



Looking for the Movie Seven or Sven from the Movie Frozen?

A Multi-perspective Strategy for Recommending Queries for Children

ION MADRAZO AZPIAZU, NEVENA DRAGOVIC, OGHENEMARO ANUYAH & SOLE PERA

Department of Computer Science

Boise State University

Boise, Idaho, USA







Research supported in part by National Science Foundation, Award # 1565937.

Background: Who is Sven?



Let's search for *sven*



[All](#) [Images](#) [Videos](#) [News](#) [Maps](#) [More](#)

[Settings](#) [Tools](#)

Search by voice

About 318,000,000 results (0.67 seconds)

Seven (1995 film) - Wikipedia

[https://en.wikipedia.org/wiki/Seven_\(1995_film\)](https://en.wikipedia.org/wiki/Seven_(1995_film)) ▼

Seven (stylized as SE7EN) is a 1995 American neo-noir crime thriller film directed by David Fincher and written by Andrew Kevin Walker. It stars Brad Pitt, Morgan Freeman, Gwyneth Paltrow, John C. McGinley, R. Lee Ermy, and Kevin Spacey. It tells the story of David Mills (Pitt), a detective who partners with the retiring ...

Music by: [Howard Shore](#) **Budget:** \$33 million
Produced by: [Arnold Kopelson](#); [Phyllis Carlyle](#) **Edited by:** [Richard Francis-Bruce](#)
[Andrew Kevin Walker](#) · [Darius Khondji](#) · [Comic book](#) · [Seven \(1979 film\)](#)

Se7en (1995) - IMDb

www.imdb.com/title/tt0114369/ ▼

★★★★★ Rating: 8.6/10 - 1,172,553 votes

Crime · Two detectives, a rookie and a veteran, hunt a serial killer who uses the **seven** deadly sins as his motives.

[Se7en \(1995\)](#) · [Parents Guide](#) · [Trivia](#) · [User Reviews](#)



Se7en (1995) Official Trailer - Brad Pitt, Morgan Freeman Movie HD ...


<https://www.youtube.com/watch?v=znmZoVkJpl> ▼


Feb 18, 2014 - Uploaded by Movieclips Trailer Vault

Se7en (1995) Official Trailer - Brad Pitt, Morgan Freeman Movie HD Subscribe to CLASSIC TRAILERS: <http://bit...>



Seven



 1995 · Drama/Crime film · 2h 8m

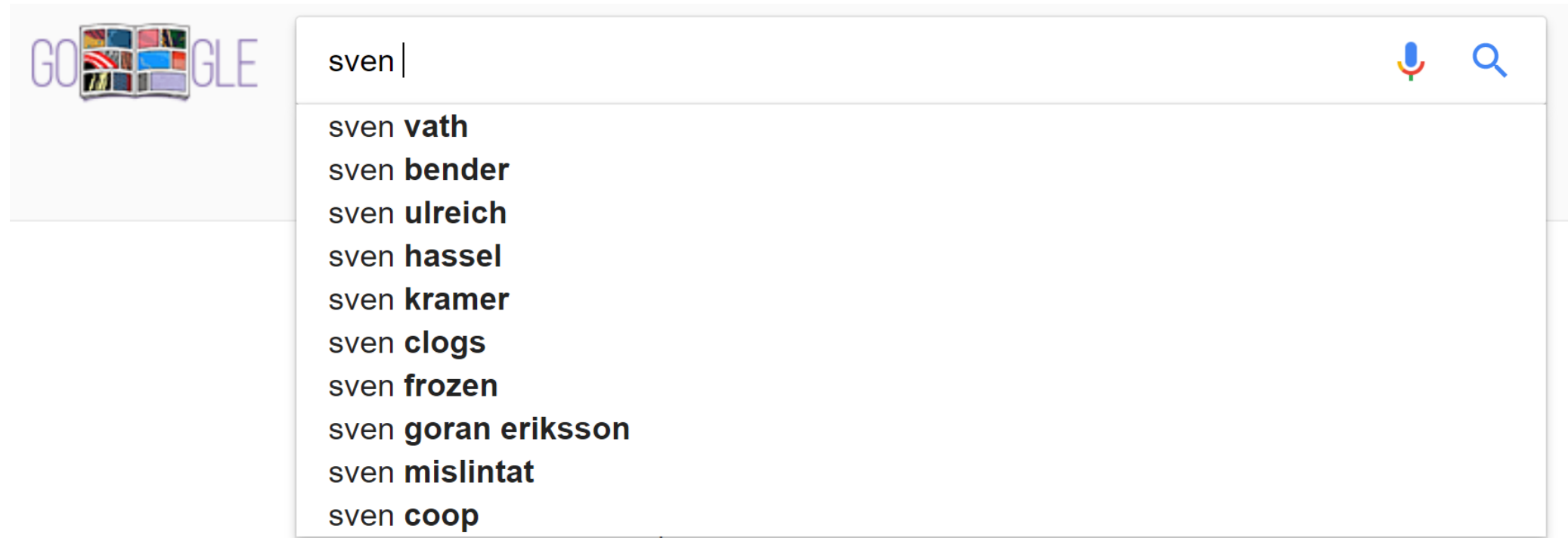
 [Play trailer on YouTube](#)

