Knowledge Discovery in E-Learning with Social Media

Doctor of Philosophy Dissertation

by

Nauman Sharif

Univ.Prof. Dr. Denis Helic

Institute of Interactive Systems and Data Science

Graz University of Technology

Graz 2019

Acknowledgments

In the name of Allah, the Most Beneficent, the Most Merciful. All praises and thanks are to Allah, the Lord of the Alamin (humankind, jinn and all that exists). I thank Allah, the almighty who allowed me to carry out this research. Surely, without Allah's blessings, it would not have been possible.
I would like to express my profound gratitude to my supervisor Prof. Dr. Denis Helic, who has been truly instrumental in my research. I could not have imagined having a better supervisor and mentor for my Ph.D. His valuable guidance and support helped me to transform vague thoughts into useful concepts. His patience, motivation, and continuous support helped me throughout my research and the writing of this thesis.

I owe a massive debt of gratitude to Prof. Dr. Tanvir Afzal. He has been a genuinely committed mentor. I am especially grateful for his constant faith in me. His patience, support, and immense knowledge were influential factors during my Ph.D.

I would also like to thank my parents, wife, children and my brothers and sisters for their help and support during hard times. I ask Allah, the Almighty, to grant them all success.

Abstract

E-Learning is one of the emerging research areas in the context of Knowledge Management. The recommendation methods in E-Learning domain are comparatively new and mostly derived from commercial recommender systems. One of the essential features that have been researched by the scientific community is how to provide relevant resources at the right time. The Technology Enhanced Learning community has been continuously developing innovative methods for learners to access the most relevant learning resources.

This dissertation starts with background and related research efforts. Over 200 recent and classical papers have been critically reviewed about E-Learning recommender systems and Twitter-based recommender systems in various domains. It establishes the need for specialized E-Learning system centered towards personalized learning. The next section explores a prototype of E-Learning system for domain specific recommendation using social bookmarking. A comprehensive study of the literature reveals a gap between the E-Learning recommender systems and Twitter-based recommendations. There are no existing E-Learning recommender systems which are utilizing the influence of social websites such as twitter.

Learner and learning resources are two pivotal entities to consider for E-Learning recommender systems. There are certain objects associated with learners for recommendation task, for example, the learner's profile as well as context and history are commonly used in the literature. Similarly, for learning resource recommendations, the essential features used in the literature are metadata (Title, Author, Category, Keyword), content and extensive vocabulary (Synonyms, Growbag which is a DBLP dataset of co-occurrences terms). However, such learning resource features have not been employed and evaluated for Twitter-based recommender systems.

The aim is to evaluate how the metadata of resources should be combined. Similarly, another important question is which combination of metadata of resources provides the best results for a Twitter-based recommendation? The metadata of resources are used as input for different recommendation techniques, namely lexical matching, semantic similarity, and extended vocabulary. Comprehensive experimentation and evaluation indicate the usefulness of twitter for providing technology-enhanced learning.

Zusammenfassung

E-Learning ist einer der aufstrebenden Forschungsbereiche im Kontext des Wissensmanagements. Die Empfehlungsmethoden in der E-Learning-Domäne sind vergleichsweise neu und meist von kommerziellen Empfehlungssystemen abgeleitet. Eines der wichtigen Merkmale, das von der wissenschaftlichen Gemeinschaft erforscht wurde, ist die Bereitstellung relevanter Ressourcen zur richtigen Zeit. Die Community des Technology Enhanced Learning hat kontinuierlich neue und notwendige Methoden entwickelt, mit denen Lernende auf die relevantesten Lernressourcen zugreifen können.

Die Dissertation beginnt mit Hintergrund- und verwandten Forschungsarbeiten. Mehr als 200 neue und klassische Artikel wurden kritisch über E-Learning-Empfehlungssysteme und Twitter-basierte Empfehlungssysteme in verschiedenen Bereichen geprüft. Das begründet die Notwendigkeit eines spezialisierten E-Learning-Systems, das auf aktives und authentisches Lernen ausgerichtet ist. Der nächste Abschnitt untersucht den Prototyp des E-Learning-Systems für die domänenspezifische Ressourcenempfehlung unter Verwendung von Social Bookmarking. Eine umfassende Studie der Literatur zeigt eine Kluft zwischen dem E-Learning-Empfehlungssystem und der Twitter-basierten Empfehlung, nämlich dass keine existierenden E-Learning-Empfehlungssysteme die Macht von sozialen Webseiten wie Twitter nutzen.

Lern- und Lernressourcen sind zwei zentrale Punkte, die für E-Learning-Empfehlungssysteme zu berücksichtigen sind. Es gibt bestimmte Dinge, die mit Lernenden für eine Empfehlungsaufgabe verbunden sind, zum Beispiel das Profil des Lerners, der Kontext und die Geschichte sind die allgemeinen Optionen, die in der Literatur verwendet werden. In ähnlicher Weise sind für Lernressourcenempfehlungen die wesentlichen in der Literatur verwendeten Merkmale Metadaten (Titel, Autor, Kategorie, Schlüsselwort), Inhalt und erweiterter Wortschatz (Synonyme, Grow bag). Solche Lernressourcenmerkmale wurden jedoch nicht für Twitter-basierte Empfehlungssysteme verwendet und ausgewertet. Ziel ist es, zu bewerten, wie die Metadaten von

Ressourcenmerkmalen kombiniert werden sollen. In ähnlicher Weise liefert die Kombination von Metadaten von Ressourcenmerkmalen das beste Ergebnis für Twitter-basierte Empfehlungen. Metadaten von Ressourcenmerkmalen werden als Eingabe für verschiedene Empfehlungsverfahren verwendet, nämlich lexikalisches Matching, semantische Ähnlichkeit und erweitertes Vokabular. Umfassende Ergebnisse zu Experimenten und Evaluierungen zeigen die Nützlichkeit von Twitter für technologisch verbessertes Lernen.

Contents

CHAPTER 1: Introduction

1.1 Research Challenges/Trends

New and innovative approaches have evolved continuously over the past decades to provide better recommendation strategies. However, most of this great success has been in the field of commercial recommender systems. Such examples include Amazon.com, for recommending books, CDs, and various products and MovieLens, for recommending movies. There is a growing need to transfer this success into technology-enhanced learning in order to provide personalized learning environments for learners [Hoic et al. 2016].

The recommendation methods in E-Learning domain are comparatively new and mostly derived from commercial recommender systems. However, E-Learning recommendation goals are different from commercial based recommendation and these goals are surely not measured in similar manner. The commercial recommender's main goal is profitability whereas E-Learning recommendation focuses on improving the learning. In E-Learning context, targeted learner and their supported tasks should be thoroughly analyzed. This process should take place before the recommender system is deployed.

New knowledge or further expansion on a concept would require diversified learning resources. Based on the learner interest and current learning task the system should be able to discover the most relevant information from internal and external sources [Henze et al. 2005]. The discovered information should be made available to learner's current context. The discovery of most relevant information from the Web is not a trivial issue because of the unstructured information on the Web [Sharif et al. 2014].

However, with the evolution of Web 2.0 (social web), the unstructured data on the Web is being structured by social community. There is a technology shift in virtual and social societies for example, a large number of social communities are now connected through the internet and learning using others' vocabulary and shared resources. In this context, E-Learning can be viewed as the transfer of Knowledge in the context of Knowledge Management, where the social web can facilitate knowledge creation, organization and archiving.

1.2 Motivations, Thesis Objectives, and Contributions

1.2.1 Motivation

Traditional E-Learning Management Systems (for example, Moodle and Blackboard) are limited in the amount of personalization which they can offer to the learner.

Educational material and courses provided in the contemporary E-Learning systems have complicated structure. In general, online course material provides freedom to the learner to choose any navigational path in their learning context without depending on the structure set by the course designer.

This unsolicited freedom may prove to be ineffective for the learner. Learners may not have enough knowledge and experience to follow an effective navigational path. Hence the learner is left to wander around a topic which may be irrelevant to the learner's learning context.

1.2.2 The scope of the Thesis

There are two entities or factors in the recommender systems:

- Learner
- Learning resources

There are certain objects associated with learners for the recommendation task, for example, the learner’s profile as well as the context and history are commonly used in many domains.

Similarly, for learning resource recommendations, the central features used in the literature are metadata (Title, Author, Category) content and extensive vocabulary (Synonyms, Grow bag).

This research will endeavor to find the possible influence of social communities (available on the social web) on the E-Learning system. Focusing mainly on learner and learning resource metadata for Twitter-based learning recommendations. Learner parameters are evaluated using a user study whereas learning resources metadata are evaluated with different techniques using n-grams.

1.2.3 Research Questions

Following research questions are investigated and evaluated in this thesis:

RQ1: How the recommender system can help the E-learner by finding the relevant knowledge from social bookmarking? *(RQ1 is addressed in Chapter 4)*

RQ2a: What E-learner's parameters can be necessary for a recommendation? *(RQ2a is addressed in Chapter 5)*

RQ2b: What is the importance of E-learner’s parameters? *(RQ2b is addressed in Chapter 5)*

RQ3a: How resource metadata and contemporary recommendation approaches lead to an effective Twitter-based recommendation for E-learners? *(RQ3a is addressed in Chapter 2 and 3)*

RQ3b: How the resource metadata should be combined? Which resource metadata combination provides the best result? *(RQ3b is addressed in Chapter 6, 7 and 8)*

1.2.4 Foundation of the thesis and contributions

The foundation of this thesis is based on numerous published works authored over a period. *Figure 1.1* illustrates the relationship between the research questions and the published work. A comprehensive study of the literature reveals a gap between E-Learning recommender systems and Twitter-based recommendation such that no existing E-Learning recommender systems are utilizing Twitter.

Similarly, literature review illustrates recommender systems can help the E-learner to find the relevant knowledge from a social network. Twitter has great potential to take the learning beyond the realm of classroom and reduce the differences between classroom-based instruction and distance learning [Gao et al. 2012].

Firstly, this thesis highlights the strengths and limitations of prominent approaches and presents challenging tasks which will be useful for the E-Learning research community to focus on future research.

Secondly, it demonstrates the need for constructing specialized (domain specific) E-Learning systems which can help learners of some common domain and can also assist them according to their particular needs, context, profiles, histories, collaborations, etc.

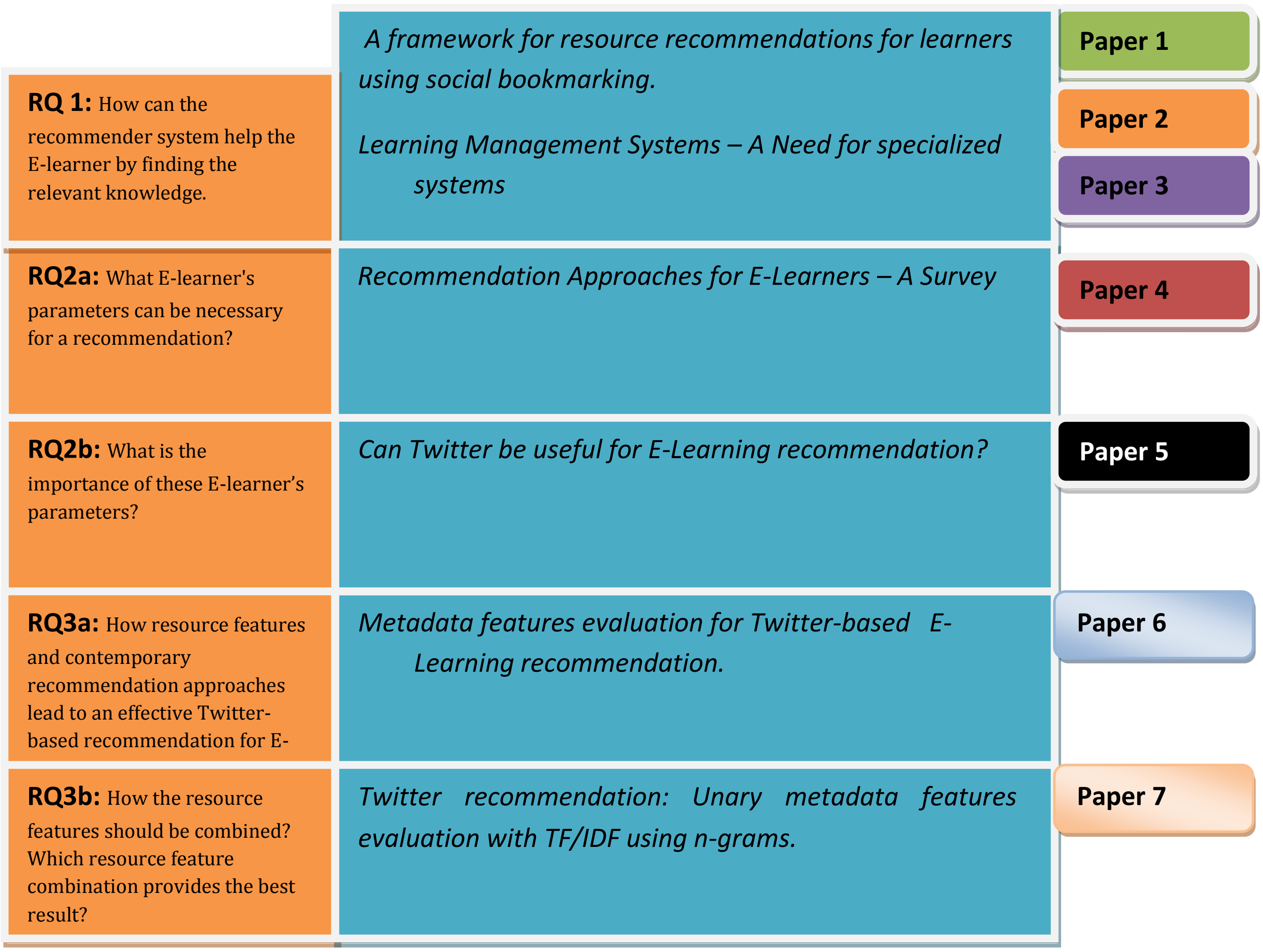

Figure 1.1 Research question & published work

Thirdly, thesis explores the potential of utilizing the Twitter for E-Learning recommendation and the parameters that should be considered when recommending E-Learning resources.

Fourthly, this thesis proposes and implements Twitter-based semantic recommendation techniques to facilitate effective learning for e-learners.

Fifthly, this thesis discusses the implementation of three distinct techniques to determine the effectiveness of resource metadata for Twitter-based E-Learning recommendation.

Sixthly, this thesis presents extensive experimentation and evaluation of metadata for Twitter-based E-Learning recommendation. In total, 102 different evaluations were conducted. The detailed evaluations reveal the effectiveness of the resource metadata for Twitter-based E-Learning recommendation.

CHAPTER 2: Background and related research efforts

The rapid introduction of new technologies within the past few decades has brought about some innovative methods for sharing and distributing knowledge and information. These innovative methods had a significant effect on entire industries worldwide, and the education sector is not an exception. In recent times over 700 universities provide distance learning programs. E-Learning systems have been used by over 200 universities [Hameed et al. 2015]. These new advancements enable us to share and manipulate information instantaneously. E-Learning technology supports the learner to access the variety of learning content any time from any place [Dwivedi et al. 2015]. This reality is shaping the new era of E-Learning.

It is, therefore, necessary to carry out a critical review of the existing E-Learning systems in the context of this new era of Knowledge Management. E-Learning is viewed as a small part of Knowledge Management [Maurer et al. 2001] where teaching and learning refer to Knowledge Transfer. A discussion in this section will provide a stepping stone for our research.

2.1 Overview of Contemporary E-Learning Systems

E-Learning aims at providing an alternative to classroom learning. E-Learning is often referred to as learning which is not restricted to a physical presence in the classroom, but rather a means of accessing the educational material with the help of a computer and internet or CD/DVD anytime anywhere in the world [Wrubel et al. 2009].

The learning management system has become an inevitable part of the E-Learning environment, and thousands of universities and colleges across the globe are using such software including WebCT, Blackboard and Moodle[1]. WebCT and Blackboard merged in early 2006 [Cheung et al. 2006]. WebCT and Moodle are the most widely used web-based learning management systems. WebCT is the first online course management system for higher education.

Moodle, on the other hand, is a free open source course management system which is widely used among web-based learning communities. There are over 1 million registered users in 215 different countries[1]. The core features which are shared among these web-based learning systems are briefly described in the following section.

2.1.1 Course design and management:

Provides the facility for the instructor to create and manage the course materials and supervise student's activities on the course. Upcoming events and announcements can be placed on the schedule.

2.1.2 Performance evaluation and feedback

It provides student assessments through online tests and quizzes. Teacher's remarks and explanations can be incorporated via a feedback mechanism.

2.1.3 Interactive communication

It provides communication and discussion tools for interactive participation among students and instructors.

2.1.4 Course evaluation

Course evaluation obtained from the students can be used to improve the delivery of the course. The features mentioned above conclude E-Learning systems as course management tools. However, learners also need more advanced functionalities. One such functionality is to obtain the most relevant resources for the resources being read in the environment of E-Learning tool.

[1]http://E-Learningindustry.com/top-10-E-Learning-statistics-for-2014-you-need-to-know

2.2 Limitation of traditional E-Learning

Traditional E-Learning systems are seen more as course management tools to facilitate the instructor in the delivery of course content rather than catering for the individual learners need [Stojanovic et al. 2001]. This has led to critics questioning the effectiveness of E-Learning systems and whether they provide an alternative to the classroom learning experience [Lennon et al. 2003]. In order to demonstrate whether the current E-Learning systems facilitate an effective learning environment, let us explore the scenario from the learner’s perspective.
Joseph is using a Learning Management System (LMS) for his "Computer Networking" course. He is currently learning about "Topologies" and feels that current online tutorial does not provide him a satisfactory understanding. In order to fully comprehend the topic, Joseph wants further clarifications on the topic. However, he could not find any further useful resources through his LMS tool. This however was relatively easier in classroom environment. He would ask the teacher and might be satisfied through further discussions, or he might receive further directions for his auxiliary explorations.

In order to find answers to his questions, he decided to explore the Web. For this task, he had different options, such as exploring through Search Engines, Citation Indexes, Digital Libraries, and Social Bookmarking sites. He tried to explore one of the best search engines, Google. On his query about "Topology," he received 28 million hits. Similarly, Citation Indexes (Google Scholar, CiteSeer, ISI), Digital Libraries (IEEE, ACM, Springer) also returned many irrelevant results. This experience left Joseph in disarray.

Joseph concluded it would take him countless hours to read through all of this content in order to fully comprehend the topic. He found this process to be daunting and extremely exhausting. Joseph felt helpless and wished his LMS system had recommended a few of the most relevant resources related to the task at hand rather than him having to go through millions of generic hits.

2.3 Emerging trends

In the above sections, contemporary E-Learning tools have been reviewed. These traditional tools have been viewed as course management systems rather than alternatives to classroom learning. Whenever learners require further clarifications on the topic, they need to visit external sources such as Search Engines, Citation Indexes, and Digital Libraries, etc. However, the introduction of new technologies in the past decades has brought about some innovative methods to share and distribute knowledge and information.

In recent times, over 700 universities provide distance learning programs. E-Learning systems have been used by over 200 universities [Hameed et al. 2015]. The advent of Web 2.0 has enabled the social community to share many essential resources with the scientific community. This reality is shaping the new era of E-Learning. [Dwivedi et al.2015]. The following sections describe some growing trends in the digital community.

2.3.1 Read Web to Read Write Web

Before the advent of Web 2.0, a Web user was primarily seen as a passive consumer of information by simply searching and reading through websites. However, the Web has changed from "the Read Web" to the "Read-Write Web," according to Tim Berners-Lee's original vision [Downes 2012]. Users are not merely the passive consumers of information but active participants which can edit, co-create and collaborate in a more dynamic environment [Cooze et al. 2007]. Social networking software (blogs, wikis etc) has played an influential role in this paradigm shift to facilitate the collaborative environment for knowledge sharing.

This has had a revolutionary effect on the education sector. There is a growing trend towards online systems which support contextual learning and support participation and interaction. This is different from simple instructor-led methods of teaching. Learners require the learning experience to be more collaborative and interactive with their peers rather than being isolated in a boxed room. Learners expect course content to be relevant to the real world [Beldarrain et al. 2006].

It can support educators in creating a vibrant, collaborative learning environment in their learning context [Lightner et al. 2007]. As a result, a more proactive learning environment takes place where resources are presented as and when required.

2.3.2 Pull technology to push technology

In a presentation at Stanford University, Yahoo's VP and Research Fellow Andrei Broder highlighted a shift from pull technology to push technology. In a visionary statement, Andrei Broder stated:

"The goal of Web Information retrieval will widen to include the supply of relevant information from multiple sources without requiring the user to make an explicit query. A prime example is the matching of ads to the content being read" [Broder et al. 2006].

Andrei Broder's idea provided the motivation to develop a system which could provide explicit and implicit information for the learner. He referred to this as push and pull syndrome. Using the concept of push technology, the system would recommend the appropriate resources based on the current context of learning without requiring the learner to request the information. It is accomplished by taking a complete account of the background knowledge and context of learners at all times. The learner profile is considered in order to personalize the presentation of information.

In-depth profiling of individual learner is applied to determine the goals and objectives of the learner. This is done by dually considering the learner's profile and learner's current activity. The system should be able to discover the most relevant information from internal and external sources based on the learner profile and current context [Henze et al. 2005]. This approach should help to reduce the effort required to produce the useful information.

More recently, personalized and intelligent E-Learning systems offer personalized learning experience by constructing the learner model based on learner aims, likes and existing knowledge. The next section will discuss such recommendation approaches for e-learners.

2.4 Recommendation approaches for e-learner

In the past few decades, the introduction of new technologies has brought about some innovative methods in web-based education. However, many of these online courses provide universal static solutions which do not cater to the individual needs of the learner. Recommending relevant content to the learner is a challenging task for any E-Learning management system.

In web-based education, it is possible to store most of the learner’s learning patterns in large-scale data sets. Personalized learning profiles can be created with the help of data mining technique [Romero et al. 2007]. The aim is to provide personalized learning activities and tasks which best suit the individual learner's needs as a result enhance the overall learning experience.

Similarly, tasks and activities are recommended to the learners who are related to previously completed tasks by the learners or their peers. Recommender systems provide an excellent opportunity for learners to have a personalized learning environment. Relevant and interesting resources are suggested to the learner from a large pool of resources. Suggestions could be based on learner’s usage history (learning resource previously visited or selected, and the ratings score provided to these resources) or the preferences and ratings of other learners about a particular resource. Often different techniques are combined to avoid the drawbacks of a single technique. Recommendations can also be in the form of an online task, tutorial or simple webpage.

Contemporary Learning Management Systems (LMS) such as Blackboard or WebCT do not provide intelligent learning environments [Rashid et al. 2013]. These LMSs do not provide personalization and dynamic learning environment. As a result, the research community has started to raise questions on the usefulness of conventional E-Learning systems [Sharif et al. 2014].

The scientific community has been continuously addressing the issue of providing relevant resources at the right time [Saaya et al. 2013]. Various techniques have been discussed in the literature to illustrate the various recommendation methods. The most widely used techniques are content-based filtering, collaborative-based filtering, trust-based, and semantic model. An E-Learning recommender system is classified based on these techniques mentioned above [Sharif et al. 2015].

The following section highlights the strengths and limitations of prominent approaches and presents challenging tasks which will be useful for the E-Learning research community to focus on future research.

2.4.1 Content-based filtering

Content-based filtering techniques are particularly useful for recommending those items which hold textual information, for example, articles, URLs, etc. Content-based recommendation systems recommend similar learning object which the learner liked previously [Pazzani et al. 2007]. New and exciting learning objects are recommended to the learner by matching the learner’s profile features with the learning object features. A user profile presents information about the user's likes and requirements. This information can be acquired through surveys and questionnaires or temporal information.

A learner's preference model is built by extracting feature information of learning object and learner profile. The similarity of each LO is calculated with learner's preference model, and the highest degrees of similar LO are recommended to the learner.

The content similarity is usually calculated using a vector space model, by applying the TF-IDF weighting mechanism. Both learner profile and learning objects are represented as weighted term vectors. Learner interest in a particular learning object is acquired by calculating the cosine similarity. Sugiyama, K et al. technique recommend relevant research papers to the users by defining user interest model through their previous publication and papers they cited. Users' profiles are modeled by creating the term frequency for weighted term vectors. Terms are extracted from the user publication history [Sugiyama et al. 2010]. Likewise, resource documents are also defined as vectors by the Term Frequency-Inverse Document Frequency (TF-IDF) method. Relevance between the user profile and resource document is calculated by the cosine similarity method.

Wu, D et al. proposed a fuzzy tree structure data model in order to represent the learner profile and activities. Learning activities matching to the learner profile were calculated using tree matching technique [Wu et al. 2014].

Kurilovas, E et al. present a novel method to recommend relevant learning path to various learner groups. When new learning objects are added in the learning path, the association of new and old learning objects needs to be updated.

In order to provide dynamic learning path selection, swarm intelligence, and modified ant colony optimization algorithm was implemented. The approach successfully supports learners to reach personalized learning objects [Kurilovas et al. 2014].

Chandrasekaran, K et al. proposed an approach where users and papers are represented as a tree concept by utilizing the ACM Computing Classification System (ACCS). Vector space classifier is trained to associate ACCS concepts to documents and users' interest. A comparison between the user profile and paper representation is calculated by tree edit distance. A similarity between two trees is calculated by the number of operations required to transform one tree into other [Chandrasekaran et al. 2008].

In Content-based recommendation systems, object information is represented in textual form, unlike structural data where feature information has a distinct value. Textual representation creates various problems when calculating the user profile model. Learner profile is extracted from the text as Key terms which do not necessarily depict user interest semantically. For example, a keyword can represent different meanings and similarly, different words can also have the same meaning. So, when the simple key term matching is performed, it is possible that essential terms with multiple meanings are found in the user profile and the learning object. As a result, the wrong learning object will be deemed relevant for this particular learner. Similarly, it is possible the relevant learning object can be overlooked if the profile does not hold a precise key term.

2.4.2 Collaborative filtering

Collaborative filtering (CF) term was invented by David Goldberg et al. The author inspiration was based on the knowledge that human influence can play a useful role in information retrieval processing [Goldberg et al. 1992]. Collaborative filtering techniques generate recommendations based on the user's rating matrix or analyzing items usage history.

