Developing Case Based Recommender System in Micro and Small Enterprises Development Agency for the Selection of Enterprises to New Members: The Case of Ethiopia

Yirga Yayeh
Head of the Department, Department of Information Technology
Faculty of Engineering and Technology, Assosa University, Ethiopia.

ABSTRACT
The main objective of this study is to develop case based recommender system for the selection of enterprise sectors and activities in Ethiopia. In order to conduct the research, the required knowledge was collected from previous sector members’ cases, domain experts, and documents through interview and document analysis methods. After all, the knowledge is modeled by hierarchical tree model and case based reasoning system has constructed. The prototype of case based reasoning for enterprise sector and activity selection is implemented by jCOLIBRI 1 programming tool. Nearest neighbor retrieval algorithm is used to measure the similarity of new case (query) with cases in the case base. As a result, if there is a similarity between the new case and the existing case, the system assigns the solution of previous case as a solution to new case. To determine the applicability of the prototype system in the domain area, the system has been evaluated by domain experts and members through visual interaction based and testing.

Key words: Micro and Small Enterprises, Federal Micro and Small Enterprise Development Agency, Enterprise Sector and Activity selection.

INTRODUCTION
Recommender systems provide advice to users about items that they might wish to purchase or examine. Recommendations made by such systems can help users navigate through large information spaces of product descriptions, news articles or other items. Specifically, recommender systems have (i) background data: the information that the system has before the recommendation process begins, (ii) input data: the information that user must communicate to the system in order to generate a recommendation, and (iii) an algorithm: that combines background and input data to arrive at its suggestions.

The Ethiopian government recognizes the significance of MSEs (Micro and Small Enterprises) sector and shows its dedication to promote the MSE Development by the Issuance of National MSEs Strategy of 1997 E.c and the establishment of the Federal Micro and Small Enterprises Development Agency (FeMSEDA). Ethiopian industrial development strategy which is issued in 2003 E.c also singled out the promotion of MSEs development as one of the important instruments to create productive and dynamic private sector. The promotion of this sector is justified on the grounds that enhancing growth with equity, creating long-term jobs, providing the basis for medium and large enterprise and promoting exports etc. The strategy puts a means to support the MSEs such as, infrastructure, financial facilities, supply of raw materials, and training. The goal of this study is to develop a recommender system in order to give recommendation to MSEs for new members that are going to be most successful in the future.

STATEMENT OF THE PROBLEM
The major problems faced in Ethiopian MSEs is the difficulty of getting speedy, consistent, clear and organized expert advice on the selection of enterprise sectors and activities. The other problem is shortage of the number of skilled experts and consulting service of enterprise sectors who can give advice on MSEs issues in the country. This study therefore aims to design a case based recommender system that helps members to get rapid and reliable advisory service to select new enterprises sectors. Based on this advice the members can match their personal characteristics and potentials.
Accordingly, this study attempts to explore and answer the following research questions:

- What is the availability of domain experts in MSEs to advice members on different enterprise sectors and activities selection?
- How does one acquire, model, represent, and implement the required knowledge for the system that is proposed.

OBJECTIVES OF THE STUDY

- To understand the concept of case based recommender system and the designing process.
- To identify and collect the previous cases, facts and rules that new members need to know in order to select MSE sectors.
- To develop a prototype recommendation system to enterprise members for the selection of MSEs that best matches with their characteristics or attribute.
- To evaluate the performance of the prototype system.

SCOPE OF THE STUDY

Due to different factors such as limitation of time and resources, the scope of the study is restricted only on developing prototype recommender system for enterprise sectors and activity selection based on different characteristics of the members. The recommender system did not include other activities of the MSEs agency such as constraints of finance, infrastructure problems, giving and renewal of enterprises license, calculation of taxation, etc.

METHOD OF DATA COLLECTION

To acquire the needed knowledge, both primary and secondary data sources were used. As Primary sources, ten (10), six (6) members and domain experts were interviewed respectively. Domain experts have been selected by using purposive sampling techniques. i.e. Based on their educational qualifications related to the domain area and their immediate position in the enterprise sector promotion 6 (six) domain experts. For both of them, the interview has been used for asking about advising system and asking about all information gathered from secondary sources. Semi structured interview was used for both members and experts. A semi-structured interview is an interview which has a guide that usually includes both closed-ended and open-ended questions.