8.6/10 IMDb	65% Metacritic	80% Rotten Tomatoes
----------------	-------------------	------------------------

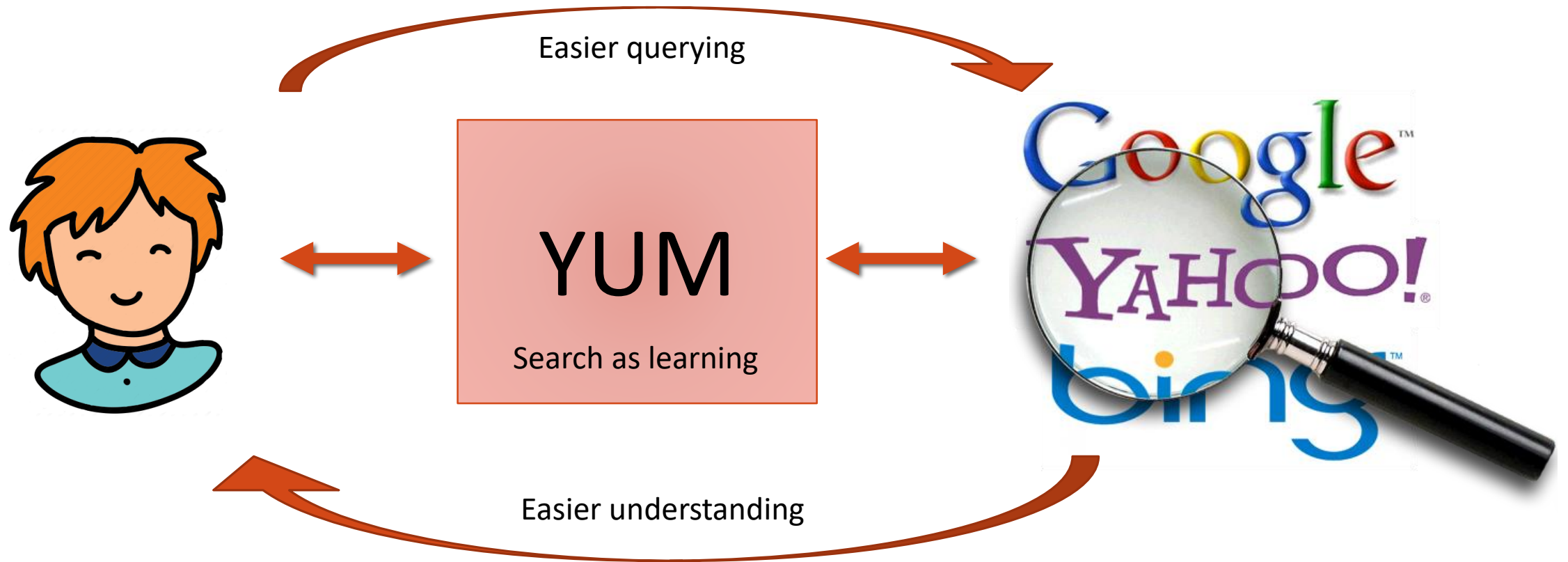
88% liked this movie
Google users  

When retiring police Detective William Somerset (Morgan Freeman) tackles a final case with the aid of newly transferred David Mills (Brad Pitt), they discover a number of elaborate and grizzly murders. They soon realize they are dealing with a serial killer (Kevin Spacey) who is targeting people he ... [MORE](#) ▼

