

**ENHANCED FUZZY-BASED MODEL FOR COMPARING KNOWLEDGE AND
KNOWER CODES IN SOCIOLOGY**

Okolai, Biriya Diripigi1, Onuodu, Friday E.2 Murphy Barituka B3

Department of Mathematics/Computer Science, Niger Delta University, Nigeria,
Department of Computer Science, Ignatius Ajuru University of Education, Nigeria1
Department of Computer Science, University of Port Harcourt, Nigeria,
Department of Mathematics/Computer Science, Niger Delta University, Nigeria,

ABSTRACT: *The study presented an optimized approach to evaluating knowledge and knower performance in Sociology via a Fuzzy-based Model. One of the concerns raised by students in the field of sociology is the overall lack of coherence among undergraduates in sociology, despite the many individual courses on offer. The study adopted Object-Oriented System Development Methodology (OOSDM) and implementation was carried out with Hypertext Preprocessor (PHP) and MySQL as backend. The results obtained showed an automated routing metrics for performance evaluation of knowledge and knower codes in sociology, which could be beneficial to knowers and other researchers in the field of study.*

KEYWORDS: fuzzy-based model, knower code, sociology

INTRODUCTION

The study presented a modified approach to the comparison of knowledge and knower code in sociology. Knowledge can be described as a useful tool to humanity. The view of knowledge is held by many sociological accounts to characterize contemporary advanced societies. However, in a circular manner, this conception is also adopted by those accounts: they treat knowledge as interchangeable tokens, like money [1]

In the field of sociology, the concept of knowledge and the knower have sparked up questions such as why some students are more academically better than the others. In this study, we intend to show how concepts from the Specialization dimension of Legitimation Code Theory (LCT) also known as specialization codes can help to shed light on study issue. At the same time the study is focused on explaining how these concepts can produce powerful descriptions, as a way into simulating why they are increasingly drawn upon by scholars and educators from systemic functional linguistics (SFL).

The concept of LCT in the field of sociology enables adequate scrutiny of both knowledge and knowers, especially in terms of knowledge practices and students' disposition. From this overview, educational experiences (or, indeed, any practices) are an outcome of what the French sociologist Pierre Bourdieu (1996), termed 'the gathering of two histories' or logics: the dispositions (ways

of acting, thinking and being) brought by actors to a social context and the nature of that context itself.

The study intends to develop a fuzzy-based model for comparing knowledge and knower code in sociology. Fuzzy logic is an optimization machine-learning oriented technique that handles the issue of partial truth. It is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false.

Statement of the Problem

One of the concerns raised by students in the field of sociology is the overall lack of coherence among undergraduates in sociology, despite the many individual courses on offer. It is the production, re-contextualization, teaching and learning of knowledge that makes education a distinctive social field of practice. However, as a growing number of educators are in dispute, knowledge has often been obscured by educational research, thanks to a false dichotomy that creates 'knowledge-blindness'. This false perception of studying either knowing or knowers results partly from the ways psychology and sociology has been re-contextualized into educational research over the past 40 years. On the other hand, psychologically-informed approaches typically focus on processes of learning and sideline differences between the forms of knowledge being learned. 'Knowledge' tends to be viewed as that which is in people's minds, and 'learning' as comprising generic processes—the central focus thus knows.

Aim and Objectives of the Study

The aim of this study is to develop an enhanced fuzzy-based model for comparing knowledge and knower in sociology. The specific objectives of the study include to:

- i) design a fuzzy-based system for evaluating knowledge and knower performance
- ii) optimize performance of the proposed system using fuzzy logic
- iii) implement the proposed system with Hypertext Pre-processor (PHP) and MySQL
- iv) evaluate performance of the existing and proposed systems for comparing knowledge and knower in sociology.

Concept of Knowledge

Knowledge is often defined as a belief that is true and justified. This definition has led to its measurement by methods that rely solely on the correctness of answers. A correct or incorrect answer is interpreted to mean simply that a person knows or does not know something. Such methods of measurement have serious deficiencies that can be alleviated by expanding the

definition of knowledge to include the test - taker's certainty. The person's certainty about the answers on a test captures important, but now neglected, dimensions of knowledge [2]

Knowledge is a familiarity, awareness, or understanding of someone or something, such as facts (descriptive knowledge), skills (procedural knowledge), or objects (acquaintance knowledge). By most accounts, knowledge can be acquired in many different ways and from many sources, including but not limited to perception, reason, memory, testimony, scientific inquiry, education, and practice. The philosophical study of knowledge is called epistemology. The term "knowledge" can refer to a theoretical or practical understanding of a subject.

It can be implicit (as with practical skill or expertise) or explicit (as with the theoretical understanding of a subject); formal or informal; systematic or particular. The philosopher Plato famously pointed out the need for a distinction between knowledge and true belief in the *Theaetetus*, leading many to attribute to him a definition of knowledge as "justified true belief". The difficulties with this definition raised by the Gettier problem have been the subject of extensive debate in epistemology for more than half a century [2]

Sociology as a Knower Code

Sociology is the study of human social relationships and institutions. Sociology's subject matter is diverse, ranging from crime to religion, from the family to the state, from the divisions of race and social class to the shared beliefs of a common culture, and from social stability to radical change in whole societies. Unifying the study of these diverse subjects of study is sociology's purpose of understanding how human action and consciousness both shape and are shaped by surrounding cultural and social structures [3]

Sociology is an exciting and illuminating field of study that analyses and explains important matters in our personal lives, our communities, and the world. At the personal level, sociology investigates the social causes and consequences of such things as romantic love, racial and gender identity, family conflict, deviant behaviour, aging, and religious faith.