The underlying assumption applied in CF is that if user A and B have similar ratings about a set of items, then their ratings should be the same on other items as well [Goldberg et al. 2010]. In order to make sound predictions and recommendations, an extensive collection of users rating dataset is required.

Model-based CF constructs the model from the training data and provides the predicted ratings on new items. The method implies various techniques from machine learning and data mining.

Collaborative filtering technique can be classified into Memory based and Model-based methods [Ren et al. 2011]. In Memory based CF methods K Nearest Neighbors (kNN) is the most commonly used algorithm [Schafer et al. 2007]. The algorithm (kNN) involves neighborhood selection rating matrix and recommendations. Neighborhood selection is computed by the similarity between two users. Similarly, rating matrix, on the other hand, will take the weighted average ratings from the neighborhood and generate predicted ratings of an item for the active user.

Users' similarity is computed through various measures such as cosine similarity, dice coefficient, and Pearson correlation and jacquard [Bobadilla et al. 2012]. Model-based CF constructs the model from the training data and provides the predicted ratings on new items. The method implies various techniques from machine learning and data mining. Model-based algorithms such as clustering model, Bayesian, dependency networks, latent semantic models have been investigated to overcome the Memory-based CF limitations [Basu et al. 1998].

Holenko Dlab, M et al. proposed a recommender system which consists of activity, student and group models, as well as a recommender module. The primary objective is to provide the methods for assessing the student's and the group's activity level based on the data collected from Web 2.0. Recommender module provides a personalized recommendation based on student group and activity models. Collaborating filtering technique was used to find the similar users' group [Dlab et al. 2014].

Wiki-Learnia is an E-Learning 3.0 solution which brings together social learning features with recommendation methods and mLearning technology [Waßmann et al. 2014]. It can be described as MOOC Meta search engine to extract the E-Learning content.

Wiki-Learning is centered towards the learner and provides the opportunity for the learner to set personalized learning targets. Each learner can search and collaborate with other learners who share her learning targets thus build up a learning community. Sevarac et al. used Neuro-fuzzy inference in order to create pedagogical convention in E-Learning [Sevarac et al 2012].

2.4.3 Hybrid-based Model

In a Hybrid based model, two or more techniques are combined in order to overcome the limitations of one technique. For example, collaborative filtering can be combined with content-based filtering [Chen et al. 2014]. Burke introduced seven different classifications of the hybrid recommendation systems [Burke et al. 2002], which are as follows:

(a) Weighted hybrid:

In this technique, scores from different recommendation components are combined using a linear formula.

(b) Switching hybrid

Recommendation components are chosen as per the situation, and recommendation components performances may differ from situation to situation.

(c) Mixed hybrid

Multiple ranked lists from different recommendation components are merged and presented together.

(d) Feature combination hybrid

Recommendation component is provided with features extracted from different sources.

(e) Feature augmentation hybrid

It is like feature combination method; however, the recommendation component produces new features. These features are used as input for next recommendation components.

(f) Cascade hybrid recommendation

An agent in the first stage generates a ranking of candidates while secondary recommender agent is used as a tiebreaker and provides more refined results.

(g) Meta-level hybrid

Two different recommendation components are combined where the first component produces a model which in turn is used as an input for the second component.

Tang et al. proposed a hybrid recommendation approach using two pedagogy attributes: learner interest and background knowledge. The content-based technique was compared with hybrid recommendation approach, and empirical results showed hybrid collaborative filtering could reduce the computational cost [Tang et al. 2005]. In the process of providing a personalized learning experience, learner's skill competency is often ignored. As a result, it leaves the learner disorientated.

A new personalized E-Learning model was proposed which combines the item response hypothesis with collaborative filtering method [Chen et al. 2005]. Individual learning path is provided to the learner in order to support effective learning. Object features are used to build the Knowledge model, and Maximum Likelihood Estimation is used to predict learner skill level. Similarly, another hybrid recommender system for learning resources was proposed [Ghauth et al. 2010] which utilizes the k-nearest algorithm and Preference Matrix. Empirical studies provide promising results to cater to cold-starts and sparsity problems.

One of the significant drawbacks with the hybrid recommendation is the time complexity. As the size of the data set increases, with time the recommender system also performs sluggishly. Similarly, the use of different data sets also affects the system performance, and as a result, the learner interest is diminished [Salehi et al. 2013].

2.4.4 Trust-based

E-Learning recommender systems are different from conventional recommender systems. More experienced learners can provide better recommendations than a beginner level learner [Helic 2007]. A trust-based recommender system assigns the trust level according to the ability of the user and the user's interaction with the system over a period [Yuan et al. 2010]. Another trust-based model was proposed [Pitsilis et al. 2008] which provides an association between current knowledge through similarity measure and common values required to ascertain trust.

Well-known fuzzy logic applications, the fuzzy inference system, and fuzzy MCDM methods are used to supervise the quality and reliability of peer learners [Li et al. 2009]. Victor, P et al. proposed a trust model based on fuzzy logic where trust scores are paired as trust and distrust [Victor et al. 2009].

Dwivedi, P et al. proposed a method where recommendations of learning resources are filtered at two levels in order to provide recommendations from most experienced and trustworthy learners. The experimental results indicate trust and experience play a vital role in the accuracy of recommendations and this collaborative framework proves better than conventional Pearson collaborative filtering [Dwivedi et al. 2011].

2.4.5 Semantic model

A semantic model can provide various advantages in personalized recommender systems. Learner’s interest in a particular domain can be dynamically contextualized [Kumar et al. 2015]. Next generation of recommenders should consider how the personalization process can take benefit from semantics as well as social data in order to improve the recommendations [Victor et al. 2010].

Semantic web provides better prospects to improve the metadata associated with learning content. It also offers an excellent opportunity to expand the existing E-Learning methods [Malik et al. 2009].

Shen, L et al. proposed an ontology-based learning content recommendation model. Different learning objects are connected by exploiting the sequencing rules which are derived from knowledge-base and skill set gap analysis [Shen et al. 2005]. ADL Shareable Content Object Reference Model (SCORM) provides and delivers XML-based interoperable specifications in order to exchange and sequence learning objects [Olivier et al. 2004].

The goal is to keep the learning content independent of a particular content provider technology. This will enable new and innovative learning experiences through existing learning objects. The original composition and skills are bonded together in a single context. Metadata and activities metaphors state this context and bond [Downes 2012].

Neil Rubens et al. suggest using artificial intelligence knowledge such as semantic filtering and recommendation systems to be used in LMSs which are geared towards E-Learning 3.0 [Rubens et al. 2014].

An ontology-based knowledge framework was proposed by Yarandi, M et al. for an adaptive E-Learning system. Learning process incorporates learner's knowledge, skill set, and learning preferences. The ontological based user profile is updated as the user progresses in her learning process. Learning content is annotated metadata from a domain and content ontology [Yarandi et al. 2013].

The proposed model based on Fuzzy Knowledge Management System takes into account the learners' profiles, learning objects and pedagogy. The goal is to provide most suitable leaning content which match learner's skill set and learning preferences.

The knowledge base includes ontologies for course and concepts as well as the knowledge of learners' profiles and learning resources. Fuzzy logic techniques are used to present and disseminate the knowledge [Salahli et al. 2012].

Table 2:1 Analysis of Recommendation Approaches for e-Learners

Analysis of Recommendation Approaches for e-Learners					
Name of Technique	**Short Description**		**Advantages**	**Disadvantages**	**E-Learning Use**
	Memory-Based	**Model-Based**			
Content-based filtering	Cluster analysis TF-IDF [Sugiyama et al. 2010]	Artificial neural network [Sevarac et al. 2012] Decision Trees [Wu et al. 2014] NaïveBayesian Classifier [Basu et al. 1998] Clustering	Domain knowledge is not required. Learn user preference. Adaptive quality of recommendations improves over time.	New user problem Does not cater for changes in users interests. Lack of serendipity in discovering interesting items by chance, e.g. Wikipedia	A learning model for each user Keep learner well-versed with learning goals
Collaborative Filtering	k-Nearest Neighbors [Schafer et al. 2007] cluster analysis [Hogo et al. 2010] Graph theory [Sun et al. 2014]	Naïve Bayesian Classifier Probabilistic models[Basu et al. 1998]	Domain-independent Adaptive and provides personalized recommendations.	Suffers from Cold Start problems, i.e. not enough user ratings to make accurate recommendations. Scalability becomes an issue when comparing millions of user ratings.	Benefit from other learners' experience Assign learners into groups based on similar learning goals
Hybrid	Combines two or more techniques to increase performance and avoid the drawbacks of each technique.		Building a unified model More accurate recommendation. Avoids cold start problems.	Scalability can become an issue as the data set grows over a period.	Useful for beginner level learners as well as experienced learners.

CHAPTER 3: LMS - A need for a specialized system

E-Learning is considered a multibillion-dollar industry today with rapid growth. There are many generalized E-Learning systems available which offer the same architecture and services for all domains of sciences. Due to a considerable expansion in the E-Learning industry, there is a need to provide subject specialized E-Learning systems to meet the individual needs of the learner rather than a generic E-Learning system across all discipline. The specialized E-Learning systems should be centered towards personalized learning.

The recent era has seen a massive proliferation in the E-Learning industry. “The Global E-Learning market is accounted for $165.21 billion in 2015 and is expected to reach $275.10 billion by 2022 growing at a CAGR of 7.5% during the forecast period.” [1]

As a result, traditional learning management systems (LMS) are faced with the challenge to provide not only a genuine alternative to brick and motor classroom environment but a more enriching learning experience. The scientists have started to question the effectiveness of traditional E-Learning systems. Although E-Learning systems have brought about great success stories, it has been investigated whether the E-Learning has been successful in bringing about the long-anticipated paradigm shift [Lennon et al. 2004].

In the next section, general E-Learning systems have been critically examined and how they have fallen short in fulfilling the need of diversity of learners. In section 3.2 scientific community contribution was investigated. Similarly, possible facet was identified which can enhance the learning process. Section 3.4 highlights the need to utilize the scientific community contribution to create a specialized E-Learning system.

[1] https://www.reuters.com

3.1 General learning management system (LMS)

In a traditional classroom, environment instructor can obtain feedback on student learning experiences in one-to-one interactions with students. The instructor can assess the learner needs by various means, for example learner's previous learning experience can provide useful information about her learning style. This feedback mechanism enables the instructor to recommend appropriate learning resources and tasks to support the individual's learning experience [Sheard et al. 2001]. With the evolution of the Web, researcher focused on developing E-Learning system, i.e. learning electronically without physically appearing in the classroom [Kurbel 2001].

There are varieties of LMS available in the market which support online learning by creating course material, designing student assessment such as quizzes, assignments, etc and provide online forums, blogs for peer learning. Examples of the most popular commercial LMS are blackboard/ WebCT, JoomlaLMS similarly free LMS are Moodle, Sakai, Docebo. [1]

These LMS are seen mainly as a mean to teach the masses adequately. In traditional LMS educational material is planned and designed by educational institutions and the instructors whereas the learner is expected to interact with the predefined pedagogical process. A generic solution is used across the different domain of education such as computers science, mathematical, and biological sciences.

Educators select the content for learning. In other words, knowledge is presented to the learner without considering the individual need of the learner [Stojanovic et al. 2001]. A universal solution is applied across culturally and linguistically distinct learners. Hence we can refer them as general LMS.

Let's look at an example of Moodle an open source Course Management System. Moodle is offering over 1.8 million courses, used by 1.7 million teachers in 270 different countries [2.]

---

[1] http://lms.findthebest.com/compare/83-228/Blackboard-Learning-System-vs-JoomlaLMS

[2] http://E-Learningindustry.com/top-10-E-Learning-statistics-for-2014-you-need-to-know

It is widely used by different universities, colleges, and training institute to provide online course content through the web. Moodle offers instructors, students, and educational institution features such as file downloading, assignment submission, online quizzes, online calendar, online news, and announcements, discussion forum, Wiki and grading. [1] Moodle has mainly been seen as a course management tool facilitating the instructor to create and manage online courses. Although Moodle has some great success stories to share, providing a useful personalized learning experience across discipline remain in question.

In E-Learning environments, a learner can discover the relevant knowledge by a search query or by following the navigational pattern set by the education provider. The relevance of content is left on the knowledge of the learner to extract according to her learning needs. The learner is often bewildered with the amount of information which is presented to the learner. The content needs to be presented in a supervise manner. The relevance of content to learning context is taken into consideration rather than learner selecting irrelevant content [Fischer 2012].

Learning management system should be more centered towards individual learner's need rather than facilitating the teacher as a course management tool. Technology should be seen as assisting individual learning experience rather than manipulating the static learning content. Many of these generic LMS provide a static approach across the different level of the learner assuming all learners are at equal intelligence quotient. Hence a generic pedagogy is adapted to meet the needs of many learners without considering their learning diversity.

In a classroom learning environment, the teacher can recommend relevant resources from diversified references. However, in a generalized LMS environment, a learner may come across a situation when the learner has some ambiguity about a particular concept within the LMS. There are two possibilities:

---

[1] http://barrysampson.com/2009/04/08/open-source-lms-10-alternatives-to-moodle/

The learner can independently go outside E-Learning realm, i.e. web (external environment) The E-Learning system can provide recommendations from the outer world.

In current E-Learning systems, the choice is left on the learner, as most of these learners are 'digital natives' they seek the help by going the outside E-Learning realm, i.e. web. When further information is presented to the learner about a particular topic learner was bewildered with the wealth of information which is offered in her learning context. The learner is faced with the challenge to find the most relevant information according to her current learning context, which is a tedious task to ask.

As a result, the learner is left in disarray and disengages in the learning process. Learner simply desired a recommendation on few of the most relevant material which suit her current learning context. This could have been easily possible in a physical classroom environment by merely requesting the instructor.

However, in the research practices; there are many proposed and somehow evaluated systems in closed settings perhaps in a specific domain which can provide recommendations from the outer world. In the next section, some of those specialized systems will be reviewed.

3.2 Adaptive and intelligent E-Learning

The emergence of web2.0 brought about innovation in the field of E-Learning.
"The Web 2.0 revolution has peddled the promise of bringing more truth to more people, more depth of information, more global perspective and more unbiased opinion from dispassionate observers" [Stove 2007].

Today's technology-oriented user who has the experience of social network where he can independently create, publish and redistribute content. The user finds traditional LMS structure inflexible as compared to the user-centered approach of Web 2.0 services [Craig 2007].

The scientific community is actively engaged to make the learning experience more and more productive and fruitful across distinct needs of the learner. There is a paradigm shift from teacher centered to learner-centered education [Lee et al. 2009].

One of the essential features which have been researched by the scientific community is to provide relevant resources at the right time [Saaya al 2013]. However, a recommendation in E-Learning is somehow different from other domains. For example, one has to keep in mind the learning context, is it a new concept or continuation of the existing knowledge and may necessitate a different type of learning resources [Manouselis et al. 2012].

Okazaki et al. employed a lexical database to discover the lexical relationship between terms which appear in a sentence. Lexical database usage provided data enhancement [Okazaki et al. 2003]. Li et al. presented a sentence similarity measurement based on a lexical database and word ordering [Li et al. 2006].

Chen et al. proposed a technique to recommend URLs on Twitter as a means to better direct user interest from the information pool. Twitter recommendation technique considered three independent elements into consideration content source, topic interest and social voting [Chen et al. 2010].

A small-scale variety of recommender systems are successfully recommending appropriate content in specific subject learning domain. Kumaran, V.et al recommender system uses learner information and domain knowledge (computer science,) to represent in the semantic net. Authors proved the method provided accurate course content recommendation [Kumaran et al. 2013].

However, this has been evaluated for one typical course. Such feature incorporation in a generalized system is a challenging task. There is an emergence of adaptive and intelligent systems where a learner can create his learning environment which best suit his learning need rather than technology provide his learning context. Learner past learning experience and current context can be used to provide personalize and adaptive learning experience.

These adaptive and intelligent systems are the joint venture of intelligent tutoring and hypermedia systems (AHS). Some examples of domain-specific ITS are SQL-Tutor, German Tutor, ActiveMath, VC-Prolog-Tutor, similar examples of AHS are AHA!, InterBook, KBS-Hyperboo WebCOBALT [Brusilovsky et al. 2003].

Adaptive and intelligent E-Learning system can be obtained by modeling a domain (using, for example ontology structure) pedagogical dataset (set of designed problems and their solution), data about user interaction and learner model, i.e. likes and dislikes [Romero et al. 2009].

A learning experience does not take place in isolation. An intelligent learning management system would keep track of the prior knowledge of the learner to suggest the knowledge according to the learner context.

Various Data mining techniques are used to mine user profile in a particular domain context. For example, Koutheaïr et al. recommend a system which mines users' web usage and learning materials during content and profiling phase in order to predict what to recommend to an active learner [Khribi et al. 2007]. However, such a feature can generate relevant recommendations if used in a specific E-Learning environment.

The personalized E-Learning content recommender system for example proposed by Lu et al. applied fuzzy matching law to determine relations between students learning need and list of learning content [Lu 2004]. They tested the system on a specific domain. The identification of learning needs and subsequently mapping the learning needs to the learning content is quite difficult in the generalized E-Learning systems.

Tane et al. applied text clustering and mining rules to arrange documents as per their topic and likeness [Tane et al. 2004]. Dwivedi et al. used a weighted hybrid scheme to recommend relevant learning content to the learner by modeling learning style and the knowledge with the collaborating filtering techniques. A learner with the same learning style and greater knowledge has greater weight in a recommendation [Dwivedi et al. 2013]. Such system when integrated with generalized E-Learning system may not produce good results because the learning style of students from the different domain may be similar and their collaboration data (co-downloads, co-views, click streams, etc) may be misleading.

Some website may not be accessible through search engines however a direct link can point to these useful resources. However these independent, unconnected (via search engine) web content belongs to the broader web [Wright 2009]. Social networking sites such as Facebook and Twitter can provide access to these buried resources.

Maloney states "social networking sites such as MySpace and Facebook have shown, among other aspects, that students will invest time and energy in building relationships around shared interests and knowledge communities”[Maloney 2007].

Wiki-Learnia is an example of such system where with the help of different semantic filtering mechanisms relevant knowledge from external sources such as Facebook, Twitter and YouTube is extracted, acquired and distributed. An E-Learning hub is designed which involve different repositories and filters, collects and allocates the information on a specific learning domain [Waßmann et al. 2014].

3.3 Generalized LMS vs. Specialized LMS

The research indicates there is a considerable gap between the general LMS and the need for ‘Digital Native' learner. These general LMS were designed by ‘Digital immigrants' who have failed to cater the ever-growing demand of ‘Digital Native' learner [Prensky 2001].

General LMS are primarily seen as course management tools. The course designer and instructors have a fixed navigational pattern in mind when designing the online course and the material. It is assumed all the learners will follow the same logical path which is set in design. However, a learner could follow different navigational paths prompting a series of learning activities [Zaíane 2002].

As we have investigated in general LMS section, the danger of providing unsolicited freedom of knowledge discovery can easily deviate the learner from his/her learning pattern, and learner can waste time in unnecessary activities.

Instead, a specific E-Learning system should be provided to the learner who will make the global information available in the local social context of the learner. It should enable the learner the elasticity to discover, organize, share information in a locally meaningful fashion which is globally accessible [Boyd 2006].

A general E-Learning system provides a generic profile for all learners. The same content is presented to diversified skill learner. However, a specialized system can dynamically build a learner profile and deliver the appropriate content at the right time to support individual learning. Dynamically profile can be build by considering learner recent navigational history and similarity and dissimilarity among the content of the learning resources [Khribi et al. 2008]. Learner profile features are evaluated against learning objects features based on the result and new promising learning objects are recommended to the learner.

Existing recommender systems have the limitation of domain dependency, cold start, overspecialization, and sparsity [Marin et al. 2014]. The research community has provided a significant contribution towards active and authentic learning. A variety of recommender agents are successfully recommending appropriate content in limited knowledge settings, perhaps specific to one course. There is a dire need of specialized E-Learning system which can model the entire knowledge domain (for example, computer science) by using techniques investigated in the section3.

The quality of a recommendation can be enhanced in various domains by combining different recommendation techniques. However, it should not be seen as a generic solution to overcome these limitations. Users who have similar preferences in one domain may not share the same in other domains.

Similarly, specialized E-Learning system should also provide a comprehensive pedagogical dataset model and different learner model. In this manner, specialized E-Learning systems can provide a more accurate recommendation according to the personalize need of the learner.

Specialized E-Learning systems can be semantically more affluent and provide better diagnostic analysis than data from general LMS [Merceron et al. 2004]. Specialized E-Learning system would model entire knowledge domain hence searching the relevant information would be more enhanced than general E-Learning system.

3.4 A Need for specialized systems

In this chapter, we investigated the functionalities of generalized LMS. Contemporary E-Learning systems are viewed as course management tools rather than facilitating learning needs. However, there have been many attempts by the scientific community, in a localized domain. These efforts prove to be quite impressive features and can be considered as real needs of the learners. Finally, it is highlighted the need for constructing a specialized (domain specific) E-Learning system. Such a system can help learners in their perspective domain learning and can assist them according to their particular needs, context, profiles, history, and collaborations.

CHAPTER 4: Resource recommendation using Social bookmarking

In this chapter, we will explore how social bookmarking can be utilized to provide domain-specific recourse recommendation.

Current Search Engines, Citation Indexes, etc heavily rely on the index of the keyword and returns all the pages where the searched keyword is matched. Yes, based on some ranking criteria while using certain factors. Many techniques have been used in the literature to extract useful knowledge from unstructured text such as search engines and Wikipedia. Some approaches use Machine Learning [Milne et al. 2008] techniques. The automated approaches solve the problem with some error margin and have inherited problems. However, the engagement of the social community to provide structure is of great importance. Therefore, the research community is striving for giving a structure to the Web using the paradigm of Social Web and Semantic Web.

Social bookmarking sites such as CiteULike, Bibsonomy, and Delicious have engaged social community from the World to provide useful Keyword (tags) to resources. In this way, on the one hand, resources are given structure, and on the other hand, social community shares exciting, up-to-date, and essential material. Social Bookmarking systems have become quite successful in a short period. One such system is known as “CiteULike” which has more than 6 millions research resources, and thousands of resources are shared on a daily basis.

Users of CiteULike annotate resources with useful Keyword termed as tags to give a structure to these resources which can be exploited for the discovery of useful resources. Therefore, there is a need to automatically extract useful resources from social sites and recommend the most relevant ones to the learner.

In the following sections, we propose and implement a method to find the most relevant resources from CiteULike to recommend the e-Learner.

4.1 Recommendation method overview

Whenever a user shares a resource, the metadata is also provided such as Title, Keyword, Author, etc. This metadata reflects the main dimensions, topics, techniques related to the resource. This metadata along with resources are stored in the database as shown in *Figure 4.1.*

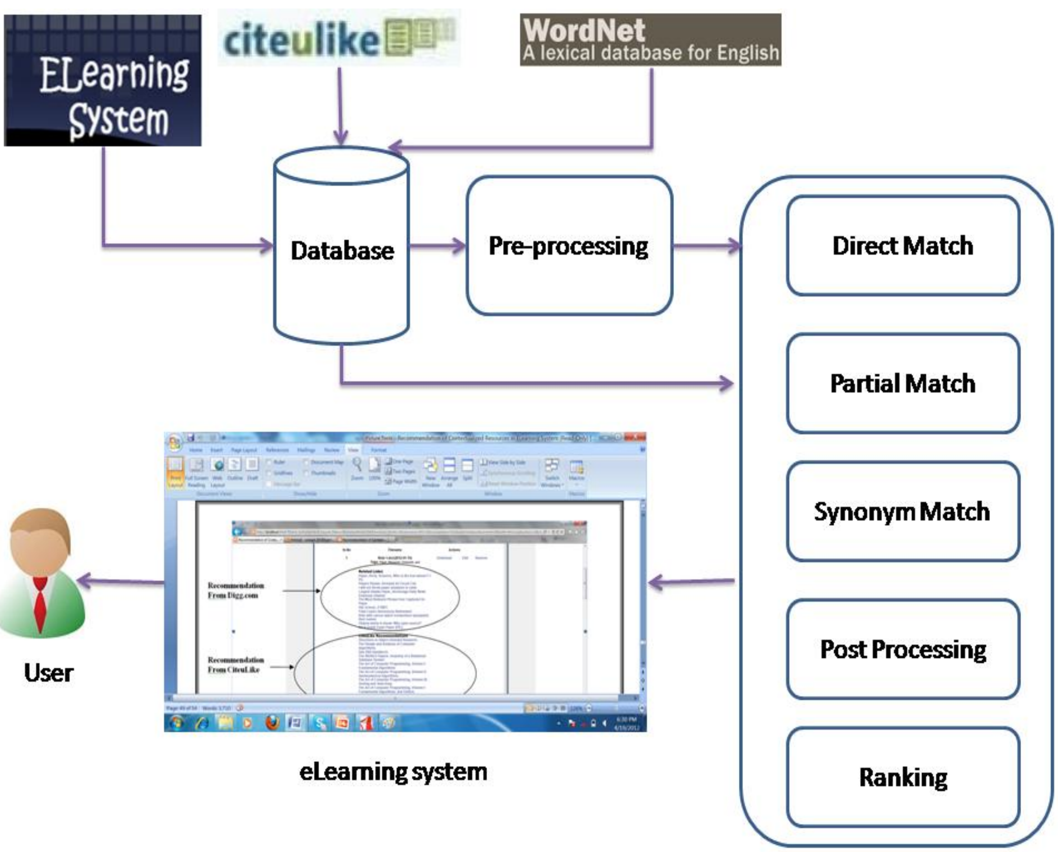

Figure 4.1 Recommendation method with social bookmarking

The proposed method extracts this metadata to find the relevant resources for the focused resource. Subsequently, the metadata are further enriched from the Wordnet database by incorporating synonyms set of each metadata. This enriched metadata set is searched within the database of CiteULike using direct match, partial match and synonym match. All of the matched resources become candidate relevant resources for the focused resource. Furthermore, the ranking is achieved by comparing a set of all metadata of the resource with the set of tags (tag-cloud) of each resource from CiteULike. The resources are ranked based on the maximum matching of metadata set and tag cloud set. The architecture of the proposed method is given in *Figure 4.1*. All of the discovered resources are made available within the learning context of the learner. Therefore learners do not need to query; instead, the learners are receiving all of the relevant, up-to-date, and essential resources about their current learning context.

