

Simplifying Machine Learning Architectures with an Event Streaming Platform

Kai Waehner

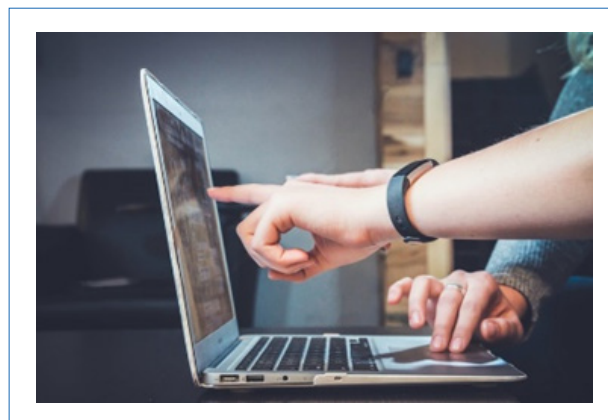
Confluent, Germany



Abstract

Machine Learning (ML) is separated into model training and model inference. ML frameworks typically use a data lake like HDFS or S3 to process historical data and train analytic models. Model inference and monitoring at production scale in real time is another common challenge using a data lake. But it's possible to completely avoid such a data store, using an event streaming architecture.

This talk compares the modern approach to traditional batch and big data alternatives and explains benefits like the simplified architecture, the ability of reprocessing events in the same order for training different models, and the possibility to build a scalable, mission-critical ML architecture for real time predictions with much less headaches and problems. The talk explains how Kai Waehner this can be achieved leveraging the open source frameworks Apache Kafka and TensorFlow.



Biography

Kai Waehner is a Technology Evangelist at Confluent. He works with customers across Europe, US, Middle East and Asia and internal teams like engineering and marketing. Kai's main area of expertise lies within the fields of Big Data Analytics, Machine Learning, Hybrid Cloud Architectures, Event Stream Processing and Internet of Things. He is regular speaker at international conferences such as Apache Con and Kafka Summit, writes articles for professional journals, and shares his experiences with new technologies on his blog.

Publication

1. Confluent Blog: Kafka-Native MQTT at Scale with Confluent Cloud and Waterstream <https://www.bbvaopenmind.com/wp-content/uploads/2015/04/BBVA-OpenMind-Big-Data-and-the-Future-of-Business-innovation-1.pdf>
2. Industrial IoT Podcast: Event streaming architectures enabling IoT applications beyond messaging

[Global International conference on Big Data Analytics & Data Mining](#) | Singapore | July 21-22, 2020

Citation: Kai Waehner: Simplifying Machine Learning Architectures with an Event Streaming Platform, Big Data 2020, Global International conference on Big Data Analytics & Data Mining, Singapore, July21-22,2020, 1