Exploring Document Features Influencing Short- and Long-term Vocabulary Learning Outcomes

Rohail Syed and Kevyn Collins-Thompson

School of Information, University of Michigan

Abstract

We know people use Web search to learn but we don't know what **types** of documents are best suited for such goals.

We train regression models on vocabulary learning outcomes using document/user features. Our models show strong ability to predict learning outcomes even when using only document features.

We further conduct the first longitudinal crowdsourced study of learning from Web documents. Our retrieval model outperforms Google search results by over **92%** in long-term retention of learning gains.

Acknowledgements

MICHIGAN INSTITUTE FOR DATA SCIENCE UNIVERSITY OF MICHIGAN

Paper Download ⬇

Introduction

Web search engines today still optimize for generic relevance, but increasingly more people use Web search for learning. There is a need to understand the types of Web documents better are suited for optimizing learning goals.

To address this, we trained regression models on learning outcomes using data **DS** from an earlier study of learning from the Web. The predictors were all document features.
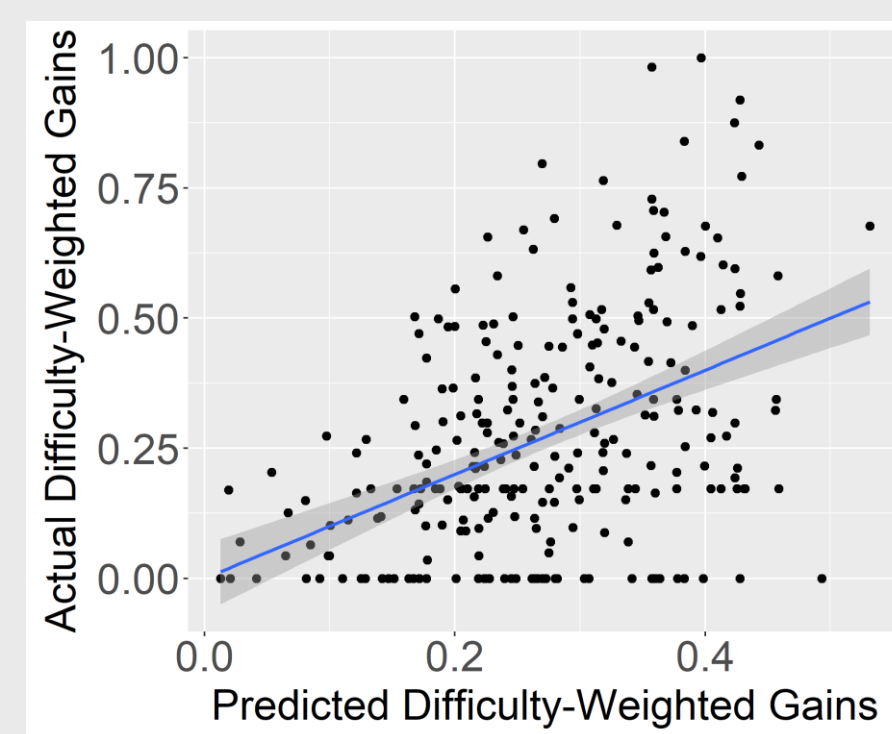
The retrieval model used in **DS** was evaluated for short-term learning gains. But what about long-term retention of those gains? We conducted a longitudinal test (after nine months) to evaluate their retention of gains.