4.1.1 Preprocessing

The data of CiteULike is pre-processed to remove noisy tags. It was found there are many irrelevant and useless tags which have been provided by users of CiteULike such as ABC, XYZ, computer, etc. It was also observed in our previous research that some tags represent the future context of use rather than the content of the paper. Some heuristics were made to remove the noisy tags, for example, the tags having a length of more than 30 or up to 3 were marked and removed after manual verification. Furthermore, the tags which are not a standard dictionary words were removed.

4.1.2 Keyword-tags matching

A number of techniques are employed for the matching of Keyword of a resource from E-Learning system with tags from CiteULike. Initially, a direct match was performed. However, manual inspection shows the partial match may be helpful. Therefore, a partial match was also performed with some heuristics. A typical example of a partial match could be the match of a keyword such as "wiki" with other tags such as "wikification," "semantic wikis" and "biological wikis."

This match was found to be of great importance because when a learner is reading a resource where the keyword was "wiki," the system can recommend resources. The recommended resources are not only deal with wikis but also related concepts such as wikification and biological wikis can be discovered. This means the system could make serendipitous discoveries for learners as well. However, the partial match resulted in some noisy tags as well which need to be removed.

Furthermore, synonyms from Wordnet were utilized to match Keyword and tags. This is because sometimes, a learner may annotate a resource with a keyword k_1; however, the social community may have used a synonym or set of synonyms for the keyword k_1. In order to achieve this, Wordnet 3.0 dictionary was used as this is the latest dictionary covering many new concepts which were not available in previous versions.

4.1.3 Post processing

After the matching has been done, it is the time to post-process the matched tags. As indicated in the previous step, during the process of a partial match and synonym match, there are times when noisy tags are found. The manual inspection helped to make heuristics to remove such tags.

4.1.4 Ranking

Once the match of Keyword and tags has been found, now the resources need to be ranked for learners. In the process of matching, the system has acquired all those resources from CiteULike where at least one tag is matched with a keyword of the particular resource from the E-Learning system. The ranking is based on a number of weights such as direct match weight, partial match weight, and synonym match weight. Suppose for a resource RI_i from E-Learning system, the candidate relevant resources from CiteULike are RC_1 to RC_n. The RCK will be ranked on the top based on a maximum match between Keyword of RI_i and tags of RCk. The ranked lists of resources are further pushed to the user's context within the E-Learning environment.

4.1.5 Case study

This section explains the working of the prototype with an example. Suppose a user of the E-Learning system has shared a resource in the E-Learning system. The title of the resource is "The Transformation of the Web: How Emerging Communities Shape the Information we consume." On sharing this resource, the user has provided a list of Keyword such as wiki, blogs, web transformations, Web 2.0, and social Web as shown in *Figure 4.2.* The system matches the Keyword with tags from CiteULike based on Direct Match, Partial Match, and Synonym Match.

Figure 4.2 CiteULike interface for the focused resource

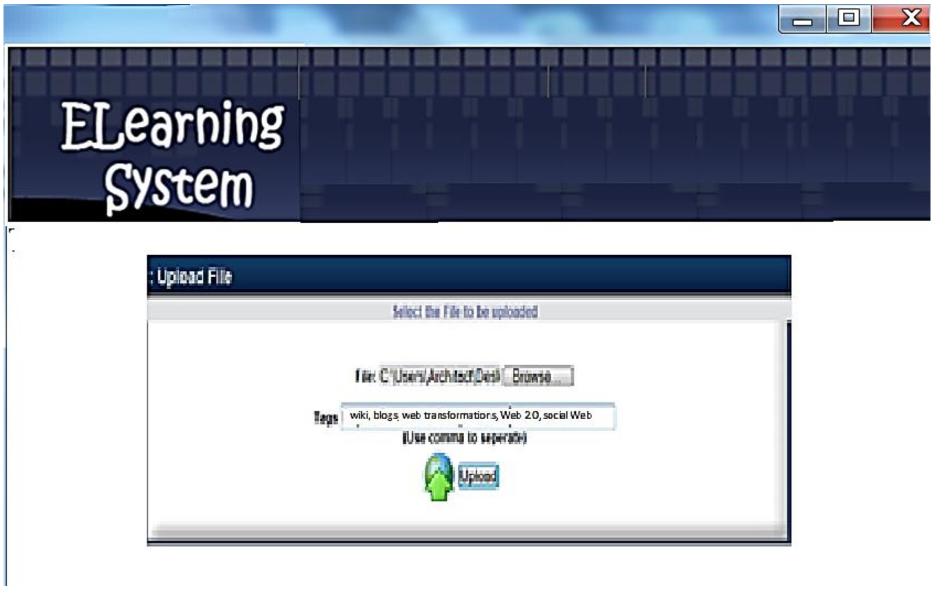

Figure 4.3 E-Learning System

For the keyword "wiki," the system identifies more than 200 papers from CiteULike. Similarly for the keyword "blogs," the system identified 200 papers, for the keyword, "web transformation," the system identified three papers, for the keyword "Web 2.0", the system identified more than 800 papers, and for the keyword "social web," the system identified 142 papers. Therefore, the total number of candidate relevant papers from CiteULike (RC_1 to RC_n) is 1345. This was just a direct match. Similarly, based on a partial match and synonym match, the system identified more than 5000 candidate relevant papers.

Subsequently, a ranking is achieved for all 5000 candidate relevant papers from CiteULike. For this purpose, all Keyword of the focused paper from E-Learning system are matched with all tags (tag cloud) of each candidate papers from CiteULike. The ranked list of resources from CiteULike is shown in *Figure 4.4*.

All of these ranked lists of resources are saved in the database of the E-Learning system. Whenever a learner is now reading the paper with the title: *"The Transformation of the Web: How Emerging Communities Shape the Information We Consume*," the user is shown ranked list of papers from CiteULike. It is obvious to see the ranked lists of resources are quite relevant with the paper being read in the local context of the user.

Additionally, these resources are giving versatile and important papers from many domains. This presents a breadth and depth knowledge related to the domain and topics of the focused paper. For example, the keyword "wiki" got a match with tags such as "wikification" and "semantic wikis."

Figure 4.4 Ranked list of resource

Papers such as: "Learning to link with Wikipedia" etc. are the resources which cannot be found without a partial match. This is giving an additional knowledge to the learning about the evolution of the field. A paper with the title "HIV/AIDS Stories on the World Wide Web and transformation perspective" is making aware the learner about the diffusion of the focused paper into BioScience.

Therefore, this list is providing enough knowledge to the learner for further exploration. Any of the discovered ranked papers can be clicked, and the learner is redirected to the CiteULike page. For example, following the link for the paper title, "Learning to link with Wikipedia", and the user is redirected to the screen. The user can read the abstract and other related information for the paper. All of the tags are also shown to the learner where the user is free to explore further.

Therefore, such recommendations may enhance the overall efficiency of the learner. The learner does not need to visit any external site. The learner does not lose her focus from the E-Learning system; rather the most relevant resources are pushed to the learner's local context.

CHAPTER 5: Can Twitter be useful for E-Learning recommendation?

The concept of big data is already present in the E-Learning domain. Social networks are facilitating the sharing of an enormous amount of learning resources through tweets, blogs, and Wikis, etc. Twitter provides an opportunity to share various learning material along with the short and precise text. Micro-blogging community is ever expanding. Discovery of the most relevant learning material becomes essential to solve the information overload problem. There is a potential need to provide twitter-based learning recommendation for the learner without directly inferring the tool. E-Learning recommender systems have been successfully supporting the individual learning process.

However, there is a gap between the E-Learning recommender system and Twitter-based recommendation such that no existing E-Learning recommender systems are utilizing the twitter. Furthermore, some Web pages may not be reachable via search engines however a direct link can point to these valuable resources. This independent web content belongs to the broader web [Wright 2009]. Social networking site such as Twitter can grant access to these obscured resources.

Personalize learning activities, and related tasks can be recommended to the learner according to the individual need of learner as a result overall learning experience can be enhanced. In order to achieve the best results, it is important to consider students learning style [Islam et al. 2015]. In literature, personalized recommender systems employ the characteristics of items, profiles of users and history of users and items interaction in order to recommend the related items. In this chapter, we shall investigate if the twitter can be useful for E-Learning recommendation and what are the parameters which should be considered when recommending E-Learning resources.

In the following section an overview of the literature survey is provided to discover twitter significance for E-Learning recommendation. Similarly, what are the current parameters used for twitter based recommendation? Furthermore, we shall empirically analyze if these parameters can be important for twitter based E-Learning recommendation.

5.1 Can Twitter be useful for E-Learning?

Twitter is a micro-blogging website which has monthly 320 million active users monthly with 500 million tweets sent per day.[1] Twitter users can exchange text messages of 140 characters limits. People post queries and reply, share ideas and resources/URLs and work together on problems of practice. Twitter is a popular social media among student and teacher. Twitter is more open to the public than Facebook and provides a fast way to exchange the ideas among peers [Ebner et al. 2010]. Twitter can provide instantaneous communication with the learning community.

In traditional LMS one has to log in and find the appropriate blog before posting a query and waiting for someone to reply. However, in the meantime interest levels may diminish. Twitter character limitation focuses the attention of questioner hence queries are precise. Similarly, the learner receives a precise answer to her query. Dunlap et al. investigated the research papers about smart learning education since 2007 by using five major search sites and highlighted the trends in smart learning. Twitter effectiveness in the educational environment was measured by experiment. Traditional learning method was compared with twitter assisted learning environment [Dunlap et al. 2009].

Twitter utilization experiment result indicates students who used twitters in their learning environment had better grades than those who did not use twitter. Almost all students with excellent grade used twitter [Ha et al. 2011]. Acar et al. analyzed twitter effectiveness in learning English as a foreign language course. Japanese students' tweets were analyzed, and it was discovered students were active using the twitter to assist in the language course [Acar et al. 2012].

[1] https://www.statista.com/topics/737/twitter/

Junco et al. evaluated student participation and collaboration in two different classes and its impact on learning outcomes. The class where Twitter was required to be used outperform than the class where it was optional. Twitter usage indicated improved learning outcome [Junco 2013].

Kassens carried out a study to conceive the role of Twitter for out of class learning. The author proves that student who receives daily tweets about course content remember course topics better in exam situation [Kassens, 2012]. An empirical study was conducted to measure the role of Twitter in the learning environment. Twitter utilization results demonstrate that it provides a useful mean for sharing information and collaboration among students. Students with more number of followers and following had better grades than those students who were not actively tweeting [Ha et al. 2014].

Chris Evan's study of Twitter usage among university students and teachers indicates positive association among them. Student's involvement in university activities was similarly encouraging [Evans 2014]. Additionally, user case study carried out by the author proves that Twitter can positively impact teaching, learning and students experience. Also, the results show 79% of students were using Facebook while 57% students were using Twitter before the beginning of the courses. It proves student already find social networks such as Twitter a helpful tool in their learning environment [Reed 2013].

Literature survey proves the importance of using Twitter in E-Learning domain. However, most of the potential benefits highlighted in the research are using Twitter as an independent tool or incorporated within existing LMS. Considering the great significance of the Twitter, there is growing need to provide recommendation to the learner without explicitly inferring the tool. The next section will discuss learner's parameters which should be considered when providing a recommendation from the twitter.

5.2 Recommendation parameters review

Kwak et al. acquired 41.7 million user profiles, 1.47 billion social contacts, 4262 trending topics, and 106 million tweets through twitter public API. In order to find out the influential users in twitter three different methods were compared namely number of followers, page-rank, and many retweets.

First two ranking methods were found to be similar whereas ranking by retweets as different. Similarly, Twitter follower graph depicts a tilted division of followers and a low number of shared ties among its users. It proves twitter as information sharing network rather than social network [Kwak et al. 2010]. Another approach was designed and implemented to recommend news articles based on popularity and user profile by using the Twitter public timeline. Similarly, a hybrid news recommendation model was implemented by combining both approaches. In the end, all three methods were evaluated, and the results were provided [Jonnalagedda et al. 2013].

In order to find the most popular articles, user's pre-processed tweets were used as a query in the SOLR. Cosine similarity method is used to determine how well tweet contents match with the news articles. As a result, each news article is assigned a weight. Thus the most popular article has the highest weight across all the tweets. Hybrid recommendation was employed with two different methods. In the first method weights of popularity and user profiles were merged by multiplying with each other.

In the second method, it uses an adjustable parameter (a) to measure the significance of recommendation based on popularity as well as user profiles methods. Hannon et al. proposed Twittomender system which recommends new users to follow on Twitter. Authors employed eight different recommendation approaches with the combination of four content-based, three collaborative and one by combining the seven approaches [Hannon et al. 2011]. Garcia et al. evaluated two features namely popularity and activity which could be useful for recommending followers.

The authors have investigated to measure the impact of these features to persuade users to follow other users [Garcia et al. 2010]. Phelan et al. presented a method for news recommendation by exploiting real-time Twitter data in order to rank and recommend articles from a collection of RSS feeds [Phelan et al. 2009]. Magnuson et al. proposed twitter based event recommender system which discovers the twitter activity related to previous events in order to steer geographic recommendations based on item-based collaborative filtering [Magnuson et al. 2015].

A personalized recommender system was proposed based on Twitter user’s timelines. Tweets are ranked as per the user's interests. User's interest is derived from user social features, interactions and the content history of user's tweets. User interest is dynamically model as time-variant in different topics in order to facilitate the change of interests over time [Elmongui et al. 2015].

User's short-term interest from Twitter and long-term interest from YouTube is integrated. User's short-term interest is extracted from users' tweets, and related videos are obtained from YouTube. Similarly, users' Long-term interest is based on the information extracted from YouTube. Videos are recommended to the user based on the ranking criteria consist of users profile in YouTube, time factor and quality factor [Deng et al. 2015]. The literature survey indicates the parameters (profile, current context, and history) are used for Twitter-based recommendations.

5.3 Recommendation parameters significance for E-Learning

Literature survey in the previous section proves that user profile, history, and current context are valuable parameters to consider when providing a recommendation from Twitter. However will these parameters also useful for E-Learning recommendation.

The user study will provide the answers for the following questions from real life learners:

- Do any of these parameters play any significance for E-Learning recommendation?
- Which of these parameters should be given more importance while ranking for a recommendation?

As per our knowledge, this is the first comprehensive effort to evaluate the effectiveness of these parameters for E-Learning recommendation. Different university level learners were selected for this study. Learners were from diversified background bachelors to Ph.D. level.

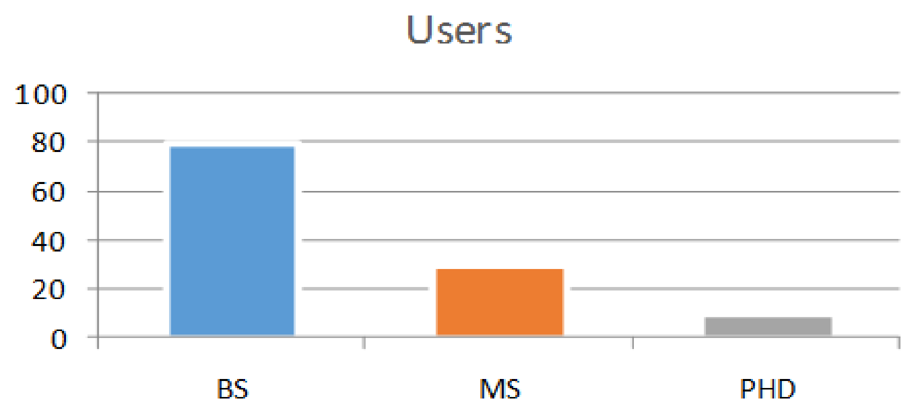

Figure 5.1 Learners status

The number of student's current status is highlighted in *Figure 5.1*. Out of 108 volunteers, 69% are undergraduate students, 23% are graduate students and 7% Ph.D., students.

The evaluation form is provided to each learner *(Appendix1)* The evaluation form contains information related to learner's educational background learning scenario, keyword based on parameters (profile, current context, and history) and tweets.

Learner reads the learning scenarios in the evaluation form and evaluates keyword relevance with tweets. For each learning scenario, nine tweets are provided to the learner related to all three user parameters Context, profile, and history.

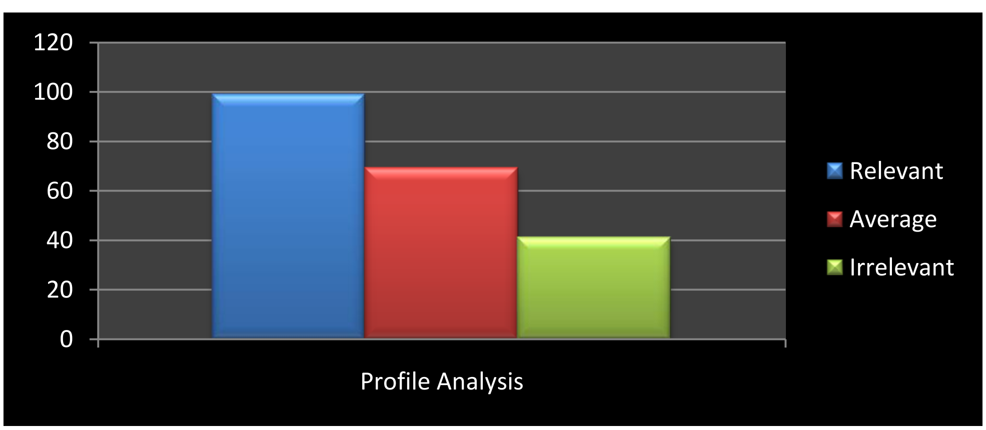

Figure 5.2 Profile Parameter

Learner reads the tweet and marks tweet relevant, average or irrelevant as per learners' parameter. Learners do the same steps for all tweets.

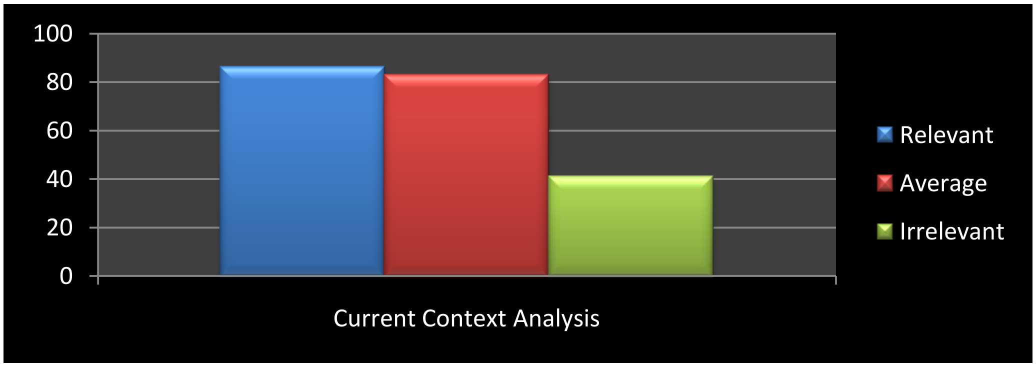

Figure 5.3 Current context parameter

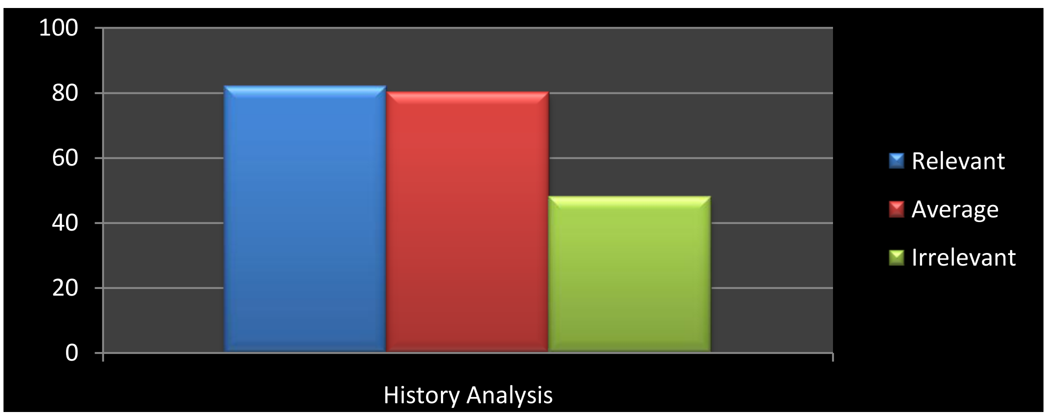

Figure 5.4 History Parameter

Table 5:1 Current context result

Current context		
Value	**Frequency**	**Percentage**
Relevant	86	40.37
Average	83	38.96
Irrelevant	41	19.24

More keywords from a category matched with tweets will return greater number which is divided with the total number of tweets. This score then multiple with weight. Finally, we will sum up values from the local context, history, and profile.

Table 5:2 Parameter *History* Result

History		
Value	**Frequency**	**Percentage %**
Relevant	82	38.49
Average	80	37.55
Irrelevant	48	22.53

Each parameters weight is determined using the following formula:

Context = *Context / (Context + History + profile)*
Profile = *Profile / (Context + History + profile)*
History = *History / (Context + History + profile)*

This formula will provide us score for each parameter. These scores will be sorted in descending order. The Figure given below shows the final results of the parameters profile local context, and history.

Table 5:3 Parameter *Profile* result

Profile		
Value	**Frequency**	**Percentage %**
Relevant	99	46.47
Average	69	32.39
Irrelevant	41	19.24

Table 5:4 Total tweets and users

Tweets & Users	
Total Users	108
Total Tweets	213

The final results indicate these parameters (profile, current context, and history) have great significance for E-Learning recommendation and all three parameters should be given equal importance when ranking twitter based E-Learning recommendations.

5.4 Twitter-based E-Learning recommendation

The social community is engaged in the sharing of a high number of essential and recent resources with each other. Social networking site Twitter has great potential to take the learning beyond the boundaries of the classroom and diminish the difference between classroom and distance learning.

However, contemporary E-Learning recommender systems have not exploited the strength of Twitter to recommend E-Learning resources. This chapter highlighted the significance of using Twitter for E-Learning recommendations. Similarly, it reveals various domains have been utilizing user parameters (profile, current context, and history) in order to provide twitter-based recommendations. However, such parameters have not been applied for Twitter-based E-Learning recommendation. A user study was carried out to evaluate the effect these parameters for E-Learning recommendation. It was identified that all of the above three parameters are equally important to consider while providing E-Learning recommendation from Twitter.

CHAPTER 6: Semantic-based E-Learning Recommendation

This chapter will present a semantic-based recommendation technique for e-learner. This is a novel alternative to conventional recommendation techniques where social network tool twitter is utilized. Relevant tweets are recommended to the learner as per the current learning topic of the learner.

The influence of the internet in education takes the learners and teacher interaction into a new realm which was previously not available. A large amount of information illustrating teacher and learner interactions is continuously produced and ubiquitously available. A wide range of learning contents is available by the press of a button. However, this vast wealth of information can also be problematic if not organized in a structured learning path.

More recently personalized and intelligent E-Learning systems offer personalized learning experience by constructing the learner model based on learner aims, likes and existing knowledge. A learner should be able to create, manage and organize the knowledge according to her personal knowledge management capabilities. Learner past learning experience and current context can be used to provide personalize and adaptive learning experience.

The current context consists of current topic viewing and keywords related to this topic. Based on the current context key terms, tweets will be fetched from Twitter and stored in a local database. Similarly, key terms will be extracted from learner's profile, history, and current context and stored in the database. Pre-processing will be performed on the extracted key terms to obtain unigrams which will be extended with the help of Growbag database. These extended keywords help us to perform semantic-based matching.

Lexical matching will be performed on tweets against extended key terms. Selected tweets are sorted in descending order concerning their occurrence frequency. Tweets are ranked by using the relevancy similarity. Tweets with high score finally recommended to e-learner.

6.1 Proposed Model

In this section, semantic model is described in detail and how the semantic model can provide the useful recommendation for the e-learner. The aim is to provide the learner personalize learning activities and tasks which suit best its individual needs. As a result overall learning experience is enhanced. Similarly, recommend related tasks and activities based on previously completed tasks by the learner or their peers.

6.1.1 Gold Set

A gold set was nonexistent in order to evaluate the effectiveness of the technique. Therefore a user study was carried out to design the gold set. We started with the ACM Computing Classification System 2012. 60 Domain Experts selected based on ACM CCS 2012. The proposed technique works on research papers, so we collected five research papers from each expert. The Proposed technique uses tweets for a recommendation; therefore, we requested ten tweets per paper from domain experts.

In total 220 research papers and 2957 tweets will be our gold set and will be treated as a benchmark. Collected research papers and tweets will be saved in a local database. After performing calculations, an evaluation will be performed concerning this dataset.

6.1.2 Input for tweet ranking

Two different types of input were provided to the proposed technique for recommendations for example, Paper's Metadata and tweets. Paper's metadata consists of research paper title, keywords and ACM CCS 2012 category from which it belongs. Besides metadata, complete tweets collection will also be provided.

6.1.3 Pre-processing

Before performing any lexical matching, some pre-processing is required to make it ready for use [Pang et al. 2015]. With the help of natural language processing, pre-processing will be performed on metadata content [Dai et al. 2015].

During pre-processing, word tokenization, normalization and stemming are performed. Non-alphanumeric characters were removed from starting and ending of a string to make it meaningful. Stop words were also removed during this process. Pre-processing phase resulted in the form of Unigrams.

6.1.4 Extend unigrams semantically

The unigrams have been extended with the help of two different databases. One is computer science domain-specific database known as Growbag [Arenas et al. 2014] and the second database is synonym-based database called Wordnet [Abdullah et al. 2015]. Each unigram is compared to these databases to obtain domain specific and synonym-based meanings of the words. These extended terms help us to perform semantic-based matching [Ley 2009].

6.1.5 Lexical matching

Each tweet is analyzed against extended key terms using Lexical matching. Selected tweets are sorted in descending order concerning their occurrence frequency. Proposed technique performs lexical matching on extended terms [Pang et al. 2015].

6.1.6 Precision / Recall calculation

Finally, precision and recall are calculated on returned rows. Matched tweets are compared with the tweets given by domain experts. On the basis of matched results, precision and recall score is received [Egghe 2015].

6.2 Evaluation

After receiving the results the evaluation was performed to find out the accuracy of the technique. Gold set was used for this purpose. The ranked tweets were compared from each metadata category with the tweets given by experts. The more tweets belong to the gold set reflects the better performance of the algorithm.

The architectural diagram of this proposed technique is illustrated in Figure 6.1.

