

The background features a series of concentric circles in a light gray color, centered on the slide. Overlaid on these are stylized circuit board traces in a light blue color, with small circles at the junctions, located in the corners of the slide.

DIMENSIONALITY REDUCTION

JENS BAETENS

WAT IS HET?

Dimensionality reduction technieken gaan proberen het aantal features dat nodig is te minimaliseren maar proberen om de behouden informatie te maximaliseren.

Veel gebruikte algoritmes:

- PCA of Principal Component Analyse
- Autoencoders → *Volgend jaar*
- Linear Discriminant Analysis
- Missing Values Ratio
- Low Variance Filter



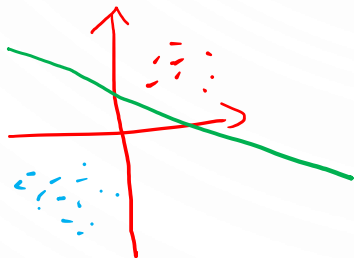
WAAROM?

Curse of dimensionality -> Verscheidene technieken werken minder goed bij een hoog aantal features (performantie / trainingstijd) —> *reprocessing step*

Data compressie

Ruis verwijderen ✓

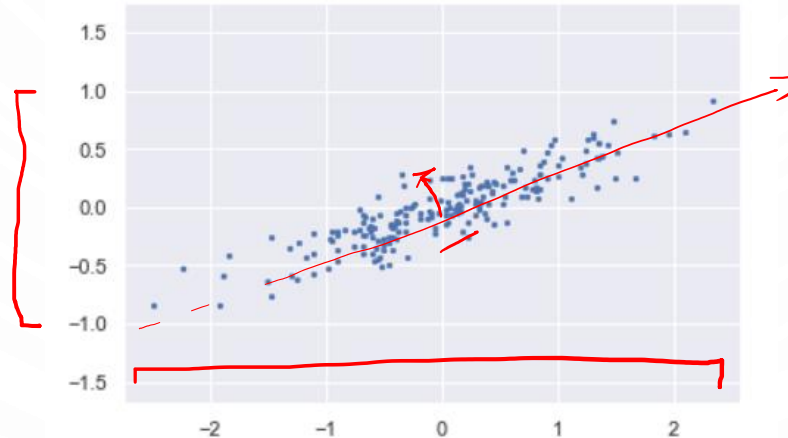
Visualisaties in 2D of 3D



PCA – PRINCIPAL COMPONENT ANALYSE

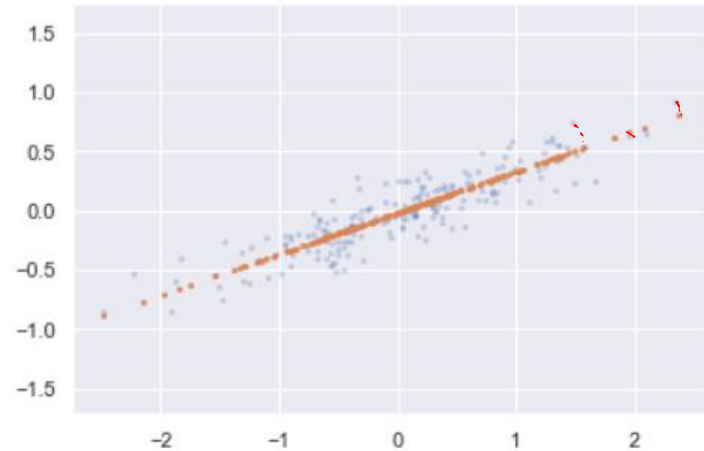
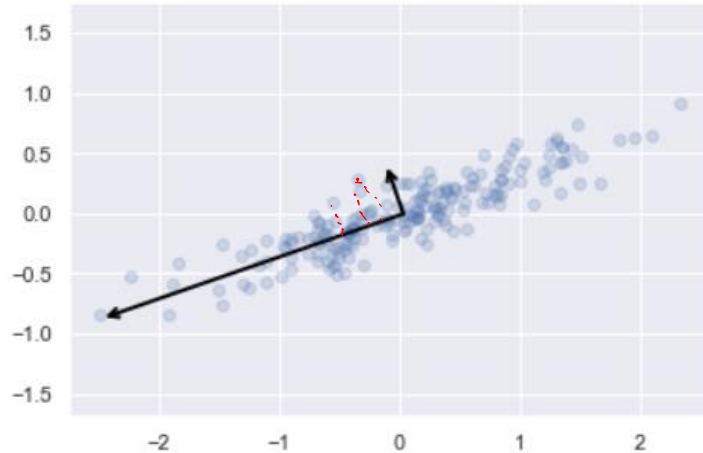
Belangrijke

