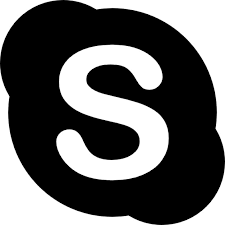
**Chiranjeevi Kolli**

An adept, technically sound Lead Big Data Engineer proficient in Architecture, Analyzing, Designing, Integrating, re-engineering and developing highly sophisticated systems including Analytics, Big data space and Data warehousing. Possessing rich and diverse experience in various technologies an in- depth knowledge of disparate platforms, spanned over 8+ years of industry experience.

* [cj.shans@gmail.com](mailto:cj.shans@gmail.com)  linkedin.com/in/chiranjeevi-kolli
* +60186698370  kolli.chiranjeevi

**TECHNICAL EXPERTISE:**

Big Data Technology: |Hadoop|Spark|Hive|HBase|Kafka|Sqoop|MapReduce|HDFS|Oozie|

Languages/Scripting: |Python|Scala|MATLAB|R|Unix Shell Scripting|

Database (RDBMS/NoSQL): |Oracle|Teradata|HBase|

System Integration Tools: |Pentaho Data Integration|DataStage|Teradata GCFR|

Visualization: |Tableau|

Operating Systems: |Linux|AIX|Windows|

Other tools: |SBT|Git|ELK|Control-M|

**PROFESSIONAL EXPERIENCE:**

Standard Chartered GBS- Client Due Diligence (CDD) Enablement Sep 2015 - Present

Lead Big Data Engineer

**Technologies:** Hadoop, Hive, Spark SQL, Python, NiFi, Unix Shell script

**Description:**

* Automating the risk rating generating system by capturing the updates of customer static information.
* Strategic consumption approach to replacing Teradata GCFR (ELT) with Hadoop Filter Transform Load (HFTL) framework.
* Projected saving of $1m per annum by replacing the HFTL framework.

**Responsibilities:**

* Define, design and develop the common Big Data pipeline framework and environment.
* Extraction, preparation and loading the data.
* Enrich and transform the data to perform data analytics.
* Performance analysis, tuning and recommendation.
* Estimation of complexity, volume of work and provide the cost estimation.
* Draw the deeper insight into the pipeline for critical discoveries to addressing the client queries.
* Analyzing data lineage of desperate source systems and provide the inputs on the requirements for proposed systems.
* Applying the statistical/mathematical approaches in area of prediction and optimization.
* Visualizing the data to indentifying the hidden data patterns and distribution.
* Capacity estimation for big data distributed applications.