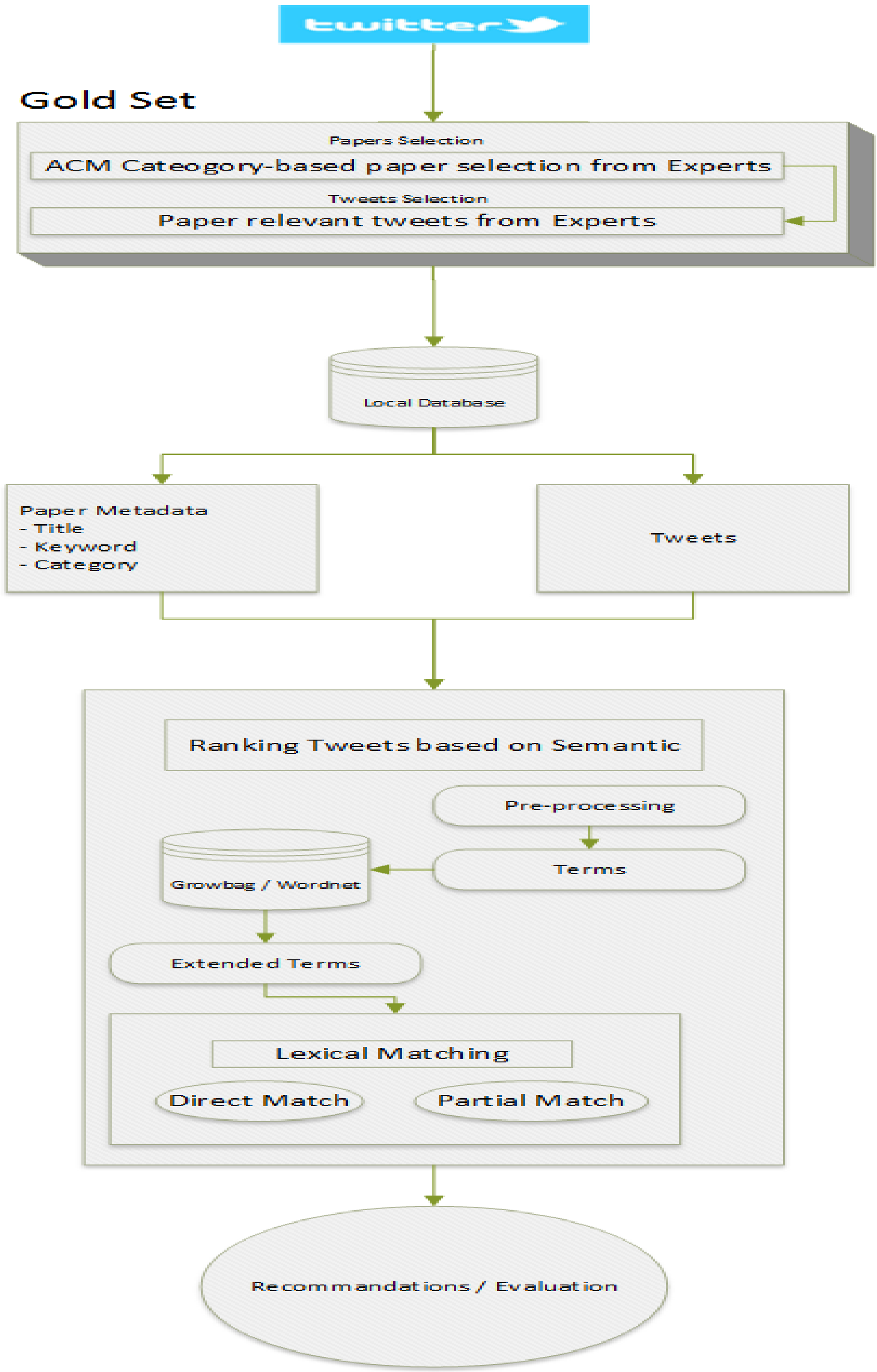

Figure 6.1 Architectural diagram of the proposed model

6.3 Results

Below find a graphical representation of the precision/recall score calculated in the previous step. The result is sorted in ascending order based on recall scores.

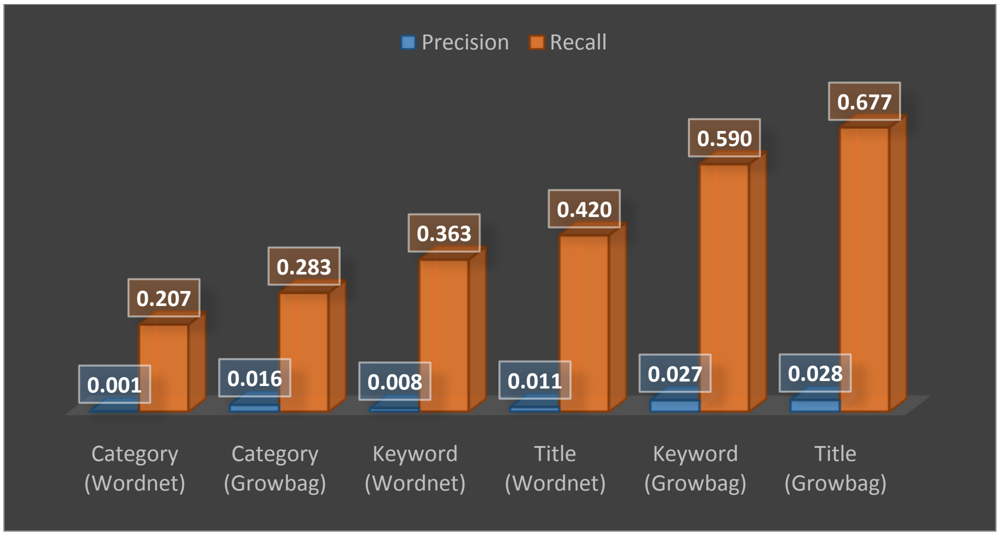

Figure 6.2 Precision / Recall score of returned records

The graph shows the recall is performing relatively better than precision. One reason is when we extended terms from Growbag and Wordnet database we received ten times more terms. As result retrieved rows increased many times. Hence precision went down. Recall and Precision are inversely related. An Empirical study of retrieval performance indicates when Precision declines Recall increases [Buckland et al. 1994].

Another important point from the graph is Growbag extended terms performed relatively better than Wordnet. This could be due to expert's tweets had more computer science specific terms. On the other hand, Wordnet extended terms are with respect to English dictionary. These terms were not commonly found in expert's tweets.

CHAPTER 7: Metadata evaluation for twitter based E-Learning recommendation.

Research community has been continuously working on innovative methods for learners to access the most relevant learning resources. There are two essential units to consider for E-Learning recommendation which are Learner and learning resources. The essential features used in the literature for learning resource recommendations, are metadata (Title, Author, Category, keyword) content and extensive vocabulary (Synonyms, Grow bag). The aim of this chapter is to provide evaluation of the resource metadata for twitter based recommendation. This entails how the metadata of resource should be combined. Similarly, what combination of metadata provides the best result for a Twitter-based recommendation?

Metadata of resource are used as input for different recommendation techniques, namely lexical matching, semantic similarity, and extended vocabulary. In the next section, an overview of learning resource metadata model is provided. Similarly evaluation of metadata for twitter based E-Learning recommendation has been discussed.

7.1 Learning resource metadata model

The aim is to identify the effectiveness of individual and group metadata features for E-Learning recommendation. Three recommendation techniques namely lexical, cosine similarity and extended vocabulary were applied to metadata (Title, Author, Keyword, and Category).

Each technique recommended the relevant tweets based on metadata. The recommendation techniques results were evaluated against the gold set. The next section will elaborate on how the gold set was formulated, and subsequent sections will discuss the recommendation techniques in detail.

7.1.1 Gold Set

A gold set was devised in order to evaluate the significance of metadata resource against each recommendation technique. There was no benchmark available in the literature for this evaluation. A comprehensive gold set was devised by extracting all the topics from the ACM classification system and selecting domain experts who have published papers in the relevant topic area. Total 60 domain experts participated in this study, and they were active authors who have published papers in the relevant topics of ACM classification. Each domain expert was requested to provide five research papers and ten tweets related to metadata of each paper. In total 220 research papers and 2957 related tweets were collected as gold set.

7.1.2 Pre-processing

Preprocessing plays a pivotal role for acquiring healthy results. At this stage lexical noise present in the form of the unique character is removed. Word Vector Tool was utilized in this regard [Rousu et al. 2006]. The details are as follows:

- Metadata (Title, Author, Keyword, and Category) of domain experts' research papers were extracted.
- Comma based split is performed on the above metadata to obtain individual entities, for example the name of each author from Author metadata.
- Term normalization was used to change all characters to lower-case and equivalent terms were normalized, for example, customize /customize.
- Greedy tokenizer was used which smartly handles digits and returns unigram.
- Stop words have been removed based on a standard list of words. Similarly, non-alphanumeric characters were removed from the start and end of a string to make it meaningful.

- The Porter stemming algorithm is applied to these tokens to obtain root words [Porter et al. 1980]. Porter stemming algorithms removes the suffixes and also stems the term.
- These stemmed words are pruned with a frequency lower than 3 in the collection.

  In this manner, preprocessed metadata items will be used in all three (lexical, cosine and extended vocabulary) recommendation techniques.

7.2 Metadata with the Lexical Approach

Metadata of fifteen different combinations were used as an input query. Metadata combinations were split into four sections namely, solo metadata, binary metadata, trio metadata, and all metadata.

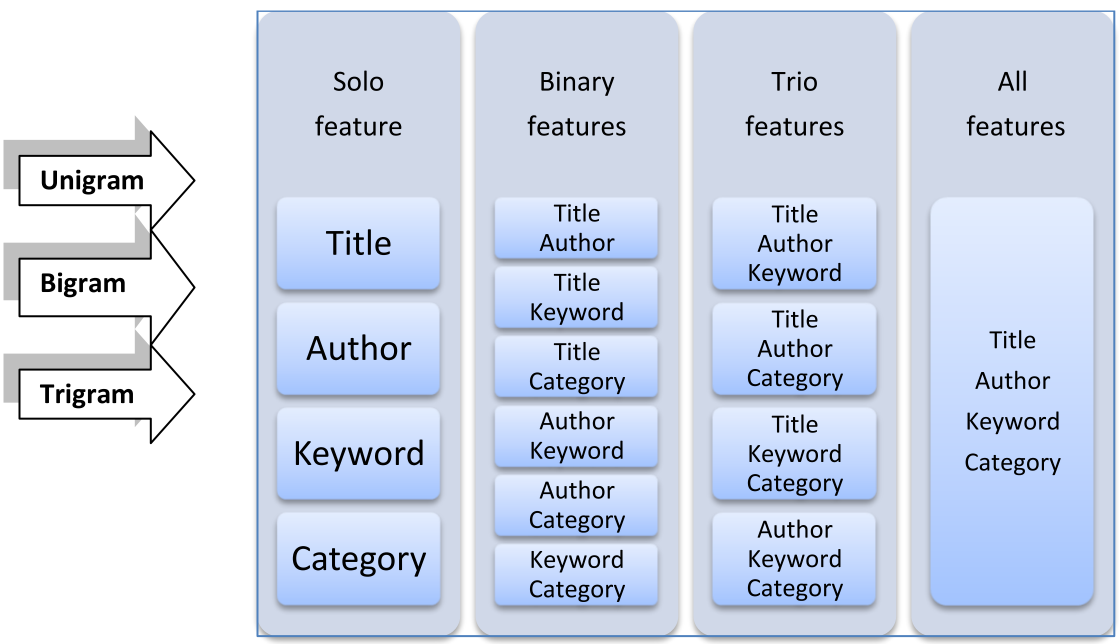

Figure 7.1 Metadata Combination for Unigram, Bigram, Trigram

All of these combinations were implemented with unigram, bigram, and trigram. An illustration of metadata combinations with n-gram is provided in *Figure 7.1*. Metadata with the Lexical approach is explained in *Figure 7.2*. The results of this technique will be discussed in the evaluation section.

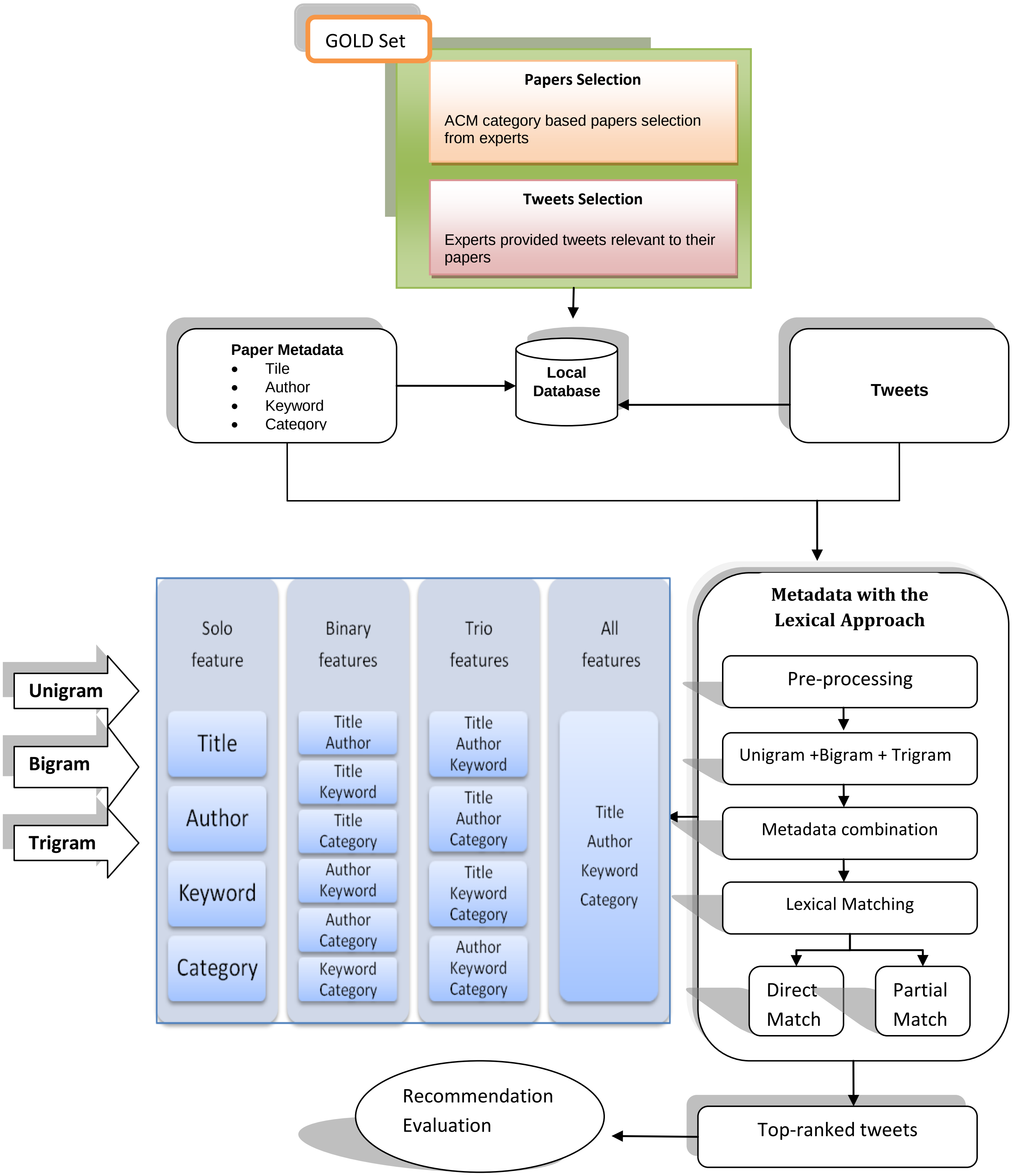

Figure 7.2 Metadata with the lexical approach

The algorithm takes domain expert's research papers metadata and its combination *(Title, Author, Keyword, and Category)* as an input query. Relevant tweets are computed as per the input query, and output is provided in the form of recommended tweets for the user. Algorithm steps are as follow:

Algorithm: Metadata with lexical Approach

1. Extract metadata (Title, Keyword, and Category) from research papers provided by domain experts.
2. Split metadata items (E.g., Comma based separation in keywords, category data)
3. Term normalization, changing all characters to lowercase, removal of accents
4. Stop word removal
5. Tokenization
6. Stemming
7. Prune in the collection of words with a frequency lower than 3 for each word from metadata
8. Obtain tweets related to metadata of each paper from gold set.
9. Fetch tweets related to the metadata using the lexical matching method.
10. Calculate ($pricision = \frac{TP}{TP+FP}$) using output from step 8 & 9
11. Calculate ($Recall = \frac{TP}{TP+FN}$) using output from step 8 & 9
12. Calculate ($F1\, Score = 2 * \frac{Precision\ *Recall}{Precision\ +Recall}$) using output from step 8 & 9
13. Repeat the steps 8 -12 for all 15 possible metadata combinations with unigram, bigram and trigram.

Figure 7.3 Metadata with lexical Approach

7.3 Metadata with TF-IDF Vector Space Model

In this section, metadata was implemented with TF-IDF vector space model. Term frequency-inverse document frequency (TF-IDF) is a renowned technique in the field of information retrieval or text mining.

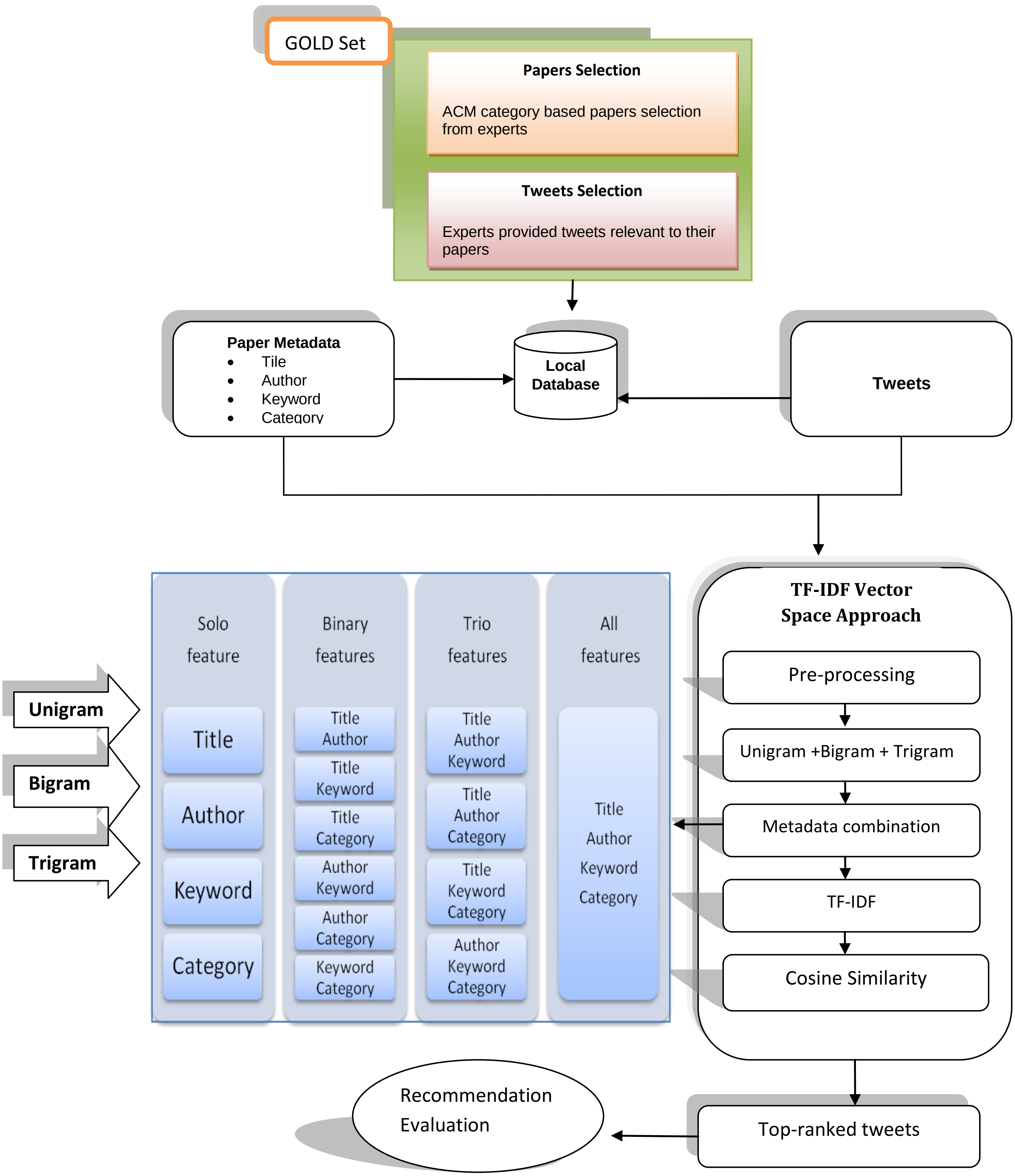

Figure 7.4 Metadata with TF-IDF Vector Space Model

This technique evaluates the significance of a word in a document. In the field of digital libraries, 83% of text-based recommender systems use TF-IDF term-weighting schemes [Beel et al. 2016].

After performing the preprocessing steps metadata query and tweets datasets were transformed into "bags of words," and were indexed. TF-IDF weighting factor was employed to measure how valuable a term (word) is to a tweet in a collection of tweets dataset. Python was used to create the language model for unigram, bigram, and trigram for 45 Metadata combinations. These Metadata combinations were split into four section solo metadata, binary metadata, trio metadata, and all metadata. Tweets are ranked based on matching score with each metadata combination.

$$score\ (q,t) = \sum_{t \in q} \mathrm{t\,f}\ \ (t,t).\log \frac{N}{tft}$$

Equation 7:1

Here, q is the metadata query, t is tweets, tf (t; t) is the term frequency of metadata query-term in tweets, N the total number of tweets (total data set), and tft the tweets frequency of query-term t in the collection. This is a commonly used scoring method. Both tweets and metadata query documents may be viewed as a set of vectors in a vector space.

In this vector space model, how do we measure the similarity between tweet and metadata query documents? A tweet with similar content to metadata query can have a significant vector difference because one is relatively longer than the other. Therefore the relative distribution of terms may be the same in both documents; however one may be far more significant than other in complete term frequencies.

In order to overcome the document length limitation, cosine similarity is used. Cosine similarity is the standard way of measuring the similarity between the query vector and tweet vector representations $\vec{v}(q)$ and $\vec{v}(t)$.

$$score\ (q,t) = \frac{\vec{v}(q).\vec{v}(t)}{|\ \vec{v}(q)\ ||\vec{v}(t)|}.$$

Equation 7:2

Vector space model (VSM) was used to compute the similarity between metadata query and tweets. Retrieval model assigns a score of relevance to tweets (t) from an index, given a metadata combination query (q). Both metadata query and each tweet were represented as weighted term vectors. As a result, a score is assigned to a (query, tweet) pair. The resulting scores can then be used to select the top scoring tweets for a query.

These tweets are matched against the gold set in order to check how well this technique performed. The same procedure was performed for all n-grams. The results will be discussed in the evaluation section.

7.4 Metadata with Extended Vocabulary

When short words are used to retrieve the relevant object, it often returns a large set of results. Filtering out relevant information from such vast results can be a laborious task. Growbag is an automatic classifier which filters out large result according to the objects' semantics content. Lightweight concept graphs are created by applying Keyword co-occurrences patterns algorithm. It uses DBLP dataset.

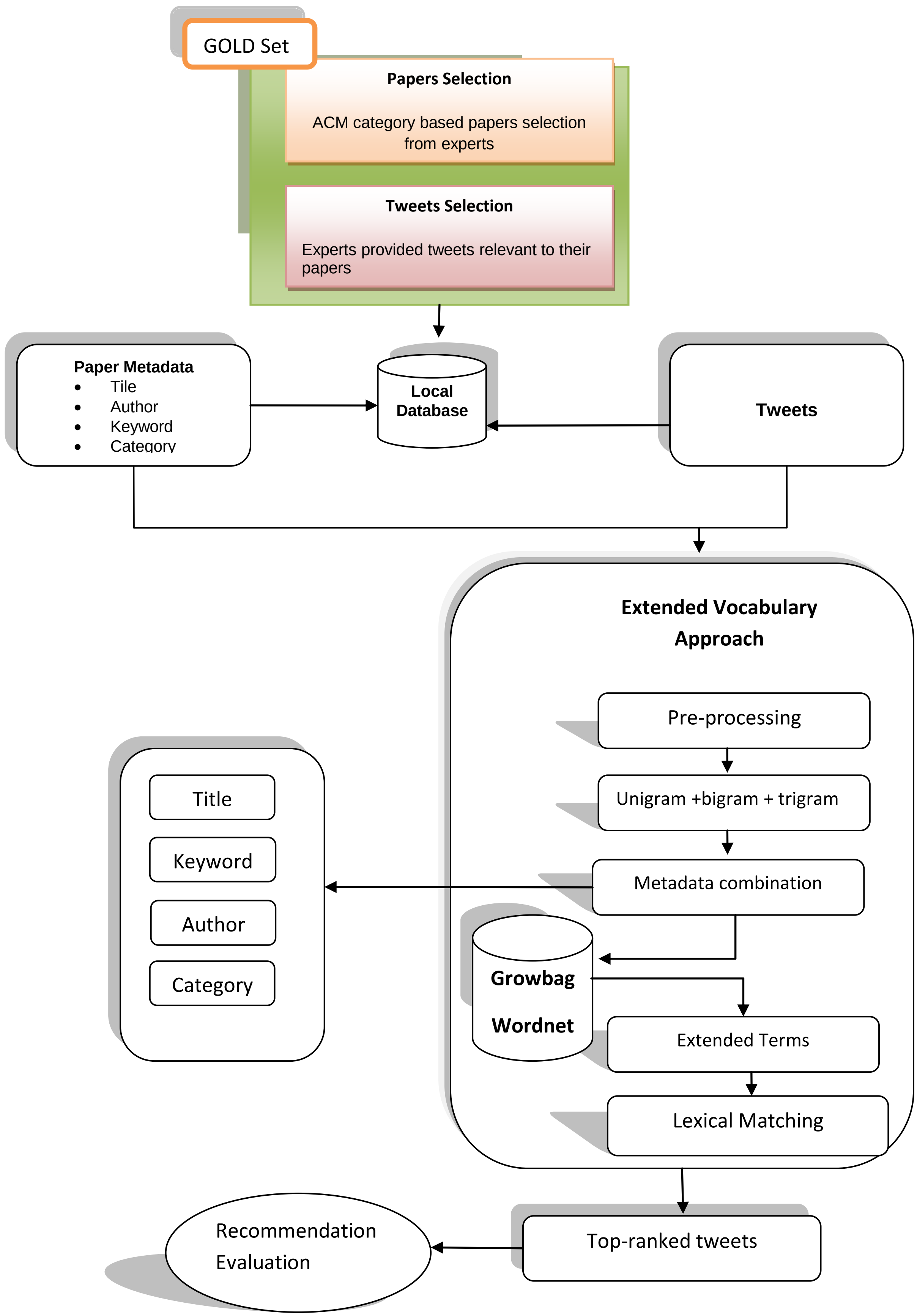

Figure 7.5 Extended Vocabulary Conceptual Model

There was substantial variation in experts' tweets to describe the same concept. Some computer science concepts co-occur with other similar concepts. For example, "recommendation" is sometimes referred to as "personalization." Additionally, different experts have used synonyms to describe the same concept.

