

Met de computer naar advertenties kijken

Melvin Wevers

November 28, 2017

Digital Humanities Group - KNAW Humanities Cluster

DH Clinic - Amersfoort

Advertenties als historische bron

- “insight into the **ideals and aspirations** of past realities. They show the **state of technology**, the **social functions of products**, and provide **information on the society** in which a product was sold” (Marchand, 1985)
- Bijna 20 miljoen krantenadvertenties (1890-1990)
- Hoe kunnen we computers inzetten om deze grote hoeveelheid advertenties te bestuderen?

Advertentie als digitaal object

- Complexe en gelaagde bron
- Metadata
- Tekst
- Afbeelding



Figure 1: *Limburger Koerier*, April 13, 1938

Metadata

- Krantentitel
- Datum
- Grootte
- Positie op pagina
- Paginanummer
- Totaal aantal pagina's in krant

Hoe groot waren de advertenties?

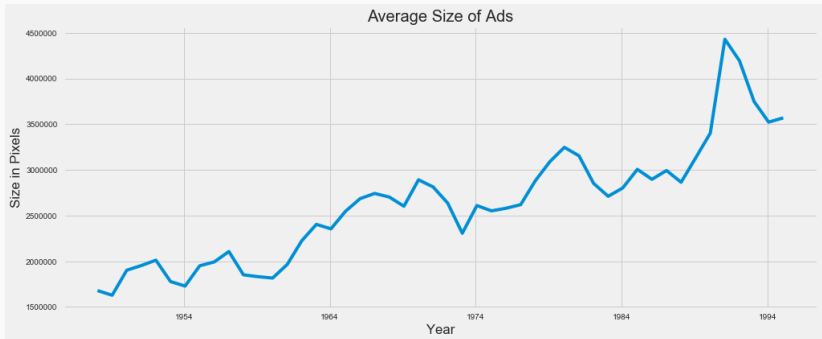


Figure 2: Gemiddelde grootte van advertenties in AH en NRC

Waar stonden de advertenties in de krant?

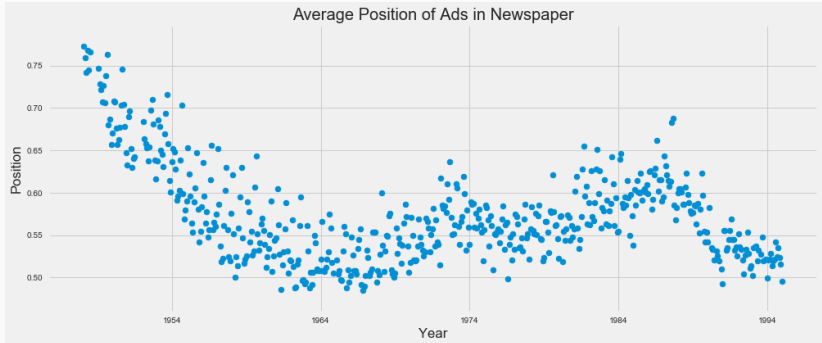


Figure 3: Gemiddelde positie van advertentie in AH en NRC

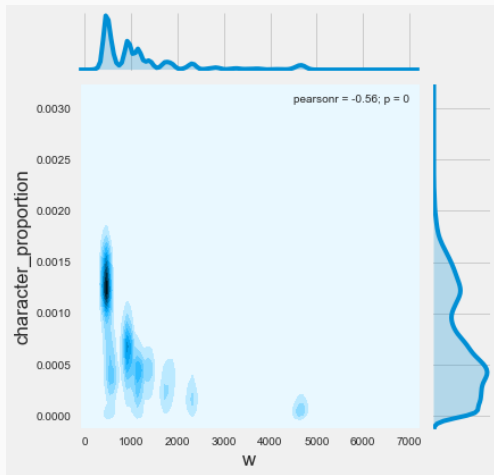


Figure 4: Tekst-beeldverhouding afgezet tegen breedte

Tekst

Zoeken op woorden in de advertentie

The screenshot shows the Delpher search interface. At the top, the Delpher logo is on the left, followed by a search bar containing 'sigaret AND camel'. To the right of the search bar are icons for 'Kranten' (newspapers), a magnifying glass, and a dropdown for 'Uitgebreid zoeken' (extended search). A 'Menu' button is in the top right corner.

Below the search bar, there are filters on the left and search results on the right.

