HYBRID DECISION SUPPORT SYSTEM FRAMEWORK FOR CROP YIELD PREDICTION AND RECOMMENDATION

Alebachew Chiche

College of Computing, Debre Berhan University, Debre Berhan, Shewa, Ethiopia, alebachew.chz@gmail.com

Abstract: In this paper, a hybrid decision support system is presented which uses both quantitative and qualitative data to provide effective and efficient decision making for crop yield prediction and suggestion. Our framework integrates KD-DSS and DD-DSS for solving complex problems by complementing the existing gap of individual decision support system in agriculture domain. For analyzing collected quantitative data of agriculture research center, our framework uses artificial neural network as a data mining technique. So, we use ANN for uncovering hidden knowledge in stored dataset. And this knowledge is further integrated with the knowledge base developed by acquiring qualitative data from expertise and represented using an IF-THEN production rule. The integration of knowledge collected from both qualitative and quantitative source of data provides a potential advantage for solving complex problems for decision makers. Finally, we will have the opportunity to enhance the framework coupling the features which can provide a group knowledge sharing among decision makers. So, this feature can present the opportunities to fill the disparity of decisions made by different decision makers.

1. INTRODUCTION

Agriculture is one of the central area in most of the developing countries in which the life majority of the rural population depends. Enhancing the production is the foremost difficulty facing the sector with the decreasing natural resources necessary for crop production. The continuing growth in demand of agricultural products to bear the living gives opportunities for the producers. Information and communication Technology (ICT) has an important role in improving the livelihoods of the poor rural people by alleviating these challenges [1].

The progress in computing and information processing in today's Internet based society have lead us to a large volume of data. With this truth, there is a challenge to acquire a knowledge from this huge amount of data with a constrained time [2].

In today's information age, every single activity is supported by information communication and technology. One of the field that most of the activities are supported by information technology is agriculture. In a fast-growing computer technology, numerous evolving technologies are becoming very familiar in recent years. One of these emerging technologies are computer-based information systems called decision support system [1].

A Decision Support System (DSS) is a computer-based information system developed to support organizational decision making in their semi-structured tasks [3-4]. In the other way it is defined as Decision Support Systems is a combination of models, data (i.e. database or knowledge base) and a user-friendly management level computer system or software which support a strategic decision making for unstructured and semi-structured decisions. As stated by [3], Decision Support System is designed and implemented to provide support for the decision makers at planning, operation and management levels to make a decision which is varying and difficult to specify. According to Jain [3], a computer based decision support system includes a knowledge based system developed form different sources. It is possible to help decision makers in solving and making a decision with the help of decision support system, collaborative software-

based system, which is designed by using useful information collected from a combination of different documents, tacit and explicit personal knowledge, business models and raw data. So, the presence of different alternative in decision making while the advance in communication networks which provides the possibility to access the global markets and the wide spread of internet and electronic commerce.

In developing countries like Ethiopia, it is difficult and costly to build a specialized agricultural research laboratory. The difficulties are not only from cost perspective for building laboratories but also it is a very challenging concern to get enough number of agricultural experts who can analyze and recommend a possible solution for a given problem. Because of these shortcomings there is no possibilities to analyze and understand the given soil properties for determining the suitability of the given land for particular crop. Moreover, it is difficult to know what kinds of natural and artificial resource are needed for the given land soil type. Though the resource, the laboratories and agriculture experts are accessible, experimenting a sample of land soil is a costly, very tedious, time consuming, and is a lot harsh part of soil selection for better yield crop production. And the consultation acquired from experts obviously vary from one experts to other experts depending on the knowledge and experience of the intended experts.

Agricultural experiment in general and crop yield prediction and soil type recommendation in special are categorized as a complicated agriculture decision making [5]. This complexity of the field is primarily because of the fact that the existence of several variables or factors should be taken in consideration. And also difficult to understand the inter dependency and interaction among these variables and factors. Because of these problematical natures of crop yield prediction and recommendation existed in agricultural problems, they are known as unstructured ill-structured problems [5].

Previous scientific works conducted on Knowledge-Driven DSS depends only on expert's knowledge (i.e. tacit and explicit), on the other hand Data-Driven DSS focus only on the use of hidden knowledge extracted from huge stored data for helping the decision making process. Hence, we need a framework that complements the problem existed in both KD-DSS and DD-DSS for agriculture domain.

From this, we can understand that, the problem in integration of KD-DSS and DD-DSS for agriculture crop yield prediction and recommendation has not been researched yet considerably. It is more significant to analyze soil properties and understand the natural resources required for the soil to get expected crop yield production. And also it has a potential advantage for reasoning the recommendation about the decision.

