QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Electronic Recommendation Based on Customer Review Using Text Mining

Asst. Prof. Dr. Saravana Balaji B
Department of Information Technology, College of Engineering and Computer Science, Lebanese French University, Erbil, Kurdistan Region, Iraq.
saravanabalaji.b@lfu.edu.krd

Eman Sabah Mustafa
Master Student, Department of Information Technology, College of Engineering and Computer Science, Lebanese French University, Erbil, Kurdistan Region, Iraq.
Imansabah5@gmail.com

ARTICLE INFO

Article History:
Received: 2/8/2021
Accepted: 20/9/2021
Published: Autumn 2021

Keywords: *Customer Review, NLP (Natural Language Processing), Python, Anaconda, Recommender Systems, Context-aware recommender systems, social networking.*

Doi:
10.25212/lfu.qzj.6.4.31

ABSTRACT

In this paper, buyers are looking for reviews of the products before purchasing them. In view of this, online shopping platforms are encouraging their customers to provide reviews on products that would help future customers and the service provider to enhance their services. These reviews are normally in natural language mostly in the English language. These are used to analyze and provide data that is used for repairing and building new products because most services are unable to review consumers' reviews at the same time regularly, so they need mining tools to learn about those reviewers, which is what consumers need for their goods. Users review assessments for upgrading their products. Reviews are analyzed by customers to decide whether to purchase or not to purchase. The main objectives of this study are to develop a recommendation system based on customer reviews, develop a dataset of customer reviews from Amazon and eBay, analyze the reviews to create a database of products, develop an algorithm for generating positive/negative scores for a product, develop a method for gathering user requirements in natural language and to identify the main product, and also develop an algorithm to match the user request with the product and generate recommendations. The study concluded that the selected

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

program was suitable for analyzing and managing the review data.

1. Introduction

Natural Language Processing (NLP) is commonly used for analyzing customers' feedback. Computers and machines are great at working with tabular data or spreadsheets (Qasim, Din, & Alyousuf, 2020). However, as human beings generally communicate in words and sentences, not in the form of tables. Much information that humans speak or write is unstructured. So, it is not very clear for computers to interpret such. In natural language processing (NLP), the goal is to make computers understand the unstructured text and retrieve meaningful pieces of information from it. Natural language Processing (NLP) is a subfield of artificial intelligence, in which its depth involves the interactions between computers and humans (Pratik Shukla & Roberto Iriondo 2021). All modification that occurred from past up to now is called generation (Kamaran Faraj 2016). Social networking is defined as the use of Internet-based social media platforms to keep in touch with friends, family, coworkers, customers, or clients. Through social media sites such as Facebook, Twitter, LinkedIn, and Instagram, among others. Social networking can be used for personal, professional, or both objectives (Mohammed, 2021). Whereas, social networking has become a significant resource for consumers wishing to engage with other consumers with over two billion users. As well as, to access social media platforms, the majority of users use web-based software on their (Desktops) and (Laptops). Or they use their smartphones and tablets to download social networking applications. Context-aware recommender systems (CARS) have emerged as a hot research topic in the field of recommendation, with the goal of enhancing recommendation quality and user loyalty by incorporating context knowledge. Integrating context information into recommendation structures is complex due to the high dimensionality of context information and the sparsity of results (Shi et al., 2017). The study looked at the evolution of travel advisory systems, as well as their usefulness and existing shortcomings. The recommendations were based on the community's collective experience over the last decade (Hashim, Hammood, & Al-azraqe, 2016). It also went into the core algorithms used in classification and recommendation systems. And it looked at metrics for evaluating the algorithms' and therefore the recommenders' performance (Shini Renjitha, et al., 2019). The negative affect is that if the customer can't get the products very well, it means there is something wrong with their products. (Kamaran Faraj, 2019) Nevertheless E-recommendation by (semi restricted search engine and restricted search engine) solve the inconvenience of traditional recommendation because of E-recommendation is paperless, less time consume, accurate, reduce health and safety hazards and etc (Mohammed, Andrii, Ahmed, Olha, & Owaid, 2019). regarding the TELOS categories E-recommender output outcome is much higher and it is feasible than the traditional recommender (Sharma & Kumar, 2021).

2. Literature Review

There is a lot of research that presents the solution to this issue that describes the methods of analyzing data. The model based on the long short-Term Memory (LSTM) network produces the best outcomes. The proposed model for collecting review similarities greatly enhanced the recommender system's efficiency. According to the scientists, the proposed model was used in Amazon.com's results (Ghasemi and Momtazi, 2021). Social media is "highly necessary" to have access to users' contributions on social media platforms. Linked Data and the Semantic Web have proven to be successful for knowledge management and data integration. (Sánchez, et al, 2020). A Text-based Price Recommendation System (TAPE) is a new tool that helps Airbnb and other short-term rental sites make better pricing decisions. TAPE was designed and implemented using deep learning techniques. The results were shown in Boston, London, Los Angeles, and New York City over the course of three months (Shen, et al., 2020). Because of its broad variety of uses, aspect-based sentiment analysis (ABSA) has recently received a lot of attention. This paper suggests a practical method for dealing with large-scale unlabeled datasets. The suggested strategy incorporates a frequency-based (word level) with a syntactic-relationship-based (syntactic level) approach (sentence level). The extracted aspects are used to generate a total review sentiment score after calculating the weight and rating of the

