It is Just a Machine that Learns

Kristoffer L Nielbo

kln@cas.dk

knielbo.github.io

chcaa.io

Center for Humanities Computing
Aarhus University, Denmark

Outline

It is Just a Machine
that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Text classifier

It is Just a Machine
that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

A transparent and very popular probability learning model that can be implemented efficiently

The probability of a document d being in class c, $P(c \mid d)$ is computed as:

$$P(c \mid d) \propto P(c) \prod_{i=1}^{m} P(t_i \mid c)$$

and the class of a document d is computed as:

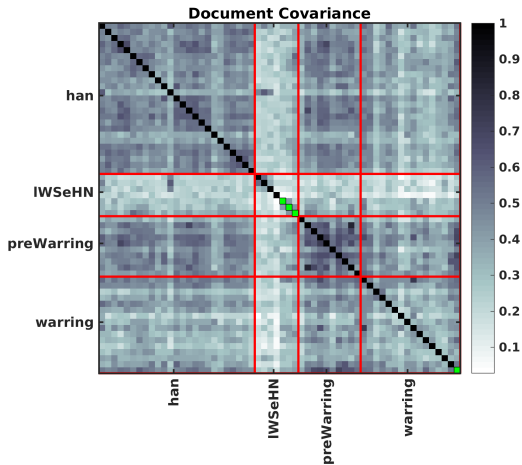$$c_{MAP} = arg\ max_{c \in \{c_1, c_2\}} P(c \mid d)$$

It is Just a Machine
that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Figure 1: some text here

It is Just a Machine
that Learns

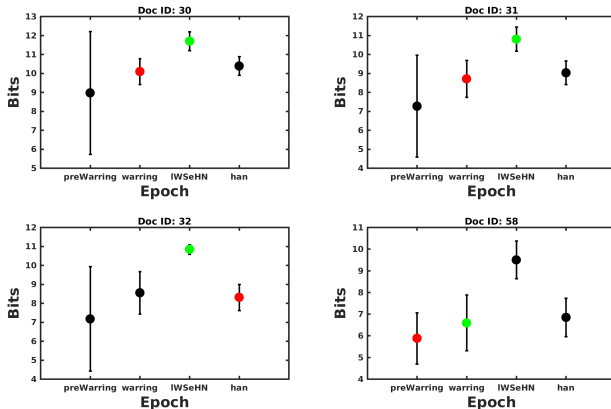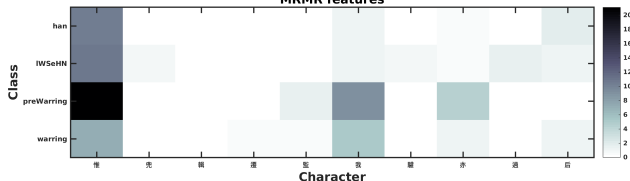Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Figure 2: Average distance to error class is shorter

$$D_{KL}(d_1 \mid d_2) = \sum_{i=1}^{n} d_{i1} \times \log_2 \frac{d_{i1}}{d_{i2}} \qquad (1)$$

It is Just a Machine
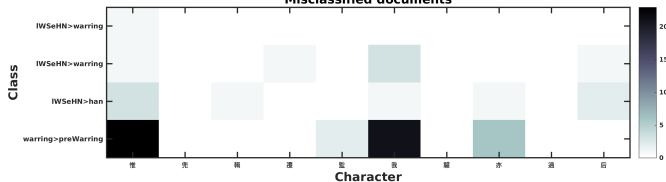that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Figure 3:

It is Just a Machine
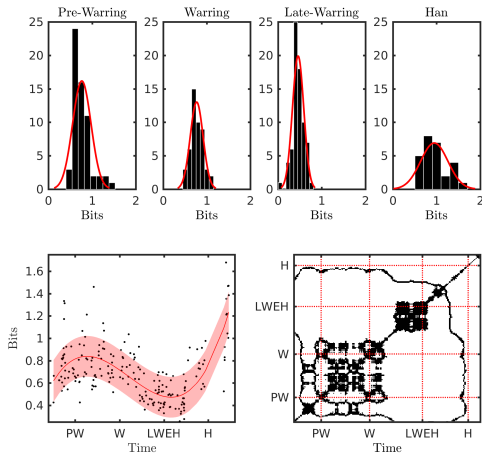that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Figure 4:

It is Just a Machine
that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Figure 5:

It is Just a Machine
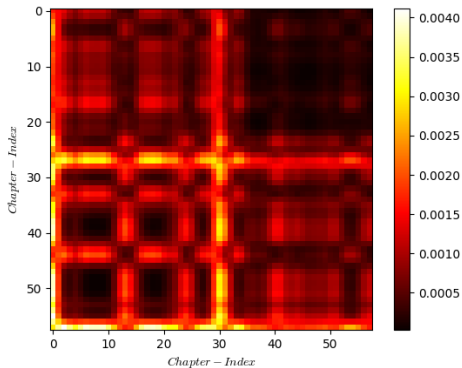that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class

Disruptive dynamics

Persistent dynamics

Concluding remarks

Figure 6:

Figure 7:

It is Just a Machine
that Learns

Kristoffer L Nielbo
kln@cas.dk
knielbo.github.io
chcaa.io

Background
CTEXT database

Erroneous class
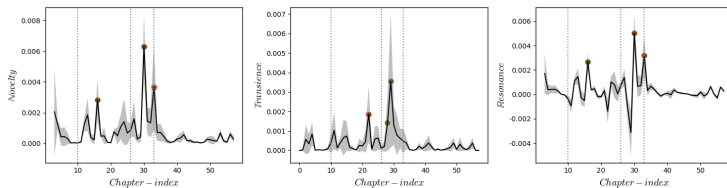
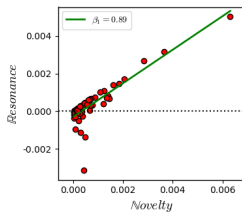Disruptive dynamics

Persistent dynamics

Concluding remarks

TAK

kln@au.dk
knielbo.github.io
chcaa.io

slides: http://knielbo.github.io/files/kln_ann101.pdf