CERN openlab

Lightning Talk

Anomaly Detection in the Elasticsearch Service

Jennifer Andersson

15/08/2019

Introduction

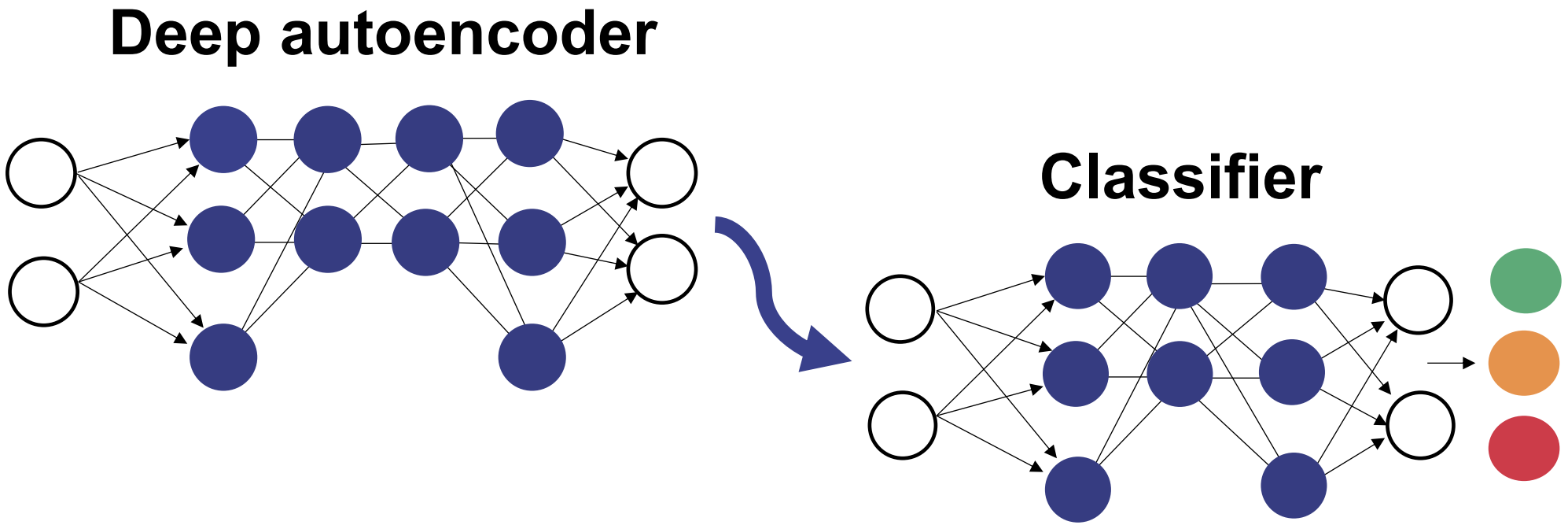
Elasticsearch Service:

- Search- & analytics engine
- 30 clusters & 160 use cases

Project goal:

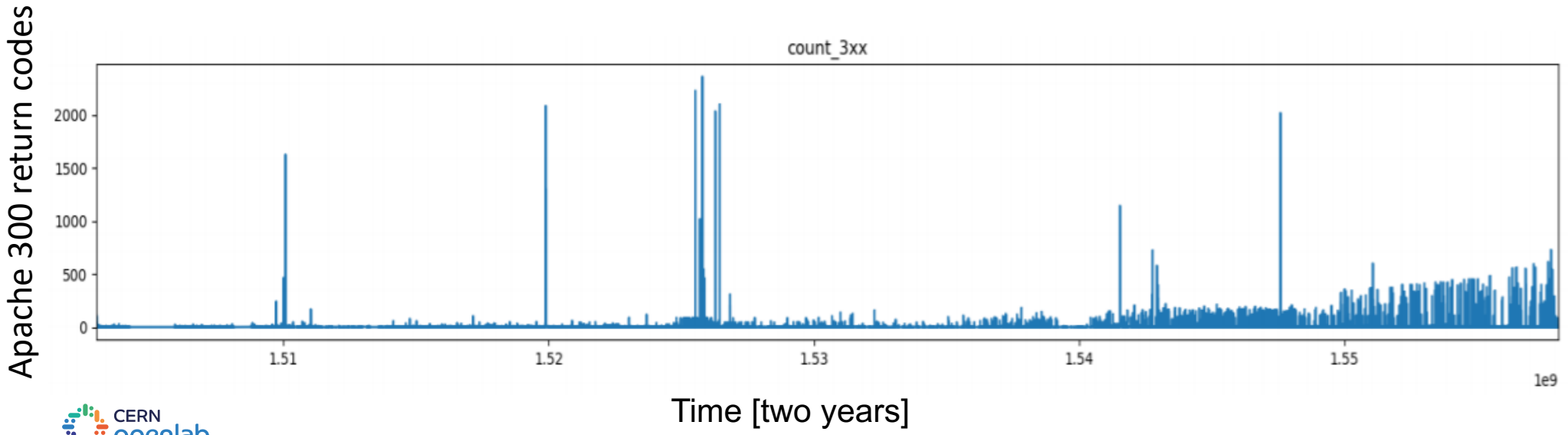
Detect service issues before they cause problems