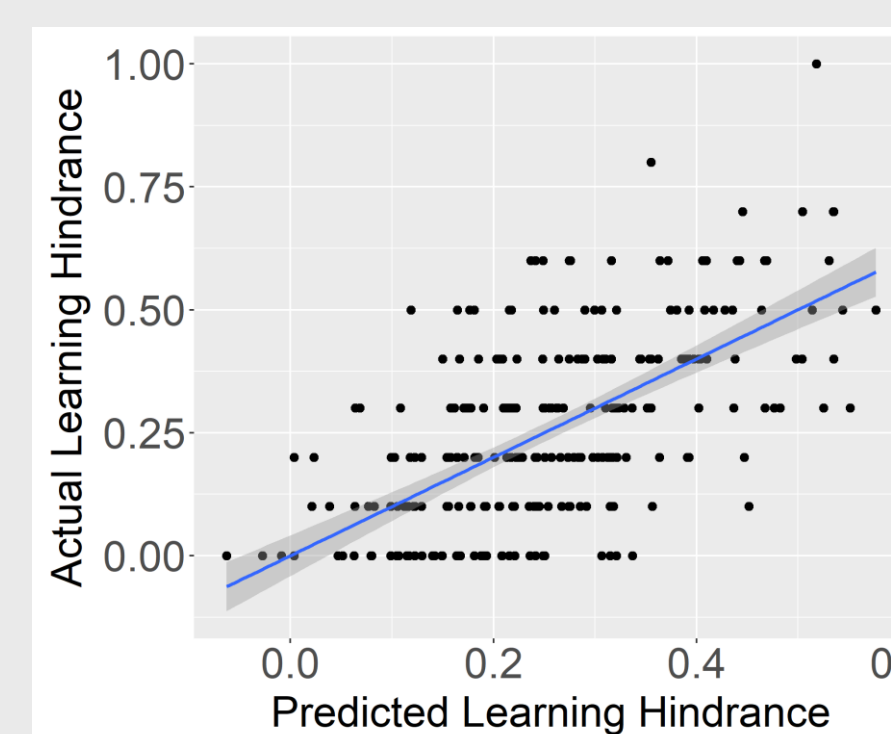We address two fundamental Research Questions:

- **RQ 1.** What document and/or user features predict short-term learning outcomes?

- **RQ 2.** Does personalized vocabulary-centric retrieval algorithm show better **long-term** retention of gains than a commercial baseline ("Google")?

