Crop Darpan: Engineering crop diagnosis framework by extending the concept of generalization/specialization - A prototype for Cotton crop

Thesis submitted in partial fulfillment
of the requirements for the degree of

MS by Research
in
Computer Science and Engineering

by

Narendra Babu Unnam
201507602
narendra.unnam@research.iiit.ac.in

Data Sciences and Analytics Center
International Institute of Information Technology
Hyderabad - 500 032, INDIA
February 2019

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled " **Crop Darpan: Engineering crop diagnosis framework by extending the concept of generalization/specialization - A prototype for Cotton crop**" by **Narendra Babu Unnam**, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Adviser: Prof. P. Krishna Reddy

To
Family and Friends

Acknowledgments

As I submit my MS thesis, I would like to take this opportunity to acknowledge all the people who helped me in my journey at IIIT-Hyderabad.

Firstly, I would like to express my gratitude to my guide, Prof. P. Krishna Reddy, for his constant support, guidance, and understanding throughout my thesis work. This work would not have been possible without him. I want to thank him for his patience and encouragement while I worked on various problems before landing on the current work. His knowledge and advises have always helped whenever I faced a problem. He always makes time for brainstorming sessions of the research. His continuous review process before submitting every research paper made me understand the way of presenting ideas in the paper. Finally, I also want to thank him for his support during the highs and lows of my journey at IIIT Hyderabad.

I want to thank my fellow labmates A. Srinivas, A. Mamatha, P. Revanth, R. Akhil, C. Saideep, and V. Satya krishna at Data Sciences and Analytics Center (DSAC), IIIT Hyderabad for their immense help towards my research work. The learning environment created in the lab with these people made my work easier and interesting. I would like to appreciate the efforts of field coordinators Prakash and Raji Reddy for their support in the field data collection. I would like to especially thank Prof. K. Narayana Reddy, Dr. G. Sreenivas Gade, Dr. Balaji Naik from Professor Jayashankar Telangana State Agriculture University (PJTSAU)for their exceptional help in creating domain knowledgebase. I want to thank all my friends for making my life peaceful and memorable at IIIT Hyderabad.

Finally, I would like to express my gratitude to my parents Jagan mohan rao and Rama devi, and my brother Nava krishna for their unconditional love, support, and continuous motivation throughout my life.

Abstract

Despite the variety of agricultural extension approaches, which include information technology based approaches, the majority of farmers in India are not acquiring actionable agricultural information. Farmers are facing difficulty in acquiring actionable agricultural advice in a real-time manner from call centers and web portals due to communication and perceptual issues. The radio, video, SMS and voice-based services push generic information to farmers. The eSagu variants like farm/location specific advisory systems suffer from generality/scalability issues. To enable the farmer to get the actionable agro-advisory in a realtime manner, we have made an effort to build a smartphone-based Crop Disease Diagnosis Guide, which is called as "Crop Darpan", by exploiting the progress in data science and mobile phone technologies.

The basic idea of Crop Darpan is as follows. A crop disease in the field can be identified by confirming the presence/absence of a set of visual symptoms. The proposed Crop Darpan contains a hierarchy of visual symptoms of field diseases that consist of generalized symptoms to fine-grained symptoms. It is assumed that the farmer possesses a smartphone with an Internet connection and visits the field. The Crop Darpan helps the farmer to identify the field problem by confirming the corresponding symptoms of the crop disease and acquire the agro-advice. A prototype was built for a Cotton crop (www.cropdarpan.in) in English and Telugu languages.

The research issue was to build a system, which resembles the process followed by an agricultural expert to detect the crop disease. This requires the embedding of experts knowledge about the visual symptoms of all diseases into the system. The components of the Crop Darpan system are Knowledge acquisition protocol, Knowledge base, Question popping model, and the User interface. We have explored three design choices: (i) number of questions to be asked per screen, (ii) algorithms followed for knowledge acquisition protocol from the domain expert, and (iii) question popping algorithm to pose questions to the farmer. The development of Crop Darpan has been carried out by following engineering research methodology. Overall, five design frameworks were developed in an incremental manner. The fifth and the final framework, which is called as Hierarchy-based protocol for knowledge acquisition and Hierarchy model for question popping with Multi-questions per screen (HHM) framework, has been conceived.

As a part of the Knowledge acquisition protocol, we have developed a protocol to collect both generalized and specialized symptoms for a given crop from agricultural domain experts. The question popping model has been developed for organizing the collected knowledge database in a hierarchical

manner. An algorithm was proposed to pose an appropriate sequence of visual symptom related questions, starting from generalized to specialized visual symptoms.

The Crop Darpan system was built and operationalized in the field for Cotton crop, It was found out that the farmers are able to identify the field problem using the Crop Darpan system. The evaluation study carried out with domain experts and farmers show that the domain experts are in agreement with diagnosis methodology followed by the system and the educated farmers are also able to diagnose most of the Cotton crop diseases. The proposed framework is a scalable and generic system. It is extendable to multiple crops, regions, and languages/dialectics.

Contents

List of Figures

List of Tables

Chapter *1*

Introduction

Farming or Agriculture is the basic source of food supply and an important sector in the economic development and poverty alleviation drive of many countries. Agriculture provides food and raw material to non-agricultural sectors of the economy. It is an important source of livelihood and provides employment opportunities for rural people on a large scale in both underdeveloped and developing countries. In most countries, agriculture is a way of life.

By 2050, the world's population will exceed nine billion and to meet the increased demand, agricultural production will have to grow by 60% [15]. Achieving this significant increase in agriculture productivity is going to be a major challenge, made more difficult by the impacts of climate change and water scarcity and the growing competition of land and agricultural resources. These challenges can only be overcome when farmers, the players in agricultural productivity are supported by respective stakeholders who can guide, and support them, resulting in addressing their areas of need.

Rising agricultural productivity has a significant effect on the economic development of a country. Efforts are being made by governments to improve agriculture to create an eco-system for better crop production, in the form of improved seeds and crop production methodologies. Agricultural extension is the application of scientific research and knowledge to agricultural practices through farmer education. Generally, an agricultural extension can be defined as the delivery of information inputs to farmers [17]. The role of extension services is invaluable in teaching farmers how to improve crop productivity. An extension is also critical to move research from the lab to the field and to ensure a return on investment in research by translating new knowledge into innovative practices [22]. The UN High-Level Panel of Experts on Food Security and Nutrition argues that national research and extension systems need full attention and investments from governments and the donor community [25]. Extension services are classified into three types [24]: Technology transfer the traditional model of the transfer of advice, knowledge, and information in a linear manner; Advisory by the cadre of experts as a source of advice in relation to specific problems faced by farmers; Facilitating the farmers to define their own problems and develop their own solutions.

In this thesis, by considering Indian context, we have explored the development of Decision Support System (DSS) to enable the farmer to get the advisory by detecting the crop problem by exploiting the latest technological developments such as the internet, smartphone revolution and the data science

related concepts like classification, decision trees, generalization/specialization, and notion of coverage set.

The rest of the chapter is organized as follows. In the next section, we explain the farming situation (agriculture and farmers) in India. Next, we give an overview of the existing IT-based systems which have been developed in India. After explaining the limitations and gaps of the existing systems, we provide an overview of the proposed approach.

1.1 Agriculture, farmers and knowledge delivery in Indian context

In the Indian context, agriculture is the key sector of the Indian economy as more than two-thirds of the Indian population continues to live in villages. Agriculture has been the major source of livelihood for these people. Although the share of agriculture in the national income has declined from 50% in 1950 to 18% in 2007-08, more than 60% of the workforce is still engaged in agriculture. Currently, the economy of the country is directly or indirectly depending on the performance of agriculture.

Another important dimension of the contemporary agriculture scenario in the country is that more than 70 percent of cultivation is taking place in small and marginal holdings. What is more important to note here is the socio-economic profile of the small and marginal farmers of today. Majority of them belong to the lower strata of society, belonging to OBC, SC and ST categories. Not many progressive farmers and early adopters of the erstwhile green revolution phase are found in agriculture today. Actual cultivation is being handled by those sections of farmers whose social, cultural and economic capital is poor. At the same time, we witness a steady increase in the intensity of agricultural technologies which are supposedly highly productive and efficient but not available with the small and marginal farmers. When compared to the era of green revolution the knowledge dissemination to the needy farmers is dismal or absent. It may be said that among several factors contributing to the agrarian crisis, the absence of knowledge/information dissemination at the small and marginal farmers has become the most important one.

The technological developments are rapid in optimum seed rate, fertilizers, and pesticides, but more than 80% of Indian farmers are ignorant of these technologies. The technology transfer system in India is outmoded, slippage, and causing serious technology gaps. Another problem is knowledge dissemination is not spreading equally across society. These gaps are high due to the poor and marginal sections, and non-uninformed or uneven distribution of knowledge in the society, and hence the gap widens [77]. On the other hand, a report [44] is in the opinion that commercial agriculture needs up-to-date information, strong and effective communication channels to make agriculture sustainable. In this context, information has become more critical in agriculture.

So, for sustained growth of agriculture productivity, every country supports three main branches of agriculture: education, research, and extension services [18]. The core objective of agricultural education is to grow the workforce with expertise to carry out education, research and extension services. Agriculture research involves developing new farming methods, solving the existing problems, and

developing new products and technologies. Agricultural extension's objective is to bring the innovations of agricultural research to the field level by educating the farmers [69].

The agricultural extension acts as a bridge between the farmers and agricultural research scientists. It disseminates the information regarding the latest technologies developed in the lab to the fields so that farmers are benefited by practicing the learned technologies. In the majority of developing countries, agricultural extensions are the vital facilitator of food security and also reduce rural poverty. Agricultural extensions are majorly run by public, private, NGOs and civil society institutions. Agriculture extension provides a wide variety of services to the farmers such as advisory regarding crop selection, management, marketing, technology transfer, training, information and knowledge [38].

- In India, since the 1950s planned agricultural extensions have been in service of farmers. In the green revolution decade of 1960-1970, the major extension service initiatives Government's Community Development programme (1952), National Extension Service (1953), Intensive Agricultural District programme (1960), Intensive Agriculture Area programme (1964), and High yielding Varieties programme (1966) were taken to educate the farmers about the improved farming methods and high yielding varieties.

- In the 1970s and 1980s, world bank's training and visiting (T&V) programme was introduced in multiple states like Punjab, Haryana, Uttar Pradesh, Andhra Pradesh, and Tamilnadu. In the T&V programme, a large number of technical workers are recruited, formally trained on agricultural technologies and deployed area-wise. The extension workers are assigned to specific areas and set of farmer groups and they visit the fields through the crop cycles and ensure that the advice is followed by the farmers appropriately. As a huge workforce with technical skills is required for this programme, it encountered funding sustainability issues.

- Indian government introduced Agriculture Technology Management Agency (ATMA, 1998) which created a platform to converge the financial and human resources from not only by the public sector but also non-government bodies such as private sector, NGOs, farm communities, and civil societies.

- In 2014, National Mission on Agriculture Extension and Technology (NMAET) introduced in which information and communication technologies (ICTs) like projectors, low-cost films, handheld devices, mobile-based services, Kisan Call Centres (KCCs) have played a key role in speeding up the information dissemination [38].

1.2 Overview of IT-based Agro-Advisory Systems in Indian context

In agriculture, crop production dynamics are influenced by multiple factors such as the type of soil, crop variety, location, weather, and management practices. To improve crop productivity, farmers need integrated farm advice that consists of advice for crop protection and production problems, and appropriate risk mitigation measures based on the weather pattern experienced and experiencing by crop.

To make correct decisions on various critical matters, farmers frequently need information and advice on many different technical and economic aspects. The information helps them to make correct decisions on various critical matters such as what crop to plant, the variety to use, the inputs to apply, and practices to follow, including how, how much and when, for the best productivity and returns. With the rapid development and scientific progress, the number of choices available and the knowledge-base of agriculture has expanded tremendously, making decision-making more complex and difficult. Further, market liberalization, globalization, and climate change are resulting in growing variability/ volatility including in the agro-climatic environment and the markets, thereby substantially increasing the risks and making the consequences of wrong decisions more severe. Farmers livelihoods depend substantially on the decisions they make and therefore on the information available to them.

Information and Communication Technologies (ICTs) are a "diverse set of technological tools and resources used to communicate, and to create, disseminate, store, and manage information. These technologies include computers, internet, and network hardware and software, satellite systems, broadcasting technologies (radio and television), and telephony (landlines and cellular). As well as the various services and applications associated with them, such as web portals, email, SMS, video-conferencing, decision support systems, and expert systems, etc" [69]. ICTs in recent decades have seen many significant developments. ICTs have the ability to disseminate the information faster and cost-effectively there by ICTs are used to improve communication in many domains such as education, agricultural, healthcare, disaster management [55]. Mass usage of ICTs in agricultural extension services enables the socio-economic development of rural areas in many developing countries like India.

In this context, the recent developments in ICTs offer a great new way and opportunity and have been harnessed by the Government of India into the initiative of Kisan Call Centres (KCC), and related systems of Kisan Knowledge Management System (KKMS), Farmers Portal, and M-Kisan Portal. These have gradually grown into action since 2004. During the last two decades, several efforts are being made to extend IT to build information delivery systems to support farmers in improving crop productivity by providing scientific crop management advisory by exploiting information and communication technologies.

We briefly provide the outline of the systems.

(i) **Kisan Call Centers:** The Department of Agriculture & Cooperation, Ministry of Agriculture, Govt. of India launched Kisan Call Centres [80] in 2004. The role of these Kisan Call Centres (KCCs) is to respond on the spot to questions related to agriculture asked by farmers, in the local language, and on a continuous basis. A toll-free number 1551 was designated for this.

(ii) **Farmers portal:** Farmers' Portal [4, 78] is an integrated portal frormer converging elements from 800 plus websites and giving information about Dealers (seed, fertilizer, pesticide, machinery), Government schemes & subsidy, Location of Storage Godowns, Insurance, Credit, Daily Market Price of commodities. The farmer can find the information through a drill-down approach from the state down to the block level.

(iii) **mKisan:** mKisan [7] is an SMS Portal for farmers enables all Central and State government organizations in agriculture and allied sectors to give information/services/advisories to farmers by SMS in their language, preference of agricultural practices and location. SMS Portal for Farmers has empowered all Central and State Government Organizations in Agriculture & Allied sectors (including State Agriculture Universities, Krishi Vigyan Kendras, Agromet Forecasts Units of India Meteorological Department, ICAR Institutes, Organization in Animal Husbandry, Dairying & Fisheries, etc.) to give information/services/advisories to farmers by SMS in their language, preference of agricultural practices and locations.

(iv) **Farm-level eSagu:** An effort has been to build a farm-specific a push-based system called, eSagu [62, 64] system (Here, the word 'Sagu' means 'Cultivation' in the Telugu language), which is a "personalized farm-specific push-based system" with the feedback. In eSagu, agricultural experts generate expert agro-advice based on the latest information about the crop situation received in the form of both digital photographs and text. The scientific agro-advice is provided to every farm (or field) at regular intervals, i.e., once every 10 days from sowing to harvesting.

(v) **Village-level eSagu:** In "Village-level eSagu system" [53], a sample number of farms are selected for each major crop of a location (a group of villages). The agricultural experts provide expert advice to these farms at regular intervals based on photographs and text. The expert advice is made available to all the farmers in the village by displaying color printed sheets of both photographs and advice text on the notice boards in the corresponding villages. The farmers access the information on the notice board and take appropriate steps.

(vi) **aAQUA:** aAQUA [61] is an online multilingual, multimedia Agricultural portal for disseminating information from and to the grassroots of the Indian agricultural community. The aAQUA system answers farmers queries based on the location, season, crop and other information provided by farmers. It makes use of novel database systems and information retrieval techniques like intelligent caching, offline access with intermittent synchronization, semantic-based search, etc.

(vii) **Digital Green:** Digital Green [33] is a research project that seeks to disseminate targeted agricultural information to small and marginal farmers in India using digital video. Digital Green works with existing, people-based extension systems and aims to amplify their effectiveness.

(viii) **mKrishi:** An Interactive Voice Response (IVR) based advisory service was developed [72]. The entire service was divided into two modules: Farmer end IVR based module and a Web console accessed by Agricultural Expert. Whenever the farmer calls the IVR service number using his registered phone number, the query is recorded and stored in the database. The agricultural expert can see this query on his "Expert web console" where all the details with regards to registered information are displayed to and provide expert advice.

(ix) **IFFCO Kisan Sanchar Limited:** IFFCO Kisan [2] is providing daily 3 free voice-based messages in local vernacular language to all the customers with the help of Green Sim Card and

empowering farmers with timely, relevant and high-quality information and services. Tt covers a variety of topics relevant to the farmers like agriculture, animal husbandry, weather, market rates, education, health, Govt. schemes, etc. specially created and tailor-made to suit the local requirement.

1.3 Research gaps and opportunity

1.3.1 Research gaps

Requirement: The requirement of an Indian farmer is that he/she needs timely, continuous and actionable information for efficient farm management.

Over three decades, as explained in the preceding section, several agro-information systems are being developed by exploiting ICTs. These systems can be divided into push-, pull- and hybrid systems.

- Under pull-based systems, it is assumed that farmers and other stakeholders pull the information from the system. Kisan call centers, farmers portal, aAQUA, Digital Green and several web portals [51] operated by the central government and state governments in India fall under this category.

  Pros: It is easy to develop and operate.

  Cons: It is difficult for the majority of illiterate and semi-literate farmers to pull the actionable agricultural advice through such systems due to low knowledge levels, communication and perceptual problems.

- Under the push-based system, it is assumed that the system pushes the agricultural information to the farmers door-step. The push-based systems vary with respect to the degree of generalization to personalization.

  - Generic push systems: In generic push systems, generic agricultural information is pushed to farmers. The radio, video, proactive SMS and voice-based services push generic information to farmers [69].

    Pros: If infrastructures like radio, TV, and the internet are available, it is easy to develop and operate such systems. It is useful to empower farmers.

    Cons: The utility is low. Unable to push the relevant information to farmers as each farmer's problem might be different. That is, it is not possible to provide timely and actionable agricultural information to a large number of farmers.

  - Farm- and location- specific push systems: In this system, actionable information is delivered to each farm in a regular manner. The farm-level eSagu system is an example of such a system.

Pros: The farm-specific system meets the requirement of farmers. It provides actionable and timely information to each farm.

Cons: the location-specific eSagu system suffers from scalability issues.

- Under hybrid systems, there are efforts to build a hybrid system by combining the concepts of push- and pull-based systems. The village-level eSagu system [53] is an example of the hybrid system.

  Pros: Provides location-specific information to farmers. It is scalable and cost-effective.

  Cons: Still, some degree of generic information is pushed to farmers.

Overall, the research gaps are as follows.

- Most of the systems are pushing only generic information to farmers.
- Several farmers are having difficulty in pulling agricultural advice from call centers due to knowledge gap and communication gap.
- Most of the Indian farmers are not getting timely actionable agricultural advice.

1.3.2 Opportunity

Overall, it can be observed that farmers are following unscientific practices since actionable agro-advisory are not reaching the majority of the farming community in a timely manner [37]. The farmers are also suffering from the issues of the knowledge gap and communication gap. We have to investigate the building of efficient systems to enable the farmer to acquire actionable agro-advise in a timely manner.

India is into a smartphone revolution. The capability of a smartphone is equal to computers. They can browse the web portals and search for information. Several smartphone-based applications are developed, which are revolutionizing the e-Commerce and gaming.

The literacy levels of India are increasing. So in the future, most farmers will be educated and start using smartphones. So, we can consider farmer who visits the field with the smartphone as a sensor.

Also, it has been observed that the notion of generalization and hierarchy has been not been exploited in building agro-information delivery systems. Farmer is also educated and has some perception capabilities to detect the visual symptoms of the crop disease. Similarly, agriculture expert analyzes the crop problem by employing sophisticated visual symptoms analysis of the crop disease. By exploiting the visual/sensor capabilities of the farmer in identifying crop diseases on one end and analyzing sophisticated visual symptom analysis of crop problems on the other end, there is a scope to develop a crop diagnosis model by extending the notions of generalization, hierarchy and visual symptom analysis.

1.4 Overview of the proposed approach

To enable the farmer to get the actionable agro-advisory, we have made an effort to build a smartphone based Crop Disease Diagnosis System by exploiting the progress in data science and smartphone technologies. Especially in India, in recent years, an increased number of rural farmers are purchasing smartphones. By exploiting the latest developments in data science technologies and increased penetration of smartphones in rural India, there is an opportunity to build a system to enable the farmer in identifying the crop problem and getting the agro-advice. It is assumed that the farmer possesses a smartphone with an internet connection and visits the field (or crop). The issue is to build a system, which we call *Crop Darpan* for guiding the farmer to identify the field problem and acquire the actionable agro advice.

The summary of the proposed methodology is as follows. It is assumed that being literate and able to use the smartphone, a farmer is able to identify the visual symptoms of the crop. In this system, the farmer visits the field with smartphone having internet connectivity, confirms the existence of the visual symptoms of the field problem through a smartphone, and identifies the field problem and acquire the agro-advice.

The research issue was to build a system, which resembles the process followed by agricultural expert to detect the crop disease. This requires the embedding of experts knowledge about the visual symptoms of all diseases into the system. We consider that a field problem is a combination of a set of visual symptoms. The field may suffer from one of the several (about more than 30) field problems. The issue is to develop a system to help the farmer to identify the field problem out of several problems. The idea is to organize the visual symptoms (building a king of index structure) of all field problems in a hierarchical manner enabling the farmer to quickly identify the problem. For this, we need a generalization-specialization (or parent-child) relationship among the field problems. Based on several observations, it has been identified that it is possible to organize the visual symptoms of field problems in a hierarchy, i.e., for each specialized visual symptom, it is possible to identify generalized visual symptoms. For example, a *leaf curl* visual symptom is a generalization of *downward leaf curl* and *upward leaf curl* visual symptoms. Also, *leaf problem* visual symptom is generalization of *leaf curl* and *holes on the leaves* visual symptoms. Both generalized and specialized visual symptoms of field problems of the given crop can be collected from the subject matter specialists (or domain experts). The proposed *Crop Darpan* database contains a hierarchy of visual symptoms of field problems of the given crop that consists of generalized and specialized visual symptoms, which are organized in a hierarchy.

The components of Crop Darpan system are Knowledge acquisition protocol, Knowledge base, Question popping model, and the user interface. We have explored three design choices: (i) number of questions to be asked per screen, (ii) algorithms followed for knowledge acquisition protocol from the domain expert, and (iii) question popping algorithm to pose questions to farmer. The development of Crop Darpan has been carried out in an incremental manner. Overall, five design frameworks were developed in an incremental manner. The fifth and the final framework, which is called as Hierarchy-based

protocol for knowledge acquisition and Hierarchy model for question popping with Multi-questions per screen (HHM) framework, has been conceived.

As a part of Knowledge Acquisition protocol, we have developed a protocol to collect both generalized and specialized symptoms for a given crop from agricultural domain experts. The question popping model has been developed for organizing the collected knowledge database in an hierarchical manner. An algorithm was proposed to pose an appropriate sequence of visual symptom related questions, starting from generalized to specialized visual symptoms, to the farmer, when he/she uses Crop Darpan to diagnose the field for the crop problem with the smartphone. We have opted for multiple questions per mobile screen.

The Crop Darpan system was built and operationalised in the field for Cotton crop, It was found out that the farmers are able to identify the field problem with the Crop Darpan system. The evaluation results was carried out with domain experts and farmers show that the domain experts are very positive about the system. The farmers are also able diagnose the most of the Cotton crop diseases. The proposed framework is a scalable and generic system. It is extendable to other crops, regions, and languages/dialectics.

1.5 Contributions

The contributions are as follows.

1. Introduced the research problem of diagnosing crop disease by extending the concept of generalization-specialization.

2. For the problem of crop disease diagnosis, we have proposed a framework of *field diagnosis guide*. As a part of this, we have developed two protocols: Knowledge acquisition protocol and Question popping algorithm.

3. By extending the framework of *field diagnosis guide*, we have built a prototype, called *Crop Darpan* for the Cotton crop, which helps the farmer to identify the crop problem through a smartphone.

4. We have carried out the user evaluation study of *Crop Darpan* system for the Cotton crop.

1.6 Thesis organization

The rest of the thesis is organized as follows.

- **Chapter 2. Related work:** In this chapter, we discuss about the work done in the literature about the ICT based systems in agricultural knowledge dissemination and the related data science approaches.

- **Chapter 3. Background, Problem Context, and Research Methodology:** In this chapter, we present the background, context of the problem, and the research methodology.
- **Chapter 4. Concepts of Crop Darpan:** We present the basic idea of the Crop Darpan system and present the design choices. Next, we present the five design frameworks based on those design choices.
- **Chapter 5. Details of Crop Darpan Prototype for Cotton Crop:** We present the details of Crop Darpan system design and operationalization for Cotton crop.
- **Chapter 6. Evaluation of Crop Darpan Prototype for Cotton:** We present the evaluation results of Crop Darpan prototype for Cotton crop,
- **Chapter 7. Conclusion and future work:** In this chapter, we present the summary of the thesis, conclusions/advantages, limitations, and share the plans for future research.
- **Appendix A:** Contains the supplementary material referred in Chapter 3.
- **Appendix B:** Contains the supplementary material referred in Chapter 5.
- **Appendix C:** Contains the supplementary material referred in Chapter 6.

Chapter *2*

Related work

There are efforts to extend ICTs to build knowledge delivery system in agricultural domain. Also, there are data science based approaches/paradigms, which could be extended to build improved knowledge delivery systems in agriculture.

In this section, we discuss the related approaches of ICT based approaches for knowledge delivery systems in agriculture and related data science approaches.

2.1 Literature related to Agriculture knowledge delivery systems

A voice call based crop advisory system named Kisan Call Centres (KCCs) [51] is reported in [80]. KCC system provides a toll free number to which a farmer can call and get information needed in local language. KCC employs a farm tele advisor who communicates with the farmer to identify the information requirement of farmer and responds with appropriate advice. In this system, the communication between the farmer and the tele advisor is the backbone of the advice delivery process. In this system it is expected that farmer communicates the details of the crop disease to the tele advisor to obtains the corresponding agriculture advice.

In [78, 80], a web portal based knowledge delivery system named Farmer's Portal (FP) [4] is reported. The FP system is a platform that integrates more than 800 websites fulfilling specific needs of farmers. The FP system warehouses multiple databases from different sources so FP provides variety of information regarding dealers (seed, fertilizer, pesticide, machinery), government schemes & subsidy, location of storage godowns, insurance, credit, daily market price of commodities for the farmers. The FP platform support both English and Hindi languages. The farmer is expected to drill down the website from state to block level to obtain/pull the specific information needed.

In [31, 16], an SMS based portal named mKisan [7] is reported. To avail mKisan portal service, the farmers are first registered in the portal by providing the details such as mobile number, location, language, and preference of agricultural practices. The central and state government organizations in agriculture and allied sectors to the block level are authorised to send advisories and information to the farmers. Based on the district or region, language and the crop details of the farmer, SMS containing

information about current problems and scientific advisories are sent in a broadcast mode. This service provides the generic information to the farmers.

In [53, 62], a personalized farm-specific advisory system named eSagu is presented. In eSagu, the farmer avails the advisory services by registering his/her crop details in the system. In eSagu, the latest information about the farmers' crop situation is captured by using both digital photographs and text at regular intervals. Agricultural experts analyzes the collected crop information and generates the corresponding agro-advice which is later delivered to the farmers through photo copies and/or SMS. The system is targeted to deliver/push the advice to illiterate farmers, so the farmer is a passive consumer in advice delivery process. The eSagu system suffers from scalability issues.

In "Village-level eSagu system" [53], a sample number of farms are selected for each major crop of a location (a group of villages). The agricultural experts provide expert advice to these farms at regular intervals based on photographs and text. The expert advice is made available to all the farmers in the village by displaying colour printed sheets of both photographs and advice text on the notice boards in the corresponding villages. The farmers access the information in the notice board and take appropriate steps. Even though this system is scalable, it pushes the generic information to the farmers.

aAQUA (almost All QUestions Answered) [61] is a multilingual and multimedia web based discussion forum. It provides a platform for the agricultural community to create, view, and manage the content. Like call centers, it expects the farmer to ask a question regarding the crop disease information.

Digital Green [33] is a research project that seeks to disseminate targeted agricultural information to small and marginal farmers in India using digital video. This system also pushes the generic information in terms of digital videos to the farmers.

An Interactive Voice Response (IVR) based advisory service mKrishi is reported in [72]. In this system, whenever the farmer calls the IVR service number using his registered phone number, the query is recorded and stored in the database. The agricultural expert can see this query on his "Expert web console" where all the details with regards to registered information are displayed to and provide the expert advice. This system also expects the farmer to convey the details of the crop condition through recorded voice.

IFFCO Kisan [2] is provides daily 3 free voice based messages in local vernacular language to all the customers with the help of Green Sim Card and empowering farmers with timely, relevant and high quality information and services. The duration of the messages is 60 seconds and it covers a variety of topics relevant to the farmers like agriculture, animal husbandry, weather,market rates, education, health, Govt. schemes etc. specially created and tailor-made to suit the local requirement. This system also pushes the generic information to the farmers.

Rice Doctor [9] is an interactive tool for extension workers, students, researchers and other users who want to learn and diagnose pest, disease, and other problems that can occur in rice and how to manage them. In this system, manually observed information about the infected plant is taken as input from the user and based on the input information of the infection a diagnosis is identified. It is a knowledge

repository aimed at all kinds of stakeholders especially for agricultural personnel who wants to acquire the domain knowledge. This is not farmer-centric system.

2.2 Literature related to Computer vision based diagnostic systems

In [50], a smartphone-assisted disease diagnosis system has been presented with the motivation of increasing global smartphone users and the advances in computer vision. In this work, convolutional neural network (CNN) models are trained on large verified data set consisting 54,306 images of diseased and healthy plant leaves. The CNN model is trained using images of the leaves as input and 38 crop disease pairs as output classes. The the trained models performed very well on the held-out test set with an accuracy of 99.35%. In [32], a deep-learning-based approach is proposed to detect diseases and pests in tomato plants. The works in [13, 28, 67], used image processing and machine learning algorithms to diseases in grape leaf samples.

In all the above mentioned works, the leaf samples of the diseased plants are collected in a very controlled conditions such as plucking the affected leaves from plant and taking photographs with plane background, good lightening conditions and clicks from constant distance. Also, the tested held-out sets are also from the same data sets. So, the diagnosis accuracies can not be expected in realistic conditions where the images are collected with complex background, varying lightening conditions and varying distances.

In [68], a work is proposed to diagnose the disease from images of grape plant leaves collected in realistic conditions using image processing techniques and neural network architectures. However, this is developed to identify only two distinguishable diseases: downy mildew and powdery mildew.

Plantix [8] is a mobile crop advisory app for farmers, extension workers and gardeners. It can diagnose pest damage, plant disease and nutrient deficiencies affecting crops and can offer corresponding treatment measures. Plantix takes advantage of deep learning technology which involves neural networks. From the daily new images sent by Plantix users worldwide, the network constantly learns more. This constantly growing database provides Plantix users with current information and alerts on plant diseases, pests and their worldwide distribution. As this system uses deep neural network architectures, it requires huge data sets to perform well. Creation of huge data sets confined to a specific domain is a tedious task. The prediction process involved in these models are not human interpretable and the predictions are probabilistic in nature. Recently, accuracy aspects of the Plantix is reported in [79]. It was reported that, over the years the accuracy of the Plantix is improved up to 50%.