As secondary data sources, the researcher used purposive data selecting methods to select sample datasets got from FeMSEDA of Ethiopia. Because there were large amount of data and the data includes all members in different status namely in startup (infant stage which needs to start an enterprise), growth (have a power to compute on the market) and maturity stage (it is profitable and have try to add other sectors by themselves). Therefore, to select successful enterprise sectors from each activity, the researcher selects enterprise sectors that are at growth and maturity stages as a case base development. These successful enterprise activities were selected from each enterprise sectors. There were totally 50,000 growth and matured enterprise sectors from the total of 100,000 start up, growth and maturity levels of enterprise sectors. But since the data set having different problems such as missing value of majority attributes, redundancy, etc the researcher selects five thousand one hundred thirty three (5133) growth and matured cases from the total of 50,000 cases. Finally these 5133 cases represented as case base in a CSV (Comma Separated Values) format that are used as previously solved cases. From 5133 cases used for this research 2,856 were from manufacturing, 1, 143 were from trade sector, 785 from service sector, 266 from construction sector, 83 agriculture sectors, the number of samples from each enterprise sector depends on the total number of cases in the case base. So the researcher selects the case from each sector proportionally.

KNOWLEDGE BASED SYSTEM (CASE BASED REASONING SYSTEM)

Case-Based Reasoning (CBR) is one of the most successful machine learning methodologies that exploits a knowledge-rich representation of the application domain. Basically, CBR is a problem solving methodology that addresses a new problem by first retrieving a case, which is already solved,
and then by reusing that case for solving the current problem. In the most straightforward application of CBR to recommendation generation, the case base models the products to be recommended and the set of suggested/recommended products are retrieved from the case base by searching for products similar to that is partially described by the user. So, in this recommender system development, previous members’ cases were used as a solved problem and used for the newly solving current selection of enterprise sectors as a solution by retrieving from the case base.

OPPORTUNITY AREAS FOR MSEs IN ETHIOPIA
Ethiopia is one of the countries relying upon MSEs helping as an engine to drive to economic growth, eradication of unemployment and poverty. The failure rate of these businesses in Ethiopia indicates that they are in entire need of assistance. To provide assistance, it is mandatory to identify the constraining factors of MSEs’ growth that they have faced. The key small enterprise sector opportunities for potential members exist in the following areas like: agriculture, agribusinesses and processing, mining and resource development, Infrastructure development, Manufacturing, industry, construction and tourism.

KNOWLEDGE MODELING
After collecting the core concepts from domain experts, members and documents, a model was built by the researcher. The researcher discussed with domain experts to validate whether the knowledge that is acquired from different sources is accurate for enterprise sector and activity selection. Mainly, the selection of enterprise sector and activity has been done by taking consideration of the attributes such as: age, gender, educational level, amount of capital, form of ownership (trade association, cooperation and private), number of members, address (region, zone, and woreda), and enterprise type (micro and small). The prototype is developed based on the model presented in the following hierarchical structure and followed the procedures presented in the hierarchical structure. Finally, the conceptual modeling technique is used to show how enterprise sector and activity selection is performed.

Figure 1: Hierarchical structure of enterprise sector selection
SIGNING THE ARCHITECTURE OF CASE BASED RECOMMENDER SYSTEM

The researcher collects enterprise members’ cases from domain experts, members and document analysis; then several preprocessing tasks such as: selecting attributes, assigning attributes, configuring connectors were performed. After performing all the tasks, it is stored as a general knowledge or case base in order to build the case base recommender system.

As the new query/problem is entered, the prototype of the system matches the new case to the solved case in the case base of the system by using similarity measurement. When relevant cases are found within the case base, then the prototype system ranks the relevant retrieved cases based on their local similarity. Next, it proposes a solution. The similarity between query and existing case is approximate. The proposed solution needs modification (adoption of solution) to fit the new case/query. At the end, the best modified solution should be stored into the case base for future use. The case base updates incrementally when the system learns from new case used by the members.

