Semantic Annotation Based Mechanism for Web Service Discovery and Recommendation

J. Brindha Merin[1], Dr.W. Aisha Banu[2*], R. Akila[3] and A. Radhika[4]

[1] Assistant Professor (Senior Grade), Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai. brindhamerin@crescent.education, Orcid: https://orcid.org/0000-0001-9736-1061

[2*] Professor, Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai. aisha@crescent.education, Orcid: https://orcid.org/0000-0001-7799-6393

[3] Assistant Professor (Senior Grade), Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai. rakila@crescent.education, Orcid: https://orcid.org/0000-0003-4000-4535

[4] Assistant Professor (Senior Grade), Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai. radhika.a@crescent.education, Orcid: https://orcid.org/0000-0001-7959-4830

Received: June 11, 2023; Accepted: July 22, 2023; Published: September 30, 2023

Abstract

Web Mining is regarded as one among the data mining techniques that aids in fetching and extraction of necessary data from the web. Conversely, Web usage mining discovers and extracts essential patterns usage over the webs which are being further utilized by various web applications. In order to discover and explore web services that are registered with documents of Web Services-Inspection, Discovery and Integration registry, Universal Description wants specific search circumstance similar to URL, category and service name. The document of Web Service Description Language (WSDL) offers a condition of the web services customers to take out operations, communications and the service format of right message. Therefore, WSDL is being utilized together with semantic explanation dependent substantiation for the extraction of different web services for related purpose, other supporting operations and attributes. The reason is that there subsist different web services having corresponding functionalities however altered or changeable attributes that are non–functional. Resultant, recognize the preeminent web service become tiresome for the user. A method is projected which caters the analysis of service resemblance with the aid of semantic annotation and machine learning (ML) algorithms depending on the analysis intended for enhancing the classification through capturing useful web services semantics related with real world. The emphasizes on the research technique of choosing preeminent web service for the user based on the semantic annotation. The research work in turn recommends a web mining technique that determines the best web service automatically thus ranking concepts in service textual documentation and classifies services on behalf of particular domains. Parallel computation is made easy with web services. The different management stages in the system of recommendation entail collection of dataset through WSDL on the semantic annotation basis, thereby recognizing the best service with

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 3 (September), pp. 169-185. DOI: *10.58346/JOWUA.2023.I3.013*

*Corresponding author: Professor, Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai.

the DOBT-Dynamic operation dependent discovering method, ranking through mechanisms MDBR - Multi-Dimensional based ranking, recommendation and classification. In this work, it has been employed a combination of fundamental ML estimators, namely Multinomial Naive Bayes (MNB) and Support Vector Machines (SVM), as well as ensemble techniques such as Bagging, Random Forests, and AdaBoost, to perform classification of Web services. It was observed from the investigate work that the adapted system of best web services recommendation defers high performance in contradiction of the existing recommendation technique regarding accuracy, efficiency in addition to processing time.

Keywords: Recommendation, Web Service, Classification, Semantic Annotation, Description Language, Web Mining, Ranking, Multi-Dimensional Based Ranking Approaches, Machine Learning.

1 Introduction

In the present situation, there has been a significant rise in the usage of Internet by means of the social network users (Suma Singh, S., 2018). Since past few years, web has integrally become a robust platform for the analysis, retrieval and distribution of information (Irfan, S., 2018). By the means of the powerful tool of Search engines, any sort of information retrieval has become extremely feasible for any user. There are three types of Mining research namely:

a) Web Content Mining - mining the web content
b) Web Structure Mining - mining the web structure and
c) Web Usage Mining - mining the web use (Ensan, A., 2012).

The approach of web usage mining refers to employing the approach of data mining across the web access logs for optimizing the web site in accord with the interest of the user (Rajeswari, B., 2018) (Djelloul Bouchiha, 2012). The principle behind association of semantics with web service elements lies in determining the most suitable semantic approach in ontology for a WSDL element. The categorization technique emphasizes on classifying WSDL (Web Service Description Language) service description to its respective domain. As a result, there is a breakdown of service description to its basic WSDL elements. With the rise in various web services, there occurs a requirement of automated tools for identifying and composing web services. The various handling stages in the recommended system involves: dataset collection via WSDL on the basis of semantic annotation, identifying the best service by the means of DOBT-Dynamic operation based discovering techniques, ranking by the means of MDBR - Multi-Dimensional based ranking mechanisms, classification and recommendation. The UCI dataset is utilized for the collection of Web services of semantic annotation in the WSDL document. During the Pre-processing, unnecessary WSDL document attributes are eliminated from the actual/original WSDL document. Computation of dynamic offered and QoS similarities, similarities between the messages and path port numbers is being performed. The users can alter their data dynamically in the web services. Semantic annotation algorithm based on the approach of DOBD (Dynamic Operation Based Discovering) for domain webpage has been proposed in the research by carrying out a comparison of ontology data of extractive webpage. The web service ranking technique is being proposed for ranking and selecting the services by the means of MDBR (Multi-Dimensional Based Ranking) as well as by analyzing the user’s behavior. Content involving rich semantic information related to operations and data types of the different WSDL files are related to the ontology abstract and concepts (Alhindi, H., 2021). By contributing the individually perceived web service information to user, the best web service can be recommended to the users on the basis of ranking approach. With the output it’s revealed that the proposed algorithm yields good performance compared to the conventional method in terms of semantic annotation rate of support.

2 Related Work

Aidan Hogan et.al, presents the traditional search engine framework of SWSE that involves crawling, indexing, data enhancing and a user interface to search, browse and fetch information. In contrast to other conventional search engines, SWSE operates over RDF Web data – also referred to as Linked Data – that portrays unique challenges concerning the system design, interface, algorithms, architecture and implementation. Overall, there exist various challenges in employing Semantic Web technologies for Web data. The existing Semantic Web standards neglect the various web challenges, like inconsistency, scale, noise and unreliability (Hogan, A., 2011).