2.3 Literature related to Expert systems

Expert system is one of the core applications of artificial intelligence which is essentially designed to emulate the experts in a particular domain [40]. Many expert systems have been developed in different domains such as healthcare [73, 30, 12, 46], education [43, 42, 29], automobile[81, 19, 14], agriculture

[57, 59, 58, 48, 66, 45, 54, 34, 49, 74, 56, 70, 11, 47]. An expert system [48] is developed to diagnose soybean diseases. It contains the both the decision rules representing the experts diagnostic knowledge and rules gathered by learning from large number of diseases cases. POMME [66] is an expert system for apple orchid management. It advises the farmers about remedies for infections, winter injuries, drought control and insect problems. COMAX [45] is an expert system for cotton crop management. It is integrated with a simulation model *Gossym* to simulate the growth of the cotton plants. It can take the weather information soil parameters, and pest damage as input and predicts the crop growth and yield. [58, 59] have discussed the importance of the expert systems and their usage as tools to speed up agricultural desert development in Egypt. CALEX [54] is an expert system shell targeted at crop management decision support. It is a generic shell and can be used for any domain. A large scale implementation of the CALEX is developed for cotton crop management and irrigation scheduling in San Joaquin Valley, California. POMI [34] is an expert system for integrated pest management of apple orchards. ETES [49] is an expert system developed in India with an intelligent front-end to calculate evapotranspiration (ET) in water irrigation systems. DIARES-IPM [47] is an integrated pest management expert system in in solanaceous crops. It helps the farmers identify pests and insects and advice appropriate treatments. [56, 74, 57] discuss the several expert systems related to different crops and different agricultural services such as irrigation scheduling, fertilization, pest control, disorder diagnosis and treatment.

The success of any expert system is majorly depends on the quality, accuracy and precision of the knowledge integrated in to it. However, knowledge acquisition is the most tedious and slow task in building a expert system. A knowledge engineer who is typically a computer scientist interviews and observes the domain experts and learns how they reason with their knowledge. The knowledge engineer then converts the acquired knowledge into computer understandable rules and complex logical inferences. But, converting expertise of an expert into rigid, pre-fixed rules is very challenging, restrictive and usually incomplete. Another major challenge of expert systems emerges when the size of the knowledge base increases. This causes the processing complexity to increase.

2.4 Literature related to object classification with human in the loop

In [1, 23], a two player spoken parlor game called 20 questions game is discussed. In this game, first player, the answerer thinks of a subject with out reveling it to the other player questioner. The goal of the questioner is to guess the subject thought by the answerer by asking maximum of 20 binary yes/no questions.

An interactive, hybrid human-computer method for object for object classification is proposed in [26]. The method applies to classes of objects that are recognizable by people with appropriate expertise (e.g., animal species or airplane model), but not (in general) by people without such expertise. It can be seen as a visual version of the 20 questions game, where questions based on simple visual attributes are posed interactively. The goal is to identify the true class while minimizing the number of questions

asked, using the visual content of the image. A general framework introduced for incorporating almost any off-the-shelf multi-class object recognition algorithm into the visual 20 questions game, and provide methodologies to account for imperfect user responses and unreliable computer vision algorithms. Through results it was demonstrated that incorporating user input drives up recognition accuracy to levels that are good enough for practical applications, while at the same time, computer vision reduces the amount of human interaction required.

Based on the framework proposed in [26], an interactive human in the loop computer vision technique for the recognition of skin lesion images is presented in [63]. In this work a dermatology "Question and Answer" bank has been developed for interactively extracting human perceptual knowledge of images in order to assist computer vision algorithms in boosting recognition accuracies. The experimental results to show that for some diseases, traditional computer vision techniques can only achieve a recognition rate of 20%, whilst with human in the loop the performance can be boosted to over 96%. It was shown that users do not require any medical knowledge to answer these questions to achieve excellent recognition rates.

The works like [26, 63] have demonstrated the possibility of developing human in the loop based question answering systems for other domains like agriculture.

2.5 How the proposed approach is different?

The proposed Crop Darpan system is different from the approaches in the literature in the following manner.

- Farmers' requirement is that "he/she needs actionable agro-advice to his/her field in a regular manner". So far, none of the agro-advisory system except eSagu has made an effort to build the system. It has scalability and sustainability problems. The proposed approach is different in the sense that it aims to help the farmer with the actionable agro-advice by exploiting recent spread of smartphone infrastructure in the country.

- The image/vision based systems prone to data collection and preparation problems. The answers are difficult to interpret. Also, the answers are probabilistic and lack accountability. The proposed system is different in the sense that it considers farmer as a sensor. Also, the process followed by the system for decision making is interpretable. The farmer owns the responsibility.

- The expert systems and deductive systems are not farmer centric. The proposed Crop Darpan system is farmer-centric, which in turn exploits the existing crop-specific other related agricultural knowledge bases.

- The proposed system is different in the sense that it is the first attempt to extend the notion of generalization to develop agricultural knowledge delivery system by assuming farmer with the smartphone as a sensor.

2.6 Summary of the Chapter

In this chapter, we have surveyed the approaches about ICT-based agricultural knowledge delivery and related data science approaches. We also presented how the proposed approach is different from the existing approaches in the literature,

In the next chapter, we present the context of the problem.

Chapter *3*

Background, Problem Context and Research Cethodology

As outlined in the introduction, the research issue is to develop a system to help the farmer to diagnose the crop disease using the smartphone. The proposed framework is due to several observations regarding issues with existing agricultural information dissemination technologies and opportunities identified from the data science domain which could be extended to the agricultural domain for building improved knowledge delivery solutions. The observations include (i) The issue of lack of efficient agricultural knowledge delivery tools in India (ii) Insights during field implementation of eSagu technology over since 2004 (iii) Trends in information technologies and (iv) concepts of generalization/specialization and human-in-the-loop systems. Based on the observations, we have conceived a few intuitions and confirmed the same through field experiments. Next, we interacted with the subject matter specialists and conceived the basic idea of generalization-specialization based crop diagnosis framework. Overall, we have extended generalization-specialization concept and other related concepts from the data science domain and made an effort to build a generalized crop diagnosis framework for all crops to help farmers, especially, Indian farmers.

In this chapter, we explain the background, which consists of a collection of observations and defines the problem. We also present the research methodology, which has been followed to develop a solution (refer to the next chapter) to the proposed problem.

3.1 Background

In this section, we start by explaining several observations regarding the issue of crop disease detection.

3.1.1 Issue of Lack of Efficient Agricultural Knowledge Delivery Systems

In [38], the agricultural knowledge delivery systems in India till 2018 have been reviewed and to form a plan for *Doubling the Farmers' Income*. In that report, the following broad conclusions are drawn after the review of public, private and civil society-led extension interventions in the country:

- *Agriculture extension services in India are predominantly centered around crop husbandry with a pronounced tilt towards terms of trade. The approach of public sector extension is to offer a one-size-fits-all product to all farmers. In a country with over 86 percent of farmers categorized as small and marginal, this is a self-limiting approach as the huge variations in resource endowment, agro-climatic conditions, and legal exigencies are not factored into the model of agriculture extension being followed either by the government.*
- *While NGO-led extension models offer far more variety and display sensitivity to local priorities and conditions, they do not have the capacity or scale to make a significant impact across large regions. They are also seriously hampered in scaling up due to paucity of resources, as public sector extension agencies rarely explore synergies or cooperation and donor support continues to be project-driven and episodic.*
- *In recent years, the growth in the High-Value Agriculture (HVA) sector has been twice or sometimes even thrice that of the crop husbandry sector. Yet, agriculture extension services for HVA sectors remain weak and disorganized.*
- *The analysis suggests that the government, private sector, NGOs and others providing agriculture extension services are working in isolated silos with little or no functional coordination at the field level. This leads to a restriction of good practices generated in each of these sectors and an opportunity for wider application is lost.*
- *Lastly, it may be concluded that a large number of players in the agriculture extension arena function without any standards or certification of quality. This leaves questions of accountability up in the air as the majority of farmers are not in a position to pursue legal remedies in case of erroneous or even harmful advice.*

In [20], an effort has been made to review the state of agricultural extension systems in India. It was reported that

> *while the broad objectives of decentralization and farmers' participation have been achieved, the reforms fall short in terms of increased accountability to farmers and being fully demand-driven. The inclusiveness of smallholder and marginal farmers has been achieved only partially. The group approach to extension remains weak and needs strengthening at the block and village levels.*

Overall, the above studies [38] and [20] reveal that there is a lack of efficient knowledge delivery systems that cover poor and marginal farmers in India. So, developing new innovations for agriculture knowledge delivery is a research issue.

3.1.2 Observations of eSagu implementation

Since 2004, at IIIT Hyderabad, research efforts are being made to build IT-based agro-advisory systems to deliver actionable agro-advice to every farm field in a timely manner. Two systems were

built: Farm-specific agro-advisory system and Location-specific agro-advisory system. We will briefly provide the outline of these systems and provide the observations.

- Farm-specific agro-advisory system [62, 64]: We have made efforts to build a push-based system called, eSagu system (Here, the word 'Sagu' means 'Cultivation' in the Telugu language), which is a "personalized farm-specific push-based system" with the feedback. In eSagu, agricultural experts generate expert agro-advice based on the latest information about the crop situation received in the form of both digital photographs and text. The scientific agro-advice is provided to every farm (or field) at regular intervals, i.e., once every 10 days from sowing to harvesting. Since 2004, the expert advice has been delivered to thousands of farms of several crops at different places in India. The impact studies show that the farmers have realized considerable monetary benefits by reducing the quantity of fertilizer application, the number of pesticide sprays, besides getting the additional yield. There are scalability issues with a farm-specific eSagu system.

- Location-specific agro-advisory system [53]: To improve the scalability of the eSagu system, a new model called "Location-specific eSagu system" is being developed. In this system, a sample number of farms are selected for each major crop of a location (a group of villages). The agricultural experts provide expert advice to these farms at regular intervals based on photographs and text. The expert advice is made available to all the farmers in the village by displaying color printed sheets of both photographs and advice text on the *notice boards* in the corresponding villages. The implementation results show that, in addition to reducing input costs and improving crop productivity, the system is also enabling community discussion and knowledge sharing/empowerment.

A sample picture of notice board is presented in Figure 3.1 and a sample advice sheet shown in Figure 3.2.

During three-year implementation (2015-18) of location-specific eSagu [53], it was observed that several farmers are seeing the *notice board*, seeing the color photographs of crop issues and them calling agricultural scientist. Several farmers have informed that their own field problem is facing similar problem or somewhat different problems and requested the agricultural scientists to provide expert advice. Based on this experience, the following observations were drawn.

Observation 1.

- Farmer is able to see the crop photos on the notice board and understand the crop disease.

- After understanding the crop disease, by seeing the color pictures, he is able to call the agricultural expert and explain the crop diseases through voice call.

The following intuition has emerged from the above observations.

Figure 3.1: Notice board

Figure 3.2: Sample advice sheet in notice board

Intuition 1. *When the farmer is able to see the notice board, understand the crop images and convey to agricultural scientist through voice call, it should be possible for the farmer to see the field and confirm the visual symptoms of the crop if we pose them to him/her in an understandable manner. So, if the farmer is in the field, and agricultural scientist calls him and asks questions concerning visual symptoms of the crop disease, the farmer has the ability to confirm the presence/absence of visual symptoms of the crop disease.*

3.1.3 Recent information technology trends

Information technology is an industry on the rise, and business structure, job growth, and emerging technology will all shift in the coming years. Current trends are improving and presenting new functions in fields like agriculture, medicine, entertainment, business, education, marketing, law enforcement and so on.

The following are recent information technology trends: Cloud Computing, Mobile Computing and Applications, Big Data Analytics, Automation, Artificial Intelligence and Smart Machines, Virtual Reality, Augmented Reality, Blockchain Data, Cyber-Privacy and Security, and Internet of things. The IT trend *mobile computing and applications* especially *smartphone and smartphone-based applications* has impacted almost all walks of human life in India, including rural farmers. In a real sense, Smartphone is a mobile phone which provides advanced features and functionality not just restricted to traditional functionalities like making voice calls and sending text messages. The Smartphone is able to carry out

most of the tasks, which are carried out by computer. The Smartphone supports wide range of functions like taking photos, playing games,watching videos, navigation, audio/video recording, send/receive an e-mail, built-in apps for social web sites and surfing the Web, wireless Internet and much more.

Observation 2

The latest trends in India indicate that the usage of smartphones is increasing in the general public, especially in rural India. This statistic [10] shows the number of smartphone users in India from 2015 to 2022. For 2022, the number of smartphone users in India is estimated to reach 442.5 million. As per the Kantar IMRB ICUBE report [5],

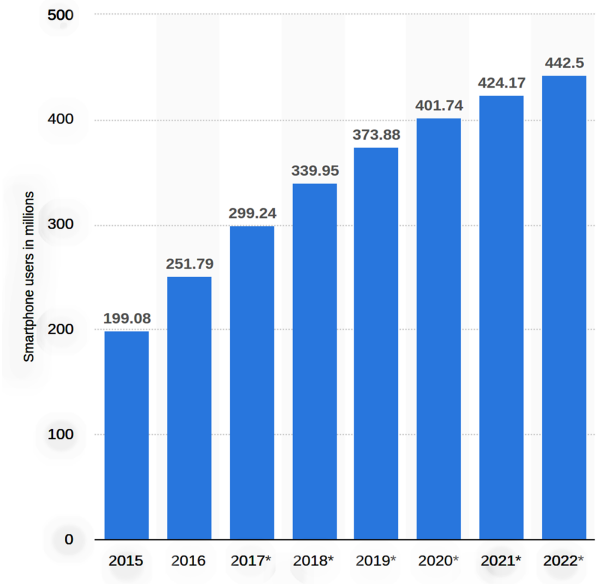


Figure 3.3: Trend of number of smartphone users in India from 2015 to 2022

the internet user base in India has exceeded 500 million mark and is likely to reach 627 million by end 2019 [6]. The number of internet users is estimated to be 566 million as of December 2018, registering an annual growth of 18%, said the report on digital adoption and usage trends in India. The report finds that 87% of the total user base, or 493 million Indians, are defined as regular users, having accessed the internet in the last 30 days. Of this, 293 million active internet users reside in urban India, while there are 200 million active users in rural India. Not surprisingly, 97% of users access the internet on their mobile devices. According to the report, digital adoption is now being propelled by rural India and figures show that it has registered 35% growth in internet users over the past year. It is now estimated that there are 251 million internet users in rural India, and this is expected to reach 290 million by the end of 2019. Increased availability of bandwidth, cheap data plans and increased awareness driven by government programs seem to have rapidly bridged the digital gap between urban and rural India. Consequently, the penetration in rural India has increased from 9% in 2015 to 25% in 2018.

The following intuition has emerged from the above observations regarding the smartphone revolution.

Intuition 2*: As a significant number of rural farmers are posssing/will possess smartphones with Internet connectivity, we can consider that it is possible for the farmer to visit the field and able to access the internet. Also, the smartphone itself has browsing and computing capabilities like a computer. Also, as per Observation 1, the farmer has the capability to confirm the existence of visual symptom, if posed in an understandable manner. So, if we build a tool which will mimic the process followed by agriculture scientist to diagnose the field problem by posing visual symptoms as questions, it will definitely help the farmer to identify the field problem.*

3.1.4 Generalizaion/Specialization and related concepts

A typical crop suffers from several diseases. For example, Cotton could suffer from 30 diseases. At a given time, typically, a crop suffers from two to three diseases. Normally, a disease or a crop disease can be captured in terms of visual symptoms. Each visual symptom can be expressed in a textual manner. For example, one of the visual symptoms of the Cotton crop disease "Aphids" is "leaf curling". When we pose the question "Are leaves curling in the crop?", we assume that the farmer is able to see the field and confirm the existence (Refer Observation 1).

So, given the list of crop diseases and corresponding visual symptoms, the issue is to help the farmer to identify one of the crop disease if it exists. As a brute force approach, we can pose all visual symptoms of each crop diseases in some order. Such a system is very inefficient.

In this connection, it is being noted that such types of problems are solved in the data science domain by extending the concepts of generalization-specialization, decision trees, and human-in-the-loop systems.

- **Generalization-specialization** In abstracting or conceptualizing the real world, perhaps *Generalization* is the most important mechanism [75]. A generalization is an abstraction which enables a class of individual objects to be thought of generically as a single named object [75].Generalization is a bottom-up approach in which a class of lower level entities are combined into a single higher entity by means of their common properties. Even the higher entities can also be combined to make further higher level entity. This kind of classification of entities into levels results in hierarchy as shown in Figure. 3.4.

  Figure 3.4 shows a sample hierarchy which models the domain of vehicles through abstraction and generalization. Here, each generic entity is abstraction that represents a class of all individual instances. For example, the generic entity *bike* represents all the physical instances of the bikes. In the figure we can see, for example, that *truck*, *bike*, and *car* can be generalized to the notion *road vehicle*. The generic notions *road vehicle* and *rail vehicle* are generalized to the notion *land vehicle*. Also, the generic notions *air vehicle*, *land vehicle* and *water vehicle* are further generalized to even higher generic notion *vehicle*. The root of of any sub-tree have the properties which

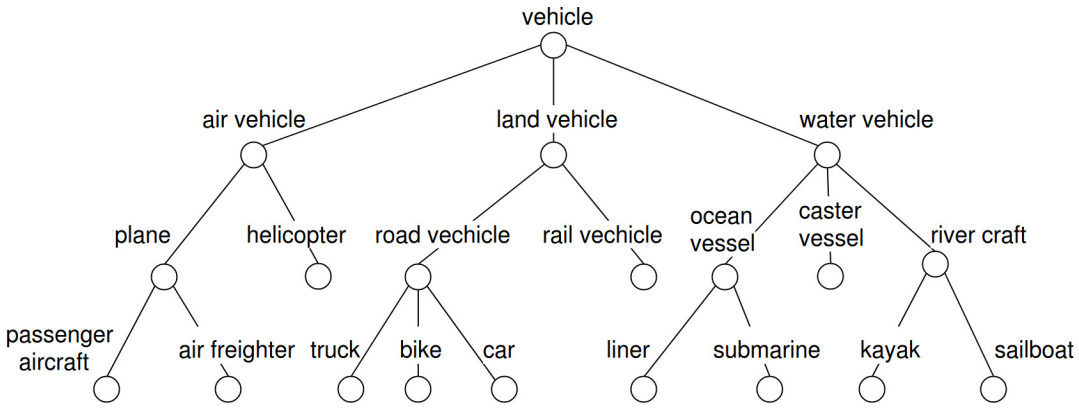


Figure 3.4: A generic hierarchy over vehicles

are common to all its descendants. For example, *water vehicle* have the property transporting over water which is common to all its descendants. Even though the descendants of a node have some shared properties, they must have distinctive properties compared their siblings and sibling's descendants. For example, air vehicle and land vehicles have common property of being transport vehicles, but they are distinguished by their medium transport.

Hierarchy can also created through the mechanism of specialization. Specialization is the reverse process of generalization. It is a top-down approach in which two are more specialized classes are formed from a generic class by identifying distinguishing features generic class objects.

Developing a hierarchy by generalization and specialization mechanisms results in a powerful data organization. Hierarchy provides easier understanding of complex models. They are easier models to develop and manage. A hierarchy can be easily extended by adding new generalizations or specializations as most of their attributes and behaviour are already available in the existing framework [41]. Due to these advantages, generalization and specialization mechanisms are widely used in databases, programming languages, and software development [27]. In Databases, the generalization or specializations hierarchies are realized through codd *is-a* relations [75]. In programming languages, object oriented paradigm and inheritance concepts are used to create hierarchies [65]. In software development, generalization specialization principles are used to acquire software specifications. Hierarchies provides a form a knowledge representation, so these generalization specialization are concepts are while acquiring the knowledge and its representation.

- **Decision tree based classification**

  To illustrate how classification with a decision tree works [3], consider a simpler version of the vertebrate classification problem. Instead of classifying the vertebrates into five distinct groups of species, we assign them to two categories: mammals and non-mammals. Suppose a new species is discovered by scientists. How can we tell whether it is a mammal or a non-mammal? One

approach is to pose a series of questions about the characteristics of the species. The first question we may ask is whether the species is cold- or warm-blooded. If it is cold-blooded, then it is definitely not a mammal. Otherwise, it is either a bird or a mammal. In the latter case, we need to ask a follow-up question: Do the females of the species give birth to their young? Those that do give birth are definitely mammals, while those that do not are likely to be non-mammals (with the exception of egg-laying mammals such as the platypus and spiny anteater). The previous example illustrates how we can solve a classification problem by asking a series of carefully crafted questions about the attributes of the test record. Each time we receive an answer, a follow-up question is asked until we reach a conclusion about the class label of the record. The series of questions and their possible answers can be organized in the form of a decision tree, which is a hierarchical structure consisting of nodes and directed edges.

The decision tree has three types of nodes: (i) A root node that has no incoming edges and zero or more outgoing edges (ii) Internal nodes, each of which has exactly one incoming edge and two or more outgoing edges. and (iii) Leaf or terminal nodes, each of which has exactly one incoming edge and no outgoing edges.

Classifying a test record is straightforward once a decision tree has been constructed. Starting from the root node, we apply the test condition to the record and follow the appropriate branch based on the outcome of the test. This will lead us either to another internal node, for which a new test condition is applied or to a leaf node.

- **Human-in-the-loop systems**

  While object recognition deals with the recognition of several objects belonging to a broader entry-level, fine-grained object recognition [26] aims to distinguish objects of subordinate categories that belong to the same entry-level object category. For example, the task of recognizing the broader entry-level categories such as a chair, airplane & bird of test objects from their respective images, whereas the task of recognizing the subordinate categories of test objects in the object category bird from their respective images. Fine-grained object recognition has numerous applications: beyond simply being able to describe the world in more detail, it can be used for improved scene understanding, studying society, and even analyzing biodiversity

  The existing fine-grained recognition methodologies can be categorized into three major recognition modes: Image-only based Recognition (IR), Question-answering based Recognition (QR), and Image and Question-answering based Recognition (IQR).

  The most popular model for fine-grained recognition is Image-only based Recognition (IR) wherein recognition is solely achieved through images by categorizing the image of a target object. However, this mode of recognition has a downside of being cost-sensitive as most often the task of the required picture-acquisition is a non-trivial task; particularly in fine-grained domains. A less popular recognition mode is Question-answering based Recognition (QR). These works utilize semantic visual attributes of objects and do recognition through interactive question

answering with a human-in-the-loop [26]. Another widely researched recognition method is Image and Question-answering based recognition (IQR) wherein recognition is realized jointly from images and humans. Unfortunately, most of the current works which employ IQR are either cost-sensitive or in-effective.

We briefly describe a picture-independent recognition method called Question-answering based Recognition (QR) which uses attribute (feature) information of an object collected through interaction with a human-in-the-loop. The QR approach exploits the notion that every visual object is characterized by its visual semantic attributes. QR does object recognition by posing questions on the presence/absence of relevant visual semantic attributes of a visible target object to the human-in-the-loop, whose perceived responses on the object to the questions collectively make recognition possible. Implicit to such a type of recognition is the assumption that lay-humans generally cannot recognize fine-grained categories (e.g., Myrtle Warbler, Thruxton Jackaroo, etc.) due to imperfect memory or limited experiences; however, they have excellent visual capabilities to recognize the attributes for recognition.

The following observation is formed based on the above approaches.

Observation 3. It can be noted that the notions of generalization-specialization, decision tree classification, and question-answering based object recognition approaches are enabling the identification of entity/record/object-based a hierarchical organization of related attributes. Normally, attributes in the respective domain are identified as per the problem requirement. Next, questions are being formed based on the attributes and these questions are organized in a hierarchical manner which is resolved to start from high-level to low-level.

The following intuition has emerged from above observations regarding generalization-specialization, decision tree classification, and human-in-the-loop systems.

Intuition 3*: As per Observation 1, the farmer has the capability to confirm the existence of visual symptom, if posed in an understandable manner. As per observation 2, the smartphone itself has browsing and computing capabilities like a computer. Based on observation 3, given crop diseases and respective visual symptoms, there is a scope to organize the visual symptoms in a hierarchical manner and build a human-in-the-loop kind of system to help the farmer to identify the crop disease in an efficient manner.*

3.2 Field Experiments

We have carried out two field experiments to understand the problems faced by the farming community in communicating the crop disease to an agricultural scientist and to understand the perceptions of farmers about the diseases of Cotton and Rice crops.

- Experiment to understand the problems faced by the farmer

In order to identify the problems faced by farming communities about communicating their crop diseases to agricultural experts through cell phones, a survey was carried out through questionnaires, and responses are being collected from a sample of farmers and agricultural experts.

(The questionnaire for the farmer is provided in Appendix A - Part 1 and the questionnaire for the scientist is provided in Appendix A - Part 2.)

The summary of the findings is as follows.

- It was identified that farmers are facing the issue of communication gap, under the existing systems such as call centers and web-based systems. On the one side, the farmer being an uneducated person or a person with minimal education was unable to express his crop diseases to agriculture experts. Another issue is the issue with the language and dialects used by the farmer. On the other side, agriculture expert being a very highly educated person was not able to identify the crop disease based on the generic symptom expressed by the farmer.
- It was identified that there is a requirement to build a system to help the farmer to frame a question to convey his/her current crop disease. This system will educate the farmer regarding the issues related to pests and other diseases of the crop with respect to various crop growth stages. The system will be designed in such a way that it will progressively unfold the questions in various crop growth stages from sowing to harvesting.
- The farmer has a mobile phone and visits the crop. It was felt that a mobile tool is required to enable the farmer to ask a question regarding crop diseases.

- Experiment to know the perceptions of the farmers about the diseases of Cotton and Rice crops.

  We have conducted the field experiment to know the real perceptions of the farmer on the diseases of Cotton and Rice crops. The data is collected from a sample of educated farmers (farmers who have completed 10th class) (Refer Appendix A - Part 3 for data collection form). Some of the visual symptoms collected on Cotton crops are given in Table 3.1.

Table 3.1: Perceptions of the farmer on the diseases of Cotton crop

Leaves are curling because I have not sprayed pesticide, Crop effected with green insects, Cotton crop is wilting, Leaf curling, Noticed sqaure/flower drop, Noticed sqaure/flower drop and flowers eaten, Green insects intensity is high, Leaves are eaten by insects, Insect infestation, Upward culrling of leaves, upward curling and reddening of leaves, upward curling and reddening backside of leaves, Leaf tip curling and reddening, Reddening backside of leaves, Cotton leaves are turning black, Spots on boll, wilting due to high temps, white insects, Leaf spots, Holes on leaves, Leaf spots, Yellowing of leaves, Cotton leaves are turning black, Spots on boll, White insect, Leaf spots, Holes on leaves, Leaf spots and holes on leaves, Yellowing of leaves, Leaves are reddening, Spots on bolls, Cotton leaf/boll tips are reddening, Cotton bolls turning red, Leaves are reddening, Leave are eating by insects, Cotton Squares are eaten by insects, Twigs are twisting, Complete wilting of cotton plant, Boll cracking, Boll infested by caterpillars, Reddening on back side of leaves.

The observations are as follows.

- The farmers are able to identify crop diseases through only generalized symptoms.
- It is difficult for an agricultural scientist to understand the specific crop disease through the general symptoms of the disease expressed by the farmer.
- For proper identification of the crop disease, the agricultural scientists need several finer visual symptoms of the corresponding crop disease.

Overall, the following are the key outcome of the field experiments.

- Farmers are unable to identify the crop disease.
- Farmer is able to identify generalized symptoms
- Agricultural expert identifies the crop disease by confirming the specialized symptoms.
- During analysis, it was observed that even though the farmer does know the name of the specialized symptom, once the scientists possess related questions, he/she is able to confirm in the field.

The following issue has been conceived based on the observations of the experiments.

Research Issue: How to build a tool to enable the farmer such that he/she starts with the generalized symptom of the crop disease and gradually identifies the crop disease by confirming specialized symptoms?

3.3 Brainstorming with agricultural scientists and Developing basic idea

A brainstorming meeting was held with agricultural scientists to develop a core idea for building a tool for farmers. The generalized symptom "Leaf curling" in the cotton crop was considered. During the discussion, when the farmer says "Leaf curling", the crop may suffer from one of the following diseases. White Flies, Aphids, Thrips, Mealy Bug, Jassids, Leaf Curl Virus, and Zinc Deficiency.

It was noted that each disease can be identified with a set of visual symptoms. The visual symptoms of each disease are given in the Table 3.2. It was also noted that there is at least one distinct specialized symptom for each disease, which has to exist in the crop for that disease to occur.

From this brainstorming session, a view was developed that it is possible to build a hierarchy of symptoms of all crop diseases and farmers could start with generalized symptoms and start confirming the existence of specialized symptoms.

Table 3.2: Leaf curling generalized symptom based data collection

Disease	Symptoms
White Fly	Leaf curling, Presence of insects, Insect color - white, Leaf discoloration, Discoloration - scattered and irregular, Black sooty growth, Small white insects flying
Aphids	Leaf curling, Presence of insects, Insect color - black, Leaf discoloration, Black sooty growth, Black sooty growth,Reduction in leaf size
Thrips	Leaf curling, Presence of insects, Leaf discoloration, Presence of white or silvery spots on lower side,
Mealy Bug	Leaf curling, Presence of insects, Insect color - white, Insect color white - with waxy lumps, Leaf discoloration, Discolorization - along with leaf margins, Black sooty growth
Jassids	Leaf curling, Presence of insects, Insect color - green, Leaf discoloration, Leaf curling from margins
Leaf Curl Virus	Leaf curling, Leaf discoloration, Discoloration - scattered and irregular, Reduction in leaf size,Leaf curling from margins, Cup shaped curled young leaves,
Zinc Deficiency	Leaf curling, Leaf discoloration, Discoloration - scattered and irregular, Reduction in leaf size, Intervenal chlorosis between veins

3.4 Proposed Problem Framework

Normally, in the agricultural domain, the given crop disease is described in terms of visual symptoms. We first explain the corresponding terminology.

- Visual symptom: A visual symptom is the feature of crop disease which can be identified by simply looking at it. For example, “Leaf curling” is a visual symptom in the Cotton crop.
- Visual symptom question: Each visual symptom can be posed as a question. For example, for “Leaf curling” visual symptom, the corresponding question is, “Are leaves curling?”.
- Crop Disease: A crop disease (or a problem) is a set of visual symptoms. A field suffers from several crop diseases. For example, “Spotted Boll Worm”, is a disease in the Cotton crop, The visual symptoms of “Spotted Boll Worm” are { Bored squares/bored flower buds, Bored bolls, Excreta in the bolls, Premature shedding of bolls, Premature opening of damaged bolls, The larva inserts its head and 3/4 of the body inside the boll, Drying of terminal buds, The caterpillars of have a number of black and brown spots on the body }
- Generalized visual symptom: A generalized symptom is an abstract symptom that expresses only broad details of its appearance. Usually, a generalized symptom exists in multiple crop diseases.
- Specialized visual symptom: A specialized symptom is a symptom with the fine-grained details about its appearance. Usually, the specific symptoms exists in very few crop diseases.

- Parent-child: The visual symptoms of the crop diseases can form a hierarchy (or a tree) through a parent-child relationship among the symptoms. Root has no parent and specialized symptoms have no children. A parent should have at least two children. A parent is called generalized symptom and its children may be either a generalized symptom or a specialized symptom.

We assume that a farmer or agricultural expert can diagnose the crop disease by confirming the specialized symptoms. The following are the assumptions concerning the proposed problem. The proposed system can not work if the assumptions fail for any crop disease.

- The existence of crop disease (or problem) is confirmed by confirming the corresponding set of visual symptoms.
- If the farmer is in the field and a question is posed for concerning visual symptoms, he/she is able to confirm the existence or absence of it.