Algorithm: Metadata with Extended Vocabulary

Input: The algorithm takes domain expert's research papers metadata and its combination *(Title, Author, Keyword, and Category)* as an input query.

Output: Relevant tweets are computed as per the input query, and output is provided in the form of recommended tweets for the user. Algorithm steps are as follow:

1. Extract metadata (Title, Keyword, and Category) from research papers provided by domain experts.
2. Split metadata items (E.g., Comma based separation in keywords, category data)
3. Term normalization, changing all characters to lowercase, removal of accents
4. Stop word removal
5. Tokenization
6. Stemming
7. Prune in the collection of words with a frequency lower than 3 for each word from metadata
8. Extend each word using Wordnet and Grow bag
9. Fetch tweets based on extended words by applying lexical matching.
10. Compare the retrieve results with the gold set.
11. Calculate ($pricision = \frac{TP}{TP+FP}$) using output from step 8 & 9
12. Calculate ($Recall = \frac{TP}{TP+FN}$) using output from step 8 & 9
13. Calculate ($F1\ Score = 2 * \frac{Precision\ *Recall}{Precision\ +Recall}$) using output from step 8 & 9
14. Repeat the steps 8 -12 for all 15 possible metadata combinations with unigram, bigram and trigram.

Figure 7.6 Extended Vocabulary Approach

For example, in the English language, a word can have multiple meanings, on the other hand, a word may mean exactly or nearly the same as another word. If we only match the word syntactical, it is quite possible we miss out some key terms which may semantically mean the same. There are many semantic similarity approaches used in literature for E-Learning recommendation [Chan et al. 2014].

However, these approaches have not been utilized for Twitter-based E-Learning recommendation. Wordnet and Growbag database were used to extend the terms in order to obtain the semantic meaning of the object.

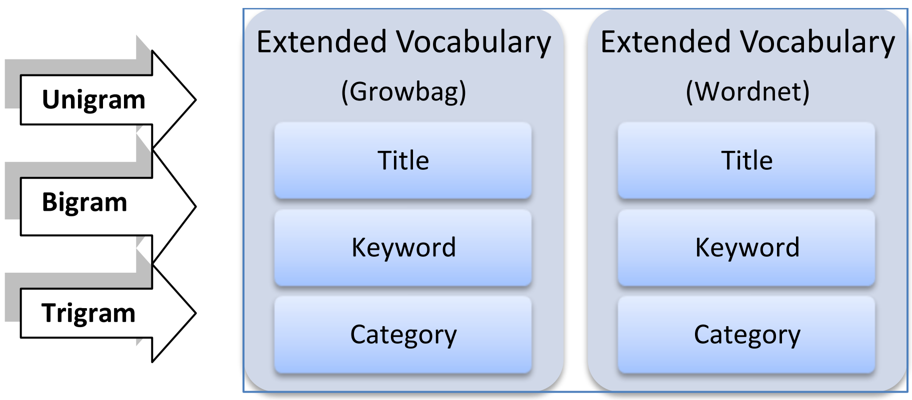

Figure 7.7 Different input for Extended Vocabulary

Wordnet is an English language lexical database. [Miller et al. 1990] It combines English words into sets of synonyms. Whereas Growbag automatically generates categorization from a collection of digital objects annotates with Keyword. [Diederich et al. 2007] After pre-processing, each word query is extended with Wordnet & Growbag datasets; as a result, one-to-many words are fetched. The dataset is computed by applying the lexical matching to find the co-occurrence of extended words in tweets.

The approach provided the top-K recommendations based on the similarity between query word and the tweets dataset. Extended vocabulary recommendation approach validity is checked against the gold set, and results will be discussed in the evaluation section. As the author name is not required to be extended semantically, therefore author metadata was eliminated for the semantic-based recommendation approach.

CHAPTER 8: Metadata evaluation with n-grams.

Comprehensive experiments evaluations established the usefulness of twitter based E-learning recommendations. Metadata (title, category, keyword, and author) was extracted from the papers provided by experts. In order to evaluate the significance of metadata for a quality recommendation, different approaches were employed. The evaluation discussion will be based on two hypotheses:

a) Which individual metadata (title, category, keyword, and author) outperforms in terms of F_1 measure.

b) How the metadata should be combined for the best quality recommendation based on the F_1 measure.

Both hypotheses (a, b) were evaluated with unigram, bigram, and trigram. The unigram method builds an assumption that each word occurs autonomously. It helps to measure the significance of the particular word. On the other hand, bigram and trigram consider local context. When DBLP and CiteSeer dataset was analyzed, it was observed keyword remains meaningful up to three words. Therefore, it was decided to evaluate metadata with unigram, bigram, and trigram.
The way different n-grams were created is demonstrated in *Figure 8.2*. The paper title "Web Evolution: From Read Web to Semantic Web." is used as an example for illustration. After employing the same method with rest of metadata lexical matching was performed with expert’s tweets.

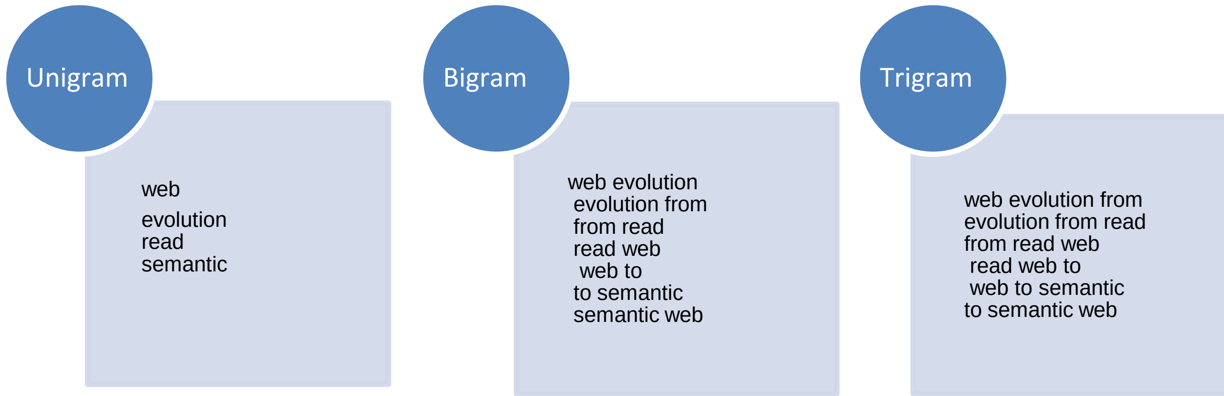

Figure 8.1 Example of Unigram, Bigram & Trigram

Each approach provided ranked tweets according to metadata (title, category, keyword, and author) using n-grams. Subsequently ranked tweets provided by each algorithm were compared with gold set in order to discover precision and recall for each metadata.

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

Equation 8:1

$$Recall = \frac{TP}{TP + FN}$$

Equation 8:2

$$F_1\,Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Equation 8:3

F_1 Score is also used, as it is defined as a harmonic mean of precision and recall. It provides a good trade-off between precision and recall metrics [Guo et al. 2018].

In the next section hypotheses (a) which individual metadata (title, category, keyword, and author) outperforms in terms of F_1 measure will be discussed.

8.1 Solo metadata evaluation using n-grams

In this section metadata title, author, category, keyword were independently evaluated in order to discover their effectiveness and contribution for twitter-based recommendation. Each metadata was evaluated with n-grams using three different techniques. As a result, 42 different evaluations were performed for hypotheses (a). In the following subsection solo metadata evaluation with three different approaches will be discussed.

8.1.1 Solo metadata evaluation with lexical

This section discusses on solo metadata evaluation with a lexical approach. In this evaluation *title* with unigram was found to be the best performing metadata. Its precision is 0.37, recall 0.88, and F_1 score as 0.53 whereas author was worst performing with precision 0.2, recall 0.10, and F_1 score as 0.4. The *title* has also performed best with bigram and trigram as well. When *title* unigram was created, the generic word ratio was higher than other metadata unigram. Co-occurrence of such generic unigram increased the possibility of finding the word in expert's tweets. Consequently, the *title* recalls also performed better than other metadata.

Whereas inversely speaking precision for *title*, *keyword*, and *category* is low. This is because domain experts provided tweets related to metadata of each paper. Some words in *title*, *keyword*, and *category* were overlapping. Therefore when lexical matching was performed for particular metadata related tweets, tweets related to other metadata were also retrieved. As a result *title*, *keyword* and *category* precision have been affected.

Let's take an example of *title to* illustrate precision. In *Figure 8.4* two different domain experts provided two separate papers. However some words are common in both research papers' title. If the tweets are fetched related to the first paper *title*, some tweets related to the second paper *title* will also be retrieved. When this result is compared with the gold set, it treats the tweets related to the second paper title as irrelevant. As a result precision for the 1st paper title will go down.

	Paper Title	Domain Expert
1	A Survey of Safety Analysis Techniques for Safety Critica Systems	Aftab Ali Haider, Aamer Nadeem
2	A Survey of Fault Tolerant CORBA Systems	Muhammad Fahad, Aamer Nadeem, Michael R. Lyu

Figure 8.2 Example for Title precision

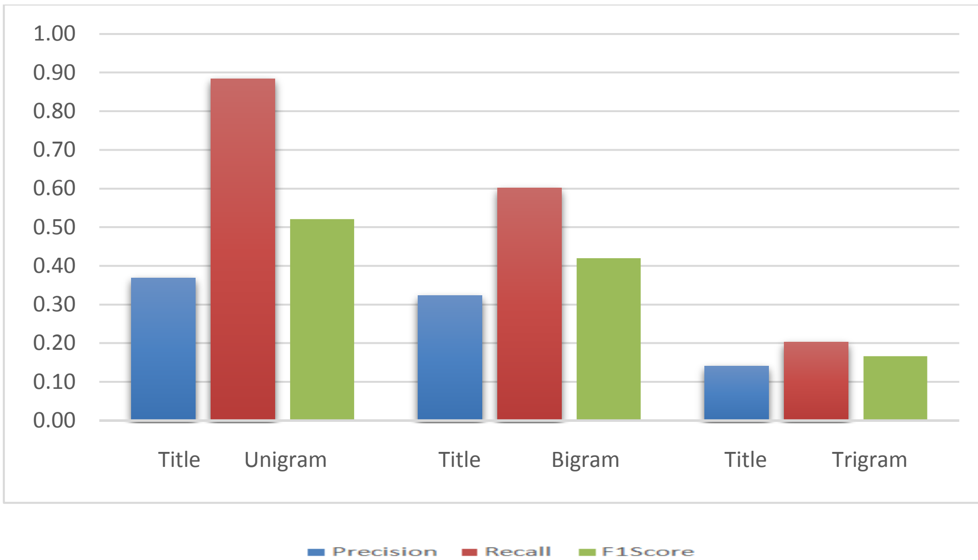

Figure 8.3 Title metadata evaluation with lexical using N-grams

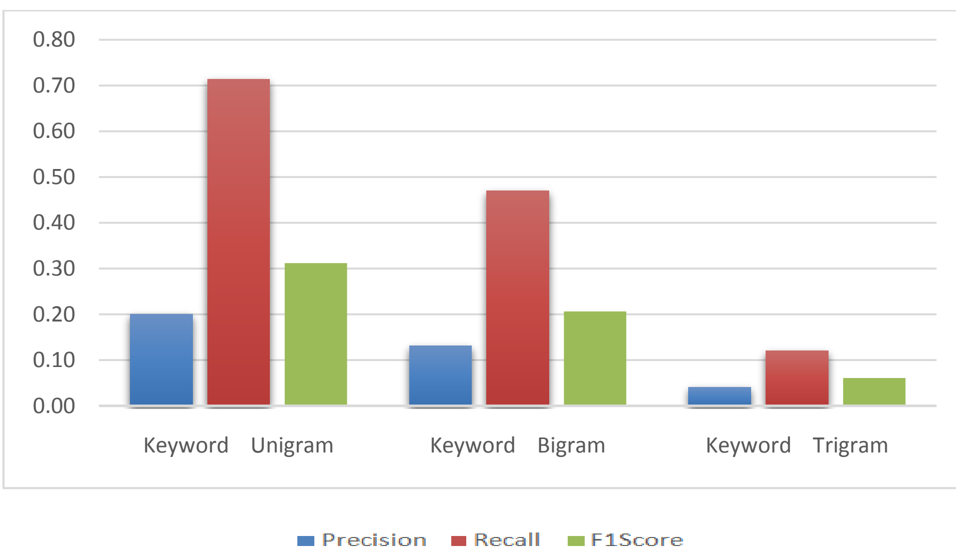

Figure 8.4 Keyword metadata evaluation with lexical using N-grams

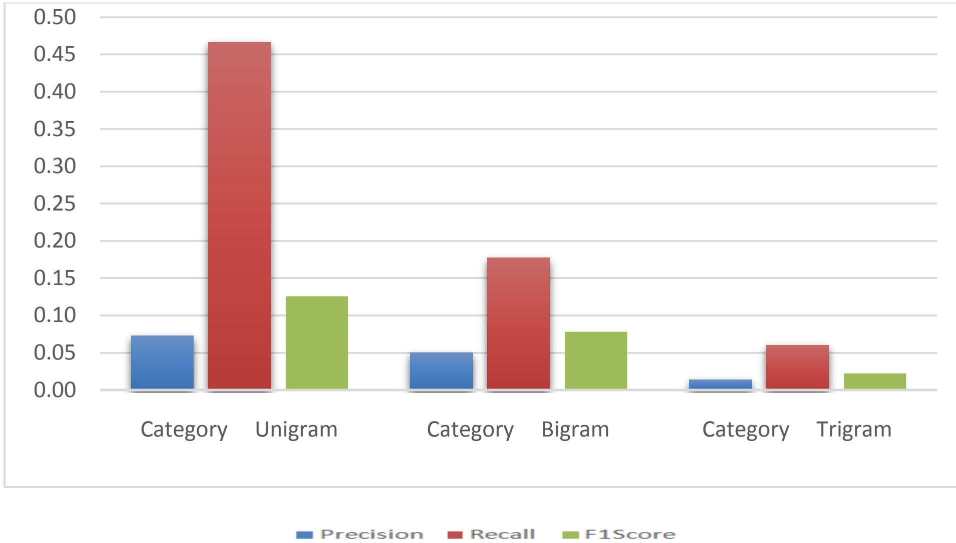

Figure 8.5 Category metadata evaluation with lexical using N-grams

It was observe with author metadata that author name ambiguity leads to irrelevant results. There are various formats of research papers, and author name has not been used uniformly. People sometimes use their full names, for example, *"Atta ur Rehman Khan*," and other time abbreviates, for example, *A.R. Khan*. In some cases, initials refer to the same name, but in reality, they are two separate individuals for example, in dataset *"K. Latif"* refers to both *"Khalid Latif"* and "*Kamran Latif.*"

A variety of techniques are used in literature for name disambiguation such as classification or clustering. These techniques offer a thorough study or profile of the person. Therefore, it is essential to examine the metadata and the text in order to make an educated guess to identify author uniquely [Torvik et al. 2009].

However such information was missing due to the limitation of text in tweets. Therefore, the absence of such information in tweets makes author name disambiguation process quite challenging.

8.1.2 Solo metadata evaluation with cosine

Each solo metadata was evaluated with unigram, bigram, and trigram. In this section, 12 different evaluations were performed. The content similarity is calculated using the vector space model by applying the TF-IDF weighting mechanism.

This approach required some hefty processing. This is due to the fact we got 220 papers and 2957 tweets. In total there are 220 x 2957 = 650,540 similarity checks for each metadata (Title, keyword, category, and author). On a standard machine, *title* metadata took 5 hours. In order to demonstrate how precision and recall were calculated let us take an example of metadata *title*.

a) Calculate the similarity score for PaperTitle_220 concerning all tweets (2957 tweets)

b) Remove the results with similarity score zeros (left with 466 tweets)

c) Sort the results in descending order.

d) Select top 15% results based on cosine similarity score. (Got top 69 tweets)

e) Calculate precision = relevant tweets /tweets retrieved

Precision = 7/69 = 0.1014

f) Calculate recall = relevant tweets /total relevant tweets (8)

g) Recall= 7/8 = 0.87

Figure 8.6 Precision and Recall calculation

A relevant tweet here refers to the tweets provided by domain experts concerning all metadata. A dataset of such tweets is referred as gold set.

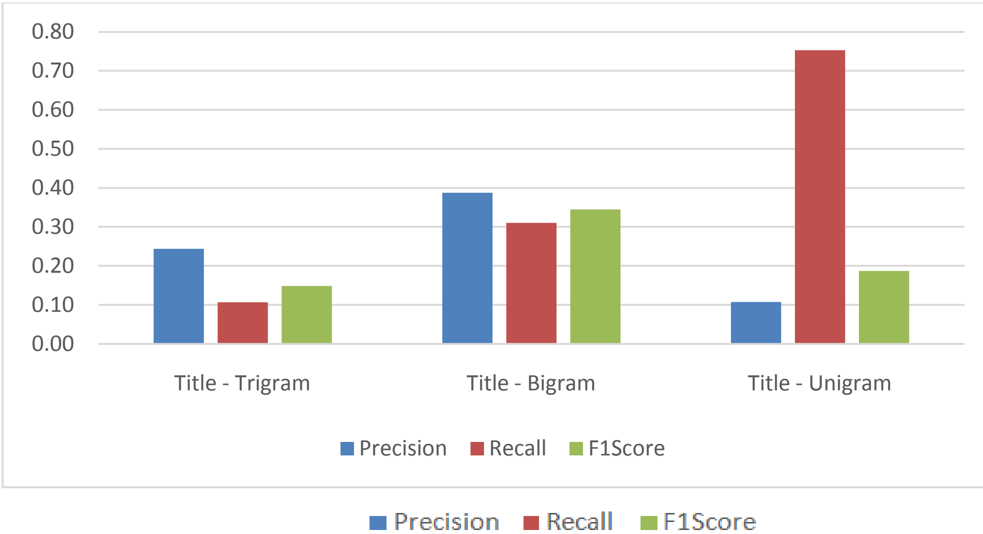

Figure 8.7 *Title* metadata evaluation with cosine using N-grams

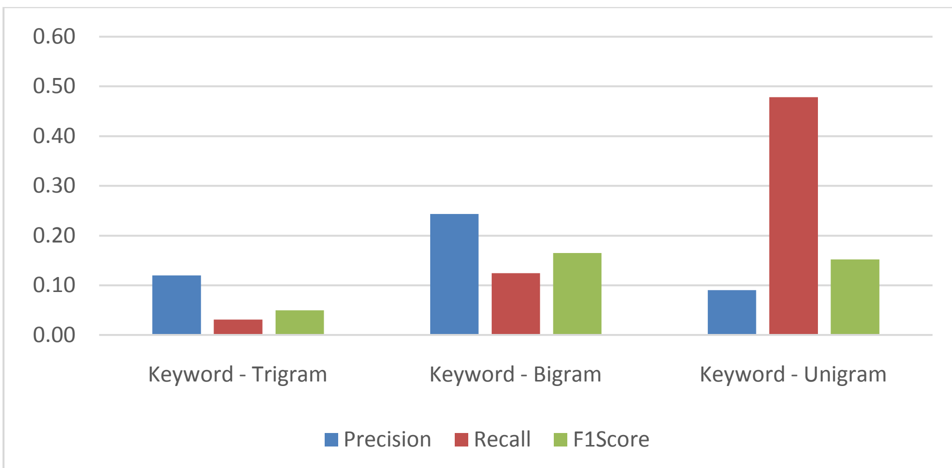

Figure 8.8 *Keyword* metadata evaluation with cosine using N-grams

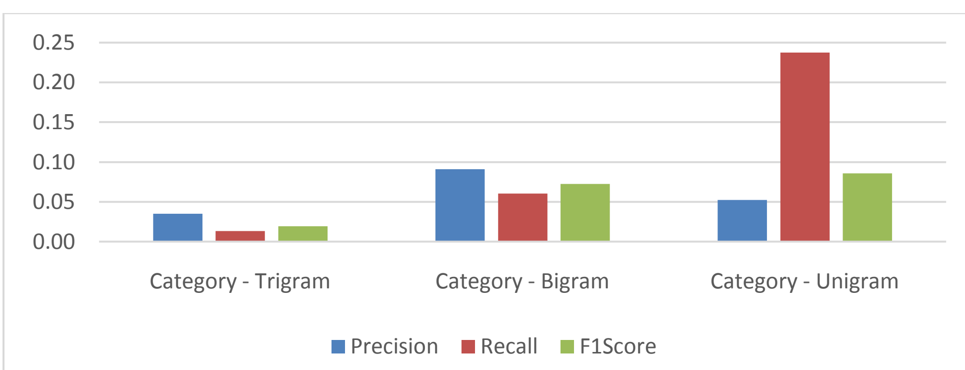

Figure 8.9 *Category* metadata evaluation with cosine using N-grams

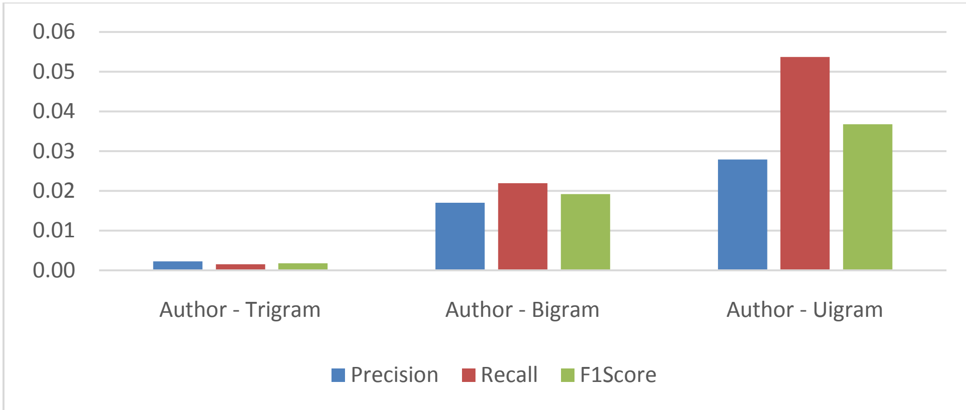

Figure 8.10 *Author* metadata evaluation with cosine using N-grams

F_1 score plays an important role to measure the accuracy of the technique. It takes both precision and recall into consideration when computing the score. Based on F_1 score *title* with bigram performed best with an F_1 score as 0.35. However, the *title* with unigram precision is 0.10, and recall 0.75. This is due to the fact different paper titles may share some common words as a result irrelevant tweets are also assigned cosine similarity score. As a result, the precision went down, and recall went up. The second best-performing metadata was the *keyword* with recall 0.12, precision 0.24, and F_1 Score 0.17.

On the hand, *author* with trigram was worst performing with F_1 Score as 0.2. When *author* unigram is matched with tweets dataset, it was observe the ratio of author names to appear in tweets is relatively low. Hence it had an adverse effect on an author's precision and recall. The next section provides discussion on solo metadata evaluation with an extended vocabulary.

8.1.3 Solo metadata evaluation with an extended vocabulary

In this section, solo metadata was evaluated with an extended vocabulary. N-grams were extended using Wordnet and Growbag databases. Wordnet is an English language lexical database consist of 115 424 synsets, which are set of synonyms. It also maintains the relations among synonym sets and their members. Growbag, on the other hand, is a classification system used in the realm of Computer Science. It is based on DBLP dataset of co-occurrence terms. Once the paper was tokenized, each term was extended using the Wordnet and Growbag databases. The ratio of each term extension concerning N-gram can be viewed in *table 8:1* and *table 8:2*.

Our evaluation result for solo metadata with extended vocabulary shows *title* unigram performed best in Growbag. Title recall was 0.48, and precision 0.1, whereas in Wordnet *title* unigram recall was 0.27 and precision remained zero.

