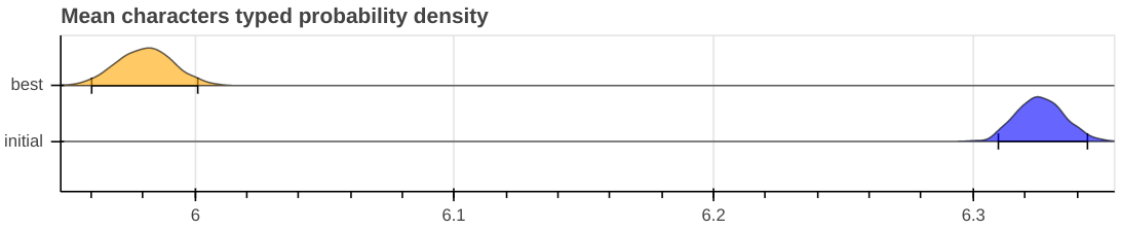
Wikidata entity completion - item / fr

Metric used

The metric used here to compare different tuning of the autocomplete algorithm represents the probablistic number of characters typed by a typical user before selecting their desired result from the autocomplete drop down. This metric first looks at real user sessions to estimate how likely a user is to continue typing even when their result is presented in the autocomplete, conditioned on the position the result is displayed at. Individual user sessions, represented in the data as (prefix_typed, page_id_clicked), are then simulated with prefixes from length 1 to the full prefix typed. From this simulation we determine the expected number of characters typed for an individual search clickthrough.

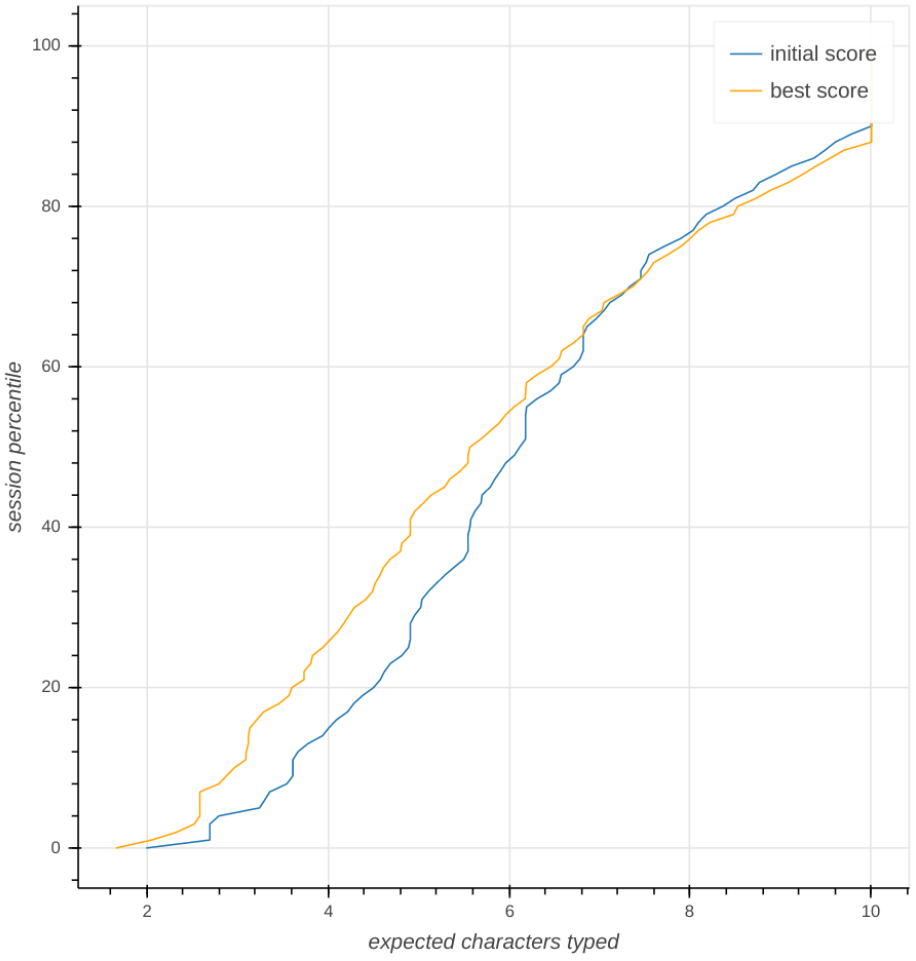
The dataset used contains 50k clickthroughs from oct 7 - dec 20 in the training set, and another 50k clicks from dec 20 - jan 7 in the test set.