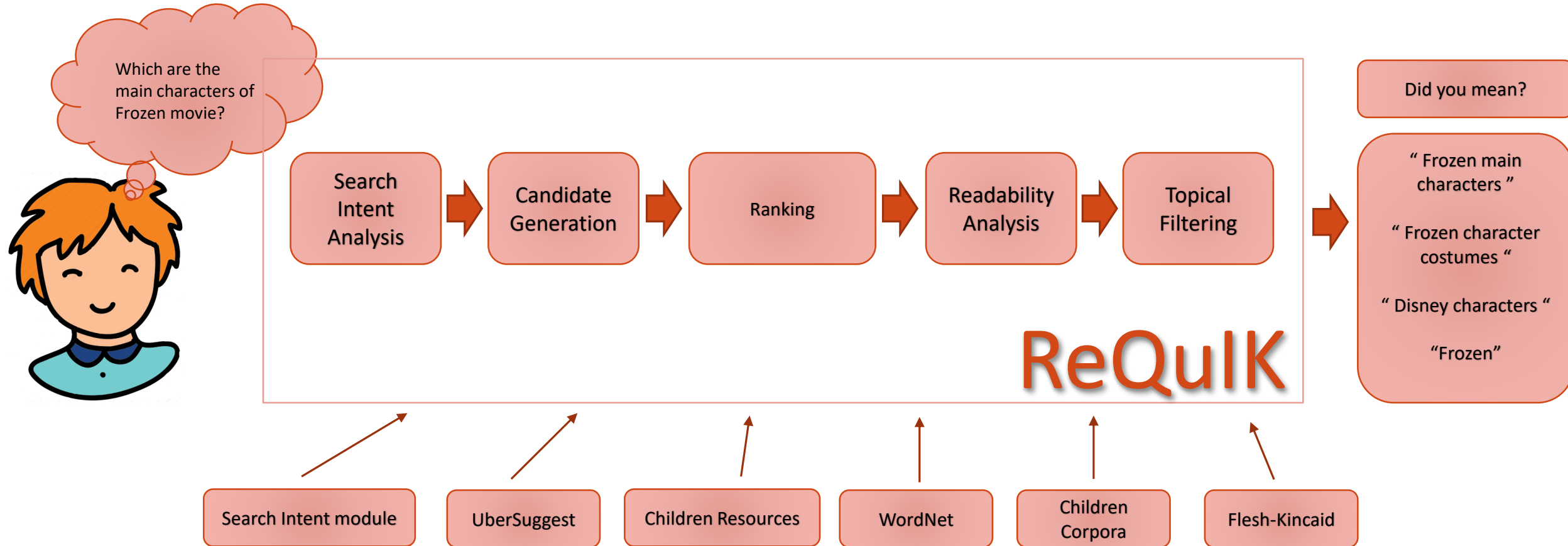
Do query suggestions help?



YUM



Proposed method



Search intent identification

Initial Query	Sven the amaaazing raindeer of Frozen
Stop-word removal	Sven ama a azing r a indeer Frozen
Spelling for children	Sven amazing reindeer Frozen
Children's relevant term detection	Sven reindeer Frozen
Final Query	Sven reindeer Frozen

QuIK

“Is Sven Seven?” A Search Intent Module for Children, Nevena Dragovic, Ion Madrazo Azpiazu, Maria Soledad Pera, SIGIR 2016

