

Big Data Project - Phase Summary

Phase 3: Development

Tasks:

- [x] Wrote Kafka producer to send `twitter_validation.csv` tweets to topic.
- [x] Trained PySpark model using `twitter_training.csv`.
- [x] Wrote Spark Streaming consumer to:
 - Read from Kafka
 - Predict sentiment
 - Save result to MongoDB
- [x] Created Django views to display:
 - Real-time plots (Chart.js, matplotlib)
 - Table of classified tweets
 - Text input for custom classification