

Konzeption und Entwicklung eines Systems zur kontinuierlichen Evaluation und Verbesserung der Sortierung von Suchergebnissen am Beispiel einer Usenet-Suche

Thema:

Konzeption und Entwicklung eines Systems zur kontinuierlichen Evaluation und Verbesserung der Sortierung von Suchergebnissen am Beispiel einer Usenet-Suche

Art:

[MA](#)

BetreuerIn:

[Prof. Christian Wolff](#)

BearbeiterIn:

Johannes Wölfel

ErstgutachterIn:

[Christian Wolff](#)

ZweitgutachterIn:

[N.N.](#)

Status:

[abgeschlossen](#)

Stichworte:

[Evaluation; Suchergebnisse; Usenet](#)

angelegt:

2017-07-04

Beginn:

2017-07-04

Anmeldung:

2019-09-13

Antrittsvortrag:

2017-07-24

Abschlussvortrag:

2019-09-24

Abgabe:

2017-09-30

Textlizenz:

[Unbekannt](#)

Codelizenz:

[Unbekannt](#)

Hintergrund

Nach dem Aufbau eines Systems zur Indexierung und Durchsuchung des Usenet, mit maßgeschneiderter und konfigurierbarer Sortierfunktion, soll es ermöglicht werden, eine Konfiguration zu finden, welche die Suchergebnisse optimal sortiert. Hierzu wird ein System benötigt, in dem neue Konfigurationen angelegt werden können und das es ermöglicht, die Qualität der Ergebnissortierung verschiedener Konfigurationen untereinander objektiv zu vergleichen.

Zielsetzung der Arbeit

Um Sortierungen objektiv vergleichbar zu machen, muss eine Funktion gefunden werden, welche die Qualität der Sortierung objektiv wiedergibt. Die Funktion muss entsprechend der domänenspezifischen Dokumenteneigenschaften und der gewünschten Prioritäten in der Sortierung gewählt bzw. angepasst werden um die Qualität der Sortierung optimal zu repräsentieren. Es muss den Entwicklern ermöglicht werden, Konfigurationen am Beispiel verschiedener Suchanfragen vergleichen zu können. Im Anschluss muss eine gefundene Konfiguration im Betrieb zusätzlich evaluiert werden können, um deren tatsächlichen Effekte messbar zu machen. Die abschließende Evaluation auf Basis von Nutzungsdaten soll Aufschluss darüber geben, ob eine offline Evaluation mit der gewählten Metrik repräsentative Vorhersagen über die tatsächliche Rankingqualität zulässt.

Konkrete Aufgaben

Es sollen ein System (Desktopanwendung) erarbeitet werden, das es Entwicklern ermöglicht, exemplarische Suchanfragen anzulegen, Rankingkonfigurationen zu erstellen, zu speichern und zu laden. Die Qualität der Ergebnissortierung der Suchanfragen soll darin über die berechnete Kennzahl repräsentiert werden. Mit Hilfe der Anwendung sollen Konfigurationen manuell optimiert werden können. Nachdem eine Konfiguration gefunden wurde, soll diese im Produktionsbetrieb eingesetzt werden und, zum Beispiel durch einen A/B-Test mit erhobener impliziten Relevanzbewertung (z.B. durch Klick) und explizitem Feedback (z.B. Positives / negatives Bewerten), evaluiert werden. Dazu muss entweder eine Nutzeranwendung, oder die Serveranwendung um die Funktionen der Datenaufzeichnung erweitert werden, und anschließend die Daten mit geeigneten Metriken ausgewertet und mit den Ergebnissen der offline Evaluation verglichen werden.

Erwartete Vorkenntnisse

tbd

Weiterführende Quellen

Clarke, C. L. A., Kolla, M., Cormack, G. V., Vechtomova, O., Ashkan, A., Büttcher, S., & MacKinnon, I. (2008). Novelty and diversity in information retrieval evaluation. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '08 (p. 659). New York, New York, USA: ACM Press. <https://doi.org/10.1145/1390334.1390446>

Schuth, A., Hofmann, K., & Radlinski, F. (2015). Predicting Search Satisfaction Metrics with Interleaved Comparisons. Sigir '15, 463–472. <https://doi.org/10.1145/2766462.2767695>

Carterette, B., & Ben. (2015). Statistical Significance Testing in Information Retrieval. In Proceedings

of the 2015 International Conference on Theory of Information Retrieval - ICTIR '15 (pp. 7–9). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2808194.2809445>

Wang, Y., Wang, L., Li, Y., He, D., Chen, W., & Liu, T.-Y. (2013). A Theoretical Analysis of NDCG Ranking Measures. In Proceedings of the 26th Annual Conference on Learning Theory (pp. 1– 30)

Moniz, N., Torgo, L., & Vinagre, J. (2016). Data-Driven Relevance Judgments for Ranking Evaluation, (December). Retrieved from <http://arxiv.org/abs/1612.06136>

Kelly, D. (2009). Methods for Evaluating Interactive Information Retrieval Systems with Users. Foundations and Trends R in Information Retrieval, 3, 1–2. <https://doi.org/10.1561/1500000012>

From:
<https://wiki.mi.ur.de/> - MI Wiki

Permanent link:
https://wiki.mi.ur.de/arbeiten/erarbeitung_und_entwicklung_eines_systems_zur_kontinuierlichen_evaluation_und_verbesserung_der_sortierung_von_suchergebnissen_am_beispiel_einer_usenet_suche

Last update: 16.09.2020 10:25