Candidate generation

Sven a

- Sven and ole
- Sven and olaf

Sven b

- Sven berg
- Sven buerki

Sven c

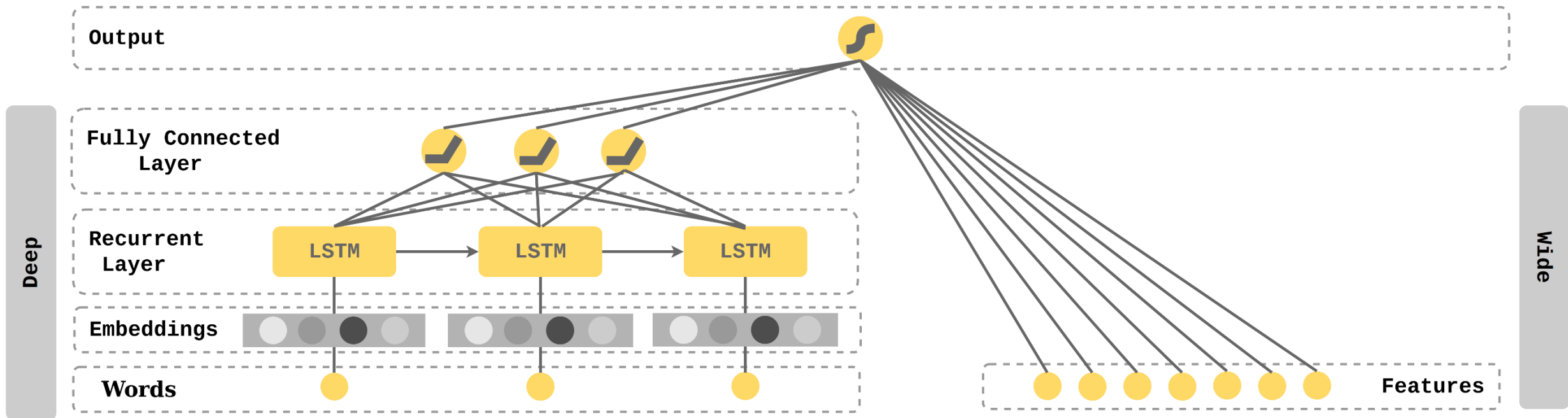
- Sven coop
- Sven counters

...

Sven z

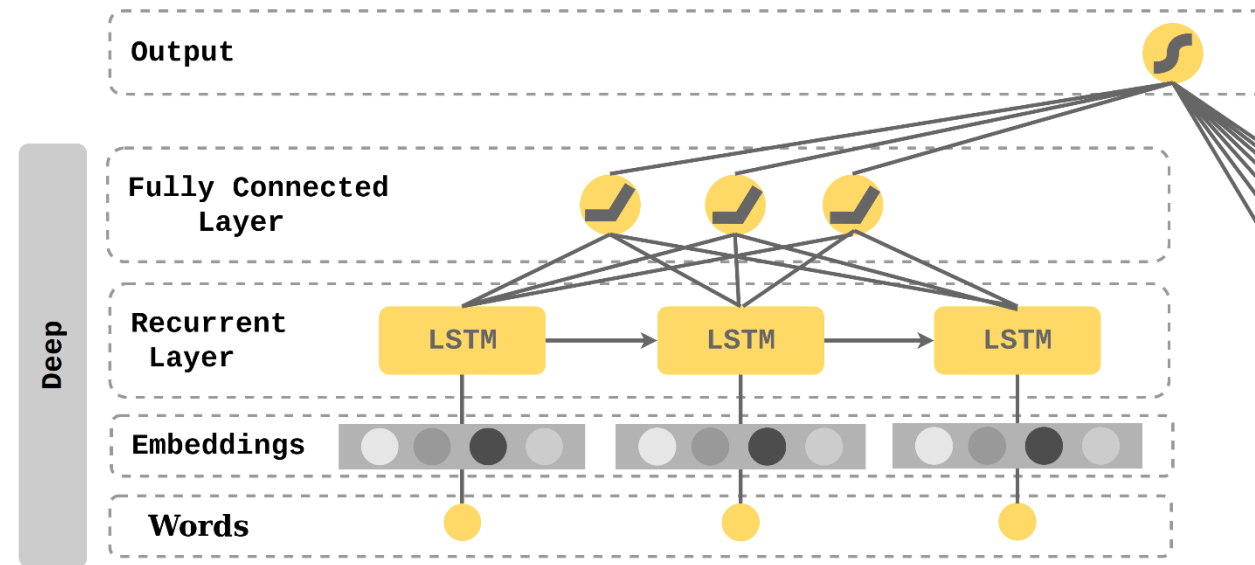
- Sven zatec
- Sven zimmer

Ranking

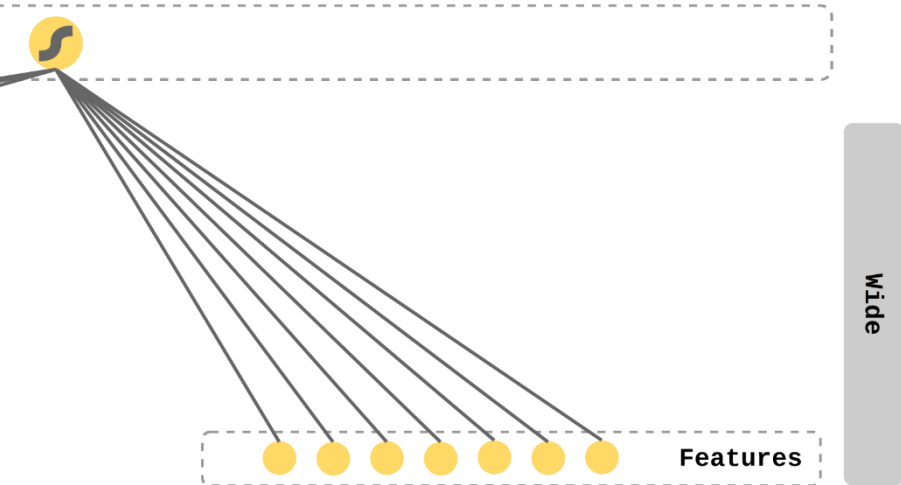


Ranking: Deep Model

- ❑ Word Embeddings
- ❑ Recurrent Long Short Term Memory Layer
- ❑ Fully Connected