Neha Sharma et.al, proposes a semantic based perfecting technique for the web browsers for surmounting the existing systems limitations. This technique makes use of anchor text that exist in the hyperlink of the webpage, for identifying and exploring patterns. By the means of anchor text, it's possible to determine the web page ranking received by the search engine. This semantic based pre fetching technique employs Decision Tree Induction for finding out the probability of the anchor text and as well as the patterns to be pre-fetched. The proposed mechanism makes use of SPRINT as the decision tree induction method (Sharma, N., 2014).

Shady F. Samara recommends a technique that helps in building a composition amidst web services at run time. This approach focuses on accelerating the web services in the educational environment. The approach appends semantic description to the existing web services present in the local server, thereafter clustering them and expanding the consumer query words via WORDNET to facilitate the web services discovery process (Maghari, A., 2018).

LI Yuan-jie et.al recommends that the general WSDL files are considered as the study object. It's not possible to directly apply the conventional method of document classification as the web service is defined by WSDL. The research work proposes and presents an approach using which automatic web service semantic annotation can be applied. It utilizes three classification methods namely, Naïve Bayes, REPTree and SVM. In addition, ensemble learning is implemented. An accuracy of 87.39% is achieved by the experiment carried over on Nine Hundred and Fifty-One WSDL files and nineteen categories (Yuan-jie, L., 2012).

A. K. Tripathy et.al, recommends clustering of web services on the basis of QoS for facilitating the process of discovery and composition. The composition process can be considered as a workflow wherein every node in the workflow symbolizes a specific job. Thereafter, during run time the algorithm obtains appropriate WS which can gain the node job with the best QOS. The algorithm is based upon the shortest path in order to meet the composition output. Manual pre-planned composition workflow is being adopted with nodes for automatically triggering the WSs during run time. Thus, the proposed algorithm can be categorized under semi dynamic web services composition (Tripathy, A.K., 2014).

Ms. Ashwini Chavhan et.al, proposes an approach of web service ranking for ranking and selecting the web services via rank aggregation method by identifying user behavior. For analyzing user's behavior, user's history of invocation is made use of. Among the various approaches for web service ranking purpose, first being the approach of Functional relevance and QoS parameter, second being the CF based score and QoS and third is taken into account all the three parameters QoS utility, FR - Functional relevance and CF score (Chavhan, M.A., 2017).

Dessislava Petrova-Antonova et.al, recommends a pattern for Web Service Description Language based testing of individual as well as composite web services. The model is certainly effective and is imbibed as a software tool which aids in automating the testing activities as suggested by the pattern.WS-

BPEL is the general language for representing the business processes that offers a list of actions elaborating the composition of web service (Petrova-Antonova, D., 2015).

M. Nandhini et.al, recommends the method of optimizing the composition of web services relying upon the algorithm based on genetic. It considers the combination functionally as well as the matching of semantics based on quality also, and thus recommends a new model which is extensible that makes use of QoS for computing the new dimension for matching the quality of semantics (quality based functional metric), which is utilized as a measure for optimizing and ranking (Nandhini, M., 2013).

Bo Cheng et.al, recommends a novel approach of Web services discovery that extracts the essential structure of the semantics used for determining the interface factors for interaction, helping the users identify and adopt Web services and match the interfaces with accuracy which is high when the interfaces parameters involve synonyms that has meaning, abbreviations, and combine fragments that are not ordered. The proposed system emphasizes on extracting the underlying semantics. Initially, a concept-based Web Service description model is proposed that involves the type of path for the interface that interacts with factors apart from the conventional text description. Eventually, a WS Operations Discovery algorithm is presented (Cheng, B., 2016).

M. Kalyani et.al, recommends a new collaborative-based filtering model for recommendation of the web service system that aids the user in selecting the service that has optimal performance of QoS. The proposed recommender system for web makes use of the information based on the location and QoS values for clustering, thus framing personalized recommendation service for users depending on the results of the cluster. In comparison to the prevailing service recommendation methods, the recommended system yields better performance on the recommendation accuracy (Suchithra, M., 2015).

Omair Shafiq et.al, presents a novel approach that permits a semantically empowered representation of logs at the time of Web Service execution and thereafter utilizes these logs for carrying out ranking and adaptation of discovered WSs. It has been observed that by blending both the approaches into a hybrid approach allows the Web Services data to be represented formally which further aids in boosting data mining and machine learning based solutions for processing such kind of data. Semantic FP-Trees oriented technique is generated for carrying out association rule learning on different attributes of Web Services (Kalyani, M., 2015).

M. Suchithra et.al, presents web services ranking relying upon the QoS parameters. The prime motto is to identify the finest WS depending on the request of the user. Six major QoS parameters are taken into account for the process of re-ranking, these are: throughput, response time, accessibility, availability, cost and interoperability. The output clearly reveals that by using the non-functional QoS parameters at the time of web service discovery process enhances the probability of choosing the best web services. The recommended method utilizes Relevancy Function that is the foundation for ranking web services (Sánchez, D., 2015).

Deivamani Mallayya et.al, presents an algorithm of UPWSR (user preference-based web service ranking for web services ranking depending on the preference of the user and QoS factors of the WS. In case the request of the user remains unfulfilled through a service, rest other prevailing services must be collected. The recommended algorithm permits the user to define the local and global constraints for composite web services that aids in enhancing flexibility. UPWSR algorithm determines best and most appropriate services for the request made by user. By selecting the count of candidate services for individual task, the time taken to produce the composition plans can be minimized (Shahidul Islam, 2018).

Lijun Duan et.al, recommends a system through fetching the hidden relationships between user and the service by considering the social link. It's possible to produce a set of candidate services using the proposed approach via complementary manner, wherein the discovery of the web service and recommendation canco-operate depending on the formalized social link (Mallayya, D., 2015).

3 Proposed Work

Overview

Among the various data mining techniques available today, Web mining stands as a very significant one. Web mining helps in extraction and fetching of essential information from the web services. Categorization and Matching are the two major processes in the annotation approach. Categorization involves classification of WSDL to its respective domain. Matching process involves mapping of WSDL entities to the already existing domain ontology. It's essential that the service providers focus on the specification of the WSDL documents as it helps in effectively discovering and comprehending Web Services from their WSDL documents.