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

extracted aspects mentioned in the review. On the Amazon and Yelp datasets, the derived aspects using the proposed approach with the domain-specific lexicon outperformed all baselines in terms of F-measure and precision (MOHAMMED et al.,2020). Preferences have been extensively researched in a variety of fields, including computer science. They are crucial in a wide range of computational activities, such as decision support, suggestion development, and device personalization. It is necessary to develop a paradigm that offers sufficient constructions for users to communicate their desires. In this article, this process suggests a preference meta-model that offers multiple preference structures, including end-user expressions (Nunes, et al., 2013). A recommender system is a type of information system that helps users make choices by recommending products that are important to them. The main goal of these services is to products shown by a user to result in conversions, or user intervention, such as buying access to watch a video (Bogdan Walek and Vladimir Fojtik, 2020). The implementation of intelligent systems for cultural visit planning was very appealing in the cultural heritage domain. We propose an innovative method for incorporating a route planner into cultural sites such as museums. The proposed solution uses a user-centered recommendation methodology to indicate the most suitable cultural items. (Flora Amatoa, et al., 2020). Product-Service System (PSS) models provide businesses with a value-added digital service approach. Customers' tastes, as well as product and service attributes, may have a significant impact on PSS contract decisions. The simulation model takes stochastic volatility into account. (Maryam Farsia, and John Ahmet Erkoyuncu, 2021). We have presented an agent-driven e-commerce framework in this paper that recommends goods to customers based on their preferences. When you get feedback from customers after they see the goods in person, you can use it to improve your business. In this paper, we present an agent-driven e-commerce system that suggests products to customers based on their preferences. When customers provide feedback after seeing the product in person, you can use it to develop your company (B.K. Mohanty, K. Passi, (2010). The Improvised Clustering Recommendation System (ICRS) will introduce a new system for online service providers. The ICRS is based on data collected from previous purpose invocations to examine the different aspects of our models. It will be available in April (Arif Ahmed Khan, and V. A. Chakkarwar, 2018). The web services are software components that provide interoperable machine-to-machine communication over the internet. The adoption of Web services as a business delivery mode has ushered in a new paradigm shift away from monolithic application development and toward dynamic business process configuration. Web services have piqued the attention of both industry and academia in recent years, and the number of public Web services continues to grow. (Xi Chen, Members, 2013). The Internet of Things (IoT) is grappling with a growing number of IoT users, resulting in a huge data transfer burden for cloud data centers. We propose a QoS conscious resource allocation approach based on user ranking implicit input in Mobile Edge Computing (MEC) (Puja Das, and Asik Rahaman Jamader, 2019). Web Services (WS) are server modules that aid in the integration of different Web-based systems. Almost all web apps make use of WS. The number of WS on the Internet is increasing. It is impossible for the customer to choose the best service from a wide range of options. We are working on a location-aware QoS-based web service recommendation framework (Blessina Gonsalves and Vandana Patil, 2017). The proposed scheme derives contextual properties from WSDL and uses them to cluster Web services based on function similarities. It then recommends services to consumers using a better matrix factorization process. In the real-world dataset of over 1.5 million Web service invocation reports (Shun Li, and Junhao Wen, 2017). Movie recommendation services provide users with ranked lists of movies based on their tastes and restrictions. LSIC model uses adversarial testing to exploit long and short-term experience for content-aware movie recommendation. The proposed model outperforms competitors and achieves cutting-edge performance. (Wei Zhao, Etal, 2019). EmoWare (emotion-aware) is a personalized and emotionally intuitive video recommendation engine. It makes use of a novel context-aware interactive filtering approach that assesses the intensity of users' nonverbal emotional reactions to the suggested video. The EmoWare framework outperforms state-of-the-art approaches and is tested with actual users for a month. It models users' emotional needs very well with stable convergence. It is focused on deep-bidirectional recurrent neural networks in both positive and negative time directions (DBRNN) (Abhishik Tripathi,

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

2019). The aim of this paper is to provide an overview of mobile context-aware recommender systems (Manikandan, et al., 2020). In a mobile setting, several intimate, social, and environmental contextual variables may be integrated into the recommendation process. This study also seeks to identify potential research directions in this area (Ben Sassi et al., 2017). The difficulty of uncovering meaningful user and product representations using only user-side pairwise ranking was highlighted in this paper. To that aim, they present BSPR, or basket-sensitive personalized ranking, a unique probabilistic pairwise technique that handles both user- and product-side pairwise ranking problems in a unified manner (Wu, B., & Ye, 2020). To boost client loyalty and cross-sell products, several online businesses utilize automated product recommender systems. According to the study, the proposed strategy beats existing recommender approaches in terms of predicted accuracy and ability to provide actionable reasons. It is also a prerequisite of artificial intelligence systems from an ethical standpoint (Marchand, A., & Marx, P. 2020). Recommendation System (RS) discovers users' interests by analyzing implicit or explicit user behavior on an e-commerce website and recommending goods that best fit the user's preferences. The suggested method combines content and user-item-based collaborative filtering to provide a single RS with a minimal number of recommendations (Tewari A. Sh.,2020).