At the societal level, sociology examines and explains matters like crime and law, poverty and wealth, prejudice and discrimination, schools and education, business firms, urban community, and social movements. At the global level, sociology studies such phenomena as population growth and migration, war and peace, and economic development [3]

Fuzzy Logic as a Tool for Enhancing Knowledge and Knower Codes

Fuzzy logic is a branch of science that is extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. Fuzzy logic may be applied to many fields, including control systems, neural networks and artificial intelligence (AI).

It is straightforward to formulate a set of fuzzy rules for this task, but it is not immediately obvious to show how to build a network to do neither the same nor how to train it. Fuzzy logic is now being

used in many products of industrial and consumer electronics for which a good control system is sufficient and where the question of optimal control does not necessarily arise. The difference between crisp (i.e., classical) and fuzzy sets is established by introducing a membership function. Consider a finite set

$$X = \{x_1, x_2, \dots, x_n\} \quad (2.3)$$

which will be considered the universal set in what follows. The subset A of X consisting of the single element x_1 can be described by then-dimensional membership vector

$$Z(A) = (1, 0, 0, \dots, 0), \quad (2.4)$$

where the convention has been adopted that a 1 at the i th position indicates that x_i belongs to A. The set B composed of the elements x_1 and x_n is described by the vector

$$Z(B) = (1, 0, 0, \dots, 1). \quad (2.5)$$

Any other crisp subset of X can be represented in the same way by an n -dimensional binary vector. But what happens if we lift the restriction to binary vectors? In that case we can define the fuzzy set C with the following vector description:

$$Z(C) = (0.5, 0, 0, \dots, 0) \quad (2.6)$$

Related Work

Clarence et al, [4], looked at knowledge and knowers in teaching. The study proposed a complementary approach that accounts for different kinds of knowledge and knower building. Using Legitimation Code Theory's concept of specialization, the paper argues that accounting for what makes a discipline 'special' in terms of its basis for legitimate achievement can enable curriculum writers to align curricula more effectively with that basis in different disciplines.

Birchoff [5], proposed a unique theory for knowledge and knower in the Department of Political Science. The study opined that a significant first step in extending constructive alignment to more adequately account for different forms of knowledge and knower construction is to consider what we mean by knowledge.

Khan et al, [6], looked at Critical perspectives on methodology in pedagogic research. The study opined that aligning a curriculum, when the underlying code of a discipline has been conceptualized and unpacked, can become less focused on connecting 'content' with 'skills' in teaching and assessment; rather it can shift to aligning teaching and students' learning with the code of the discipline itself. In the case of a knower code, what need to be aligned across and between years of study are the underlying critical, imaginative and analytical dispositions

Pham et al, [7], looked at Mining Patterns in Source Code Using Tree-Mining Algorithm for Knowledge and Knower Codes. According to the authors, “discovering regularities in source code is of great interest to software engineers, both in academia and in industry, as regularities can provide useful information to help in a variety of tasks such as code comprehension, code refactoring, and fault localization”. In addition, the authors proposed FREQTALS, a new algorithm for mining patterns in source code based on the FREQT tree mining algorithm. First, they introduced several constraints that effectively enable them to find more useful patterns; then, they showed how to efficiently include them in FREQT. To illustrate the usefulness of the constraints, the authors carried out a case study in collaboration with software engineers, where they identified a number of interesting patterns in a repository of Java code. However, their analyzed Frequent Sub-tree Mining Technique was not implemented with a machine learning model which could have further enhanced their adopted case study.

Miltiadis and Charles [8], looked at Analysing Knowledge and Knower Codes via Mining Source-Code Repositories at Massive Scale using Language Modeling. The authors built the first giga-token probabilistic language model of source code, based on 352 million lines of Java. Furthermore, the authors called the metrics data-driven complexity metrics. They also proposed new metrics that measured the complexity of a code module and the topical centrality of a module to a software project. However, the datasets used for training and designing their Source-Code Repository Mining Model was not trained with machine learning algorithms which further resulted to latencies in their model performance.

Chris and Daniel [9], looked at Structured Generative Models of Natural Knower Code. The authors studied the problem of building generative models of natural source code (NSC); that is, source code written by humans and meant to be understood by humans. Their primary contribution is to describe new generative models that are tailored to NSC. The models are based on probabilistic context free grammars (PCFGs) and neuro-probabilistic language models which are extended to incorporate additional source code-specific structure. However, the authors failed to show how the developed generative model can be fragmented in order to manage memory in a repository.

Ruchika and Anuradha [10], looked at Knowledge and Knower Codes in Software Reusability: Systematic Literature Review and Current Trend. The authors performed a systematic review of the existing studies related to software maintainability from January 1991 to October 2015. In total, 96 primary studies were identified out of which 47 studies were from journals, 36 from conference proceedings and 13 from others. All studies were compiled in structured form and analysed through numerous perspectives such as the use of design metrics, prediction model, tools, data sources, prediction accuracy, etc. According to the reviewed results, the authors found that the use of machine learning algorithms in predicting maintainability has increased since 2005. The use of evolutionary algorithms has also begun in related sub-fields since 2010. In addition, they observed that design metrics is still the most favoured option to capture the characteristics of any

given software before deploying it further in prediction model for determining the corresponding software maintainability. However, the review was not implemented with a machine learning model which could have further enhanced their adopted case study.

Madhulatha [11], looked at an overview on clustering methods for comparing Knowledge and Knower Codes. The study covered about clustering algorithms, benefits and its applications. According to the author, clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is the process of grouping similar objects into different groups, or more precisely, the partitioning of a data set into subsets, so that the data in each subset according to some defined distance measure. However, analysis of the study showed that the author failed to implement the discussed clustering methods to a model for further clarification and understanding.