Figure 8.11 *Title* with Wordnet

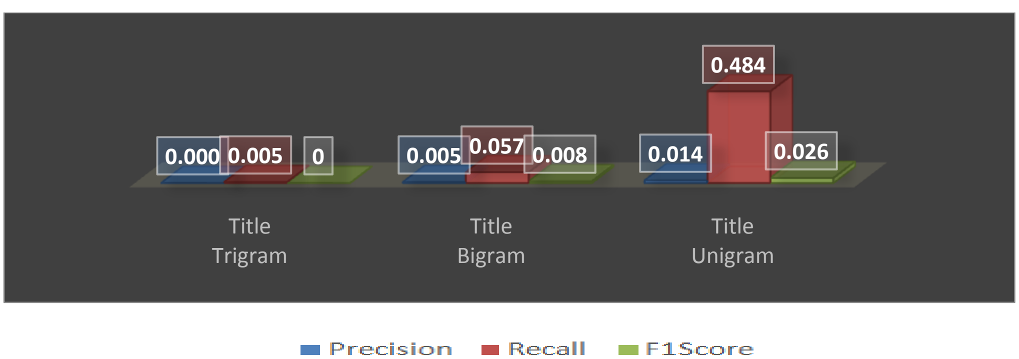

Figure 8.12 *Title* with Growbag

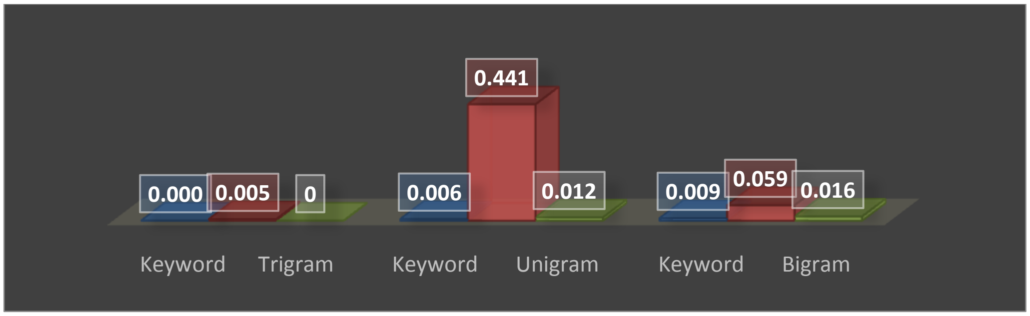

Figure 8.13 *Keyword* with Growbag

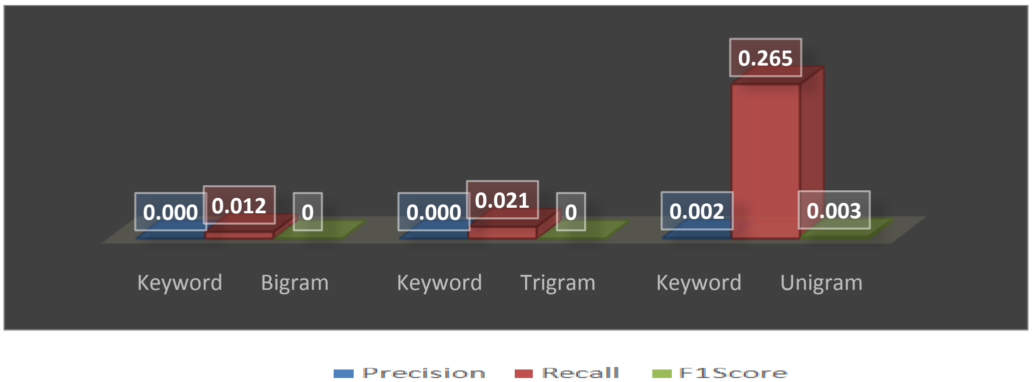

Figure 8.14 *Keyword* with Wordnet

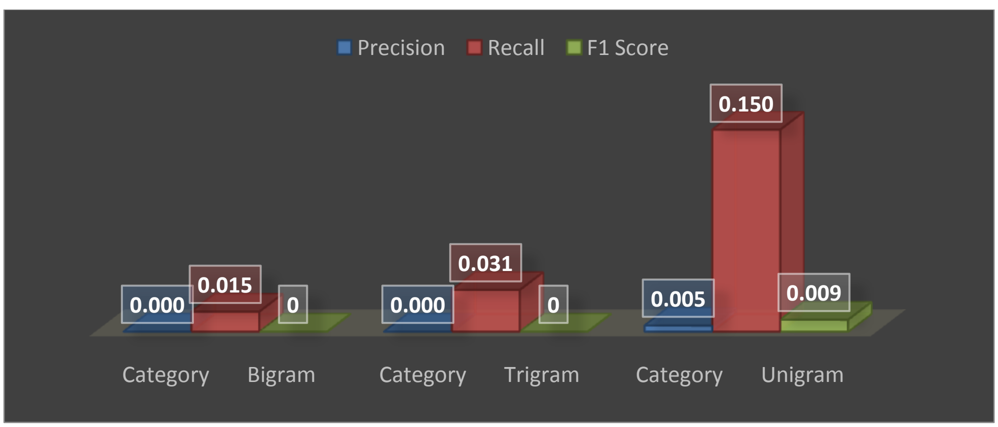

Figure 8.15 *Category* with Wordnet

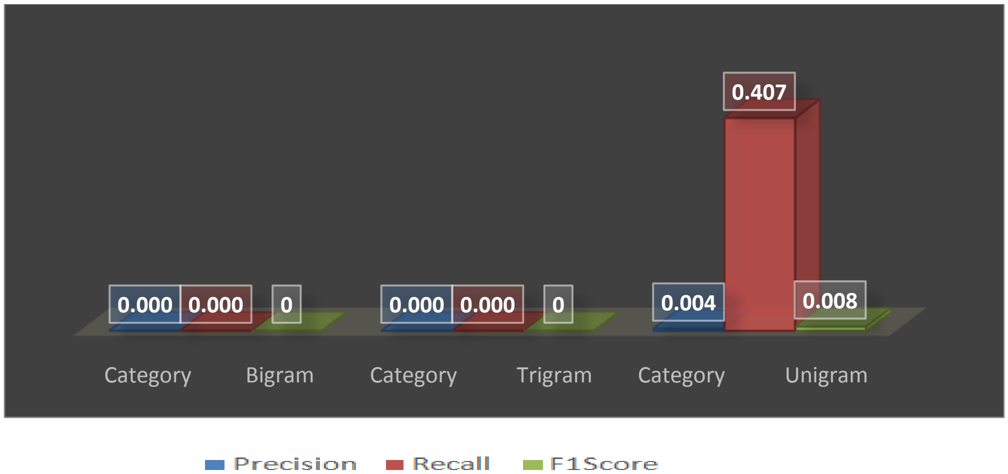

Figure 8.16 *Category* with Growbag

Evaluation result indicates unigram solo metadata performed better with Growbag as compare to Wordnet. This is due to fact Growbag provides computer science domain specific co-occurrence dataset whereas Wordnet provides comprehensive English language synonyms dataset. When unigram solo metadata were queried for extension in Growbag more comprehensive results were retrieved than Wordnet as illustrated in *table 8:1*.

However, on the hand, it also extracted irrelevant information which affected the precision to decline.

Table 8:1 Wordnet Extension

Wordnet Extension	Unigram Presence (%)	Bigram Presence (%)	Trigram Presence (%)
Title	28.05	0	0
Keyword	24.62	0	0
Category	8.35	0	0

Table 8:2 Growbag Extension

Growbag Extension	Unigram Presence (%)	Bigram Presence (%)	Trigram Presence (%)
Title	254.30	1.1	0.02
keyword	294.36	0.95	0.064
Category	122.41	0.05	0

Overall unigram solo metadata performed best. However, solo metadata with bigram and trigram did not do well with Growbag and Wordnet. This is due to fact Wordnet, and Growbag datasets did not provide a valuable extension for bigram and trigram. As *Table 8:2* indicates the presence of bigram and trigram extension were significantly low in Wordnet and Growbag datasets. Bigram and Trigram are useful to extract phrases and sub-phrases in a sentence. However, information in Wordnet and Growbag dataset is not represented in this manner.

8.1.4 Overall solo metadata evaluation

Overall solo metadata evaluation can be concluded as Title and Keyword has independent significance for twitter based recommendation. On the other hand, *Author* metadata demonstrates the least significance for a quality recommendation.

Table 8:3 Overall solo metadata evaluation

Best in each Approach	Best Metadata	Best N-gram	Precision	Recall	F1 score
Lexical	Title	Unigram	0.37	0.88	0.52
Cosine	Title	Bigram	0.39	0.31	0.34
Extended Vocabulary with Growbag	Title	Unigram	0.014	0.484	0.026

8.2 Hybrid metadata evaluation using n-grams

This section of evaluation provides a discussion on how Metadata should be combined for the best quality recommendation. The evaluation discussion will be in three parts binary, trio, and all metadata evaluation. In this section, 66 different evaluations were performed. In the following subsection binary metadata evaluation with three different approaches will be discussed.

8.2.1 Binary metadata evaluation with lexical

In this section, binary metadata were evaluated with a lexical approach using n-grams. Binary metadata possible pair combination is illustrated in *Table 8:4*.

Table 8:4 Binary metadata combination

1	*Title & Category*
2	*Title & Keyword*
3	*Title & Author*
4	*Keyword & Category*
5	*Author & Category*
6	*Author & Keyword*

All of the above pairs were evaluated with unigram, bigram and trigram and altogether 18 different evaluations were performed. Binary metadata combination title and keyword performed best with bigram. Their recall was 0.63 precision 0.44 and F1 score 0.52. The title and keyword unigram recall were 0.75 whereas the precision was 0.16 and F1 score 0.26. This is well below the bigram precision and F_1 score. This illustrates many irrelevant records were retrieved along with the relevant record.

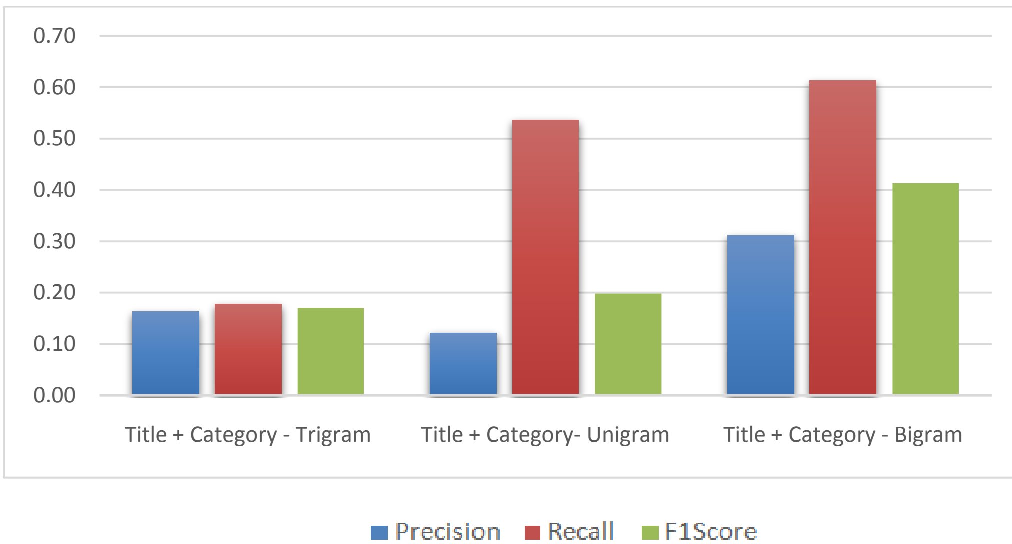

Figure 8.17 Title & Category evaluation with lexical using N-grams

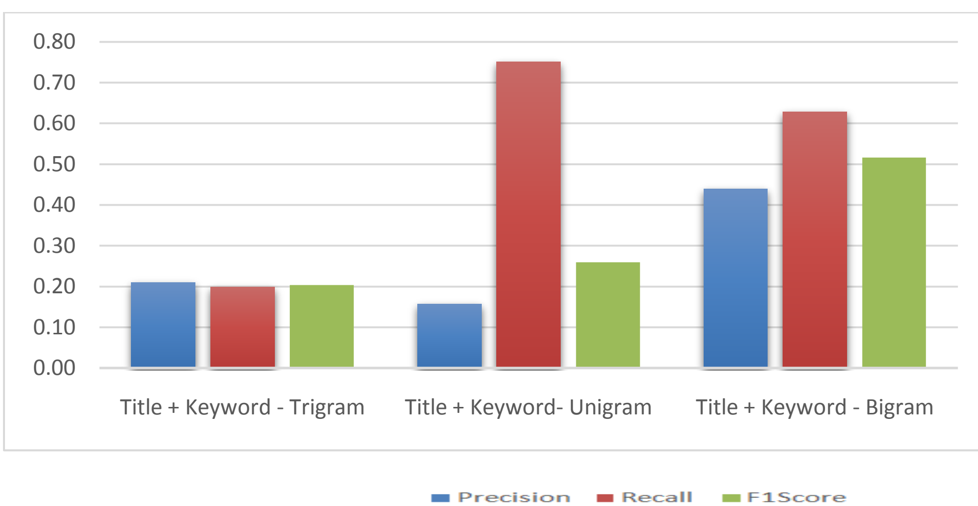

Figure 8.18 Title & Keyword evaluation with lexical using N-grams

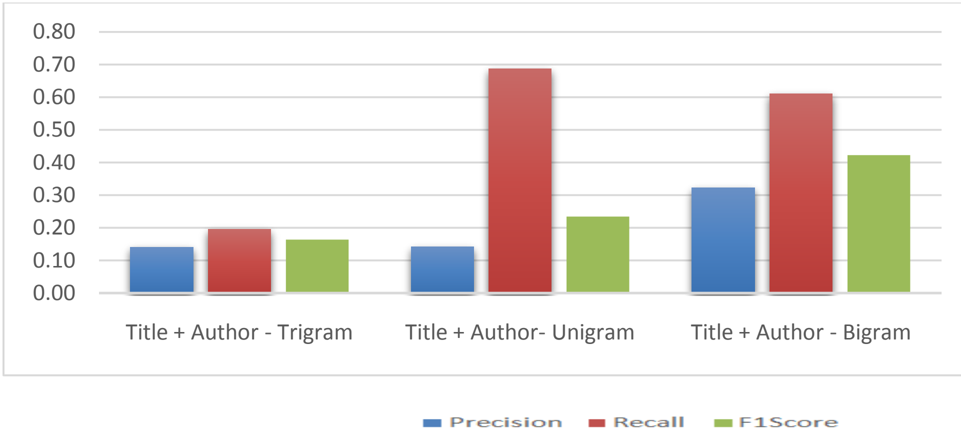

Figure 8.19 Title & Author evaluation with lexical using N-grams

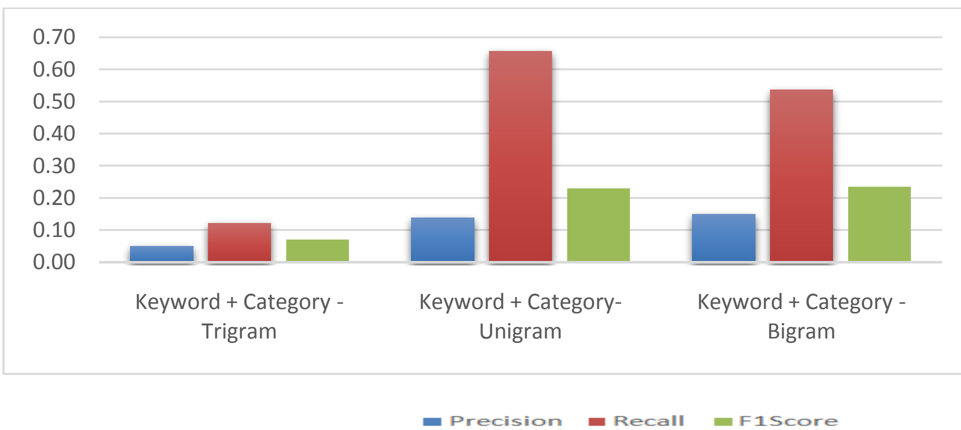

Figure 8.20 Keyword & Category evaluation with lexical using N-grams

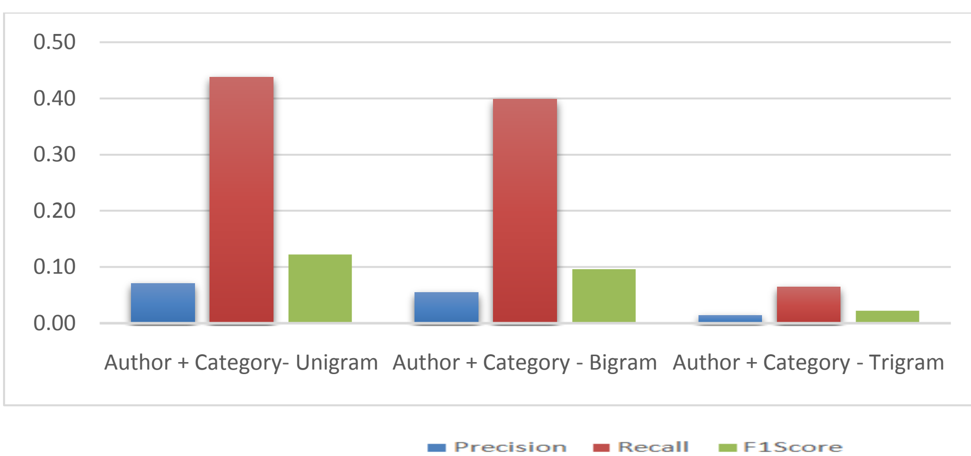

Figure 8.21 Author & Category evaluation with lexical using N-grams

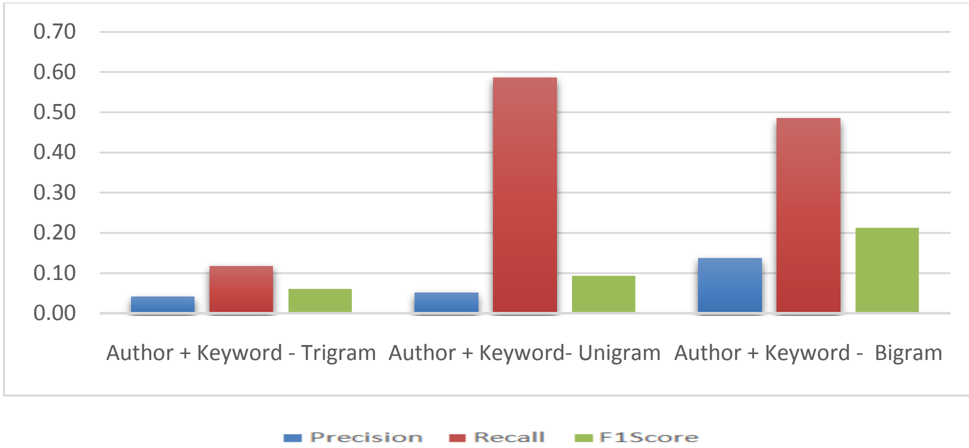

Figure 8.22 Author & Keyword evaluation with lexical using N-grams

It was observed the title and keyword metadata share some generic words. The ratio of such generic words is further increased when words are tokenized into unigram. Let us take an example of a paper titled "Web Evolution: From Read Web to Semantic Web." One of the unigrams in this title is "web." Now, this is a broad term that can be web technology, web interface, semantic web and so forth. On the other hand, when bigram is used it narrows it down to more meaningful information. In the above scenario bigram *"semantic web"* is more precise than unigram *"web."* It was this reason unigram fetched more records however it fetched irrelevant records as well. Our decisive factor is F1-Score as it provides a good trade-off between precision and recall metrics. F1-Score of title and keyword indicates it outperformed from the rest of the results.

It was observed both metadata did complement each other. Both title and keyword performed best in solo features evaluation as well. However, keyword are often derived from the title, so due to this partial repetition, both features combination reduced its result.

8.2.2 Binary metadata evaluation with cosine

Binary metadata combinations in Table 8:4 were evaluated with cosine approach. All of the pairs were evaluated with unigram, bigram and trigram and altogether 18 different evaluations were performed.

The evaluation results indicate binary metadata combination title & keyword performed best with bigram. The title & keyword precision was 0.43 recall 0.16, and the F_1 score was 0.24.

The title & keyword metadata also performed best in solo metadata features evaluation. This is an indication if these individual features are combined together it will complement each other. On the other hand, author & category were the worst performing combination with trigram.

However title & keyword precision and recall are relatively low with cosine approach. This is because both approaches use a matching process; however, the matching process itself is different. The lexical approach applies direct match where the value is either true or false, whereas cosine approach calculates the result based on similarity score. Each tweet is assigned a score based on its relevance to the user query.

Consequently, numbers of irrelevant records retrieved with cosine approach are higher than the lexical approach. As a result precision and recall have decreased with cosine approach.

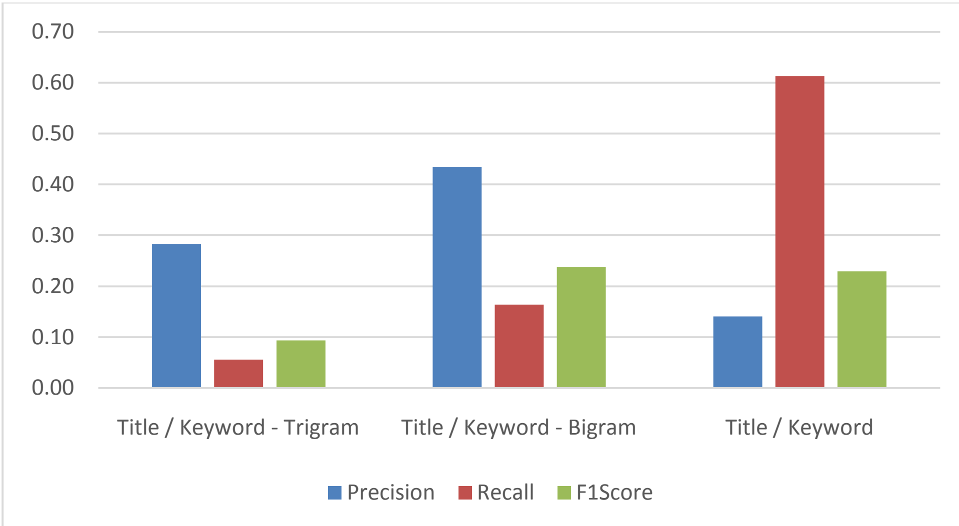

Figure 8.23 Title & Keyword metadata evaluation with cosine using N-grams.

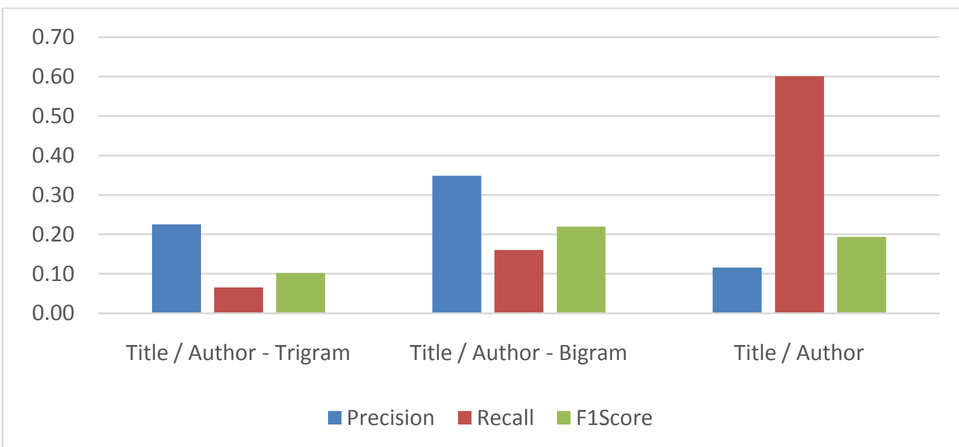

Figure 8.24 Title& Author metadata evaluation with cosine using N-grams.

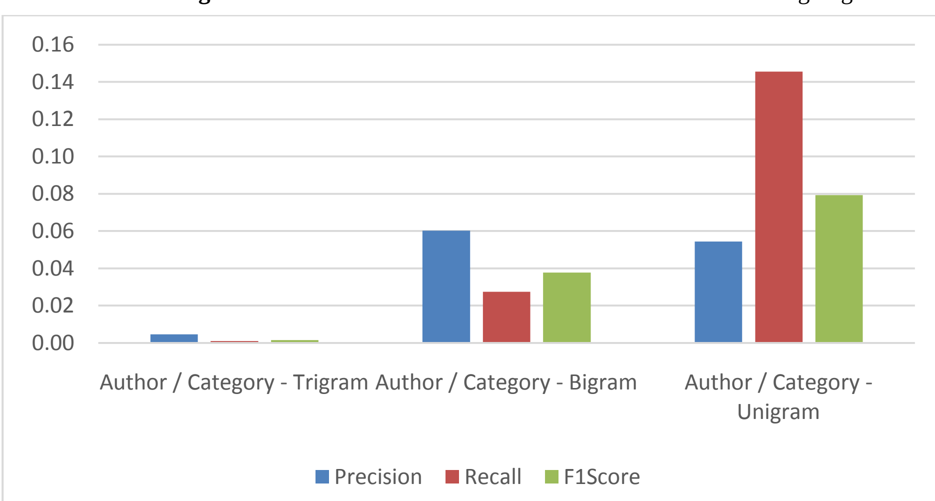

Figure 8.25 Author & Category metadata evaluation with cosine using N-grams.

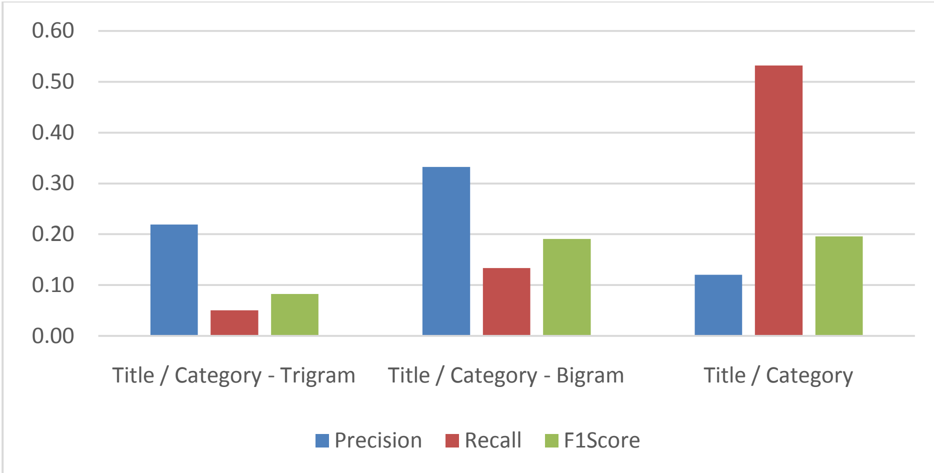

Figure 8.26 Title & Category evaluation with cosine using N-grams.

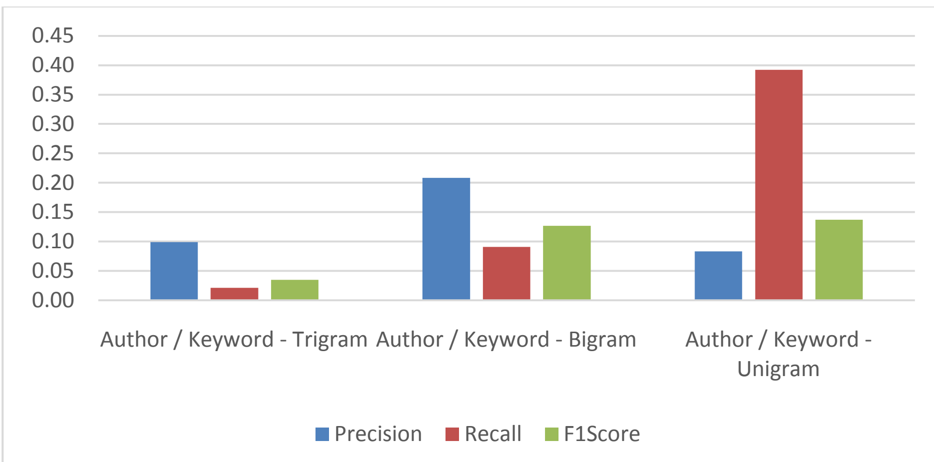

Figure 8.27 Author& Keyword metadata evaluation with cosine using N-grams.