In this article, we present a framework, a hybrid decision support system to predict crop yield and recommend suitable soil for a specific crop type. The proposed hybrid decision support system comprehends a knowledge based decision support system (KBDSS) and data driven decision support system (DDSS) to provide a strategic decision making process taken by various users such as experts, managers efficiently and effectively. So, the proposed approach interestingly supports a decision making by means of both prediction (i.e. the use of data driven analytical model) and recommendation (i.e. the use of Knowledge based system) techniques.

2. COMPUTER SYSTEM IN DECISION MAKING

Nowadays computer has a vital role in our day to day activities. One of the major role of computer in today convergence enterprises is decision making. There are various number of application areas of computer such as in data processing system, management information systems and knowledge based system in decision making process but here we focus on application of knowledge based system in decision making to handle semi-structured and unstructured problem solving.

2.1 ROLE OF KBS IN DECISION MAKING

As defined by [6-7] a knowledge based system is a computer system used by the system users like experts and manager to support in various problem solving and considerable decision making. Knowledge bases (Expert systems in specific) are a collection of knowledge from different sources based on the problem which the knowledge engineer has defined to help expert and managerial decision making.

According to Szeghegyi [6, 8] in artificial intelligence world, knowledge and knowledge base play an important role in decision making processes in form of expert system, decision support system etc. The wider application of knowledge based systems is in problem solving and decision making than information processing activities. The esteemed author in [6] semi-structured problems can be managed and solved by the use of computer tool called knowledge based systems effectively is most effective in the managing of semi-structured problems. Szeghegyi indicated that the most commonly application of knowledge based system is in activities like strategic planning in managerial level.

According to [3, 8-10], decision support system has four fundamental components incorporated within it. These are: (1) The Model (user criteria and the decision context), (2) The knowledge base, (3) The User interface and finally (4) the Users.

3. CLASSIFICATION OF DSS

There are several classifications of DSS depending on different criteria [11]. As mentioned by Kyungyong [11-12] based on the data and knowledge used for reasoning, Decision Support System is classified into knowledge Driven Decision Support system, Model Driven Decision Support System, Data-Driven Decision Support System, Communication Driven Decision Support System, Web-Based Decision Support System, Document Driven Decision Support System and Hybrid Decision Support System.

Many researchers like in [13-14] asserted that in managerial decision making, integration of analytical results (DSS) and intuitive knowledge of human experts (Expert system) has important advantage to realize problem solving. This idea is attracting the attention of many researchers in recent years. And this integrated system has been commonly termed as Expert DSS (EDSS), Expert Support System (ESS), Intelligent DSS (IDSS), Intelligent Support System (ISS), or Knowledge-Based DSS (KBDSS).

As defined earlier in [3], there are about seven (7) class of decision support system. Most of the decision support systems are proposed and designed by following only one of the classes presented by author [11] among all.

In most of proposed decision support system for agricultural domain for different purposes and decision making processes, one of the major shortcomings is that it only uses either qualitative data or quantitative data for supporting the managerial or expert level problem solving and decision making.

For example, if a specific land soil type is considered as not suitable to a specific crop type, the reason for being unsuitability or inappropriateness of soil for a given crop type or what a soil needs to be a suitable for a crop cannot be justified the details to the users (i.e. experts or decision maker/manager). And no recommendation is included to improve the condition of soil for a given crop.

In the conceptual framework resulted from this work, the expert/decision maker can validate the suitability of a soil in a given section of farming land for required crop type. If the soil type of the given land is valued as "unsuitable" for the required crop, the reason can be explained for unsuitability and forwarded to the user (i.e. expert or manager).

Furthermore, the possibility of providing recommendation to improve the soil properties to come up with a "suitable" soil conditions.

4. HYBRID DSS FRAMEWORK

The Hybrid-DSS framework could be described as a sophisticated solution for complex problems to provide useful information about the significant elements in crop yield cultivation in general. This conceptual framework presents an explanation concerning the settings, modules, process and outputs involved in the proposed framework. Besides, Hybrid-DSS is intended to provide potential Knowledge and information for the decision makers about the specified problem domain. In Fig. 1, the proposed conceptual framework is presented.
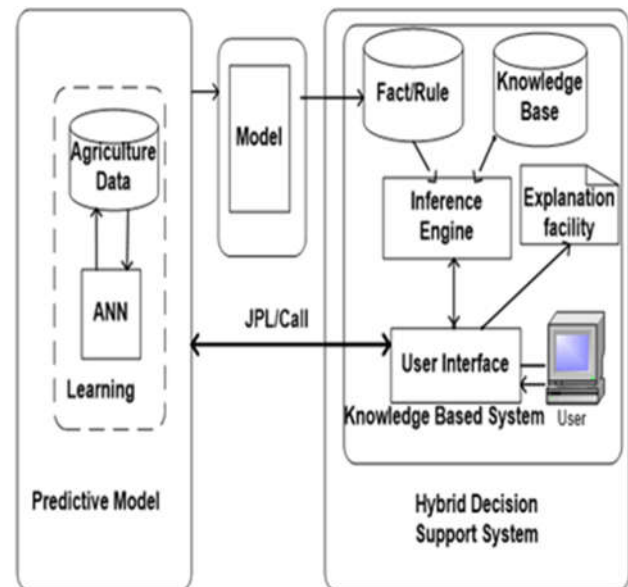