3. Proposed Technique

This research is about the system architecture of the customer reviews and how they send the query to the process. Customer reviewers are an important source of "voice of the customer" information, and they may reveal details of what customers want or dislike about a service or product. This diagram included the customer's sent query then analyzed for (PoS Tagging, Annotation, Positive and negative reviews, then clustering), The review dataset, which will receive online data after that in the process, can classify the negative and positive sides of customer reviews. Then they will be in some clusters. In (figure 3.2), the diagram illustrates the system architecture for a user:

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

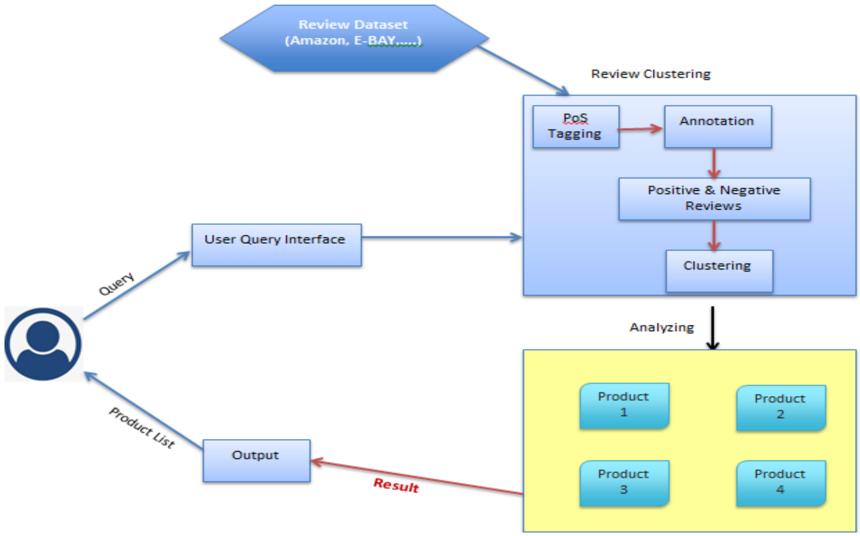

Figure (1): Recommender System Architecture

There are three main components they are:

3.1 Review Clustering

Review clustering matches similar names into one cluster. For instance, if TV and television have the same mining, then they will be in the same cluster. There are four main subcomponents. They are:

- **PoS Tagging**

Parts of Speech Tagging (POS Tagging) is a method of marking up words in text format for a specific section of a speech. Counting nouns can help determine how many different topics are being discussed at the same time. The aim of our script will be to count how many adjectives and nouns appear in the positive subset of the review samples.

- **Annotation**

Annotations are a new feature that allows users to add and manage notes for products. Annotations mean extra information. For example, when a customer searches for a product, the product name automatically appears and gives the user

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

many options until it is clearer. For instance, the abbreviation for television is TV and the abbreviation for electronic mail is E-mail.

- **Positive and Negative Reviews**

used to find if a word is positive or negative. It is clear that every device has negative and positive sides during applications. For instance, the customer asked for the product, but during receiving the product they see the draw backs. However, the customer can get the requirements fast and soon without any problems. In the end, the customer can decide whether the product is good or bad. For example: Excellent is a positive word and slow is a negative word.

- **Review Clustering**

Clustering is a well-known data mining Technique for identifying hidden patterns. Classifying and analyzing text for products which means that it breaks up into many groups, the aim of this work is to the group of the customer reviews depends on their opinions on different products, for this purpose we design an algorithm which is shown in the table.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Algorithm for Clustering

Table (3.1): Clustering Algorithm

```
Start:
  Input: Review Dataset R_All
  Output: Clustered Reviews
   FOR R_i Є R_All
                    APPLY PoS TAGGING → PoS(Ri)
                    SELECT NOUNS (Ni) FROM PoS(Ri)
                    FIND SYNSET (Si) of Ni Є Ri
  FOR ALL (S_i)
                     FIND MATCHED CONCEPT C of (S_i)
                     SELECT C (S_i) AS MAJOR CONCEPT
  END
   FOR ALL PoS_i IN PoS(R_i)
                     IF PoS_i(R_i) Є POSITIVE
                      POSITIVE ++
           ELSE
                       NEGATIVE++
  END
                 IF POSITIVE > NEGATIVE
                    Mark (PoSi(Ri)) = POSITIVE
                    ELSE Mark(PoSi(Ri)) = NEGATIVE
                    FORM CLUSTER[C]
                 END

            FOR All Ci Є [C]
          FOR ALL Cj Є [C]
                     FIND SYNSET of C_i & C_j
                     IF C_i && C_j MATCH
                    GROUP C_i && C_j
          END
          END
          REFIN [C] && PRODUCT OUTPUT
          END
```

The methodology of the study includes Python programming elements, which start by inputting data such as Review Dataset and the output data of clustered reviews.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

The main steps of the program start with:

Apply PoS Tagging then select "Noun"

For instance, if a user searches for a product and another user searches for the same product, the process will match the meaning of those words. If they have the same meaning, they will be in one group. The aim of this algorithm process is to match the meaning of all clusters. If they have the same meaning, the program matches the meaning.

