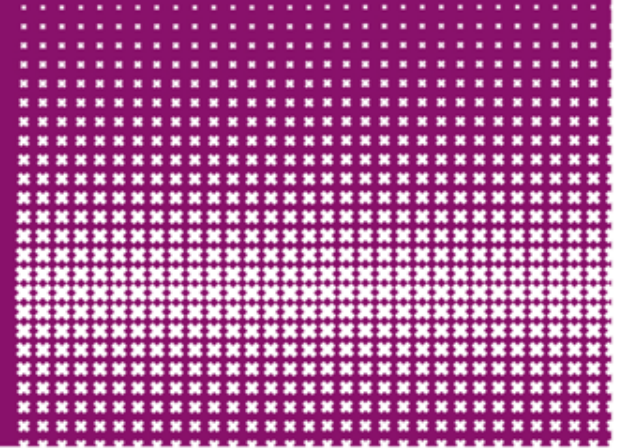




**Richard Berendsen, Bogomil Kovachev, Edgar Meij,
Maarten de Rijke, Wouter Weerkamp**



Classifying queries for people

from a people search engine

People search

leonhard euler - Google Search | Geert Wilders - Persoonlijke g...

www.google.com/#q=Leonhard+Euler&hl=en&biw=1600&bih=762&fp=1&cad=b

Web Images Videos Maps News Shopping Gmail more

iGoogle | Search settings | Sign in

Leonhard Euler

Search

About 978,000 results (0.19 seconds) [Advanced search](#)

Everything

- Images
- Videos
- News
- Shopping
- Books
- More

Any time

- Latest
- Past 24 hours
- Past week
- Past month
- Past 3 months
- Past year
- Custom range...

All results

- Sites with images
- Wonder wheel
- Timeline
- More search tools

Something different

- carl friedrich gauss
- david hilbert
- daniel bernoulli
- isaac newton
- leonard euler

Leonhard Euler - Wikipedia, the free encyclopedia

Leonhard Euler was a pioneering Swiss mathematician and physicist. He made important discoveries in fields as diverse as infinitesimal calculus and graph ...

[Life - Contributions to mathematics ... - Personal philosophy and ...](#)
en.wikipedia.org/wiki/Leonhard_Euler - Cached - Similar

[Euler biography](#)

Leonhard Euler's father was Paul Euler. Paul Euler had studied theology at the University of Basel and had attended Jacob Bernoulli's lectures there. ...
www.groups.dcs.st-and.ac.uk/~history/.../Euler.html - Cached - Similar

[Euler summary](#)

Leonhard Euler (1707-1783) ... Leonhard Euler was a Swiss mathematician who ...
www.history.mcs.st-and.ac.uk/Mathematicians/Euler.html - Cached - Similar

[Show more results from st-and.ac.uk](#)

LEONHARD EULER

Leonhard Euler (1707-1783) was arguably the greatest mathematician of the eighteenth century (His closest competitor for that title is Lagrange) and one of ...
www.usna.edu/Users/math/meh/euler.html - Cached - Similar

[Leonhard Euler \(1707 - 1783\)](#)

Leonhard Euler was born at Båle on April 15, 1707, and died at St. ... Euler wrote an immense number of memoirs on all kinds of mathematical subjects. ...
www.maths.tcd.ie/pub/HistMath/.../Euler/.../RB_Euler.html - Cached - Similar

[Images for Leonhard Euler - Report images](#)

[Leonhard Euler](#)

Swiss mathematician, born at Basel on the 15th of April 1707, his father Paul Euler, who had considerable attainments as a mathematician, being Calvinistic ...
www.nndb.com/people/954/000046810/ - Cached - Similar

[Leonhard Euler](#)

Gives overview of Euler's upbringing, a later appointment to the faculty at St. Petersburg, and his work in fluid mechanics. Page includes list of math ...
www.centennialoflight.gov/essay/.../Euler/Dit144.htm - Cached - Similar

Ads

[Leonhard Euler](#)

Exclusieve limited edition horloges van de bekende merken!
www.qpx.nl

[See your ad here >](#)

A people search engine

The screenshot shows a web browser window with the URL `wieowie.nl/personen?fullName=Leonhard+Euler&country=nl`. The search bar contains the name "Leonhard Euler" and the country is set to "Nederland". A green "ZOEKEN" button is visible.