The graph below shows bootstrapped probability densities for each bucket. Tick marks are shown at the 95% confidence levels.



Mean characters typed by percentile

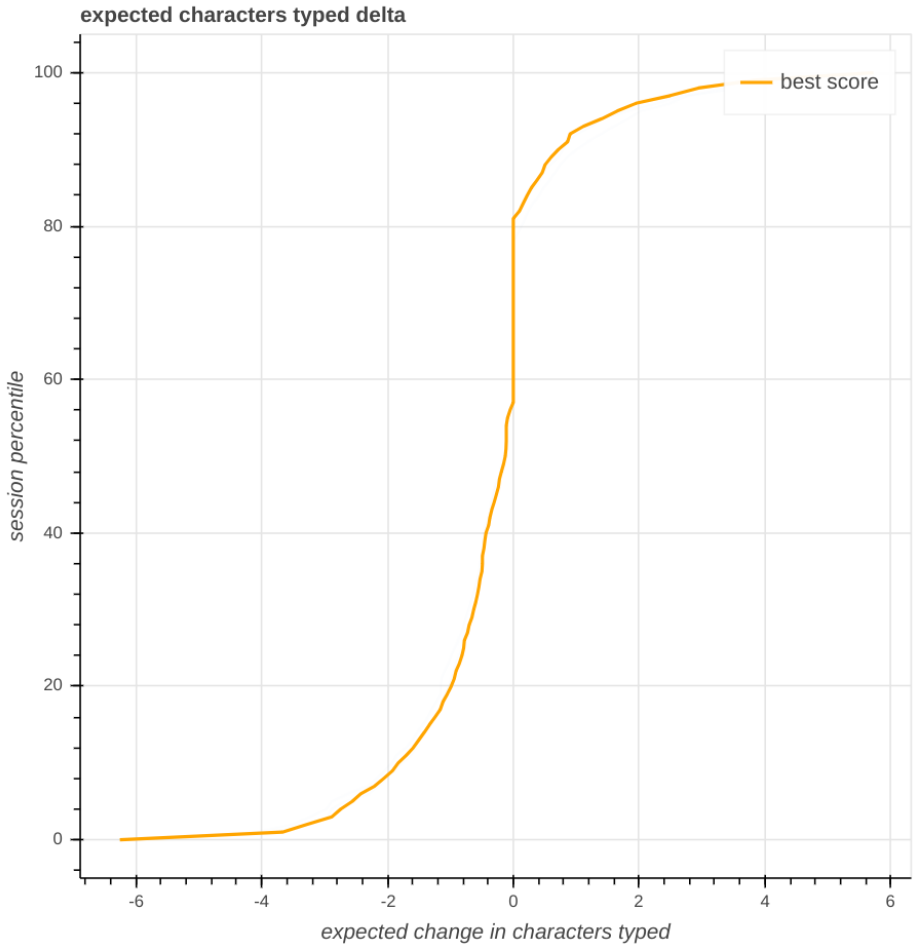
The following graphs show the before and after effects of tuning. This shows strong improvement, up to a full character, from percentiles 0-65. Tail queries show some decline in performance, but only slightly.



Per-session delta in expected characters typed

This shows on a per-session basis the change in number of characters typed between the baseline scoring and the scoring after parameter tuning. This suggests up to 20% of sessions save between 1 and 6 characters typed. Another 35% save a fractional character, and 20% have no impact. Around 20% of sessions are impacted negatively, future inspection of what makes these sessions different may be useful for investigating new scoring signals.

The orange line shows the training run with the best score. The faint blue lines show other training runs that have a mean score less than the upper 95% CI of the best score.