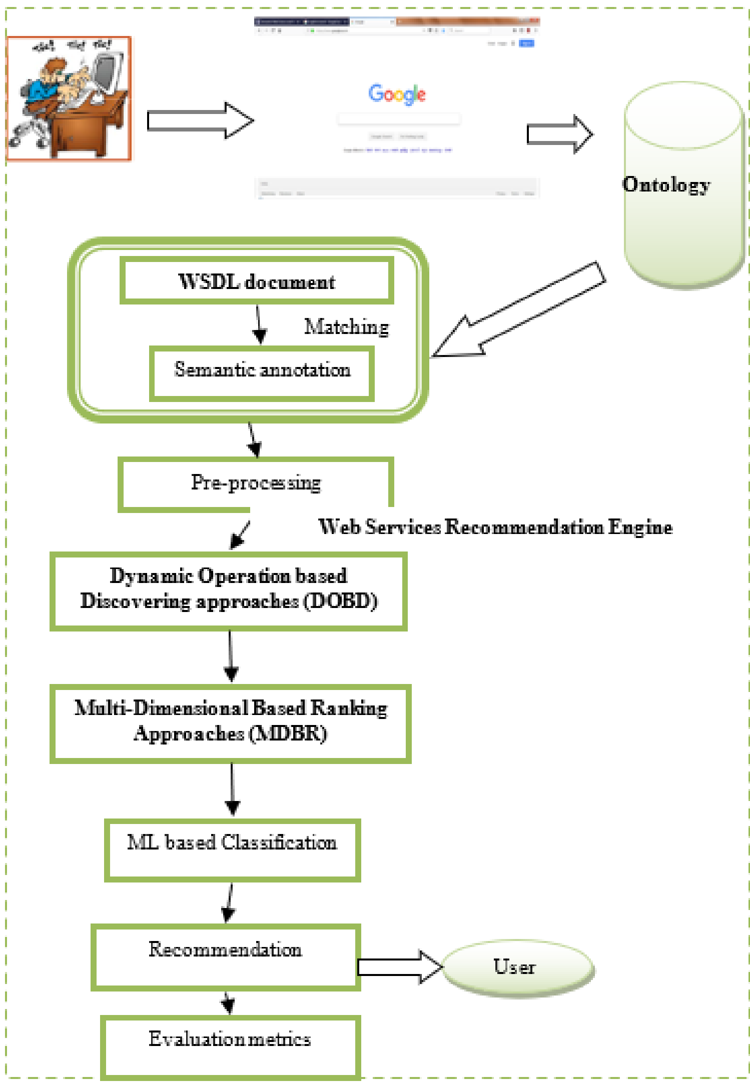

Figure 1: Architecture of Proposed System

For the growth of service-oriented applications, Web Service discovery stands as a crucial factor. For building a system that triggers external WSs, the service registries must be thoroughly inspected by the developers for discovering and retrieving by the means of DOBD (Dynamic Operation based Discovering) approaches as well as for linking the service descriptions WSDL documents with functionality, it must be adopted in the system. MDBR (Multi-Dimensional Based Ranking Approaches), the most effective search engine is based on this algorithm. MDBR is well integrated into the web services analysis and offers ranking of each web services based on the number of links from other Web services directing to those services. The historical usage data classifies and proposes the users, the web services

Fig 1. Architecture of Proposed System operations that are most appropriate according to their requirements. Based on the comprehensive assessment the effectiveness, efficiency and usability of their commended approach is presented.

Dataset Collection

According to the WSDL document, services refer to the collections of network ports or end points. Semantically annotated Web services in the WSDL file is assembled from the UCI dataset. The existing approach refers ontology as a group of entities /concepts, WSDL file and entities that includes interface, XSD data types, tag name, operations, port numbers, messages, and path.

Semantic Annotation

Annotation process involves two phases: first is the Categorization phase that classifies WSDL documents with their respective domain (2) second is the Matching phase that involves mapping of WSDL entities to the already existing ontology of the domain. The process of annotation is based on the techniques that match ontology which utilizes few similarity measures. The process used to match aids in mapping the entities of the WSDL files to the ontology concepts. Similarities amidst WSDL entities and the chosen ontology concepts are then computed for determining the concept that is to be attached to the initial element of WSDL. The same operation is carried out for all the WSDL entities. The concept of domain ontology is used for every annotation of the element of the WSDL document

Data Pre-processing

There exist lots of noisy, incomplete and inconsistent data in the real word. Hence there is a necessity to eliminate unnecessary and junk data from the original data. Web service duplication involves removal of tags and extraction of purely text contents for performing duplicate evaluation. Thereafter discarded WSDL document attributes are eliminated from the actual WSDL document. For few transactions, there can be inconsistencies recorded in the WSDL document. In addition, there can be inconsistencies because of WSDL document integration, in which the same WSDL document is denoted by different names or a given characteristic may possess different names in different databases.

Dynamic Operation Based Discovering Approaches (DOBD)

DOBD (Dynamic Operation based Discovering) approaches coordinate with Semantic Web technologies for enhancing the Web services performance. To offer semantics for Web service descriptions, ontologies, user-defined dictionaries or corpora were leveraged. User give list of required web services that is being chosen from the Service Discovery. Besides, each user can offer reliability rating that can be assessed. Generally, users have one ratting values, though web services provide certain dynamic offering in order to leverage the business. Users provide attribute values according to the increasing and decreasing rating in Web services. Eventually, on the basis of ratings, discovery-based web services can be analyzed.