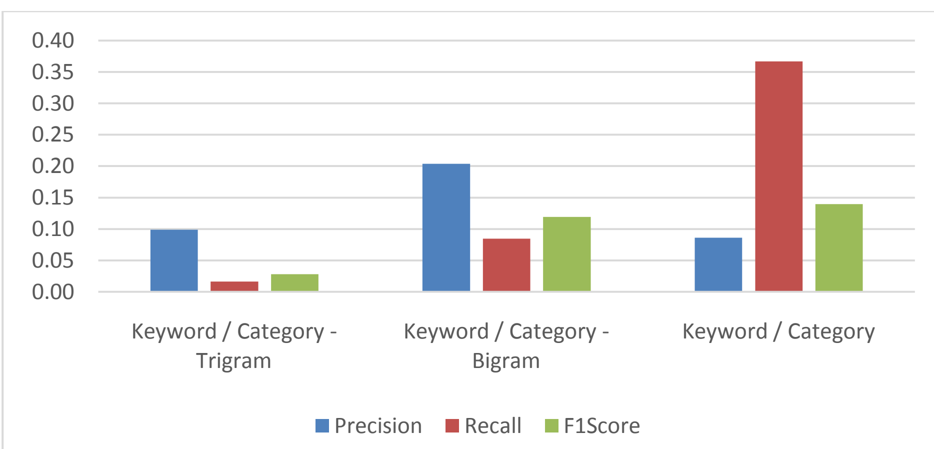

Figure 8.28 Keyword& Category metadata evaluation with cosine using N-grams.

8.2.3 Binary hybrid metadata evaluation overview

The binary metadata combination of title & keyword with bigram performed best in both lexical, and cosine approaches. The precision in both approaches is almost similar however the recall for lexical is relatively higher than cosine. This is an indication the lexical approach retrieved more relevant record than cosine.

In the cosine approach, each tweet is assigned a score based on its similarity with the query. As a result, numbers of irrelevant records are relatively higher than the lexical approach. As the lexical approach will only return records if there is a match, i.e., a match is found. Therefore recall for the lexical approach is comparatively higher than cosine.

The overall binary metadata evaluation results indicate Title & Keyword with bigram combination will provide the most significant contributions for a quality recommendation. On the other hand, author & keyword combination will be the least significant for a quality recommendation.

Table 8:5 Binary Metadata overview

Approach	Best Metadata	Precision	Recall	F_1 Score
Lexical	Title + Keyword Bigram	0.44	0.63	0.52
Cosine	Title + Keyword Bigram	0.43	0.16	0.24

8.3 Trio hybrid metadata evaluation using n-grams

In this section trio, hybrid metadata evaluation will be discussed with lexical and cosine approach. Both approaches results are provided in the subsequent subsections.

8.3.1 Trio hybrid metadata evaluation with lexical.

In this section trio, metadata were evaluated with a lexical approach using n-grams. The trio metadata possible combinations are illustrated in *Table 8:6*.

Table 8:6 Trio Metadata Combination

1	Title +Author + Keyword
2	Title + Author+ Category
3	Title + Keyword+ Category
4	Author + Keyword+ Category

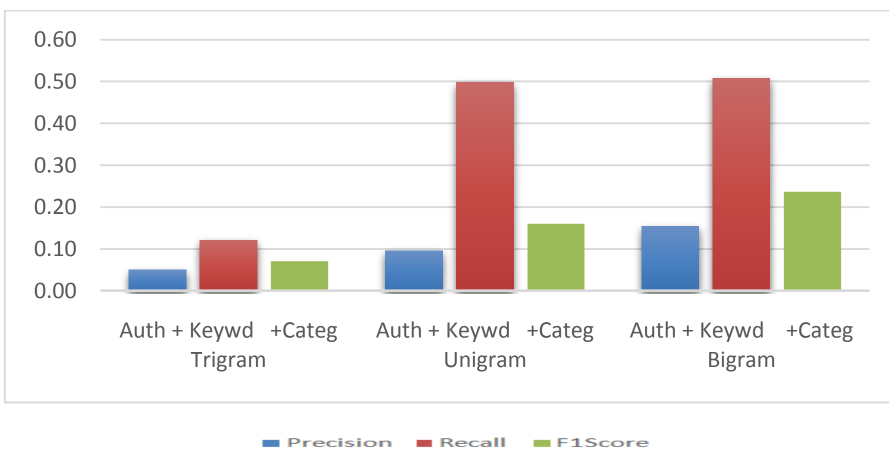

Figure 8.29 Lexical evaluation of Author, Keyword & Category using N-grams

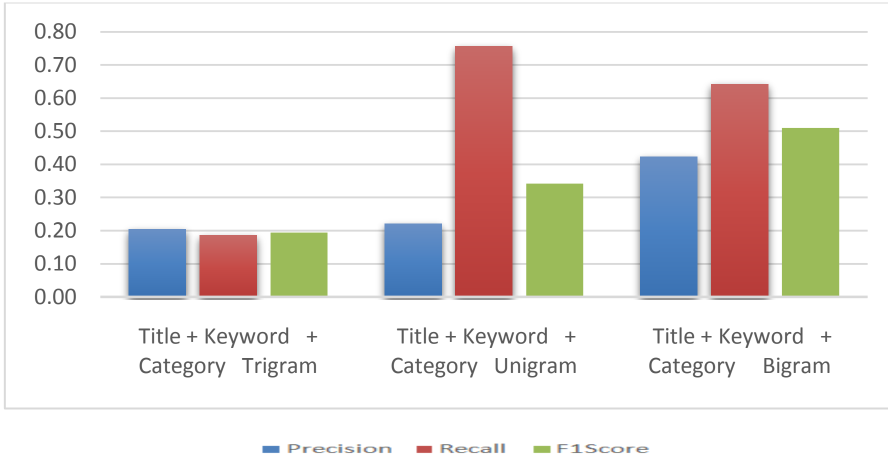

Figure 8.30 Lexical evaluation of Title, Keyword& Category using N-grams

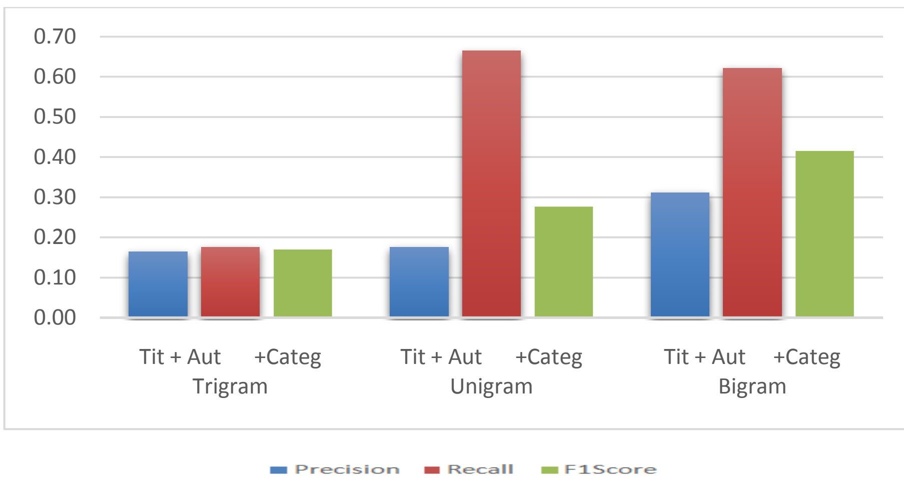

Figure 8.31 Lexical evaluation of Title, Author & Category using N-grams

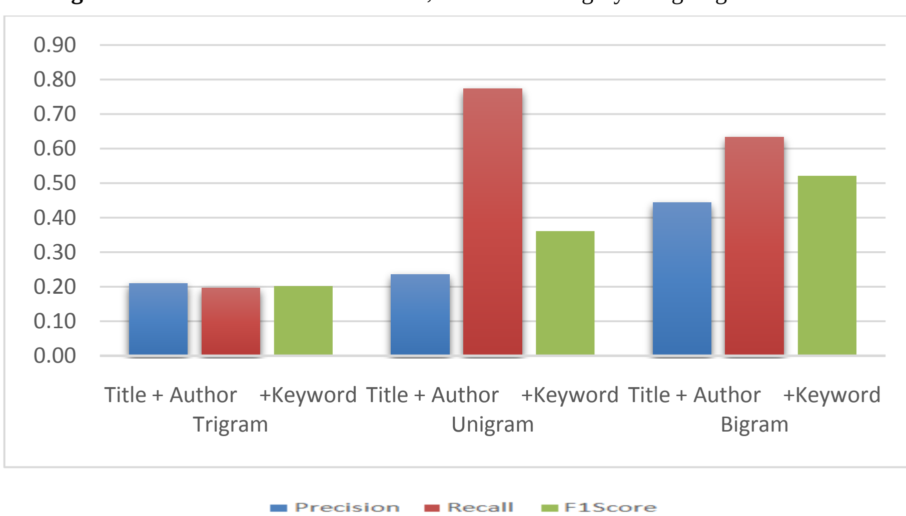

Figure 8.32 Lexical evaluation of Title, Author & keyword using N-grams

All of the above combinations were evaluated with unigram, bigram, and trigram. Altogether 12 different evaluations were carried out. The evaluation results indicate *Title+Author+keyword* combination performed best with bigram. All three of the metadata individual recall was also good; however, on the other hand, the precision has gone down. This is due to the fact a high ration irrelevant records were fetched alongside relevant records for each feature. As a result when irrelevant records of all three features were combined precision went down.

8.3.2 Trio metadata evaluation with cosine.

Trio metadata combinations in Table 8:6 were evaluated with cosine approach. All of the combinations were evaluated with unigram, bigram, and trigram. Altogether 12 different evaluations were performed. The evaluation results indicate trio metadata combination *Title + Keyword+ Category* performed best with unigram. The precision was 0.15, recall 0.53; F_1 score was 0.23.

In this approach unigram model performed best in trio metadata evaluation whereas trigram remained the worst with F_1 as low as 0.3. Unigram treats each word as it appears autonomously. As a result, the likelihood of a word sequence becomes the product of the likelihood of the individual words.

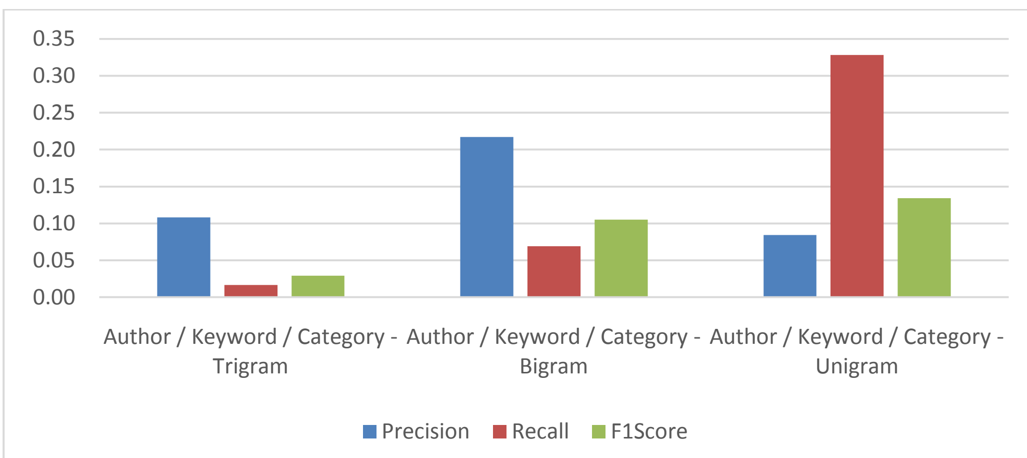

Figure 8.33 Cosine evaluation of Author, Keyword& Category using N-grams

With bigram and trigram, the likelihood of a new word relies on the likelihood of the earlier words.

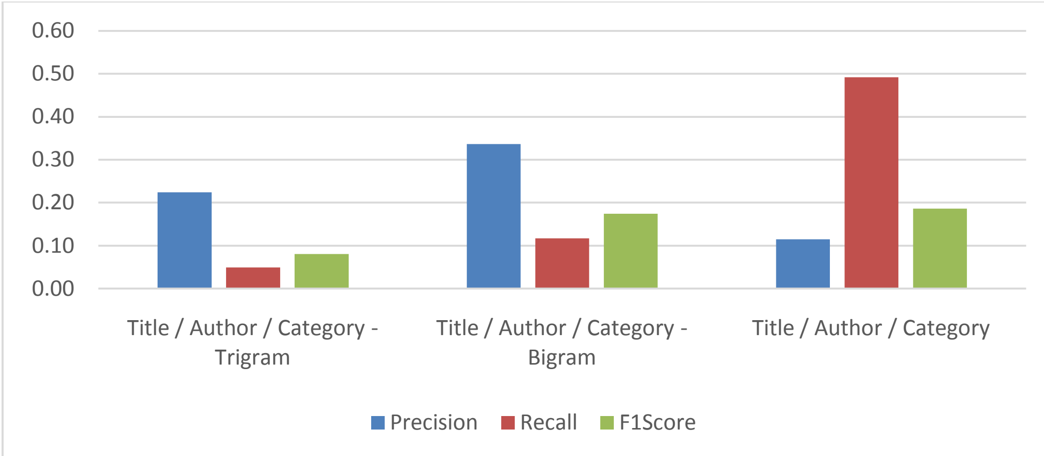

Figure 8.34 Cosine evaluation of Title, Author & Category using N-grams

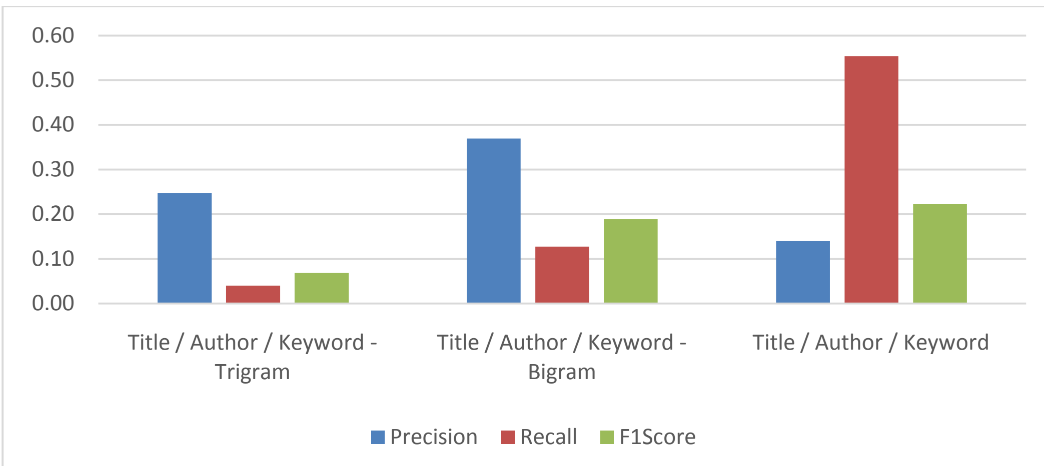

Figure 8.35 Cosine evaluation of Title, Author & Keyword using N-grams

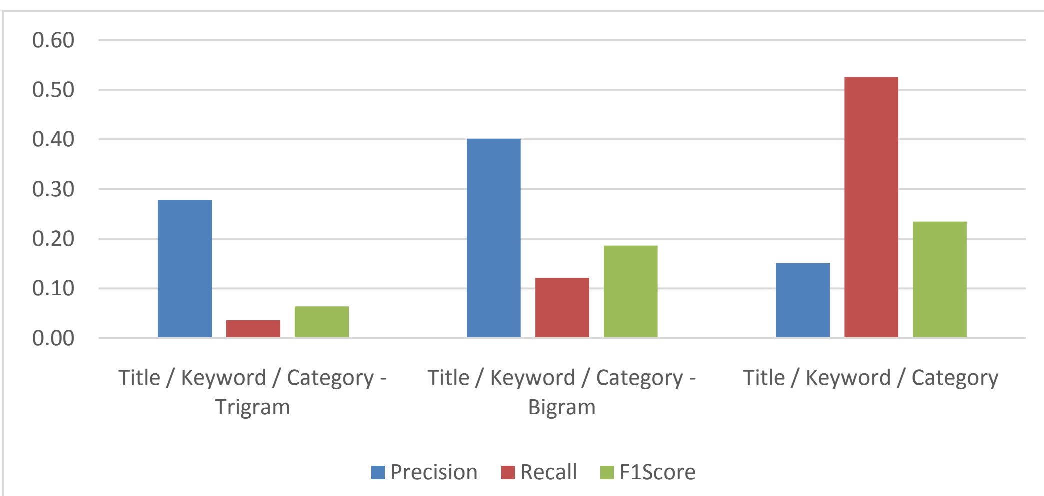

Figure 8.36 Cosine evaluation of Title, Keyword& Category using N-grams

8.3.3 Trio metadata evaluation overview

In trio metadata evaluation four different combinations were evaluated with unigram, bigram, and trigram. With lexical and cosine approached 24 different evaluations were performed. *In lexical approach Title + Keyword +Category performed best with bigram whereas in cosine approach Title + Keyword+ Category performed best with unigram. Overall in both* approaches *Title + Keyword, +Category performed best with bigram. The precision was 0.42, recall 0.64 and F_1 Score 0.51.*

Table 8:7 Trio Metadata Combination

Approach	Best Metadata	Precision	Recall	F_1 Score
Lexical	Title + Keyword+ Category Bigram	0.42	0.64	0.51
Cosine	Title + Keyword+ Category Unigram	0.15	0.53	0.23

8.4 All metadata evaluation using n-grams

In this section, all metadata evaluation using n-grams will be discussed with lexical and cosine approach. Each approach result is provided in the subsequent subsections.

8.4.1 All metadata evaluation with lexical

In this section, all 15 different metadata combinations were evaluated with a lexical approach. All of the combinations were evaluated with unigram, bigram and trigram and altogether three different evaluations were performed. All metadata combination with bigram performed best with precision 0.43, recall 0.65 and F_1 Score 0.51.

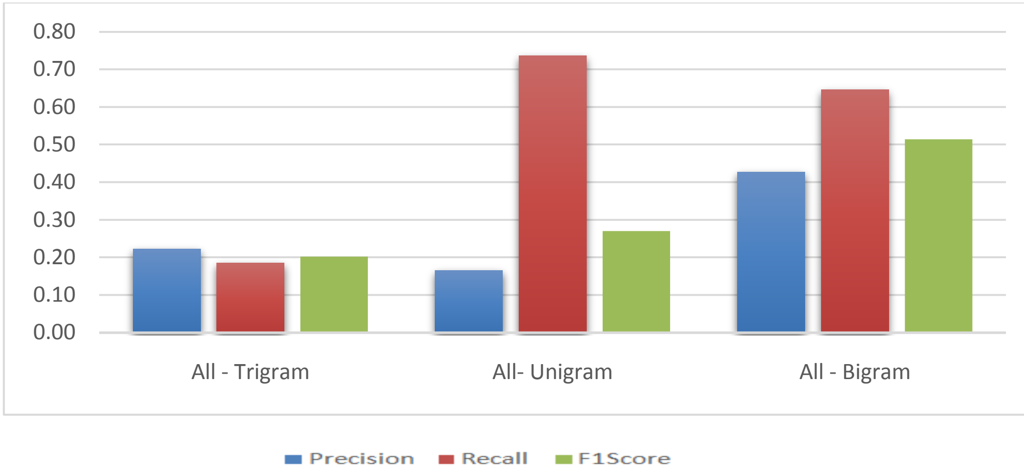

Figure 8.37 All metadata evaluation with lexical

8.4.2 *All metadata evaluation with cosine*

All 15 different metadata combinations were evaluated with a lexical approach. All metadata evaluation was carried out with unigram, bigram, and trigram. In this section all metadata combination with unigram performed best with precision 0.14, recall 0.50 and F_1 Score 0.21.

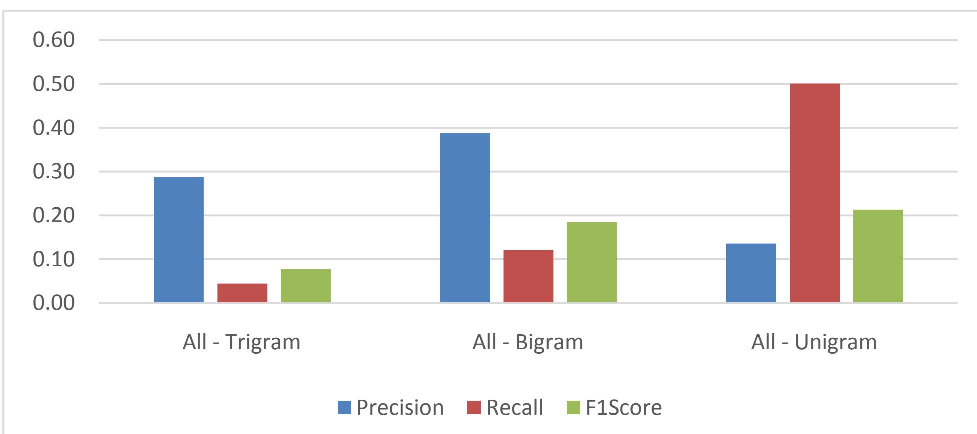

Figure 8.38 All metadata evaluation with cosine

8.4.3 *All metadata evaluation overview*

With all the metadata combination bigram with lexical approach performed best.

The precision was 0.43, recall 0.65 and F1 Score 0.51.

CHAPTER 9: Conclusion and future work

The dissertation investigated the various aspects of E-Learning recommendation. The work demonstrates its significance in the field of technology-enhanced learning. The thesis contributions can be structured into the following areas:

1. Contemporary E-Learning systems & recommendation approaches review.
2. A need for specialized E-Learning system.
3. Domain-specific recourse recommendation.
4. Twitter-based E-Learning recommendation.

9.1 Contemporary E-Learning systems & recommendation approaches review

Firstly, the thesis investigated the functionalities of contemporary E-Learning systems. It was essential to have an imperative analysis of the existing E-Learning systems in the context of a new era of knowledge management. Conventional E-Learning systems are viewed as course management tools rather than facilitating the individual learning need. As a result research community questions the usefulness of the E-Learning system.

Over 200 recent and classical papers have been critically reviewed about E-Learning recommender systems and Twitter-based recommender systems in various domains. Literature study reveals no existing E-Learning recommender systems are utilizing the twitter.

It was highlighted recommendation techniques in E-Learning domain are comparatively new. Moreover, many of these recommendation techniques are derived from commercial recommender systems.

The thesis highlights the strength and limitations of prominent approaches and presented challenging tasks which will be useful for the E-Learning research community to focus for future research. Discussion in this section provided the stepping stone for the research.

9.2 Need for specialized E-Learning system

The thesis establishes the need for constructing a specialized (domain specific) E-Learning system. Such systems can help learners in a particular domain and assist them according to their particular needs, context, profiles, histories, and collaborations. A specialized E-Learning system will be centered towards active and authentic learning rather than providing a universal fit for all solution.

The research community has been continually developing new and imperative methods for active and authentic learning. However such efforts of the scientific community have been on a restricted level. These efforts have proven to be quite inspiring features which can be incorporated towards specialized (domain specific) E-Learning system.

A specialized E-Learning system would support global information to be available in the local social context of the learner. It should enable the learner the elasticity to discover, organize, share information in a locally meaningful fashion which is globally accessible.

9.3 Domain specific recourse recommendation

The thesis also explored how social bookmarking and semantic method can be utilized to provide domain-specific recourse recommendation. Thesis described remedies for the problems which arise from the traditional E-Learning system. Such as static learning resources provided to all the learners though learners' individual needs may significantly be different from each other. Thesis contributed by devising two distinct techniques in order to proactively discover the most relevant knowledge resources from the social Web.

9.4 Twitter-based E-Learning recommendation

Thesis discovers the potential of twitter-based E-Learning recommendation. It also establishes learner and learning resources are two pivotal entities to consider for twitter-based E-Learning recommendation. A significant contribution of the thesis has been the evaluation of the learner parameters and learning resource metadata evaluation.

Literature study reveals the parameters associated with the user for recommendation task are profile, context, and history. However, such parameters for the learner have not been applied for Twitter-based E-Learning recommendation. A user study was carried out to evaluate the effectiveness of these parameters for E-Learning recommendation. It was identified all of the above three parameters are equally important to consider while providing E-Learning recommendation from Twitter.

Similarly, for learning resource, the metadata associated with resource recommendation are title, author, category, keyword, content, and extended vocabulary (Synonyms, Grow bag).

The aim was to evaluate how the metadata of resource metadata should be combined. Similarly, which combination of metadata of resource metadata provides the best result for a twitter-based recommendation?

Extensive experimentation and evaluations were conducted to determine the effectiveness of resource metadata for twitter based E-Learning recommendation. Three distinct recommendation techniques were used for evaluation. Recourse metadata were evaluated individually as well as with different combination using unigram, bigram, and trigram. In total 102 different evaluations were conducted. Each technique recommended the relevant tweets based on recourse Metadata. The detailed evaluation reveals the effectiveness of the resource metadata for twitter based E-Learning recommendation.

Bibliography

Abdullah, N., & Ibrahim, R. (2015). Managing Information by Utilizing WordNet as the Database for Semantic Search Engine. International Journal of Software Engineering and Its Applications, 9(5), 193-204.

Acar, A., & Kimura, N. (2012). Twitter as a tool for language learning: the case of Japanese learners of English. Special Issue of International Journal of the Computer, the Internet and Management, 19, 1-14.

Adelsberger, H. H., Bick, M., & Pawlowski, J. M. (2000). The Essen learning model-a step towards a standard model of learning processes. In Proc. of ED-MEDIA 2000, World Conference on Educational Multimedia, Hypermedia & Telecommunications, Charlottesville, VA: AACE.

Arenas, M., Cuenca Grau, B., Evgeny, E., Marciuska, S., & Zheleznyakov, D. (2014). Towards semantic faceted search. In Proceedings of the companion publication of the 23rd international conference on World Wide Web companion (pp. 219-220). International World Wide Web Conferences Steering Committee.