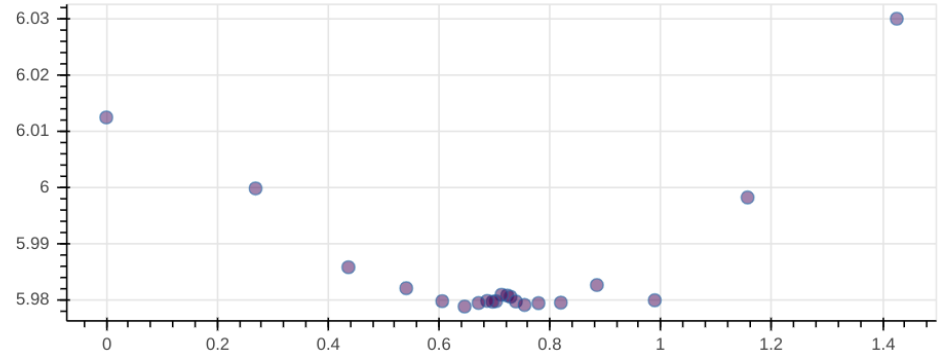
Final Tuned Values

```
  0.71 : query/dismax/0/constant_score/labels.fr.near_match/boost:0
  0.48 : query/dismax/1/constant_score/labels.fr.near_match_folded/boost:0
  0.97 : query/dismax/2/constant_score/labels.fr.prefix/boost:0
  0.08 : query/dismax/3/constant_score/labels_all.near_match_folded/boost:0
  0.47 : query/dismax/4/constant_score/labels.en.near_match/boost:0
  0.60 : query/dismax/5/constant_score/labels.en.near_match_folded/boost:0
  0.75 : query/dismax/6/constant_score/labels.en.prefix/boost:0
  0.50 : query/dismax/tie_breaker:0
  0.60 : query/title.keyword/boost:0
  0.36 : query/title.keyword/k1:0
  0.34 : rescore/0/function_score/0/satu/incoming_links/a:0
 54.16 : rescore/0/function_score/0/satu/incoming_links/k:0
  0.77 : rescore/0/function_score/0/weight:0
  0.29 : rescore/0/function_score/1/satu/sitelink_count/a:0
223.80 : rescore/0/function_score/1/satu/sitelink_count/k:0
  0.32 : rescore/0/function_score/1/weight:0
  0.14 : rescore/0/query_weight:0
  0.53 : rescore/0/rescore_query_weight:0
```

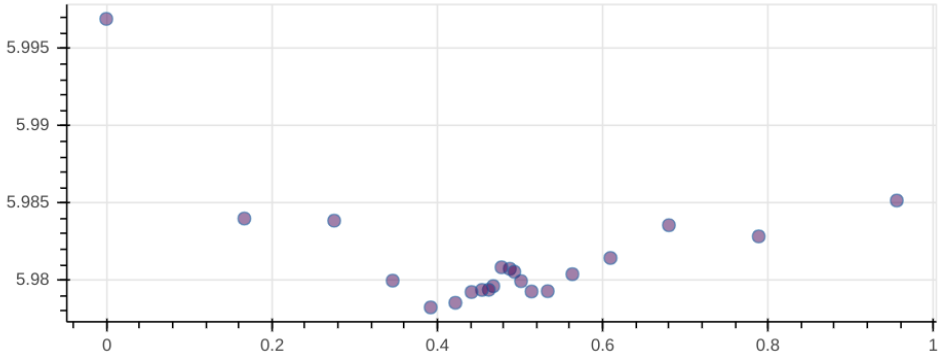
Sensitivity of chosen parameters

To get an idea of how much influence individual parameters have on the final score, and to estimate how sensitive those variables are to small changes, the graphs below plot the sensitivity of individual parameters. This is performed by holding all variables except one as a constant, and sweeping a set of values around the chosen point. The graphs then show how the final score changes based on changes to that variable. Dots on the graphs are additionally colored by their score. A graph of a single color suggests the variable in question has a relatively small influence on the final output.

