Integration of Data Mining with Knowledge Based System for Diagnosis and Treatment of peppercorn Crop

Melsew Belachew [1]
University of Gondar
mlswblchw@gmail.com

Muluneh Atinaf [2]
Addis Ababa University
mulunehatinaf@yahoo.com

INTRODUCTION & MOTIVATION

Agriculture is the main source of economy and the means of employment for more than 82% of the population of Ethiopia. In developing country farmers did not get disease advisory system about their crops. Peppercorn crop is the most important cash crop. The sector is sensitive distinctly affected with the aid of climate condition, poor agronomic practice and other factors.

Today, agricultural organizations work with large amounts of data. Processing and retrieving of significant information in this abundance of agricultural data is necessary. Utilization of data mining technology enables extracting significant patterns and knowledge from the data set. Data Mining is important to discovering and analyzing data from different perspectives.

Integrating data mining with knowledge base system is essential to deploy the knowledge extracted from data mining models. This used to utilizing interesting and previously unseen knowledge extracted from data mining models for knowledge base system.

In Ethiopia, farmers crop are continuously vulnerable by a range of pest and disease problems. Among those problems:-

- Irrational use of pesticides among the farmers.
- Unable to diagnosis and treat peppercorn crop disease.
- Shortage of domain experts
- Limitation of knowledge on agricultural workers.

To tackle this problem, we integrate knowledge based system with data mining to diagnosis and treat peppercorn crop disease. To addressed the problem, we design the following research questions.

- How to extract knowledge from peppercorn crop data set?
- How could DM and KBS techniques integrated to diagnosis peppercorn crop?
- Dose the advisory system, diagnoses and treat peppercorn crop disease?

OBJECTIVE

- To create predictive model from peppercorn crop dataset for extracting knowledge
- To build integrator that automatically acquire knowledge from the predictive model
- To construct Knowledge base system using rule-based knowledge representation
- To design advisory system for diagnosis and treatment of peppercorn crop disease
- To evaluate the performance of the advisory knowledge based system.

METHODOLOGY

- **Research Design:** Empirical type of research design
- **Hybrid process model** is selected. One such model is a six-step KDP model.
  - Rule based approach
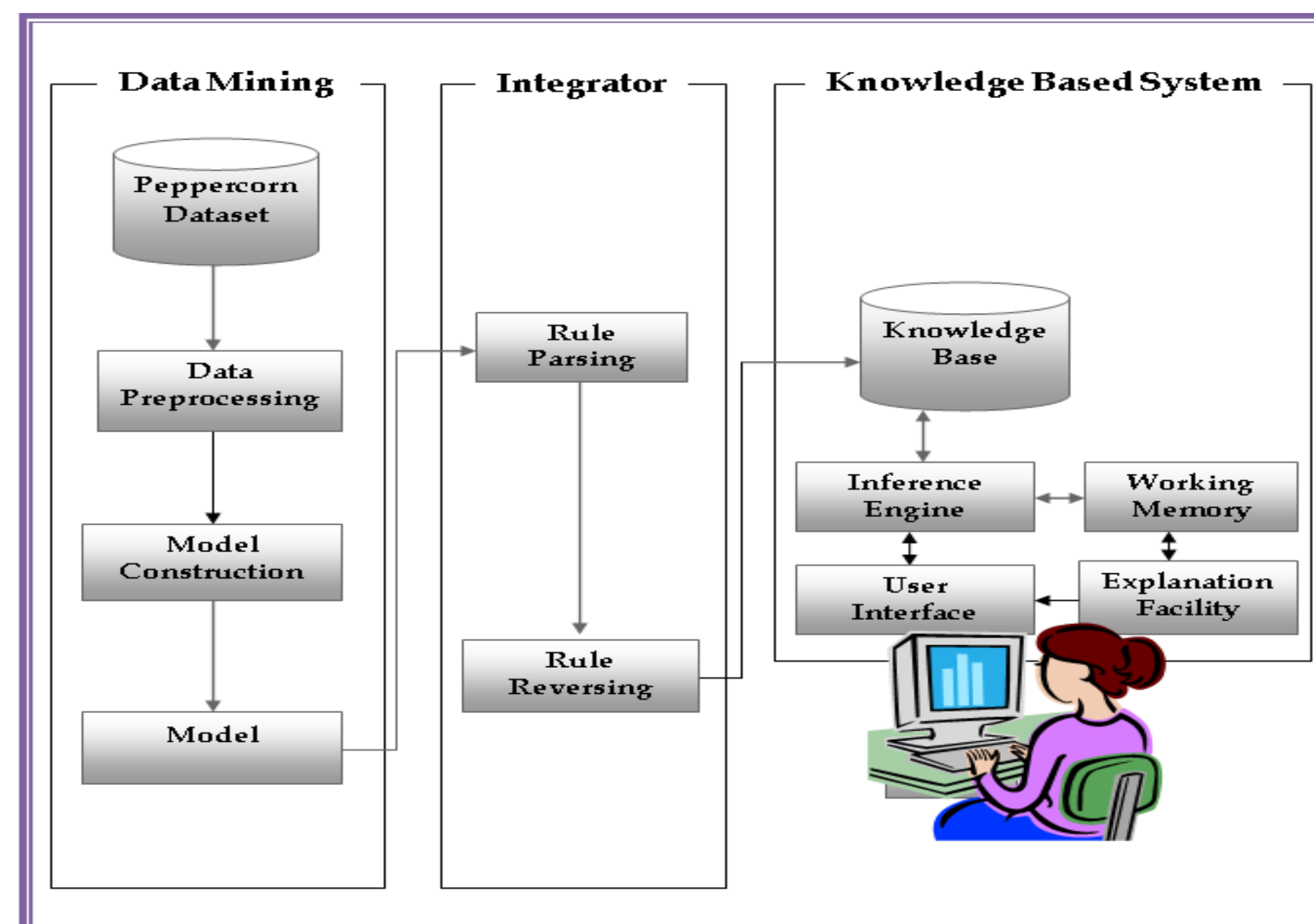 - Classification Algorithm