Input: *Web services giving to WSDL document*

Output: *Web services Dynamic offering to discovering*

```
Initialized set of web services (WS) →{WS1, WS2, WS3…..WSn}
Initialized web services rating from Users(R) →{R1, R2, R3…..Rn}
Begin
{
Discovering (Web services WS, Rating R)
Compute to frequency of rating given by user per Web Service;
Dynamic changing by attribute values (A1, A2, A3……An);
Initialize i;
(WS (i) =1; WS (i) <=WS length (); i++)
{
If (Whole WS (i)>= (threshold))
        If (Attribute value (i) <=max (i))
                {
                Values=max for given Web Services
        Ratting++
        Else
        Ratting--
Display the sorting order to discovering;
Finally discovered in ratting based Web Services
End
                }
        }
}
```

Multi-Dimensional Based Ranking Approaches (MDBR)

With the existence and availability of huge number of WS, choosing the best and most appropriate web service can be tedious enough. As a result, it becomes highly essential to build a robust algorithm for ranking the web services so as to filter and choose the best one. By incorporating the MDBR approach this issue can be sorted by imbibing the technique of best web service selection which utilizes QoS and is an energy effective approach. The approach of Multi-Dimensional Based Ranking Approaches accepts user request that involves set of web services. The user places request for the suggestion and Service selection (QoS). The similarities attributes are evaluated on the basis of dynamic offered and QoS similarities. Similar set of web services are identified based on similar values. Basic service ranking are extracted based on the prior usage experience of same users. For improvising the accuracy of the ranking, relational degree can be employed which aids in gaining successful and best service ranking results. It also helps in determining maximum threshold values and most alike web services to rank. Subsequently, there is ranking on the basis of maximum QoS attributes value as well as maximum threshold and ratings for ranking the web services. The ranking result is provided to the active user.

Input: *Web services Dynamic offering to discovering*

Output: *Ranking to best web services*

```
Initialized set of web services (WS) →{WS1, WS2, WS3…..WSn}
Initialized web services rating from Users(R) →{R1, R2, R3…..Rn}
Initialized set of QoS Attributes (QA)→ {QA1, QA2, QA3……QA18}
Begin
{
Discovered in rating based web services
WS→Offering some attributes;
This 18 attributes are represented by QA1, QA2, QA3……QA18;
Path;
Port number;
Data type;
Threshold=30;
For (WS=1;WS<=QA;Ws++){
if (WS(QoS==QA) || (QA >= threshold)||(Ratting (R>=Max))){
To selecting WS
Forward another layers
i++
When the WS == minimum;
Based on User_QA1();User_QA2();User_QA3();
Repeat step QA;
Finally→ single output of (one WS)
End
}}}
```

Classification

Classification can enhance the potential of processing WSDL file after continuous training process. Emphasis is laid upon training various classifiers referred to as weak classifiers via the same training set which are combined to produce an improvised classifier referred to as strong classifier. This can be achieved by modifying the distribution of the WSDL file. Weight of each sample is stated by the classification output of each sample and the general classification accuracy of the prior classification process. To produce the final decision classifier, the new dataset with modified weight is utilized as the next layer classifier's training set, by combining the trained classifiers.

The method of web service classification involves the automated categorization of a service into preset categories, utilizing feature vectors and similarity measures. Supervised classification encompasses two distinct stages, namely training and classification. During the training phase, a subset of the classified document dataset pertaining to services is selected to train the classifier in order to acquire knowledge about the different classes. During the classification step, the classifier infers one or more categories for a new document with varying levels of accuracy. The categorization of Web services and their descriptions is influenced by the terminology used by service designers. This results in dependencies between the service category and its description. By leveraging this latent information, classifiers are able to accurately group services into functionally similar categories. The class prediction accuracy of a single classifier is contingent upon its inductive bias and generalization error. This is due to the fact that several classifiers may select distinct patterns within the input data in order to do the classification. The primary aim of employing ensemble learning methodologies is to amalgamate the collective forecasts generated by diverse basic machine learning classifiers, with the intention of attaining an enhanced model that may mitigate the constraints inherent in each individual classifier. The classification accuracy of the service dataset was assessed using both base estimators, including Multinomial Naive Bayes (MNB) and Support Vector Machines (SVM), as well as ensemble approaches such as Bagging Classifier, Random Forests Classifier, and AdaBoost Classifier. The scikit-learn machine learning package for Python was utilized in our study.

The Naive Bayes classifier is a type of supervised classifier that use Bayes' theorem to analyze provided data, making the assumption that the features are independent. The Multinomial Naive Bayes (MNB) technique is utilized to classify texts that are represented as word or frequency vectors. It is specifically designed for multinomially distributed data.

The Support Vector Machine (SVM) is a supervised learning technique that use hyperplanes to categorize data with large dimensions (Yuan-jie, L., 2012). Various kernel functions are offered for utilization in the decision function, allowing for customization based on the specific characteristics of the data.

Ensemble techniques are distinguished by the manner in which they aggregate the predictions generated by the individual machine learning algorithms that comprise them. There are two distinct types that may be discerned in this context, namely averaging methods (such as Bagging and Random Forests) and boosting methods (such as AdaBoost and Gradient Tree Boosting). Three distinct ensemble approaches were employed.

In the Bagging Classifier, the ensemble consists of individual classifiers that are trained on a randomly redistributed version of the training set (Breiman, L., 1996). In the process of training a classifier, a training set is created by randomly selecting N samples with replacement from the original dataset, which also has a size of N. As a consequence of the random selection process, there is a possibility of either overlap or exclusion of some training samples. This implies that the outcome of each individual classifier is produced using a distinct random sampling technique applied to the training set.

Random Forests (Breiman, 2001) is a machine learning technique that constitutes an ensemble of several decision trees. Random Forests are a classification algorithm that utilizes decision trees. In this algorithm, each decision tree is built using a random subset of the training data. During the classification process, each decision tree contributes to the final prediction by voting for the most commonly occurring class label. As a result, Random Forests operate as a majority voting-based classifier.