However, as observed from preceding sections, it is difficult for the farmer to perceive and confirm the visual symptoms. But, normally, if the corresponding question of visual symptom is posed, the farmer is capable to confirm the existence of the visual symptom.

Suppose, we identify all specialized visual symptoms and convert them into questions. Assuming that the farmer has a smartphone, the issue is how to pose the symptom-based questions to the farmer. In the worst case, we have to pose all questions of the corresponding crop diseases to confirm the existence of the crop disease. Such a system is very inefficient.

On the other hand, as explained above, several specialized symptoms can be mapped to one generalized symptom. By considering the available symptoms as bottom-level, we can map bottom-level symptoms to the next-level visual symptoms. These next-level symptoms can be mapped to subsequent-level symptoms. In this way, by forming a parent-child relationship between the symptoms, it is possible to form a hierarchy of symptoms. The idea is to build a system that starts with a few high-level symptoms and asks the farmer to confirm the existence of these symptoms on the field through a smartphone. Based on the confirmation, the system starts posing next-level questions and finally asks the farmer to confirm the bottom-level symptoms. The final disease is identified based on the confirmation of bottom-level symptoms.

Let us elaborate on higher-level symptoms of the crop disease. Normally, in the agriculture literature concerning crop disease diagnosis, the high-level symptoms are not provided clearly.

With generalized and specialized symptoms for each crop disease, it is possible for the farmer to start with generalized systems and the ability to identify the crop disease by pruning the symptoms concerning other crop diseases.

The question is, how to identify generalized systems given the specialized symptoms? The problem is stated as follows.

Research Issue 1: Given the crop, the list of crop diseases and specialized symptoms of each crop disease, develop a protocol to gather the high-level symptoms from agricultural scientists.

Suppose, we have obtained all generalized and specialized systems of all diseases of the given crop. Normally, the farmer starts to detect the crop disease by selecting the crop. Next, he will identify the part of the plant which is appearing as problematic. Based on his/her perception, different farmers follow different paths. the issue is how to develop an efficient question-popping algorithm to enable the former to identify the crop disease through a smartphone? The problem is stated as follows.

Research Issue 2: Given the crop, the list of crop diseases and specialized/generalized symptoms of the crop disease, develop a question popping algorithm to enable the farmer to identify the crop disease.

3.5 Research methodology

A research methodology is a set of guidelines for solving a problem. The common components of a research methodology are phases, tasks, methods, techniques and tools [21]. Two research paradigms have been mentioned in [21]: The scientific method and the mathematical method. The scientific method: Observe the world, propose a model of the theory of behavior, measure and analyze, validate the hypothesis of the model or theory, and if possible repeat the procedure. The mathematical method: Propose a formal theory or set of axioms, develop a theory, derive results and if possible compare with empirical observations.

The scientific method can be divided into two paradigms.

- The engineering method: Observe existing solutions, propose better solutions, build/develop, measure and analyze, and repeat the process until no more improvements appear possible.
- The empirical method: Propose a model, develop statistical/quantitative methods, apply to case studies, measure and analyze, validate the model and repeat the procedure.

The choice proposed for this research was the engineering method. We have observed the existing methods, proposed solutions, refined in an iterative manner until a workable solution with good results.

3.6 Summary of the Chapter

In this section, we have explained the background which consists of several observations regarding problem context. Next, we have discussed the field experiment which has been conducted to have insights into the problem. Lastly, we have presented the proposed research problems of developing approaches for collecting the visual symptoms of the crop from subject matter specialists and developing a system to help the farmer to identify the crop disease with a smartphone. In the next section, we present the proposed approaches.

Chapter *4*

Concepts of Crop Darpan

As explained in the preceding chapter, the goal is to develop a system, which we call as Crop Darpan system, to help the farmer to diagnose the crop disease when the farmer visits the field with a smartphone. (Here, the word "Darpan" means "mirror" in the Sanskrit language. So, Crop Darpan means "Crop Mirror"). As a part of this, two research issues were identified: First is to develop a protocol to collect the knowledge of disease symptoms from agricultural domain experts and the Second, given the visual symptoms of crop diseases of a given crop, develop a question popping algorithm to enable the farmer to identify the crop problem.

In this chapter, we present the details of Crop Darpan. In this connection, we present the Knowledge acquisition protocol and Question Popping model to enable the development of the Crop Darpan system for a given crop. The Crop Darpan framework has been developed in an incremental manner. In total, five frameworks have been developed.

This chapter is organized as follows. In the next section, we explain the overview of the basic idea of Crop Darpan and the overview of the proposed frameworks. Next, we explain the details of each framework. This includes the final framework, which is called a Hierarchy-based protocol for knowledge acquisition and Hierarchy model for question popping with Multi-questions per screen (HHM) protocol.

4.1 Basic idea, System components, and Overview of Proposed Frameworks

In this section, we present the basic idea and explain the components of Crop Darpan. Next, we explain the design choices and list the proposed frameworks.

4.1.1 Basic idea

Crop Darpan is a visual symptom-based human-computer interactive crop disease diagnostic tool. The ideal system should give the farmer an impression that he/she is interacting with a human expert [71]. That is, the system should try to respond like an expert in the diagnostic conversations between the

farmer and expert. So, we have made an effort to build a system that could mimic the human agricultural expert.

The basic idea of Crop Darpan is as follows. Given the list of crop diseases and the corresponding visual symptoms, the issue is to pose the corresponding questions to the farmer (he/she is having a smartphone and in the field) so that he/she can confirm the existence of symptoms by seeing the crop. As explained in the preceding chapter, it is assumed that the farmer has the ability to correctly examine a symptom present in the crop plants. The objective of the Crop Darpan system is to resemble the agricultural expert. The expert's knowledge about the visual symptoms of all diseases should be embedded in the system.

Usually, in the disease diagnostic conversations, we can fairly assume that initially both the farmer and expert don't know the name of the crop disease; expert enquires the farmer about the crop symptoms through a series of questions and finds the disease through deductive reasoning. Finally, the expert informs the farmer about the disease and treatment measures. We found this closely analogues to the 20 Questions Game (see Section 2.4). Table 4.1 shows the analogy between the 20Q game and the Disease diagnosis process. Agricultural expert acts as the questioner and asks the farmer (answerer) a series of yes or no questions and finds the disease name. Except that, in the case of Crop Darpan, the farmer (answerer) also doesn't know the disease (subject). Also, in 20Q game, the questioner can name a subject to ask whether it is correct or not, for example, the questioner can ask a question like *is it a chair?*, but in case of disease diagnostic conversations, an expert can not name a disease and ask whether it is the actual disease or not, an expert can ask questions only in terms of visual symptoms.

In the 20Q game, the questioner should know the properties/attributes for every subject in the domain of all subjects and the answerer must have the ability to correctly examine the presence of an attribute of the subject. No two subjects can have all the same attributes i.e every subject must be able to differentiate by at least one combination of one or more symptoms. In case of crop diagnosis also we can expect that expert knows all the diseases and their visual symptoms.

Analogy	
20 Question Game	**Crop Darpan**
Questioner	Agricultural Expert
Answerer	Farmer
Unknown Subject	Unknown Disease
Subject Visual Attributes	Disease Symptoms

Table 4.1: Analogy between the 20 Question Game and Crop Darpan

4.1.2 Components of Crop Darpan

The Figure 4.1 shows the components of the Crop Darpan system. The components are Knowledge acquisition protocol, Knowledgebase, Question popping model, and user interface. We explain these components in detail.

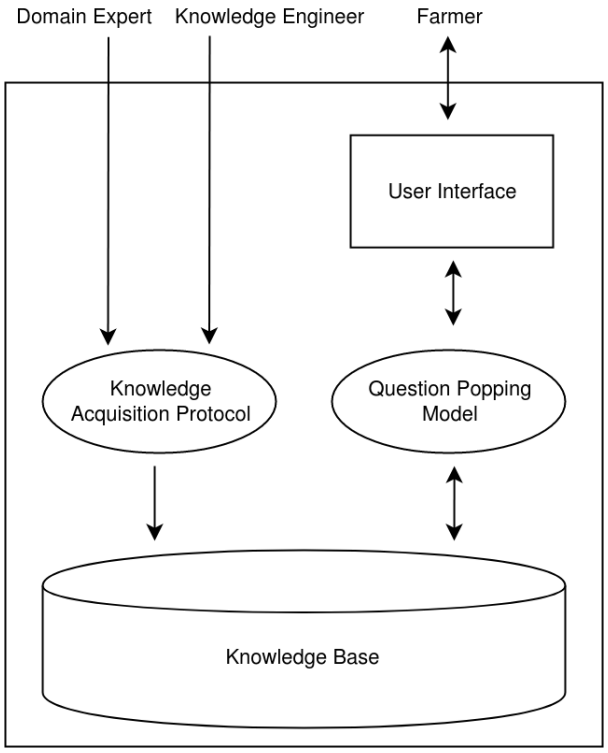


Figure 4.1: Components of Crop Darpan system

- **Knowledge acquisition protocol:** Knowledge acquisition is the process of collecting knowledge from various sources such as domain experts, textbooks, reports, etc. Knowledge engineer analyzes the information requirements and interacts with the domain experts to extract the knowledge by following a protocol. The Knowledge acquisition protocol may include multiple techniques to deduce the knowledge from the expert such as interviews/interactions, diagram based techniques, list-based techniques, matrix-based techniques, hierarchy-generation techniques [39, 52].

- **Knowledge-base:** The collected knowledge usually contains facts, rules, entities and their relationships, heuristics, etc. The knowledge extracted from the expert is organized using appropriate data structures and stored/updated in the knowledge base.

- **Question popping model:** It uses the knowledge available in knowledge-base and interacts with the farmer through the user interface. The Question popping model displays/ask the user and receives the responses. The Question popping model contains the algorithm, which dynamically decides the questions to ask based on the farmer's responses and the information at the disposal.

- **User Interface:** Farmer interacts with smartphone. An appropriate user-interface is required to enable the farmer to use Crop Darpan in a simplified manner.

4.1.3 Overview of the Proposed Frameworks

The development of Crop Darpan has been carried out in an incremental manner. Overall, the following design choices were identified: Number of questions displayed per mobile screen, Knowledge acquisition protocol, and Question popping model. We explain these choices briefly.

i **Number of questions displayed per mobile screen:** There are two options for this design choice. Based on the number of questions displayed on the user interface, the framework can be classified as either a single question framework or a multi-question framework. Based on the type of framework, the knowledge requirements may vary and so is the Knowledge acquisition protocol. We have started designing the system with a single question per screen. Later, we have opted for multiple-questions per screen.

ii **Knowledge acquisition protocol:** The issue is to develop a protocol to collect the list of diseases and the corresponding visual symptoms of the given crop from the domain experts. The objective is to minimize the effort of domain experts in providing symptoms. In this connection, we have developed two protocols. Initially, we have followed a set-based protocol, which allows collecting a set of visual systems from domain experts. To improve performance, as an improved version, we adopted a hierarchy-based protocol, which allows collecting an hierarchy of visual systems from the domain experts.

iii **Question popping model:** After collecting the knowledge of visual systems of crop diseases for a crop from the domain expert, the issue is to develop a model to organize the knowledge base and develop an algorithm to the display visual symptoms to the farmer in an efficient manner. The objective is to enable the system to mimic the agricultural expert. It should render the questions regarding visual symptoms like an agricultural expert. In this connection, we have developed five models: Entropy-based, Frequency-based, Hybrid, Coverageset-based and Hierarchy-based models. We have started with an entropy-based model and finally evolved into a hierarchy-based model based on the feedback from the farmers and the stakeholders, through continuous refinement.

The entropy-based model uses the concept of "entropy" as a criterion to decide the questions to ask the farmer. Similarly frequency-based model uses the "frequency of each symptom" to decide the questions. The hybrid model employs both "entropy" and "frequency" criteria to decide the questions asked at each time step. These models treat each disease as a set of visual symptoms. So, we have developed a set-based knowledge acquisition protocol to deduce the disease symptom relationships from the expert.

Under the multi-question framework, we have proposed two models named coverageset-based model and the hierarchy-based model. The coverageset-based model extracts multiple questions at each time step which ensures that no diseases are missed from the diagnostic procedure. This model uses the disease symptom relationships obtained under set-based knowledge acquisition

protocol. The hierarchy-based model aims to maintain the relatedness between the questions asked in the successive time steps. This model requires symptom-symptom relationships along with disease-symptoms relationships. So, we have developed a hierarchy based knowledge acquisition protocol that extends the previously mentioned set based knowledge acquisition protocol to deduce both the disease-symptoms relations and symptoms-symptom relations.

Table 4.2: Frameworks and their design choices

Name of the framework	Knowledge acquisition protocol	Question popping model	# Questions displayed
Set-based protocol for knowledge acquisition and Entropy-based model for question popping with Single-question per screen (SES)	Set-based	Entropy-based	Single question
Set-based protocol for knowledge acquisition and Frequency-based model for question popping with Single-question per screen (SFS)	Set-based	Frequency-based	Single question
Set-based protocol for knowledge acquisition and Hybrid model for question popping with Single-question per screen (SHS)	Set-based	Hybrid	Single question
Set-based protocol for knowledge acquisition and Coverage-set-based model for question popping with Multi-questions per screen (SCM)	Set-based	Coverage set-based	Multi question
Hierarchy-based protocol for knowledge acquisition and Hierarchy-based model for question popping with Multi-questions per screen (HHM)	Hierarchy-based	Hierarchy-based	Multi question

By combining the three design choices, in total, five frameworks have been developed. The details of frameworks design choices are given in Table 4.2.

i. Set-based protocol for knowledge acquisition and Entropy-based model for question popping with Single-question per screen (SES)

ii. Set-based protocol for knowledge acquisition and Frequency-based model for question popping with Single-question per screen (SFS).

iii Set-based protocol for knowledge acquisition and Hybrid model for question popping with Single-question per screen (SHS)

iv Set-based protocol for knowledge acquisition and Coverage-set-based model for question popping with Multi-questions per screen (SCM)

v Hierarchy-based protocol for knowledge acquisition and Hierarchy-based model for question popping with Multi-questions per screen (HHM)

In the remaining sections, we present these frameworks.

4.2 Set-based protocol for knowledge acquisition and Entropy-based model for question popping with Single-question per screen (SES)

This is the first framework we have developed and operated in the field and got the feedback. As a part of the SES framework, we have developed a set-based knowledge acquisition protocol to collect the data of the visual symptoms of crop diseases for a crop. We have developed the entropy-based question popping model for organizing the visual symptom knowledge database and the corresponding question popping algorithm to enable the farmer to identify the crop disease. We also got feedback from the stakeholders.

In this section, we present the set-based knowledge acquisition protocol, entropy-based question popping model and the feedback.

We have considered a Cotton crop and developed Crop Darpan. So, in all frameworks, we present the examples concerning Cotton crop.

To put the things in perspective, the design choices which have been opted under SES framework are as follows.

- Single question per screen.
- A set-based protocol is followed for knowledge acquisition.
- Entropy-based model for question popping algorithm

We first explain the aspect of a single question per screen. Next, we present the set-based model for knowledge acquisition. Subsequently, we present the entropy-based model and the corresponding question popping algorithm.

4.2.1 Single question per screen

We have considered one question per screen. It means, the system displays the single question concerning visual symptom at a time to the farmer. The farmer with the smartphone examines the field and clicks the YES/NO option based on the presence/absence of the symptom in the field.

Table 4.3: Data Collection: Sample Disease-Symptoms data collected from Experts

Expert	Disease	Symptoms set
Expert 1	White flies	curling of leaves, discoloration of leaves, scattered discoloration, stunted growth, premature shedding of bolls, presence of insects, insects on the lower side of the leaf, black sooty on leaves, chlorotic spots on leaves
	Aphids	curling of leaves, stunted growth, discoloration of leaves, dried leaves, black sooty on leaves, presence of insects, insects on the lower side of the leaf, black color insects, reduction in leaf size

	Bacterial blight	spots on leaves, shedding of leaves, blackening of leaves veins, splitting of stems, staining of lint, reduced seed size, premature shedding of bolls, dark black irregular spots, yellowing of lint
	Thrips	curling of leaves, discoloration of leaves, dried leaves, bolls turn black, curling of leaves only upward, presence of insects, green color insects
	Zinc deficiency	curling of leaves, discoloration of leaves, upward cupping of leaves, shedding of flowers, bronzing and interveinal chlorosis appear, scattered irregular discoloration, brittle leaves
	Mealy bugs	curling of leaves, spots on leaves, white waxy growth on stems, dried leaves, presence of insects, white insects in waxy lumps, discoloration along leaf margins, black sooty on leaves
Expert 2	Angular leaf spot	dark black irregular spots leaves, shedding of leaves, blackening of leaves veins, splitting of stem, yellow color lint, reduced seed size
	Thrips	curling of leaves, fading of leaves, dried leaves, bolls turn black, upwards curling of leaves, green insects
	Zinc deficiency	curling of leaves, fading of the leaves, Upward cupping of leaves, brittle leaves, shedding of flowers, interveinal chlorosis between veins
	Mealy bugs	curling of leaves, yellow spots on leaves, powdery growth on stems, white insects in waxy lumps, fading of leaves along leaf margins, sooty mold development on leaves
	White flies	curling of leaves, fading of leaves, stunted growth, premature shedding of bolls, white insects, sooty mold development on leaves, chlorotic spots on leaves
	Aphids	curling of leaves, stunted growth, fading of leaves, dried leaves, sooty mold development on leaves, black insects, reduction in leaf size
Expert 3	Mealy bugs	leaves curling, yellow spots on leaves, white insects in waxy lumps, black sooty on leaves
	Thrips	leaves curling, discoloration of leaves, wilting of leaves, bolls turn black, leaves curl upwards, presence of green insects
	Angular leaf spot	presence of black spots on leaves, shedding of leaves, reduced seed size blackening of leaves veins, splitting of stems, lint turns yellow
	Aphids	leaves curling, discoloration of leaves, stunted growth,reduction in leaf size, discoloration of leaves, wilting of leaves, black sooty on leaves, presence of black color insects on the lower side of the leaves
	White flies	leaves curling, scattered discoloration of leaves, stunted growth, premature shedding of bolls, presence of white color insects on the lower side of the leaves, black sooty on leaves, chlorotic spots on leaves

4.2.2 Set-based knowledge acquisition protocol

We have developed a knowledge acquisition protocol that consists of three stages: Data collection, Data transformation, and Data pre-processing.

By considering the Cotton crop, we explain the stages.

- **Data Collection:** To collect visual symptoms for the diseases of the Cotton crop, We have interacted with multiple agriculture domain experts to collect the Diseases-Symptoms data.

  The protocol for collecting diseases-symptoms data for a given crop from an agricultural scientist has two steps: firstly, list out all the diseases that occur in the given crop; Secondly, for each disease in the above-created list, enumerate all the related visual symptoms of it.

  As visual symptoms of a crop disease is a collection of visual symptoms, we call it as a set-based knowledge acquisition protocol

  Table 4.3 shows a sample disease-symptoms data collected from multiple experts.

- **Data Transformation:** As we have collected the Diseases-Symptoms data for the Cotton crop from multiple agricultural experts, we have to consolidate the data from multiple sources to a single coherent data repository. Usually, while merging the data from multiple sources we come across multiple issues such as duplication, logic conflicts, inconsistent data, and missing data [60].

  In the context of disease-symptom data, duplication problems can be defined as - a real-world entity (disease or symptom) being referred with multiple different names even though they have the semantically same meaning. For example, “Bacterial blight” and “Angular leaf-spot” are two different names for the same disease. Different experts can use different phrases to express the same symptom. For example “Leaf Curling” and “Curling of Leaves” express the same meaning but they will be interpreted to be different by string comparison operation. Sometimes, the data may result in violating the basic logical premise that diseases must have a unique set of symptoms and no two diseases should have equal symptom sets or subsets to one another. Different experts may list out symptoms at different abstraction levels. For example, in the sample data (Table 4.3), Expert 1 mentioned that white flies diseases have insects on the lower side of the leaf without mentioning the color of insects, whereas Expert 2 mentioned that the insects are of white color. Some experts may not consider minor diseases or symptoms but some may include them. Issues of this kind can be resolved by converting the unstructured data into structured data as the structured data gives the ability to compare the entities involved.

  The collected data in the form of sets has transformed into a matrix or a table; we refer to this table as Disease-Symptoms Co-occurrence table. The union of all the visual symptoms sets of all diseases collected from all the experts results in a list of all possible symptoms which can be exhibited by a crop. Each visual symptom in this list is allocated to a row in the Diseases-Symptoms Co-occurrence table. Similarly, for each disease in the list of all the diseases collected

from all the experts, a column is allocated in the Disease-Symptoms Co-occurrence table. The cell in the table is filled either by 1 or 0. A cell is filled with 1 if the symptom of the corresponding row occurs in the disease of the corresponding column and if the symptom doesn't occur in the disease the cell is filled with 0. While using the previously collected data and by contacting the agricultural experts all the cells in the Disease-Symptoms co-occurrence table are filled. Table 4.4 shows the Disease-Symptoms co-occurrence table for the sample data in Table 4.3. From Table 4.4, we can observe that this data structure results in a vector for each entity. Each disease gets a binary column vector representation and each symptom gets a binary row vector representation.

- **Data pre-processing:** The data cleaning has been carried out with the help of the experts. We examine and modify the transformed data iteratively until all the following conditions are satisfied.

  1. All the diseases should have a semantically unique meaning.
  2. No two disease vectors should be equal.
  3. No disease vector should be a subset of another disease vector.
  4. All the symptoms should have a semantically unique meaning.
  5. Two symptom vectors can be equal.
  6. A symptom vector can be a subset of another symptom vector.
  7. A disease should have at least one symptom.
  8. A symptom should occur in at least one disease.
  9. The diseases and the symptoms can engage in many to many relationships. I.e a disease can have multiple symptoms and a symptom can occur in multiple diseases.

  Using bit-wise $XNOR$ and $Implies$ ($\implies$) operations we can find if two-bit vectors are equal or not and subsets or not respectively. Two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ are equal if the result of their bit-wise $XNOR$ operation is $\overrightarrow{1}$ (all elements are 1). The magnitude of the resultant vector of $\overrightarrow{a} XNOR \overrightarrow{b}$ also signifies the similarity between $\overrightarrow{a}$ and $\overrightarrow{b}$ vectors. And if the result of $\overrightarrow{a} \implies \overrightarrow{b}$ is a $\overrightarrow{1}$ (all elements are 1), then $\overrightarrow{a}$ is subset of $\overrightarrow{b}$. Now that the entities are comparable, we can address the issues of duplication and logic conflicts. The task of resolving the duplicates is called "De-duplication or "Entity resolution" [35]. Entity resolution works based on the principle that even though the real-world entity referred with multiple names its properties or attributes will be the same everywhere it is referred. So, to remove the duplicate diseases, we find the pairwise similarities between the diseases using their vector representation. The most similar pairs are examined and eliminate the duplicate records if there are any. A similar process is followed to remove duplicate symptoms as well. Conditions 1, 2, 4, 5 ensures the entity resolution in both diseases and symptoms. Conditions 3, 6, 7, 8, 9 resolves the fundamental logical premises conflicts. The issues of data inconsistency and missing data are dealt with in the second stage: data transformation. Table 4.5 is the result of data cleaning process on the sample Diseases-Symptoms

Table 4.4. We refer to the final resultant table as Disease Fingerprint Table (DFT) as in this table each disease gets a unique binary vector representation. The diseases-symptoms co-occurrence table (Table 4.4) contains 7 diseases and 51 symptoms, which are compressed to 6 diseases and 30 symptoms in DFT (Table 4.5) through duplicate elimination and merging processes in data cleaning stage.

Table 4.4: Data Transformation: Disease-Symptoms Co-occurance Table

Symptom/Disease	White flies	Aphids	Bacterial blight	Thrips	Zinc deficiency	Mealy bugs	Angular leaf spot
curling of leaves	1	1	0	1	1	1	0
discoloration of leaves	1	1	0	1	1	1	0
scattered discoloration	1	0	0	0	1	0	0
stunted growth	1	1	0	0	0	0	0
premature shedding of bolls	1	0	1	0	0	0	1
presence of insects	1	1	0	1	0	1	0
insects on the lower side of the leaf	1	1	0	0	0	0	0
black sooty on leaves	1	1	0	0	0	1	0
chlorotic spots on leaves	1	0	0	0	0	0	0
dried leaves	0	1	0	1	0	0	0
black color insects	0	1	0	0	0	0	0
reduction in leaf size	0	1	0	0	0	0	0
spots on leaves	1	0	1	0	0	1	1
shedding of leaves	0	0	1	0	0	0	1
blackening of leaves veins	0	0	1	0	0	0	1
splitting of stems	0	0	1	0	0	0	1
reduced seed size	0	0	1	0	0	0	1
staining of lint	0	0	1	0	0	0	1
dark black irregular spots	0	0	1	0	0	0	1
yellowing of lint	0	0	1	0	0	0	1
bolls turn black	0	0	0	1	0	0	0
curling of leaves only upward	0	0	0	1	0	0	0
green color insects	0	0	0	1	0	0	0
upward cupping of leaves	0	0	0	0	1	0	0
brittle leaves	0	0	0	0	1	0	0
shedding of flowers	0	0	0	0	1	0	0

bronzing and interveinal chlorosis appear	0	0	0	0	1	0	0
scattered irregular discoloration	1	0	0	0	1	0	0
white waxy growth on stems	0	0	0	0	0	1	0
white insects in waxy lumps	0	0	0	0	0	1	0
discoloration along leaf margins	0	0	0	0	0	1	0
dark black irregular spots leaves	0	0	1	0	0	0	1
yellow color lint	0	0	1	0	0	0	1
fading of leaves	1	1	0	1	1	0	0
upwards curling of leaves	0	0	0	1	0	0	0
green insects	0	0	0	1	0	0	0
yellow spots on leaves	0	0	0	0	0	1	0
powdery growth on stems	0	0	0	0	0	1	0
fading of leaves along leaf margins	0	0	0	0	0	1	0
sooty mold development on leaves	1	1	0	0	0	1	0
white insects	1	0	0	0	0	1	0
black insects	0	1	0	0	0	0	0
leaves curling	1	1	0	1	1	1	0
scattered discoloration of leaves	1	0	0	0	1	0	0
presence of white color insects on the lower side of the leaves	1	0	0	0	0	1	0
wilting of leaves	0	1	0	1	0	0	0
presence of black color insects on the lower side of the leaves	0	1	0	0	0	0	0
presence of black spots on leaves	0	0	1	0	0	0	1
lint turns yellow	0	0	1	0	0	0	1
leaves curl upwards	0	0	0	1	0	0	0
presence of green insects	0	0	0	1	0	0	0

Table 4.5: Data cleaning: Disease Fingerprint Table (DFT)

Symptom/Disease	White flies	Aphids	Bacterial blight	Thrips	Zinc deficiency	Mealy bugs
curling of leaves	1	1	0	1	1	1
upward curling of leaves	0	0	0	1	0	0

discoloration of leaves	1	1	0	1	1	1
scattered irregular discoloration	1	0	0	0	1	0
discoloration of leaves along leaf margins	0	0	0	0	0	1
presence of insects	1	1	0	1	0	1
presences of insects on the lower side of the leaf	1	1	0	0	0	0
black color insects	0	1	0	0	0	0
green insects	0	0	0	1	0	0
white insects	1	0	0	0	0	1
white insects in waxy lumps	0	0	0	0	0	1
black sooty on leaves	1	1	0	0	0	1
spots on leaves	1	0	1	0	0	1
dark black irregular spots on leaves	0	0	1	0	0	0
chlorotic spots on leaves	1	0	0	0	0	0
yellow spots on leaves	0	0	0	0	0	1
dried leaves	0	1	0	1	0	0
brittle leaves	0	0	0	0	1	0
reduction in leaf size	0	1	0	0	0	0
shedding of leaves	0	0	1	0	0	0
blackening of leaves veins	0	0	1	0	0	0
bronzing and interveinal chlorosis appear	0	0	0	0	1	0
premature shedding of bolls	1	0	1	0	0	0
bolls turn black	0	0	0	1	0	0
splitting of stems	0	0	1	0	0	0
reduced seed size	0	0	1	0	0	0
yellowing of lint	0	0	1	0	0	0
powdery growth on stems	0	0	0	0	0	1
shedding of flowers	0	0	0	0	1	0
stunted growth	1	1	0	0	0	0

4.2.3 Entropy-based model for question popping

The input to question popping algorithms is the DFT table of crop diseases and the output is the sequence of questions regarding visual symptoms to be displayed to the farmer. On seeing the question on a smartphone, the farmer confirms the presence/absence by clicking the YES/NO option.

We present the three approaches

– **Rudimentary Approach:** We can follow the following strategy. The question popping algorithm decides the sequence of questions to be asked to find the disease. The most basic question popping strategy is to ask all the questions about all the symptoms/attributes. Based on the responses of the answerer for all the questions, we can form a binary vector which we call as "response vector". If the answer is "Yes", we use 1 as an element corresponding to the symptom in the vector and If the answer is "No", we use 0 as the element. So, after receiving the responses for all the questions about all the symptoms, we compare the response vector with all the disease vectors and we choose a disease as diagnosis whose vector is equal to the response vector. In this strategy, the number of questions asked is equal to the number of all possible symptoms which is typically a large number.

– **Binary Approach** In the above strategy while choosing a question to ask, the responses of previous questions asked are not utilized. The search space of diagnostic diseases contains all the diseases until the diagnosis is completed. So we can improve the question popping algorithm based on the deductive reasoning so that it utilizes the previous responses. The basic idea is: every time response received, the search space of all diseases is reduced to only feasible diseases. In the beginning, the search space contains all the diseases and we can choose any random question to ask. Once the response is received, the search space is divided into two sub-spaces: one corresponding to the response "Yes" and other corresponds to the response "No". For the next question, the sub-space corresponding to the response (Yes/No) of the current question will become the feasible search space and we eliminate all the diseases which contradict the current response and all the symptoms which are unique to those diseases. For example, if the first question is chosen is "Do you notice curling of leaves in the crop?" (refer first symptom in the sample DFT : Table 4.5), and the answer is "Yes", then we remove the "Bacterial bight" from the search space and we remove "dark black irregular spots on leaves", "shedding of leaves", "blackening of leaves veins", "splitting of stems", "reduced seed size", "yellowing of lint" from the feasible symptoms to be asked. For the second question, we can choose any random question from the reduced feasible symptoms set. This process continues until the search space of feasible diseases contains only one disease and finally, we choose that disease as the diagnosis. Compared to the previous strategy the number of questions asked will be fewer in this strategy.

– **Entropy-based Approach:** In the previous strategy the information gathered from the responses is efficiently used to reduce the search space of feasible diseases with each response and converge to the final disease. But, it is not using the information already available in DFT to converge faster. While choosing a question from feasible symptoms it chooses a random question. But, careful selection of questions can greatly reduce the number of questions to be asked to converge to the disease. As discussed earlier, with each question and response pair, questioner gains some information about the unknown subject. So, each time

the questioner should the question that provides the maximum new information about the subject.

