Career Selection For Students In The Information Technology Sector Using A Hierarchical Fuzzy-Based Ontological Recommender System Approach

Ahsan Zubair¹, Asim Shahzad¹, Muhammad Zeeshan¹, Jamaluddin Mir², Muhammad Zubair Tahir¹, Aamer Khan¹, Muhammad Sibt ul Hassan¹, Muhammad Naeem¹

Mehmood Ahmed³

¹ Department of Computer Science, Abbottabad University of Science & Technology, Hevellian, Pakistan.

² Faculty of Computer Science & Information Technology, Universiti Tun Hossein Onn Malaysia, Batu Pahat, Malaysia.

³ Faculty of Computer Science, University of Haripur, KPK, Pakistan.

Abstract: Recommendation frameworks are usually utilized over the Internet to direct clients to discover the services or items that best fit their inclinations. Picking a profession between students is critical because of various human capabilities. Numerous students have chosen their career paths without legitimate guidance from appropriate experts or university administrations. This may cause conflict between the student's abilities, personality, academic achievements, and interests. To suggest students in career choices, it is fundamental to assemble a recommendation framework that gives guidance and direction to students in picking their profession. In the view of traditional techniques applied in previous works, such as e-learning, they center on one career option without looking into different possibilities in other work professions that likewise can be coordinated with their skills or capabilities. Consequently, this paper proposes a career development system driven by ontologies and fuzzy logic strategy. To improve the process of the suggestion in the career recommendation segment, the proposed framework gives a most realistic method of dealing with the client's question to limit the students alternative to assist with choosing the best career option utilizing fuzzy logic combined with ontology and M-tree hierarchy. So proposed problem statement puts the thought of a more accurate and precise career development system. Using semantic relations between ontology and fuzzy logic helps students choose a profession according to their skill sets.

Key Words: Ontology, Fuzzy Logics, Recommender Systems, E-learning Recommenders, M-Tree
1. Introduction

University or college is a period of significant change in choosing their career path for most students. Research has discovered that 75% of students enter school, college, or university without having the ultimate choices about their field of specialization. Cuseo et al. [1] and [2] calculated that 50% to 75% of students will change their major subjects at any rate once during their school or college career. A significant test for students is the trouble of solidifying information from unmined data sources. Here comes the enormous issue where students at secondary school battle to discover online valuable and exact data that can affect their arrangements plan to select a career. If we distinguish students' decision and interest, we can guide them toward enhanced opportunities like in which major to study and assets to investigate their forward career path. Subsequently, the general reason for this framework is to empower students to make choices that are well tailored and suitable according to their previous and forward education and career goal. Which is custom fitted to their general profession and instructive objectives. Students will utilize our recommender system to get instant access whenever to a cutting edge tailored made solution, easy to understand (user friendly), intelligent and interactive framework that gives them clear foresight, extensive data they should explore their way toward their next career path.

Ontologies are properties of the concept, claims of disjoint-ness, limitations of value, and descriptions of logical relationships between objects. Ontologies offered a method to systematically model a system's structure based on the relationships arising from its observation. The term taxonomy is used when ontology includes only "IS-A" relationships. Generally, the use of the word ontology is confined to systems that support various relationships between concepts, including logical propositions that formally define the relationship. Domain Ontology, Upper Ontology, and Application Ontology can be divided into three groups. Domain Ontology describes the terminology related to a common domain such as learning, medicine, automobiles, or any specific function or operation such as selecting or diagnosing by varying the words introduced in top-level Ontology. Upper Ontology, also known as Foundation Ontology, is a concept of everyday objects that can be extended to various ontology applications. An Application Ontology defines a specific domain and task concepts. Upper Ontology and Domain Ontology can be combined with Application Ontology in the application context.

Concepts, relationships, axioms, and instances are the main elements of ontologies. The fundamental abstract part of a domain is a concept (also known as a term or class). A category is usually a group of common properties owned by many members. In addition, groups are organized on two levels in hierarchical graphs. The concepts may have many distinctive properties. In the ontology structure, a Relation describes the relationships between concepts in a domain. To define the two classes involved in a specific relationship, one of them is designated as a "domain" and the other as a range. An Axiom (sometimes referred to as a facet or function restriction) is used in
ontology to impose constraints on both class and instance values. An Instance (also known as an individual) is a relation between ontology concepts regarding their actual value.