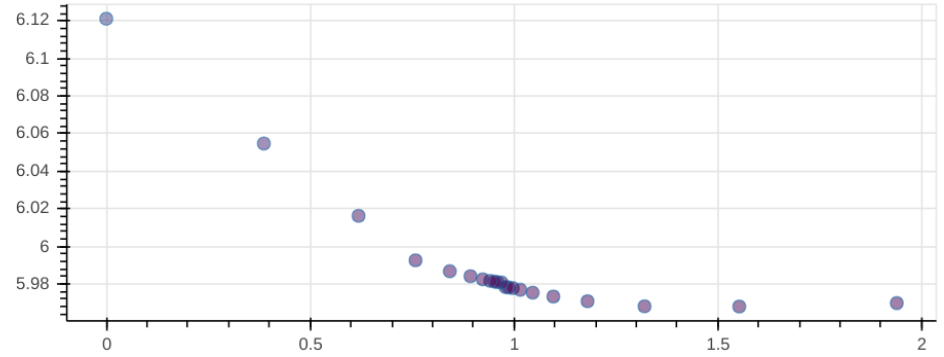
query/dismax/0/constant_score/labels.fr.near_match/boost:0



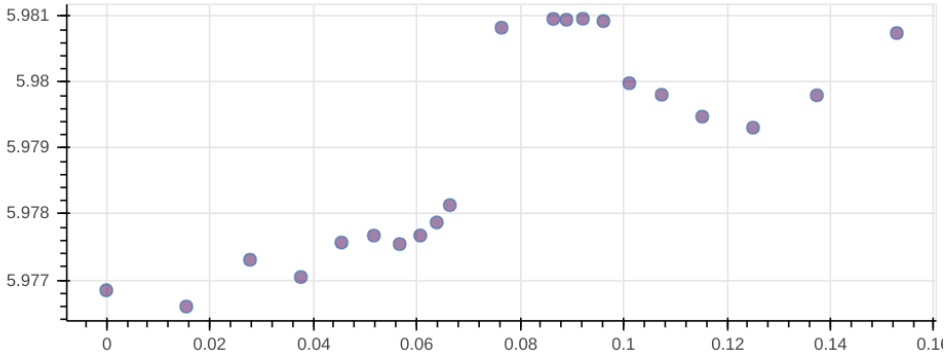
query/dismax/1/constant_score/labels.fr.near_match_folded/boost:0



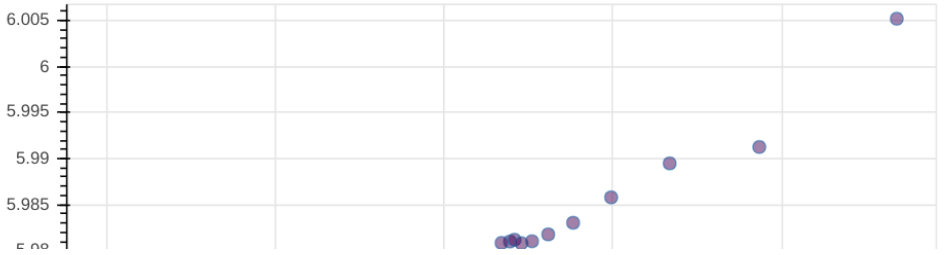
query/dismax/2/constant_score/labels.fr.prefix/boost:0