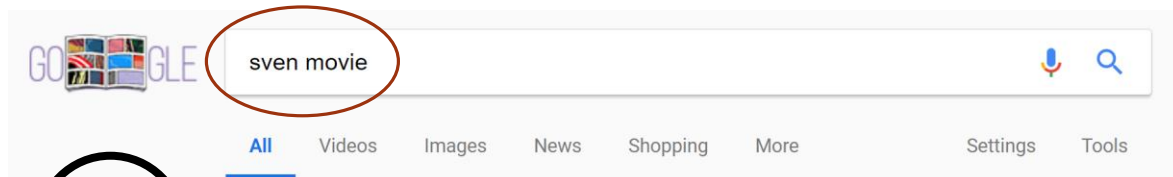
Ranking: Wide Model



Feature Name	Description
Trendy Terms	$tt(q) = \frac{\sum_{i=1}^{ q } t(q_i)}{ q }, t(q_i) = \begin{cases} 1, & \text{if } q_i \text{ is recognized as a trendy term} \\ 0, & \text{otherwise} \end{cases}$ <p>To determine if q_i is a child-related trendy term, <i>ReQuIK</i> examines its existence in children related pages on well-known websites, such as Amazon.com and CommonSenseMedia.org.</p>
Entities	$et(q) = \frac{\sum_{i=1}^{ q } e(q_i)}{ q }, e(q_i) = \begin{cases} 1, & \text{if } q_i \text{ is recognized as an entity by Stanford NER tool [35]} \\ 0, & \text{otherwise} \end{cases}$ <p>To determine if q_i is an entity, <i>ReQuIK</i> uses well-known CoreNLP tools.</p>
Children Dictionary	$ct(q) = \frac{\sum_{i=1}^{ q } c(q_i)}{ q }, c(q_i) = \begin{cases} 1, & \text{if } q_i \text{ is included in a children dictionary} \\ 0, & \text{otherwise} \end{cases}$ <p>We created our own children dictionary, comprised of 100,000 non-stop lemmatized terms, extracted from texts retrieved from a sample of various children-related websites.</p>
Flesch-Kincaid	$FK(q) = 0.39 \left(\frac{totalwords}{totalsentences} \right) + 11.8 \left(\frac{totalsyllables}{totalwords} \right) - 15.59$ <p>In our case <i>total words</i> is the total number of words in q_i and <i>total sentences</i> is always equal to 1.</p>
Enhanced Spache	$ES(q) = (0.121 \times AvgLengthOfq) + (0.082 \times NumberOfUniqueUnfamiliarWordsInq)$ <p>We implemented Enhanced Spache formula by updating the existing dictionary of common words with the children dictionary in C.</p>
Difficult terms	$dt(q) = \frac{\sum_{i=1}^{ q } d(q_i)}{ q }, d(q_i) = \begin{cases} 1, & \text{if } q_i \text{ is recognized as a difficult term} \\ 0, & \text{otherwise} \end{cases}$ <p><i>ReQuIK</i> treats q_i as a difficult term if it is not included in the children dictionary, or Spache dictionary of common words, or trendy terms list in A</p>

Wide model features; where q represents a query, q_i is the i^{th} non-stop word, lemmatized term in q , and $|q|$ is the length of non-stop words in q .

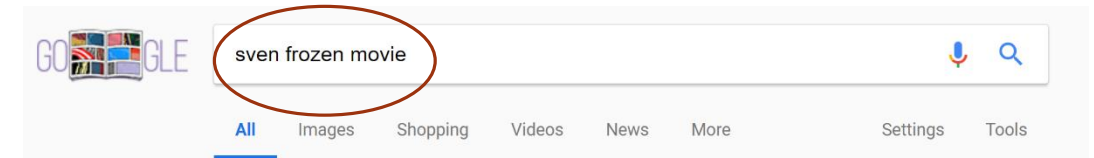
Reading level filtering



12th Grade

15th Grade

11th Grade



8th Grade



→ More images for sven frozen movie

Report images

People also ask

Is Sven from frozen in Moana?

What is Sven in frozen?

What is the meaning of the name Sven?

Is Sven a reindeer?