Fuzzy logic techniques are the presently utilized strategies to model complex, non-linear, uncertain systems. A fundamental quality of Fuzzy logic techniques is the division of the space of framework variables into fuzzy layers utilizing fuzzy sets [3]. In every layer, the quality of the framework can be defined by only using a standard rule. Generally, a Fuzzy logic technique comprises a common rule correlation with regulations based on layers, where the data is accessible, easily readable, and transparent. This attribute of the fuzzy frameworks has been utilized in numerous fields, including clinical field, recommendation systems, pattern recognition, decision support, agriculture, and engineering. One of the qualities of Fuzzy logic techniques is their comprehensibility. Comprehensibility shows how effectively individuals can perceive a Fuzzy logic technique. In the past few years, the enthusiasm of scientists to get more interpretable fuzzy models has grown. Nonetheless, the decision of a proper interpretability measure is an open conversation because of its emotional nature and numerous variables included.

To help with personalization, in this paper, we proposed a recommender system that can help students find a significant career path by utilizing a fuzzy logic system and ontology strategies. The precise and formal detail of a common conceptualization is known as ontology. It has concepts related to a specific domain like properties, entities, and attributes. Utilizing ontologies encourages the reusing, reasoning, parsing, and sharing of information, and using ontology improved personalization results in many recommender systems [4] [5], proving its utility regardless of the application domain. Our framework will use ontologies to demonstrate students' attributes, education attributes, and career jobs in the IT sector.

The rest of the paper is divided into different sections. Section 2 describes the related work about the proposed methodology, and section 3 walks through the problem statement, proposed framework, and results. And the last section is the conclusion and direction toward the future.

2. Related Work

2.1. Recommender Systems

A recommender framework is a mining application and approach which assist clients with discovering better decisions from among an enormous inventory of options. These frameworks can help the client by prescribing related to various decision-making techniques. The instances of dynamic strategies that are broadly utilized are what item to buy, what book to study, etc. Clients can decide without much of a stretch by using a recommender framework. The recommender framework typically gives recommendations and suggestions to the clients when they are confronting various decisions in choosing. A framework that can provide individualized suggestions or could help clients in a customized way to distinguish compelling data on things with regards to a significant space of potential other options. These frameworks are valuable to
help clients coordinate things to provide simplicity in the details and act like instructors, advisors, and guidance mentors [6]. Different investigations show that few recommendation applications offer suggestions to clients. For instance, a book suggestion for web-based shopping recommends alluring sites or helps the client look for music and motion pictures. This recommender framework shows the client the best approach to construct a framework to discover clients that share a nonappearance to make a suggestive framework dependent on the things that will likewise be popular with different clients [7].

Approaches toward recommender systems are different from one another. In article [8], there are different types of classes of a recommender system that are defined:

1. **Technique Based on the Content**: The framework figures out how to suggest things that have similarities that the user liked in the past. Calculating similarities of an item compared to another object is based on the feature association method.
2. **Filtering Based on Collaboration**: In the past, what similar items were liked by the users? The likeness in the taste of two clients is determined dependent on the similitude in the rating history of the clients. Those items are then recommended to the active users.
3. **Location Based**: This framework suggests to users the content based on their location profile. The generation of the recommendation is supposedly dependent on the location. For different users, location recommendations will be different.
4. **Recommendations based on Knowledge**: These types of recommender systems suggest domain-specific content. A different domain can have various features related to them and can have different preferences and needs. How are that needs or preferences valuable for the users? These types of systems are known as the case-based framework.
5. **Recommendations based on Community**: This framework suggests things dependent on the inclinations of the client’s companions.
6. **Hybrid System**: This type of recommender system is a combination of the above-described techniques.

As our goal is to arrange students in advanced their career path based on the information of graduated class students, the class graduate can have different parameters skill set retained, relating subjects to technology and exact field subjects related to students skills. In our approach, we decided to join the two techniques ontologies derived from knowledge for facilitating the client-system interaction, and users’ preferences are improved by using fuzzy logic models.

### 2.2. Recommender Systems Using Ontology Semantics

Different types of methods based on ontology have been developed for users to make better decisions. For example, in press review and its distribution, actual SARL has specialized in this field. The company has developed a recommender framework that suggests custom-built reviews for every client to diminish the diversity among data. Werner et al. [9] introduced a recommender framework dependent on the semantic specifications of articles and client profiles. The design of
the framework depends on two layers. Layer one is the intelligence layer, used to oversee data extraction. The second layer is known as the ontology layer, used to manage certain information to present the profiles and articles in a manner acceptable to the system.