Filters on the left:

- ☐ Duitse (207)
- ☐ Verenigde Staten (35)
- Soort bericht**
 - ☒ **Advertentie (497)**
 - ☐ Artikel (216)
 - ☐ Illustratie met onderschrift (1)
- Krantentitel**

Kies krantentitel...
- Plaats van uitgave**

Kies plaats van uitgave...
- Herkomst**

Kies herkomst...
- Toevoegingen in Delpher**
 - ☐ Toon alleen op 13 april 2016 toegevoegde kranten
- [Meer informatie](#)

Search Results:

The search results are displayed in a list. Each result includes a thumbnail image of a newspaper page, the title 'Advertentie', a star icon, a description, and metadata.

Result 1:

- Advertentie** ☆
- sigaret**/0 rokers. **CAMEL** dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van **CAMEL**
- Krantentitel**: Algemeen Handelsblad
- Datum**: 27-01-1967
- [Meer details ▼](#)

Result 2:

- Advertentie** ☆
- sigaret**/0 rokers. **CAMEL** dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van **CAMEL**
- Krantentitel**: Algemeen Handelsblad
- Datum**: 09-05-1967
- [Meer details ▼](#)

Result 3:

- Advertentie** ☆
- sigaret**/0 rokers. **CAMEL** dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van **CAMEL**
- Krantentitel**: Algemeen Handelsblad
- Datum**: 28-09-1967
- [Meer details ▼](#)

Figure 5: Zoeken in Delpher op 'Sigaret' en 'Camel'

- Woorden tellen
- Co-occurrence analyse
- “You shall know a word by the company it keeps” (Firth, 1957, p. 179)

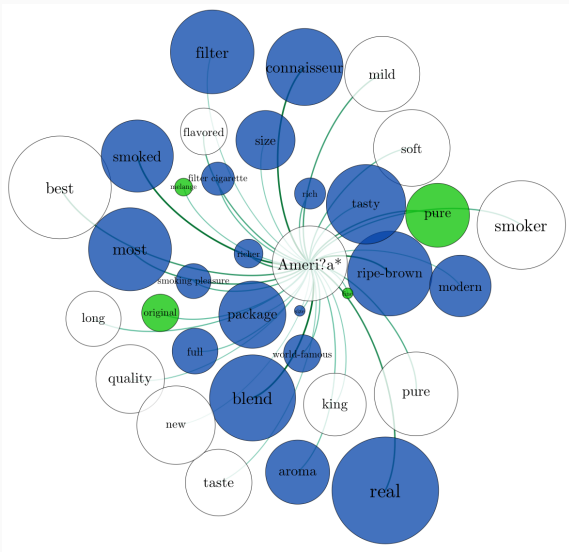


Figure 6: Kenmerken van de Amerikaanse sigaret

Waar gaan we op zoeken?

Golden Shadow
Een voorbeeld uit deze originele Amerikaanse collectie is dit feestjurkje, waarin de Amerikaanse vrouw haar onweerstaanbare charme ten toon stelt. Het jurkje is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **14⁹⁵**.

Twilight
Ook voor de gouden jaren heeft C&A originele Amerikaanse modellen. Speciaal voor hen hebben wij in de prijzen van 1955 het jurkje 'Twilight'. Dit jurkje van satijn heeft een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. Het is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **19⁹⁵**.

Miss Young America
Een cocktailjurk, zoals de Amerikaanse meisjes graag dragen, waarvan zij zeggen: 'De stijl is een perfecte afspiegeling van de Amerikaanse levenswijze'. C&A heeft voor de prijs van **24⁷⁵** een jurkje van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. Het is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **24⁷⁵**.

Little Princess
Hoeveel vrolijkheid heeft een jurkje met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. C&A heeft voor de prijs van **39⁷⁵** een jurkje van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. Het is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **39⁷⁵**.

High noon
Een heren cocktailjurk, zoals de Amerikaanse mannen graag dragen, waarvan zij zeggen: 'De stijl is een perfecte afspiegeling van de Amerikaanse levenswijze'. C&A heeft voor de prijs van **29⁷⁵** een jurkje van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. Het is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **29⁷⁵**.

Vanity Fair
Ook voor de gouden jaren heeft C&A originele Amerikaanse modellen. Speciaal voor hen hebben wij in de prijzen van 1955 het jurkje 'Vanity Fair'. Dit jurkje van satijn heeft een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. Het is gemaakt van satijn met een lange ruitelgordel, een rits aan de achterkant en een rits aan de voorkant. De originele prijs is **49⁷⁵**.

