# 04 Smarter and instant analysis for huge amounts of video
## ～ Jubatus, a scalable big data real-time analysis framework ～

### Abstract

Jubatus is an online machine learning framework on distributed environments. This is the first framework in the world. Jubatus learns optimal models without prior knowledge and analyzes complex big data like surveillance videos on memory. Several useful algorithms are provided on Jubatus, such as clustering, regression, recommendation, classification, and graph mining. We can immediately analyze big data arrived at high-speed by using these algorithms. Jubatus archives high performance and scalability by adding low-cost servers for increasing amount of data. How do you analyze increasing big data? The goal of Jubatus is to achieve the world that everyone can easily utilize machine learning. We will promote the development of easy-to-use machine learning frameworks.

## Real-time state estimation for persons from huge amount of video datasets using a classification algorithm



**Fast response**

Quick inference for complex data

- With tie
- Black clothes
- Not reading

- Without tie
- Blue clothes
- Reading

- Without tie
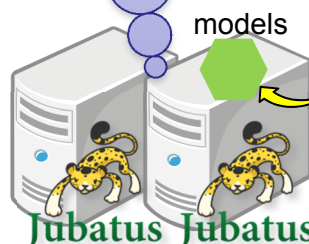- Blue clothes
- Reading

Immediate correction if it made a mistake

**Stable performance**

Only adding commodity PCs to ensure scalability and performance

models

**Online learning**

Jubatus Jubatus

Jubatus is a joint development with Preferred Infrastructure, Inc.   http://jubat.us
NTT Software Corp. provides commercial supports for Jubatus.

### Related work

[1] S. Hido, S. Tokui, S. Oda, "Jubatus: An open source platform for distributed online machine learning," *NIPS Workshop on Big Learning*, 2013.
[2] S. Oda, S. Nakayama, K. Uenishi,  S. Kinoshita, "Jubatus: distributed processing technique enabling realtime processing of big data," IEICE Tech. Rep., Vol. 111, No. 409, IN2011-126, pp. 35-40, 2012 (in Japanese).
[3] H. Makino, "Jubatus: scalable distributed processing framework for realtime analysis of big data," *XLDB*, 2012.

### Contact

**Keitaro Horikawa**     Distributed Computing Technology Project, NTT Software Innovation Center
E-mail：horikawa.keitaro{at}lab.ntt.co.jp (Please replace {at} with @)