Moreover, the idea of the semantic web is utilized to improve e-learning platforms. This Paper [10] developed an eLearning framework to help students find objectives according to their learning needs, and the approach is built on the semantic level. Also, the framework can help the educators to propose new assets to improve the course schedule. In this framework, they have used an extension known as query keyword, ontology-based reasoning, and semantic relations. Like the above two methodologies, valencia et al. [11] introduced personalized services to users, a customized recommender system based on ontology to present the clients profiles and elements according to system understanding under the influence of semantic web application. When semantic enhanced methods are applied in the proposed framework it presented high accuracy. The data set used for the experiment was taken from the Netflix movies dataset. The assessment shows that the semantics-based techniques for the recommender framework improved the suggestions' precision. A recommender framework dependent on an ontology can likewise tackle the cold start issue that relies upon lacking client data from a past time [12]. This issue happens because of an underlying absence of ratings for new clients. Subsequently, it gets difficult to make dependable suggestions. To avoid this issue in our framework, we will welcome new clients to finish their profiles with their data and preferences by reacting to specific inquiries. At that point, we make the client profile model dependent on the client ontology model.

### 2.3. Recommender Systems Using Fuzzy Logics

These days, recommender frameworks are utilized for a bigger scope in e-learning to suggest the things necessary to online clients. In the e-learning field, recommender systems have been developed using fuzzy logic. A few parts of career selection experience can be improved regularly by using fuzzy logic's essential mechanisms. For example, the collaborative orientation approach considers the uncontrolled development of data delivered. Every member, such as teachers and students, can embed and share new information, for instance, remarks and messages concerning a college/university course. All the information added incrementally can be assessed in a few different ways: by the mediation of an "instructor" (for example, the educator), utilizing a rating structure, or in any event, as indicated by how many times the information is accessed. A concept data overload is defined, which means that the data quickly becomes unusable due to the rapid increase of knowledge in the system. A different measure of fuzzy logic can be used to control the overload of information in the system. Those measures can be a measure of fuzziness across and membership function of the system. This measure can recognize what relevant and trivial information is. In this paper [13], the methodology leads to an understandable treatment of the data overload because numerous clients' movements steadily generate a solitary wellspring of information without the need to take out any substance. Shockingly better, this sort of portrayal permits us to feature and point out which highlights of the considered topics are strong and very
much assessed and gives us a method to figure out which subjects are the most dubious and complex, using measures of fuzzy logic. It is astounding that this procedure can be applied not exclusively to college classes and courses but additionally to circumstances where one needs to assess a lot of information from various sources – for example, arts, films and surveys, and analysis of books.

The proposed methodology in this paper utilizes a fuzzy similarity family tree to choose the key concepts that are more related to the information to online end users. The authors target producing a rundown of suggestion options with the highest rankings among the expectant appraisals of different vital concepts that are fit to be perused by online clients. Empirical assessments demonstrate that the proposed procedure is proficient and achievable in remembering the key concepts for the recommendation list, which would somehow be forgotten about in the traditional similarity tree method. This work manages similarity family tree fuzzy logic-based e-learning recommender framework dependent on the learning profile of the targeted online client and learning exercises ranked by the other online clients. Top comparable online clients are picked depending on semantic and parameters similarity and afterward weighed evaluation dependent on the UKCR network matrix. In this proposed technique, the modified algorithm tentatively demonstrated a more precise accuracy rating to the fundamental concepts essential to the online client, consequently bringing about more solid and requested proposal choices, suggestions, and upgrades; generally speaking learning experience of the selected clients. Compared to traditional algorithms based on a similarity tree, this framework's performance is more efficient and takes less time to execute.

### 3. Proposed Framework

You can find vast amounts of information on the Internet. It may, therefore, not be an easy task to find relevant information. Consequently, it is essential to have an efficient and effective strategy to coordinate and collect relevant data. With the rapid increase in available documentation from the WWW network, more information about users’ needs is embraced. However, a large amount of information makes it challenging to classify relevant details. For instance, web crawler standard tools have low accuracy as some significant website pages are typically returned but are joined with numerous irrelevant pages mainly because of topic-specific elements that might happen in various contexts. A suitable structure is therefore required to coordinate the vast number of documents on the Internet. When web content expands, it will become progressively troublesome, especially for students looking to find and arrange the relevant, pertinent, helpful, instructive, and educational information like data about course providers, course contents, and vocation information. Until now, there has been no systematic way for finding, aggregating, and using educational content using a search engine crawler to retrieve data from many web pages. In addition, this can also be useful as a way of finding a variety of internet information. Since the aim is to find accurate data on the Internet, given the current size of the web, this detailed approach may not automatically retrieve the required information.
A large portion of the ongoing retrieval procedures depends on keywords. There is no doubt that keywords or index terms neglect to catch data adequately, bringing about numerous immaterial outcomes and causing low quality in the retrieval process. Another web crawler strategy in view of ontology is presented in this paper that is utilized to gather detailed data within online IT courses. For instance, if a client tries to gather all the data about AI courses, the crawler will attempt to collect all the data about the course connected with the particular ontology intended for the domain of computer science. The proposed system in this paper centers around a crawler that can separate data by estimating the similarity for a space between the query terms of the client and the conceptualization in the reference ontology.