Figure 1 – A Typical Architecture for Hybrid DSS Framework

From the proposed architecture we can understand that the proposed framework encompasses two broad modules: Data driven DSS and Knowledge driven DSS. So, in this section we have discussed the major components, phases, tools, and their roles in two modules.

4.1 DATA-DRIVEN DECISION SUPPORT SYSTEM

Data-driven decision support system are becoming sophisticated with the growing in the use technology [15]. Data-driven decision support today is used for a range of purposes such as operational and strategic decision making process [12]. Now, in emerging big data environment, big data and analytics methods and tools are becoming

commonplace in various fields [16]. Likewise, the fast growth of big data technology and analytics pave a way in emerging of Data-Driven decision making [17]. Brynjolfsson et al [18] has established a data-driven decision making by collecting and analysis of external and internal data of enterprise business practice. Then, author Power [12] presented that, a wide range of operational and strategic purposes are helping with a data-driven decision support system.

With this reality on hand, about the evolving complication of big data environment [16] and its importance in decision making and problem solving activities at present. This could lead the researcher to develop a Decision Support System using Artificial Intelligence tools for knowledge discovery from huge amount of stored data [17]. This feature of the proposed framework can provide both the decisions carried out and an adaptive decision making and problem solving with an automatic learning capability. Therefore, this learning capabilities are realized by the limitless opportunities provided by machine learning techniques.

Consequently, we have followed that incorporating machine learning capabilities, applied on the large data for automatic learning and discover hidden patterns used to predict future uncertainty, using DSS which is center on expertise knowledge for decision making. This incorporated DSS, called Hybrid DSS, so can have the great importance in updating the predictive model and the knowledge base as well. Such an integration of two single DSS can offer an effective and efficient decision making for decision makers in different level. There are prominent artificial intelligence approaches used in complex problem such as genetic algorithms and artificial neural networks used in machine learning applications [13]. For that reason, in this work, we chose artificial neural network (ANN) artificial intelligence tool for data mining task. Finally, we can get the predictive model resulted from machine learning trained on the given agriculture data.

In this module of the research also one has to follow two phases to predict a crop yield of the given soil properties:

Phase # 1: Data Collection and Preparation: In this phase of the research the required data is collected form selected agriculture research center. To understand the collected data, we need to conduct some kinds of discussion with crop farming expertise. After collecting the required data from the selected research site again we need to do activities such as cleaning irregularities, fill missing values and handling some outliers to make ready for further data mining tasks [19]. The duplication of the instances should also in this phase. Finally, in this stage the pre-processed data is converted as Comma

Separated Values (CSV). For testing the framework, we collected 3281 instances of soil dataset. The dataset preparation is part of this work which are carried out for the current study. The datasets have three class labeled as maximum, medium and minimum. For the purpose of this study, soil dataset is collected from north shoe agricultural research center in Ethiopia to explore the impact of soil property on Teff yield production. This soil dataset was a primary data collected from the specified area. This soil dataset consists of information about collected number of instances. So, the dataset has 3281 instances with the selected 10 attributes of soil sample. The following table describes the collected soil dataset.

In the following table the attributes information is described:

Table 1. Type and description of attributes

No.	Soil attributes	Descriptions
1.	Olsen_P	It is an attribute to describe a measure of plant-available phosphate in soil
2.	SoilPH	Soil pH describes a measure of soil acidity and basicity
3.	Initial_P_ppm	Total phosphorus T_P describes the sum of all phosphorus available in different forms.
4.	Total_Nitrogen	Total_Nitrogen is an attribute which describes a sum of all forms of nitrogen nutrients available for plants
5.	Organic_Matter	Organic matter describes the non-mineral content of the sampled soil. Organic matter has a water holding, contribution to soil structure, cycling a nutrient, and CEC functions.
6.	CEC real	CEC real is measured in mill equivalent / 100 grams of sampled soil. CEC describes an electrostatic charge sites measure of the soil. Such as Ca^{2+}, Mg^{2+}, Mn^{2+}, Zn^{2+}, K^+, H^+, and Al^{3+}.
7.	Clay	It describes texture of the soil.
8.	N	NH_4^+ and NO_3^- form of available nutrients.
9.	P2O5	It is an attribute which describe an available phosphorus in plants an estimate of P that is available to plants. And which is an important soil element in plants.
10.	H2O	It describes soil water available and water handling capacity in terms of mm.