For instance, TV and television have the same meaning. The program will match and it will be in the same cluster.

Below the table is the description of Part-of-Speech:

S.No	Part of Speech	Description
1	CC	Conjunction, joining words
2	ADJ	Adjective to describe a noun
3	DT	Determiner
4	IN	Preposition
5	JJ	Adjective, like big
6	JJR	Adjective, comparative, like bigger
7	JJS	Adjective, superlative, like biggest
8	NNS	Plural noun, chocolates
9	NN	Singular noun, chocolate
10	NNP	Proper noun, like Donald Trump
11	PRP	Personal pronoun, I, HE, SHE
12	PRP$	Possessive pronoun, his, her
13	UH	Interjection
14	RB	Adverb
15	RBR	Comparative adverb
16	RBS	Superlative adverb
17	VB	Verb
18	VBD	Verb, past tense
19	WRB	Wh -adverb where, when

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

20	FW	Foreign Word
21	LS	List Marker
22	MD	Modal cloud, will
23	PDT	predeterminer 'all the kids'
24	TO	To go 'to' the store
25	CD	Cardinal Digit

In this progress, the Parts_of_Speech (PoS) is categorized into Noun, Verb, Adjective, Conjunction, Pronoun, Number, Preposition, Determiner, and Adverb classes.

Ni: means the nouns for the review of products and (i) means, for example, product number one, two, three.

Si: Synset, which means the synonyms for all words, match the meaning. If they have the same meaning, they will be in the same cluster. Otherwise, they will be in a different cluster.

PoS: it is an abbreviation for Part-of-Speech. As mentioned before, it is a technique for breaking up words into Noun, Verb, Adjective, Conjunction, Pronoun, Number, prepositions, determiners, and Adverb classes.

i: means how many numbers it will add to the program.

Є: means elements

Match: for matching synonyms in the program. If they have the same meaning, the program matches the meaning and they will be in one cluster.

Ci: Product number one, for example, TV.

Cj: Product number two, for example, is television.

[C]: Cluster

C: Concept

Positive: means it mentions a good word in the review for the customer's feedback.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

For instance, it is an amazing dress.

Here the word (amazing) is a positive word.

Negative: means negative words in the customer's review

For instance, I received a smartphone, but it was broken.

Here the word (Broken) is a negative word.

Cluster: means the group of products that we classify in progress. For example, there are a thousand clusters including the customer’s review.

Input: Reviews dataset that the user receives, for instance, Amazon’s review dataset.

Output: means the cluster reviews that result when it appears after all the progress.

The step below will collect all reviews. Then each review will be selected for finding a synset. Then for each synset, it will find the major concept (it will find the review concept).

```
FOR Ri Є RAll
    APPLY PoS TAGGING  →  PoS(Ri)
    SELECT NOUNS (Ni) FROM PoS(Ri)
            FIND SYNSET (Si) OF Ni Є Ri
```

The second step below in the algorithm, the review contains all about television so major concept is Television.

```
FOR ALL (Si)
        FIND MATCHED CONCEPT C of (Si)
        SELECT C (Si) AS MAJOR CONCEPT
END
```

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

In the third step, each word in this PoS will be calculate with positive and negative, and after it checked all algorithm, it will check the total number for positive and the total number negative word.

```
FOR ALL PoSi IN PoS(Ri)
            IF PoSi(Ri) Є POSITIVE
                POSITIVE ++
           ELSE
             NEGATIVE++
END
```

In the fourth step in algorithm, it will check the total number of positive and the total number of negative reviews. If positive word is higher, it means this review will be positive, but if the negative review is higher, it means this review will be negative. The user has a concept of positive and negative .

```
IF POSITIVE > NEGATIVE
          Mark (PoSi(Ri)) = POSITIVE
               ELSE Mark(PoSi(Ri)) = NEGATIVE
            FORM CLUSTER[C]
END
```

This step below will be repeated for all reviews.

```
FOR All Ci Є [C]
FOR ALL Cj Є [C]
```

From each concept every other concept all be matched with synset

```
FIND SYNSET OF  Ci & Cj
            IF Ci && Cj MATCH
            GROUP  Ci && Cj
END
```

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Find output with all grouped reviews and positive or negative rank.

REFIN [C] && PRODUCT OUTPUT

3.2 Data Processing

- Algorithm for Clustering

3.3 User Interface

In a device, the user interface (UI) is the point of interaction and communication between Humans and computers. Display screens, keyboards, mice, and the appearance of a desktop are all examples of this.