Feedback

Search
Intent

Candidate
Generation

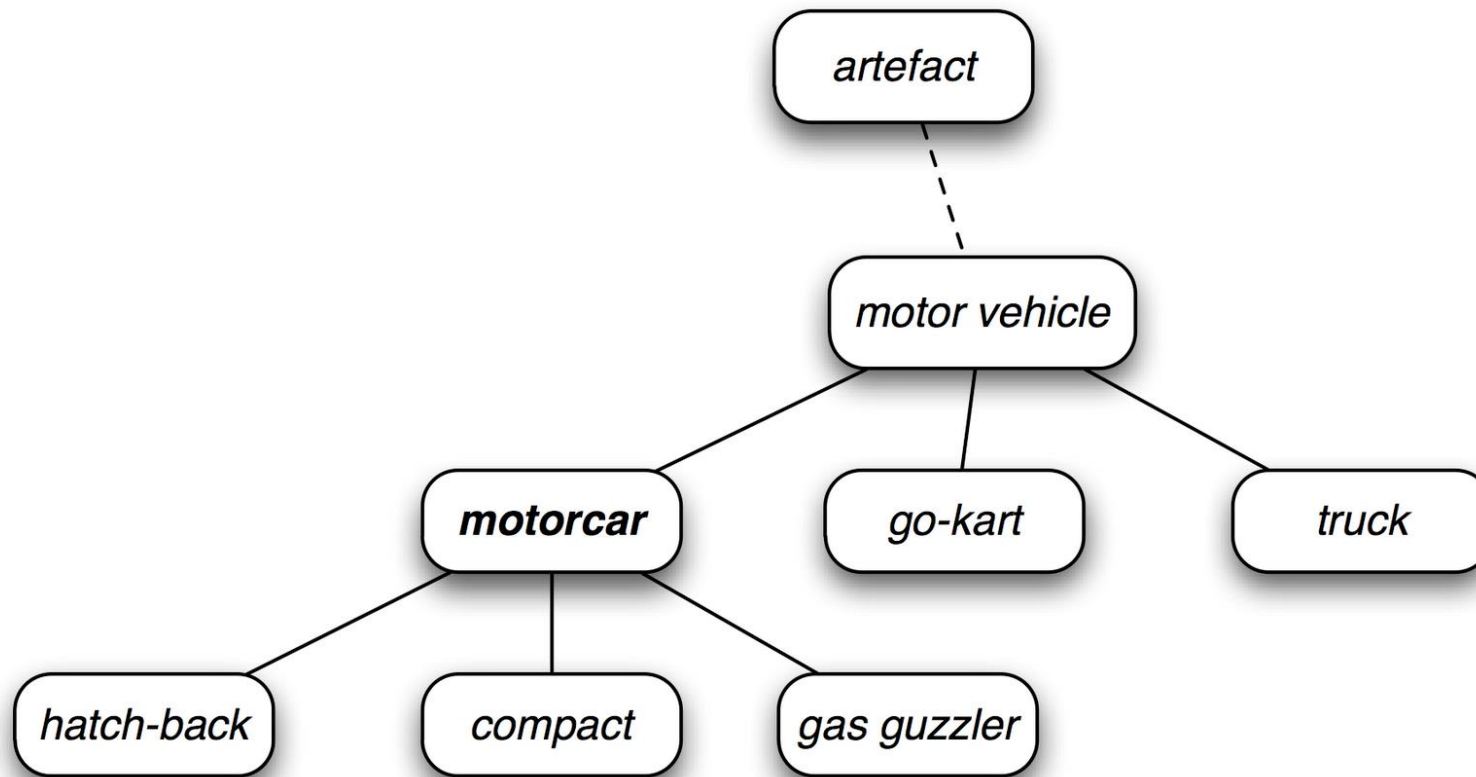
Ranking

Reading
Level

Diversity

13

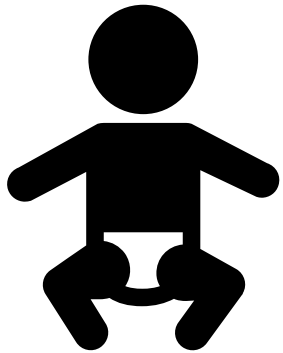
Ensuring diversity



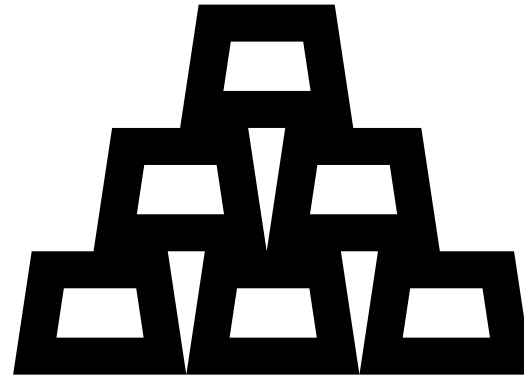
Wordnet Based Similarity

Image courtesy: FreeSandal.org

Evaluation



Not an option



Module-wise evaluation

Data: Query gathering



K-9 Idaho Teachers



50 Children



300 open queries



Publicly Available

Dataset available at: http://scholarworks.boisestate.edu/cs_scripts/5/

Data: Other resources



1M sentences based on children content

Compared strategies



Evaluation: Search intent

Google	Bing	Yahoo	Kidzsearch	CQS	ReQuik
46%	36%	65%	76%	57%	94%

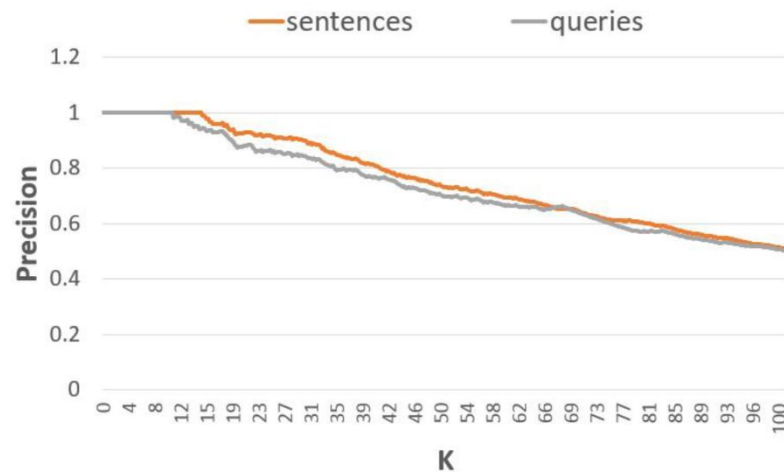