The proposed solution can be derived directly from a retrieved case that matches exactly or partially to the problem of the new case. Partial match of retrieved cases means some attribute values of the existing case and new cases/query are the same and some attribute values are different. Using the proposed solutions directly may have a risk since some attribute values need of editing/changing based on different conditions. As a result, the user of the system has made an adaptation on the proposed solution having differences between the proposed case and the new case. Additionally, case contradictions are revised because there are situations where previous cases attribute values are not similar with the new case/query attribute values. There is no similarity between the existing case and new case: means there is no previous stored case having similarity with the new case/query in all attribute values. Therefore if there is no similarity between the existing and new case, the proposed solution cannot give recommendation to new cases. In such cases, this new case or problem can be revised and stored in the case base. Finally, the revised solution or stored cases is retained in the case base for problem solving for the next time.

Figure 2: Architecture of case based recommender system for enterprise sector and activity selection
RESULT AND DISCUSSION

DEPLOY THE CASE BASE RECOMMENDER SYSTEM

After defining and configuring all the necessary steps required for the case base recommender system in jCOLIBRI, the following figure shows testing the recommender system application.

Figure 3: Sample query window into the case base for the selection of enterprise sector

USER ACCEPTANCE TESTING

After developing the prototype of enterprise sector and activity selection, user’s acceptance, for applicability of the prototype is evaluated by potential users of the system. The FeMSEDA six (6) domain experts from each enterprise sector which were actively participating from the beginning to the end of the system are evaluating based on the following criteria’s. During testing, experts are requested to rank each parameter by assigning value for poor=1, fair=2, good=3, very good=4, excellent= 5.

Table 1: The system performance evaluation by the domain expert

<table>
<thead>
<tr>
<th>No.</th>
<th>Evaluation criteria</th>
<th>Performance value</th>
<th>average</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is the system is efficient in time</td>
<td>1 2 3 4.3 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Is the system is easy to use</td>
<td>2 2 2 4 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Is the interface is interactive for users</td>
<td>1 2 3 4.3 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Is retrieved cases for decision making are Relevant</td>
<td>3 3 4.5 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Is the system is adequacy and clear for decision support</td>
<td>4 2 4.3 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Is attributes relevant for representing members case</td>
<td>2 4 4.7 94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Is the explanation facility is clear</td>
<td>2 4 4.7 94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Is the system is significant in the domain area</td>
<td>1 5 4 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total average</td>
<td>4.3</td>
<td>87%</td>
<td></td>
</tr>
</tbody>
</table>

As it has shown in the above table, regarding to the efficiency of the system, 16% of the respondents said as good, 34% respondents responded as very good and 50% of the respondents said as excellent. Similarly, use of the recommender system 33% of the respondents said as good, 33% said as very good and 34% said as excellent. In the case of interface interactive, 16% of the respondents responded that good, 34% said as very good and 50% said as excellent. For the relevance of retrieved cases for decision making, 50% of the respondents responded that very good and 50% of the respondents said as
excellent. The clarity and adequacy of decision support, 68% of the respondents said as very good and 32% of the respondents said as excellent. For relevancy of the attributes, 32% of the respondents responded that as very good and 68% and the clearness of explanation facility, 32% of the respondents said as very good and 68% of the respondents responded that as excellent. Finally for significance of the system in the domain area, 16% said as very good and 84% of them said as excellent.