The search results are displayed under the heading "Leonhard Euler (Nederland)". The interface includes several tabs: "Sociale netwerken", "Zoekmachines", "Foto en video", "Persoonlijke info", and "Acties".

The "Sociale netwerken" tab is active, showing results from various social media and search engines:

- Dizenof Tower**: Rent Luxury Apartment in Tel Aviv. Short-Stay Furnished Apartments. www.dizenof-tower.co.il
- Wieowie profielen**: 0 resultaten gevonden
- Hyves**: 2 resultaten gevonden
- Twitter**: 12 resultaten gevonden
- Facebook**: 2 resultaten gevonden

The Facebook results section shows two entries for "Leonhard Euler":

- Profile 1: Includes a photo of a man, the name "Leonhard Euler", and options to "0 opmerkingen", "Link toevoegen aan profiel", and "Mail dit resultaat naar een vriend". A link to "Bekijk profiel" is provided.
- Profile 2: Includes a photo of a man and the name "Leonhard Euler".

At the bottom of the Facebook results, there is a "results by bing" logo and a link to "Meer resultaten van Facebook".

On the right side of the search results, there are two additional panels:

- Beschikbare domeinen**: A section for checking domain availability. It includes a text box "Stel je online identiteit veilig door nu je domein te claimen!" and a list of domains with their availability status: `leonhardeuler.nl` (beschikbaar), `leonhard-euler.nl` (beschikbaar), `leonhardeuler.com` (bezet), `leonhard-euler.com` (beschikbaar), `leonhardeuler.net` (beschikbaar), `leonhard-euler.net` (beschikbaar), `leonhardeuler.org` (beschikbaar), and `leonhard-euler.org` (beschikbaar).
- Filter met een keyword**: A section for filtering results by keyword. It includes a text box "Filter met een keyword" and a list of keywords: `mathematician`, `one`, `swiss`, `important`, `people`, `basel`, `april`, `petersburg`, `werd`, and `lived`.

Popularity of people search

- One can search for:
 - Film stars
 - Politicians
 - School teachers
 - Friends
 - Yourself
 - ...
- An estimated 11-17 % of web search queries contain a person name (2004 data)

Taxonomy

- Non-famous
- High-profile
 - Event-based
 - Regular

Non-famous



Event-based: “Eli Nikki”



Regular: “Geert Wilders”



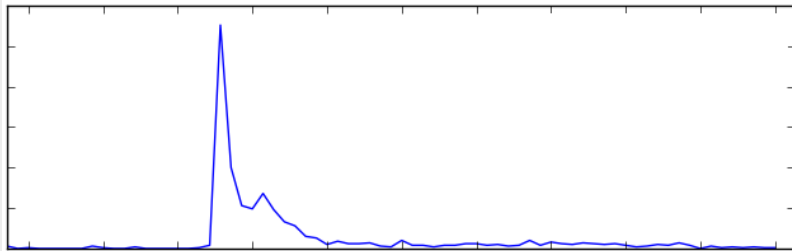
Motivation

- Show different results
- Apply a different ranking algorithm
- Show results in a different way

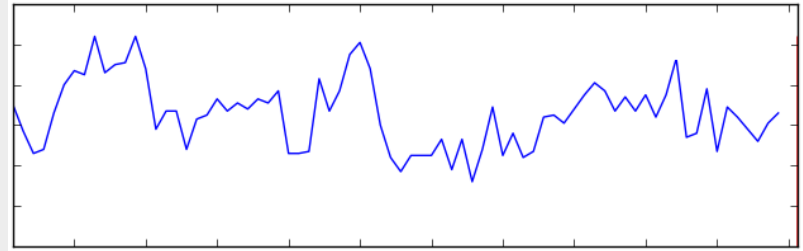
Data

- 4 months from a Dutch people search engine:
 - 13 million searches
 - 4 million clickouts
- Annotated set: 216 total
 - 132 non-famous
 - 60 event-based
 - 24 regular

