* + - Correlated ORC, Parquet, CSV data sets using Scala and Spark to generate Viewership reports for Live TV and Video on Demand (SVOD, TVOD, EST, FVOD) content watched using internet which will help the client to understand the customers viewing preferences across a range of devices, age groups, programs. Automated the reports generation using shell scripting.

 Video Audience Measurement, Capgemini Jan 2019 – Mar2020

* + - Ingested high volumes of event logs for Live TV and Recordings generated by client Set-top boxes from elastic search in batch mode and also from Amazon Kinesis Streaming, Kafka brokers at real time to Hadoop using Spark and NiFi.
		- Correlated different data sets of around 1TB data each day using Scala and Spark to generate viewership reports which will help client to monetize the reports to content providers and advertising partners.
		- Developed a data quality framework using hive queries, which can detect data supply issues.
		- Reported a revenue leakage on customer subscriptions which received appreciations from client.

 Kafka Topic Ingestion, Capgemini Aug 2018 – Dec 2018

* + - Developed a framework using Spark, which is capable of consuming real-time streaming data from Kafka topics and stores the data in the enterprise Hadoop environment.
		- Implemented complex transformations using Spark SQL to read nested Avro messages and loaded the data into hive tables to report key metrics.
		- Added a Scala functionality that is capable of detecting Avro schema evolution and sends an automated e-mail of what columns got newly added, removed, or got their data type changed. This reduced manual effort of around 300 hrs. in a year.

 TV Everywhere, Capgemini Apr 2018 – July 2018

* + - Generated viewership reports for content providers and automated the process using the oozie fork concept to run 24 actions parallelly at a time every week to send the data via SFTP or e-mail or to AWS S3 Bucket adhering to data security, governance, and encryption rules.

 COX Enterprise Analytics, Capgemini Feb 2018 – Mar 2018

* + - Analyzed and improved the effectiveness of the HiveQL scripts by 20% to pull the data necessary for predicting avoidable truck rolls based on the calls received to the customer support center and accessing the Network health information of the nodes in real-time and automated these scripts.

 Video on Demand-Dynamic Ad Insertion, Capgemini Dec 2017 – Jan 2018

* + - Helped functional users on resolving functional defects and made enhancements to a java code, which will convert a JSON message to a HQL query to get the counts or the customer information who watched a particular video on demand program and fine-tuned the query to run within 120 secs using hive LLAP.

 **SQOOP Framework,** Capgemini Oct 2017 – Dec 2017

* + - Created a reusable shell script that is capable of generating the job.properties file dynamically to trigger an oozie workflow for doing a sqoop import and export between oracle and HDFS and added the functionality of monitoring the oozie flow on its own.

##  ACHIEVEMENTS

* Received Extra Mile Award (2020) in appreciation for reporting a revenue leakage to the business at Capgemini.
* Received Rising Star Award (2019) in appreciation for implementing an Avro Schema Evolution detector at Capgemini**.**