The AdaBoost classifier, also known as Adaptive Boosting, conducts classification by training a single classifier on the original dataset (Al-Stouhi, S., 2011). In the future phases of the process, many iterations of the classifier are employed on the identical dataset. However, the weights assigned to instances that were improperly categorized are modified. This adjustment aims to prioritize challenging situations, hence enhancing the overall accuracy of the subsequent classifiers. In summary, averaging approaches include the utilization of many independent base estimators, with the ultimate outcome being determined as the average of the data obtained from each estimator. Boosting-based techniques employ a sequential arrangement of constituent classifiers, resulting in the cumulative effect of each classifier contributing to the ultimate outcome. Therefore, in this study, we employ a combination of fundamental machine learning estimators, namely Multinomial Naive Bayes (MNB) and Support Vector Machines (SVM), as well as ensemble techniques such as Bagging, Random Forests, and AdaBoost, to perform classification of Web services. We evaluate the impact of these approaches on the accuracy of classification by utilizing several service feature vector models that have been suggested in previous research.

Recommendation

The recommendation system relies upon its learning by making use of the user's feedback, prepares the list and finally provides it to the user. List of web services presented to the user can be a single service or may be a composition of services. From the given service list, the user can opt for any service. Post execution, the user can give a rating to the selected service via given metric which basically resembles the satisfaction level of the user. This rating is made store in a repository which can be later utilized as an input for the recommendation.

Evaluation Metrics

There are certain parameters that are computed and evaluated for carrying out the performance assessment and computing system stability. These are listed as following:

The performance of the proposed Multi-Dimensional Based Ranking Approaches (MDBR) is calculated using Root Mean Square Error (RMSE), recall sensitivity, precision, probability misclassification error (PME), F-Score and accuracy of the ranked set, rank set and overall performance is evaluated through the Equations 1to7 correspondingly. Here Yi denotes the actual and R_i denotes the outcome of the achieved web service detection attribute (num), TN - True Negative represents the detection for web services that actually had web services, FN - False Negative (FN) represents the detection for web services that had web services, TP - True Positive (TP) depicts the detection for web services that had web services, and FP - False Positive (FP) denotes detection for web services that had web services.

- True Positive (TP): If it is positive instance classification is positive
- False Negative (FN): If it is positive instance classification is negative
- True Negative (TN): If it is negative instance classification is negative
- False Positive (FP): If it is negative classification is positive

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - R_i)^2} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Recall = \frac{TN}{TN + FP} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$F - score = 2 * \frac{Recall * Precision}{Recall + Precision} \tag{5}$$

$$PME = \frac{FP + FN}{TN + FP + TN + FN} \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{7}$$

4 Results and Discussion

A general framework is proposed in the research work for the ranking, discovery and recommendation of composite web services. Ontology is adopted for carrying out the semantic annotation of the WSDL document discovery of composite web services and for proposing the best optimum discovered composite web services. In addition, there is an elaboration of ranking and recommendation as well as architecture and the components of web service discovery. The research extends the web service discovery and ranking algorithm for including constraint specifications.UCI dataset acts as the basis for collecting data and JAVA environment helps in the overall development process. Complete dataset is maintained using MYSQL and results obtained are also stored in MYSQL.

Table 1 depicts and compares performance of the recommended discovery techniques with the available techniques such as the CSDRM (Collaborative Web Service Discovery and Recommendation Mechanism), TRTSD (Trustworthy Service Discovery based on Trust and Recommendation relationships) and REM (Random Ergodic Matching). The recommended approach of DOBD (Dynamic Operation Based Discovering approach) yields in high performance in contrast to other prevailing techniques.

The DOBD is particularly noteworthy for its remarkable efficiency, as it attains an amazing efficiency rating of 75.49%. This finding suggests the efficacy of the system in rapidly discerning pertinent online services. In addition, the DOBD system has exceptional computing efficiency, with a processing time of about 0.66 milliseconds. On the other hand, the CSDRM demonstrates a moderate efficiency level of 50.53%, while concurrently ensuring a comparatively short processing time of 0.75 milliseconds. The efficiency values of the TRTSD and REM approaches are 29.91% and 33.58%, respectively. TRTSD requires 1.5 milliseconds, whereas REM requires 1.2 milliseconds, making TRTSD slightly more time-consuming. These numbers play a crucial role in evaluating the balance between efficiency and computing time, hence aiding in the decision-making process for selecting the most suitable web service discovery approach based on unique application requirements.

Fig.2 compares the web services discovering techniques of DOBD (Dynamic Operation Based Discovering approach) with rest of the techniques such as CSDRM, TRTSD and REM. The recommended discovery approach of DOBD yields in high performance in contrast to other prevailing techniques.

Table 1: Web Services Discovering Techniques Comparison

S. No	Discovering Techniques	Efficiency (%)	Time(ms)
1	Collaborative Web Service Discovery and Recommendation Mechanism (CSDRM)	50.53	0.75
2	Trustworthy Service Discovery based on Trust and Recommendation relationships (TRTSD)	29.91	1.5
3	Random Ergodic Matching (REM)	33.58	1.2
4	Dynamic Operation Based Discovering approach (DOBD)	75.49	0.66

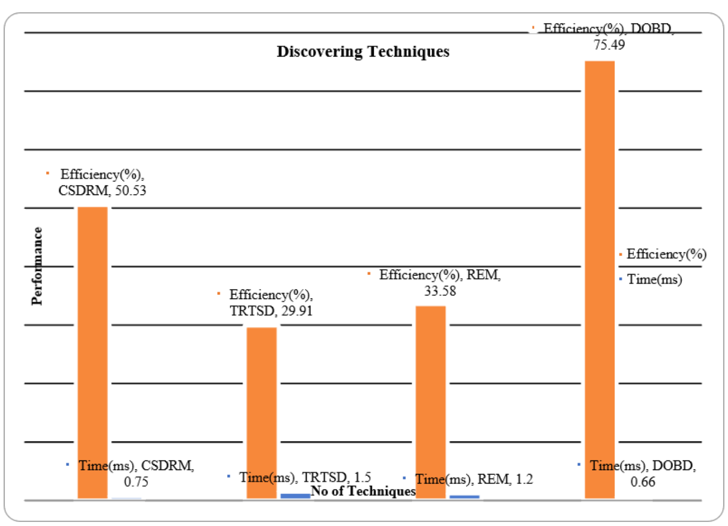

Figure 2: Web Services Discovering Techniques Performance

Table 2: Web Services Ranking Algorithm Comparison

S. No	No of Techniques	Accuracy (%)	Time (ms)
1	Skyline	55	5
2	User Preference based Web Service Ranking (UPWSR)	72	1
3	Multi-Dimensional Based Ranking Approaches (MDBR)	85	0.6