3.3.1 Find features of product as per query

Natural language processing (NLP) is a specialized field for analysis and generation of human languages. NLP provides the ability to comprehend natural language input and produce natural language output appropriately.

RANKING ALGORITHM

START
INPUT: USER QUERY POS(U_Q)
OUTPUT: LIST OF PRODUCTS BASED ON RANK (PRO_L)
FOR ALL PoS(U_Q)
  SELECT N_i(U_Q)
  FOR ALL S_i = SYNSET (N_i(U_Q))
   SELECT MAJOR CONCEPT S_i
END
FOR ALL C_j of [C]
    IF$\sum_{x=1}^{n} PoS(U_Q) == \sum_{Y=1}^{n} PoS(R_i)$
    weight (R_i) = TF-IDF ((U_Q) && PoS(R_i))
END
END

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

UQ: User Query, which means the user sends a query through Interface for example: (“The user search for the product”).

PROL: Product List, in the last step the output is product list for the user will appear.

Cj : Concepts

The step below will be collected all reviews. Each query applied with Part-of-Speech (PoS) tagging from the output each noun will be selected and their meaning are generated using Synset, the Synset method meaning are taken from the library, this library called WordNet which is Lexical Database for the English language. Then each user will ask about his product. Then for each Synset, it will find the major concept (it will find the concept of the review). And each word in this PoS will be calculated with positive and negative.

FOR ALL PoS(U_Q)
    SELECT $N_i(U_Q)$
    FOR ALL $S_i = $ SYNSET ($N_i(U_Q)$)
      SELECT MAJOR CONCEPT S_i
END

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

The second step below the user search, which is means for all (n) number x=1 means the product number one for instance (TV) and the (y=1) means another product name (for instance: Television). Spärck Jones, K. (1972), TF-IDF is an information retrieval technique that weights a term's frequency (TF) and its inverse document frequency (IDF). Each word or term that occurs in the text has its respective TF and IDF score. The product of the TF and IDF scores of a term is called the TF*IDF weight of that term. Put simply, the higher the TF-IDF score (weight), the rarer the term is in a given document and vice versa.

FOR ALL C_j of [C]
IF$\sum_{x=1}^{n} PoS(U_Q)==\sum_{Y=1}^{n} PoS(R_i)$
weight (R_i) = TF-IDF ((U_Q) && PoS(R_i))
END

co-(process where people are made priests, rabbis, etc.) levels, and averaging over the set of requests is by plain/honest/easy average of numbers. (high) quality at ten standard recall values is then (figured something unknown out based on things that are already known). The same relationship between full-term matching and this restricted matching with non-frequent terms only is shown by the other collections: the recalled ceiling is lowered by at least 30%, and in fact, for the well-developed the collection is reduced from 75% to 25%, though the (high) quality is maintained in the step, which means for all review products is an element in concept, and then it checks each weight of Review products and it selects with the top weight.

FOR ALL PoS(R_i) ∈ C;
CHECK EACH WEIGHT OF (R_i) WITH R;
SELECT MAJOR WEIGHT(R_i) AND DECREMENT EACI
END

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

In the last step, for all Review products which are elements in clustering the program will check the Positive and negative sides, and the output for the product list will appear.

```
FOR ALL RANKED(Ri) Є [C]
        CHECK EACH POSITIVE & NEGATIVE
         OUTPUT REVIEW WITH IMPACT
END
```

4. Performance Evaluations

This paper discusses the programming language and the libraries which used in the process. Anaconda is a distribution of the Python and R programming languages for scientific computing. The Anaconda Navigator is a desktop graphical user interface (GUI) that allows users to launch applications and manage conda packages, environments, and channels without using command-line commands. Python is a popular programming language. This programming will use the Anaconda program that involves (Spyder IDE, Pandas Library, NLTK Library, Scikit-learn package, and Numpy Library). The review dataset has been generated using the reviews from Amazon and eBay which consists of 1000 reviews. These reviews are variable in length and in the English language (Miran & Kadir, 2019b). For evaluating the performance of the proposed system (PR-CT), three existing systems namely BSPR [1], APR [2], and GIR [3] were compared through three performance metrics namely: Precision, Recall, and F1 Score. For evaluating the performance, ten different queries were given to the systems with different levels of language usage. Then the retrieved answers were compared for evaluating the performance of all four systems.

Precision is the fraction of the correct answers retrieved and the total number of retrieved answers, and the formula for the precision is given in equation (1)

$$Precision = \frac{TP}{TP+FP} \qquad (1)$$

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Recall is the fraction of the correct answers retrieved from the total number of relevant answers, and the formula for the recall is given in equation (2)

$$Recall = \frac{TP}{TP+FN} \qquad (2)$$

F1 Score is the harmonic mean of precision and recall and the formula for the F1 score is given in equation (3)

$$F1\ Score = 2 * \frac{Precision*Recall}{Precision+Recall} \qquad (3)$$

The precision values of the proposed technique (PRB) and the other existing techniques is given in Figure 2.

