International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Survey on Evaluating Query with Uncertainty

Akshay Harish[1*], Aryan Shiv[2*]

*School of Engineering, Presidency University Bengaluru, Karnataka, India

1. 201910101608@presidencyuniversity.in
2. 201910100446@presidencyuniversity.in

ABSTRACT: - The Internet is an endless ocean of information with the solution to every query searched. The pattern of searching over the years has changed drastically, from 2-year-old to 100-year-old every person has a distinct pattern of searching. A query searched by a kid and an adult is immensely contrasting, an adult query might be “Apple iPhone 126GB Rose Gold” and the same query by a kid would be "Gold Phone". The structure of the query has degraded over the decade by the millennials, they have very limited knowledge of a structured query. These are the main problems the internet face, “Uncertainty” and “Vagueness” of the query. User query lacks clarity which leads to information not being relevant to the user. This paper discusses about various evaluating techniques and systems that are helpful for the user to have a satisfying experience on the web.

Keywords- Uncertainty, vagueness, structured query.

I. Introduction

A query is a a “question, especially one expressing doubt or requesting information” and in “Information Retrieval” this is a certain way on how to “retrieve” a specific piece of information and it uses a medium such as “search engines” to express it on the Web. Queries never have a certain structure in the context of the “Web” to retrieve a “piece of information” or a document because in the collection of documents or “corpus” is actually very huge ranges in around billions to trillions of it and is unstructured in the Web and queries might have many meanings and semantics and might not be relevant to the user and that’s why it needs to be meaningful and semantically accurate in order to retrieve the document to fulfill the query. The term “Query evaluation” comes in and it is defined to evaluate the query written by the user and rank the documents according to a “score” of the query and this paper is to survey those techniques which currently plays role in query evaluation.

II. Related Work

Jian Pei et al, in their paper “Query Answering Techniques on Uncertain and Probabilistic Data," have mentioned how to rank the queries, and how to answer the query on uncertain data(this is not based on search engine queries). Semantics and efficient algorithms such as knn, reverse nearest neighbor search, continuous nearest neighbor search, etc. can be used.

Mehdi Hosseini et al, in their paper “An Uncertainty-aware Query Selection Model for Evaluation of IR Systems” have mentioned certain theoretical models which could be used for evaluating uncertain queries like uncertainty-aware selection, and adaptive query selection.

Martin Grohe et al, in their paper "Probabilistic query evaluation with bag Semantics" have mentioned a technique to use bag semantics to evaluate query.

Mingang Chen et al, in their paper "Performance Evaluation of Recommender Systems." have talked about how the recommendation system can be refined.

Rashid Ali et al, in their paper "An overview of Web search evaluation methods." Have discussed about how the web query can be evaluated to users' satisfaction.

Santosh Kumar Bharti et al, in their paper "Automatic Keyword Extraction for Text Summarization: A Survey" have talked about techniques used to measure performance of the query.

Christopher D. Manning et al, in their textbook "Introduction to Information Retrieval" have talked about various evaluation metrics that are used.

Tom Fawcett in his paper "An Introduction to ROC Analysis" have mentioned about how roc accuracy curve is helpful in evaluation.

David W M Powers in his paper "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation" has talked about various mathematical evaluation techniques.

III. Current Work

The Internet is a pool of information with never-ending results for the query searched. A query is the most important aspect of retrieving information on the internet, but the results might not be satisfactory to the user. This could be due to an uncertain or vague query.

Uncertainty is a growing issue in terms of retrieving information. Lack of patience and lack of structured query knowledge has led to uncertainty on the internet which leads to an unsatisfied user who blames the search engine's efficiency for not retrieving relevant data.

How do uncertainty and vagueness affect the search? Suppose a user query is *"Apple rates"* now the search engine will be confused as to which apple the user is talking about, is it the rate of the Apple iPhone, or is the user asking about the rate of apple the fruit. If the user was expecting to see the rates of apple the fruit but the search result was the cost of the Apple iPhone, then the user is disappointed with the search engine.

The use of structured queries always helps in retrieving relevant information. Use of operators such as AND, OR, NOT, and "" (quotes) help in fetching the exact information expected by the user.

These searches(queries) are evaluated in a few methods, Result evaluation, Recommendation evaluation, Search engine evaluation-query categorization, and search evaluation.

[4] Result Evaluation: - once any result is fetched it is ranked based on your previous searches on the search engine. The evaluators rank them in many ways positive or negative, rank, scale, side-by-side. The more a user searches the more the algorithm will learn, and relevant data will be fetched.

[4] Recommendation Evaluation: - certain results are recommended by the search engine at the time of fetching relevant information, if the user cannot find any relevant data, then the recommendation is very important. A recommendation is done based on the search and it plays a very vital role.