Zhiting et al, [12], looked at the unification of Knowledge and Knower Codes via Deep generative models. According to the authors, “deep generative models have achieved impressive success in recent years. Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), as powerful frameworks for deep generative model learning, have largely been considered as two distinct paradigms and received extensive independent studies respectively”. Furthermore, the study aimed at establishing formal connections between GANs and VAEs through a new formulation of them. The authors interpret sample generation in GANs as performing posterior inference, and show that GANs and VAEs involve minimizing KL divergences of respective posterior and inference distributions with opposite directions, extending the two learning phases of classic wake-sleep algorithm, respectively. The unified view provides a powerful tool to analyse a diverse set of existing model variants, and enables to transfer techniques across research lines in a principled way. However, the authors failed to adopt a supervised learning approach to their designed paradigm for deep generative models.

Jungang et al, [13], looked at an overview of deep generative models for comparing Knowledge and Knower Codes. The study analyzed three important deep generative models including DBNs, deep autoencoder, and deep Boltzmann machine are reviewed. In addition, some successful applications of deep generative models in image processing, speech recognition and information retrieval are also introduced and analyzed. However, analysis of the study showed that the authors failed to adopt a supervised learning approach to their designed paradigm for deep generative models.

David and Graham [14], researched on Knowledge design science in decision support systems research. According to the authors, “design science has been an important strategy in decision support systems (DSS) research since the field’s inception in the early 1970s”. Recent reviews of DSS research have indicated a need to improve its quality and relevance. DSS design-science research has an important role in this improvement because design-science research can engage industry and the profession in intellectually important projects. In addition, the analysis carried out

in the study highlights major issues in DSS research that need attention: research design, evaluation, relevance, strategic focus, and theorizing. However, analysis of the study showed that the author failed to implement the discussed decision support system methods to a model for further clarification and understanding.

Smita and Patel [15], looked at a study of graph storage database of Knowledge and Knower Codes. The study described what is big data storage management, dimensions of big data, types of data, what is structured and unstructured data, what is NoSQL database, types of NoSQL database, basic structure of graph database, advantages, disadvantages and application area and comparison of various graph database. However, the authors failed to implement the discussed big data concepts with NoSQL database.

Wasiwasi and Zaipuna [16], looked at the design and development of web-based digital repository for scholarly communication. The aim of the study was to design and develop a web - based digital repository for scholarly communications using NM - AIST as a case study. The system was developed using open source software. Findings obtained from system validation tests show that the system is a viable solution to the major challenges encountered in the management and sharing of scholarly information at the institution. The authors did a good job. However, a major limitation of their study is that the developed model was deficient in benchmarking and cost benefits analysis.

MATERIALS AND METHODS

Analysis of the Existing System

The existing system is an LCT-based Model for measuring Knowledge and Knower Code Satisfaction as illustrated in Figure 3.1. The concept of LCT (Legitimation Code Theory) is to measure education performance in especially in terms of the educator and the knower. This further encompasses scrutiny of learning structure, tools and so on. Also, the Legitimation Code Theory (LCT) also known as specialization codes can help to shed light on study issue. At the same time the study is aimed at illustrating how these concepts can generate powerful explanations, as a way into understanding why they are increasingly drawn upon by scholars and educators from systemic functional linguistics (SFL).

From figure 3.1, a new system user initializes the LCT-based model and queries for specific performance metrics for shared knowledge and impartation effect of a specific field. There is a virtual repository for interfacing sources of information requested by the user. The existing system is also a framework for exploring practices in terms of their organizing principles or 'legitimation codes'. It is also the basis of an international and multidisciplinary community of scholars, educators and practitioners who use the approach to shape their research and practice. Research and teaching enacting LCT are motivated by concerns with social justice and knowledge-building. LCT concepts reveal the 'rules of the game' shaping different arenas of social life, such as education and law.

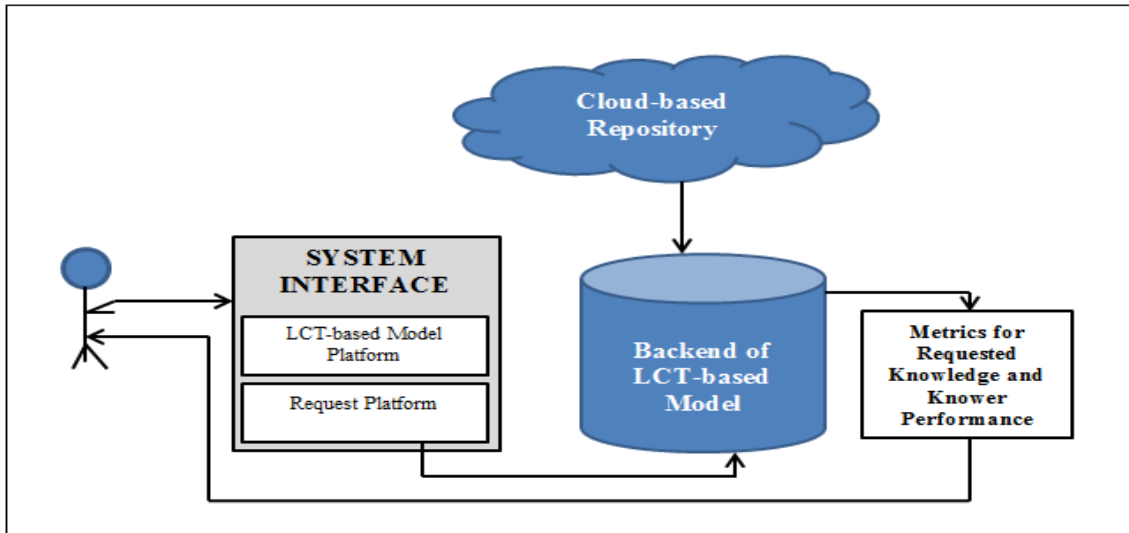


Figure 3.1: Existing System Architecture of an LCT-based Model for Knowledge and Knower Performance (Source: [4])

