|  |  |
| --- | --- |
| **PAVAN KUMAR**  **DATA ENGINEER**  [**pavankumar@bigdataeng.com**](mailto:pavankumar@bigdataeng.com) **| (415)579-8820** |  |

**PROFESSIONAL SUMMARY:**

**3.5+ years** of **Software Engineering** experience**.** Involved in various SDLC methods from analysis, design, development, testing,

implementation and maintenance with timely delivery against aggressive deadlines in both **Agile/Scrum** and **Waterfall** methodology. 

Good experience in project Leadership with excellent Communication, Management and Presentation skills.

**SKILLS:**

**DATA INGESTION: Sqoop, Azure Data Factory, Azure Poly Base**

**DATA PROCESSING: Spark, Hive, Map Reduce**

**HADOOP ECOSYSTEM: HDFS, Spark (PySpark), Hive, Oozie, Sqoop, NiFi**

**LANGUAGES: Python, Scala, Java, Shell, SQL**

**RELATIONAL DATABASES: MySQL, Oracle, Azure SQL DB**

**NoSQL DATABASES: HBASE, MongoDB, Monet DB, CosmosDB**

**MONITORING: Cloudera Manager**

**DISTRIBUTIONS: Cloudera, Hortonworks**

**VERSION CONTROL: GIT**

**BUILD TOOLS: Maven, SBT**

**IDE: IntelliJ, PyCharm**

**CLOUD: Azure, AWS**

**PROFESSIONAL EXPERIENCE:**

**HEXADUS Corp | Role: Data Engineer June 2018 – Present**

**Client: 84.51°**

**Overview: As a part of AC/DC team, helped in making STRATUM (a data analytics solution) better by adding functionality,**

**improving data quality and data availability speed for customers.**

**As a part of DataGazers team, helped in migration of data, daily and weekly processes to Azure**

**Environment: PySpark, Spark SQL, Spark Core, Scala, Python, HDFS, Automic, Hive, Oracle, Sqoop, Oozie, Cloudera.**

* Supported **ETL** of transactions, dimensions and segmentations data.
* Helped in Improving the speed of data availability from **16 hours to 8 hours** and making sure the **SLA** is met on every Monday.
* Helped in reprocessing efforts of the segmentations data.
* Provided **On-Call support** for our daily and weekly processes.
* **Lead** the **migration** of daily and weekly processes from **Spark 1.6** to **Spark 2.3**
* Supported in adding new functionality and columns to **Spark** jobs written in **Scala**.
* Developed **ETL** applications using **Python, Spark (PySpark)** and **Shell scripting** based off business requirements.
* **Lead** the **migration** of converting a few **SAS** processes to **PySpark**.
* Worked with **CSV**, **JSON** and **PARQUET** files.
* Supported **Oozie** workflows for scheduling and orchestrating the **ETL** process
* Used **JIRA** for **Agile** development.
* Used **Confluence** to document and share process related information.
* Worked on developing automation scripts using **Python**
* Automated processes using **Shell** and **Python** scripting.

**IKON TECH LLC | Role: Software Engineer Sep 2017 – May 2018**

**Client: FBHS (Fortune Brands Home and Security)**

**Environment: Hortonworks, Spark, Hive, Ambari.**

* Worked on setting up Hadoop cluster, access and security.
* Helped setting up **hive** tables for facts and dimensions in **HDFS.**
* Ingested data using Sqoop into Hadoop (HDFS).
* Monitored applications, optimized and troubleshooted using **YARN** logs and **Ambari**.
* Used **PySpark** to transform data in staging tables before loading into production tables.

**EXTARC Software Solutions | Role: Software Engineer Jan 2015 – Nov 2015**

* Supported day to day operations of **ETL** processes.
* Debugged overnight **ETL** failure.
* Identified, Tested and Corrected changes to data using **sqldeveloper**.
* Independently supported the preproduction environment.
* Developed shell scripts to automate administration duties

**EXTARC Software Solutions | Role: Software Engineering Intern Mar 2014 – Sep 2014**

* Completed a proof of concept of the project and followed with business users to qualify the deliverable.
* Integration of company software from conception to production release.
* Followed an **Agile** methodology to deliver frequent and well-tested code to production users.
* Interfaced and collaborated with **development, testing, and operations** teams throughout the company.
* Utilize APIs to prototype and develop interfaces to existing production systems.
* Learned to design **software architecture**, work within short deadlines, and adapt to the company culture.

**EDUCATION:**

MASTERS IN EMBEDDED ELECTRICAL AND COMPUTER ENGINEERING - **3.55 -** Aug 2015 - Jan 2018 **-** San Francisco State University - USA

BACHELoR’S IN ElECTRONICS AND COMMUNICATION SYSTEMS – Jun 2011 - Apr 2015 - Gandhi Institute of Technology - India