Transformeren van assenstelsel voor de features om zoveel mogelijke informatie in de datapunten in de eerste features te krijgen zodat de andere weggelaten kunnen worden zonder belangrijke informatie te verliezen.



PCA – PRINCIPAL COMPONENT ANALYSE

2 features naar 2 principle components



PCA IN COMPUTER VISION

Voorbeeld van handgeschreven cijfers (scaled MNIST)

Afbeeldingen van 8x8 pixels met cijfers van 0 t.e.m. 9 (64 features)

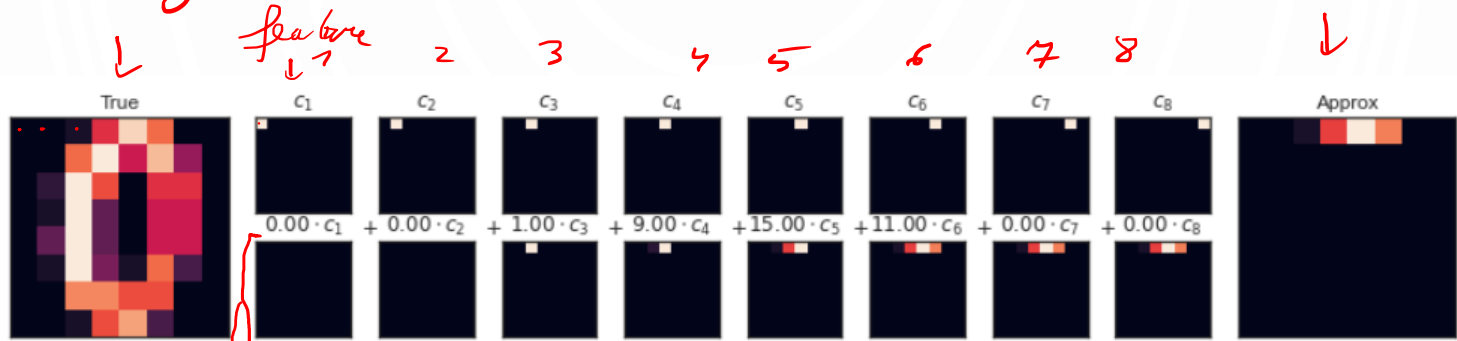


Doel van PCA: Reduceer 64 features maar behoud zoveel mogelijk van de informatie

PCA IN COMPUTER VISION

Startimage

gereduceerd
naar 8 features



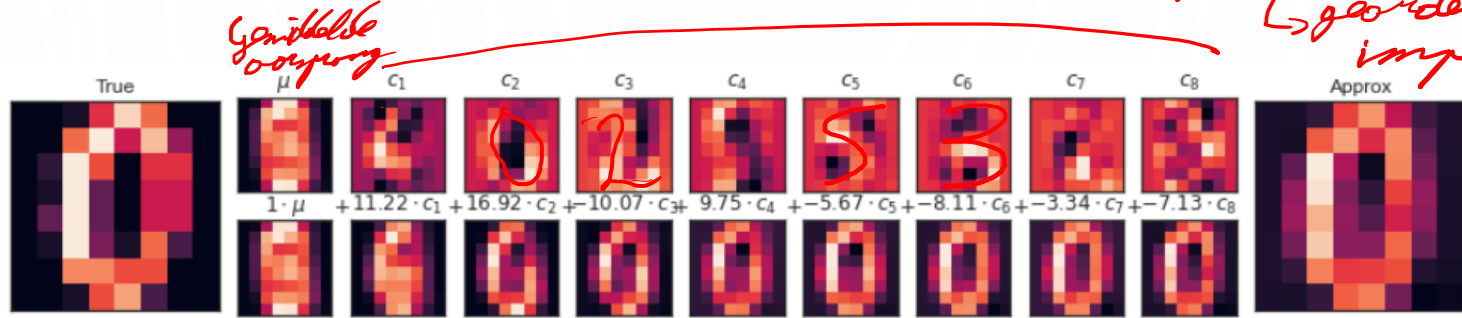
busse bijbre
resultaat

PCA IN COMPUTER VISION

"Nieuwe features"

