STREAMANALYTIX

BIG DATA TRENDS FOR 2018

Data is a powerful corporate asset that enterprises are now beginning to fully harness. Enterprises are looking to derive breakthrough value through investments in cloud-migration, data lakes, in-memory computing, modern business intelligence, and data science technologies.

Following predicts represent views of multiple Fortune 500 companies that are actively investing to transition into future-ready data-driven real-time Enterprises in 2018.

Trend 1: Move to Public Cloud

Availability of pre-built, reliable, and scalable PaaS for development and deployment will accelerate shift to cloud

Public cloud services are more cost effective compared to running large application DevOps on private infrastructure

With the public cloud becoming more robust and secure, primary security concerns are diminishing

The shift will enable designing and operating applications without any coding hassles, focusing on the business logic

Trend 2: Year of Real-time Analytics & Stream Processing

Early adopter success in 2017 will drive large scale adoption of Stream processing and advanced real-time analytics as part of core data processing infrastructure

Key business drivers include:

- Competitive pressure
- Need for fast data processing
- Need to predict and act on business opportunities in real-time
- Demand for contextual and time-relevant customer experiences
- Growing use of connected devices and sensors (IoT)
- Need for data driven operational efficiency

Increased demand for off-the-shelf vertical end-user applications like:

Customer 360 Frameworks	Personalized Recommendation	Pre-built Churn Prediction	Call Centre Analytics
IoT, Sensor Data, Log Analytics	Financial Fraud and Risk Analytics	Telecom Network Maintenance	Anomaly and Pattern Detection

Focus will shift to derive a higher value from data lake investments

- Transactional platforms will shift to big data lakes to process transactions in real-time
- Direct BI solutions running on top of data lakes spanning very large data sets will gain critical mass adoption

Trend 3: Dominance of Apache Spark

Apache Spark™ will remain the de-facto big data processing engine

Apache Spark deployments will encompass a wider range of use cases

Self-service functionality of new Apache Spark productivity tools will drive big data and fast data analytic applications

Increased adoption of "build once, deploy both as batch and streaming jobs" feature of Apache Spark Structured Streaming APIs

Penetration levels can surpass Hadoop adoption, driven by cloud-based approaches and non-Hadoop usage of Apache Spark

Trend 4: Emerging Streaming Engines

Apache Flink™ will begin to rise as a "true" low-latency streaming engine

Kafka™ Streams to be the only real competitor to Flink

Apache Apex™, Apache Samza™, etc. to stay as small players

Trend 5: Data Science at the Center of Analytics

Data Science model management will take shape as a mature and out-of-the-box feature

More enterprise will start using machine learning, advanced predictive and prescriptive analytics, and AI at scale

Rule-based and manual approaches will transform into machine learning driven solutions

Deep learning will replace manual retraining of aging or low performing traditional machine learning models

Data science work benches, self-service, and productivity tools for easier model building, training, and execution will gain popularity and higher adoption

IMPETUS

Read Cloud + Streaming Analytics + Data Science = Five Big Data Trends Now for a more detailed view on the Big Data Trends for 2018

Author: Anand Venugopal- AVP and Business Head of StreamAnalytix

StreamAnalytix is an open-source enabled, enterprise-grade multi-engine stream processing and machine learning platform that empowers enterprises across industries to make smart decisions, and act on them in real-time.

Disclaimer: The views and opinions expressed in this article are those of the author's and do not necessarily reflect the official position of Impetus Technologies.