

Vertical Selection for Heterogeneous Search Engine Result Pages

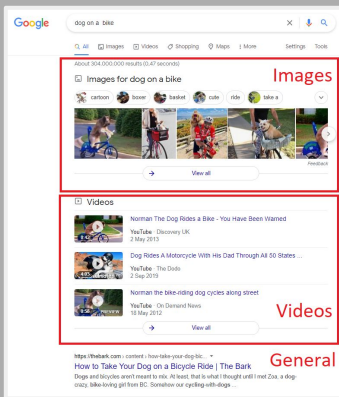
Responsible professor and Supervisor: Claudia Hauff

Augustas Vilčinskas

a.vilcinskas@student.tudelft.nl

Goal

When searching the web, sometimes results for queries should include more than just a list of webpages.



1

Research questions

- (1) How can each vertical be identified using the content inside it?
- (2) How to make a vertical selection for any query?
- (3) How to display the selected verticals in a SERP?

2

Background

Vertical - a single category of search engine results. For example, news vertical only containing news, blog verticals that contains only blogs.

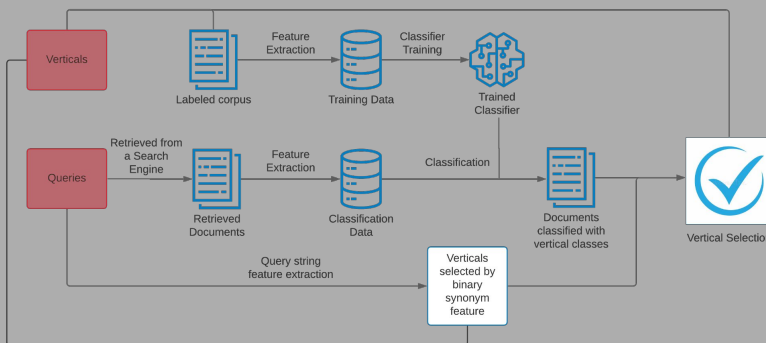
Heterogeneous Search Result Pages - Result pages, that allow users to see results within multiple verticals of results in one page.

Combining verticals has been proven to be **beneficial** in the work of Bron et al. [1]. This motivates use of heterogeneous search engine result pages in search engines.

Verticals and **queries** used for evaluation and implementation were taken from the 2014 TREC [2] dataset, the federated web search track.

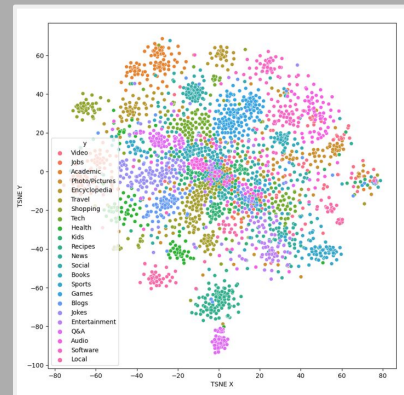
3

Approach



4

Results



Vertical document class visualization

Classifier	F-Score
Random Forest	0.4921
XGBoost	0.4814

Evaluation results

5

Future Work

From document class visualization above, it can be seen that certain verticals are unidentifiable. Future work should focus on improving the corpus retrieval and labelling approach.

6