ARCHITECTURE DESIGN & INTEGRATION

Fig:1 System Architecture of the proposed Integration system for Peppercorn crop diagnosis and treatment
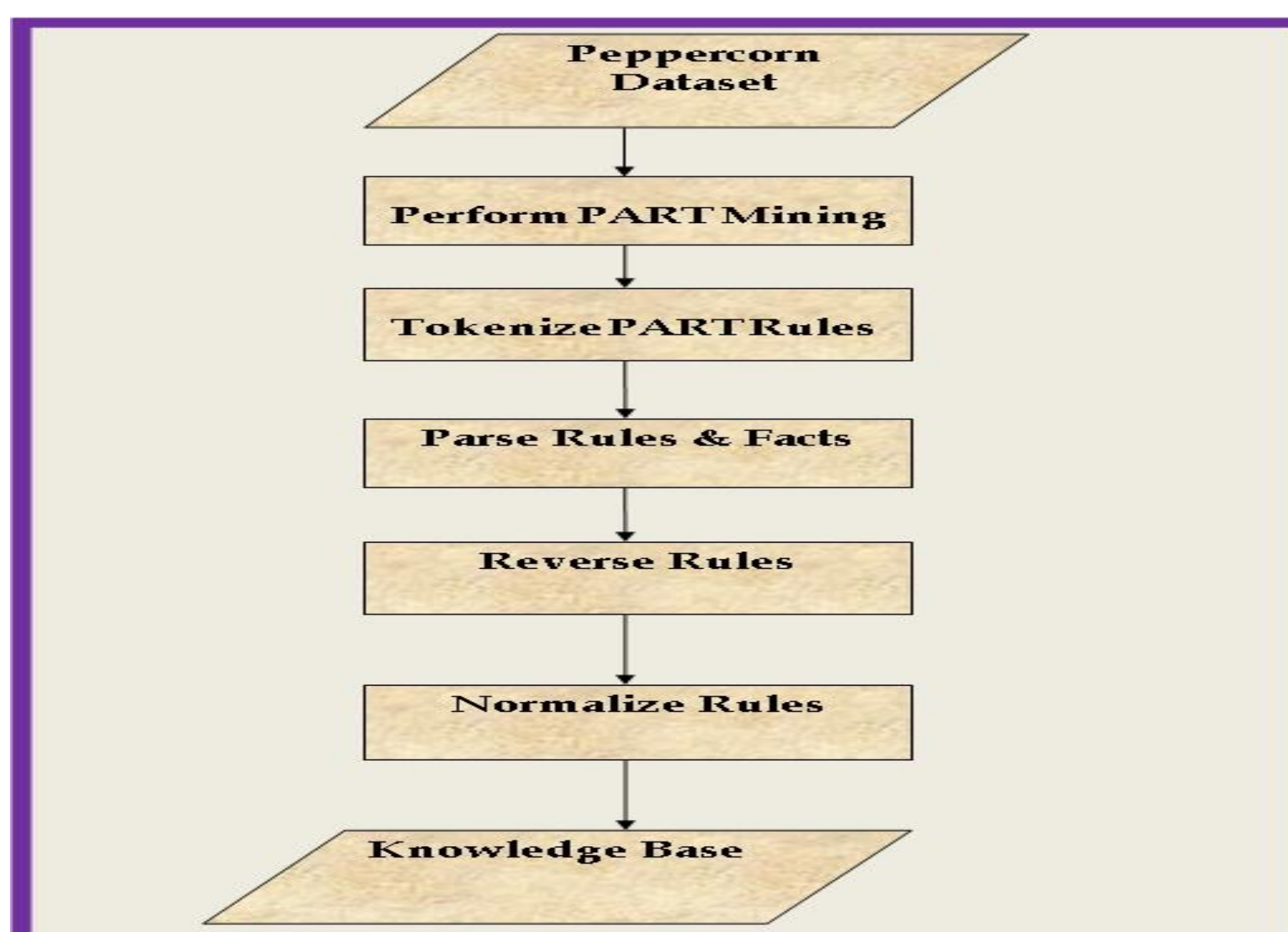

