**BIG DATA**

**Latifur Khan**

**Website Finger Printing**

**Allen**

**Anvesh Ragi**

**Appa Rao**

**Gbadebo Gbadero Ayoade**

**Jheng-Hao Chen**

**Linda**

**Satyam Kotikalapudi**

**Table of contents**

Project Description

Design

Flow Chart

Analysis

source code

References

**Project Description:**

**Design:**

**Sniffer:**

**Kafka:**

**Spark Streaming:**

**Learning Algorithm:**

**Flow Chart:**

START

Sniffer – Data collection

Zookeeper & Kafka Server

Creating the topic

Spark Clean & Package

Spark streaming to generate ######

Machine Learning Training

STOP

**Analysis:**

**Implementation specifics:**

The development is done in Java, Scala, Apache Kafka, Spark.

**Algorithm:**

**Programming Language**: JAVA

**Input:**

• #############.

**Execution:**

1. ###############################################

**Output:**

• ###############################################

**Source code:**

**References:**

#### <https://github.com/satya592/WebsiteFingerPrinting>

#### <http://spark.apache.org/>

#### <http://kafka.apache.org/documentation.html#uses>