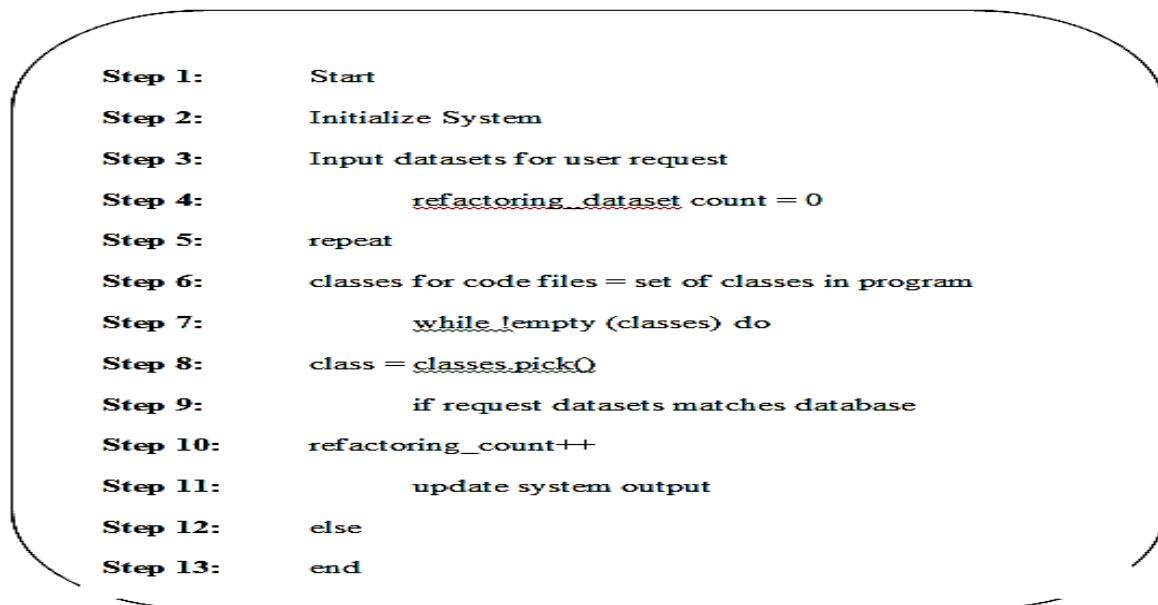


Figure 3.2: Algorithm of the Existing System (Source [4])

Table 3.1 shows the explanation of the existing system components

Components	Description
End-User	This component illustrates any individual that uses the system
System Interface	This component illustrates a graphical user platform which consist of sub-components such as the LCT-based model platform and the request platform for knowledge and knower codes performance
Cloud-based Repository	This component illustrates a powerful physical or virtual infrastructure that performs application- and information-processing storage.
Backend of the LCT-based Model	This component illustrates a centralized storage location of information of the system
Performance Metrics	This component illustrates the requested results by the user on performance of knowledge and knowledge codes

Advantages of the Existing System

The following advantages of the Existing System are:

- i) a unique graphical user interface for flexible end-user interaction with the system
- ii) easy request interpretation of the LCT-based model
- iii) quick search links and metrics for performance of knowledge and knower codes

Disadvantages of the Existing System

The following disadvantages of the existing system are

- i) poor data security for stored information on knowledge and knower codes performance
- ii) complex interpretation of the requested metrics by the user
- iii) latency in the entire process
- iv) absence of a machine-learning oriented techniques which would have keep the system updated in the long run.

Proposed System for Measuring Knowledge and Knower Code Satisfaction

The proposed system is an improvement of the existing system as illustrated in Figure 3.3. The improvement encompasses a fuzzy-based approach to the measurement of knowledge and knower

code satisfaction. The fuzzy-based approach will also optimize the system performance especially in terms of security, update and speed. This is because, the Fuzzy Logic in the design process is a system which consists of a knowledge base, which includes the information, given in the form of linguistic control rules, and a fuzzification interface, which has the effect of transforming crisp data into fuzzy sets. Additionally, an Inference System is included which works together with the knowledge base to provide inference features.

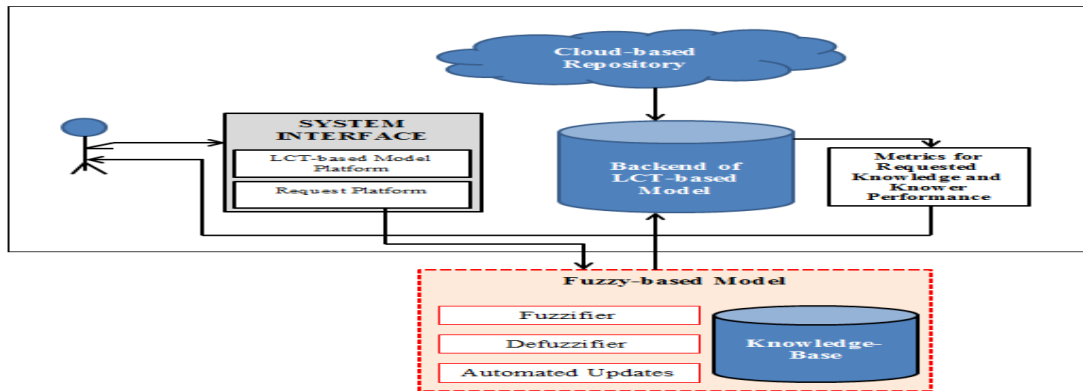


Figure 3.3: Proposed System Architecture

Table 3.1: Sample Datasets (Source: Github, 2020)