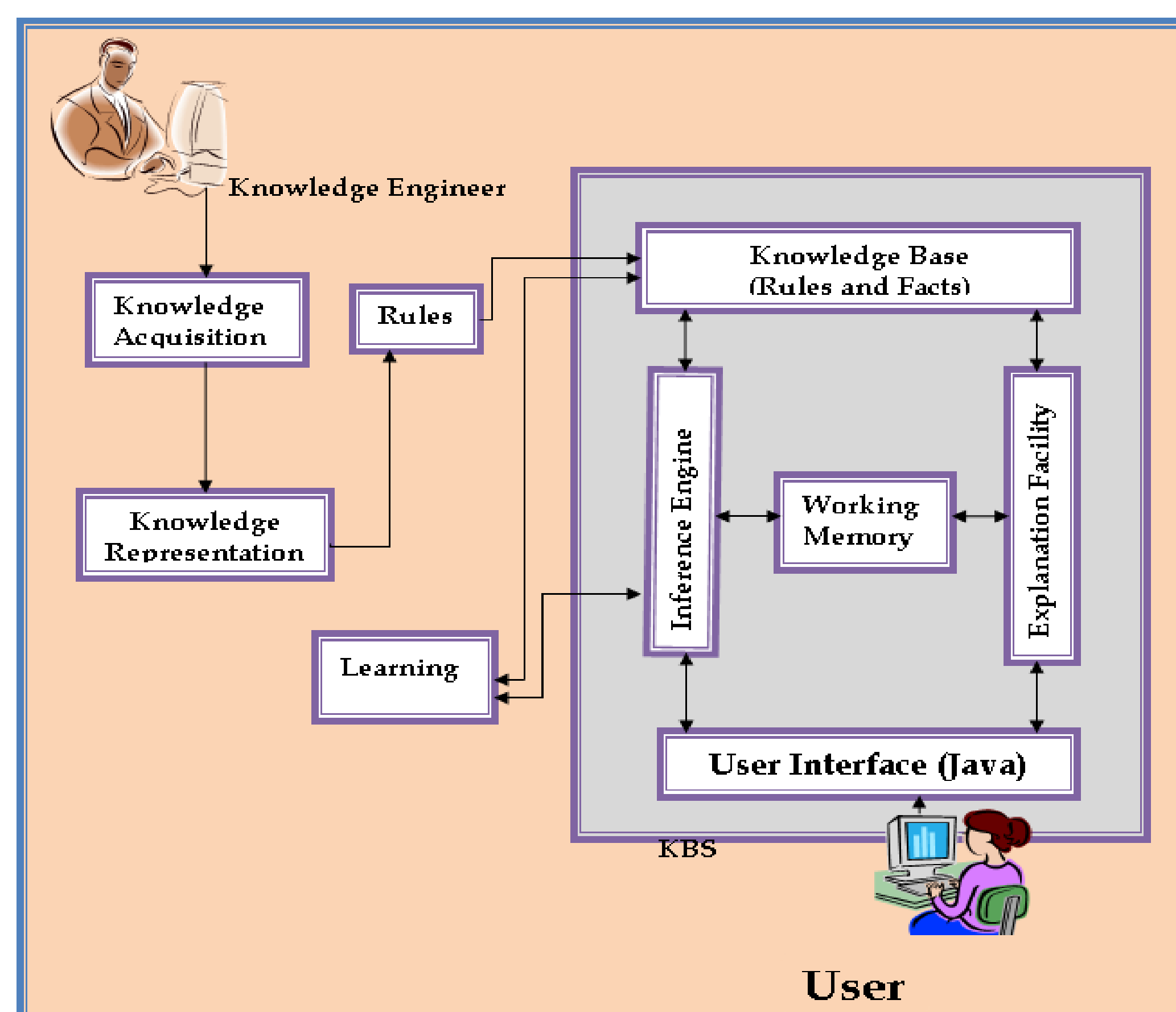**Fig:2** Work flow diagram for Rule mapping from PART to prolog.