3.1. Problem Statement

Cause of the quick advancement of society, graduates need guidance to empower them to pick a promising career. Choosing a profession has gotten progressively more intricate because of various human capabilities, which means every individual or human has their capabilities or ability in a specific area and can be applied to different occupation careers. Moreover, the progressive advancement experienced by the nation has brought about particular fields of work development. Development region of specialization is generally because the nation's economic framework is more unpredictable. The fundamental issue of trouble in selecting career decisions among understudies is that they don't know how to settle on choices, and information and data are absent about positions and professions. To suggest students in vocation particular, it is essential to assemble a Career Path Selection Recommendation System with an ability to address all the issues where it gives guidance and direction to students in picking a lifelong career that suits their aptitudes and capacities.

3.2. Overview of The Proposed Career Development System

Career selection among students aspiring to enter the field of information technology has become very tedious. It creates a dilemma since the IT sector comprises a wide range of job profiles, each requiring a different skill set altogether. A recommendation system, also known as a recommender system, is a tool to help the user provide a suggestion for a specific dilemma. The IT sector is utterly diverse, catering to hundreds of roles working on various technological platforms. In contrast, the studies, both at undergraduate and postgraduate levels, have been designed such that the output product is not a single specialized expert, rather a student who has almost negligible knowledge of intrinsic technological and attitudinal needs of various job jobs profiles in the industry. The students are almost in a fix on how to train themselves and, more importantly, how to pick a job role for themselves. The enthusiasm for building a suggestion framework in numerous fields has recently grown. Fuzzy Logic systems (FLSs) and ontologies are those approaches that can be utilized to model the suggestion frameworks as they can manage vulnerable and uncertain data. One of the primary purposes of using ontologies is to model knowledge at the semantic level. An essential resource in the emerging IT sector is the intrinsic knowledge potential
of the students studying computer science at various levels. For the potential of this resource to be fully realized, we believe there should be a systemically designed career development system (CDS). A CDS can help channel each IT student into a career best suited to that individual while considering career choice and IT sector needs. This paper discusses a CDS founded on evolving fuzzy-based Ontological recommender system framework, which can handle extensive data, stipulate analytics, and logistically suggest a perfect career pathway for an individual.

3.3. Proposed Career Development System Methodology

The proposed crawling system matches the ontology concept, giving the desired result. Upon crawling definition words, the crawled information is ranked by a similarity ranking system. It exposes highly relevant pages that focused standard web crawlers crawling for educational content might have been missed while simultaneously filtering redundant pages and preventing other paths. Ontology mapping is also referred to as the alignment of ontology and ontology. Mapping or matching ontology is distinct from combining ontology. Ontology mapping attempts to make the ontologies of the origin coherent and consistent with each other while keeping them separate.

On the other hand, an ontology merger aims to create a single coherent ontology that incorporates all sources of information. Ontology mapping is used to "establish correspondences among the source ontologies, and to determine the set of overlapping concepts, concepts that are similar in meaning but have different names or structure, and concepts that are unique to each source [14]. It is also defined by Kalfoglou and Schorlemmer [15] as follows, "Given two ontologies O1 and O2, mapping one ontology onto another means that for each entity (concept C, relation R, or instance I) in ontology O1, it tries to find a corresponding entity, which has the same intended meaning, in ontology O2". This research defines ontology mapping as finding a set of semantic correspondences between similar elements in different ontologies. An ontology is considered a taxonomy of concepts. The matching problem is reduced to "for each concept node in one taxonomy, find the most similar node in the other taxonomy."

Profession trait is a fundamental component that impacts the determination and selection of careers among graduate students. Profession choice is one of the indispensable choices students must make in choosing their future. One territory concerning students’ future professional improvement is professional decisions related to one's professional decision-making. The profession trademark, which may apply to students in picking their career, is typically ordered into three gatherings: compensation or security, the organization or work environment, and the career itself. Besides, the professional direction offered in universities ought to satisfy the specialized data needs of university students at various degrees of their career advancement. It is essential to convey career direction in a few different ways, for example, through courses, preparing workshops that offer gathering encounters for future professional planning, and gathering specialized group studies or individual advising exercises.
3.3.1. System Design

The proposed technique advances a thought of career selection framework by M-tree ranking, controlled with ontology for semantic correlation by utilizing fuzzy logics. The proposed framework provides suggested recommendations on the following steps:

1. In the proposed framework, preprocessing is the initial phase wherein preprocessing the client inquiry. Preprocessing implies when a client gives some input as an inquiry, then eliminates words like an, was, so, is, and on, called stop-words expulsion. After stop-words, expulsion stemming is accomplished. For instance, the going word is changed over into go. Then apply the term frequency-inverse document frequency to compute the frequency of weight.