Sn	Fname	Lname	Gender	Email	Address	Phone No:
1.	Mafuzu	Clark	Male	mc@gmail.com	Null	Null
2.	Sarah	Williams	Female	sw@gmail.com	Null	Null
3.	Olivia	Shane	Female	os@gmail.com	Null	Null
4.	Hannah	Ethan	Female	he@gmail.com	Null	Null
5.	Megan	Cameron	Male	mec@gmail.com	Null	Null
6.	Chloe	Tim	Male	ct@gmail.com	Null	Null
7.	Jess	David	Male	jd@gmail.com	Null	Null
8.	Caitlin	Andre	Female	ca@gmail.com	Null	Null
9.	Tallulah	Lawrence	Male	tl@gmail.com	Null	Null
10.	Ammaarah	Joshua	Female	aj@gmail.com	Null	Null
11.	Michelle	Johann	Female	mj@gmail.com	Null	Null
12.	Jenna	Ryan	Female	jr@gmail.com	Null	Null
13.	Haajarah	James	Male	hj@gmail.com	Null	Null
14.	Emma	Matt	Male	em@gmail.com	Null	Null

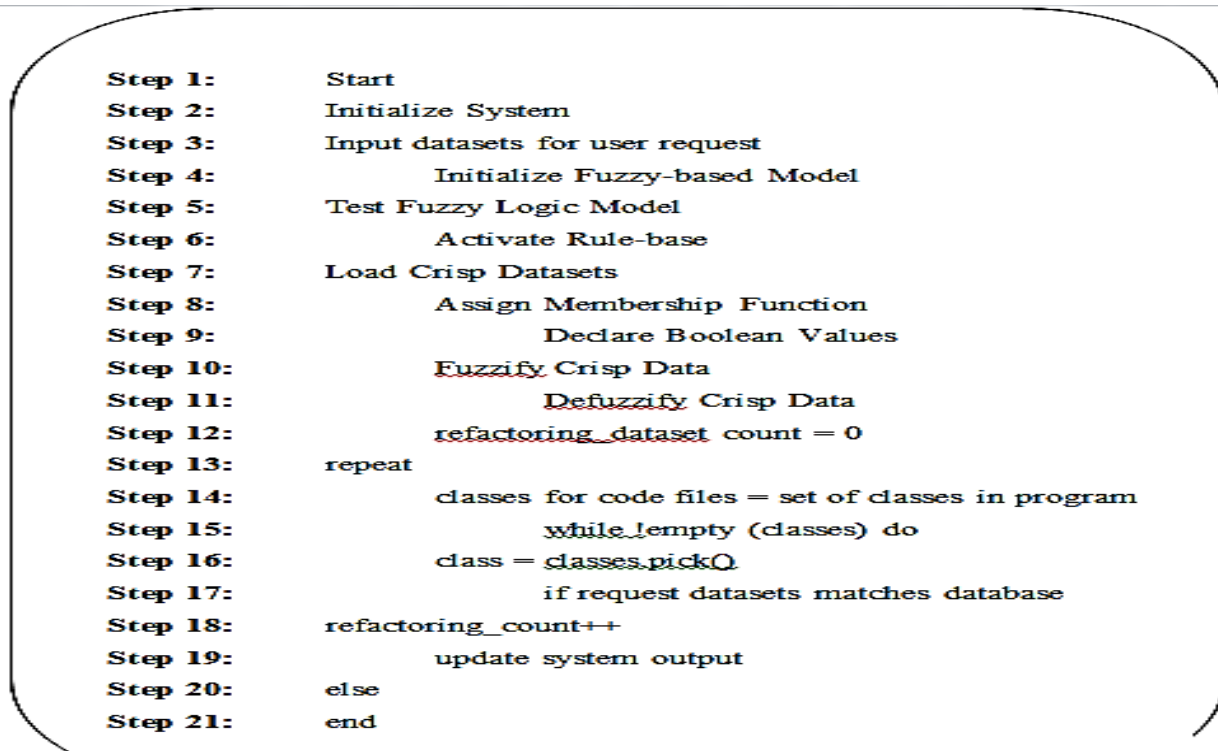


Figure 3.4: Algorithm of the Proposed System

METHODOLOGY

Many system methodologies are available to structure, plan and control the overall activities involved in the optimization of the existing system or development of the new system. These methodologies combine set of principles, practices and processes that allows the development of systems quickly and properly. This research work on Enhanced Fuzzy-based Model for comparing knowledge and knower codes in Sociology will be achieved following the Object-Oriented System Development Methodology (OOSDM). This is aimed at viewing, modeling and implementing the proposed system as a collection of interacting classes and objects. OOSDM is adopted because it is more effective, efficient, reliable, reusable and a faster way of developing systems. Object-Oriented System Development Methodology (OOSDM) is a technical approach for analyzing and designing an application, system, or business by applying object-oriented programming, as well as using visual modeling throughout the software development process to guide stakeholder communication and product quality.

OOSDM in modern software engineering is typically conducted in an iterative and incremental way. The outputs of OOSDM activities are analysis models (for OOA) and design models (for

OOD) respectively. The intention is for these to be continuously refined and evolved, driven by key factors like risks and business value.