Entropy is the measure of expected or average information content conveyed by finding the outcome of an event. The entropy H of an event X is calculated as shown in equation 4.1, where $P(x_i)$ is probability of outcome x_i and n is number of possible outcomes. In the case of the 20Q game, the event is questioner asking a question and its possible outcomes are answerer responding either yes or no. So for a question Q, the entropy is given equation 4.2, where p is the probability of response yes. The probability of answerer responding yes to a question regarding a symptom is the number of diseases in the feasible set in which the symptom is presently divided by the total number of diseases in the feasible set. Similarly, the probability of response no is calculated as the number of diseases in the feasible set in which the symptom is absent divided by the total number of diseases in the feasible set.

$$H(X) = -\sum_{i=1}^{n} P(x_i)\ log_2\ P(x_i) \tag{4.1}$$

$$\begin{aligned} H(Q) &= -P(yes)\ log_2 P(yes) - P(no)\ log_2\ P(no) \\ &= -p\ log_2 p - (1-p)\ log_2\ (1-p) \end{aligned} \tag{4.2}$$

The entropy of the question Q in equation 4.2 is maximized when $p = 1/2$. The question Q provides the maximum information about the unknown disease when its possible responses both yes and no have an equal probability $1/2$. So, at each time-step, it is best to choose the question which will split the search space of feasible diseases into two subspaces where each subspace contains exactly half the diseases from previous feasible diseases set. Using, this strategy, irrespective of the response to the question asked, at each time step, on average, we can eliminate the maximum number of diseases from the search space and converge to the actual diagnostic disease faster with the minimal number of questions.

Figure 4.2, depicts a sample decision tree produced by applying the entropy-based question popping algorithm for the sample data in Table 4.5. There are six diseases in DFT, so the first symptom (question) to be asked should be selected such that it should present exactly in three diseases. Black sooty on leaves and Spots on leaves are the candidate's symptoms for the first question as they present in exactly three diseases. When there is a tie among the symptoms, any symptom can be chosen as a question, however, their consequent decision tree (sub-tree) can be very different. In Figure 4.2, Black sooty on leaves is selected as the first symptom. Selected root node symptom splits the diseases into two subsets: {Mealy bugs, White flies, Aphids}, {Thrips, Bacterial blight, Zinc deficiency} corresponding to the response yes and no respectively. When there is an odd number of diseases in the

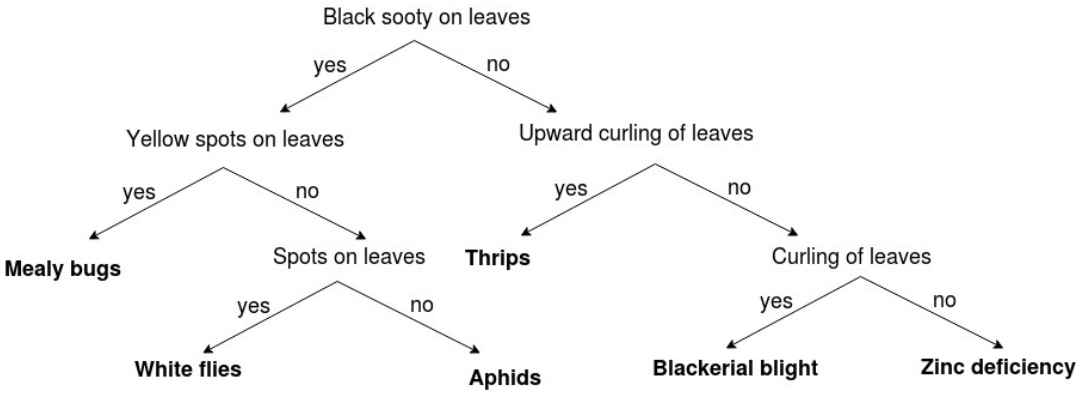


Figure 4.2: Entropy based algorithm decision tree

feasible disease set, a symptom that would split the set into two approximately equal subsets. In Figure 4.2, symptom Yellow spots on leaves is selected to split the set {Mealy bugs, White flies, Aphids} into two sets {Mealy bugs} and {White flies, Aphids} of sizes 1 and 2 respectively. A similar process is continued until all the internal nodes are symptoms and all the leaf nodes are diseases, i.e, each disease is uniquely identified by the path of symptom and response pairs.

4.2.4 Feedback

The entropy-based question popping algorithm aims to narrow down to the diagnostic disease using the minimal number of symptoms/questions. So, the algorithm chooses only the most discriminative questions to ask the user. Generally, the symptom set contains some generic symptoms and some specific symptoms. For example, upward curling of leaves and yellow spots on leaves are specific symptoms and their corresponding generic symptoms are curling of leaves and spots on leaves respectively. The entropy-based algorithm doesn't consider this generalization-specialization nature among the symptoms. Usually, in the disease diagnostic conversations, the experts first inquire about the generic symptoms and after confirming their presence experts would inquire about the specific symptoms. For example, the expert first confirms the presence of the insects and only after that he/she would inquire about the color of the insects. whereas, the entropy-based algorithm doesn't pay attention to the order of the questions being asked. It eliminates and doesn't ask most of the generic symptoms as they present in many diseases and are not discriminative in nature. Also, it is possible that the question popping algorithm asks a specific symptom first and then asks its corresponding generic symptom. For example, in decision tree Figure 4.2, generated from the sample data, the symptoms yellow spots on leaves and upward curling of leaves are asked prior to their corresponding generic symptoms spots on leaves and curling of leaves respectively. Also, as it eliminates some diseases from a feasible set every time it receives a response, even if the farmer responds no to every question asked the algorithm will

end up with a diagnosis. For example, in Figure 4.2, the diseases Zinc deficiency is identified as the diagnosis even if all the responses are given the user is no.

4.3 Set-based protocol for knowledge acquisition and Frequency-based model for question popping with single-question per screen (SFS)

The design choices for the SFS framework are as follows.

- Single question per screen.
- A set-based protocol is followed for knowledge acquisition. It is already explained in the preceding section.
- Based on the feedback of the entropy-based model, to improve performance, we have developed a frequency-based model.

The notion of "Single question per screen" and "set-based protocol" have been explained as a part of the SES framework. The frequency-based model for question popping is as follows.

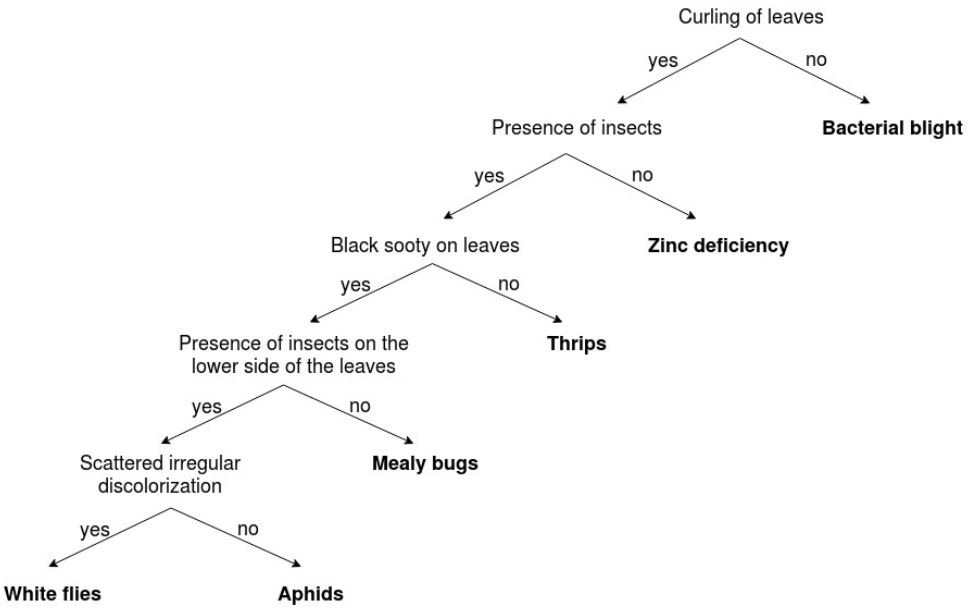


Figure 4.3: Frequency based algorithm decision tree

4.3.1 Frequency-based model for question popping

In the previous strategy question, the order does not care. There is a need for developing a question popping algorithm that minds the order of questions being asked. The motivation is to develop an algorithm that would pop up generic symptoms before their corresponding specific symptoms. By analyzing the generic and specific symptoms, it is observed that specific symptoms always co-occur with their generic symptoms. example: If the specialized symptom scattered irregular discoloration is present in a set of diseases, then its generalized symptom discoloration of leaves will definitely be present in that set of diseases. The bit vector of a specific symptom will be a subset of generic symptom bit vector and the frequency of generic symptom occurring in the diseases will be greater than or equal to the frequency of occurrence of its specific symptoms. based on this observation there is scope to develop an algorithm based on the frequency of the symptoms. The data needed is the same as in the previous strategy.

Algorithm

The algorithm follows the same framework of the 20Q game. In this, the strategy is to ask the most frequent symptom occurring in a feasible disease set every time step. Initially, the most frequent symptom in all the diseases is popped as the question. based on the response of the user, the feasible disease set is reduced by selecting the subset of diseases corresponding to the presence (response: yes) or absence (response: no) of the symptom. For the next question, the symptom which is most frequent in the reduced feasible diseases set is selected. This process is continued until a feasible disease set contains only one disease which is eventually identified as the crop diagnosis. Note that, at each step, while choosing the most frequent symptom, if a symptom is present in all the diseases in the feasible set then that symptom is ignored and next most symptom is selected as the symptom present in all the disease either not reduce the feasible set or eliminates all the diseases in the feasible set based on the user response.

Figure 4.3 shows the decision tree generated by the frequency-based question popping algorithm for the sample data in disease fingerprint Table 4.5. In DFT, both curling of the leaf and discoloration of leaves are the most frequent symptoms so anyone of these symptoms can be selected as the question to be posed. In 4.3, the curling of the leaf is chosen to be the first question so it becomes the root node of the decision tree. It is present in five diseases and absent in a single disease. So if the user response to this question is yes then the next most frequent symptom in those five diseases is chosen as the next question. In this case, discoloration of leaves is the next most frequent symptom, however, it is present in all the five diseases and it will either pass all the five diseases or none of the diseases to next step so it is ignored and second most frequent symptom presence of insects is selected. If the response of the root node question is no then the only disease in the feasible disease set Bacterial blight is identified as the diagnosis. Note that,

here, if there were multiple diseases corresponding to the response no then there would have been a left sub-tree to the root node instead of a single leaf node. This process is continued until all the diseases are identified by a unique path.

4.3.2 Analysis and feedback

Assumed equal probability of occurrence for all the diseases, the algorithm asks the most probable (frequent) symptoms in the crop. That is it tries to choose the questions which have a high probability of occurrence in the crop. It aims to choose a question that would yield a maximum number of farmers yes response. It chooses the question which has a high probability of getting a response yes from the farmer.

To compare the entropy-based algorithm with frequency-based algorithm, the former one tries to build the best possible balanced full binary decision tree using greedy approach while every internal node having better discriminative power than its internal descendent nodes, whereas the later one tries to build the best possible skewed full binary decision tree using greedy approach while every internal node having higher frequency than its internal descendant nodes. In decision trees formed by the both algorithms, the number of leaf nodes (diseases) are same (6 in sample DF Table 4.5), so total number of internal nodes (total number of questions) are same (5 in Figures 4.2 and 4.3). i.e., the average number of questions asked to diagnose a disease is the same in both the algorithms. However, in entropy-based algorithm, an almost equal number of question are employed to reach all the diseases as it builds a balanced decision tree, whereas in the frequency-based algorithm, the number question asked to reach diseases are stretched over the range of minimal to maximal number questions as it builds a skewed decision tree. Due to this, the frequency-based algorithm introduces biases and prefers diseases with rare symptoms over diseases over diseases with very common symptoms. i.e., diseases with rare symptoms are easy to reach and diseases with common symptoms takes a large number of interactions with the user and hard to reach. In decision tree Figure 4.3, Bacterial blight which doesn't have the most common symptom curling of leaves is identified with only one question-response pair but White flies which has the most common symptoms requires five question-response pairs to reach. Overall, a frequency-based question popping algorithm rectifies the counter-intuitive question ordering drawback of an entropy-based question popping algorithm. However, in the case of a frequency-based algorithm, like in an entropy-based algorithm, even if the user responds no to all the questions asked, the algorithm will end up with disease diagnosis. In the case of Figure 4.3, the disease bacterial blight will be identified as a diagnosis just if the user responds no to the root node question without confirming the presence of any of its symptoms.

4.4 Set-based protocol for knowledge acquisition and Hybrid model for question popping with single-question per screen (SHS)

The design choices for the SHS framework are as follows.

- Single question per screen.
- A set-based protocol is followed for knowledge acquisition. It is already explained in the preceding section.
- Based on the feedback of the frequency-based model, to improve performance, we have developed the Hybrid model.

The aspect of "Single question per screen". and "Set-based model for data acquisition" have already explained as a part of the SES framework. The Hydrid model for question popping is as follows.

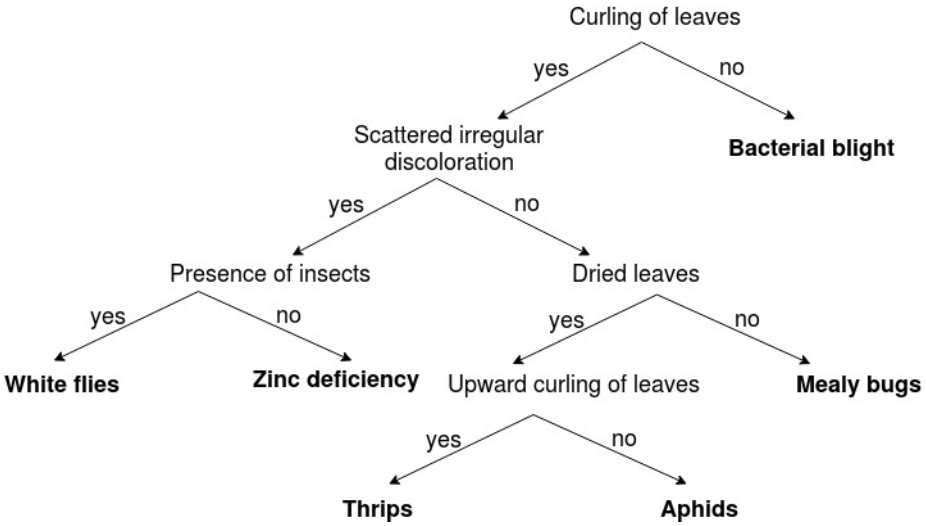


Figure 4.4: Hybrid algorithm decision tree

4.4.1 Hydrid Model for question popping

The entropy-based algorithm chooses discriminative questions and produces almost equal length traversal paths for all the diseases in the decision tree however it ignores the order in which questions asked. The frequency-based algorithm chooses common symptoms and intuitive question order is maintained however, it favors some diseases over others by producing varied length traversal paths for the diseases. There is scope to develop a hybrid algorithm that minds the order question being asked and attempts to produce almost equal traversal paths for all the diseases by

tapping advantages both entropy-based and frequency-based question popping algorithms. The basic idea of the algorithm is to ask the most discriminative generalized symptom of the most discriminative specialized symptom at each step.

Algorithm

The algorithm has two main stages. In the first stage, for each symptom, its generalized symptoms are extracted. In the second stage, the best discriminative generalized symptoms are extracted iteratively and posed to the user as a question.

In the previous sections, it is established that a symptom always co-occurs with its generalized symptoms. So a symptom's frequency will be lower than the its generalized symptoms' frequency and its bit vector will always be sub-vector[1] of its generalized symptoms' bit vectors in DFT. In the first stage, for every symptom, a set of symptoms are identified such that the symptom's bit vector is a sub-vector of every symptom's bit vector in that set. we call that set of symptoms as its super symptoms. A symptom's super symptoms contain its generalized symptoms however, not all the symptoms in the super symptoms are its generalized symptoms. For example, in Table 4.5, the super symptoms set of upward curling of leaves is curling of leaves, discoloration of leaves, presence of insects, and dried leaves in which only curling of leaves is a generalized symptom and rest of them are the symptoms that always co-occurred with the upward curling of leaves. To distinctly identify the generalization symptoms, external human input is required but using the statistical information from DFT it is only possible to find the superset of generalized symptoms but not the exact generalized symptoms. Anyone of the statistically derived super symptoms or human-generated generalized symptoms can be passed as input to the second stage of the algorithm. Even if the super symptoms are passed to the algorithm, it treats all the symptoms in the super symptoms set as generalized symptoms.

Fundamental condition to follow in the second stage, symptoms selection stage is that a symptom can be inquired if and only if all of its generalized symptoms (super symptoms) are inquired previously. The algorithm finds the best discriminative symptoms to ask while following the previous condition. At each time step, from the previously unasked symptoms, the algorithm finds a symptom with maximum entropy as a candidate symptom to be asked to the user. If all of the candidate symptom's super symptoms are asked previously then the candidate symptom has posed a question. Here, note that, if a candidate's symptom's super symptom presents in all the diseases in feasible disease set, then that super symptom need not be asked previously as it would never help to reduce the feasible diseases set, so it will never be asked as a question. If some of the super symptoms of the candidate symptom are not asked previously, then a symptom

[1]Sub-vector is defined as follows: a bit vector $\overrightarrow{a}$ is sub-vector of a bit vector $\overrightarrow{b}$ if for every element of value 1 in $\overrightarrow{a}$ there is a corresponding element in $\overrightarrow{b}$ whose value is 1 and $\overrightarrow{a} \neq \overrightarrow{b}$. Using logical implies operation, one can say $\overrightarrow{a}$ is sub-vector of $\overrightarrow{b}$ by checking the conditions $(\overrightarrow{a} \implies \overrightarrow{b}) = \overrightarrow{1}$ and $\overrightarrow{a} \neq \overrightarrow{b}$.

with maximum entropy among the super symptoms is selected as the new candidate symptom. Similarly, the candidate symptom is updated iteratively until a symptom with all of its super symptoms are previously inquired. The final updated candidate symptom has posed a question and based on the user response feasible disease set is reduced and a similar process is continued until the crop diagnosis is found.

Figure 4.4 shows the decision tree generated by the hybrid question popping algorithm. In DFT, symptom black sooty on leaves have the highest entropy. So, it is selected as the candidate symptom for the first question. Its super symptoms set is curling of leaves, discoloration of leaves, presence of insects. Among the super symptoms, the presence of insects has the highest entropy, so it is updated as the new candidate symptom. The super symptoms set of the presence of insects is curling of leaves, discoloration of leaves. Both the symptoms in the super symptoms set have equal entropy, so anyone of them can be chosen as the new candidate symptom. curling of leaves symptoms is selected as a new candidate symptom and it doesn't have super symptoms. So curling of leaves is chosen as the first question to ask. If the user response is yes then the updated feasible diseases set contain all the diseases except Bacterial blight. Now, with the reduced feasible diseases set, scattered irregular discoloration is identified as the candidate symptom as it has the maximum entropy. Its super symptom set is curling of leaves, discoloration of leaves in which curling of leaves is already asked before and discoloration of leaves presents in all the feasible diseases so the candidate symptom fulfills the fundamental conditions so it is posed as a question to the user. This process continues until all the diseases in the feasible set have a unique traversal path.

4.4.2 Analysis and feedback

The hybrid algorithm tries to strike the balance between maximizing the entropy of questions and the frequency of the questions. It acts as a middle ground between the entropy-based algorithm and frequency-based algorithm. It selects the most discriminative symptoms with their generalized symptoms asked previously. It tries to build a nearly balanced decision tree with the condition that a symptom should never have its generalized symptom as its descendant. The height of the decision tree built by this algorithm will always be greater than or equal to the height of the decision tree built by an entropy-based algorithm and will always be lesser than the height of the decision tree built by a frequency-based decision tree. So the variance among the lengths of the traversal paths of the diseases is higher and lower compared to the variance among the diseases traversal paths generated by the entropy-based and frequency-based algorithms respectively. Similar to its parent algorithms, this hybrid algorithm also suffers from all no responses problem.

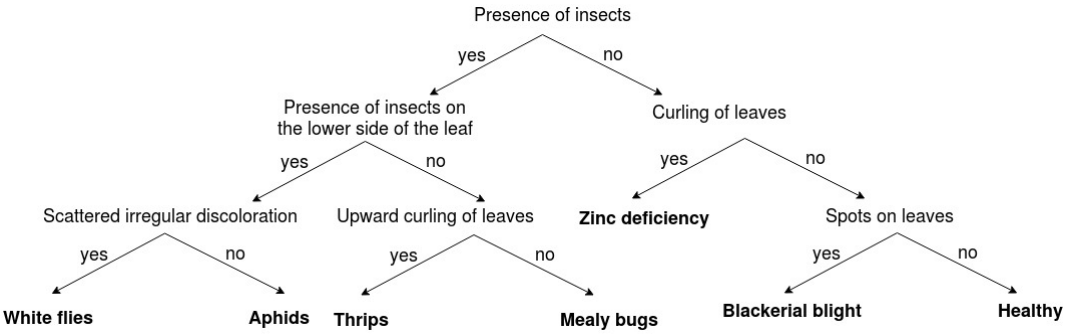


Figure 4.5: Entropy based algorithm decision tree+

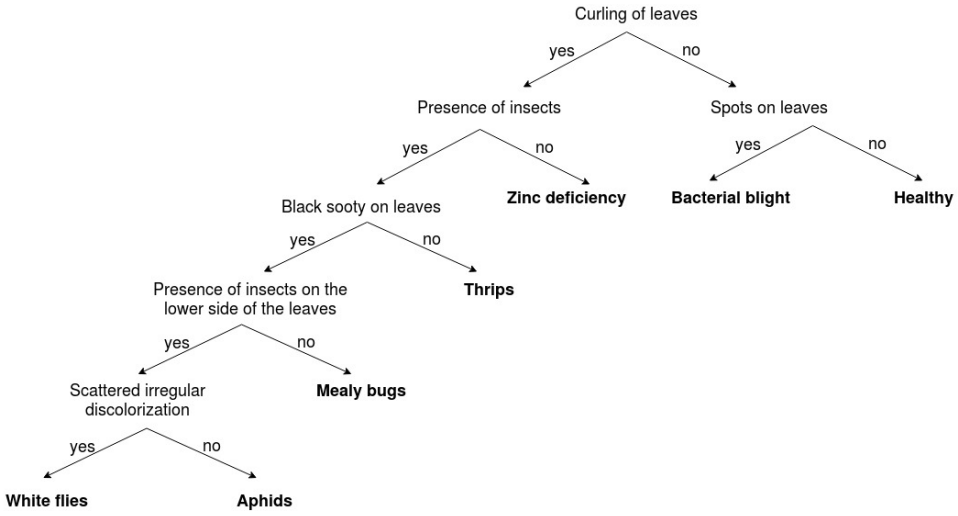


Figure 4.6: Frequency based algorithm decision tree+

4.4.3 Issue of Handling all no responses from the user

In the previous strategies, when the user is responded either yes or no, the feasible diseases set is reduced by eliminating the diseases based on the presence or absence of the symptom. As diseases are eliminated with each response, even when the user responds no to all the symptoms the algorithm would end up with a disease diagnosis which could be considered as the contradiction to common-sense expectation. In disease diagnostic conversations, when farmer acknowledges absence of all the symptoms inquired by the expert, the expected reply from the expert is that the crop is in healthy condition, as the disease diagnostic tool is expected to mimic the experts, there is need to develop a scheme which could model this common-sense behavior into the question popping algorithms.

From the previous strategies, we can observe that this issue is rising because of the elimination of the diseases from the feasible diseases set when response no is received from the user. So the simple solution is to identify diagnostic diseases based on only the presence of symptoms and

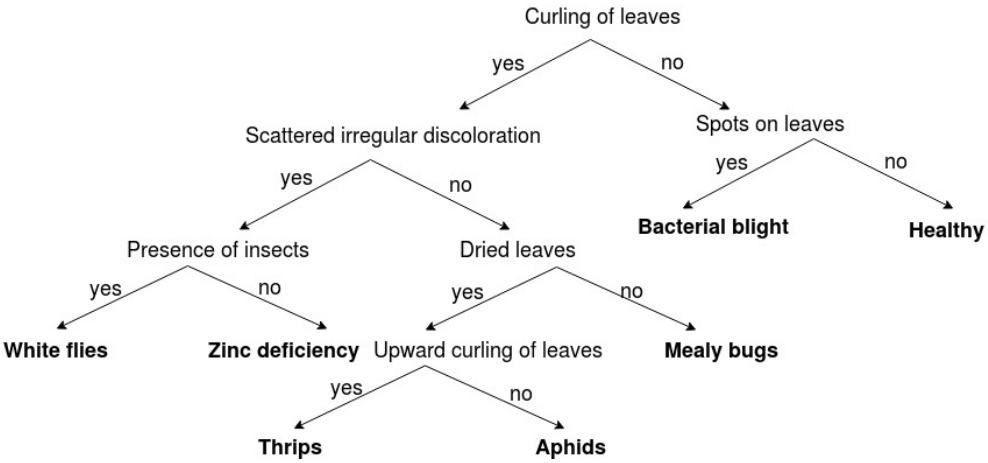


Figure 4.7: Hybrid algorithm decision tree+

not the absence of symptoms. i.e the feasible disease set is reduced based only based on yes responses and the no responses are ignored. So, if the user responds no to every question asked, all the symptoms in the DFT will be exhausted and this condition can be notified as to the crop's health condition to the user. However, this scheme may lead to the counter-intuitive question order because even after a generalized symptom's absence is confirmed by the user, its corresponding diseases are not eliminated from the feasible set so its specialized symptoms are still candidate symptoms to ask the user. For example, even after no presence of insects is confirmed, as this response is not used to eliminate the diseases White flies, Aphids, Thrips, and Mealy bugs from the feasible diseases, the symptoms black insects, white insects,.. etc are still candidates for the next symptoms.

Another solution is to treat the health condition also as a diagnostic class as ill conditions have diagnostic classes (diseases). A healthy condition of crop signifies the absence of all ill symptoms. So, a diagnostic class for the healthy condition can be created by adding an additional column vector for DFT with all of the values of its elements as 0. Now, the previously proposed three-question popping algorithms can be applied to the modified DFT. Figures 4.5, 4.6, and 4.7 shows the decisions derived by the entropy-based algorithm, frequency-based algorithm and hybrid algorithm respectively on the modified DFT. From the figures, it can be observed that if the user responds no to all the asked question then the health condition is reported as a diagnosis.

4.5 Set-based protocol for knowledge acquisition and Coverage-set-based model for question popping with Multi-questions per screen (SCM)

To put the things in perspective, the design choices for the SCM framework are as follows.

- Multiple questions per screen.
- A set-based protocol is followed for knowledge acquisition.
- Coverage-set model for question popping algorithm

The set-based model for knowledge acquisition is explained as a part of the SES framework. Next, we present the aspect of multiple questions per screen and them coverage-set model and the corresponding question popping algorithm. We also provide feedback.

4.5.1 Multiple questions per screen

All the previous question popping algorithms follow the 20 question game framework. This framework allows the interactive system to ask a single question and receive one of the yes or no responses from the user. In this framework, to ask the next question, the current question must be answered by the user. So, it compulsories the user to respond to the current question to reach the diagnosis even if he/she is not confident about the current symptom's presence/absence. The main reason for this issue in a single question framework is that for each disease there is only one traversal path to follow by the user to reach it as a diagnosis. From all the symptoms a disease has, this framework chooses a very limited number of those symptoms and uses them to identify the disease and ignores the rest of the symptoms. So, there is a need to develop a framework that would allow the user to reach the disease through multiple paths using multiple symptoms.

4.5.2 Coverage set based model for question popping

A simple solution to this issue is to display multiple symptoms on the interactive system's monitor and allow the user to pick a symptom whose presence is user most confident about. Now, one has to decide the criteria to select the symptoms/questions to display. With multiple interactions with the farmers who are using the diagnostic system, it is observed that farmers are usually confident about answering the generic symptoms than very specific symptoms. So, it is appropriate to interact with the farmer using generic symptoms first and then move to specific symptoms that distinguish the diseases well. As established earlier, the generic symptoms are usually frequent symptoms in DFT. So, the frequency criteria to choose the questions to display are compatible with the common-sense expectation of the user.

Now, one has to decide the number of questions to display on the monitor. Here, the fundamental condition that should be followed is that the set of symptoms/questions displayed must represent or act as pointers to all the diseases in the feasible set. we refer to this set of symptoms/questions as coverage sets. i.e for each disease from the feasible disease set, there should be at least one symptom in the coverage set that occurs in that disease. For a given feasible disease set, there

can be many coverage sets. However, we discussed earlier that generalized symptoms i.e most frequent symptoms are preferred by the farmers. So, we select the coverage set formed by the most frequent symptoms. As a frequent symptom covers many diseases, coverage set formed by most frequent symptoms has the advantage of having few symptoms. Here, note that coverage set with the minimal number of symptoms can be obtained by considering the most frequent symptoms with the least overlap of diseases they cover. However, usually, high-frequency symptoms have a high overlap of diseases and as a result of considering the least overlap among the coverage symptoms, the generalized symptoms get bypassed and specialized symptoms are introduced into the coverage set. On the other hand, considering the no overlap among the symptoms results in coverage set with a large number of overlapping generalized symptoms which is undesirable considering the number of questions can be displayed on the system's monitor comfortably. So, while assembling/constructing the coverage set, a symptom is added to it if and only if it covers at least one new disease. So, this scheme ensures that the number of symptoms in the coverage set doesn't exceed the number of diseases in the feasible set. It is often possible that the resultant coverage set of this scheme has a lesser number of symptoms than the number of questions a system's monitor/screen can accommodate comfortably. In this case, to utilize the screen size properly, along with the coverage symptoms next most frequent symptoms are also displayed. So, if n is the number of symptoms that the system's monitor can show comfortably and l is the size of the coverage set, then the number of symptoms to be displayed on the screen is equal to $max(n, l)$.

Figure 4.8 shows the decision tree produced by the coverage-based multi-question framework for the sample data in Table 4.5. Initially, all the six diseases in the DFT are in the feasible set. The most frequent symptom curling of leaves covers five diseases. The next most frequent symptoms discoloration of leaves, presence of insects, black sooty on leaves, etc .. do not cover any new diseases so, they are bypassed and spots on leaves are selected as it covers the only disease that is left uncovered. So, the coverage set contains curling of leaves, Spots on leaves (l=2). In this sample case, the number of questions screen accommodates comfortably (n) is 4. Therefore, the two most frequent symptoms from the symptoms list which are not in the coverage set are selected to display. The screen $[a]$ in the figure shows the final set of questions displayed. Let us say, if the farmer is confident about presenting the symptom curling of leaves, then he/she acknowledges it by selecting that symptom out of the four displayed questions. Note that, here acknowledging the presence of a symptom doesn't mean the absence of the rest of the three symptoms (unlike multiple-choice questions). Based on the symptom that the farmer responded to, the feasible set is reduced and a new set of symptoms are identified to display on the interface using both frequency and coverage criteria on the updated feasible disease set. For the confirmation of the symptoms curling of leaves, spots on leaves, discoloration of leaves and presence of insects the corresponding next screens are [b], [c], [d], and [e] respectively. This process continues until all the diagnosis is identified.