Percentage of queries that trigger a recommendation in each compared system.

Classifying children queries

	Wide	Deep	Wide and Deep
Accuracy	68%	92%	94%

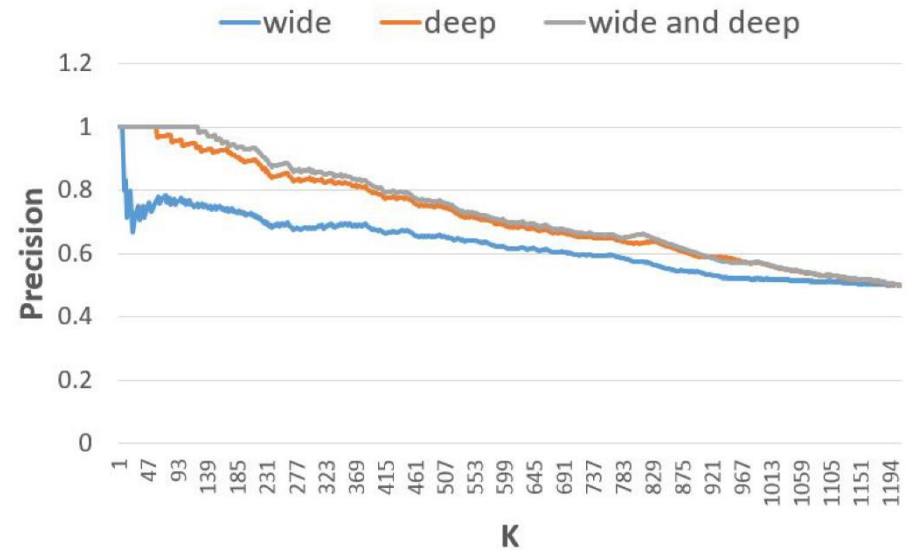
Performance of diverse ranking strategies.

Can we use this model for ranking?



Precision@K, where K is defined as the percentage of queries and sentences analyzed.

(K is defined as a percentage, as the raw counts of sentences and queries are not comparable otherwise)



Model assessment based on Precision@K, where K is the number of queries examined.

Readability of retrieved documents

ReQuik	Google	Bing	Yahoo!	CQS
7.71	12.46	19.96	11.42	11.82
AskKids	lpl2	Sweet Search	KidRex	Kiddle
13.3	10.9	12.3	12.7	12.83

Average Readability of top-3 documents retrieved for test query recommendations.

