

PHILIPS

Tool Play Live

dealing with ambiguity in artist similarity

mining from the web

Gijs Geleijnse

Observation

- Web mining for artist classification and similarity give impressive results on dedicated test sets
- But give strange results on real-life collection

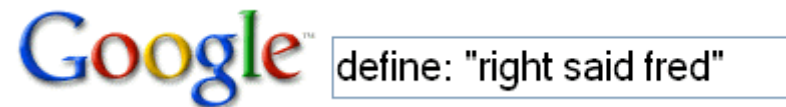
Analysis

- Artist name homonyms without dominate meaning give problems
- Texts on Web hard to parse
- Mutual Information Model needs to be repaired



Solution Approach

We estimate the number of times **Play** refers to the Swedish Band



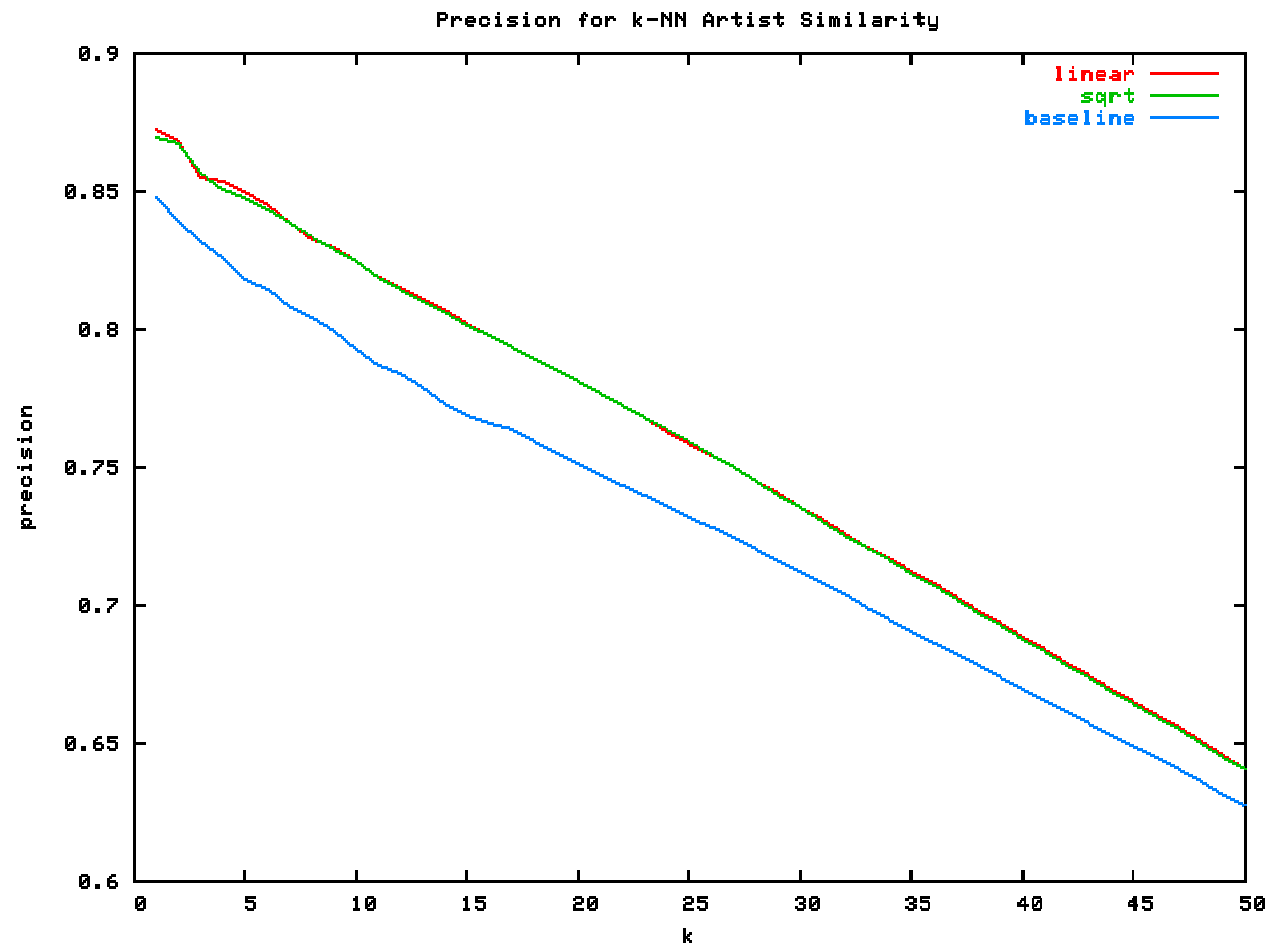
Definitions of **"play"** on the Web:

- participate in games or sport; "We played hockey"
- act or have an effect in a specified way or with development played into her hands"; "I played piano"
- play on an instrument; "The band played all night"
- act: play a role or part; "Gielgud played Hamlet"
- be at play; be engaged in playful activity; amuse her husband's master"
- replay (as a melody); "Play it again, Sam"; "She played the record"
- perform music on (a musical instrument); "He played the guitar"
- act: pretend to have certain qualities or state of mind"
- move or seem to move quickly, lightly, or irregularly"
- bet or wager (money); "He played \$20 on the roulette"

Definitions of **"right said fred"** on the Web:

- Right Said Fred are a British pop group made up of Bernard Cribbins. Fairbass has previously found success with the band. See en.wikipedia.org/wiki/Right_Said_Fred

Experiments



Babylon Zoo:	B12:
Tool	Tool
Live	Live
Fish	Fish
Juli	Juli
Chumbawamba	Play
Play	Japan

Babylon Zoo:	B12:
Juli	Juli
Chumbawamba	Carl Craig
Jamiroquai	Jamiroquai
Shakira	Autechre
Sonic Youth	Shakira
Right Said Fred	Speedy J.

Conclusions

- There's work to be done in Web MIR!
- Outstanding results on the evaluation sets are misleading
- Need for more realistic evaluation set for Web MIR