De gehele eigen Amerikaanse mode op het gebied van postleiders heeft nu voor het eerst haar eigen plaats in Nederland. In onze Amerikaanse collectie vindt u vele prachtige en ook nieuwe Amerikaanse stijlen van een volkomen nieuwe en in Nederland onbekende stijl. De originele prijs is **29⁷⁵** en **39⁷⁵**.

Figure 7: Advertentie C&A

Beeld

Van plaatje naar data

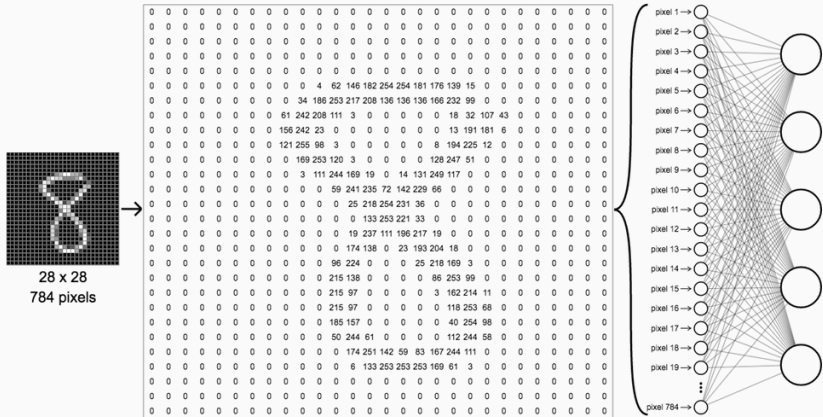


Figure 8: Input layer

Neuraal network architectuur

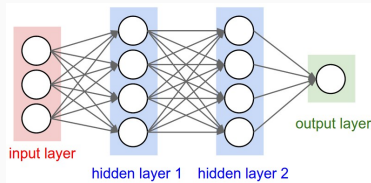


Figure 9: Neuraal netwerk

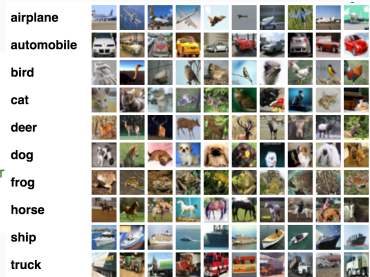


Figure 10: Output layer



Figure 11: Precisie van de beeldherkenning

SIAMESE

- Dataset: 1.6 miljoen advertenties uit *Algemeen Handelsblad* en *NRC Handelsblad* tussen 1948-1995
- Verwijderen:
 - Kleine afbeeldingen (op basis van metadata)
 - Advertenties met specifieke afmetingen, bv. filmladders & rubrieksadvertenties
 - Advertenties met voornamelijk tekst
 - Advertenties met auteursrecht
- Resultaat: 426,777 voor de periode 1945-1994

- Afbeeldingen omzetten in vectoren
- Groeperen op basis van 2,048 abstracte visuele aspecten
- Meest gelijkende afbeeldingen vinden (nearest neighbors)
- Tijdslijn met de meest vergelijkbare afbeelding per jaar

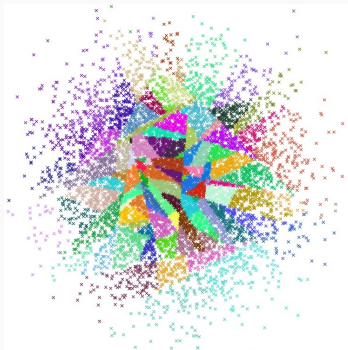


Figure 12: Annoy:
Approximate Nearest
Neighbor

- Exploratief door het beeldarchief navigeren op basis van visuele aspecten
- Inzicht krijgen in visuele trends
- Hypotheses genereren, zoekwoorden vinden
- Afbeelding bekijken in context van de krant
- <http://lab.kb.nl/tool/siamese>
- Samen met Juliette Lonij

Volgende stappen

Hoe kan dit handig zijn voor bibliotheken?

- Classifiers trainen voor specifieke advertenties
- Op basis van meerdere features, classifiers trainen voor advertenties
- Verrijken van collecties
- Combineren van tekstueel en visueel zoeken

Sneak preview: gezichten genereren



Figure 13: Celebrity

Figure 14: Gezichten

Vragen/Opmmerkingen?

Email: `melvinwevers@icloud.com`

`http://www.melvinwevers.nl`

`http://lab.kb.nl/tool/siamese`