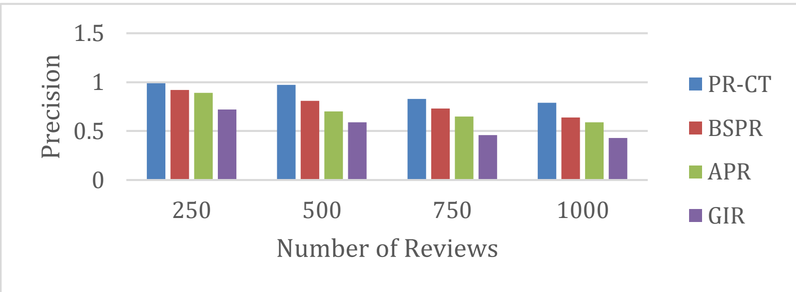

Figure 2: Precision Values

The figure 2 gives the precision values of the all the systems compared and the results shows that the proposed technique is having higher precision than the other systems. Figure 3 gives the recall values of the proposed system and the other techniques.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

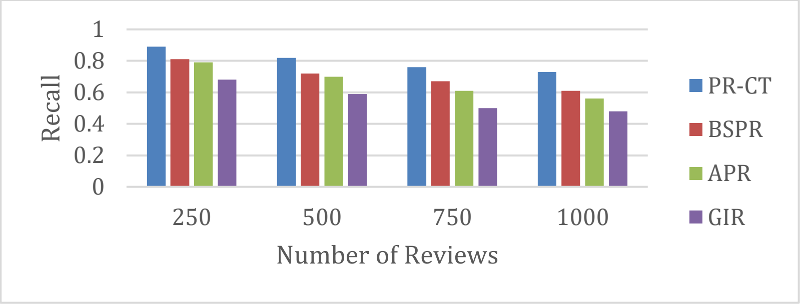

Figure 3: Recall Values

Figure 4, gives the F1 Score values of the all four systems compared.

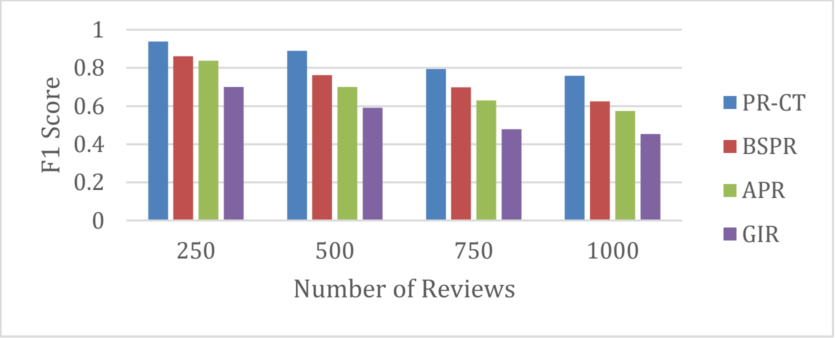

Figure 4: F1 Score Values

From all the performance evaluation operations, it is proved that the proposed technique PRB is performing well than the other techniques discussed in the literature.

5. Conclusion:

The primary goals of this research are to construct a recommendation system based on customer reviews from Amazon and eBay, evaluate the reviews to create a product database and provide suggestions. This paper discusses the recommender system architecture which received the dataset and analyzed and divided the

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

products. For instance, there are a thousand clusters matching the meaning in the same cluster and analyzing the positive and negative reviews by using “PoS tagging, annotation, and clustering”. Then they were divided into several clusters. The output is a list of products that the user got it. The performance evaluation shows that the proposed technique works better than the other techniques discussed in the literature.

References:

Al-Ghuribi, S. M., Noah, S. A. M., & Tiun, S. (2020). Unsupervised Semantic Approach of Aspect-Based Sentiment Analysis for Large-Scale User Reviews. *IEEE Access*, *8*, 218592-218613.

Amato, F., Balzano, W., Cozzolino, G., De Luca, A., & Moscato, F. (2019, March). Textual Processing in Social Network Analysis. In *Workshops of the International Conference on Advanced Information Networking and Applications* (pp. 822-832). Springer, Cham.

Das, P., Jamader, A. R., Acharya, B. R., & Das, H. (2019, May). HMF Based QoS aware Recommended Resource Allocation System in Mobile Edge Computing for IoT. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (pp. 444-449). IEEE.

Farsi, M., & Erkoyuncu, J. A. (2021). An agent-based approach to quantify the uncertainty in Product-Service System contract decisions: A case study in the machine tool industry. *International Journal of Production Economics*, *233*, 108014.

García-Sánchez, F., Colomo-Palacios, R., & Valencia-García, R. (2020). A social-semantic recommender system for advertisements. *Information Processing & Management*, *57*(2), 102153.

Ghasemi, N., & Momtazi, S. (2021). Neural text similarity of user reviews for improving collaborative filtering recommender systems. *Electronic Commerce Research and Applications*, *45*, 101019.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Gonsalves, B., & Patil, V. (2017, March). LoQoS location and QoS sensitive web service recommender. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1-4). IEEE.