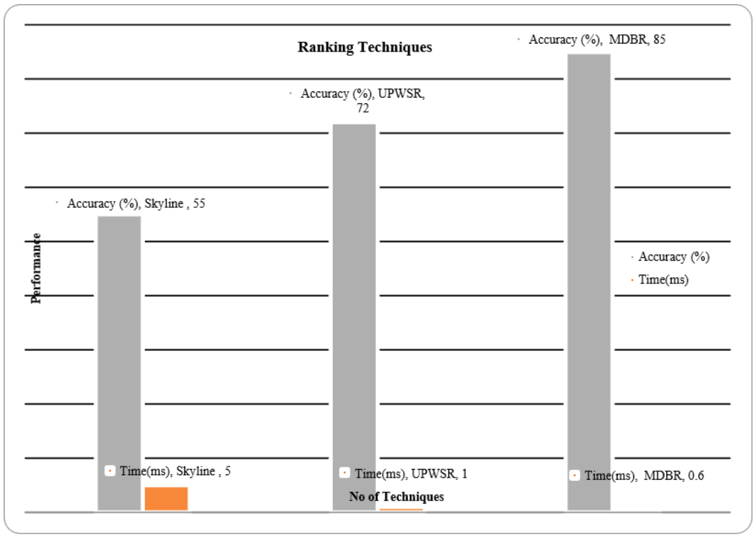

Figure 3: Ranking Algorithm Performance

Figure 3. compares the performance of MDBR (Multi-Dimensional Based Ranking) Approaches in contrast to the Skyline and UPWSR (User Preference based Web Service Ranking). The recommended approach of MDBR yields high efficiency in comparison to the rest of the existing techniques.

Table 2compares the performance of MDBR (Multi-Dimensional Based Ranking) Approaches in contrast to the Skyline and UPWSR (User Preference based Web Service Ranking). The recommended approach of MDBR yields high efficiency in comparison to the rest of the existing techniques. Among the tested strategies, the MDBR emerges as the most notable, boasting a remarkable accuracy rate of 85%. This result serves as a strong indication of the efficacy of MDBR in the identification of pertinent online services. In addition, the MDBR system exhibits notable efficiency, with a processing time of about 0.6 milliseconds. On the other hand, it is worth noting that the Skyline method exhibits a diminished level of precision, amounting to 55%, however accompanied by a little protracted execution duration of 5 milliseconds. The UPWSR approach achieves a trade-off between precision, with a success rate of 72%, and computational efficiency, since it only requires 1 millisecond for calculation. These values offer significant insights into the trade-offs between accuracy and computing efficiency in the context of web service ranking algorithms. They assist in the decision-making process of selecting the best appropriate approach based on unique application needs.

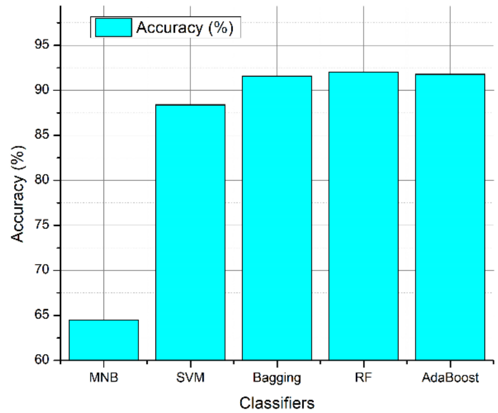

Figure 4: Accuracy for Various Classifiers Used in Web Service Classification

Figure 4 presents a comprehensive overview of the effectiveness of diverse classifiers in the classification of web services, as evaluated based on their accuracy. Among the studied classifiers, Random Forest (RF) demonstrates notable performance with an accuracy rate of 92%, hence highlighting its efficacy in accurately classifying web services. The Bagging and AdaBoost classifiers demonstrate notable levels of accuracy, with respective scores of 91.6% and 91.8%. both results suggest that both classifiers work reliably in the context of web service categorization tasks. The Support Vector Machine (SVM) has a high level of accuracy, achieving a robust 88.4% accuracy rate. On the other hand, the Multinomial Naive Bayes (MNB) classifier offers a decent accuracy rate of 64.5%. The accuracy numbers provided give significant insights into the appropriateness of each classifier for web service classification applications. This information can assist in the decision-making process of selecting the most suitable approach, taking into consideration the desired degree of accuracy.

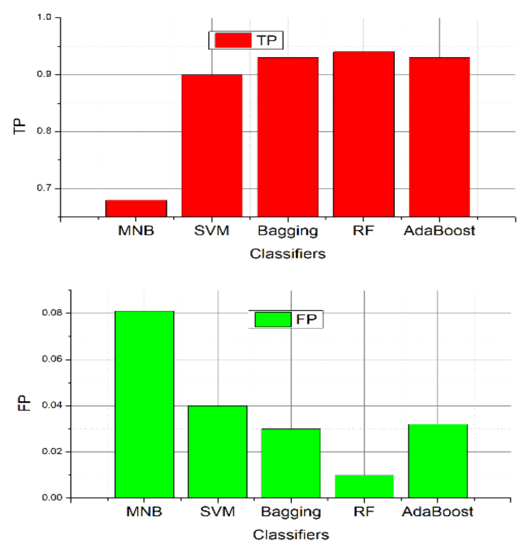

Figure 5: Classifier Performance in Terms of TP and FP Rate in Web Service Recommendation

Figure 5 offers valuable insights into the effectiveness of different classifiers in the context of web service recommendation. This is achieved by measuring their rates of TP and FP outcomes. RF is particularly noteworthy due to its high TP rate of 0.94, which indicates its ability to accurately suggest suitable online services. The RF model also has a remarkably low FP rate of 0.01, which underscores its precision in mitigating erroneous recommendations. The Bagging and AdaBoost classifiers demonstrate comparable performance, with true positive rates of 0.93 and reasonably low false positive rates of 0.03 and 0.032, respectively. The SVM has a high true positive rate of 0.9, whilst the MNB reveals a moderate TP rate of 0.68, accompanied with a false positive rate of 0.081. These values facilitate the evaluation of classifiers' efficacy in accurately suggesting pertinent online services while mitigating the likelihood of advocating irrelevant ones, hence assisting in the identification of the best appropriate classifier for web service recommendation jobs.

