**Priyanka N Rajan**

|| [priyankanava@gmail.com](mailto:priyankanava@gmail.com) || +91-9698244499 ||

# EDUCATION

1. **St. Joseph’s College of Engineering-Anna University Chennai GPA: 7.2/10**

Bachelors, Computer Science and Engineering, May ‘17

1. **SBOA School and Junior College, Chennai GPA: 8.3/10**

Higher Secondary School, May ‘13

1. **SBOA School and Junior College, Chennai GPA: 9/10**

Secondary School, May ‘11

# Work Experience

1. **Wolters Kluwer (May 2017 – November 2018)**

* Role: Associate software engineer in Automation
* Responsibilities:
* Used cloudera environment to load files into Hadoop HIVE tables and performed ETL aggregations in Hadoop HIVE
* Used sqoop to import and export data from SQL server to Hadoop ecosystem
* Did Scheduling and monitoring the console outputs through Jenkins
* Testing and bug fixing and providing support the production

1. **Mobius Knowledge Services (November 2018 - present)**

* Role: Machine Learning Analyst
* Responsibilities:
* sequence data pre-processing, extraction, model fitting and validation
* file format transitions Text, Sequence, ORC, Avro, JSON, Parquet

# TECHNICAL SKILLS

* Big Data / Hadoop : HDFS, MapReduce, HBase, spark, PIG, HIVE, Sqoop, Impala and Flume
* Real time/Stream Processing : Apache, Spark
* Operating Systems : Windows, Unix and Linux
* Programming Language : Java, Scala, SQL, Python
* Database : SQL Server, MS Access
* Methodologies : Agile, Scrum and Waterfall
* Data Ingest:
* Import and Export data from a MySQL database into HDFS using Sqoop
* Change the delimiter and file format of data during import using Sqoop
* Ingest and process real-time and near-real-time streaming data into HDFS using Flume or Spark Streaming
* Load data into and out of HDFS using the Hadoop File System command line.
* Transform, Stage and Store:
* Load, Read and Write RDD data from HDFS for use in Spark applications using Spark RDD and Spark DF
* Perform standard extract, transform, load (ETL) processes on data using Spark RDD, Spark DF and Hive
* Data Analysis:
* Use metastore tables as an input source or an output sink for Spark applications using Spark RDD, Spark DF, Spark SQL, Hive, Impala
* Query and Filter datasets in Spark using Spark RDD, Spark DF, Spark SQL
* Write queries that calculate aggregate statistics using Spark DF, Spark SQL, Hive and Impala
* Join disparate and produce ranked or sorted data datasets using Spark using Spark RDD, Spark DF and Spark SQL

# FINAL YEAR PROJECT

Sentiment Analysis of Product Details

• System rates the product, based on the weightage of the keywords in database.

# EXTRA-CURRICULAR

* + Participated in India’s biggest networking championship 2015 grand finale at IIT-Delhi (CISCO).
  + Participated in ICTACT digital youth summit
  + Zonal winner for India’s biggest networking championship conducted by CISCO
  + Cambridge English level 1 certificate – Business English (vantage)