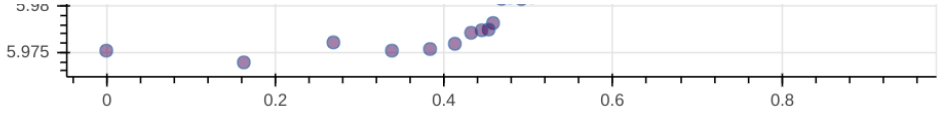


query/dismax/3/constant_score/labels_all.near_match_folded/boost:0

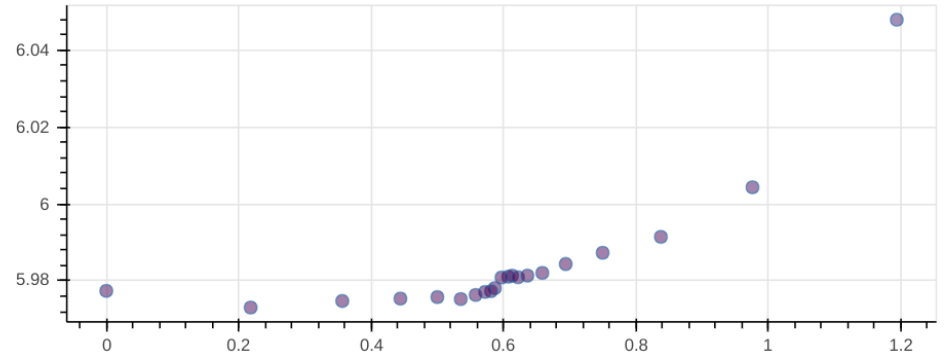


query/dismax/4/constant_score/labels.en.near_match/boost:0

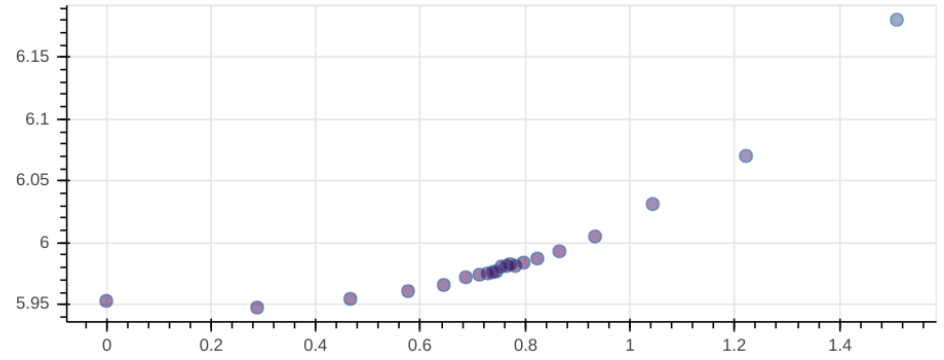




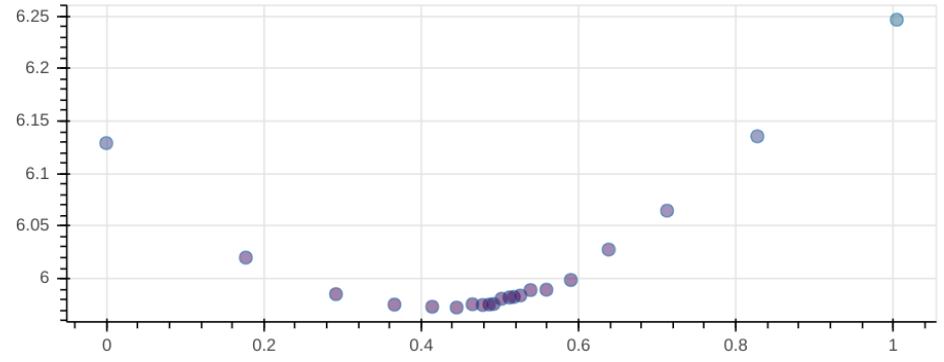
query/dismax/5/constant_score/labels.en.near_match_folded/boost:0



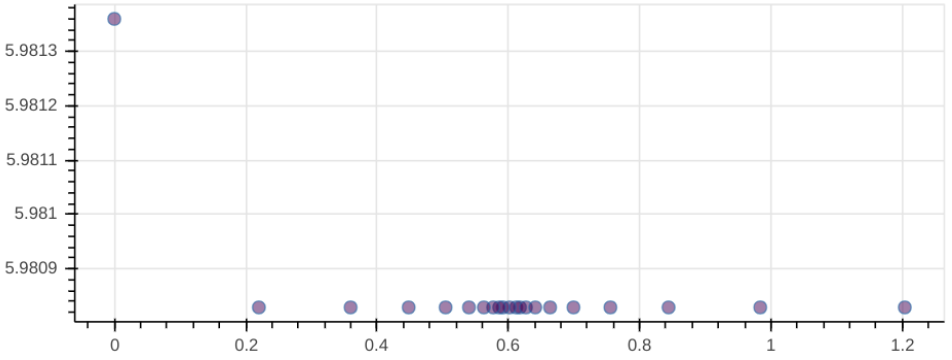
query/dismax/6/constant_score/labels.en.prefix/boost:0