As indicated in Table 2 we can understand that class Medium covers the majority number of instances (43.45%) out of 3281 instances. And class Maximum is followed next to class Medium with (27.8%) of share in total number of records. In class Minimum there are the smallest number of instances with (24.75%) of the total instances. So, the class distribution of the dataset shows the masses of the instances are labelled as class Medium (Fig. 2).

Table 2. Share of classes in dataset

No.	Classes	Number of shares of classes	Share of Classes (%)
1.	Maximum	912	27.80
2.	Medium	1556	43.45
3.	Minimum	813	24.75
	Total	3281	100

Phase # 2: Data Mining: In the second phase data mining technique needs to be employed on data which was collected and prepared in the previous phase of the research. In this work, a predictive (supervised learning) data mining [19] is employed to analyze historically stored data. So, artificial neural network (ANN) [12, 20] machine learning tool is applied to uncover new pattern or hidden knowledge which later used for prediction.

In this phase, data mining techniques named Artificial Neural Networks is used for the analysis of soil dataset. So, it focuses on the purpose of predicting Teff crop yield; and make ready for integration with the knowledge base for automated decision support system. The learning process of neural network can be either supervised learning or unsupervised learning mode. As stated in [12] with a supervised learning mode of ANN a training dataset is used for training the algorithm about the given problem. During dataset preparation for neural network, the dataset contains an input and the output which repeatedly feed to the ANN algorithm. Then the output presented by the algorithm is computed and compared to the output anticipated by the user. The following figure (Fig. 3) shows the supervised working process of artificial neural network is presented.
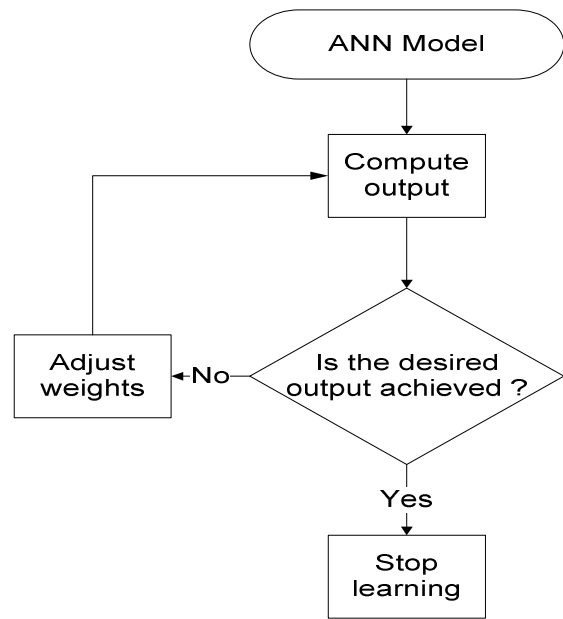

Figure 2 – A Typical supervised learning process of an ANN

4.2 ARTIFICIAL NEURAL NETWORK

The three layer perceptron neural network which consists input, hidden and output layer has been used for modeling the predictive model.

As stated in [12, 20] Artificial Neural Network is one of the very powerful and flexible tool to help agriculture experts to analyze complex soil data across many of applications of neural network. So, it used in crop yield prediction of the given soil type because of it is effective classifier for complex tasks and generality, besides to environment learning.

An artificial neural network is an arrangement of tremendously intersecting elements working in parallel to process the tasks. A collection of processing elements is called a layer within the interconnected network. The first layer which the user input, is the input layer and the layer which is processing result of the algorithm is the output layer. Between the two layers (input and output), there may be another extra layer, which is called hidden layers. The complexity of the data decides the number of neurons in a layer and the number of layers. So, there is no limitation in determining the number of middle layers.

As described above a minimum of three layers is required in an ANN model: the input, hidden and output layers (Figure. 3).
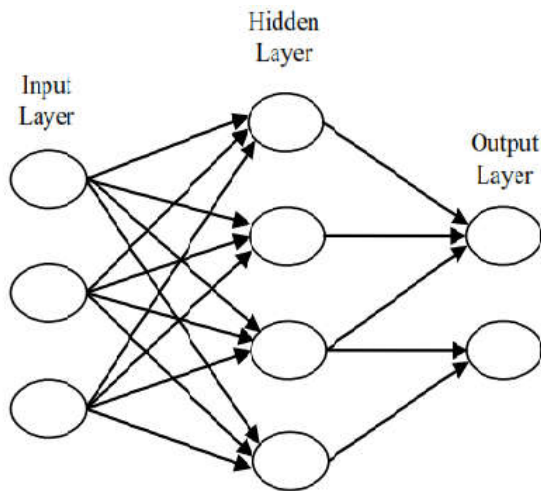
Figure 3 – Artificial Neural Network [12]