Basu, C., Hirsh, H. and Cohen, W. (1998). Recommendation as classification: Using the social and content-based information in a recommendation. In *AAAI/IAAI* (pp. 714-720).

Beel, J., Gipp, B., Langer, S., & Breitinger, C. (2016). Paper recommender systems: a literature survey. International Journal on Digital Libraries, 17(4), 305-338.

Beldarrain, Y. (2006). Distance Education Trends: Integrating new technologies to foster student interaction and collaboration. Distance Education, 27, 139–153.

Bobadilla, J. E. S. U. S., Serradilla, F., & Hernando, A. (2009). Collaborative filtering adapted to recommender systems of E-Learning. Knowledge-Based Systems, 22(4), 261-265.

Bobadilla, J., Ortega, F., and Hernando, A. (2012). A collaborative filtering similarity measure based on singularities. *Information Processing & Management*, *48*(2), 204-217.

Bobadilla, J., Serradilla, F. and Hernando, A. (2009). Collaborative filtering adapted to recommender systems of E-Learning. *Knowledge-Based Systems*, *22*(4), 261-265.

Boyd, D. (2006). Friends, friendsters, and myspace top 8: Writing community into being on social network sites.

Broder, A. (2006). The Next generation Web Search: From Information Retrieval to Information Supply, IEEE talk, Stanford University, November 16, 2006

Brusilovsky, P., & Peylo, C. (2003). Adaptive and intelligent web-based educational systems. International Journal of Artificial Intelligence in Education (IJAIED), 13, 159-172.

Buckland, M. K., & Gey, F. C. (1994). The relationship between recall and precision. JASIS, 45(1), 12-19.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.

Chan, N. N., Roussanaly, A., & Boyer, A. (2014). Learning resource recommendation: An orchestration of Content-based filtering, word semantic similarity and page ranking. In European Conference on Technology Enhanced Learning (pp. 302-316). Springer, Cham.

Chandrasekaran, K., Gauch, S., Lakkaraju, P. and Luong, H. P. (2008). Concept-based document recommendations for Citeseer authors. In Adaptive Hypermedia and Adaptive Web-Based Systems (pp. 83-92). Springer Berlin Heidelberg.

Chen, C. M., Lee, H. M.and Chen, Y. H. (2005). Personalized E-Learning system using item response theory. Computers & Education, 44(3), 237-255.

Chen, J., Nairn, R., Nelson, L., Bernstein, M., & Chi, E. (2010). Short and tweet: experiments on recommending content from information streams. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1185-1194). ACM.

Chen, W., Niu, Z., Zhao, X. and Li, Y. (2014). A hybrid recommendation algorithm adapted in E-Learning Environments. World Wide Web, 17(2), 271-284.

Cheung, K.S., (2006). A Comparison Of WebCT, Blackboard And Moodle For The Teaching And Learning Of Continuing Education Courses, International Conference On ICT In Teaching And Learning, Enhancing Learning Through Technology, pp 219-228.

Cooze, M., & Barbour, M. K.. (2007). Learning styles: A focus upon E-Learning practices and pedagogy and their implications for successful instructional design, Journal of Applied Educational Technology, 4 (1)

Craig, E. M. (2007). Changing paradigms: managed learning environments and Web 2.0. Campus-Wide Information Systems, 24(3), 152-161.

Dai, X., & Bikdash, M. (2015). Hybrid classification for tweets related to infection with influenza. In SoutheastCon 2015 (pp. 1-5). IEEE.

Deng, Z., Yan, M., Sang, J., & Xu, C. (2015). Twitter is faster: Personalized time-aware video recommendation from Twitter to YouTube. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 11(2), 31.

Diederich, J., & Balke, W. T. (2007). The semantic growbag algorithm: Automatically deriving categorization systems. In International Conference on Theory and Practice of Digital Libraries (pp. 1-13). Springer, Berlin, Heidelberg.

Dlab, M. H., & Hoic-Bozic, N. (2014). Recommender system for web 2.0 supported E-Learning . In Global Engineering Education Conference (EDUCON), 2014 IEEE (pp. 953-956). IEEE.

Downes, S. (2012). Learning objects: Resources for learning worldwide. In Online education using learning objects (pp. 47-57). Routledge.

Dunlap, J. C., & Lowenthal, P. R. (2009). Instructional uses of Twitter. The CU online handbook, 45.

Dwivedi, P. and Bharadwaj, K. (2011). Effective resource recommendations for E-Learning: A collaborative filtering framework based on experience and trust. In *Computational Intelligence and Information Technology*.Springer Berlin Heidelberg.

Dwivedi, P., & Bharadwaj, K. K. (2013). Effective trust-aware E-Learning recommender system based on learning styles and knowledge levels. Journal of Educational Technology & Society, 16(4), 201.

Dwivedi, P., & Bharadwaj, K. K. (2015). E-Learning recommender system for a group of learners based on the unified learner profile approach. Expert Systems, 32(2), 264-276.

Ebner, M., Lienhardt, C., Rohs, M., & Meyer, I. (2010). Microblogs in Higher Education–A chance to facilitate informal and process-oriented learning?. Computers & Education, 55(1), 92-100.

Egghe, L. (2015). Quantitative analysis of the Recall-Precision relationship in information retrieval.

Elmongui, H. G., Mansour, R., Morsy, H., Khater, S., El-Sharkasy, A., & Ibrahim, R. (2015). TRUPI: Twitter Recommendation Based on Users' Interests. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 272-284). Springer International Publishing.

Evans, C. (2014). Twitter for teaching: Can social media be used to enhance the process of learning?. British Journal of Educational Technology, 45(5), 902-915.

Fischer, G. (2012). Context-aware systems: the'right'information, at the'right'time, in the'right'place, in the'right'way, to the'right'person. In Proceedings of the international working conference on advanced visual interfaces (pp. 287-294). ACM.

Gao, F., Luo, T., & Zhang, K. (2012). Tweeting for learning: A critical analysis of research on microblogging in education published in 2008–2011. British Journal of Educational Technology, 43(5), 783-801.

Garcia, R., & Amatriain, X. (2010). Weighted content-based methods for recommending connections in online social networks. In Workshop on Recommender Systems and the Social Web (pp. 68-71).

Ghauth, K. and Abdullah, N. (2010). Measuring learner's performance in E-Learning recommender systems. *Australasian Journal of Educational Technology*, *26*(6), 764-774.

Goldberg, D., Nichols, D., Oki, B. M. and Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, *35*(12), 61-70.

Goldberg, K., Roeder, T., Gupta, D., and Perkins, C.(2001). Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, *4*(2), 133-151.

Grosseck, G., & Holotescu, C. (2008). Can we use Twitter for educational activities? In 4th international scientific conference, e-Learning and software for education, Bucharest, Romania.

Guo, J., Nomura, A., Barton, R., Zhang, H., & Matsuoka, S. (2018). Machine Learning Predictions for Underestimation of Job Runtime on HPC System. In Asian Conference on Supercomputing Frontiers (pp. 179-198). Springer, Cham.

Ha, I., & Kim, C. (2014). The research trends and the effectiveness of smart learning. International Journal of Distributed Sensor Networks, 2014.

Ha, I., & Kim, C. (2014). Understanding User Behaviors in Social Networking Service for Mobile Learning: a Case Study with Twitter. Malaysian Journal of Computer Science, 27(2).

Hameed, T. M., Hassan, Z. B. H., & Sulaiman, R. (2015). Is Social Network an Effective E-Learning Tool: A Survey. Middle-East Journal of Scientific Research, 23(1), 119-126.

Hannon, J., McCarthy, K., & Smyth, B. (2011). Finding useful users on Twitter: twittomender the followee recommender. In European Conference on Information Retrieval (pp. 784-787). Springer Berlin Heidelberg.

Helic, D. (2007). Managing collaborative learning processes in E-Learning applications. In Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on (pp. 345-350). IEEE.

Henze, N. (2005). Personalized Learning Object Readers for the Semantic Web. In: Looi, Proc. of 12th International Conference on Artificial Intelligence in Education, AIED'2005. IOS Press 2005 274-281.

Hogo, M. (2010). Evaluation of E-Learning systems based on fuzzy clustering models and statistical tools. *Expert systems with applications*, *37*(10), 6891-6903.

Hoic-Bozic, N., Dlab, M. H., & Mornar, V. (2016). Recommender System and Web 2.0 Tools to Enhance a Blended Learning Model. IEEE Transactions on education, 59(1), 39-44.

Islam, N., Beer, M., & Slack, F. (2015). E-Learning Challenges Faced by Academics in Higher Education: A Literature Review. Journal of Education and Training Studies, 3(5), 102-112.

Jonnalagedda, N., & Gauch, S. (2013). Personalized news recommendation using twitter. In Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 03 (pp. 21-25). IEEE Computer Society.

Junco, R., Elavsky, C. M., & Heiberger, G. (2013). Putting Twitter to the test: Assessing outcomes for student collaboration, engagement, and success. British Journal of Educational Technology, 44(2), 273-287.

Kassens-Noor, E. (2012). Twitter as a teaching practice to enhance active and informal learning in higher education: The case of sustainable tweets. Active Learning in Higher Education, 13(1), 9-21.

Khribi, M. K., Jemni, M., & Nasraoui, O. (2007). Toward a hybrid recommender system for E-Learning personalization based on web usage mining techniques and information retrieval. In E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 6136-6145). Association for the Advancement of Computing in Education (AACE).

Khribi, M. K., Jemni, M., & Nasraoui, O. (2008). Automatic recommendations for E-Learning personalization based on web usage mining techniques and information retrieval. In Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on (pp. 241-245). IEEE.

Kumar, S. M., Anusha, K., & Sree, K. S. (2015). Semantic Web-based Recommendation: Experimental Results and Test Cases. International Journal of Emerging Research in Management & Technology, 4(6), 215-222.

Kumaran, V. S., & Sankar, A. (2013). Recommendation system for adaptive E-Learning using Semantic Net. International Journal of Computer Applications, 63(7).

Kurbel, K. (2001). Virtuality on the students' and on the teachers' sides: a multimedia and internet-based international master program. In Proceedings of the 7th International Conference on Technology Supported Learning and Training. Online Educa (pp. 133-136).

Kurilovas, E., Zilinskiene, I. and Dagiene, V. (2014). Recommending suitable learning paths according to learners' preferences: Experimental research results. *Computers in Human Behavior*.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media?. In Proceedings of the 19th international conference on World wide web (pp. 591-600). ACM.

Lee, B. C., Yoon, J. O., & Lee, I. (2009). Learners' acceptance of E-Learning in South Korea: Theories and results. Computers & Education, 53(4), 1320-1329.

Lennon, J., & Maurer, H. A. (2003). Why it is difficult to introduce E-Learning into schools and some new solutions. J. UCS, 9(10), 1244.

Ley, M. DBLP. (2009). Some Lessons Learned. Proc. VLDB Endow. 2(2), 1493–1500.s

Li, Y. M. and Kao, C. (2009). TREPPS: a trust-based recommender system for peer production services. *Expert systems with applications*, *36*(2), 3263-3277.

Li, Y., McLean, D., Bandar, Z. A., O'shea, J. D., & Crockett, K. (2006). Sentence similarity based on semantic nets and corpus statistics. IEEE transactions on knowledge and data engineering, 18(8), 1138-1150.

Lightner, S., Bobler, M.J., Willi, C. (2007). Team-based activities to promote engaged learning. College Teaching, Winter 2007, 55(1).

Lu, J. (2004). A personalized E-Learning material recommender system. In International Conference on Information Technology and Applications. Macquarie Scientific Publishing.

Magnuson, A., Dialani, V., & Mallela, D. (2015). Event recommendation using twitter activity. In Proceedings of the 9th ACM Conference on Recommender Systems (pp. 331-332). ACM.

Malik, H. (2009). Visual semantic web: Ontology-based E-Learning management system. *Blekinge Institute of Technology, Sweden*.

Malik, W. (2008). Visual Semantic Web. Ontology-based E-Learning management system.

Maloney, E. (2007). What Web 2.0 can teach us about learning. Chronicle of higher education, 53(18), B26.

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology-enhanced learning. In Recommender systems handbook (pp. 387-415). Springer, Boston, MA.

Marin, L., Moreno, A., & Isern, D. (2014). Automatic preference learning on numeric and multi-valued categorical attributes. Knowledge-Based Systems, 56, 201-215.

Markellou, P., Mousourouli, I., Spiros, S., Tsakalidis, A. (2005). Using semantic web mining technologies for personalized E-Learning experiences. In Proceedings of the web-based education, Grindelwald, Switzerland (pp. 461–826).

Maurer, H., & Sapper, M. (2001). E-Learning has to be seen as part of general knowledge management. In EdMedia: World Conference on Educational Media and Technology (pp. 1249-1253). Association for the Advancement of Computing in Education (AACE).

Merceron, A., & Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. Journal of Interactive Learning Research, 15(4), 319.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. International journal of lexicography, 3(4), 235-244.

Milne, D. Witten, I.H. (2008). Learning to link with Wikipedia. In Proceedings of the ACM Conference on Information and Knowledge Management CIKM'2008.

Okazaki, N., Matsuo, Y., Matsumura, N., & Ishizuka, M. (2003). Sentence extraction by spreading activation through sentence similarity. IEICE TRANSACTIONS on Information and Systems, 86(9), 1686-1694.

Pang, C., Hendriksen, D., Dijkstra, M., van der Velde, K. J., Kuiper, J., Hillege, H. L., & Swertz, M. A. (2015). BiobankConnect: software to rapidly connect data elements for pooled analysis across biobanks using ontological and lexical indexing. Journal of the American Medical Informatics Association, 22(1).

Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In The adaptive web (pp. 325-341). Springer, Berlin, Heidelberg.Phelan, O., McCarthy, K., & Smyth, B. (2009, October). Using Twitter to recommend real-time topical news. In Proceedings of the third ACM conference on Recommender systems (pp. 385-388). ACM.

Pitsilis, G. and Marshall, L. (2008). Modeling trust for recommender systems using similarity metrics. In Trust Management II. 103-118.

Porter, M. F., & Codclams, C. C. (1980). Porter,“. An algorithm for suffix stripping”, Program, 14(3), 130-137.

Prensky, M. (2001). Digital natives, digital immigrants part 1. On the horizon, 9(5).

Rashid, S., Khan, R. and Ahmed, F. (2013). A Proposed Model of E-Learning Management System Using Semantic Web Technology.

Reed, P. (2013). Hashtags and retweets: using Twitter to aid Community, Communication and Casual (informal) learning. Research in Learning Technology, 21.

Ren, Y., Li, G. and Zhou, W. (2014). A survey of recommendation techniques based on offline data processing. *Concurrency and Computation: Practice and Experience*.

Romero, C. and Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, *33*(1), 135-146.

Rousu, J., Saunders, C., Szedmak, S., & Shawe-Taylor, J. (2006). Kernel-based learning of hierarchical multilabel classification models. Journal of Machine Learning Research, 7(Jul), 1601-1626.

Rubens, N., Kaplan, D., & Okamoto, T. (2014). E-Learning 3.0: anyone, anywhere, anytime, and AI. In New Horizons in Web-Based Learning (pp. 171-180). Springer Berlin Heidelberg.

Saaya, Z., Rafter, R., Schaal, M., & Smyth, B. (2013). The curated web: a recommendation challenge. In Proceedings of the 7th ACM conference on Recommender systems (pp. 101-104). ACM.

Saaya, Z., Rafter, R., Schaal, M., & Smyth, B. (2013). The curated web: a recommendation challenge. In Proceedings of the 7th ACM conference on Recommender systems (pp. 101-104). ACM.

Salahli, M. A., Ozdemir, M. and Yasar, C. (2012). Building a fuzzy knowledge management system for personalized E-Learning. *Procedia-Social and Behavioral Sciences*, *46*, 1978-1982.

Salehi, M., Pourzaferani, M. and Razavi, S. A. (2013). Hybrid attribute-based recommender system for learning material using a genetic algorithm and a multidimensional information model. *Egyptian Informatics Journal*, *14*(1), 67-78.

Schafer, J. B., Frankowski, D., Herlocker, J.and Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web* (pp. 291-324). Springer Berlin Heidelberg.

Sevarac, Z., Devedzic, V. and Jovanovic, J. (2012). Adaptive neuro-fuzzy pedagogical recommender. Expert Systems with Applications, 39(10), 9797-9806.

Sharif, N., & Afzal, M. T. (2015). Recommendation approaches for e-learners: a survey. In Proceedings of the 7th International Conference on Management of computational and collective intelligence in Digital EcoSystems (pp. 137-141). ACM.

Sharif, N., Afzal, M. T., & Helic, D. (2012). A framework for resource recommendations for learners using social bookmarking. In Computing and Networking Technology (ICCNT), 2012 8th International Conference on (pp. 71-76). IEEE.

Sharif, N., Afzal, M.T. and Helic, D. (2014). Learning Management Systems - A Need for Specialized Systems. In *IPSI Bgd Internet Research Society, New York, Frankfurt, Tokyo, Belgrade July 2014 Volume 10 Number 2 (ISSN 1820-4503).*

Sheard, J., Ceddia, J., Hurst, J., & Tuovinen, J. (2003). Inferring student learning behavior from website interactions: A usage analysis. Education and Information Technologies, 8(3), 245-266.

Shen, L. P., and Shen, R. M. (2005). Ontology-based learning content recommendation. *International Journal of Continuing Engineering Education and Life-Long Learning*, *15*(3-6), 308-317.

Stojanovic, L., Staab, S., & Studer, R. (2001). E-Learning based on the Semantic Web. In WebNet2001-World Conference on the WWW and Internet (pp. 23-27).

Stove, R. J. (2007). The Cult of the Amateur: How Today's Internet Is Killing Our Culture [Book Review]. National Observer, (74), 64.

Sugiyama, K.and Kan, M. Y. (2010). Scholarly paper recommendation via user's recent research interests. In *Proceedings of the 10th annual joint conference on Digital libraries* (pp. 29-38). ACM.

Sun, X., Li, Z. and Hu, S. (2014). Directed-Hypergraph Based E-Learning Process Modeling Supporting Dynamic-Personalized-Combined Resource Optimization. In *Digital Home (ICDH), 2014 5th International Conference on* (pp. 324-330). IEEE.

Tane, J., Schmitz, C., & Stumme, G. (2004). Semantic resource management for the web: an E-Learning application. In Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters (pp. 1-10). ACM.

Tang, T. and McCalla, G. (2005). Smart recommendation for an evolving E-Learning system: Architecture and experiment. *International Journal on E-Learning*, *4*(1), 105-129.

Torvik, V. I., & Smalheiser, N. R. (2009). Author name disambiguation in MEDLINE. ACM Transactions on Knowledge Discovery from Data (TKDD), 3(3), 11.

Victor Codina and Luigi Ceccaroni. (2010). Taking Advantage of Semantics in Recommendation Systems. René Alquézar, Antonio Moreno, and Josep Aguilar (Eds.) IOS Press, Amsterdam, The Netherlands, 163-172

Victor, P., Cornelis, C., De Cock, M. and Da Silva, P. (2009). Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, *160*(10), 1367-1382.

Waßmann, I., Schönfeldt, C., & Tavangarian, D. (2014). WIKI-LEARNIA: SOCIAL E-LEARNING IN A WEB 3.0 ENVIRONMENT. Engineering Sciences & Technologies/Nauki Inzynierskie i Technologie, 4(1).

Waßmann, I., Schönfeldt, C., & Tavangarian, D. (2014). WIKI-LEARNIA: SOCIAL E-LEARNING IN A WEB 3.0 ENVIRONMENT. Engineering Sciences & Technologies/Nauki Inzynierskie i Technologie, 4(1).

Wright, A. (2009). Exploring a ‘deep Web that Google cannot grasp. New York Times, 23, B1.

Wrubel, J., White, D., Allen, J. (2009). High Fidelity E-Learning: SEI's Virtual Training Environment (VTE), Technical Report, CMU/SEI-2009-TR-005 ESC –TR-2009-005, CERT program. Hanscom AFB, MA 01731-2100, Carnegie Mellon University.

Wu, D., Zhang, G. and Lu, J. (2014). A fuzzy tree matching-based personalized E-Learning recommender system. In *Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on* (pp. 1898-1904). IEEE.

Yarandi, M., Jahankhani, H and Tawil, A. R. (2013). A personalized adaptive E-Learning approach based on semantic web technology. *Webology*, *10*(2), Art-110.

Yuan, W., Guan, D., Lee, Y. K., Lee, S. and Hur, S. J. (2010). Improved trust-aware recommender system using small-worldness of trust networks. *Knowledge-Based Systems*, *23*(3), 232-238.

Yuan, W., Guan, D., Lee, Y. K., Lee, S., & Hur, S. J. (2010). Improved trust-aware recommender system using small-worldness of trust networks. Knowledge-Based Systems, 23(3), 232-238.

Zaíane, O. R. (2002). Building a recommender agent for E-Learning systems. In Computers in education, 2002. Proceedings. International Conference on (pp. 55-59). IEEE.

Appendices

List of Publications

Sharif, N., & Afzal, M. T.(2018) Metadata features evaluation for twitter based E-Learning recommendation. *In Information Systems Education Conference (ISECON)* Proceedings *San Antonio, Texas USA*

Sharif, N., Afzal, M. T, & Ahmed Usman. (2018)Twitter recommendation: Unary metadata features evaluation with TF/IDF using n-grams. *In Information Systems Education Conference (ISECON)* Proceedings *San Antonio, Texas USA*

Sharif, N., & Afzal, M. T. (2016). Can Twitter be useful for E-Learning recommendation? In *Information Systems Education Conference (ISECON) Proceedings. Baltimore, MD*.

Sharif, N., & Afzal, M. T. (2015). Recommendation approaches for e-learners: a survey. In Proceedings of the 7th International Conference on Management of computational and collective intelligence in Digital EcoSystems (pp. 137-141). ACM.

Sharif, N., Afzal, M. T., & Muhammad, A. (2015). SEMANTIC BASED E-LEARNING RECOMMENDER SYSTEM. European Scientific Journal, ESJ, 11(10).

Sharif, N., Afzal, M.T. and Helic, D. (2014). Learning Management Systems - A Need for Specialized Systems. In IPSI Bgd Internet Research Society, New York, Frankfurt, Tokyo, Belgrade July 2014 Volume 10 Number 2 (ISSN 1820-4503).

Sharif, N., Afzal, M. T., & Helic, D. (2012). A framework for resource recommendations for learners using social bookmarking. In Computing and Networking Technology (ICCNT), 2012 8th International Conference on (pp. 71-76). IEEE.

List of Figures

List of Tables

List of Equations

Parameters Evaluation Forms

Learn Center & Twitter Recommendation

Group# 2 | Task# 1

Name: ________________ **Age:** _______________

Gender: M ☐ F ☐ **E-Learning Usage Experience:** __________

Student of: BSCS ☐ MS ☐ PHD ☐ **Semester No:** ______________

Scenario:

Suppose a user is searching, the information on **"Cloud computing"** gets nine most recent tweets. You are requested to rank the tweets considering user's Context, Profile and History.
Context: We mean the current focused topic which user is viewing. User context is **"Cloud computing"**.

History: We mean the previous topic user have viewed or downloaded for learning. User history is **"Client-server architectures"**.
Profile: We mean the areas/topics of interest which user have selected during registration. User profile is **"Distributed systems"**.

Check Keyword Relevance with Tweet:

Keywords	**Recent Tweets**	**Ranking**
Client-server architectures	**Jason Jurotich** RT @ForbesTech: Red Hat's CEO sees a huge business opportunity in the shift from client server to cloud architectures: http://t.co/Tfonnpec... http://t.co/TfonnpecCD 6/10/2014 8:48:32 PM	Relevant ☐ Average ☐ Irrelevant ☐
Client-server architectures	**Asim** "These systems are often implemented as multi-tier client server/architectures (discussed in Chapter 18)" sejak bila SEAD ada chap 18?? ._. 6/4/2014 11:07:41 PM	Relevant ☐ Average ☐ Irrelevant ☐

Client-server architectures	**Matt** @frazerbw Multiplayers a big topic! I'd start by reading up on server-client architectures, and poking about with sending/receiving TCP msgs 6/7/2014 11:53:27 PM	Relevant ☐ Average ☐ Irrelevant ☐
Cloud computing	**Johannes Lenz** RT @IFBlueprint: The Microsoft Approach to Compliance in the Cloud - Cloud Computing | Microsoft Trustworthy Computing TechNet Blogs: http:... http://t.co/ojKqdSSMI0 6/11/2014 10:19:15 AM	Relevant ☐ Average ☐ Irrelevant ☐
Cloud computing	**Hostgator Coupon** #CloudComputing Initiatives Could Save Federal Agencies Billions: Cloud computing, diversification and virtual... http://t.co/G5k6MDmIwe http://t.co/G5k6MDmIwe 6/11/2014 10:13:29 AM **VoipTown Voip Forum** New_iAppsComplete Manual: Cloud Computing - Imagine Publishing: Complete Manual: Cloud Computing - Imagine Pub... http://t.co/zOqpZ0Bu4n http://t.co/zOqpZ0Bu4n 6/11/2014 10:05:43 AM	Relevant ☐ Average ☐ Irrelevant ☐ Relevant ☐ Average ☐ Irrelevant ☐
Cloud computing	**greygatch** RT @IBMcloud: A layman's guide to cloud computing: http://t.co/rDtn18THtM via @HuffingtonPost #Cloud101	Relevant ☐ Average ☐ Irrelevant ☐

	http://t.co/rDtn18THtM 6/11/2014 10:11:21 AM	
Distributed systems	**inthecloud247** Yes. NSA distributed systems are so redundant they can't find, preserve or delete data. Codename: HUNGRYHIPPO http://t.co/bq81ross4M http://t.co/bq81ross4M 6/11/2014 8:46:55 AM	Relevant ☐ Average ☐ Irrelevant ☐
Distributed systems	**DBWorld Updates** http://t.co/rPL6juA3dm Geographical Asynchronous Information Access for Distributed Systems. (arXiv:1406.2644v1 [cs.DC]) #databases http://t.co/rPL6juA3dm 6/11/2014 5:51:23 AM	Relevant ☐ Average ☐ Irrelevant ☐
Distributed systems	**Matt Cottingham** Dapper, Google's distributed systems tracing infrastructure: http://t.co/Ak0lwFHR42. Interesting to note how it's used within the org. http://t.co/Ak0lwFHR42 6/11/2014 3:23:58 AM	Relevant ☐ Average ☐ Irrelevant ☐