RQ1: Document Features are Strong Predictors of Learning

- Common **positive features** for multiple learning measures:
  - Use more images (excluding ads and navigational)
  - Document sets should have longer paragraphs.
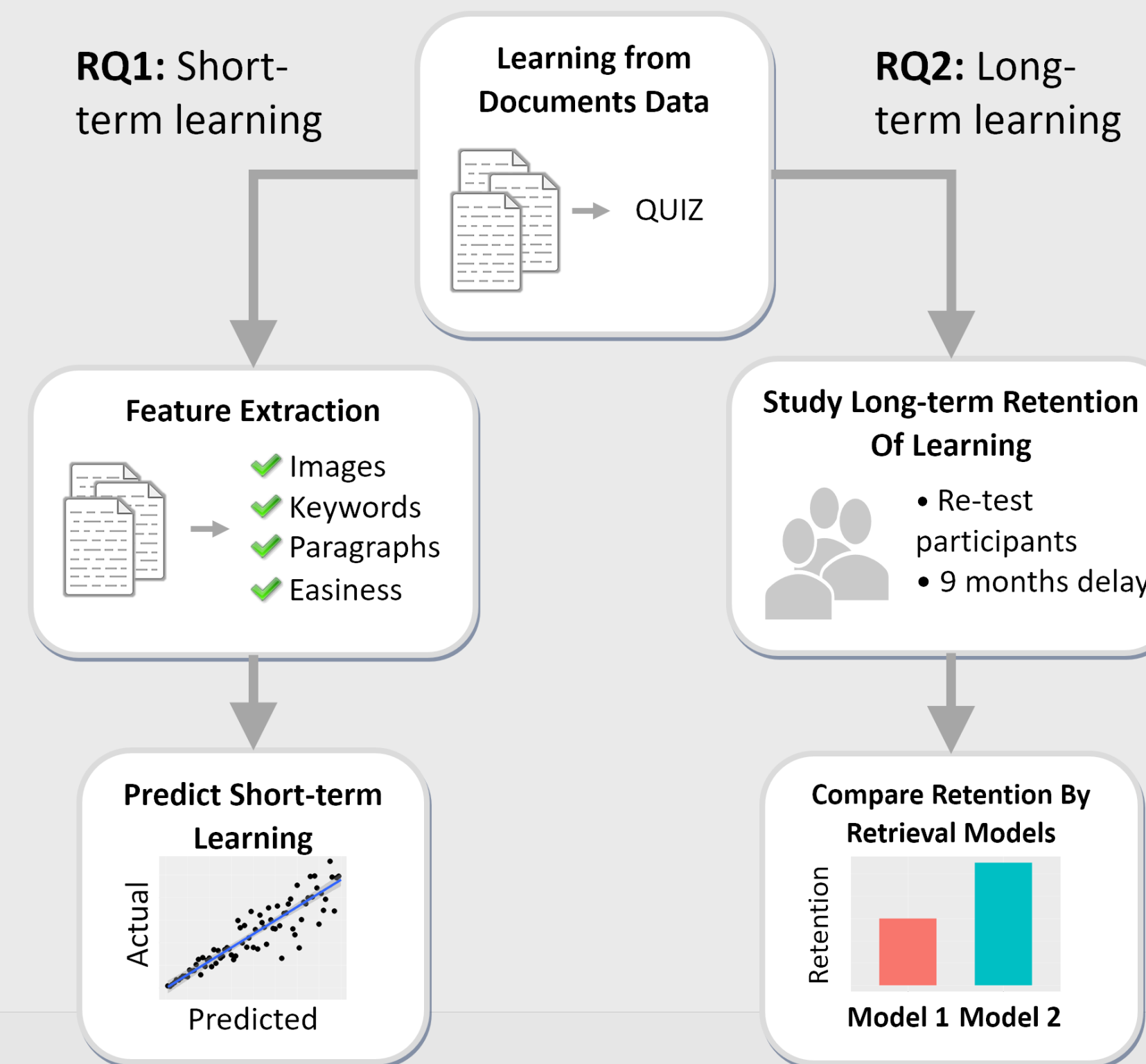 - Document sets should have more coverage of relevant keywords compared to other words.