Data flow between the layers transversely within some weighted connections. A node which accepts an input data form the previous layers and can compute a weighted sum of all given inputs, t:

$$ti \sum_{j=1}^{n} WijXj,$$

where w is the weight of the connection among the node i and j, n is the number of given inputs, and x stands for the input from node j. And a transfer function is then applied to the weighted value, t, to compute the output for the node, O_i:

$$Oi = f(ti).$$

In this work, Waikato Environment for Knowledge Analysis (WEKA) is used for data mining. WEKA is a machine learning tool for data pre-processing and classification. The results of data mining phase are the model for predicting the grain yield productivity.

Phase # 3: Knowledge Deployment: - In the third and last phase of this module the result of data mining phase can be used for further prediction of crop yield for a specific soil. So, this later incorporated with knowledge base to predict and recommend crop yield for decision makers.

4.3 DATA-DRIVEN DECISION SUPPORT SYSTEM

Decision support system is one of the emerging area which has a sub-specialty of knowledge-based decision support systems (KBDSS), which incorporates a decision support system and expert system that support a broad range of problem solving in an organization [6].

For today's convergence industry knowledge oriented decision support systems are designed by using timely data and information for more precise and effective decision-making [11]. As described in [16-17], by the analysis of both qualitative and quantitative data has been highly applied to solve a wide-range of practical managerial problems and decision making.

Bakari [13] mentioned that few researchers would disagree with the benefit of integrating decision support system and expert system, which is the integrated system termed as Intelligent Support System or Expert Support System which comprises the knowledge of selected experts from organization because of the difficulties faced during the knowledge acquisition of knowledge based system development [21].

According to [21], in recent time plenty of computer systems are integrating modelling, domain knowledge and analysis part of systems to assist the user intelligently. The knowledge base part of the systems used in formulating real-world problems and formulating decision model, and in analyzing and interpreting the outputs. And in most of systems knowledge-based module is used to mimic human decision making. But the majority of managerial decision making have been taken in future uncertainty prediction. As stated in [17], knowledge and data intensive decision making are rapidly growing. As a result, more effort is needed to interpret and use this huge amount of data and knowledge for decision making.

An artificial intelligence tools such as case based learning and reasoning, and machine learning are some knowledge management components which can be included in knowledge based decision support system. So, these tools contribute to the implementation of decision support system to help in strategic decision making and to solve complex problems by learning from stored historic data, past experiences [21].

In KD-DSS part of the research, we have followed three phases of knowledge engineering paradigm which involves in KBS development nowadays to get a recommendation and explanation about crop yield production.

Phase # 1: Knowledge Acquisition: - In the first phase of this module we can acquire tacit and explicit knowledge from different source such as domain experts, secondary literatures etc.

Phase # 2: Knowledge Representation and Modelling: In this work we have used production rule to develop a knowledge base system because production rule is one of the dominant method to represent a huge amount of one domain knowledge. In KBS knowledge can be represented in the form of "situation" -action', i.e. 'IF a certain condition match THEN take a suitable action' [22]. For example:

If Grain Yield = Maximum

THEN Action = provide list of improvement techniques for maximum productivity of soil type AND detail explanation for better decision.

ELSE stop the given crop cultivation on the given soil type

If Grain Yield = Medium

THEN Action = provide list of improvement techniques for maximum productivity of soil type AND detail explanation for better decision

ELSE stop the given crop cultivation on the given soil type

If Grain Yield = Minimum

THEN Action = provide list of improvement techniques for maximum productivity of soil type AND detail explanation for better decision

ELSE stop the given crop cultivation on the given soil type

Phase # 3: Use of knowledge: - In this phase of the research the decision makers can use the knowledge and information provided by the integrated DSS through the user interface. To use the knowledge and information generated from ANN [20] learning, Java Prolog Library (JPL) tool is employed. This tool is concerning about connecting both analytical (machine learning result) and intuitive modules (reasoning and explanation) of the DSS. So, the decision maker can automatically invoke analytical model while he/she use the user interface of knowledge based system.

Fig. 1 shows the typical architecture of Hybrid DSS demonstrating its main modules and subsystems. As described in section 4., depicted in Fig. 1, the framework of hybrid DSS is a construct of two main Modules, the analytical (i.e. the data Driven DSS) and the intuitive module (i.e. Knowledge Driven DSS). The Knowledge Driven part of the hybrid DSS can be implemented using swi-prolog. Since prolog is one of the leading backward reasoning logical programming language, which can represent the expertise knowledge using first order predicate logic. And also WEKA classifiers are java oriented development environment we can use JPL (Java interface to prolog), a bidirectional java prolog library, to integrate and invoke the analytical part of the DSS.