query/dismax/tie_breaker:0

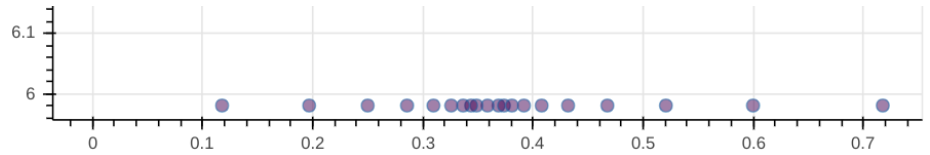


query/title.keyword/boost:0

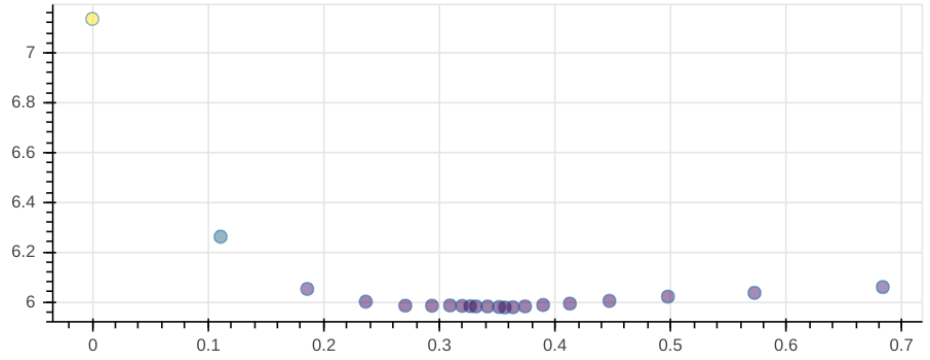


query/title.keyword/k1:0

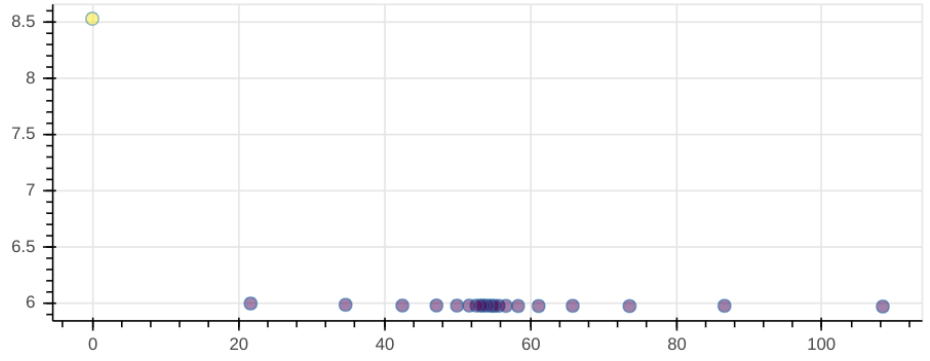




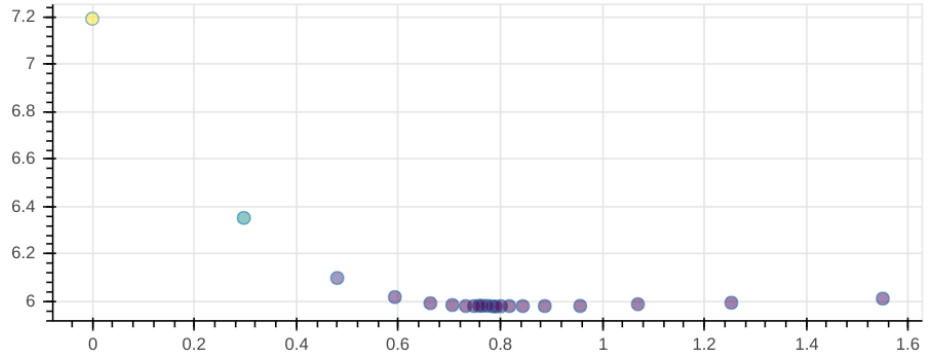
rescore/0/function_score/0/satu/incoming_links/a:0