Table 2 The system performance evaluation by the members

<table>
<thead>
<tr>
<th>No.</th>
<th>Evaluation criteria</th>
<th>Performance value</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5 average</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Is the system is efficient in time</td>
<td>1 4 5 3.4</td>
<td>68</td>
</tr>
<tr>
<td>2</td>
<td>Is the system is easy to use</td>
<td>2 2 6 2 4.4</td>
<td>88</td>
</tr>
<tr>
<td>3</td>
<td>Is the interface is interactive for users</td>
<td>1 4 3 2 3.6</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>Is retrieved cases for decision making are Relevant</td>
<td>5 5 4.5</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td>Is the system is adequacy and clear for decision support</td>
<td>2 4 4 4.2</td>
<td>88</td>
</tr>
<tr>
<td>6</td>
<td>Is attributes relevant for representing members case</td>
<td>2 8 4.8</td>
<td>96</td>
</tr>
<tr>
<td>7</td>
<td>Is the explanation facility is clear</td>
<td>3 7 4.7</td>
<td>94</td>
</tr>
<tr>
<td>8</td>
<td>Is the system is significant in the domain area</td>
<td>2 8 4.8</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Total average</td>
<td>4.3</td>
<td>86.5%</td>
</tr>
</tbody>
</table>

As shown in the above table, for the efficiency of the system 10% of the respondents said as good, 40% respondents responded as very good and 50% of the respondents said as excellent. Next, the use of the recommender system 20% of the respondents said as fair, 20% of the respondents said as good and 40% said as very good and 20% said as excellent. In the case of interface interactive, 10% of them said as fair, 40% of the respondents said as good, 30% said as very good and 20% said as excellent. For the relevance of retrieved cases for decision making, 50% of the respondents said as very good and 50% of the respondents said as excellent. The clarity and adequacy of decision support, 20% of the respondents said as fair, 40% of the respondents said as very good and 40% of the respondents said excellent. For relevancy of the attributes, 20% of the respondents said as very good and 80% and the clearness of explanation facility, 30% of the respondents said as very good and 70% of the respondents said as excellent. Finally, for significance of the system in the domain area, 20% said as very good and 80% of them said as excellent. Based on eight close ended questions and five parameters both domain experts and members were evaluated after developing the system. As shown in above two tables, the result is above average (86.5% and 87%) to both members and experts respectively. This indicates users are satisfied by the proposed system.

FINDINGS
- The applicability of case based recommender system for enterprise sector and activity selection haven been proved with evaluations.
- The result of system performance indicated that users are satisfied with proposed system and the performance of the system validation result showed that the system recommends highly acceptable enterprise sector and activities for members.
- To enhance the performance of the prototype case based recommender system of the system, hybrid strategy approaches can be investigated which combines rule based reasoning and case based reasoning. Because, combining rule based and case based reasoning systems can avoid the limitations of both of them.
- The system was developed in English language and is difficult to understand by some members. Further investigation can be conducted by developing the CBRSESAEAS system in different local language. Because, it helps members to communicate using their own language with the case based recommender system.
CONCLUSION

In developing countries like Ethiopia, the advising system on MSEs remains at lower stage. Different factors affect the advising system of enterprise sectors in Ethiopia. These factors include lack of guideline or criteria to assign members in different enterprise sectors, shortage of skilled manpower in the area, lack of consistency advising system, and lack of awareness of members about the purpose of advising systems for the selection of enterprise sectors and activities.

The system aimed to assist both the domain experts and new members in the processes of making proper enterprise sector and enterprise activity selection decisions from already solved cases. In order to, acquire knowledge domain experts, members and documents were used. Mainly previous member’s cases were collected from FeMSEDA. The acquired knowledge was conceptually modeled using hierarchical structure conceptual modeling method. The case representation method that is used in this study is feature value case representation method. Feature value case representation is applied to represent the knowledge before it has been codified using the jCOLIBRI tool. The prototype was developed by using jCOLIBRI 1.1 Programming tool.

Finally, the evaluation result shows that, as retrieval performance of the prototype registers an average value of 86% recall and 64% precision, while its reuse performance registers an average value of 90% accuracy. The user acceptance evaluators (both domain experts and members) assign more than average value for all parameters that are used in the user evaluation form for the prototype. So the average user acceptance evaluation achieved 87% and 86.5 % performance by domain experts and members respectively.

REFERENCES

- Biazen Getnet, “Application of Case Based Recommender System to Advice Students in Field of Study Selection at Higher Education in Ethiopia,” 2013.

Robin Burke, Hybrid Recommender Systems: Survey and Experiments California State University, Fullerton Department of Information Systems and Decision Sciences, 2002.