As a result, the integration of the data Driven DSS modules and Knowledge Driven DSS resulted from JPL can be usable in one interface.

5. EXPERIMENTAL RESULTS

As discussed above the main goal of the Hybrid Decision Support system is to support a decision maker in complex decision making and problem solving activities which can be considered as a decision maker's problem. The data driven decision making process of the approach is used to predict the grain yield of the soil type as Maximum, Medium and Minimum. This prediction of crop yield has been done by artificial neural network.

In this approach the following steps are followed to construct a predictive model for decision support systems.

Algorithm: ANN modeling for Data Driven Decision Support System

Input: Agriculture dataset

Output: Prediction of Grain Yield

　　Step 1: Load the Agri_dataset

　　Step 2: Apply pre-processing tasks

　　Step 3: Select the best set features using feature subset selection

　　Step 4: Select Tenfold Cross validation evaluation technique

　　Step 5: Agri_dataset fed to ANN for training and testing

　　Step 6: Compute Detection accuracy, TPR, FPR, F-Measure, Recall

In this study WEKA machine learning tool was used for our experimental analysis. As discussed in section 4.1.1 we collected and prepared an agricultural dataset used for this experimental analysis. In this dataset 3281 instances and 10 attributes were included which consists of three (3) classes. For this we tested artificial neural network algorithm on the dataset.

In this work the following parameters are evaluated to measure the performances of the model using 10 fold cross validation evaluation mode [19]. The main reason to use tenfold cross-validations instead of other validation techniques is that there is not enough data available to partition it into separate training and test sets. In cross validation the data is split into some number of partitions of the data, in most cases it is segregated into 10 equal portions, and each of these portions of the dataset is used for testing purpose in the same time the remaining portion of the dataset is used for training purpose. This process of tenfold cross-validation evaluation method is repeated 10 times. Finally, we can consider the mean result of tenfold cross-validation.

Despite the existence of the criteria for measuring performance of the prediction model TP Rate, FP Rate, Precision, Recall, F-Measure and MCC was used for the study. So, the experimental results acquired from WEKA or ANN model can be analyzed and interpreted with respect the following measuring aspects as depicted in Table 3 ANN model achieves an average of 92.9% TP Rate, 4.3% FP Rate 94.5% of precision, 92.9% of recall, 93.5%

of F-Measure and 88.1% of MCC on the given dataset using 10-fold cross-validation testing mode. Due to a large number of instances labelled as Medium, we found that ANN model score better TP Rate than Maximum and Minimum, a fact which relates to an imbalanced share instances in each class.
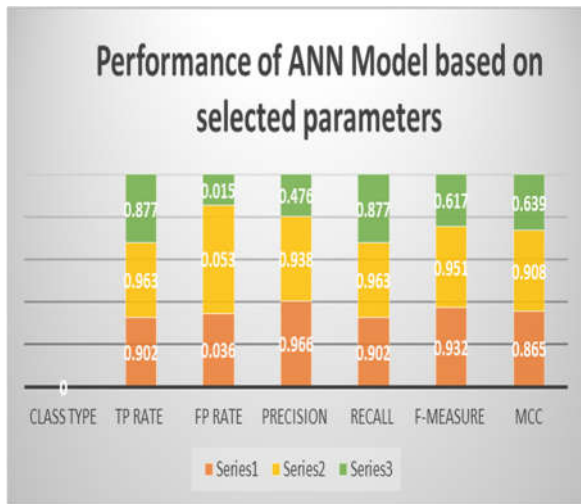

Figure 4 – Performance of ANN Model based on evaluation parameter

As stated in this section before the purpose of the study was to design and test the new framework that can be used for predicting the grain crop yield of the soil sample, interpret the result and suggest a possible course of actions either for improving the grain yield or productivity of the crop or recommend the suitable crop for the given soil type in agriculture problem domain. Specifically, the study was examined the effect of sampled soil properties that directly or indirectly contributes to the productivity of the crop using neural network machine learning algorithm. So, ANN was applied for our data driven part of the experiments to predict grain crop yield using the agriculture dataset. ANN predictive model correctly predicted about 3,046 instances out of the total 3281 instances while the predictive model incorrectly predicted about 235 instances. Table 4 shows that the average accuracy of ANN model using 10-fold cross-validation mode of evaluation on given dataset.

