SENTIMENTAL ANALYSIS

WIREFRAME DOCUMENTATION

INTRODUCTION

The customer review data processing pipeline is a project aimed at extracting valuable insights from customer reviews using Apache Spark 3.2.1 and Hadoop 3.3.1. The objective of this project is to analyze customer sentiment expressed in reviews and provide businesses with actionable information to enhance their products and services. The purpose of the wireframe documentation is to visually represent the user interfaces, system components, and data flow within the pipeline. It serves as a guide for designing and implementing the pipeline, ensuring a clear understanding of the architecture and facilitating effective communication among stakeholders. The wireframe documentation helps in visualizing the pipeline's structure, user interactions, and overall flow, enabling efficient development and seamless collaboration.

ARCHITECTURE

Data Source (Amazon S3)

- The pipeline starts by accessing the customer reviews stored in an Amazon S3 bucket.
- The reviews are typically in CSV format and located at "s3://sentimental-analysis-project-ineuron/customer_reviews.csv".

Data Ingestion (Apache Spark)

- Apache Spark, running on a Spark cluster, is responsible for ingesting the customer reviews data from the S3 bucket.
- Spark utilizes the "s3://" file system protocol to efficiently retrieve and load the data into its distributed memory.

Storage in HDFS

- The processed customer reviews, along with the sentiment analysis results, are stored in HDFS (Hadoop Distributed File System) for persistent storage.
- The HDFS location for storing the data is typically "hdfs://localhost:9000/customer_reviews".

Iterative Execution and Scheduling

- The pipeline is designed to run iteratively, triggered automatically after each hour.

DATA

CUSTOMER_REVIEW DATA SAMPLE

{

    "\ufeffid": "AVpgNzjwLJeJML43Kpxn",

    "dateAdded": "2015-10-30T08:59:32Z",

    "dateUpdated": "2019-04-25T09:08:16Z",

    "name": "AmazonBasics AAA Performance Alkaline Batteries (36 Count)",

    "asins": "B00QWO9P0O,B00LH3DMUO",

    "brand": "Amazonbasics",

    "categories": "AA,AAA,Health,Electronics,Health & Household,Camcorder Batteries,Camera & Photo,Batteries,Household Batteries,Robot Check,Accessories,Camera Batteries,Health and Beauty,Household Supplies,Batteries & Chargers,Health, Household & Baby Care,Health Personal Care",

    "primaryCategories": "Health & Beauty",

    "imageURLs": "https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81ZTa1BrkzL._SL1500_.jpg,http://ecx.images-amazon.com/images/I/41pSoRu7UlL._SS40_.jpg,https://www.upccodesearch.com/images/barcode/0841710106442.png,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%252525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/51DfBuqy4qL._SL1189_.jpg,https://images-na.ssl-images-amazon.com/images/I/51TshsEJTIL.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2525252525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%25252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/51g3uIwzOIL._SL1050_.jpg,http://ecx.images-amazon.com/images/I/613LUdOc1eL._SS40_.jpg",

    "keys": "amazonbasics/hl002619,amazonbasicsaaaperformancealkalinebatteries36count/b00lh3dmuo,0841710106442,841710106442,amazonbasicsaaaperformancealkalinebatteries36count/b00qwo9p0o",

    "manufacturer": "AmazonBasics",

"manufacturerNumber": "HL-002619",

"reviews.date": "2017-03-02T00:00:00.000Z",

"reviews.dateSeen": "2017-08-28T00:00:00Z",

"reviews.didPurchase": "",

"reviews.doRecommend": "",

"reviews.id": "",

"reviews.numHelpful": "",

"reviews.rating": "3",

"reviews.sourceURLs": "https://www.amazon.com/product-reviews/B00QWO9P0O/",

"reviews.text": "I order 3 of them and one of the item is bad quality. Is missing backup spring so I have to put a pcs of aluminum to make the battery work.",

"reviews.title": "... 3 of them and one of the item is bad quality. Is missing backup spring so I have ...",

"reviews.username": "Byger yang",

"sourceURLs": "https://www.barcodable.com/upc/841710106442,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-Count/dp/B00LH3DMUO/,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-Count/dp/B00LH3DMUO,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-36-Pack/dp/B00LH3DMUO/,http://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-36-Pack/dp/B00LH3DMUO"
},
{
"\ufeffid": "AVpgNzjwLJeJML43Kpxn",

"dateAdded": "2015-10-30T08:59:32Z",

"dateUpdated": "2019-04-25T09:08:16Z",

"name": "AmazonBasics AAA Performance Alkaline Batteries (36 Count)",

"asins": "B00QWO9P0O,B00LH3DMUO",

"brand": "Amazonbasics",

"categories": "AA,AAA,Health,Electronics,Health & Household,Camcorder Batteries,Camera & Photo,Batteries,Household Batteries,Robot Check,Accessories,Camera Batteries,Health and Beauty,Household Supplies,Batteries & Chargers,Health, Household & Baby Care,Health Personal Care",

"primaryCategories": "Health & Beauty",

"imageURLs": "https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81ZTa1BrkzL._SL1500_.jpg,http://ecx.images-amazon.com/images/I/41pSoRu7UlL._SS40_.jpg,https://www.upccodesearch.com/images/barcode/0841710106442.png,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%252525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/51DfBuqy4qL._SL1189_.jpg,https://images-na.ssl-images-amazon.com/images/I/51TshsEJTIL.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2525252525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%25252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/81qmNyJo%2525252BkL._SL1500_.jpg,https://images-na.ssl-images-amazon.com/images/I/51g3uIwzOIL._SL1050_.jpg,http://ecx.images-amazon.com/images/I/613LUdOc1eL._SS40_.jpg",

"keys": "amazonbasics/hl002619,amazonbasicsaaaperformancealkalinebatteries36count/b00lh3dmuo,0841710106442,841710106442,amazonbasicsaaaperformancealkalinebatteries36count/b00qwo9p0o",

"manufacturer": "AmazonBasics",

"manufacturerNumber": "HL-002619",

"reviews.date": "2016-08-31T00:00:00.000Z",

"reviews.dateSeen": "2017-08-28T00:00:00Z",

"reviews.didPurchase": "",

"reviews.doRecommend": "",

"reviews.id": "",

"reviews.numHelpful": "",

"reviews.rating": "4",

"reviews.sourceURLs": "https://www.amazon.com/product-reviews/B00QWO9P0O/",

"reviews.text": "Bulk is always the less expensive way to go for products like these",

"reviews.title": "... always the less expensive way to go for products like",

"reviews.username": "ByMG",

"sourceURLs": "https://www.barcodable.com/upc/841710106442,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-Count/dp/B00LH3DMUO/,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-Count/dp/B00LH3DMUO,https://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-36-Pack/dp/B00LH3DMUO/,http://www.amazon.com/AmazonBasics-Performance-Alkaline-Batteries-36-Pack/dp/B00LH3DMUO"