8 Principal Components

↳ geordend op impact



Na Pca geven de eerste 8 features al een goede benadering van het origineel

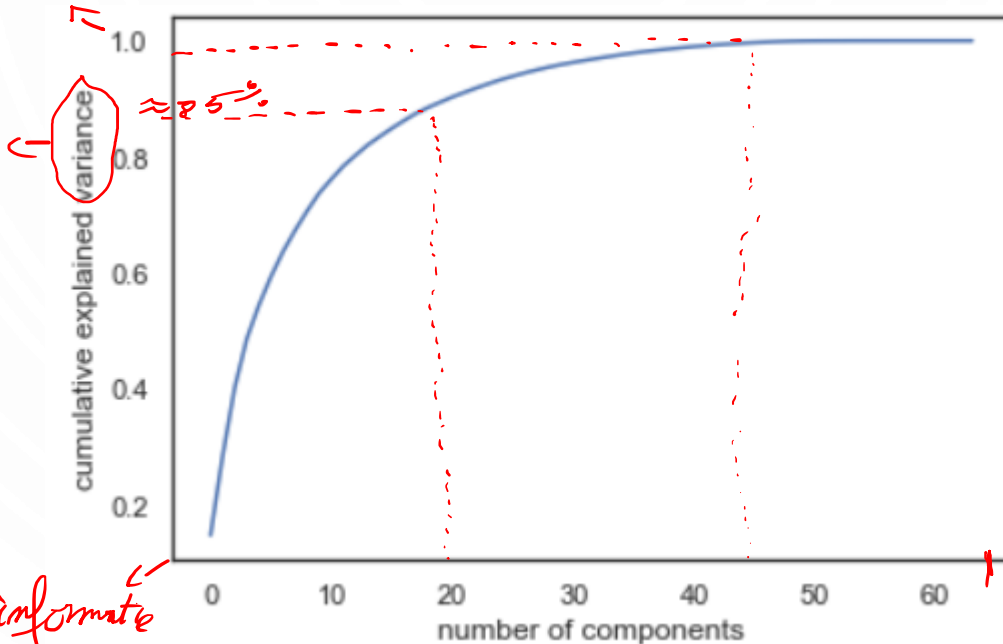
- Sterke compressie van de data (factor 8)
- Verbeterd performantie van de modellen

In computer visie worden deze principal components ook Eigenfaces genoemd

PCA IN COMPUTER VISION → Elbow - methode

alle informatie

Hoeverd % v.b
informatie
blijft
~~over~~
over



geen informatie

↳ # PC's v.d 64 features

PCA VOOR NOISE FILTERING

Aangezien de ruis willekeurig is gaat PCA vooral focussen op de zaken die gemeenschappelijk is tussen de beelden van dezelfde klasse.

Hierdoor wordt de ruis eruit gefilterd

PCA VOOR NOISE FILTERING



Origineel



Met ruis

Na PCA