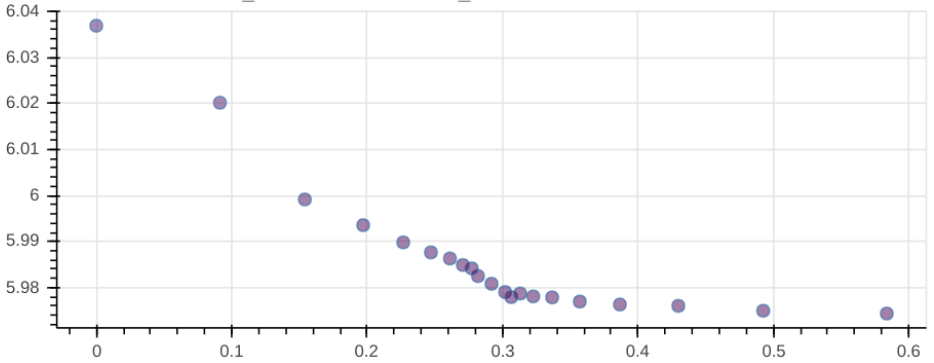
rescore/0/function_score/0/satu/incoming_links/k:0



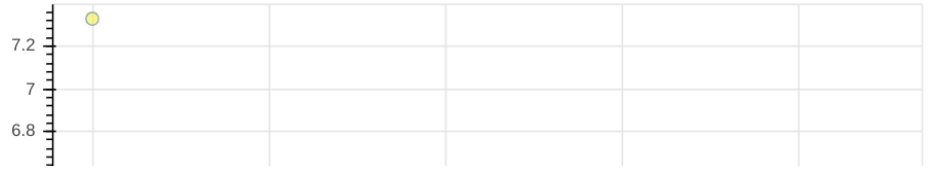
rescore/0/function_score/0/weight:0

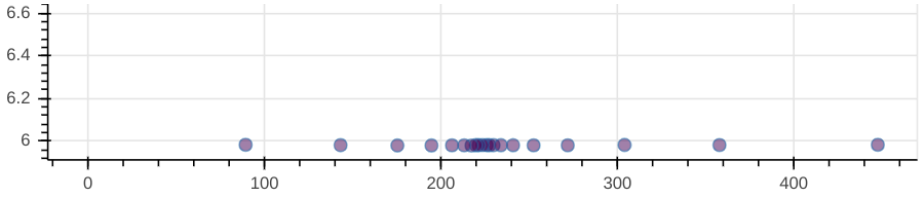


rescore/0/function_score/1/satu/sitelink_count/a:0

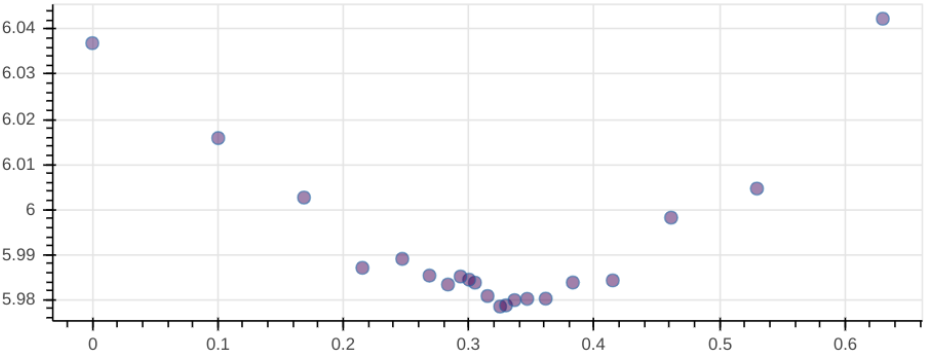


rescore/0/function_score/1/satu/sitelink_count/k:0

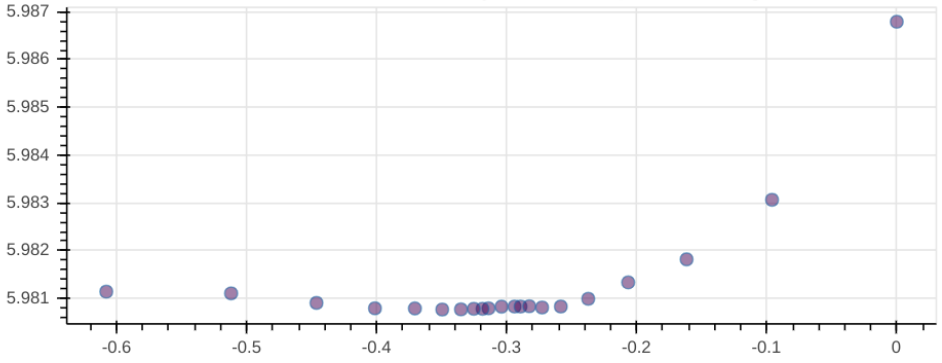