Hashim, E. W. A., Hammood, M. O., & Al-azraqe, M. T. I. (2016). A Cloud Computing System Based Laborites' Learning Universities: Case Study of Bayan University's Laborites-Erbil. *Book of Proceeding*, 538.

Kamaran faraj. (2014). E-recruitment Tiered Architecture in Feasibility Study Role". International Journal of Computer Sciences and Engineering ,2015.

Kamaran faraj. (2016). Web-based Teaching in Particular Developing Counties, Experience at "Sulamani University". International Journal of Computer Sciences and Engineering Vol.-4(3), PP(01-04) Mar 2016, E-ISSN: 2347-2693.

Kh, T., & Hamarash, I. (2020). Model-Based Quality Assessment of Internet of Things Software Applications: A Systematic Mapping Study.

Li, S., Wen, J., Luo, F., Gao, M., Zeng, J., & Dong, Z. Y. (2017). A new QoS-aware web service recommendation system based on contextual feature recognition at server-side. *IEEE transactions on network and service management*, *14*(2), 332-342.

Li, Zhen, and Aoi Shimizu. "Impact of online customer reviews on sales outcomes: An empirical study based on prospect theory." *The Review of Socionetwork Strategies* 12.2 (2018): 135-151.

Manikandan, V., Gowsic, K., Prince, T., Umamaheswari, R., Ibrahim, B. F., & Sampathkumar, A. (2020, September). DRCNN-IDS Approach for Intelligent Intrusion Detection System. In *2020 International Conference on Computing and Information Technology (ICCIT-1441)* (pp. 1-4). IEEE.

Marchand, A., & Marx, P. (2020). Automated product recommendations with preference-based explanations. *Journal of retailing*, *96*(3), 328-343.

Miran, A., & Kadir, G. (2019c). Enhancing AODV routing protocol to support QoS. *International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE)*, *8*(5), 1824–1830.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Mohammed, A. G. (2021). A Study of Scheduling Algorithms in LTE-Advanced HetNet. *QALAAI ZANIST SCIENTIFIC JOURNAL*, *6*(3), 945-968.

Mohammed, A. S., Andrii, S., Ahmed, A. N., Olha, D., & Owaid, S. R. (2019, December). Multiparametric Assesment of the Conditional Channel of Multitarent Radio Communication Systems using Fuzzy Sets. In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 374-380). IEEE.

Mohanty, B. K., & Passi, K. (2010). Agent based e-commerce systems that react to buyers' feedbacks–A fuzzy approach. *International Journal of Approximate Reasoning*, *51*(8), 948-963.

Pratik Shukla, Roberto Iriondo 2021. Natural Language Processing (NLP) with Python. Tutorial on the basics of natural language processing (NLP) with sample coding implementations in Python. (pp. 11-12). IEEE.

Qasim, A. J., Din, R., & Alyousuf, F. Q. A. (2020). Review on techniques and file formats of image compression. *Bulletin of Electrical Engineering and Informatics*, *9*(2), 602-610.

Sassi, I. B., Mellouli, S., & Yahia, S. B. (2017). Context-aware recommender systems in mobile environment: On the road of future research. *Information Systems*, *72*, 27-61.

Selmi, A., Brahmi, Z., & Gammoudi, M. M. (2017, June). PACT: A new trust prediction method for multi-agents recommender systems. In *2017 IEEE 26th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 9-14). IEEE.

Sharma, A., & Kumar, S. (2021). Network Slicing and the Role of 5G in IoT Applications. In Evolution of Software-Defined Networking Foundations for IoT and 5G Mobile Networks (pp. 172-190). IGI Global.

Shen, L., Liu, Q., Chen, G., & Ji, S. (2020). Text-based price recommendation system for online rental houses. *Big Data Mining and Analytics*, *3*(2), 143-152.

Shi, Y., Lin, H., & Li, Y. (2017, August). Context-aware recommender systems based on item-grain context clustering. In *Australasian Joint Conference on Artificial Intelligence* (pp. 3-13). Springer, Cham.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Tewari, A. S. (2020). Generating items recommendations by fusing content and user-item based collaborative filtering. *Procedia Computer Science*, *167*, 1934-1940.

Tripathi, A., Ashwin, T. S., & Guddeti, R. M. R. (2019). EmoWare: A context-aware framework for personalized video recommendation using affective video sequences. *IEEE Access*, *7*, 51185-51200.

Walek, B., & Fojtik, V. (2020). A hybrid recommender system for recommending relevant movies using an expert system. *Expert Systems with Applications*, *158*, 113452.

Wu, B., & Ye, Y. (2020). BSPR: Basket-sensitive personalized ranking for product recommendation. *Information Sciences*, *541*, 185-206.