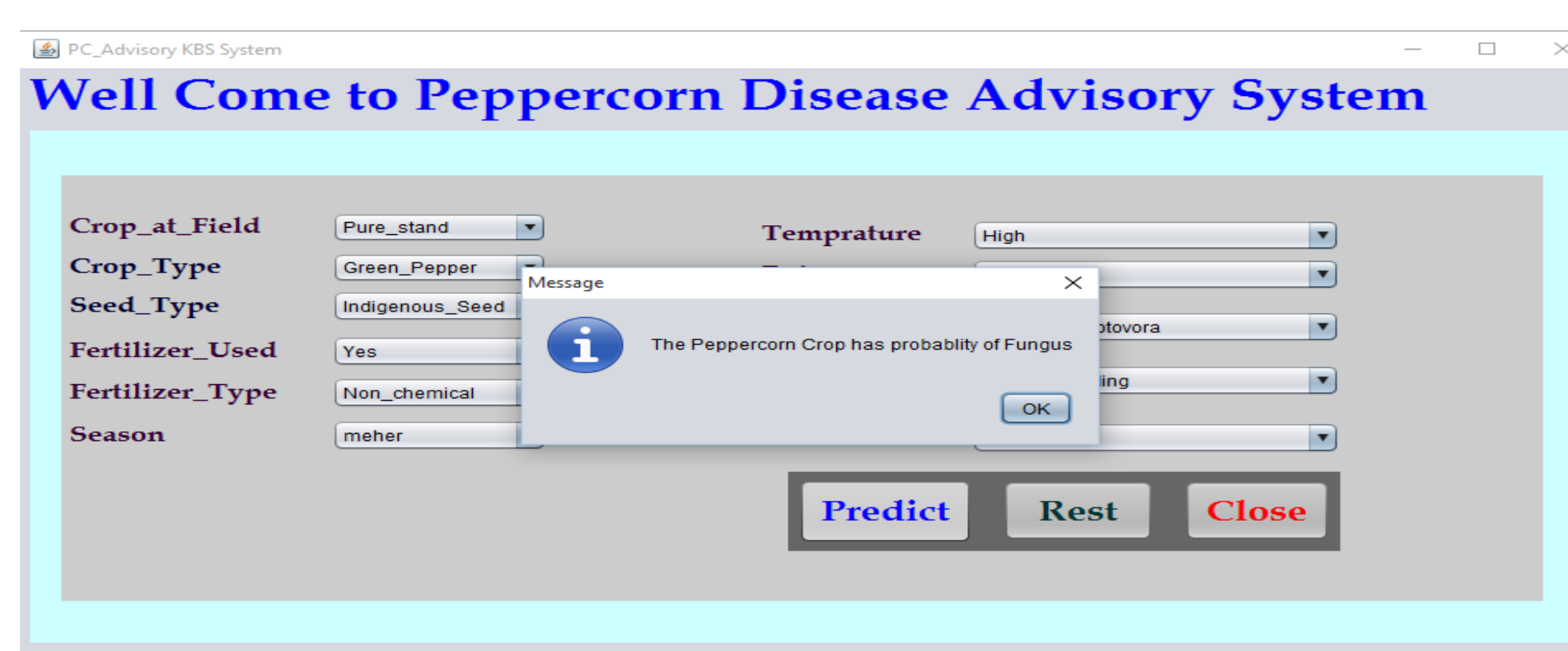**Fig:3** System Architecture of PCA-KBS for Diagnosis and Treatment of Peppers

Fig:4 User Interface for PCA-KBS

EXPERIMENTS & RESULTS

A total of four experiments (JRip, PART, J48 and REPTree) are aimed to building predictive models.

