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

Center for Humanities Computing Aarhus University, Denmark

Outline

- 1 Background CTEXT database
- 2 Erroneous class
- 3 Disruptive dynamics
- 4 Persistent dynamics
- 5 Concluding remarks

It is Just a Machine that Learns Kristoffer L Nielbo

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

> Ckground FEXT database

roneous class

uptive dynar

rsistent dynamics



Text classifier

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

CTEXT databas

Erroneous class

Disruptive dynamics



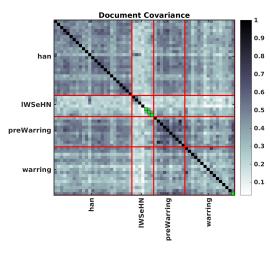


Figure 1: some text here

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

Background

Erroneous class

Disruptive dynamic

· ·



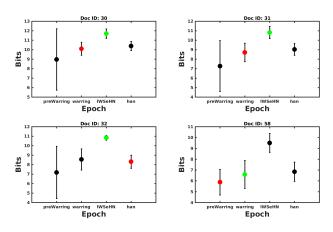


Figure 2: Average distance to error class is shorter

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

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

CTEXT databas

Erroneous class

Disruptive dynamics

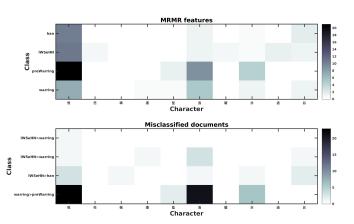


Figure 3:



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

Background

Erroneous class

Disruptive dynamics

Persistent dynamics

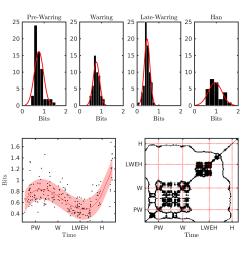


Figure 4:

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

CTEXT database

Disruptive dynamics

Persistent dynamics



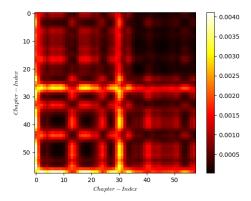


Figure 5:

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

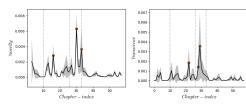
Sackground CTEXT databas

Erroneous cl

Disruptive dynamics

Persistent dynamics





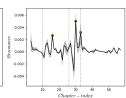


Figure 6:

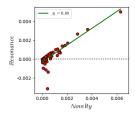


Figure 7:



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

CTEXT databas

Disruptive dynamics

Persistent dynamics Concluding remarks

TAK

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

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

It is Just a Machine that Learns

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

CTEXT database