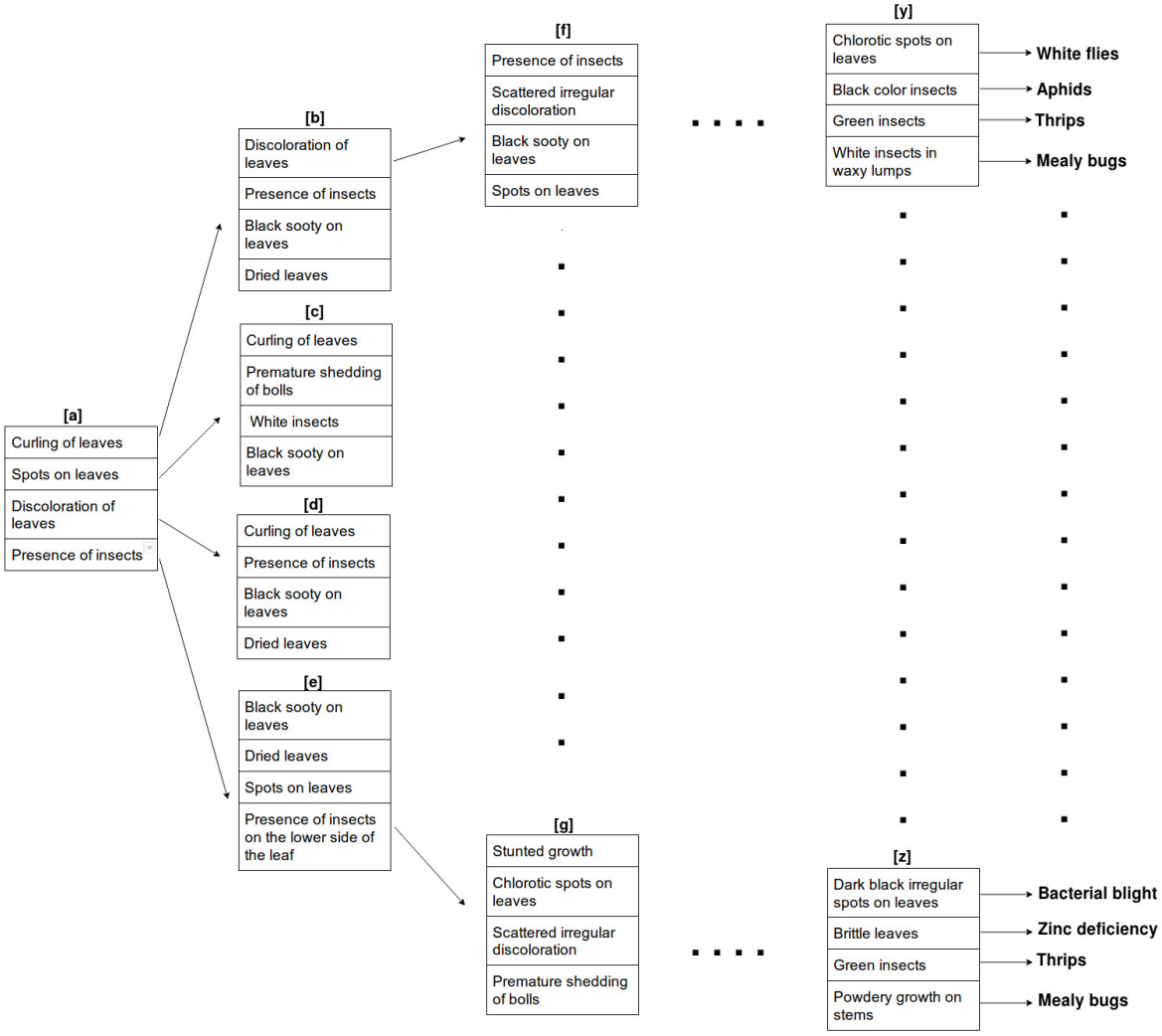


Figure 4.8: Coverage-based multi-question framework decision tree

4.5.3 Analysis and feedback

In multi questions framework, the frequency criterion ensures that the specialized symptoms are asked only after the generalized symptoms asked and coverage criterion provides the correctness and completeness as it ensures that no disease is missing from the diagnostic process. This framework realizes the diagnostic process only by considering the presence of the symptoms and not based on the absence of the symptoms (The framework provides only yes response but not no response). As the initial set of symptoms/questions displayed covers all the diseases, the absence of all these symptoms indicates the healthy condition of the crop. After multiple interactions with the farmers, it is observed that the farmers are able to respond to the question easily compared to the previous single question framework as they can respond to the question they are most confident. However, farmers are preferring a sense of relativity between the successive questions. For

example, If the farmer acknowledged the presence of insects in the crop, then he/she would expect the next question to be about its specialization details such as the color of insects is white, black or green. Even though, the system asks specialized symptoms only after the confirmation generalized symptoms, they are not asked immediately after the corresponding generalized symptoms. For example, in decision tree Figure 4.8, after confirming spots on the leaves in $[a]$, which kind of spots (dark spots, chlorotic spots, or yellow spots) are not inquired, instead the next most frequent symptoms displayed in $[c]$ are asked.

4.6 Hierarchy-based protocol for knowledge acquisition and Hierarchy-based model for question popping with Multi-questions per screen (HHM)

As established earlier, the fundamental objective of the diagnostic system is to mimic a human expert. To provide practical applications in real-world scenarios, systems such as this should be able to model real-world data efficiently in its knowledge base. As discussed in section **??**, in abstracting or conceptualizing the real-world domain data, generalization and specialization are the most important mechanisms. As generalization and specialization relations are mostly observed among real-life objects, Using these concepts to build diagnostic systems will enable efficient human-computer interaction and realize commonsensical expectations. In previously discussed models, statistical information of symptoms such as frequency and subset-vectors are used to derive the generalized symptoms and specialized symptoms relations among the symptoms. However, due to the limited information available in the DFT, for a symptom, we can derive only possible generalized symptoms (defined as super symptoms in section 4.4) but not the exact generalized symptoms. So, external human-expertise is needed to provide the generalization/specializations relations among the symptoms.

By embedding the hierarchy (generalization/specialization relations) of symptoms in a multi-question framework, we can achieve the relativity among the successive questions by asking the specialization symptoms immediately after generalization symptoms.

To put the things in perspective, the design choices for the HHM framework are as follows.

- Multiple questions per screen.
- A hierarchy-based protocol is followed for knowledge acquisition.
- A hierarchy-based model is followed for question popping algorithm

We have presented the aspect of multiple questions per screen as a part of the SCM framework. We now explain about hierarchy model for knowledge acquisition and question popping algorithm.

4.6.0.1 Hierarchy-based protocol for knowledge acquisition

In the previous models, only problems-symptoms relations are required, so a problem is considered a set of symptoms and followed a protocol (see section 4.2.2) to create data and conditions to obey. however, in the hierarchical model, besides the problem-symptom relations, symptoms-symptoms relation is required.

The relations among the symptoms can be captured by a hierarchy. A hierarchy can be realized by a rooted tree data structure. In this symptom hierarchy, the nodes represent symptoms and the edges between the nodes represent the generalization/specialization relation between the symptoms. The parent node represents a generalized symptom and its child nodes represent the corresponding specialized symptoms. The nodes at lower-level contain further specialized symptoms than the symptoms at higher-level and with each level the specificity(fine-grained details) of the symptoms increases. This kind of hierarchy is called subsumptive containment hierarchy, or taxonomic hierarchy, or IS-A hierarchy where the object classes are classified from general to specific.

The number of levels in the hierarchy depends on the criteria of the classification. After multiple interactions with the agricultural experts and studying the symptoms from the previously collected data, we found that it is appropriate to classify the symptoms firstly based on the anatomical part of the plant it belongs to and then further classify the symptoms based on their fine-grained details. So, the constructed hierarchy consists of four levels. Fig. 4.9 shows a sample tree representing the symptoms hierarchy. The first level (L_0) consists of a crop name as the root node. The second level (L_1) consists of the parts of the plant. The third level (L_2) consists of symptoms corresponding to each part of the plant. The last level (L_3) consists of the fine-grained details of the symptoms from the above level.

The direct materialization of a hierarchy involving several symptoms occurring in multiple diseases is a difficult process. So, we have developed a protocol for an agricultural scientist to easily construct the hierarchy in a modular manner. The steps involved in the protocol are as follows.

Step 1: Specify the name of the crop and list out all the parts of the plant. This step provides information about the hierarchy levels L_0 and L_1.

Example:- Crop name: Cotton; Parts: Leaves, Bolls, Stems, Flower, Whole plant, etc.

Step 2: For each part, list out the visual symptoms of all possible damages that can occur for the part. In this step, we get the information about symptoms in the L_2 level. The symptoms in this level are generalized symptoms which can be easily identified by the farmer.

Example:- Leaves: Curling of leaves, Holes in leaves, Spots on leaves, Leaves color change, etc. Stems: Bored stem, wilting of the stem, Galls on the stem, etc.

Step 3: List out all the diseases that can occur in that specific crop.

Example:- Aphids, Leaf hoppers, Bacterial blight, Zinc deficiency, etc.

Step 4: For each disease, list out all the specific symptoms that occur in the disease. This step provides information about symptoms in the L_3 level. These symptoms are generally used by experts to identify crop diseases. As this information is available in many agricultural textbooks (knowledge sources), this step can be carried out agricultural experts very easily.

Example:- Leaf hoppers: Upward curling of leaves, Reddening of leaves from margins, Brown necrotic patches on leaves, Stunted growth.

Step 5: Link the symptoms from Step 4 (L_3 level) as children to the corresponding generic symptoms obtained from Step 2 (L_2 level). This step may involve in creation of new generic symptoms if some of the specific symptoms in L_3 doesn't have any related generic symptoms in L_2 and removal of generic symptoms in L_2 if there are no corresponding specific symptoms in L_3.

Example:- Upward leaf curling $\rightarrow$ Curling of leaves; Downward leaf curling $\rightarrow$ Curling of leaves; Reddening of leaves from margins $\rightarrow$ Leaves color change; Dark black irregular spots $\rightarrow$ Spots on bolls; etc.

The above 5 steps are followed iteratively until the refined version of the symptom hierarchical data is satisfactory. The collected data is transformed and pre-processed by following the similar process mentioned in section 4.2.2. The problems-symptoms data validation conditions mentioned in section 4.2.2 (see page 39) are also applicable in this case also. Along with these conditions, to validate the symptoms-symptoms relation data the below conditions are applied.

1. Except root node, each node must have exactly one parent node. This condition ensures that there are no dangling nodes (symptoms with no parents) in the hierarchy.

2. Each parent symptom should have at least two child symptoms. This condition eliminates redundant symptoms in the hierarchy. If a parent (generic symptom) has only one child (specific symptom) then the parent symptom can be replaced by the child symptom.

3. The sibling nodes[2] in levels L_1 and L_2 need not be mutually independent. However, sibling nodes in level L_3 must be mutually exclusive. For example, the symptoms (nodes) Leaves damage & Bolls damage are siblings in level L_1. The presence of leaf damage need not necessarily mean the absence of boll damage and vice versa. Similarly, curling of leaves and holes in leaves are sibling nodes in level L_2 and they are not mutually exclusive to one another. However, a level L_3 symptoms, Upward curling of leaves and downward curling of leaves are children of leaf curling that are mutually exclusive. The presence of symptom Upward curling of leaves indicates the absence of the symptom Downward

[2]Sibling nodes are the nodes on the same hierarchical level under the same parent node.

curling of leaves. Similarly, Lint color turn to black, Lint color turn to pink, and Lint color turn to yellow are the specialized symptoms of Lint color change and they are mutually exclusive to one another.

4.6.1 Hierarchy-based model for question popping

In the previous models, only problem-symptoms relations are available. So, based on the statistical information such as frequency and coverage derived from the problem-symptom relations data is used in the algorithm to determine the questions to ask the user. However, the question sequences are counter-intuitive. So, the question popping algorithm must be provided with the symptom-symptom relations (hierarchy) along with the problem-symptom relations (DFT) to produce commonsensical question ordering. So, the previous multi-question popping algorithm which uses frequency and coverage criteria must be adapted to use the hierarchy information to decide the next questions to display at each step.

The fundamental principle is to ask the specialized symptoms (child nodes) immediately after asking the corresponding generalized symptoms (parent nodes). Based on this principle, we have developed a tree (hierarchy) traversal algorithm to identify the diagnosis.

The algorithms' control starts at the root node (level L_0) and then traverses the hierarchy (tree) until a diagnosis is identified. Note that, the diseases are not part of the tree however based on the nodes (symptoms) visited the feasible disease set is reduced to a singleton which is to be reckoned as the crop diagnosis. The basic conditions that the traversal algorithm follows are as below.

1. A question can not be asked/displayed again if the farmer already confirmed it.
2. A question can be asked if and only if all of its ascendant questions are already asked. This condition ensures that the generalized symptoms are asked before narrowing down to the corresponding specialized symptoms.
3. Once a symptom is confirmed by the farmer, the next questions asked must be its closest relatives. Here, relatedness between two symptoms/nodes can be estimated by finding the distance between them [76]. The shorter the distance between the nodes the more the relatedness between them. The distance between two nodes is defined as the number of edges along the shortest path between two nodes.

The algorithm's control first visits the root node and confirms the crop name from the farmer. For a node the closest nodes are its parents and children. If a question is confirmed it is expected that all its parents are already confirmed. So, after confirming a question, the children of that question are preferred. So, all of root node's children, the part level symptoms are displayed initially. Out of all these symptoms, the farmer can confirm the symptom that he/she is most confident about. Once

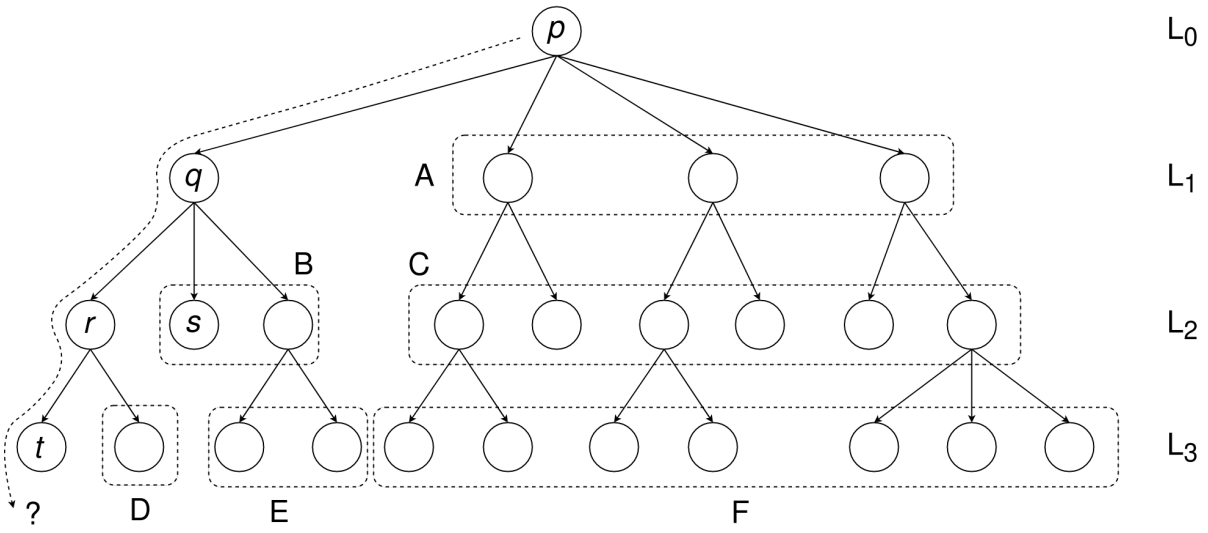


Figure 4.9: A sample tree structure representing the symptoms hierarchy

that farmer acknowledges a part level symptoms, the disease set is reduced. The candidates for the next set of questions are its children. Out of the candidate questions, the symptoms/questions forming coverage set based on the criteria of frequency and coverage are selected. This process of narrowing down from generic to specific symptoms continues until the algorithms' control reaches a leaf node. At this stage, if the feasible disease set is a singleton then the diagnosis is presented to the farmer. Otherwise, i.e, that leaf node and its parents occur in more than one disease, the needs to be reduced by jumping the algorithms control to another set of symptoms. For example, as shown in Figure 4.9, the algorithms control reached to node *t* after confirmation of nodes p, q, r, and t. Now, the candidates for next set of questions are A, B, C, D, E, and F. Here,

A: Sibling nodes in level L_1 (part level symptoms); Siblings of grand parent node *q* of node *t*.

B: Sibling nodes in level L_2 under the currently asked part level symptom *q*; Siblings of parent node *r* of node *t*.

C: Sibling sets in level L_2 under the currently not asked part level symptoms; Children of nodes in A.

D: Sibling nodes of currently asked L_3 level symptom *t*.

E: Sibling sets in level L_3 under currently asked part level symptom; children of nodes in B.

F: Sibling sets in level L_3 under currently not asked part level symptoms (A); children of nodes in C.

Based on the condition that children should not be asked before asking their parents/ancestors the sets C, E, F are eliminated from the candidate sets as their respective parents A, B, C are not asked

till now. Now, the candidates set contain A (*t*'s grand parent's siblings), B (*t*'s parent's siblings), and D (*t*'s siblings). However, *t* is in L_3 level so the presence of *t* confirms the absence of nodes in D.

The distance between *t* and the nodes in A and B are 4 and 3 respectively. So, the first preference to jump is given to B and then to A. Now, the coverage set from the B which covers all the diseases in the feasible disease set exists, then the control jumps to B and symptoms in that coverage set are displayed. If a coverage set from B cannot be derived the then nodes in A are checked for coverage set. If coverage set exists in A, then the control jumps to A and nodes in that coverage set are displayed. However, if a coverage set can not be derived either from A or B individually, that means few diseases in the feasible set are covered by nodes in B and other diseases are covered by nodes in A then the algorithm falls back on the unique symptoms of each disease. i.e a unique symptom from each disease is identified and displayed on the screen. Whichever unique symptom is confirmed by the farmer the corresponding disease is identified as the diagnosis.

If the leaf node is from the level L_2, then its siblings and its parent's siblings are the candidate sets for the jump. In the figure, node *s* is reached after confirming *p* and *q*. As this node is in L_2 its siblings are not mutually exclusive, so if nodes in B other than *s* covers the feasible disease set its siblings are displayed next. Otherwise, a coverage set is searched in A. If it is available the corresponding questions are asked. Otherwise, similar to the above case the algorithm falls back on to the unique symptoms. The algorithm flow chart is shown in 4.10.

The illustrative example of the hierarchical model is as follows. Table 4.6 shows a sample problem-symptom relations and Figure 4.11 shows a sample symptom-symptom relations (hierarchy) obtained by following the above-discussed knowledge acquisition protocol. In Table 4.6, there are 8 diseases with their corresponding symptoms. Each symptom is given a unique identification number (ID) for ease of reference. For example, the ID of symptom Boll's damage is 2 as shown in the table. The Figure 4.11 depicts corresponding hierarchy of symptoms obtained from Table 4.6 using their IDs. The L_1 level symptoms Leave damage, Bolls Damage, Stems damage, etc. has IDs from 1 to 6. The symptom Leaves damage with ID 1 has seven children (Brittle leaves, Curling of leaves, Leaves color change, etc.) with IDs from 11 to 17. The symptom Curling of leaves with ID 13 has two children Downward curling of leaves and Upward curling of leaves with respective IDs 131 and 132.

Now, we explain the question popping algorithm using four cases based on the data available in Table 4.6 and Figure 4.11.

Case 1: Initially algorithm control is at root node 0. Once the user confirms the crop name, all the L_1 (parts level) symptoms are displayed on the interface. At this time step, the feasible disease set contains all the eight diseases. Now, let's say that the user acknowledged the presence of Leaves damage (1). The diseases Pink bollworm, Boron deficiency are eliminated from the feasible disease set as they don't have leaves damaging symptoms. Now

from the children of Leaves damage (1), the coverage set {Spots on leaves (17), Leaves color change (15), Curling of leaves (13)} covers all the six diseases in the feasible disease set. So, the symptoms with IDs 17, 15, and 13 are displayed. Let's say the farmer confirmed the presence of symptom Spots on leaves (17), then the feasible disease set is reduced to {Leaf hoppers, Anthracnose, Bacterial blight, Mites}. Now, the symptoms with Ids 171, 172, 173, and 174 (the children of 17) cover all diseases in a feasible disease set. Let us say the user confirmed the presence of symptom Brown necrotic patches on leaves with ID 171. Now, the feasible disease set has reduced a singleton Leaf hoppers which is to reckon as the diagnosis. Now, the algorithm's control is at a leaf node (171), however, the diagnosis is found so there is no need to jump to any other nodes.

Case 2: Let's say the algorithm's control is at leaf node 21 as the farmer confirmed the symptoms 2 and 21 sequentially. At this time step, the feasible disease set is {Pink bollworm, Anthracnose, Bacterial blight, Mites}. As the feasible disease set is not a singleton, the algorithm checks whether it can form a coverage set from symptom 21's (L_2 level) siblings (22, 23, 24, 25) that covers the diseases in feasible disease set. Symptoms set {22, 23, 24} from L_2 level covers the feasible disease set so the control jumps to these sibling nodes by displaying them to a farmer for confirmation of any one of them. This process continues until the feasible disease set is reduced to a singleton.

Case 3: Let's say the algorithm's control is at leaf node 32 as the farmer confirmed the symptoms 3 and 32 sequentially. At this time step, the feasible disease set is {Anthracnose, Bacterial blight}. The sibling of symptoms 32 is 31 and it doesn't cover the diseases in feasible disease set. So, the algorithm checks for the coverage set in 32's parents' siblings (1,2,4,5,6). Leaf damage (1) covers both diseases in feasible disease set, So Leaves damage and few next most frequent L_1 level symptoms are displayed to utilize the screen size.

Case 4: Let's say the algorithm's control is at leaf node 132 (L_3 level) as the farmer confirmed the symptoms 1, 13, and 132 sequentially. At this time step, the feasible disease set is {Leaf hoppers, Thrips, Zinc deficiency }. Now, the algorithm checks whether leaf node parent's (node 13) siblings (nodes 11, 12, 14, 16, and 17) cover all the diseases in feasible disease set. As the parent's siblings cover only Leaf hoppers and Thrips out of three diseases in feasible disease set, the algorithm checks whether the grandparent's siblings cover all the diseases in the feasible disease set. The previously not confirmed level L_1 symptoms cover only Leaf hoppers and Zinc deficiency out of three diseases in feasible disease set. So, the algorithm falls back onto the unique symptoms. It extracts the unique symptoms for each disease that differentiates it from the other diseases in feasible disease set. In the current case, the unique symptoms 171, 12, and 42 of Leaf hoppers, Thrips, and Zinc deficiency respectively are displayed.

4.7 Summary of the chapter

In this chapter, we have presented the development of the Crop Darpan system. Mainly, we have presented the Knowledge Acquisition Protocol and Question Popping model for designing the Crop Darpan system. The proposed frameworks have been developed in an incremental manner. In total five frameworks have been developed. After presenting the first four frameworks, the final framework, which called hierarchy-based protocols for knowledge acquisition and Hierarchy protocol for question popping with Multi-questions per screen (HHM) protocol has been presented. The final framework has been developed based on the feedback from earlier frameworks.

In the next chapter, we present the system design and explain the operationalization of Crop Darpan for Cotton Crop.

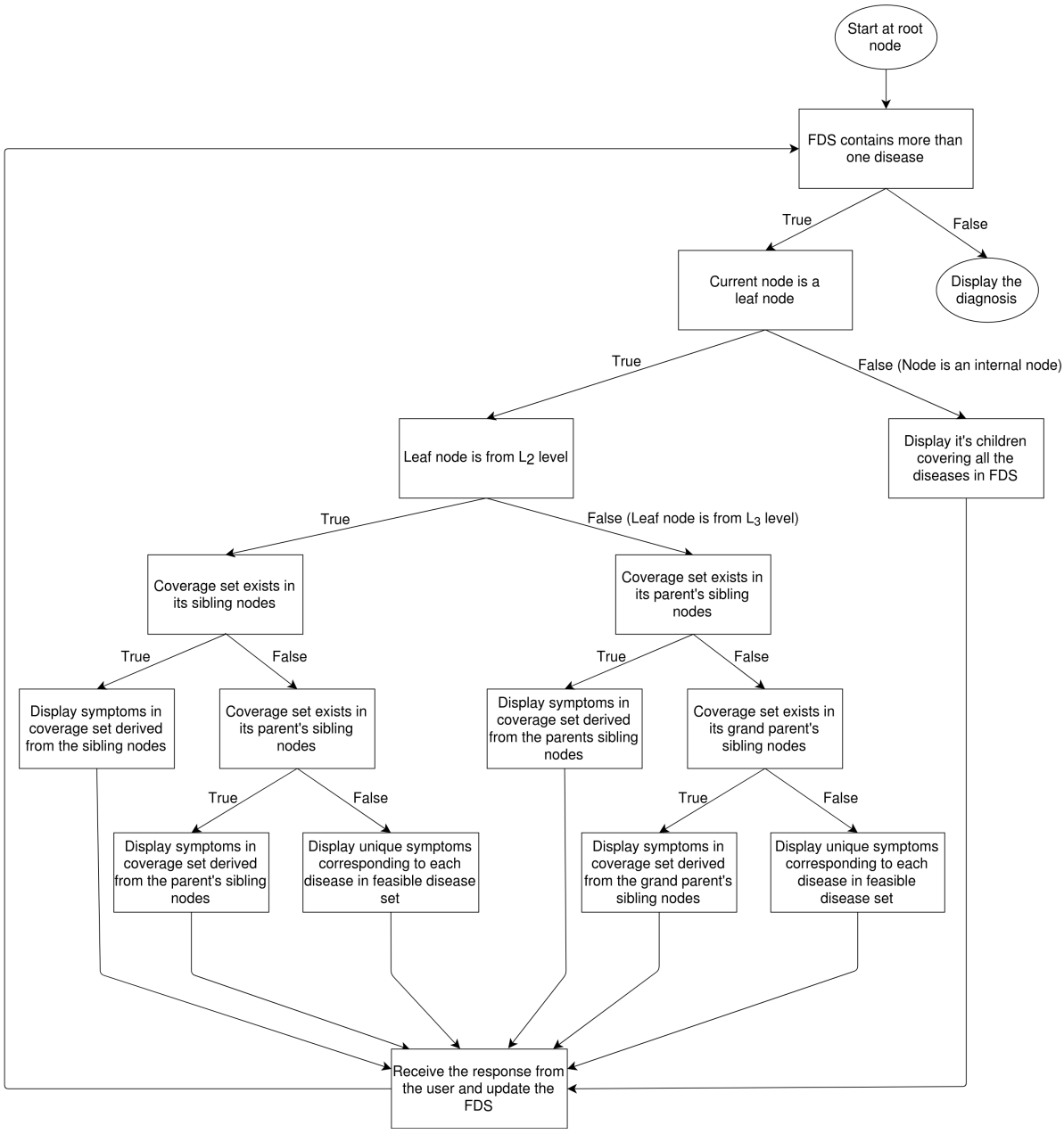


Figure 4.10: Hierarchy model's algorithm flow chart

Table 4.6: Illustrative example problems-symptoms relation table

Disease	Symptoms
Pink bollworm	Bolls damage (2), Premature opening of bolls (21), Fed bolls but holes not visible (24), Presence of small and pink color larvae with brown head inside the boll (25), Flowers damage (4), Unopened and twisted flowers in rose shape (43), Shedding of flowers (42), Lint color change [5], Pink color lint (52)
Leaf hoppers	Leaves damage (1), Curling of leaves (13), Upward curling of leaves (132), Leaves color change (15), Reddening of leaves from margins (153), Spots on leaves (17), Brown necrotic patches on leaves (171)
Thrips	Leaves damage (1), Groups of small size larvae on the lower side of the leaf (14), Curling of leaves (13), Upward curling of leaves (132), Shiny silver color coating on the lower side of the leaves along veins (12)
Anthracnose	Leaves damage (1), Spots on leaves (17), Reddish circular Spots on leaves (172), Bolls damage (2), Premature opening of bolls (21), Spots on bolls (23), Small water soaked-circular-reddish brown depressed spots appear on the bolls (232), Stems damage (3), Splitting of stem and Shredding of bark (32), Lint color change (5), Black color of lint (51), Yellow or brown color of lint (53)
Bacterial blight	Leaves damage (1), Shedding of leaves (16), Stems damage (3), Lint color change (5), Bolls damage (2), Leaves color change (15), Blackening of veins in leaves (151), Splitting of stem and Shredding of bark (32), Oozing of gum over the stems (31), Black color of lint (51), Premature opening of bolls (21), Premature shedding of bolls (22), Spots on leaves (17), Reddish-brown color angular spots on both sides of leaves (174), Spots on bolls (23), Dark black irregular spots (231)
Boron deficiency	Flowers damage (4), Whole plant damage (6), Bolls damage (2), Drying of terminal buds (41), Bushy appearance of the plant (61), Premature shedding of bolls (22), Shedding of flowers (42)
Zinc deficiency	Leaves damage (1), Curling of leaves (13), Upward curling of leaves (132), Flowers damage (4), Shedding of flowers (42), Whole plant damage (6), Stunted growth (62)
Mites	Leaves damage (1), Bolls damage (2), Brittle leaves (11), Premature shedding of bolls (22), Premature opening of bolls (21), Groups of small size larvae on lower side of the leaf (14), Curling of leaves (13), Leaves color change (15), Spots on leaves (17), Shedding of leaves (16), Downward curling of leaves (131), Reddening of leaf from middle portion and drying (152), Yellowish spots on leaves (173)

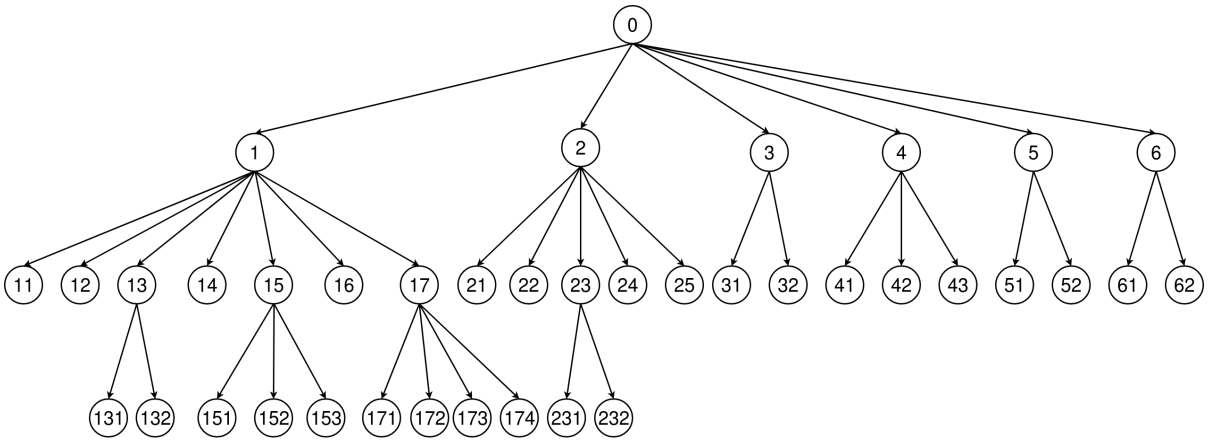


Figure 4.11: Illustrative example symptom-symptom relations (hierarchy)

Chapter *5*

Details of the Crop Darpan prototype for Cotton crop

We have built a generic Crop Darpan system and operationalized the system for the Cotton crop. The developed prototype has the flexibility to extend to other crops as well. In this chapter, we first discuss the Crop Darpan system design and then present the operationalized of the prototype for the Cotton crop.

5.1 Crop Darpan system design

In this section, we first present the functionalities and operating environment of the Crop Darpan system and then present the system design using the Operating environment, System context design, Use case diagrams, ER diagrams and Data flow diagrams.

5.1.1 Functionalities

The principal functionalities of the system prototype are as below:

- The system prototype supports the operationalization of multiple crops.
- The system prototype should support the usage in multiple languages.
- The farmer should be provided with an application user interface to interact with and diagnose the disease.
- The domain expert should be provided with an application user interface to construct the knowledge base for multiple crops in multiple languages.
- The system should have an interface for the admin to manage the users and data.
- The prototype should support both a computer-based interface and a mobile-based interface.