Performance Comparison of Classifiers

Classifier	Correctly classified instances		Incorrectly classified instances		Time take to build model
	No.instances	Percentage	No. instances	Percentage	Second
J48	9539	96.09%	388	3.91%	0.06
REPTree	9615	96.86%	312	3.14%	0.33
PART	**9664**	**97.35%**	**263**	**2.65%**	**0.91**
JRip	9628	96.99%	299	3.01%	3.47

Performance of classifiers with respect to classes

Classifier	Class	Fungus	Insect	Virus	Bacteria
J48	Precision	97%	93.4%	96.7%	100%
	Recall	99.9%	100%	96.9%	84.8%
	F-measure	98.5%	96.6%	96.8%	91.8%
	TP Rate	99.9%	100%	96.9%	84.8%
PART	**Precision**	**97.3%**	**96.4%**	**96.8%**	**99.9%**
	Recall	**100%**	**99.9%**	**99.9%**	**88.1%**
	F-measure	**98.6%**	**98.2%**	**98.3%**	**93.6%**
	TP Rate	**100%**	**99.9%**	**99.8%**	**88.1%**
REPTree	Precision	97.2%	96.2%	96.6%	97.9%
	Recall	99.1%	99.3%	99.4%	88.2%
	F-measure	98.1%	97.7%	98%	92.8%
	TP Rate	99.1%	99.3%	99.4%	88.2%
JRip	Precision	97.5%	96.1%	96.3%	98.9%
	Recall	99.5%	99.1%	99.6%	88.5%
	F-measure	98.5%	97.6%	97.9%	93.4%
	TP Rate	99.5%	99.1%	99.6%	88.5%

- Based on experiment result, PART is the best classification algorithm to recommend the peppercorn crop disease type. The algorithm generates a model with 27 rules.
- This study find out that: disease, cause, fertilizer used, fertilizer type, seed type, and symptoms are the main factors that determine peppercorn crop disease type
- We used Precision, Recall, F-measure, and True Positive rate to evaluate the results and accuracy of the data mining model.
- To evaluate the performance of this study, system performance testing and user acceptance testing were performed.
- According to the agricultural domain experts evaluation 90.5% accuracy were registered for system performance testing and 86.8 % performance registered by user acceptance evaluation

CONTRIBUTIONS

- It used to identify basic crop behaviors and disease factors.
- It created a general understanding of DM with KBS integration process.
- It helps for agriculture workers to treat and diagnosis peppercorn crop disease.
- Motivate future researchers to work on integrating data mining models and knowledge based system in other fields of studies.

FUTURE WORKS

- Enhance and expand the capabilities of the prediction algorithms, by further parameter setting is required.
- Build hybrid knowledge based system which is capable of employing case based reasoning and rule based with integration of data mining techniques.
- Design Mobile based application for Self-diagnosis at any time and any place that provides an option of local languages.
- Apply integration of DM with KBS disease diagnosis system in other crops rather than peppercorn crop

REFERENCE
[1] Berhanu Aebissa. "Developing Knowledge based system for coffee disease diagnosis and treatment". MSC Thesis, Addis Ababa University, Ethiopia, 2012.
[2] Abdulkerim Mohammed. "Towards Integrating Data Mining with Knowledge Based System: The Case of Network Intrusion detection" Addis Ababa University, Addis Ababa, M.Sc. Thesis, June, 2013.
[3] TesfamariamMulugeta. "Integrating Data Mining Results with the Knowledge Based System for Diagnosis and Treatment of Visceral Leishmaniosis". MSC Thesis, University of Gondar, Ethiopia, 2015.
[4] Federal democratic republic of Ethiopia Central statistical agency agricultural sample survey: report on area and production of major crops: volume I, 2014 / 2015 (2007 E.C).
[5] Ejigu Tefera. "Developing knowledge based system for cereal crop diagnosis and treatment: the case of kulumsa agriculture research center". MSC Thesis, Addis Ababa University, Ethiopia, 2012.
[6] Bethlehem Asferi. "Exploring Pest Management Practices and Development of Knowledge Base System for Pepper Disease Diagnosis". MSC Thesis, Addis Ababa University, Ethiopia. August, 2010.
[7] Abdi Chali. "An Integration of Prediction Model with Knowledge Base System for Motor Insurance Fraud Detection: The Case of Awash Insurance Company S.C" Addis Ababa University, Addis Ababa, M.Sc. Thesis, February 2016.