ڕاسپاردەی ئەلکترۆنی لەسەر بنەمای پێداچوونەوەی کڕیار بە بەکارهێنانی داڕشتنی (دەرهێنانی) دەق

پـوختـە:

لەم پەڕەیەدا کڕیارەکان بە دوای پێداچوونەوەی بەرهەمەکان دەگەڕێن پێش کڕینیان. لە ڕوانگەی ئەمە، سەکۆکانی کڕیاری سەرهێڵ هانی کڕیارەکانیان دەدەن بۆ دابینکردنی پێداچوونەوە لەسەر بەرهەمەکان کە یارمەتی کڕیارانی داهاتوو و دابینکەری خزمەتگوزاری دەدات بۆ بەرزکردنەوەی خزمەتگوزاریەکانیان ئەم پێداچونەوانە بە شێوەیەکی ئاسایی لە زمانی سروشتیدا بە زۆری لە زمانی ئینگلیزیدا. ئەمانە بەکاردێت بۆ شیکردنەوە و دابینکردنی داتا کە بەکاردێت بۆ چاککردنەوە و دروستکردنی بەرهەمی نوێ چونکە زۆربەی خزمەتگوزاریەکان ناتوانن بەبەردەوامی پێداچوونەوەی بەکارهێنەران بکەن، بۆیە پێویستیان بە ئامرازی کانکاری هەیە بۆ فێربوون دەربارەی ئەو پێداچوونەوەکەرانە، ئەمەش ئەو شتەیە کە بەکارهێنەران پێویستیان بە شمەکە هەڵسەنگاندنەکانی پێداچوونەوەی بەکارهێنەران بۆ بەرزکردنەوەی بەرهەمەکانیان پێداچوونەوەکان شیدەکرێنەوە لەلایەن کڕیارەکانەوە بۆ بڕیاردان لەسەر کڕین یان نەکڕین ئامانجی سەرەکی ئەم لێکۆڵینەوەیە ئەوەیە کە سیستەمی ڕاسپاردە لەسەر بنەمای پێداچونەوەکانی کڕیارەکان پەرەپێبدەن، داتاسێتێکی پێداچونەوەکانی کڕیارەکان لە ئامازۆن و ئی بیەوە پەرەپێبدەن، پێداچوونەوەکان شی بکاتەوە بۆ

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (6), No (4), Autumn 2021
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

دروستکردنی بنکەی زانیاری بەرهەمەکان، پەرەپێدانی میتۆدێک بۆ کۆکردنەوەی پێداویستیەکانی بەکارهێنەر بە زمانی سروشتی و دیاریکردنی بەرهەمی سەرەکی، هەروەها پەرەپێدانی لۆگاریتمێک بۆ هاوتاکردنی داواکاری بەکارهێنەر لەگەڵ بەرهەمەکە و دروستکردنی ڕاسپاردەکان، تویژینەوەکە بەو ئەنجامە گەیشت کە بەرنامەی دیاریکراو گونجاوە بۆ شیکردنەوە و بەڕێوەبردنی داتای پێداچوونەوەکە.

التوصية الإلكترونية على أساس مراجعة العملاء باستخدام استخراج النص

الملخص :
في هذه الورقة ، يبحث المشترون عن مراجعات للمنتجات قبل شرائها. وفي ضوء ذلك، تشجع منصات التسوق عبر الإنترنت عملائها على تقديم مراجعات حول المنتجات التي من شأنها مساعدة العملاء المستقبليين ومزود الخدمة على تعزيز خدماتهم. هذه الاستعراضات عادة ما تكون في اللغة الطبيعية في الغالب في اللغة الإنجليزية. وتستخدم هذه لتحليل وتوفير البيانات التي تستخدم لإصلاح وبناء منتجات جديدة لأن معظم الخدمات غير قادرة على مراجعة مراجعات المستهلكين في نفس الوقت بانتظام، لذلك هم بحاجة إلى أدوات التعدين لمعرفة المزيد عن هؤلاء المراجعين، وهو ما يحتاجه المستهلكون لسلعهم. يستعرض المستخدمون التقييمات الخاصة بترقية منتجاتهم. يتم تحليل المراجعات من قبل العملاء لتحديد ما إذا كانوا سيشترون أو لا يشترون. الأهداف الرئيسية لهذه الدراسة هي تطوير نظام توصية يستند إلى مراجعات العملاء ، وتطوير مجموعة بيانات من مراجعات العملاء من Amazon و eBay ، وتحليل المراجعات لإنشاء قاعدة بيانات للمنتجات ، وتطوير خوارزمية لتوليد درجات إيجابية / سلبية للمنتج ، وتطوير طريقة لجمع متطلبات المستخدم باللغة الطبيعية وتحديد المنتج الرئيسي ، وكذلك تطوير خوارزمية لمطابقة طلب المستخدم مع المنتج وتوليد توصيات. تطوير طريقة لجمع متطلبات المستخدم باللغة الطبيعية وتحديد المنتج الرئيسي، وكذلك تطوير خوارزمية لمطابقة طلب المستخدم مع المنتج وتوليد التوصيات. وخلصت الدراسة إلى أن البرنامج المختار مناسب لتحليل وإدارة بيانات المراجعة.