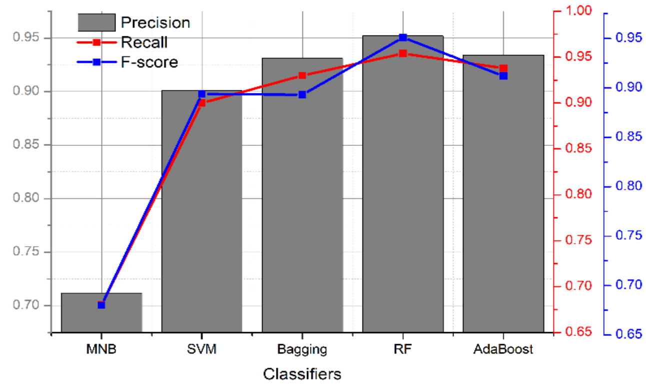

Figure 6: Classifier Performance in Terms of Precision, Recall, and F-score in Web Service Recommendation

Figure 6 presents a thorough evaluation of several classifiers within the domain of web service recommendation in terms of Precision, Recall, and F-score. The assessment takes into account the precision, recall, and F-score metrics. The Random Forest (RF) algorithm demonstrates superior performance, exhibiting a high accuracy value of 0.952 and recall value of 0.954. Consequently, it achieves an exceptional F-score of 0.951. The Bagging and AdaBoost classifiers have notable performance, achieving accuracy values of 0.931 and 0.934, along with high recall scores of 0.93 and 0.938, respectively. Consequently, these classifiers provide F-scores of 0.893 and 0.912, respectively. The Support Vector Machine (SVM) exhibits a noteworthy level of accuracy (0.901) and recall (0.9), resulting in a good F-score of 0.894. The Multinomial Naive Bayes (MNB) classifier has a somewhat reduced accuracy of 0.712, although it sustains a satisfactory recall of 0.68, resulting in a matching F-score of 0.68. The aforementioned values provide significant insights into the performance of classifiers, enabling informed decision-making in the selection of a classifier for web service recommendation. This consideration takes into account the trade-offs between accuracy and recall.

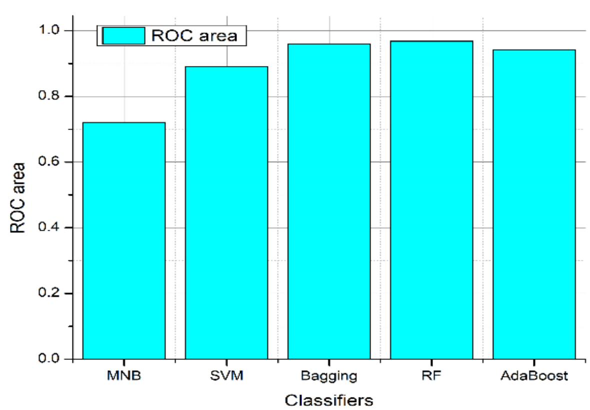

Figure 7: ROC Area for Various Classifiers in in Web Service Recommendation

Figure 7 presents a quantitative assessment of the classifiers' discriminatory capability and their aptitude for differentiating between positive and negative occurrences within the domain of web service recommendation. The Random Forest (RF) model has superior performance, with a remarkable ROC area of 0.968, which signifies its robust discriminating skills. Bagging has a high level of efficacy in discriminating between useful and irrelevant online services, as evidenced by its ROC area of 0.96. The Support Vector Machine (SVM) has a significant Receiver Operating Characteristic (ROC) area of 0.89, so highlighting its efficacy in accurately categorizing online services. AdaBoost has a robust performance, with a ROC area of 0.942. On the other hand, Multinomial Naive Bayes (MNB) reveals a somewhat lower, though still commendable, ROC area of 0.721. These values offer valuable insights into the capabilities of classifiers in accurately recommending and classifying online services. They assist in the process of selecting the best appropriate classifier for jobs related to web service recommendation.

5 Conclusion

For utilizing the real Web services technology merits, the research work recommends a method for annotate web services WSDL semantic descriptions through ontological model. Parallel computation makes the usage of web services easy. There are two main benefits of employing such models. The first benefit is that, it offers a dynamic discovering WSDL document portrayal regarding domain ontologies. The subsequent benefit being, the projected model permits the ranking dependent services of semantic web. Using a robust tool, the projected technique is being introduced. By means of this approach the user gets support in selecting preeminent optimum web services in case there happens a impediment. Furthermore, there is a time complexity reduction for the reason of the optimal value which ascertains QoS-Aware Services. It’s elucidated clearly form the yield that the suggested technique is quite competent on comparing the rest of techniques.

References

[1] Alhindi, H., Traore, I., & Woungang, I. (2021). Preventing Data Loss by Harnessing Semantic Similarity and Relevance. *Journal of Internet Services and Information Security (JISIS)*, *11*(2), 78-99.

[2] Al-Stouhi, S., & Reddy, C.K. (2011). Adaptive boosting for transfer learning using dynamic updates. *In Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2011, Athens, Greece. Proceedings, Part I 11*, 60-75. Springer Berlin Heidelberg.

[3] Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*, 123-140.

[4] Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

[5] Chavhan, M.A., & Badre, M.R. (2017). Web Service Ranking and Recommendation Using User History and Profile. *IJETT*, *1*(2), 8208-8212.