Furthermore, the Object-Oriented System Development Methodology (OOSDM) is a technical approach for analyzing and designing an application, system, or business by applying object-oriented programming, as well as using visual modeling throughout the software development process to guide stakeholder communication and product quality. In the early days of object-oriented technology before the mid-1990s, there were many different competing methodologies for software development and object-oriented modeling, often tied to specific Computer Aided Software Engineering (CASE) tool vendors. No standard notations, consistent terms and process guides were the major concerns at the time, which degraded communication efficiency and lengthened learning curves.

Datasets

A dataset (or data set) is a collection of data, usually presented in tabular form. Each column represents a particular variable. Each row corresponds to a given member of the dataset in question. It lists values for each of the variables, such as height and weight of an object. Each value is known as a datum. The dataset may comprise data for one or more members, corresponding to the number of rows. There are many ways in which data can be collected for example, as part of service delivery, one-off surveys, interviews, observations, and so on.

Table 3.2 shows the description of the proposed system component

Component	Description
Fuzzy-based Model	This component illustrates the application of fuzzy logic to the existing system improvement, in order to achieve optimization of the system especially in terms of security, update and speed. This is because, the Fuzzy Logic in the design process is a system which consists of a knowledge base, which includes the information given in the form of linguistic control rules

Advantages of the Proposed System

The following advantages of the proposed system are:

- i) ability of the system to be updated in the long run
- ii) improved security of the system due to the applied machine learning technique

- iii) a unique graphical user interface for flexible end-user interaction with the system
- iv) easy request interpretation of the LCT-based model
- v) quick search links and metrics for performance of knowledge and knower codes

RESULTS AND DISCUSSION

Outputs and Discussion

Figure 4.1 shows the welcome page of the proposed system. The welcome page consist of three navigation links namely home, register new user and evaluate performance of the knowledge and knower. Figure 4.2 shows the registration page for a new user, while Figure 4.3 illustrates the test-set input page for evaluating performance of the knowledge and knower. Figure 4.4 shows the evaluation result of the system. The concept of the proposed system is optimizing academic performance of students in the field of sociology. An early prediction of students' academic performance is what this assessment tends to pursue. This is supposed to be a very important practice as it tries to obtain an ideal students' level of learning, level of teaching, decides the success or failure of students in the course of enrolment, to inform low performing students to put in more effort and overcome weaknesses and encourages students to either continue or choose an appropriate course of carrier.

The rate of final cumulative of such students first becomes necessary to study the obvious very important parameter for entry into these institutions and see whether the right parameters are being used. This may address the global call for reduction of mass failure and drop out in our institutions of higher learning. It may also address the "pass by all means" syndrome and the high rate of examination malpractices currently seen in our society.