Online assessment

- ☐ 11 Teachers
- ☐ 5 schools
- ☐ 10 initial queries
- ☐ Each teacher could select at most 2 queries

Query Suggestions for Children

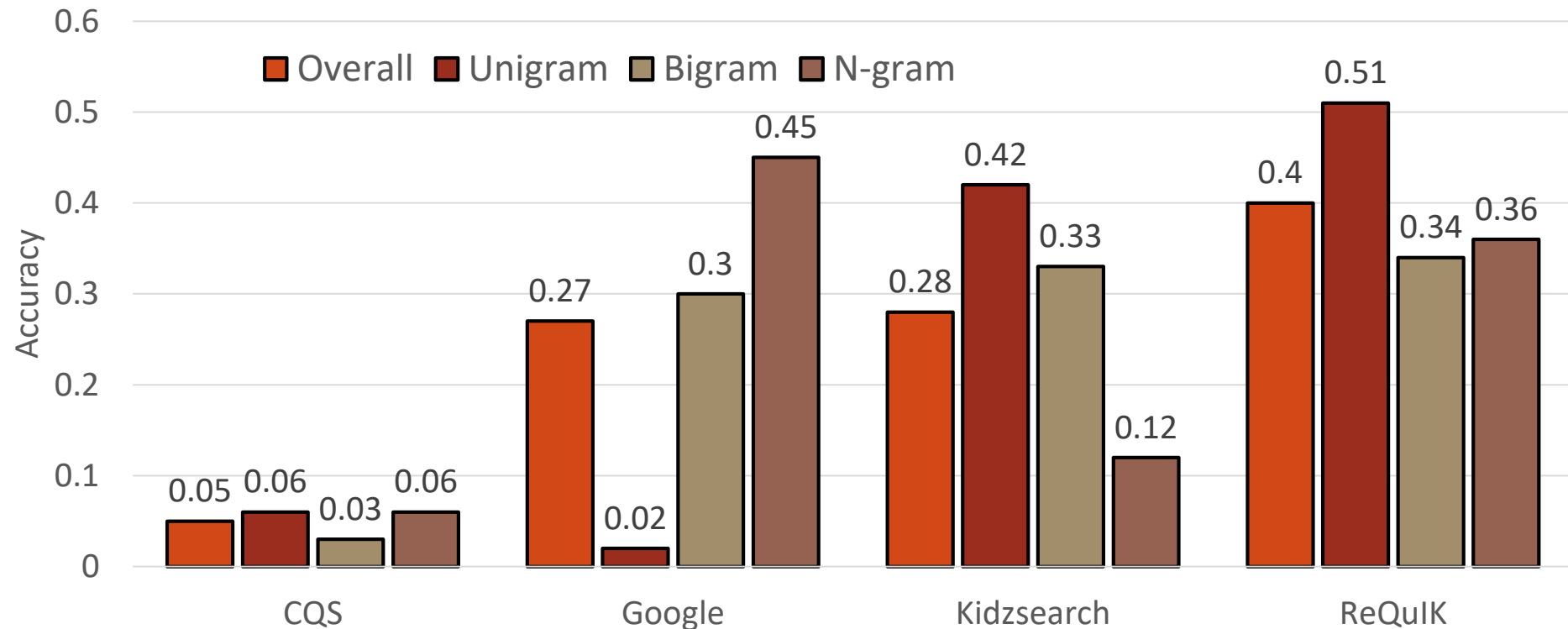
Imagine "Sven" is a young child using a search engine. To start a search, Sven uses the phrases below. In each case, which two suggestions will lead Sven to find child friendly and interesting web pages?

* Required

Sven wrote "Math", I would encourage him to use: *

- ☐ math games played for germany national-sports
- ☐ mathway
- ☐ math games
- ☐ math playground
- ☐ math games played for brazil national-sports
- ☐ math riddles for kids
- ☐ mathletics
- ☐ math worksheets for kids

Online assessment



Comparison of query recommendation systems.

Conclusions And Future Work

Main Contributions:

- ❑ ReQuik
 - ❑ Deep & wide model – trained on children sentences; to account for limited query log availability
 - ❑ Readability-based prioritization of queries
- ❑ Children query dataset

Future Work

- ❑ More exhaustive evaluation
- ❑ Gather more queries
- ❑ Better suitability analysis

Thank You!

LOOKING FOR THE MOVIE SEVEN OR SVEN FROM THE MOVIE FROZEN?
*A Multi-perspective Strategy for Recommending Queries for
Children*

Ion Madrazo Azpiazu
Nevena Dragovic
Oghenemaro Anuyah
Maria Soledad Pera



**Research supported in part by National
Science Foundation, Award # 1565937.**