5.1.2 Operating environment

The software and the hardware specifications of Crop Darpan system are as follows:

Software specifications

- Operating system: CentOS
- Database: My SQL
- Framework: CodeIgniter, MVC architecture: The MVC architecture design separates the application into three main components namely model, view and controller. The Model component corresponds to all data-related logic. The view component corresponds to the UI design logic. The Controller component acts as an interface between Model and View to process all business logic and incoming requests.
- Language: PHP
- Web server: Apache

Hardware specifications

- Primary memory: 16GB
- Hard-disk memory: 8TB
- Number of Processors: 16
- Processor: Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz

5.1.3 System context diagram of Crop Darpan

The system context diagram provides a high-level view of the system and its boundaries. It explains the interactions of the system and external entities through the expected inputs and outputs. This diagram contains entities and relationships. The entities are the system and the external actors who interact with the system. The entities are depicted as labeled boxes with no details of its interior structure. In the diagram, the system is drawn at the center surrounded by the actors. The relationships between the actors and the system are represented by the labeled directed lines. The direction of the labeled line represents the flow of the information.

Figure 5.1 shows the system context diagram for Crop Darpan. The external entities involved in the Crop Darpan are farmer, domain expert, and admin. The external entities are shown in rectangular boxes, whereas the Crop Darpan system is represented using a circle. Farmer initializes the interaction with the system by selecting the crop & language and then responds to the questions displayed by the system. The system provides the questions (symptoms) and the final diagnosis

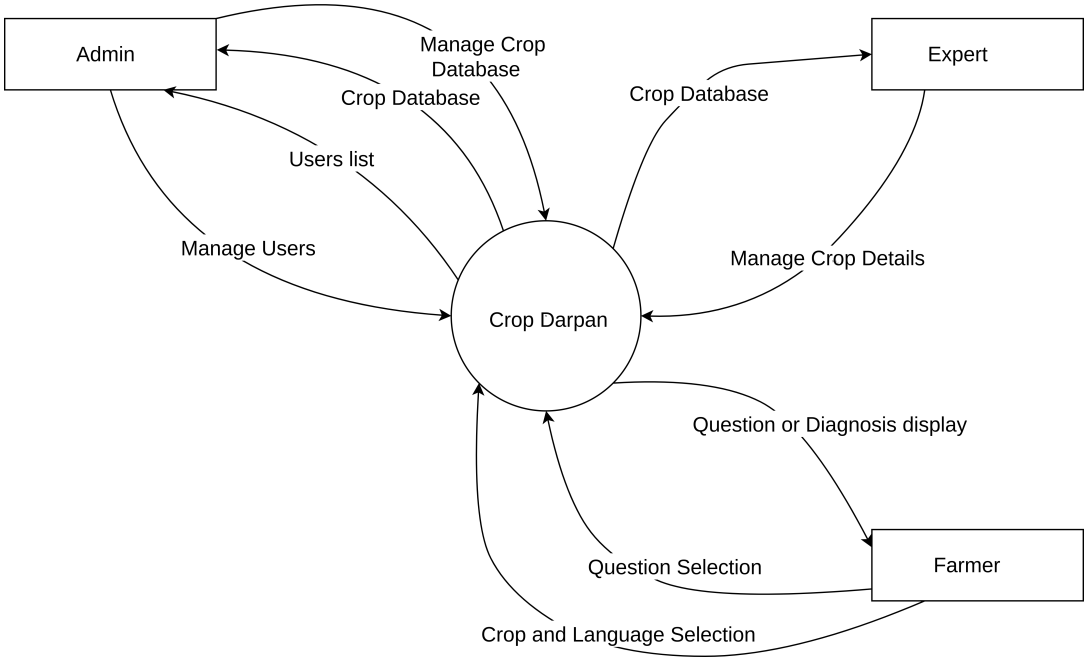


Figure 5.1: System context diagram

information to the farmer. The expert manages the data of all crops by adding, deleting and updating the crop information provided by the system. The admin manages both the users (farmers and experts) and the crop database.

5.1.4 Use case diagrams of Crop Darpan

Use case diagrams help us in visualizing the functionalities of the system. A use case diagram contains three components: actors, use case, relationship. The actor is an entity that is outside of the system boundary and interacts with the system. A use case is the functionality of the system. A system can have multiple use cases. Notions used to represent the components of use case diagrams are shown in Table 5.1. Actors are represented by a stick figure, use cases are represented by an oval shape and the relationship between the actor and use case is represented by a straight line.

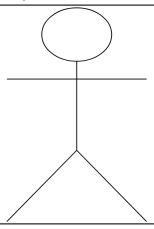
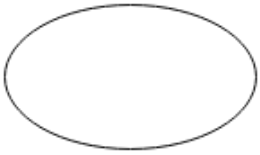
In the Crop Darpan system, there three actors who interact with the system: Farmer, Expert, and Admin. We present the use case diagrams corresponding to each actor below.

5.1.4.1 Farmer use case diagram

Figure 5.2 shows the use case diagram for the farmer. The use case is explained below.

- **Register:** This use case represents the registration functionality of the farmer.

Table 5.1: Notations used in use case diagrams

Name of the Element	Symbol
Actor	
Use Case	
Relation between actor and use case	

- **Login:** Farmers login to their accounts using email-id/user name and password provided during registration.
- **Edit Password:** Farmers can reset the password.
- **Update Profile:** Farmers can edit the basic information provided by them during the registration.
- **Select Crop:** Farmer can select the crop for which he/she wants to find the diagnosis.
- **Select Language:** Farmer can select the language in which he/she wants to use the system.
- **Select Question:** Farmer acknowledges the presence of a symptom by selecting the corresponding question.
- **View Diagnosis:** Farmer views the diagnosis identified by the system based on the symptoms he/she confirmed.

5.1.4.2 Expert use case diagram

Figure 5.3 shows the use case diagram for the expert. The use case is explained below.

- **Login:** Experts login to their accounts using email-id/user name and password.
- **Edit Password:** Experts can reset the password.
- **Update Profile:** Experts can edit his/her basic information.

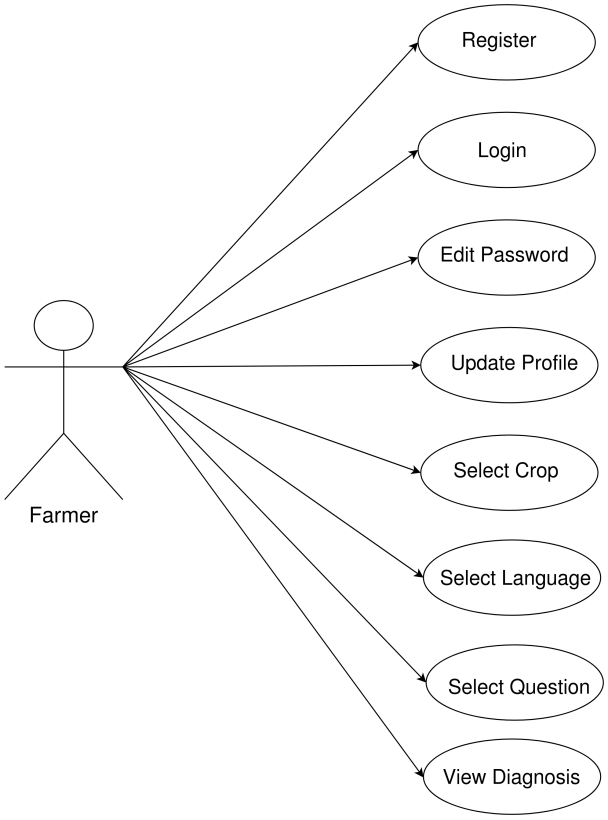


Figure 5.2: Use Case Diagram for Farmer

- **Add Crop:** Experts can add information about new crop which isn't present in the database.
- **View Crop:** Experts can view the information of crops existing in the database.
- **Edit Crop:**Experts can edit the information related to the crops.
- **Add Disease:**Experts can add information about the diseases corresponding to a crop.
- **View Disease:**Experts can view the information of the diseases corresponding to a crop.
- **Edit Disease:**Experts can edit the information of the diseases corresponding to a crop.
- **Add Hierarchy:**Experts can add the symptoms hierarchy for a crop.
- **View Hierarchy:**Experts can view the symptoms hierarchy for a crop.
- **Delete Hierarchy:** Experts can delete the symptoms hierarchy for a crop.

5.1.4.3 Admin use case diagram

Figure 5.4 shows the use case diagram for the admin. The use case is explained below.

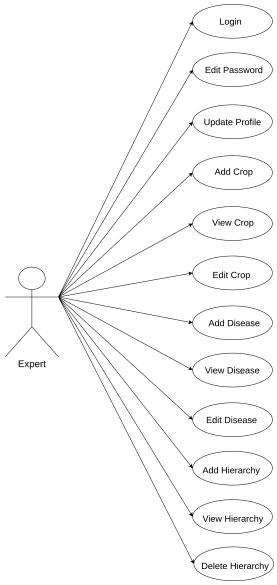


Figure 5.3: Use Case Diagram for Expert

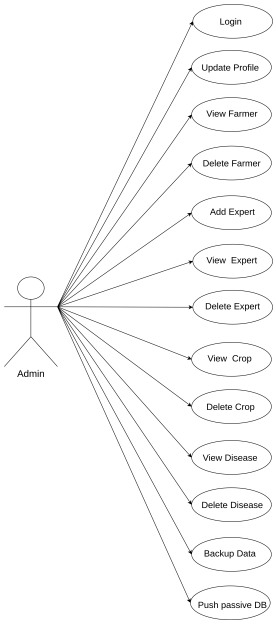


Figure 5.4: Use Case Diagram for Admin

- **Login:**Admin login to their account using email-id/user name and password.
- **Update Profile:**Admin can edit the basic information related to him/her.
- **View Farmer:**Admin can view the information related to the farmers.
- **Delete Farmer:**Admin can delete the information related to the farmers.
- **Add Expert:**Admin can add a new expert to the system.
- **View Expert:**Admin can view the information related to the expert.
- **Delete Expert:**Admin can delete an expert from the existing expert's list.
- **View Crop:**Admin can view the information related to the crops.
- **Delete Crop:**Admin can delete a crop from the database.
- **View Disease:** Admin can view the information of disease.
- **Delete Disease:**Admin can delete the information related to the disease.
- **Backup Data:**Admin can backup the data for safety and security reasons.
- **Push Passive DB:**Admin can push the passive DB to active DB. Here passive DB is the database that can be modified by the expert and active DB is the database that is accessed live by the farmer. The admin can push the Passive DB to active DB if he/she is confident

that modifications done by the expert can be reflected in the live system accessed by the farmer.

5.1.5 ER diagram of Crop Darpan system

An Entity-Relationship Diagram (ER diagram) is a graphical representation of the information related to entities in the system, their attributes, and the relationships between those entities. The major components of the ER diagram are as follows:

- **Attributes:** Attributes are the components of the table in the relational database. A group of attributes combinedly can form a table. In the ER diagram, each attribute is represented by an ellipse with the attribute name inside it. The name of the attribute is underlined if it is a key-attribute for the table.
- **Entities:** Entities in the ER diagram takes form as tables in the relational database. The Entity must have at least one attribute along with the Unique key. Each Entity is represented with the rectangle in the ER diagram.
- **Relationships:** Relationship represents the association among entities. In the ER diagram, the relationship is denoted by a diamond shape. The participation of an entity in a relationship is represented by a line. The compulsory participation of an entity in a relationship is represented by a thick straight line and similarly, non-compulsory participation is represented by a dashed straight line. Two entities can participate in one to one, one to many, many to one, or many to many relationships. In the ER diagram, these 4 types of relationships are represented by specifying either *1* or *m* on the straight lines connecting the relationship and the entities involved in the relationship.

Figure 5.5 shows the ER diagram for Crop Darpan system. Here, the entities are Expert, Farmer, Admin, Login, Crop, Disease, and Symptoms (of levels L_1, L_2, L_3). From the figure it can be understood that the three entities Farmer, Expert, and Admin participate in a *has* relationship with the Login entity. The *has* relationship between the users (Farmer, Expert, Admin) and Login comes under the category of one to one relationship with Login being a non-compulsory participant and users being a compulsory participant. This signifies that every row in a user (Farmer, Expert, Admin) table must have the corresponding row in Login table but a row in Login table may or may not a corresponding row in a user (Farmer, Expert, Admin) table.

Expert participates in *manages* relationship with Crop. Here the relationship is one-to-many and non-compulsory participation from Expert end. It signifies that in the database, for each crop record the details of the expert who updated that crop record recently is stored. The expert also *manages* the information related to Diseases and corresponding L_1, L_2, L_3 Symptoms.

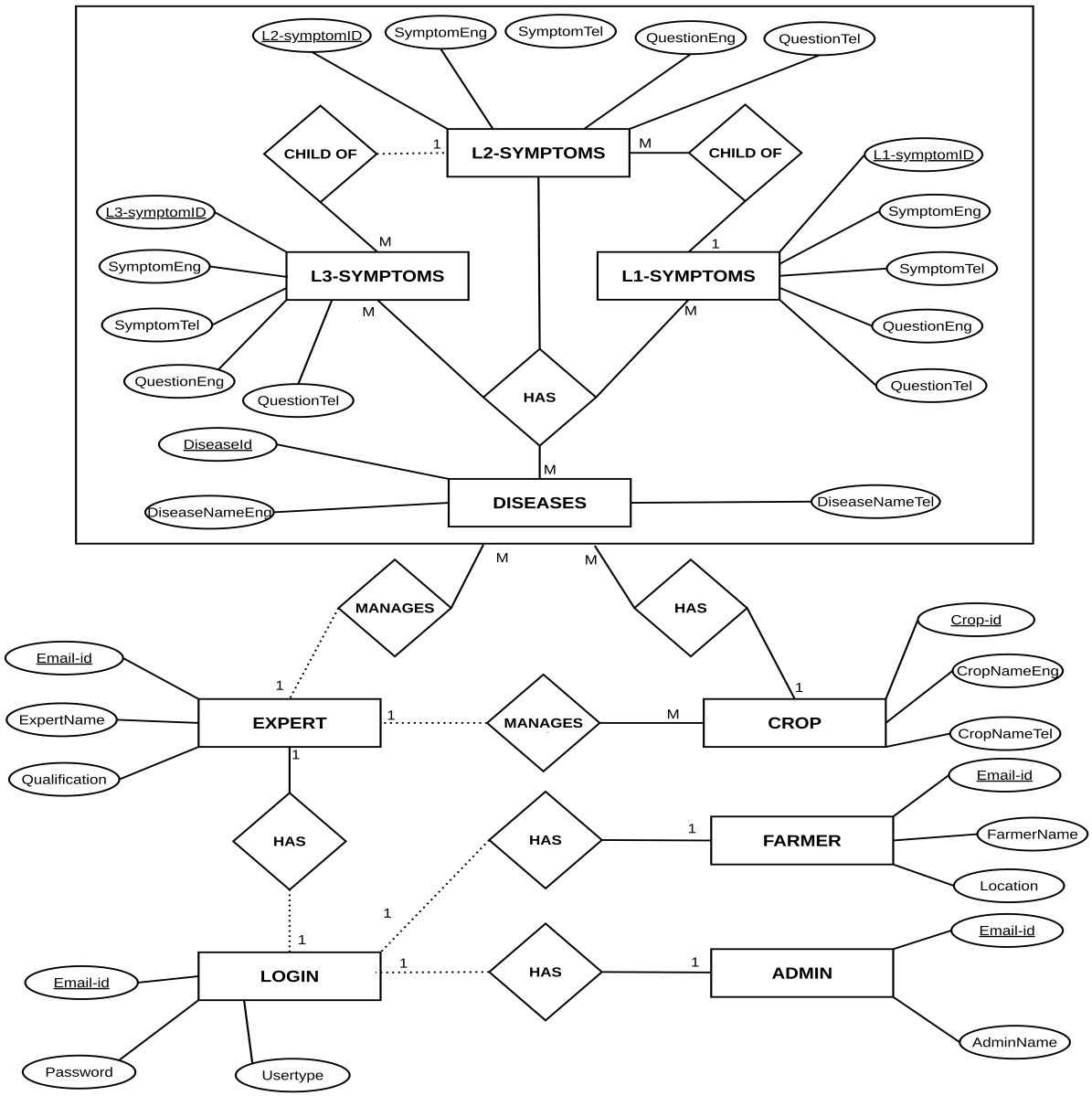


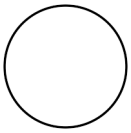
Figure 5.5: ER-diagram of Crop Darpan system

Crop participates in *has* relationship with Diseases and corresponding L_1, L_2, L_3 Symptoms. This signifies that every crop has multiple Diseases but a Disease can relate to the only Crop. The Disease and L_1, L_2, L_3 Symptoms participate in many to many relationships. It signifies that a Disease can have multiple symptoms and vice versa. The hierarchy among the L_1, L_2, and L_3 Symptoms is captured by many to one *child of* relationship from child level to the parent level. This signifies that parent symptoms can have multiple child symptoms but a child symptom can have only one parent symptom.

5.1.6 Data flow diagrams of Crop Darpan

Data Flow Diagrams (DFDs) graphically represents the flow of data from one component to another in the system. DFD includes a mechanism to model the data flow. It supports decomposition to illustrate details of the data flows and functions. DFD cannot present information on the operation sequence. As it is mainly used to express the data transformation in a system, it is not a process or procedure modeling method. The notations used in data flow diagrams are shown in Table 5.2. An entity is represented by a rectangle, process (function) is represented by a circle, the data flow is represented by an arrow and datastore is represented by an open rectangle. We present the data flow diagrams for the three actors in the Crop Darpan system below.

Table 5.2: Notations used in data flow diagrams

Name of the Element	Symbol
Entity	
Process	
Data Flow (Link to show the flow of information)	
Data Store	

5.1.6.1 Farmer data flow diagram

Figure 5.6 shows the data flow diagram for farmer. The farmer has to go through login, which validates the ID and password with the data in the login table. The system rejects if the validation fails. After login, the farmer can update their profiles in Farmer_details data store. Farmer is facilitated with interaction of the system even without login. While the crop data displayed in the application interface is retrieved from Crop table. After the selection of crop and language, a series of questions are displayed to the farmer, which are extracted from CropID_Symptoms table (Here CropID implies the primary key ID of the selected crop). After confirming multiple symptoms, the diagnosis result is shown to the farmer, which contains details about the disease of the crop, retrieved from the Diseases table. The data pulled by the farmer from the system is from the active database.

5.1.6.2 Expert data flow diagram

Figure 5.7 shows the data flow diagram for expert. First, the expert has to go through login, which validates the ID and password with the data in the login table. After login, the expert can update their profiles in the Expert_details data store. After login, experts manage the data of each and every crop. Managing the data implies adding, deleting and updating the crop information. Adding or updating of L_1symptoms is the sub-process in the process of adding or updating crop. Crop and CropID_L1symptoms tables are changed during this process. Similarly, adding or updating of L_2symptoms and L_3symptoms is the sub-process in the process of adding or updating disease. Tables accessed during this process are Diseases, CropID_symptoms, CropID_L2symptoms, CropID_symptomrelations. Finally, the expert can add or edit or delete hierarchy by accessing CropID_L3symptoms table. All these changes made by the expert are present in the passive database.

5.1.6.3 Admin data flow diagram

Figure 5.8 shows the data flow diagram for admin. First, the admin has to go through login, which validates the ID and password with the data in the login table. Managing the experts and the databases (both active and passive) is the responsibility of the admin. Admin can add or delete experts. Login and Expert_profile tables are accessed during this process. Admin can delete the crop and the disease present in the database. Diseases, CropID_symptoms, CropID_L2symptoms, CropID_L3symptoms, CropID_symptomrelations are changed during deleting of both crop and disease. In addition to the above-mentioned tables, Crop and CropID_L1symptoms are also changed during the deletion of the crop. Admin can backup the existing database (Active database). Admin can also push the passive database to the active database.

5.2 Operationalization of the prototype for Cotton crop

We have built a generic Crop Darpan system and operationalized the system for the Cotton crop. The developed prototype has the flexibility to extend to other crops as well. In this section, we present the details of the prototype for the Cotton crop.

As discussed in Chapter 4, the Crop Darpan prototype has been evolved with iterative feedback tuning. As part of this process, we have developed two knowledge acquisition protocols and five-question popping models. In the initial phases, we have collected Cotton disease data using the set-based knowledge acquisition protocol with the help of agricultural experts and textbooks. Using the data collected and single question popping models (entropy-based, frequency-based, and hybrid), we have operationalized the single question framework for Cotton

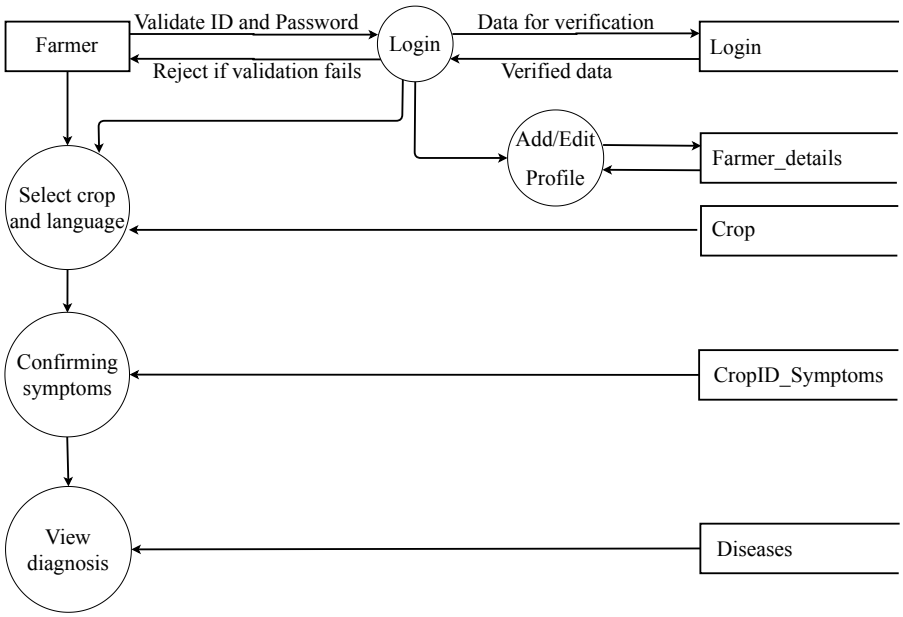


Figure 5.6: Data Flow Diagram for Farmer

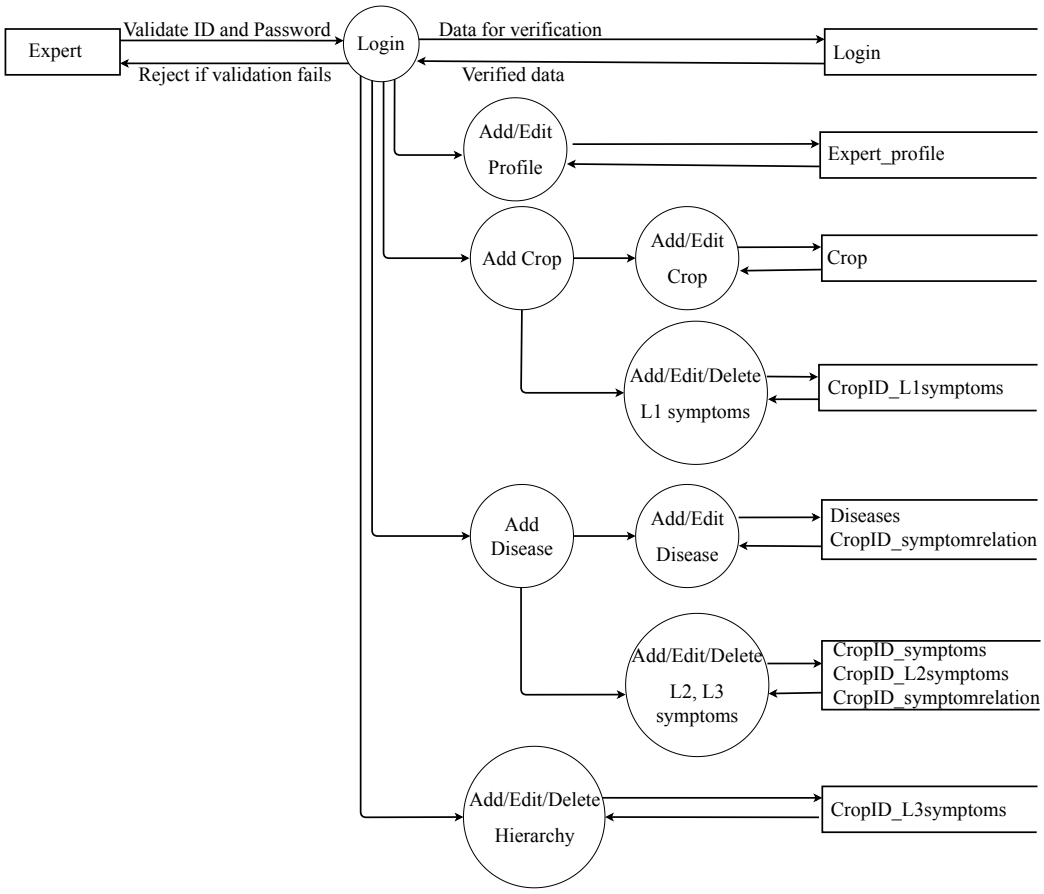


Figure 5.7: Data Flow Diagram for Expert

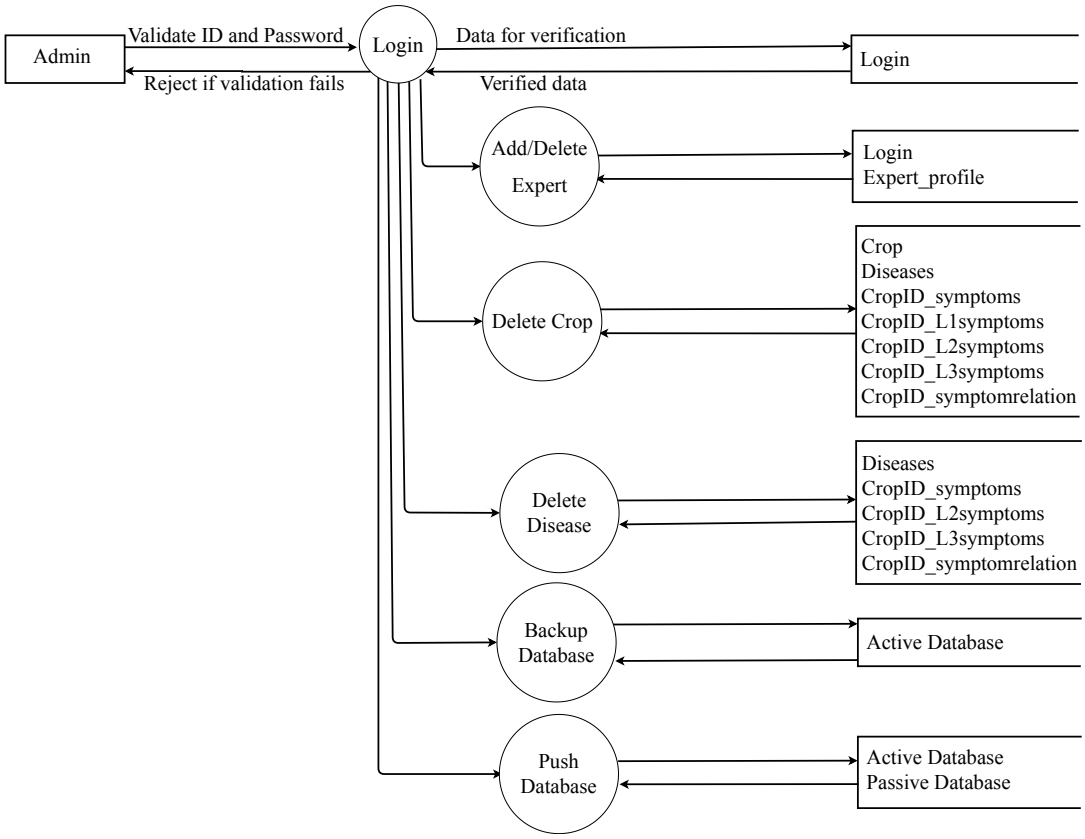


Figure 5.8: Data Flow Diagram for Admin

crop both Telugu and English. The developed single question framework is available at URL `http://eagromet.in/fpds`.

We have developed the multi-question based prototype for Cotton Crop using the hierarchy based knowledge acquisition protocol and hierarchy based question popping model. The collected cotton data contains 30 diseases and their corresponding symptoms. The collected symptoms data form a hierarchy of 4 levels as discussed in Chapter 4. Based on the visual symptoms collected, we have developed corresponding questions in both English and Telugu languages. The collected data is presented in Appendix B. The developed system supports both computer-based interface and mobile-based interface. The developed cotton prototype is available at URL `https://www.cropdarpan.in`. The developed web portal interfaces of all the users: farmer, expert, and admin are presented below.

- Farmer Interface

  * The mobile-based farmer interface for cotton crop in English and Telugu languages are shown in Figure 5.9 and Figure 5.10 respectively.

* The computer-based farmer interface for cotton crop in English is shown in Figure 5.11.
* Now, we present the sample diagnosis procedure for Red cotton bug disease which occurs in Cotton crop using the following figures. From the farmer interface shown in Figure 5.11, let's say farmer selected English language, then the screen is shown in Figure 5.12 appears. The Figure 5.12 shows all the part level (L_1) symptoms/questions. Red cotton bug disease damages the bolls so the corresponding question is selected with a click on *Yes*. Now, the screen shown in Figure 5.13. The Figure 5.13 shows L_2 level symptoms. Red cotton bug damages the bolls by making spots on them so the corresponding question is selected by clicking *Yes*. The next screen is as shown in Figure 5.14, which contains specialized symptoms (from L_3 level) of spots on bolls symptom. The red cotton bug makes the white or yellow color spots on the bolls so the corresponding question is selected. The final screen shown in Figure 5.15 shows the diagnosis found, advice to follow, symptoms confirmed by the farmer, other symptoms of the disease.
* As discussed earlier, the multi-question framework gives the flexibility to farmers to choose the symptom that he/she is most confident about. So, using this multi-question framework a disease can be reached by multiple paths. Figure 5.16 shows the multiple paths for diagnosing the Red cotton bug.

– Expert Interface: Figure 5.17 shows the expert interface. From the figure, we can observe the functionalities of an expert such as managing crop, disease, and hierarchy information.

– Admin Interface: Figure 5.18 shows the Admin interface. From the figure, we can observe the functionalities of admin such as managing crop and user information.

5.3 Summary of the Chapter

In this chapter, we have presented the design of the Crop Darpan system which can be used to develop the Crop Darpan system for any crop. We have also explained the operationalization of Crop Darpan for the Cotton crop.

In the next chapter, we present an evaluation study.