Anomaly Detection & Degradation Prediction



Project Challenges

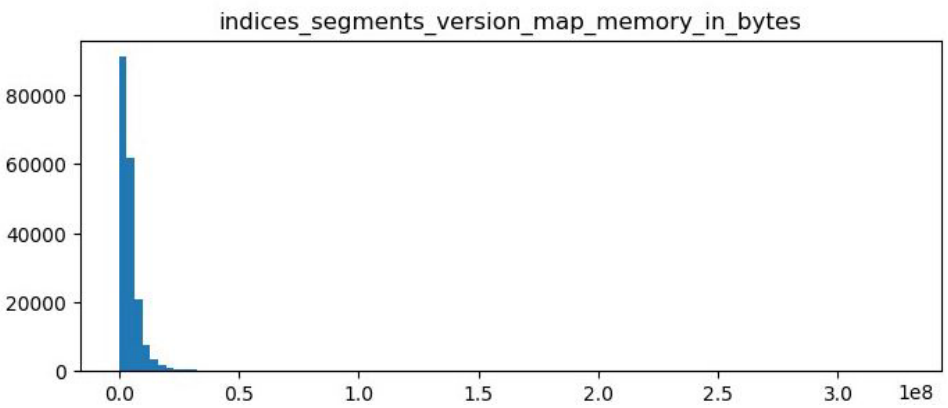
- Better preprocessing of the input data required
- Evolution of cluster characteristics over time requires frequent retraining
- Convergence issues makes retraining hard (vanishing gradient problem)