Table 3. An Average Accuracy of Ann Model on 10-Fold Cross-Validation (%)

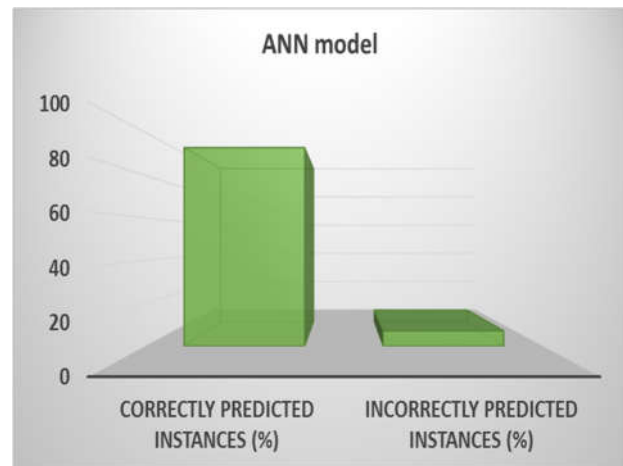| Algo-rithm	Correc-tly predic-ted instan-ces (%)	Incorrec-tly predic-ted instan-ces (%)	Correctly predicted instances (No)	Incorrectly predicted instances (No)
ANN	92.86	7.14	3,046	235

Figure 5 – Performance of ANN Model based

On the use of tenfold cross-validation mode, ANN model predicted the soil grain yield or productivity of crop on sampled soil with 92.86 % average detection accuracy (which correctly predicting 3,046 instances out of 3281 total instances) on the given dataset. On the other hand, ANN predictive model predicted 7.14 % error rate (the model incorrectly predicted 235 instances out of 3281 total instances) on the given agriculture dataset.

Table 4. Confusion Matrix for Framework

Actual class	Model predicted value		
	Maximum	Medium	Minimum
Maximum	822	46	44
Medium	56	1498	2
Minimum	8	93	712

Confusion matrix was used for evaluating the performance of the framework. Based on this the framework can predict 822 instances correctly as maximum and 46 instances and 44 instances incorrectly as Medium and Minimum respectively but expected to be Maximum. Next to this, it predicts correctly 1498 instances as Medium, incorrectly 56 instances as Maximum, 2 instances as Minimum. And the framework also predicts 712 instances correctly as Minimum, incorrectly classify 93 instances and 8 instances as Medium and Maximum respectively.

From the confusion matrix we can understand that the proposed framework can both predict and provide an appropriate recommendation for predicted grain yield of the crop.

Empirical results of the framework depict that the ANN predictive model gives better predictive accuracy within the framework for predicting grain yield of the crop as Maximum, Medium and Minimum. Then we can ask the knowledge base for further information and interpretation on the

prediction by the predictive model. So, based on the ANN model prediction value on the knowledge base interface, the knowledge base automatically provides options for explanation about the prediction result, Thereby, the decision maker can use the knowledge base for further interpretation in recommending suitable decisions for improvement and change in either soil type or crop type.

The findings of this study show that decision makers in either at tactical or strategic level could analyze and infer the above research results for predicting grain yield productivity of a crop on sampled soil dataset. In the mean while deciding the input needed for improving soil fertility and what types of crop are suitable for sampled soil type are significant concerns to develop tactical and strategic plans. The empirical results presented by the new decision support system can be advantageous in a wide-ranging category of decision making process such as health and monitoring, software product evaluation and so on.

Policy makers and experts could also use the findings of this kinds of study to predict which soil type is suitable for a particular crop type cultivation and which soil type will strongly unsuitable for particular crop cultivation.

6. CONCLUSION

In this study, we present a conceptual framework of Hybrid-DSS consists KD-DSS and DD-DSS for the potential information and knowledge about crop yield prediction and suggestion in decision making. This approach has a potential advantage over the former works to provide a solution toil-structure and semi-structured problem of proper crop yield farming because of the integration of both qualitative and quantitative data together.

This article presents the required components and steps for decision making in crop yield cultivation as a general way. We identify the necessary parts of hybrid-DSS involves in decision making process. Additionally, we forward a typical Hybrid-DSS architecture possibly integrates and organized several modules and tasks.

Because of the integration, all the modules and tasks in one single system that is accessible with a single interface. For constructing a predictive model on the stored historic data, we chose artificial neural network because of the complexity nature of soil dataset. For representing and developing collected agricultural knowledge base we chose If Then production rule using SWI-PROLOG. Based on the experimental result the proposed framework achieves a better prediction accuracy with 92.86 percent on 3281 number of instances collected from a sampled agricultural land use. So, this result

articulates that the framework is encouraging works in agricultural problem domain. This framework allows development of better decision Support System to increase the reasoning power in problem solving and Decision making process. In this study weaknesses are noticed in instructiveness of interface for using the decision support systems. In the future, better and interactive interface can be used as a solution to simplify problem solving.