Difficulty-Weighted Gains
(r=.463, p<.001)

Learning Hindrance
(r=.643, p<.001)

Methodology

RQ1: Short-term learning

Learning from Documents Data → QUIZ

RQ2: Long-term learning

Feature Extraction
- ✔ Images
- ✔ Keywords
- ✔ Paragraphs
- ✔ Easiness

Study Long-term Retention Of Learning
- Re-test participants
- 9 months delay

Predict Short-term Learning
Actual / Predicted

Compare Retention By Retrieval Models
Retention / Model 1 Model 2

RQ2: Personalization Improves Long-term Retention

- Our personalized retrieval model (**P**) showed **92%** better long-term retention of gains (for high-difficulty vocabulary) compared to Google baseline (**Web**)
  - **Retained Gains**: # words learned 9 months prior and still known.
  - **Robust Gains**: # words unknown 9 months prior and now known.
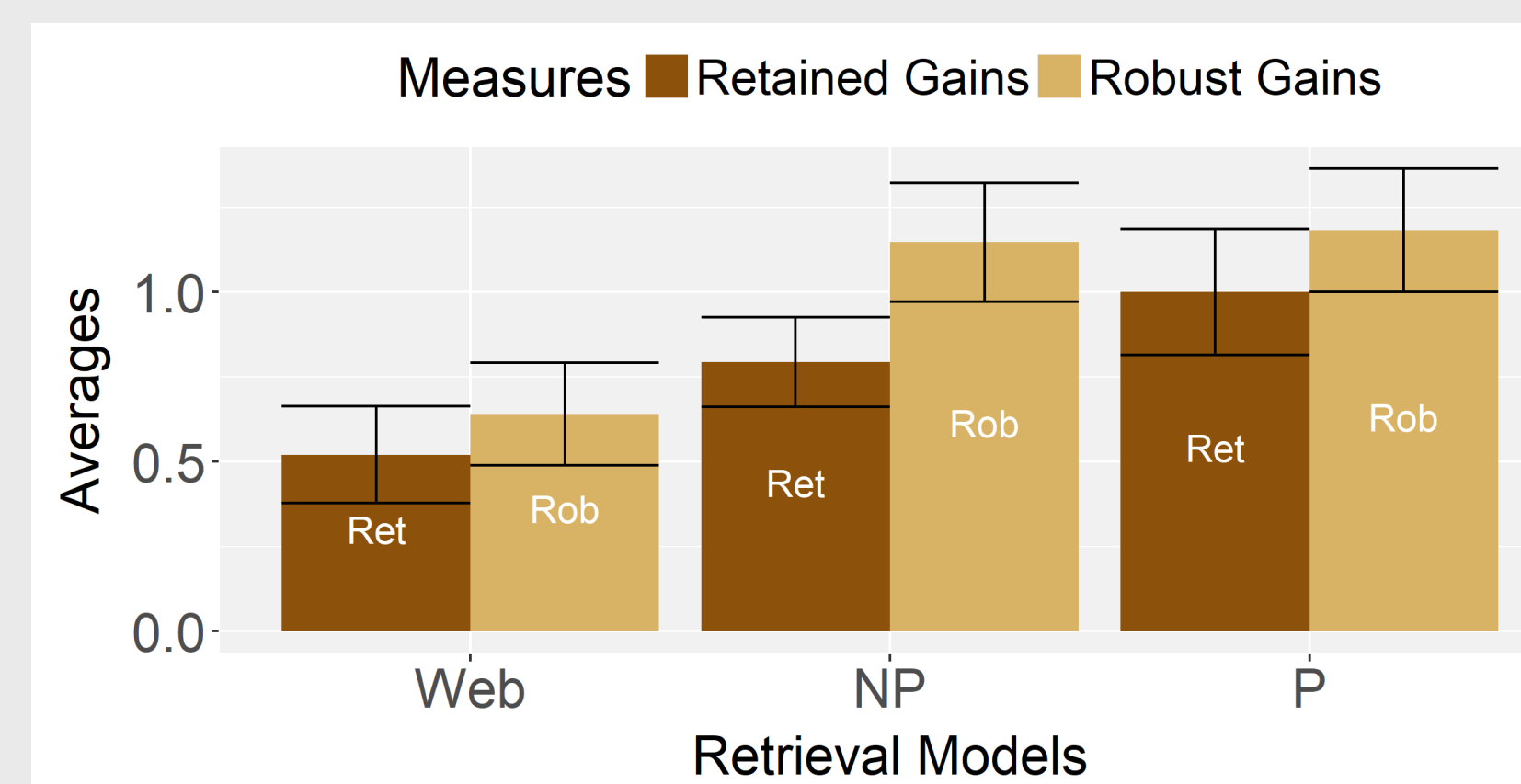
Discussion

In evaluating our two Research Questions, we found the following general results:

RQ 1 Supported: We were able to fit models with/without user-specific features that **showed strong predictive power**. Trained on multiple types of learning outcomes.

RQ 2 Supported: Strong evidence that personalized retrieval model yields far better **long-term learning retention.**

Applications
- Possible integration with existing search systems to better serve queries with learning intents.
- Intelligent tutoring systems can use trained models to predict what learning material to provide students.
- May guide website designers to structure and write content that may improve learning outcomes.

Conclusions

1. Personalized retrieval model could outperform commercial baseline in long-term retention of knowledge.
2. Regression models showed strong ability to predict short-term learning, even without user-specific features.

References

1. Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. 2014. Lessons from the journey: a query log analysis of within-session learning. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining* (WSDM '14). ACM, New York, NY, USA, 223-232.

2. Victor Kuperman, Hans Stadthagen-Gonzalez, and Marc Brysbaert. 2012. Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods* 44, 4 (2012), 978–990.

3. Peter Pirolli and Sanjay Kairam. 2013. A knowledge-tracing model of learning from a social tagging system. *User Modeling and User-Adapted Interaction* 23, 2-3 (2013), 139–168.

4. Rohail Syed and Kevyn Collins-Thompson. 2017. Retrieval Algorithms Optimized for Human Learning. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR '17). ACM, New York, NY, USA, 555-564.