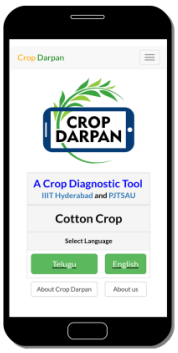


Figure 5.9: Mobile based farmer interface in English

Figure 5.10: Mobile based farmer interface in English

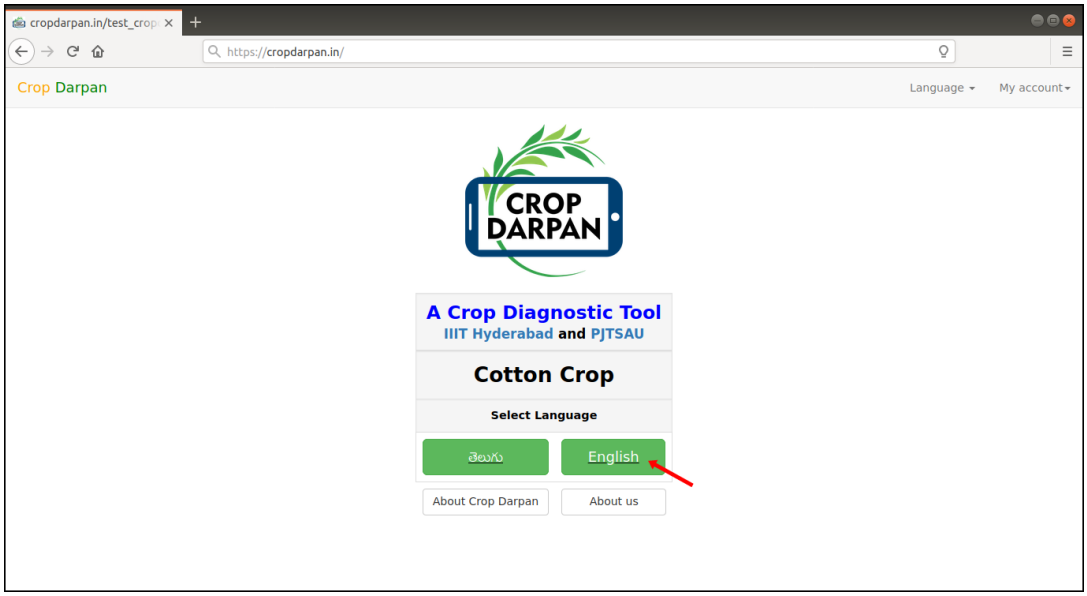


Figure 5.11: Farmer Interface

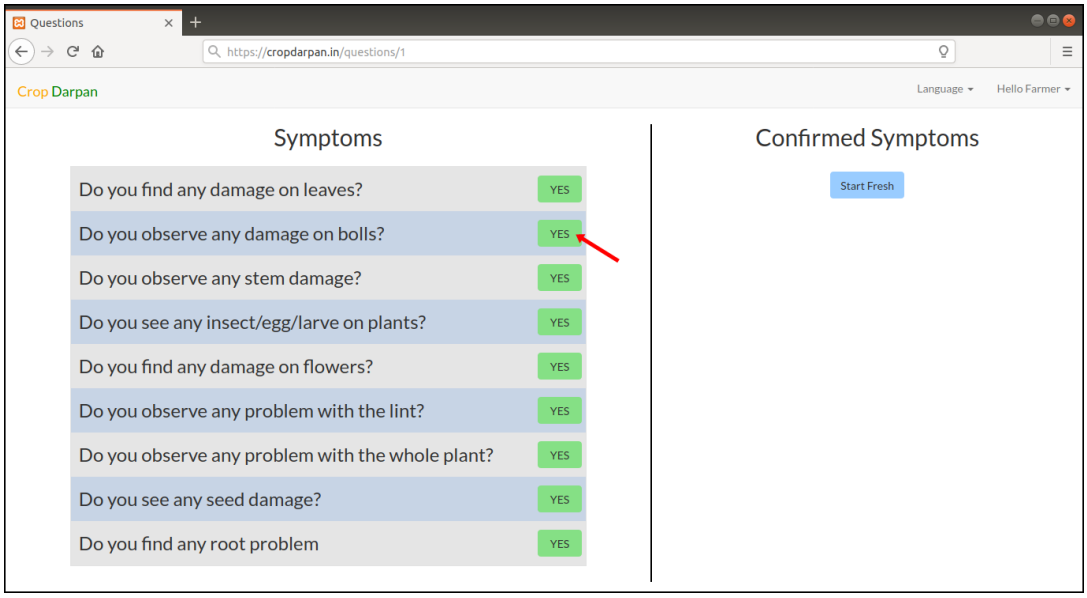


Figure 5.12: Diagnosis for Red cotton bug - Screen 1

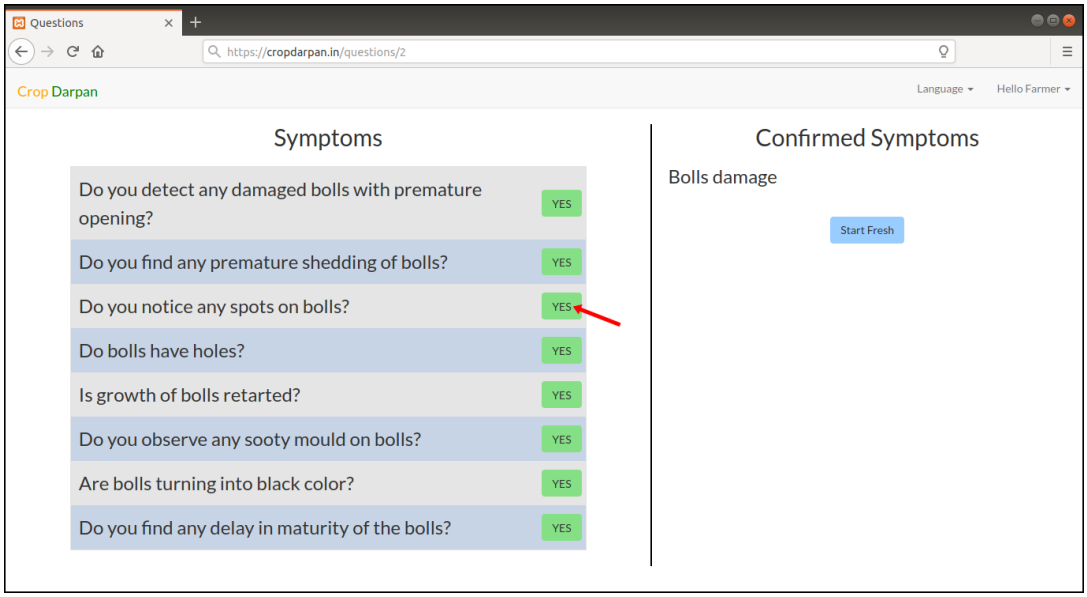


Figure 5.13: Diagnosis for Red cotton bug - Screen 2

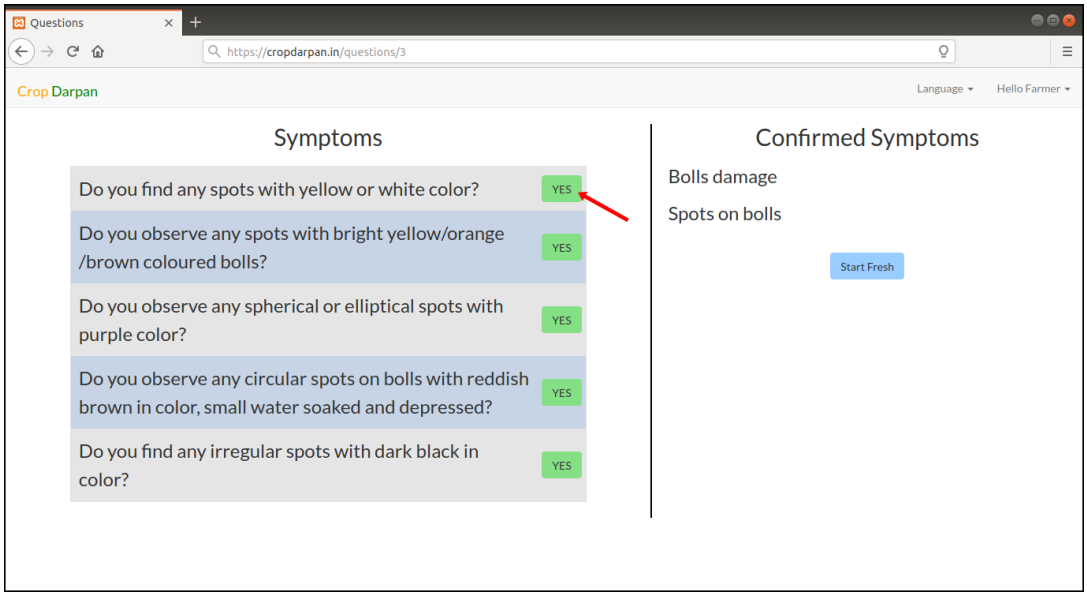


Figure 5.14: Diagnosis for Red cotton bug - Screen 3

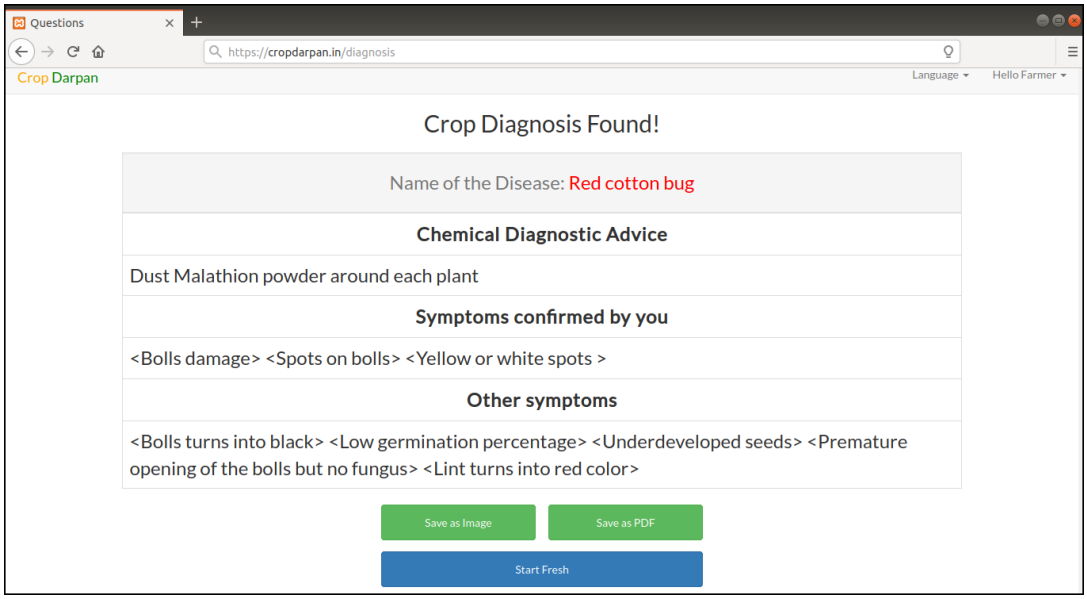


Figure 5.15: Diagnosis for Red cotton bug - Screen 4

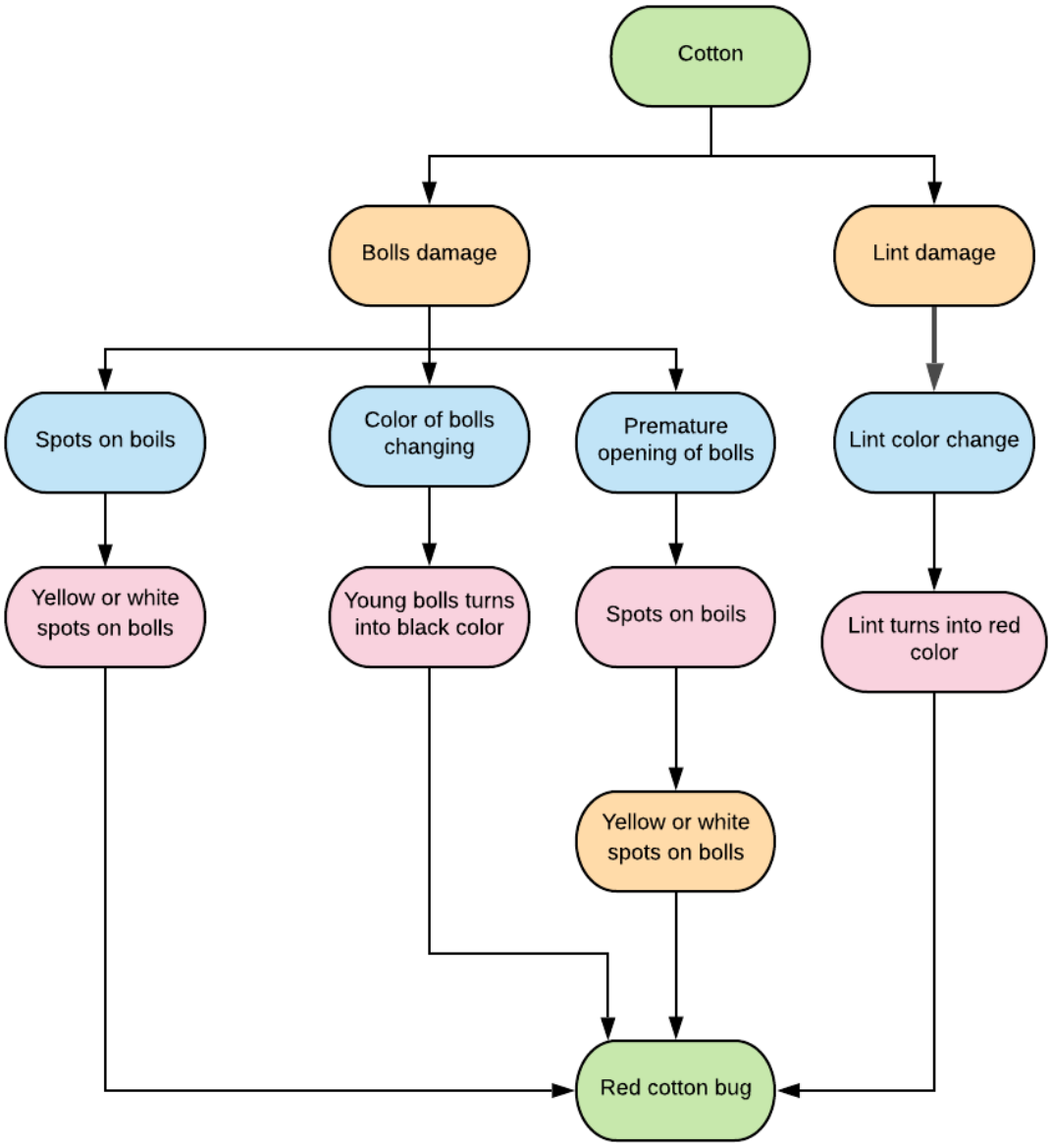


Figure 5.16: Diagnosis in Red cotton bug using multiple paths

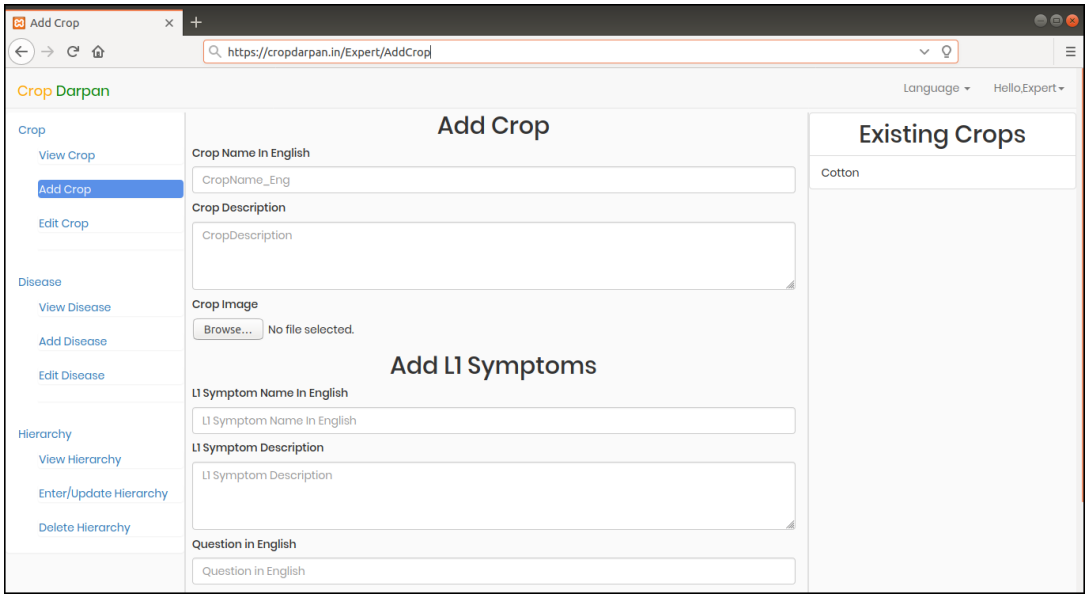


Figure 5.17: Expert Interface

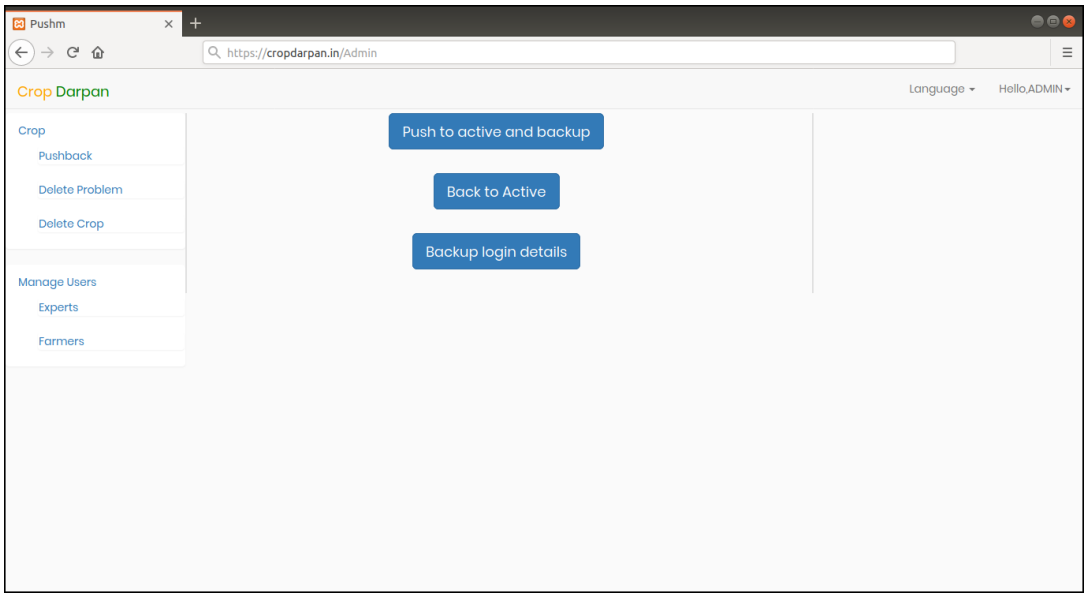


Figure 5.18: Admin Interface

Chapter *6*

Evaluation of Crop Darpan prototype for Cotton crop

6.1 Context of the evaluation

The goal of this research is to build a diagnosis system, called Crop Darpan, to help the farmer to identify the disease in a given crop using a smartphone. In this connection, we have explained the context of the problem, concepts of the Crop Darpan, details of system design and operationalization details for Cotton Crop. We have built the system for Cotton crop and operationalized during 2019-20. The system is being used by a sample of farmers for diagnosing the crop diseases (pests, bacterial and fungal diseases, nutritional deficiencies) in the Cotton Crop in both English and Telugu languages.

- Pests: Spotted boll worm, Helicoverpa, Gram caterpillar, Pink boll worm, Tobacco caterpillar, Leaf hoppers, White flies, Aphids, Mites, Thrips, Red cotton bug, Dusky cotton bug, Leaf roller, Stem weevil, Mealy bugs.
- Bacterial and Fungal diseases: Rust, Helmenthosporium spots, Cercospora leaf spot, Alternaria leaf spot, Anthracnose, Grey mildew, Root rot, Verticellium wilt, Fusarium wilt, Bacterial blight.
- Nutritional deficiencies: Nitrogen deficiency, Phosphorus deficiency, Potassium deficiency, Magnesium deficiency, Boron deficiency, Zinc deficiency.

The Crop Darpan system supposes that farmer should possess the smartphone and visit the Cotton crop and accesses `www.cropdarpan.in` portal. The system poses questions corresponding to high-level symptoms. The farmer checks the symptoms of the disease visually in the field and confirms the existence of the corresponding symptom. Based on the confirmation of visual systems, the system possesses next-level questions. Finally, the system displays the diagnosis of the crop and advice to be followed based on the symptoms confirmation done by the farmer.

Overall, the Crop Darpan system has been developed with the following suppositions.

- Suppositions related to the domain expert
  - It is possible to organize the symptoms of crop diseases based on the generalization-specialization framework.
  - The proposed generalization-specialization framework is modeling the process of crop disease diagnosis procedure followed by a typical domain expert.
- Suppositions related to the educated farmer
  - The farmer has the ability to understand the question, visit the crop and confirm the existence of the corresponding symptom.
  - The farmer has the ability to identify the crop diagnosis by navigating and confirming the existence of high-level and low-level symptoms presented by the Crop Darpan tool.

The objectives of the evaluation are as follows:

Objective 1: To understand the perspectives of domain experts about the usage of the generalization-specialization framework to conceptualize the knowledge regarding diseases and symptoms.

Objective 2: To understand the perspectives of domain experts about the crop diagnosis procedure followed by Crop Darpan.

Objective 3: Validating the practicality, interpretability, usability, data accuracy, extendibility, necessity aspects of Crop Darpan.

Objective 4: Validating the supposition that the educated farmers has the ability to read and understand and respond to the questions regarding the existence of a symptom.

Objective 5: Validate the supposition that the farmer has the ability to diagnose the crop diseases by navigating and confirming the existence of high-level and low-level symptoms presented by the Crop Darpan tool.

6.2 Evaluation Methodology

6.2.1 Scope and Methodology

The present study has adopted different methods to collect the required data for analysis. The study is confined to the issues relating to the scope of evaluating the Crop Darpan prototype of the Cotton crop. We have employed the primary data resources for this evaluation. The primary data is collected from two sources: Domain experts and Educated farmers.

6.2.1.1 Methodology of data collection from domain experts

The purpose of data collection from domain experts is to realize the above-mentioned Objectives 1, Objective 2, and Objective 3 which reflects the suppositions made regarding domain experts.

We have followed the survey research methodology [36] to collect the primary data. A survey methodology involves asking questions to a pre-defined group of respondents through a questionnaire and gather insights regarding the topics of interest. As the topic of interest for this study is modeling the behavior agricultural scientists in the diagnostic process, we have selected a group of subject matter specialists who are knowledgeable about diseases and symptoms and have experience in diagnosing the crop diseases.

- **Questionnaire preparation:** We have employed semi-structured questionnaire to collect data from domain experts. The questionnaire consists of close-ended multiple-choice questions for quantitative analysis. The questions in the questionnaire prepared such that each question and its responses draw insights about the characteristics of the Crop Darpan that we would like to assess. Specifically, The questionnaire consists of questions concerning (i) generalization-specialization framework, (ii) crop diagnosis procedure followed by both domain expert and Crop Darpan, and (iii) practicality, interpretability, usability, data accuracy, extendability, necessity aspects of Crop Darpan. Along with multiple-choice questions, we have requested the domain experts for their comments corresponding to each question for the qualitative analysis. The questionnaire is presented in 6.1. A sample questionnaire filled by an expert is presented in Appendix —BEGIN—.
- **Data collection:** We have followed the following procedure to collect the data. Firstly, we presented the concepts and operationalization of the Crop Darpan framework to all the experts/participants. The queries raised by the participants are clarified. Secondly, the participants are requested to visit the URL: `www.cropdarpan.in` to access the crop diagnostic tool. Experts are requested to use the tool and identify four diseases using multiple paths. Here, a path is a sequence of questions/symptoms confirmed by the user. Finally, the questionnaires are distributed to the participants and requested to fill them.

6.2.1.2 Methodology of data collection from farmers

The purpose of data collection from educated farmers is to realize the above-mentioned Objective 4 and Objective 5. These objectives reflect the validation of suppositions made regarding the educated farmer that the farmer has the ability to (i) identify the visual symptoms of the crop and respond to the question (ii) diagnose the crop problem using the Crop Darpan tool.

The suppositions made regarding the educated farmers can be validated by verifying whether the farmers can use the system to correctly identify the crop disease. So, we have followed the typical

ground truth comparison methods to evaluate the farmer's crop diagnosis accuracy. So, we have selected two educated farmers for the field level data collection. The farmers are asked to use the Crop Darpan system to diagnosis the cotton crop diseases and capture the diagnostic process in digital form (Photographs and Screenshots). Next, we have identified subject matter specialists who have about 20 years of experience in diagnosing crop problems and shown this digital data. The comparison results were presented.

Data collection methodology and comparison experiments are carried out in the following manner.

- **Collecting field data from farmers:** To collect the farmers' diagnostic activity while using the tool, two educated farmers are selected who has the ability to read and write in the regional language Telugu or English. The farmers are trained and smartphones with internet facilities are provided to carry the data collection activity. Firstly, the farmers visit a sample cotton crop and use the tool by reading the questions displayed on the interface and maps with the visual symptoms exhibited by the crop and responds to the questions. After a series of questions/symptoms confirmed by the farmer, the system presents the final diagnosis on the interface. Secondly, after completion of the diagnostic process using the tool, the farmer takes the screenshot of the interface in which the diagnosis of the crop and the symptoms confirmed by the farmer are shown. Finally, the farmer takes the photographs of the visual symptoms exhibited by plants in crop using camera with multiple viewing angles and closeup shots. The photographs of the crop and corresponding crop diagnosis screenshots are consolidated to create a data record. By following this procedure, multiple data records are created by the educated farmer by visiting multiple cotton crops in the local region.

- **Comparing with the ground truth:** The photographs of the cotton crops taken by the trained farmer capturing the visual symptoms of the diseases are supplied to the experienced agricultural scientists. By analyzing and observing the visual symptoms captured in the crop photographs, the agricultural scientists are able to find the actual diagnosis [53]. Finally, for each data record, the diagnosis found by the agricultural expert is considered as ground truth and compared it with the diagnosis found by the educated farmer using the Crop Darpan tool to find the percentage of the correctly diagnosed crop samples.

6.2.2 Analysis of questionnaire data collected from the domain experts

The educational qualification details of the domain experts who participated in the questionnaire survey are tabulated in Table 6.1.

The analysis of responses of questions posed in the questionnaire and their corresponding objectives is as follows.

Table 6.1: Domain experts data collection details

Qualification	Number of participants
MBA Agri. Bussiness	1
Ph.D Agronomy	2
M.Sc Agronomy	5
M.Sc Plant Pathology	7
Total	15

– **Question 1:** We requested the participants to use the Crop Darpan tool and diagnose four diseases using multiple paths. This task ensures that participates are familiarized with the tool. We posed the question 1 in the questionnaire to quantitatively understand well participants are familiarized with the system. The response details of question 1 are presented in Table 6.2. From the table, we can understand 9 out of 15 participants diagnosed four diseases using their knowledge about diseases and symptoms. Whereas, 4 out of 15 participants diagnosed three diseases. So, overall it can be noted that all the participants diagnosed at least two diseases using multiple paths.

Table 6.2: Responses of Question 1 - Out of four crop diseases, how many of the crop diseases are correctly identified?

Option	1	2	3	4
Responses in number	0	2	4	9
Responses in percentage	0%	13.3%	26.7%	60%

– **Question 2 (Objective 3 - Practicality):** We posed question 2 in the questionnaire to understand the opinion of domain experts regarding the practicality/viability aspect of the Crop Darpan system. The response details for question 2 are tabulated in Table 6.3. From the table, we can notice that out of 15 participants, 7 participants opined that after providing the training, the system will definitely help the educated farmers to diagnose the crop disease and 8 participants opined that the system will help the educated farmer to some extent. None of the participants countered the viability of the system. Participants also provided comments that the development of the Crop Darpan system in regional languages along with English improves the reach of the system among the farmers. The participants emphasised on the need for training programs for the farmers. The participants also opined that the uneducated farmers will not be able to use the system. The current version of the Crop Darpan system provides an only web interface for the educated farmers to use. By extending this framework to a regional language speech-based interface it can be made available to the uneducated farmers.

– **Question 3 (Objective 1, Objective 3 - Interpretability):** We posed question 3 to understand the opinion of domain experts regarding the adaption of generalization/specialization framework to conceptualize the disease and symptoms relationships. The details of the par-

Table 6.3: Responses of Question 2 - After providing training, do you agree that proposed system will help the educated farmer to identify the crop problem using the smartphone through Internet?

Option	It will not help the educated farmer.	It will help the educated farmer to some extent.	It will definitely help the educated farmer to identify the crop problem.
Responses in number	0	8	7
Responses in percentage	0%	53.3%	46.7%
Comments	1. Training programs should be conducted to train uneducated farmers also as it is also available in Telugu. 2. It helps the educated farmers but uneducated farmers are not ready (to use the system). 3. Most of the farmers, about 60% are illiterate (uneducated).		

ticipant's responses are tabulated in Table 6.4. From the information presented in the table, we can observe that 13 out of 15 participants responded that the adaptation of generalization/specialization framework is logical and also interpretable. Also, two participants expressed that they are unable to judge as they doubt the ability of farmers to differentiate between the symptoms. They have suggested that the supplement of the images along with a text description of the symptom will be helpful in the diagnostic process.

Table 6.4: Responses of Question 3 - The proposed system extends the generalization/specialization framework (starting with generic symptoms to specialized symptoms) to build the crop diagnostic system to help the farmer. What is your view on the methodology?

Option	The methodology does not make sense.	The methodology seems logical and it is as per common sense.	Unable to judge
Responses in number	0	13	2
Responses in percentage	0%	86.7%	13.3%
Comments	1. The proposed system providing the specialized symptoms will definitely help the farmer to diagnose a disease. 2. I doubt the ability of the farmers to differentiate between the symptoms. So, maybe an image should also be included as an option. 3. Very good development. It should be developed in many crops as soon as possible.		

– **Question 4 (Objective 3 - Usability):** The question 4 in the questionnaire is asked to find the view of domain experts regarding the ease of use aspect of the Crop Darpan system. The responses of the participants are tabulated in Table 6.5. From the table, we can see that 53% of the participants are opinioned that the Crop Darpan system is easily usable, 46% of the

www.cropdarpan.in

Expert Evaluation Form Questionnaire

1. Pls. try to detect four diseases using the Crop Darpan system using multiple paths. Out of four crop diseases, how many of the crop diseases are correctly identified? (a) 1 (b) 2 (c) 3 (d) 4	If you are unable to identify any disease, provide the reason:
2. After providing training, do you agree that proposed system will help the educated farmer to identify the crop problem using the Smartphone through Internet? (a) It will not help the educated farmer. (b) It will help the educated farmer to some extent. (c) It will definitely help the educated farmer to identify the crop problem.	Provide your comment, if any:
3. The proposed system extends generalization / specialization framework (starting with generic symptoms to specialized symptoms) to build the crop diagnostic system to help the farmer. What is your view on the methodology? (a) The methodology does not make sense. (b) The methodology seems logical and it is as per the common sense. (c) Unable to judge.	Provide your comment, if any:
4. You have used the system to identify few crop problems. How difficult is it for you to use? (a) Difficult to use. (b) Moderately difficult to use. (c) Easy to use.	Provide your comment, if any:
5. We have employed subject matter specialists to identify the crop disease symptom data. As a subject matter specialist, what is your opinion on the data? (a) The data is inaccurate. (b) The data is accurate. (c) The data is satisfactory and could be refined.	Provide your comment, if any:
6. The system is intended to resemble the real world diagnostic procedure followed by a typical agricultural scientist to identify the crop problem if he/she visits the field. What is your view on methodology? (a) Typical agricultural scientist follows a different methodology to identify the crop problem. (b) To some extent, the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem. (c) In general (overall), the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem.	Provide your comment, if any:
7. Do you think that similar systems can be developed for other crops? (a) No, it is difficult. (b) Yes, definitely	Provide your comment, if any:
8. Do you think this kind of tool is needed in agriculture? (a) No (b) Yes	Provide your comment, if any:

Figure 6.1: Domain expert evaluation Questionnaire

participants opined that the system is moderately difficult to use and none of the participants opined that system is difficult to use. Participants who opined that the system is moderately difficult to use commented that the text description of the symptoms is helpful only when farmers understood it correctly and found the exact damage in the crop and also opined that the time taken by the farmer should be lesser. The ease of use of the system can be improved by designing better user interfaces and supplying the images and videos of the symptoms along with the text description.