Features: Attention



Event-based



Regular

Attention = Volume

- Search volume, Clickout volume, News volume
 - Total
 - Last week
 - Trend
 - Burstiness

Other features

- Clickout entropy
 - over urls
 - over top level domains
- Wikipedia presence (Dutch)
 - Title match
 - Overall frequency

Experiment setup

- Two experiments:
 - 2-way: High vs Low
 - 3-way: Regular vs Event vs Low
- Stratified 10-fold cross-validation
- Classifiers:
 - C4.5
 - Naive Bayes
 - SVM

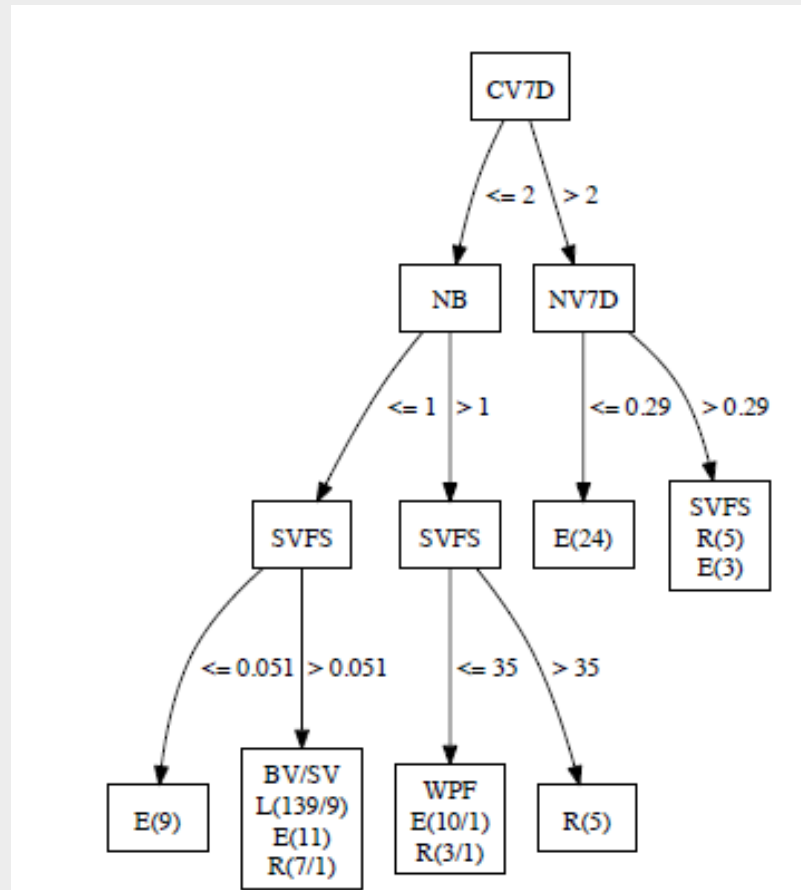
Results: High vs Low

Query Type	C4.5		Naive Bayes		SVM	
	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
High Profile	0.85	0.82	0.89	0.64	0.88	0.60
Low profile	0.89	0.91	0.81	0.95	0.79	0.95

Results: Regular vs Event vs Low

Query Type	C4.5		Naive Bayes		SVM	
	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
Event based	0.83	0.87	0.74	0.62	0.85	0.55
Regular	0.57	0.54	0.53	0.33	0.45	0.38
Low profile	0.92	0.90	0.81	0.92	0.80	0.96

Decision tree C4.5



Error analysis

- Points misclassified – 33 out of 216 (C4.5)
- Annotator agreement – 0.70 (Cohen's kappa)
 - disagreement 40 out of 216 or 19%
- Identified reasons for error for the 15 'hard' points:
 - overlooking of evidence by an annotator
 - ambiguity
 - evidence not strong enough for classifiers
 - small size of annotated set

Future work

- Try to address the ambiguity problem (hard) using:
 - the document collection
 - evidence from the search logs
 - evidence from the click logs
 - external evidence