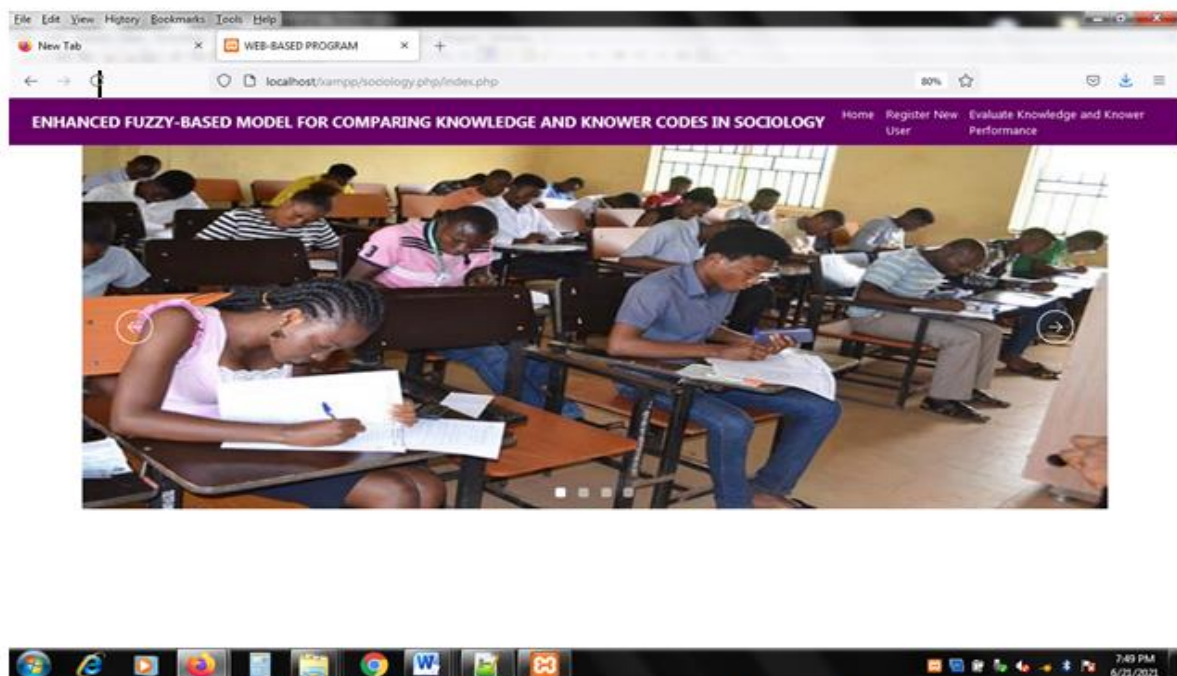


Figure 4.1: Welcome Page of the implemented Proposed System

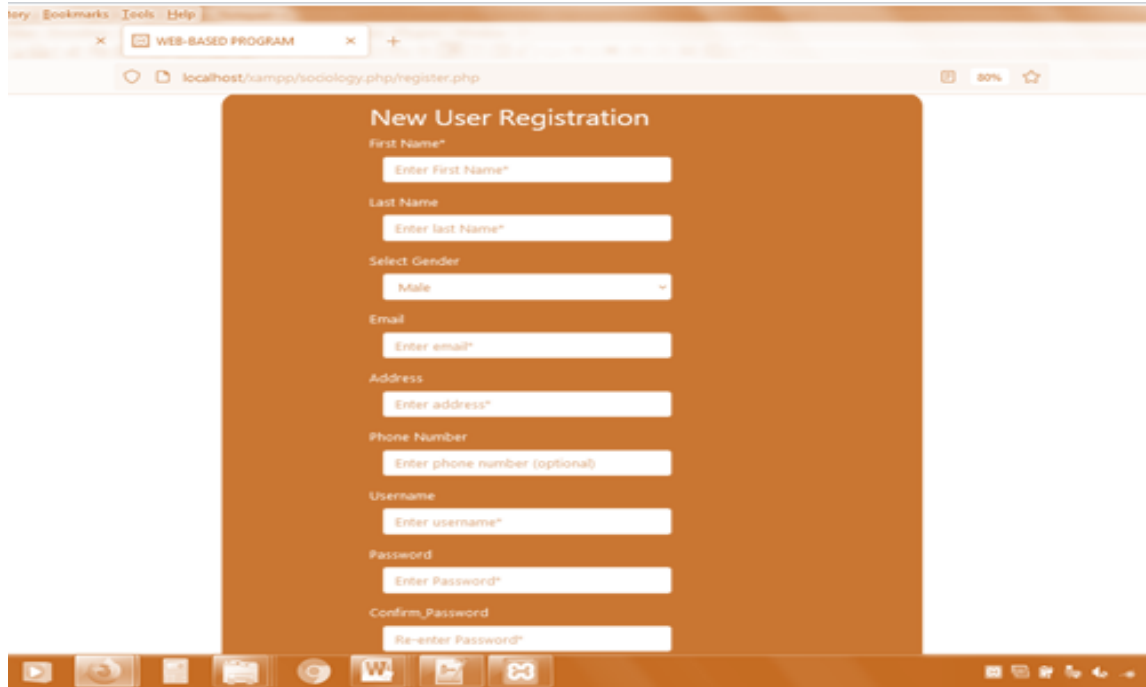


Figure 4.2: New User Registration Page of the Proposed System

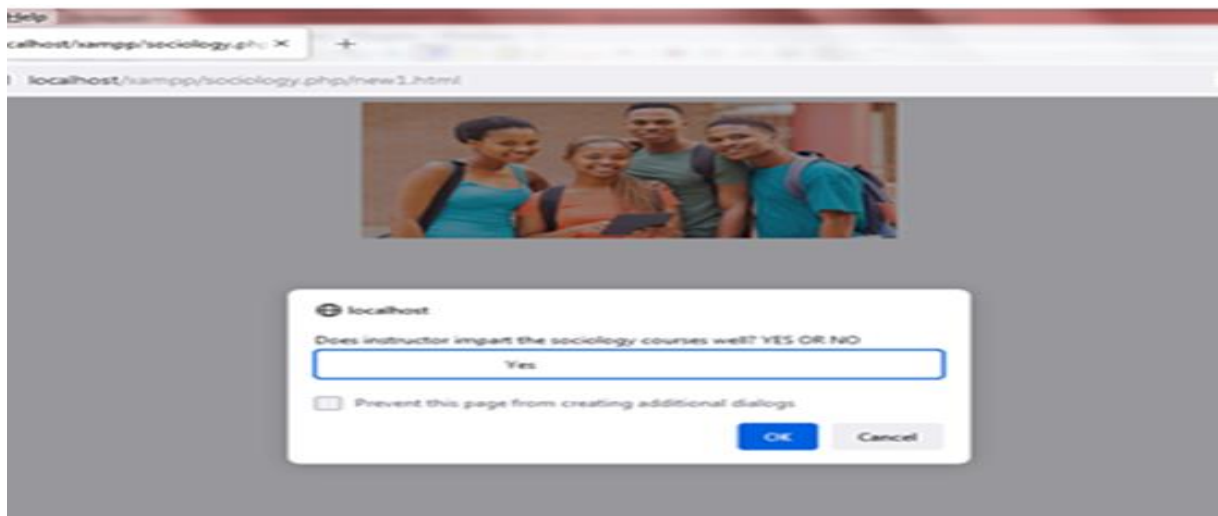


Figure 4.3: Test-set Input Page

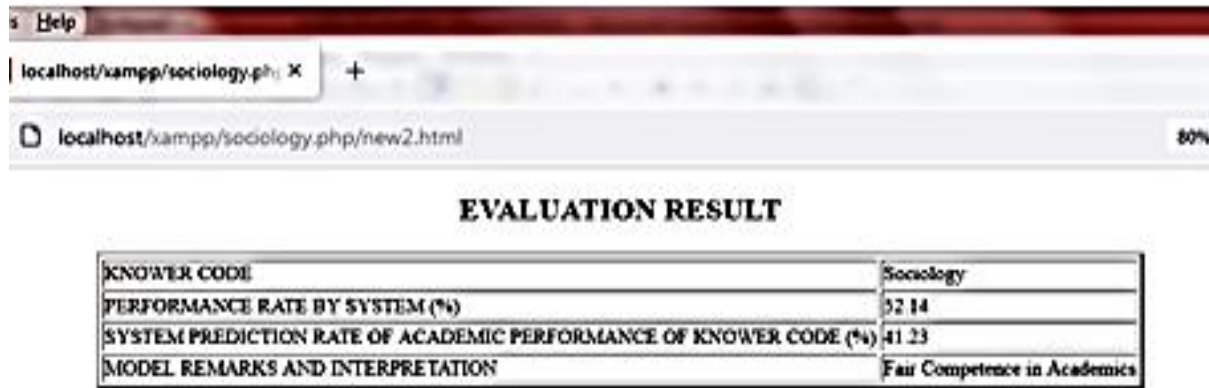


Figure 4.4: Result Page

Values



Table 4.2 Comparative Analysis of the Existing and Proposed Systems Result

(Source: Existing and Proposed Programs Implementation)