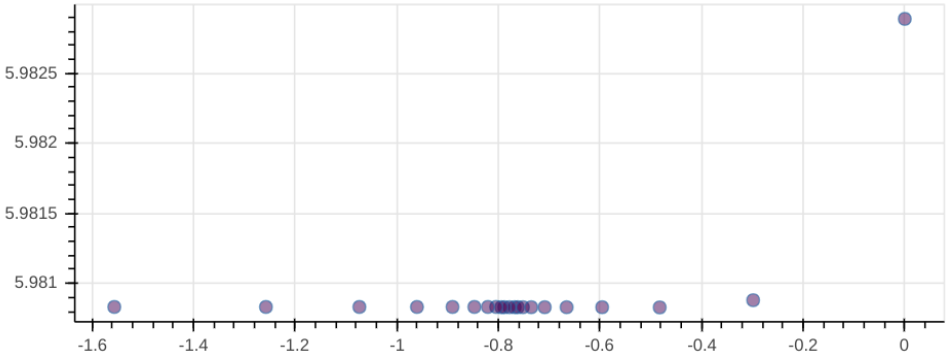
rescore/0/function_score/1/weight:0



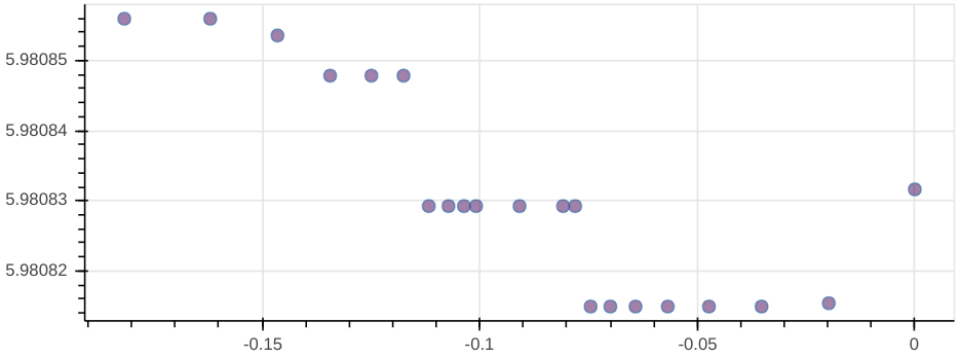
rescore/0/function_score/2/statement_keywords-P31-Q4167410/weight:0



rescore/0/function_score/3/statement_keywords-P31-Q13442814/weight:0

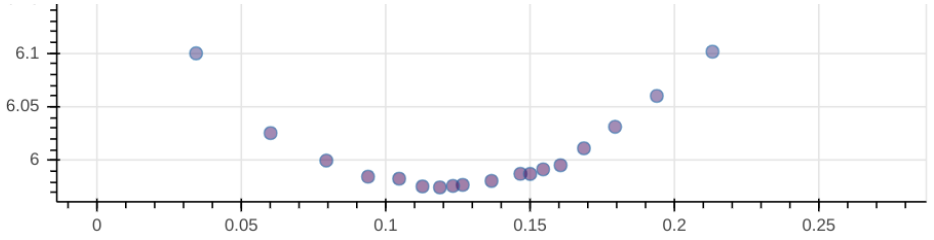


rescore/0/function_score/4/statement_keywords-P31-Q18918145/weight:0



rescore/0/query_weight:0





rescore/0/rescore_query_weight:0