[6] Cheng, B., Zhao, S., Li, C., & Chen, J. (2016). A web services discovery approach based on mining underlying interface semantics. *IEEE Transactions on Knowledge and Data Engineering*, *29*(5), 950-962.

[7] Djelloul Bouchiha & Mimoun Malki (2012). Semantic Annotation of Web Services. *Proceedings ICWIT*, 60-69.

[8] Duan, L., & Tian, H. (2017). Collaborative web service discovery and recommendation based on social link. *Future Internet*, *9*(4), 1-12.

[9] Ensan, A., & Biletskiy, Y. (2012). Matchmaking through semantic annotation and similarity measurement. *In 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 1-5.

[10] Gohar, P., & Purohit, L. (2015). Discovery and prioritization of web services based on fuzzy user preferences for QoS. *In IEEE International conference on computer, communication and control (IC4)*, 1-6.

[11] Hogan, A., Harth, A., Umbrich, J., Kinsella, S., Polleres, A., & Decker, S. (2011). Searching and browsing linked data with swse: The semantic web search engine. *Journal of web semantics*, *9*(4), 365-401.

[12] Irfan, S., & Ghosh, S. (2018). Web Mining for Information Retrieval. *International Journal of Engineering Science*, 17277-17283.

[13] Kalyani, M., & Veera Kishore, K. (2015). Personalized QOS-Aware Web Service Recommendation Via Collaborative Filtering. *International Journal of Electrical, Electronics and Computer Systems,* 54-58.

[14] Maghari, A. (2018). Dynamic Web Service Composition Method Based on OWL-S for Educational Environment. *In Samara, SF, & Maghari, AY. Dynamic Web Service Composition Method Based on OWL-S for Educational Environment. In The First International Conference on Engineering and Future Technology (ICEFT)*, 134-142.

[15] Mallayya, D., Ramachandran, B., & Viswanathan, S. (2015). An automatic web service composition framework using QoS-based web service ranking algorithm. *The Scientific World Journal*, *2015,* 1-14.

[16] Nandhini, M., & Sadish M. Sendil, (2013). Web Service Quality Composition Determination Using Genetic Algorithm in Sematic Web. *International Journal of Computer Science and Information Technologies*, *3*, 4704–4706.

[17] Petrova-Antonova, D., Ilieva, S., & Stoyanova, V. (2015). A pattern for WSDL-based testing of web service compositions. *In Proceedings of the 20th European Conference on Pattern Languages of Programs*, 1-6.

[18] Rajeswari, B., & Nisha, S.S. (2018). Web page prediction using web mining. *International Research Journal of Engineering and Technology*, *5*(05), 4234-4237.

[19] Sánchez, D., Batet, M., Martínez, S., & Domingo-Ferrer, J. (2015). Semantic variance: an intuitive measure for ontology accuracy evaluation. *Engineering Applications of Artificial Intelligence*, *39*, 89-99.

[20] Shafiq, O., Alhajj, R., & Rokne, J.G. (2015). Reducing search space for web service ranking using semantic logs and semantic FP-tree based association rule mining. In *Proceedings of the IEEE 9th International Conference on Semantic Computing (IEEE ICSC)*, 1-8.

[21] Shahidul Islam, Britto Ramesh Kumar, S., Ab Rashid Dar (2018). An Approach for an Effective Web Service Selection using Filtering and Skyline Method. *International Journal of Computer Sciences and Engineering, 6*(7).

[22] Sharma, N., & Dubey, S.K. (2014). Semantic based Web prefetching using decision tree induction. *In IEEE 5th International Conference-Confluence the Next Generation Information Technology Summit (Confluence)*, 132-137.

[23] Suchithra, M., & Ramakrishnan, M. (2015). Efficient discovery and ranking of web services using non-functional QoS requirements for smart grid applications. *Procedia Technology*, *21*, 82-87.

[24] Suma Singh, S., Sangeetha, G. (2018). A Survey of Web Search from Web Documents Based on Semantic Ontology Technique. *American Journal of Engineering Research*, *7*(2), 284-287.

[25] Tripathy, A.K., Patra, M.R., Khan, M.A., Fatima, H., & Swain, P. (2014). Dynamic web service composition with QoS clustering. *In IEEE International Conference on Web Services*, 678-679.

[26] Yuan-jie, L., & Jian, C. (2012). Web service classification based on automatic semantic annotation and ensemble learning. *In IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*, 2274-2279.

Authors Biography

Dr.J. Brindha Merin is a Asst. Professor (Sr. G) in B.S. Abdur Rahman Crescent Institute of Science and Technology. She did her B.E from Anna University in Computer Science and Engineering. She completed her M.Tech. from Anna University. She is pursuing her Phd in the field of Data mining and Web Services. She has 11 years of experience in teaching and years of experience in industry.

Dr.W. Aisha Banu is a professor in B.S. Abdur Rahman Crescent Institute of Science and Technology. Her qualifications are as mentioned. Ph.D. (Computer Science and Engineering) from Anna University, M.E. (Computer Science & Engg.) from Anna University, B.E. (Computer Science & Engg.) from Madras University. She has 25 years of teaching experience and her areas of interest include Big Data, Cloud Computing, Social Network Analysis and Information Retrieval Total number of research publications are 30.

Mrs.R. Akila is a Asst. Professor (Sr. Grade) in B.S. Abdur Rahman Crescent Institute of Science and Technology. She did her B.E from Manonmaniam Sundaranar University in Computer Science and Engineering. She completed her M.E. from Anna University. She is pursuing her Phd in the field of Data mining and machine learning. She has 21 years of experience in teaching.

Mrs.A. Radhika Asst. Professor (Sr. Grade) in B.S. Abdur Rahman Crescent Institute of Science and Technology. She did her B.E from Anna University in Computer Science and Engineering. She completed her M.E. from Bharath University. She is pursuing her Phd in the field of Data mining and analysis. She has 12 years of experience in teaching.