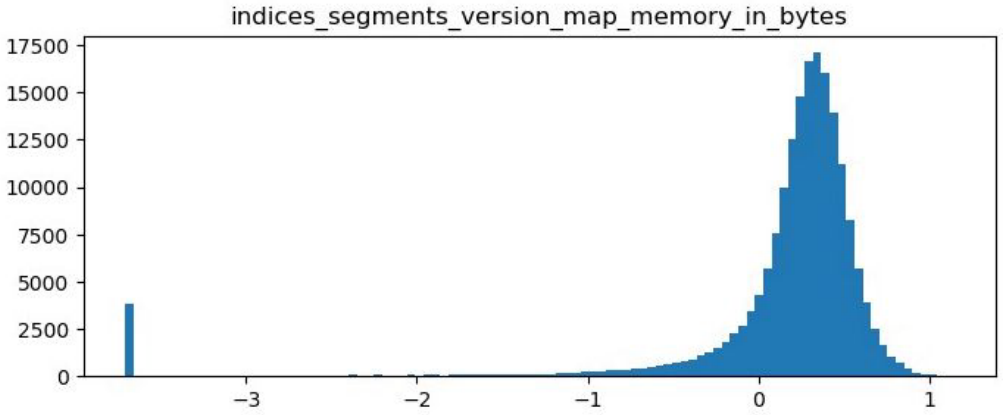
CERN openlab

Data Preprocessing

Raw data



Preprocessed data

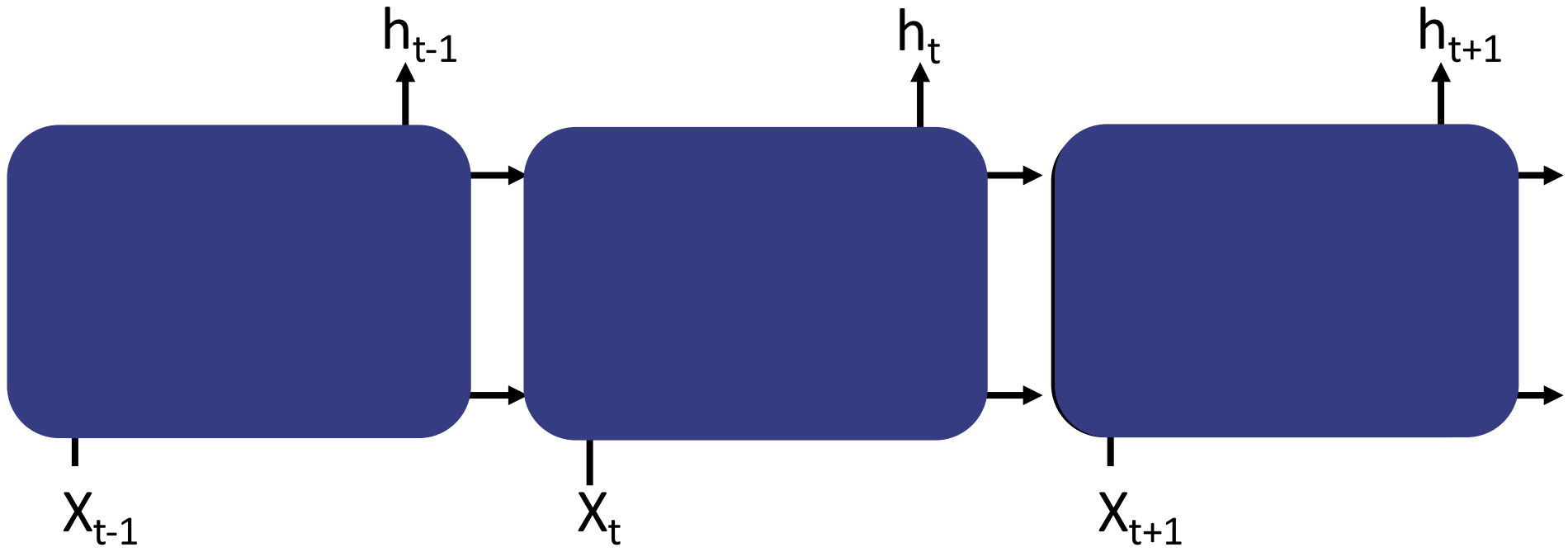


CERN openlab

Long Short-Term Memory Neural Networks

Colah's blog post:

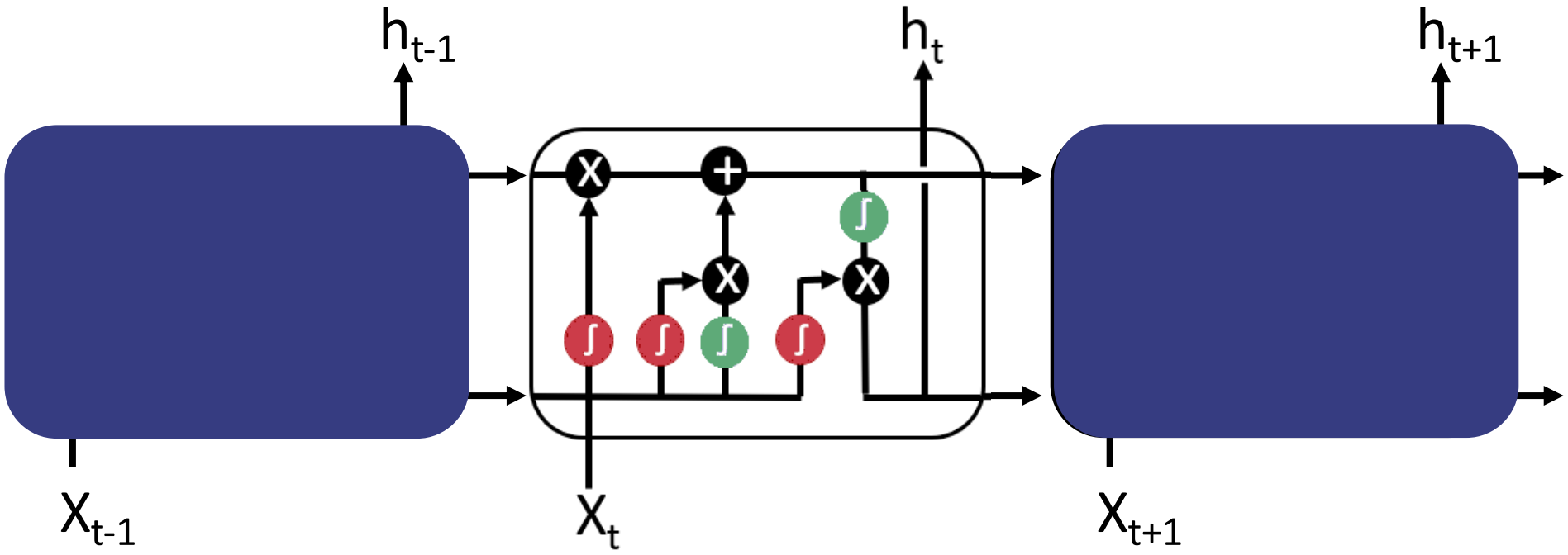
Understanding LSTM Networks



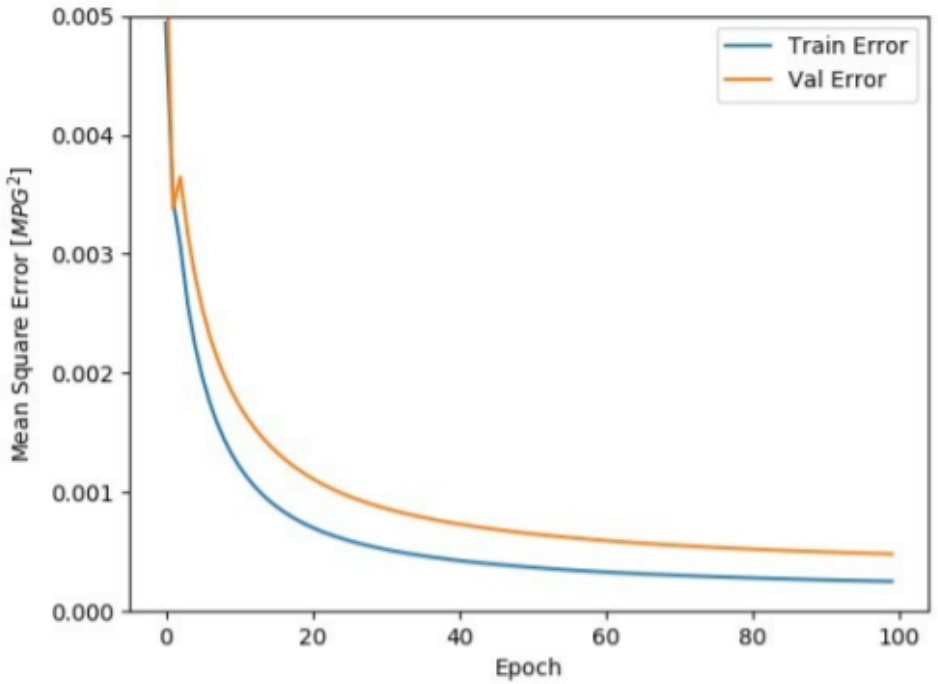
Long Short-Term Memory Neural Networks

Colah's blog post:

Understanding LSTM Networks



Model Evaluation



Ongoing work:

- Compare to
  - non-ML methods, e.g. moving average
  - other ML-methods , e.g. (Extended) isolation forests
- Index anomaly scores to Elasticsearch and create a dashboard

Achievement:

Much better convergence than the previous model

Future work:

- Classification of anomalous events
- Apply the method to other services

Thank you for listening

Contact me:

Jennifer Andersson

jennifer.r.andersson@gmail.com

CERN openlab