

# Big Data Project - Phase Summary

## Phase 2: Design

### Tasks:

#### - [ ] Designed system architecture diagram:

- Kafka topic `twitter`
- Spark Streaming pulls data from Kafka
- ML model classifies sentiment
- Results stored in MongoDB
- Django reads from MongoDB

#### - [ ] Created folder/module structure for:

- Kafka producer
- Spark consumer
- PySpark ML model
- Django dashboard

#### - [ ] Designed basic UI mockups:

- Dashboard page with bar/pie chart
- Table with tweet text and labels
- Form to classify custom text