Parameters			Parameters
	Existing System	Proposed System	
Speed in Registering a new user	11	7	Speed in Registering a new user
Speed in Authenticating newly registered user	22	14	Speed in Authenticating newly registered user
Accuracy in evaluating knower codes	7	18	Accuracy in evaluating knower codes

CONCLUSION AND RECOMMENDATIONS

In this study, we developed an Enhanced Fuzzy-based Model for comparing knowledge and knower codes in sociology. In the field of sociology, the concept of knowledge and the knower have sparked up questions such as why some students are more academically better than others. In this

study, we intend to show how concepts from the Specialization dimension of Legitimation Code Theory also known as specialization codes can help to shed light on study issue. At the same time the study is aimed at illustrating how these concepts can generate powerful explanations, as a way into understanding why they are increasingly drawn upon by scholars and educators from systemic functional linguistics. The concept of LCT in the field of sociology enables adequate scrutiny of both knowledge and knowers, especially in terms of knowledge practices and students' disposition.

Contributions to Knowledge

The study contributed the following:

- i) An LCT-based Model for evaluating the performance of knowledge and knower
- ii) A modified approach to Knowledge Codes and Understanding Knower Data

Suggestion for Further Studies

The study also suggested the importance of machine-learning techniques to the evaluation of knowledge and knower codes. This is because, Machine learning (ML) techniques enable systems to learn from experience. ML refers to a system's ability to acquire and integrate knowledge

through large-scale observations and to improve and extend itself by learning new knowledge rather than by being programmed with that knowledge.

References

1. Bernstein N. (2010), An analyzed approach to Legitimization Code Theory (LCT), *Journal of Social and Philosophical Sciences*, 7(3), 1 – 22
2. Darwin L. (2013), Approximate Query Processing for Knowledge and Knower Codes using deep Generative Models, *arxiv:1903.10000v2[cs.DB]*, 45-51
3. Chen G. (2012), The Knowledge and Knower Code Theory, *ACM Journal of Computing and Sociological Research*, 2(11), 1 – 7
4. Clarence O.W, Brendan Y. and Seedorf L. (2019), An LCT-based Model for evaluating the performance of knowledge and knower, *Journal of Social and Philosophical Sciences*, 6(2), 1 – 10
5. Birchhoff B. (2017), Knowledge Codes: Understanding Knower Data, Engineering and Applied Science, Aston University, England, Research Gate Publications, <https://www.researchgate.net/publication/291229189>, 56 - 61
6. Khan O., Murat K., and Sharad M. (2015), Secure and Efficient Processing for knowledge evaluation over Hybrid Clouds, *2017 IEEE 33rd International Conference on Data Engineering*, DOI 10.1109/ICDE.2017.125
7. Pham H.S, Nijssen S. and Mens K. (2019), Mining Patterns in Source Code Using Tree Mining Algorithms, *DS2019-22nd International Conference on Discovery Science, Split Croatia du 28/10/2019, au 30/10/2019, In: Lecture Notes in Computer Science, 11828(11828)*, 471 – 480, <http://hdlhandle.net/2078.1/221356>, DOI:10.1007/978-3-030-33778-0-35
8. Miltiadis A. and Charles S. (2020), Mining Source Code Repositories at Massive Scale for evaluating knowledge and knower codes using Language Modeling, *School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, UK*, 1 – 10
9. Chris M. and Daniel T. (2014), Structured Generative Models of Knower Codes in Sociology, *Proceedings of the 31st International Conference on Machine Learning, Beijing, China, JMLR: W and CP, V.32, arXiv:1401.0514v2[cs.PL]*, 1 – 14
10. Ruchika M. and Anuradha C. (2016), Knowledge and Knower Codes: Systematic Literature Review and Current Trend, *International Journal of Software Engineering and Knowledge Engineering*, 26(8), 1221 – 1253, www.worldscientific.com, doi:10.1142/S021819406500431
11. Madhulatha T. (2012), An Overview on Clustering Methods for evaluating knowledge performance, *IOSR Journal of Engineering*, 2(4), 719 – 725, <https://arxiv.org/abs/1205.1117>
12. Zhiting H., Zichao Y., Ruslan S., and Eric X (2018), On Unifying Deep Generative Models, *Published as Conference Paper at ICLR (2018)*, 1 – 19
13. Jungang Xu, Hui Li and Shilong Zhou (2015), An Overview of Knower Codes in Science, *IETE Technical Review*, 32(2), 131 – 139, doi:10.1080/02564602.2014.987328, <http://dx.doi.org/10.1080/02564602.2014.987328>

14. David A. and Graham P. (2012), Design Science in evaluating knowledge and knower codes Research: An Assessment using the Hevner, March Park, and Ram Guidelines, *Journal of the Association for Information Systems (JAIS)*, 13(11), 923 – 949
15. Smita A. and Patel A. (2016), A Study on Graph Storage Database of NoSQL, *International Journal on Soft Computing, Artificial Intelligence and Application (IJSCAI)* 5(1), 33 – 40
16. Wasiwasi J.M and Zaipuna O.Y (2014), Design and Development of a web-based Digital Repository for Scholarly Communication: A Case of NM-AIST Tanzania, *International Journal of knowledge content development and technology* 4(2), 97 – 108