7. REFERENCES

[1] J. Stienen, W. Bruinsma, F. Neuman, *How ICT can Make a Difference in Agricultural Livelihoods Information & Communication Technologies*, The Commonwealth Ministers Reference Book, 2007, pp. 2-4.

[2] L. Armstrong, D. Diepeveen, and R. Maddern, "The application of data mining techniques to characterize agricultural soil profiles," *Proceedings of the sixth Australasian Conference on Data mining and analytics AusDM'07*, Darlinghurst, Australia, 2007, pp. 85-100.

[3] R. Jain, *Decision Support Systems: An Overview*, Decision Support System in Agriculture using Quantitative Analysis, Udaipur, Agrotech Publishing Academy, 2016, pp. 42-50.

[4] P. Zarate, S. Liu, "A new trend for knowledge-based decision support systems design," *Int. J. Information and Decision Sciences*, vol. 8, no. 3, pp. 305–324, 2016.

[5] B. Manos, Th. Bournaris, J. Papathanasiou, Ch. Moulogianni and K.S. Voudouris, "A DSS for agricultural land use, water management and environmental protection," *Proceedings of the 3rd IASME/WSEAS Int. Conf. on Energy, Environment, Ecosystems and Sustainable Development*, Agios Nikolaos, Greece, July 24-26, 2007, pp. 340-345.

[6] A. Szeghegyi, "Investigation of Decision-Making Process by the Use of Knowledge-based System," *Proceedings of the 5th International Conference on Management, Enterprise and Benchmarking*, Budapest, Hungary, June 1-2, 2007, pp. 209-222.

[7] S.P. Anan, "Importance of knowledge base in decision making processes," *International Journal of Innovations & Advancement in Computer Science*, vol. 6, no. 5, pp. 217-221, 2017.

[8] A. Agnar, "A case-based answer to some problems of knowledge-based systems," *Proceedings of the Scandinavian Conference on Artificial Intelligence. IOS Press*, 1993, pp. 168-182.

[9] G.M. Maraka, *Decision Support Systems in the Twenty-First Century*, Upper Saddle River, London: Prentice Hall, 2003.

[10] S. Haag, *Management Information Systems for the Information Age*, Arlington: McGraw-Hill, 2006.

[11] C. Kyungyong, B. Raouf, H. Salim, "Knowledge based decision support system," *Inf. Technol. Manag.*, vol. 17, pp. 1-3, 2016.

[12] D.J. Power, *A Brief History of Decision Support Systems*, 2007, [Online]. Available: http://DSSResources.COM/history/dsshistory.html

[13] B. Nahim, J.A. Aguilar, A. Perez Pulido, N. Fernandez, "Composition intelligent frameworks," *Proceedings of the Conference on Artificial Intelligence and Agents*, 2017, pp. 1-9.

[14] F. Witlox, "Expert systems in land-use planning: An overview," *Expert Systems with Applications*, vol. 29, no. 2, pp. 437-445, 2005.

[15] D.J. Power, "Understanding data-driven decision support systems," *Information Systems Management*, vol. 25, pp. 149–154, 2008.

[16] D. Laney, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, 2001, [Online]. Available: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

[17] C. Brandas, C. Panzaru, F.G. Filip, "Data driven decision support systems: an application case in labour market analysis," *Romanian Journal of Information Science and Technology*, vol. 19, no. 1-2, p. 65–77, 2016.

[18] E. Brynjolfsson, L.M. Hitt and H.H. Kim, "Strength in numbers: how does data-driven decision-making affect firm performance?" in *Thirty Second International Conference on Information Systems (ICIS),* Shanghai, 2011, pp. 1-18.

[19] J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, San Francisco, CA, itd: Morgan Kaufmann, 2012.

[20] M.D. Odom, R. Sharda, "A neural network model for bankruptcy prediction," *Proceedings of the IEEE International Joint Conference on Neural Networks IJCNN'1990*, vol. 2, pp. 163-168.

[21] A. Ultsch, D. Korus, "Integration of neural networks with knowledge-based systems," *Proceedings of the IEEE Int. Conf. Neural Networks*, Australia, 1995, pp. 1-6.

[22] R.V. Ramani and K.V.K. Prasad, "Applications of knowledge based systems in mining engineering," *Proceedings of the Twentieth International Symposium on the Application of Computers APCOM'87*, Johannesburg, 1987, pp. 17-180.

Alebachew Chiche, *has received his BSc. in Information Systems from Hawassa University in 2012 and MSc. in Computer Networking from Jimma University in 2016. He is currently working as lecturer in College of Computing, Debre Berhan University since 2016. His current research interests include Network Security, data mining, and artificial intelligence and information systems.*