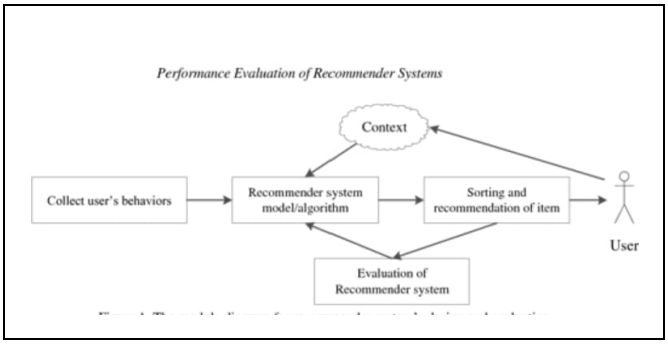

Fig 1.[5] Performance evaluation of recommender systems

[4] Search Engine Evaluation-query categorization: - all the queries are categorized into different fields. iPhone is categorized into a separate field and apple is categorized into different fields and so on. When the user searches for a query a certain category is chosen, and results are displayed accordingly.

[4] Search Evaluation: - the content of the user searches is evaluated. This means the type of query the user searches play a key role in the recommender systems. Based on the user queries from the past relevant results are fetched. Suppose a user's searches are only based on research papers, all his future query results will be based on a research paper only.

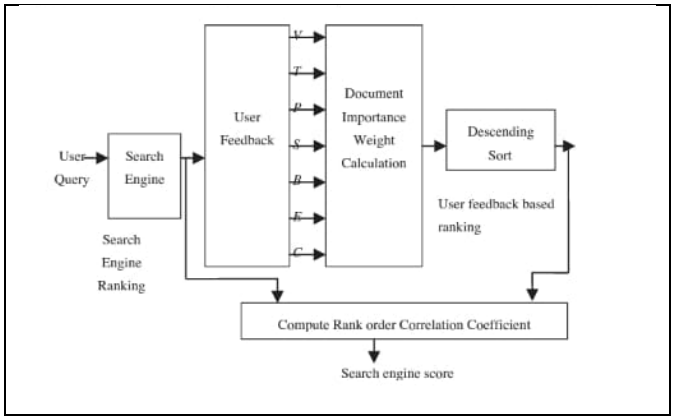

Fig 2.[6] web search evaluation system based on user satisfaction

IV. Conclusion and Future Work

In this paper we've described a lot of techniques such as Result Evaluation, Recommendation Evaluation, Search Engine Evaluation, Search Evaluation and more research is going on "Query Evaluation" and new more techniques have arrived such as "Probabilistic Query Evaluation with Bag Semantics" [3] . A relative new technique and currently with the development of AI in Information Retrieval we can fine tune the query evaluation by various parameters with the help of "reinforcement learning" where the IR system takes feedback and fine tunes itself to provide better results and the development of evaluating using these queries but the problems that are currently which are being faced are the "uncertainty" and "vagueness" cannot be solved by simple algorithms currently because these are based depending on the user and cannot be solved by simple algorithms but can be manipulated using the user's searching and semantics behavior.

References

[1]. Mehdi Hosseini, Ingemar J.Cox "An Uncertainty-aware Query Selection Model for Evaluation of IR Systems"

[2]. Jian Pei, Ming Hua, Yufei Tao, Xuemin Lin "Query Answering Techniques on Uncertain and Probabilistic Data"

[3]. Martin Grohe, Peter Lindner, and Christoph Standke "Probabilistic Query Evaluation with Bag Semantics."

[4].https://www.telusinternational.com/articles/how-to-improve-search-relevance-5-types-of-search-evaluation

[5]. Mingang Chen, Pan Liuc, "Performance Evaluation of Recommender Systems."

[6]. Rashid Ali, M.M. Sufyan Beg, "An overview of Web search evaluation methods."

[7]. Santosh Kumar Bharti, Korra Sathya Babu, and Sanjay Kumar Jena "Automatic Keyword Extraction for Text Summarization: A Survey"

[8]. Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze (2009). "Chapter 8: Evaluation in information retrieval"

[9]. Tom Fawcett. "An Introduction to ROC Analysis"

[10]. David M W Powers. "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation"

Authors Profile

AKSHAY HARISH currently pursuing Bachelor of Technology from Presidency University, Itgalpur, Rajanakunte, Bengaluru, Karnataka, 560064, INDIA.

ARYAN SHIV currently pursuing Bachelor of Technology from Presidency University, Itgalpur, Rajanakunte, Bengaluru, Karnataka, 560064, INDIA.