2. In the next step, semantics are provided to the words based on ontology reasoning.

3. When ontology reasoning is done, the next step is creating a hierarchy based on semantic correlations provided by the M-tree method.

4. In the next step, based on fuzzy classification, fuzzy logic takes a decision.

5. The last step is finding the client's best-case scenario decision, which can be done using the Pearson correlation method.

Figure 1. explains how the proposed framework effectively examines the user's necessity more semantically by utilizing ontology. As per this, the final decision is made using the Fuzzy tree blend. At that point, a hybrid suggestion will be given utilizing filtering based on collaboration controlled with the Pearson relationship method. The content-based recommendation is effectively provided to the client with high accuracy results.

3.3.2. Algorithms

This section includes algorithms that define how the system is designed and how its proportions work together to recommend valuable suggestions to help the users.

| 1. Career Development System Preprocessing Algorithm |
|---|---|
| Step 1: | Begin |
| Step 2: | Interpret the string |
Step 3: $V \rightarrow$ is a vector to store words after spaces separate them from a string
Step 4: Eliminate Special Symbols
Step 5: Recognize Stop-words
Step 6: Eliminate Stop-words
Step 7: Recognize Stemming Substring
Step 8: Distinguish Stemming Substring
Step 9: Link the words together to form a string
Step 10: End

A career development system algorithm is used to calculate and form a tree and the time complexity of the proposed career development system M-Tree algorithm is $O(\log(n))$, where $n$ is various careers in the IT sector.

### 3.3.3. Results and Discussion

In this investigation, we utilize a contextual analysis of the career development system framework to exhibit the utilization of various hierarchical fuzzy logic and ontologies frameworks for suggesting careers. The CDS is a recommendation framework that gives guidance and direction to students in picking their career path through ability evaluation that depends on different decision question strategies or rank ratings. The CDS utilized fuzzy logic procedures for mapping the aptitudes appraisal input to the career recommendations for graduate students. To show the efficacy of the proposed framework, a few analyses are conducted on java-based windows machine
utilizing NetBeans as an IDE. A metrics Mean Absolute Error is generally used to quantify the quality of different recommender system suggestions.

\[ \text{MAE} = \frac{\sum_{n,m} |C_{n,m} - C'_{n,m}|}{N} \]

(1)

In equation (1) Where \( C_{n,m} \) symbolizes the predicted career m in clients' recommendation choice n, whereas \( C'_{n,m} \) describes the recommended values, and N is the total no of recommendations by the system. So different career paths can have various ranges in value cause of the additional recommendation's properties. So, we have to normalize the mean absolute error. The normalized mean absolute error is defined in equation (2).

\[ \text{NMAE} = \text{ME}\left(\frac{\sum_{n,m} |C_{n,m} - C'_{n,m}|}{N}\right) \]

(2)

The quality of the recommendation will be high when NMAE value is smaller. In figure 2, we can see that our career development system has an inclined line which means that the proposed method has a smaller value of NMAE, so according to NMAE metrics, CDS has better accuracy. Whereas the career path NMAE values are more significant and increasing, that means recommendations are not too accurate. They are false from the user's point of view, but sometimes the suggestion by CPR makes sense in the user's opinion but is not helpful in reality.

![Graph of System Performance](http://www.webology.org)
4. Conclusion and Future Work

Conclusion of proposed work, the results of this investigation will profit university students in deciding which profession to choose, and the career selection process will become faster, easier, and more flexible. This is because how self-testing should be possible without the need for overall coaching by the instructor. The proposed method also measures students' personalities, capabilities, skill sets, and strengths and suggests viable profession decisions using ontologies and fuzzy logic. This article is an attempt to improve the idea of authorizing a career development system to utilize fuzzy tree hierarchy and ontology by presenting Pearson correlation, Preprocessing, M-Tree creation, Ontology Reasoning

For future work, the current approach can be embedded with AI techniques or machine learning algorithms to check the capabilities and to study more about career development systems under different scenarios to see their behavioral changes due to various circumstances.

5. References


[2] G. E. Steele et al., "Advising Undecided College Students The Global Community for Academic Advising Learning outcomes • What are the characteristics of undecided • What are effective advising strategies for working with them ? • How do we help them develop their academic and," 2016.