Table 6.5: Responses of Question 4 - You have used the system to identify few crop problems. How difficult is it for you?

Option	Difficult to use	Moderately difficult to use	Easy to use
Responses in number	0	7	8
Responses in percentage	0%	46.7%	53.3%
Comments	1. It is simple, but unless and until he is aware of the symptoms and damage exactly. 2. Along with the question, if pictures of symptoms are provided it would help in correct diagnosis by farmers. 3. Because farmers don't want to take this much time.		

- **Question 5 (Objective 3 - Data accuracy):** The question 5 is posed to find the view of domain experts regarding the data accuracy of the Crop Darpan system. The responses of the participants are presented in Table 6.6. From the table, it can be observed that 5 out of 15 participants opined that the provided data is accurate, 10 out of 15 participants opined that data is satisfactory and it could be refined, and none of the participants mentioned that the data is inaccurate. Participants who suggested the refinement of the data emphasized the addition of pictures of the symptoms. The current version of the Crop Darpan system contains the diseases related to the pests, bacterial, fungal, and nutritional deficiencies. So, participants suggested the inclusion of damages caused by the external environment such as heavy rains and water stress into the system to bring the completeness of the data.

- **Question 6 (Objective 2):** question 6 in the questionnaire inquires about the modeling of domain expert behavior by the Crop Darpan system in the diagnostic process. This question is posed to obtain the domain expert's opinion on the resemblance between the diagnostic methodology of Crop Darpan and the domain experts. The responses of the participants are tabulated in Table 6.7. From the table it can be observed that 6 out of 15 participants opined that the Crop Darpan system captures the diagnosis methodology followed by a typical expert to some extent, 8 out of 15 participants opined that the overall system captures the methodology followed by a typical expert. One participant mentioned that a typical scientist follows different methodology compared to Crop Darpan. The participant opined that the scientists

Table 6.6: Responses of Question 5 - We have employed subject matter specialists to identify the crop disease symptom data. As a subject matter specialist, what is your view on the data?

Option	The data is inaccurate.	The data is accurate.	The data is satisfactory and could be refined.
Responses in number	0	5	10
Responses in percentage	0%	33.3%	66.7%
Comments	1. In addition to the list of diseases and symptoms, if pictures/photos showing the symptoms are included, then it will be much easier to farmer for identification. 2. Because, in some cases, diseases or pest damage symptoms change or vary according to the external environment. 3. The images of disease samples, pest will help to better judging the infestation.		

classify the disease symptoms on the basis of an organism that causes the disease such as fungal or bacterial rather than on the basis of part of the plant infected and fine-grained details of the visual symptoms. However, it can be noticed that the Crop Darpan system captures the methodology followed by the major number of agricultural experts.

Table 6.7: Responses of Question 6 - The system is intended to resemble the real-world diagnostic procedure followed by a typical agriculture scientist to identify the crop problem if he/she visits the field. What is your view on methodology?

Option	Typical agricultural scientist follows a different methodology to identify the crop problem.	To some extent, the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem.	In general (overall), the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem.
Responses in number	1	6	8
Responses in percentage	6.7%	40%	53.3%
Comments	1. If scientist is able to go to the field, he/she will be able to look deep into the problem. 2. To detect disease, scientist use different methods: either it is fungal or bacterial, etc.		

– **Question 7 (Objective 3 - Extendability):** The question 7 is posed in the questionnaire to understand the view of domain experts regarding the extendability aspect of the Crop Darpan system. Currently, the Crop Darpan tool is developed for the cotton crop using generalization/specialization concepts. As the domain experts have the knowledge about multiple crops

and their corresponding diseases, we inquired the domain expert's opinion about the adaption of the Crop Darpan framework to other crops like Rice and Chilli through this question. The response details are presented in Table 6.8. From the table, it can be observed that all the domain experts opinioned that the Crop Darpan framework can be easily extended to other crops.

Table 6.8: Responses of Question 7 - Do you think that similar systems can be developed for other crops?

Option	No, it is difficult.	Yes, definitely.
Responses in number	0	15
Responses in percentage	0%	100%
Comments	1. But, provide the pictures of the symptoms.	

- **Question 8 (Objective 3 - Necessity):** The question 8 in the questionnaire is posed to draw the opinions of domain experts regarding the necessity aspects Crop Darpan like disease diagnostic systems in the agriculture domain. The details of the responses are tabulated in Table 6.9. From the table, it can be observed that all the participants emphasized the requirement of disease diagnostic systems like Crop Darpan.

Table 6.9: Responses of Question 8 - Do you think this kind of tools are needed in the agriculture domain?

Option	No	Yes
Responses in number	0	15
Responses in percentage	0%	100%

6.3 Analysis of field data collected from the farmers

6.3.1 Details of field data collected from the farmers

The details of the field data records collected by the farmers are presented in Table 6.10. From the table, we can observe that two educated farmers, Farmer 1 and Farmer 2 are selected for data collection. Farmer 1 visited more than 15 villages from three mandals from Telangana named Nawab pet, Momen pet, and Vikarabad for data collection. From three mandals, Farmer 1 visited 61 cotton crops and created a data record for each crop sample. Farmer 2 also visited more than 15 villages from three mandals from Telangana named Chilpur, Station ghanpur, Velair for data collection. From three mandals, Farmer 2 visited 68 cotton crops and created corresponding data records. Altogether, we have collected 129 data records for the cotton crop. We have supplied these 129 samples to two agricultural scientists from Professor Jayashankar Telangana State Agricultural University (PJTSAU), Hyderabad for analysis.

Table 6.10: Field data collection details

ID	Mandals	Collected samples
Farmer 1	Nawab pet, Momen pet, Vikarabad	61
Farmer 2	Chilpur, Station ghanpur, Velair	68

Agricultural scientists analyzed the photographs of all 129 data records and provided the diagnosis for each disease which are to be considered as ground truth. From the ground truth diagnosis details, we observed that 14 cotton crop samples are damaged by external factors such as abnormal seed sowing and heavy rains. Out of 14 samples, 4 samples are identified with abnormal seed sowing and 10 samples are identified with damage due to heavy rains. As these factors are not covered by the Crop Darpan system, the farmers could not diagnose the disease correctly. We omitted these 14 records identified with external factors that are not covered by the Crop Darpan and used the remaining 115 records for the simplicity of comparative study. From the data collected, we observed that, for few crop samples, agricultural scientist identified multiple diagnoses for a single record. This means that it is possible for a crop to exhibit visual symptoms corresponding to multiple diseases. Whereas, the current version of Crop Darpan identifies only one diagnosis for an attempt. So, we consider a diagnosis identified by the farmer correct if the disease identified by the farmer matches with one of the multiple diseases exhibited by the crop.

Table 6.11: Ground truth comparison results

Disease Type	Correctly diagnosed	Wrongly diagnosed	Total
Pests	64.34% (74)	21.73% (25)	86.08% (99)
Bacterial and fungal	6.08% (7)	5.21% (6)	11.30% (13)
Nutritional deficiency	2.60% (3)	0% (0)	2.60% (3)
Total	73.04% (84)	26.95% (31)	100% (115)

6.3.2 Realization of Objectives 4 and Objective 5

Objectives 4 and 5 of this evaluation study validate the suppositions made regarding the educated farmer. The suppositions say that an educated farmer has the ability to (i) identify the visual symptoms of the crop and respond to the question and (ii) diagnose the crop problem using the Crop Darpan system. So, these objectives 4 and 5 can be realized by conducting a ground comparison study of diagnoses found by the farmers. The diagnosis accuracy of the farmers reflects the validity of the suppositions.

We have presented the comparison results of the farmer's diagnosis with ground truth diagnosis identified by the agricultural scientist in Table 6.11. From Table 6.11, it can be noticed that out of 115 samples, farmers correctly diagnosed 84 samples and wrongly diagnosed 31 samples. So, the farmer's diagnosis accuracy is 73.04% and the error percentage is 26.95%. As the diagnosis

accuracy of the farmers is at an acceptable percentage, it supports the validity of the suppositions made regarding the educated farmer.

From the table, we can also observe that pest diseases are the most common diseases and have a share of 86.08% in all the samples. Bacterial and fungal diseases have the 11.30% share of all samples. Nutritional deficiencies are the rarest type of diseases and have only a 2.6% share of all samples. This phenomenon can be reasoned by the fact that the farmers are accustomed to the regular usage of fertilizers to grow crops.

After analyzing the data sample records which are wrongly diagnosed by the farmer, we draw the following conclusions. Bacterial and fungal diseases are the most misdiagnosed type of diseases. From the table, we can notice that 6 out of 13, about 46% of the bacterial and fungal diseases are misdiagnosed. This phenomenon can be explained by the fact that most of the bacterial and fungal diseases exhibit visually very similar symptoms. For example, leaf spots diseases such as Helmenthosporium spot, Cercospora leaf spot, and Alternaria leaf spot have very few fine-grained distinguishable symptoms. A similar case can be observed in leaf wilt diseases like Root rot, Verticellium wilt, and Fusarium wilt. However, agricultural experts notified that these diseases have the same treatment to follow. So, there is an opportunity to improving the ease of use of the Crop Darpan system by combining all the diseases with visually similar symptoms and the same treatment into a single entity.

The pest disease are the most easily identifiable disease type as many of the pests such as White flies, Aphids, Jassids, etc. have distinguishable properties. However, there are few diseases of pests which have similar visual symptoms. For example, Spotted bollworm and Helicoverpa have almost similar effects on the plants and insects also have moderately similar features. To differentiate this kind of diseases easily, pictures and videos capturing the visual symptoms can be supplemented to the text description already available in the Crop Darpan system.

6.4 Limitations of the Evaluation

- The study is undertaken on the preliminary basis to validate the suppositions of the Crop Darpan Framework.

- Agriculture domain deals with the biological systems which can be affected by the weather, region and season. The sample chosen has limitations of time, space and regional considerations. Had it been bigger and wider sample, it may have provided a more comprehensive data.

- The results of this evaluation can be seen or considered as a confirmatory step in developing a full-fledged system.

– Absence of the baseline data, farmer's access to and utilization of disease diagnostic services hindered the comparative assessment of the Crop Darpan system with other systems such as Kisan call centers.

6.5 Summary of the Chapter

In this chapter, we have carried the evaluation study of Crop Darpan prototype for Cotton Crop to understand the perspectives of domain experts about the usage of the generalization-specialization framework and crop diagnosis procedure followed in Crop Darpan along with the practicality, interpretability, usability, data accuracy, extendibility, necessity aspects of Crop Darpan.

We have also reported the evaluation results about the supposition that the educated farmers has the ability to read and understand and respond to the questions regarding the existence of a symptom and the farmer has the ability to diagnose the crop diseases by navigating and confirming the existence of high-level and low-level symptoms presented by the Crop Darpan tool.

Overall, the results are encouraging. The domain experts are very positive about the system. The farmers are also able diagnose the crop problems in most of cases.

There is a suggestion about adding images to improve the performance.

In the next chapter, we provide summary, conclusions and future work.

Chapter *7*

Conclusions and Future Work

7.1 Summary

Despite the variety of agricultural extension approaches which include IT-based approaches, the majority of farmers in India are not acquiring actionable agricultural information. The farmers are also suffering from the issues of the knowledge gap and communication gap. On the other hand, India is into the smartphone revolution. The capabilities of smartphones are comparable to computers. The smartphone are providing the facility of browsing the web portals and searching information. Several smartphone based applications are developed, which are revolutionizing the e-Commerce and gaming. The literacy levels of Indian population is constantly growing. So, in coming future, it can be expected that most of the farmers will be educated and have the ability to use smartphones.

To enable the farmer to get the actionable agro-advisory, we have made an effort to build a smartphone based crop disease diagnosis guide, called Crop Darpan system, by exploiting the progress in data science and mobile phone technology. With an assumption of farmer possessing a smartphone and visiting the field, we have investigated a framework of Crop Darpan system for helping the farmer to identify the crop disease and acquire the corresponding solution.

The research issue was to develop a system to help the farmer to diagnose the crop problem using the smartphone. The proposed framework has been conceived due to several observations regarding issues with existing agricultural information dissemination technologies and opportunities identified from data science domain which could be extended to agricultural domain for building improved knowledge delivery solutions. The observations include (i) The issue of lack of efficient agricultural knowledge delivery tools in India (ii) Insights during field implementation of eSagu technology over since 2004 (iii) Trends in information technologies and (iv) concepts of generalization/specialization and human-in-the-loop systems. Based on the observations, we have conceived a few intuitions and confirmed the same through field experiments. Next, we interacted with the subject matter specialists and conceived the basic idea of generalization-specialization based crop diagnosis framework. Overall, we have extended generalization-specialization concept

and other related concepts from data science domain and made an effort to build a generalized crop diagnosis framework for all crops to help farmers, especially, Indian farmers.

The objective of the Crop Darpan system is to resemble the process followed by an agricultural expert to detect the crop disease. This requires the embedding of expert's knowledge about the visual symptoms of all diseases into the system. The components of Crop Darpan system are Knowledge acquisition protocol, Knowledgebase, Question popping model, and the User interface. We have explored three design choices: (i) number of questions to be asked per screen, (ii) algorithms followed for knowledge acquisition protocol from the domain expert, and (iii) question popping algorithm to pose questions to a farmer. The development of Crop Darpan has been carried out by following engineering research methodology. Overall, five design frameworks have been proposed based on the feedback of previous development. The final framework, which is called as Hierarchy-based protocol for knowledge acquisition and Hierarchy model for question popping with Multi-questions per screen (HHM) framework has been conceived.

In this thesis, we have presented the concepts of Crop Darpan, design of Crop Darpan and Operationlization for Cotton crop. Also, we have presented the evaluation study and showed the effectiveness of the system.

7.2 Conclusions/Advantages

Based on the experience of developing Crop Darpan and interaction with multiple stakeholders during the development and evaluation study with domain experts and farmers, we conclude the following.

- The evaluation study results show that the domain experts are in agreement with the diagnostic methodology of the system. The farmers are also able to diagnose most of the crop diseases.
- We have proposed a generic framework. That is, it is possible to design a full-fledged system for any given crop to enable the farmer to identify crop diseases.
- Normally, agricultural experts diagnose most of the crop diseases based on visual symptoms. The system mimics the behavior of agricultural experts and helps the farmers to identify the crop disease.
- The system assumes that the farmer does not know the name of the disease. So, if the farmer uses the proposed system and unable to detect a crop disease, his/her crop is not facing any problem, i.e., the system is acting as an auditing tool. Majority of the other systems, like call centers and web-based system are based on the assumption that farmer knows the name of the crop disease, which is not valid for the majority of Indian farmers.

- The system can be extended to all languages and dialectics. Once developed in one language, it only requires translation effort to other languages.
- The proposed system is salable to all crops and regions in India and abroad.

7.3 Limitations

We have explored the design of Crop Darpan system by considering Cotton crop. The following are the limitations of the proposed system.

- It is assumed that farmer is able to identify the visual symptom of the crop disease by seeing the crop. As a result, the system can be used only by progressive farmers. It may also require a short duration training (for a few days).
- Currently, the system is able to identify only one crop disease. However, a crop can suffer from multiple diseases at the same time.
- We have only presented a proof of concept and the obtained results shows the feasibility of the system. Further refinement is required with a rigorous evaluation study and the feedback.

7.4 Future Work

The following are the list of future works.

- The system is being designed by considering that all crop diseases occur with equal probability. But, normally, certain crop diseases occur more often as compared to other diseases. So, by giving priority to the corresponding visual symptoms based on domain knowledge, it is possible to develop more efficient system.
- Extending the proposed system to enable the farmer to detect multiple diseases.
- Current system describes the visual symptoms using text only. Ease of use of the system can be improved by providing the multimedia descriptors such as images, video, and voice recording along with the text description.
- Currently, when the farmer says "YES" to the occurrence of a visual symptom in the field, we are considering it as a definitive response. However, in any system with involvement of human is error prone (*to err is human*). In the proposed framework, as we are involving human element, there is a possibility of misidentification of the symptom due to lack of

sufficient knowledge about the perception of the symptom. To make it more robust against such mistakes, we are planning to propose improved framework by dividing the diagnosis process into two steps: Detection step and Confirmation step. The Detection step is similar to the method presented in this thesis. For Confirmation, we will ask other detailed and specific information by providing appropriate images and other related symptoms. If the farmer fails in the confirmation step, he/she has to start using the system again or consult the domain expert.

- Evaluation study with large number of farmers and stakeholders.

Related Publications

1. Krishna Reddy, U. Narendra Babu, A. Mamatha, Vaishnavi Gutta, N.Lavanya, Balaji Naik Banoth, G.Sreenivas, Seishi Ninomiya, Towards Building a Field Diagnosis Guide for Farmers, In the Proceedings of Research Frontiers in Precession Agriculture, The Asia-Pacific Federation for Information Technology in Agriculture (AFITA) and World Congress on Computers in Agriculture (WCCA) (AFITA/WCCA 2018) Conference, IIT Bombay, India, October 24-26, 2018.

2. P.Krishna Reddy, U.Narendra Babu, P. Revanth Rathan, A.Srinivas Reddy, Balaji Naik Banoth, Seishi Ninomiya. A Smart Phone Based Field Diagnosis Guide for Farmers. 4th International Conference on Agriculture & Animal Husbandry, University of Hyderabad, Hyderabad, India, August 2019.

3. P Krishna Reddy, U Narendra Babu, P Revanth Rathan, A Srinivas Reddy, BalajiNaik Banoth and Seishi Ninomiya, Building a smart phone based field diagnosis guide for farmers by extending the concept of generalization, XIX International Plant Protection Congress (IPPC 2019), 10-14, November 2019, Hyderabad, India.

Acknowledgement: This work is supported by India-Japan Joint Research Laboratory Project entitled "Data Science based farming support system for sustainable crop production under climatic change (DSFS)", funded by Department of Science and Technology, India (DST) and Japan Science and Technology Agency (IST).

Bibliography

[1] Twenty questions. https://en.wikipedia.org/wiki/Twenty_Questions. Last Accessed: 04-12-2019.

[2] Iffco kisan: Empowering farmers with timely, relevant and high quality information. http://www.iffcokisan.com/home/services. Last Accessed: 04-12-2019.

[3] Classification: Basic concepts, decision trees, and model evaluation. https://www-users.cs.umn.edu/~kumar001/dmbook/ch4.pdf. Last Accessed: 04-12-2019.

[4] Farmers portal: One stop shop for farmers. https://farmer.gov.in/M_Aboutus.aspx. Last Accessed: 04-12-2019.

[5] Kantar imrb icube report. https://imrbint.com/images/common/ICUBEE284A2/_2019/_Highlights.pdf. Last Accessed: 04-12-2019.

[6] India's internet base crosses 500 million mark, driven by rural india. https://www.livemint.com/industry/telecom/internet-users-exceed-500-million-rural-india-driving-growth-report-1552300847307.html. Last Accessed: 04-12-2019.

[7] mkisan: A portal of government of india for farmer centric mobile based services. https://mkisan.gov.in/. Last Accessed: 04-12-2019.

[8] Plantix: The mobile doctor for your pocket. https://plantix.net/en/. Last Accessed: 04-12-2019.

[9] Rice doctor, 2003. international rice research institute, philipines. howpublished = http://www.knowledgebank.irri.org/decision-tools/rice-doctor, note = Last Accessed: 04-12-2019,.

[10] Number of smartphone users in india from 2015 to 2022 (in millions). https://www.statista.com/statistics/467163/forecast-of-smartphone-users-in-india/. Last Accessed: 04-12-2019.

[11] S. Abu-Naser, K. Kashkash, and M. Fayyad. Developing an expert system for plant disease diagnosis. *Journal of Artificial Intelligence*, 3(4):269–276, 2010.

[12] J. S. Aikins, J. C. Kunz, E. H. Shortliffe, and R. J. Fallat. Puff: an expert system for interpretation of pulmonary function data. *Computers and biomedical research*, 16(3):199–208, 1983.

[13] D. Al Bashish, M. Braik, and S. Bani-Ahmad. Detection and classification of leaf diseases using k-means-based segmentation and. *Information Technology Journal*, 10(2):267–275, 2011.

[14] A. T. Al-Taani. An expert system for car failure diagnosis. *IEC (Prague)*, 5:457–560, 2005.

[15] N. Alexandratos and J. Bruinsma. World agriculture towards 2030/2050: the 2012 revision. 2012.

[16] N. Anandaraja, N. Sriram, R. Venkatachalam, and H. Philip. mkisan short message services: A cross sectional analysis in tamil nadu.

[17] J. R. Anderson and G. Feder. Agricultural extension. *Handbook of agricultural economics*, 3:2343–2378, 2007.

[18] M. Anderson. Research, extension, and education. 1996.

[19] N. Y. Asabere and S. Kusi-Sarpong. Mves: A mobile vehicle expert system for the automobile industry. *International Journal of Engineering Research and Applications (IJERA) ISSN*, pages 2248–9622, 2012.

[20] S. C. Babu, P. Joshi, C. J. Glendenning, A.-O. Kwadwo, and S. V. Rasheed. The state of agricultural extension reforms in india: strategic priorities and policy options. *Agricultural Economics Research Review*, 26(347-2016-17086):159–172, 2013.

[21] V. R. Basili. The experimental paradigm in software engineering. In *Experimental Software Engineering Issues: Critical Assessment and Future Directions*, pages 1–12. Springer, 1993.

[22] D. Baulcombe, I. Crute, B. Davies, J. Dunwell, M. Gale, J. Jones, J. Pretty, W. Sutherland, and C. Toulmin. *Reaping the benefits: science and the sustainable intensification of global agriculture*. The Royal Society, 2009.

[23] P. Bazilinskyy. The turing test and the 20 question game. 04 2013. doi: 10.13140/RG.2.1.4376.6245.

[24] J. Beynon, S. Akroyd, A. Duncan, and S. Jones. *Financing the future: Options for agricultural research and extension in sub-Saharan Africa*. Oxford Policy Management, 1998.

[25] P.-M. Bosc, J. A. Berdegué, M. Goïta, J. D. v. d. Ploeg, K. Sekine, and L. Zhang. Investing in smallholder agriculture for food security: a report by the high level panel of experts on food security and nutrition. 2013.

[26] S. Branson, C. Wah, F. Schroff, B. Babenko, P. Welinder, P. Perona, and S. Belongie. Visual recognition with humans in the loop. In *European Conference on Computer Vision*, pages 438–451. Springer, 2010.

[27] M. L. Brodie, J. Mylopoulos, and J. W. Schmidt. *On conceptual modelling: Perspectives from artificial intelligence, databases, and programming languages*. Springer Science & Business Media, 2012.

[28] A. Camargo and J. Smith. An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosystems engineering*, 102(1):9–21, 2009.

[29] Y. Cheng, Q. Hu, and J. Yang. An expert system for education: Ipts. *IEEE Computer*, 1988.

[30] W. J. Clancey and R. Letsinger. *NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching*. Department of Computer Science, Stanford University, 1982.

[31] C. Costopoulos, M. Ntaliani, and S. Karetsos. Studying mobile apps for agriculture. *IOSR J. Mob. Comput. Appl*, 3:44–49, 2016.

[32] A. Fuentes, S. Yoon, S. Kim, and D. Park. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9):2022, 2017.

[33] R. Gandhi, R. Veeraraghavan, K. Toyama, and V. Ramprasad. Digital green: Participatory video for agricultural extension. In *2007 International conference on information and communication technologies and development*, pages 1–10. IEEE, 2007.

[34] A. Gerevini, A. Perini, F. Ricci, D. Forti, C. Ioriatti, L. Mattedi, A. Monetti, et al. Pomi: an expert system for integrated pest management of apple orchards. 1992.

[35] L. Getoor and A. Machanavajjhala. Entity resolution: theory, practice & open challenges. *Proceedings of the VLDB Endowment*, 5(12):2018–2019, 2012.

[36] P. A. Glasow. Fundamentals of survey research methodology. *Retrieved January*, 18:2013, 2005.

[37] C. J. Glendenning, S. Babu, and K. Asenso-Okyere. Review of agricultural extension in india. *Are farmers information needs being met?, IFPRI Discussion Paper 01048*, Dec. 2010.

[38] A. Gulati, P. Sharma, A. Samantara, and P. Terway. *Agriculture extension system in India: Review of current status, trends and the way forward*. Indian Council for Research on International Economic Relations, 2018.

[39] J. Hua. Study on knowledge acquisition techniques. In *2008 Second International Symposium on Intelligent Information Technology Application*, volume 1, pages 181–185. IEEE, 2008.

[40] P. Jackson. *Introduction to expert systems*. Addison-Wesley Longman Publishing Co., Inc., 1998.

[41] D. Kafura. Generalization. `http://people.cs.vt.edu/~kafura/cs2704/generalization.html`, June 21, 1996.

[42] J. Kaiser. The development of cai: An expert system in education. 1985.

[43] S. Khanna, A. Kaushik, M. Barnela, et al. Expert systems advances in education. In *Proceedings of the National Conference on Computational Instrumentation NCCI-2010. CSIO*, pages 109–112, 2010.

[44] T. E. T. E. Lele Uma, Pretty Jules and K. Maggie. Transforming agricultural research for development. Report for the Global Conference on Agricultural Research (GCARD), 2010.

[45] H. Lemmon. Comax: An expert system for cotton crop management. *Science*, 233(4759): 29–33, 1986.

[46] R. K. Lindsay, B. G. Buchanan, E. A. Feigenbaum, and J. Lederberg. Dendral: a case study of the first expert system for scientific hypothesis formation. *Artificial intelligence*, 61(2): 209–261, 1993.

[47] B. Mahaman, H. Passam, A. Sideridis, and C. Yialouris. Diares-ipm: a diagnostic advisory rule-based expert system for integrated pest management in solanaceous crop systems. *Agricultural Systems*, 76(3):1119–1135, 2003.

[48] R. Michalski, J. Davis, V. Bisht, and J. Sinclair. A computer-based advisory system for diagnosing soybean diseases in illinois. *Plant Disease*, 67(4):459–463, 1983.

[49] S. Mohan and N. Arumugam. An intelligent front-end for selecting evapotranspiration estimation methods. *Computers and electronics in agriculture*, 12(4):295–309, 1995.

[50] S. P. Mohanty, D. P. Hughes, and M. Salathé. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7:1419, 2016.

[51] D. of Agriculture and G. o. I. Cooperation, Ministry of Agriculture. Kisan Call center. `https://dackkms.gov.in/account/login.aspx`, Nov. 2017.

[52] J. R. Olson and H. H. Rueter. Extracting expertise from experts: Methods for knowledge acquisition. *Expert systems*, 4(3):152–168, 1987.

[53] M. K. P.Krishna Reddy, B.Bhaskar Reddy. Village-level esagu: a scalable and location-specific agroadvisory system. *In Proceedings of 3rd National Conference on Agro-Informatics and Precision Agriculture 2012*, 2007.

[54] R. Plant. Calex: an integrated expert decision support system for agricultural management. *American Society of Agricultural Engineers (Microfiche collection)(USA)*, 1988.

[55] J. Pramanik, B. Sarkar, and S. Kandar. Impact of ict in rural development: Perspective of developing countries. *American Journal of Rural Development*, 5(4):117–120, 2017. ISSN 2333-4770. URL `http://pubs.sciepub.com/ajrd/5/4/5`.

[56] G. Prasad and B. D. A. Vinaya. A study on various expert systems in agriculture. *Computer Sciences and Telecommunications*, (4):81–86, 2006.

[57] A. Rafea. Expert system applications: Agriculture. *Central Laboratory for Agricultural Expert Systems, PO Box*, 100, 1998.

[58] A. Rafea and A. El-Beltagy. Expert systems as an appropriate technology to speed up agriculture desert development in egypt. In *Proceedings of the Second International Conference on Desert Development, Harwood Academic Publisher*, 1987.

[59] A. Rafea and M. Warkentin. Knowledge engineering: Creating expert systems for crop production management in egypt. chapter 7: 89-104 in expert systems in developing countries: Practice and promise, stephen r. ruth and chip mann (eds.), boulder, co, 1991.

[60] E. Rahm and H. H. Do. Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, 23(4):3–13, 2000.

[61] K. Ramamritham, A. Bahuman, and S. Duttagupta. aaqua: a database-backended multilingual, multimedia community forum. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*, pages 784–786. ACM, 2006.

[62] K. R. P. Ratnam, B.V. and G. Reddy. esagu: An it based personalized agricultural extension system prototype-analysis of 51 farmers' case studies. *International Journal of Education and Development using ICT*, 2(1), 2006.

[63] O. Razeghi, G. Qiu, H. Williams, K. Thomas, and I. VIPLAB. Skin lesion image recognition with computer vision and human in the loop. *Medical Image Understanding and Analysis (MIUA), Swansea, UK*, pages 167–172, 2012.

[64] P. K. Reddy, G. V. Ramaraju, and G. S. Reddy. esagu^{tm}: a data warehouse enabled personalized agricultural advisory system. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, Beijing, China, June 12-14, 2007*, pages 910–914, 2007.

[65] A. J. Riel. *Object-oriented design heuristics*, volume 335. Addison-Wesley Publishing Company, 1996.

[66] J. Roach, R. Virkar, M. Weaver, and C. Drake. Pomme: a computer-based consultation system for apple orchard management using prolog. *Expert Systems*, 2(2):56–69, 1985.

[67] S. S. Sannakki, V. S. Rajpurohit, V. Nargund, A. Kumar, and P. S. Yallur. Leaf disease grading by machine vision and fuzzy logic. *Int J*, 2(5):1709–1716, 2011.

[68] S. S. Sannakki, V. S. Rajpurohit, V. B. Nargund, and P. Kulkarni. "diagnosis and classification of grape leaf diseases using neural networks". In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pages 1–5, July 2013. doi: 10.1109/ICCCNT.2013.6726616.

[69] R. Saravanan. *ICTs for agricultural extension: Global experiments, innovations and experiences*. New India Publishing, 2010.

[70] S. K. Sarma, K. R. Singh, and A. Singh. An expert system for diagnosis of diseases in rice plant. *International Journal of Artificial Intelligence*, 1(1):26–31, 2010.

[71] A. P. Saygin, I. Cicekli, and V. Akman. Turing test: 50 years later. *Minds and machines*, 10 (4):463–518, 2000.

[72] S. Shinde, D. Piplani, K. Srinivasan, D. Singh, R. Sharma, and P. Mohnaty. mkrishi: simplification of ivr based services for rural community. In *Proceedings of the India HCI 2014 Conference on Human Computer Interaction*, page 154. ACM, 2014.

[73] E. H. Shortliffe and B. G. Buchanan. *Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing Company, 1985.

[74] P. K. Singh and M. S. Hussain. Expert system for crop selection in agriculture: A critical review. *International Journal of Advanced Research in Computer Science*, 9(Special Issue 2):143, 2018.

[75] J. M. Smith and D. C. Smith. Database abstractions: aggregation and generalization. *ACM Transactions on Database Systems (TODS)*, 2(2):105–133, 1977.

[76] M. K. Swamy and P. K. Reddy. Improving diversity performance of association rule based recommender systems. In *Database and Expert Systems Applications*, pages 499–508. Springer, 2015.

[77] P. J. Tichenor, G. A. Donohue, and C. N. Olien. Mass media flow and differential growth in knowledge. *Public opinion quarterly*, 34(2):159–170, 1970.

[78] N. Uphoff. Empowerment of farmers through ict. In *Draft paper for ECOSOC Expert Group Meeting on Promoting Empowerment of People in Advancing Poverty Eradication, Social Integration, and Decent Work for All, UN HQ, NYC, Cornell University*, 2012.

[79] A. Uzhinskiy, G. Ososkov, P. Goncharov, and A. Nechaevskiy. Multifunctional platform and mobile application for plant disease detection. *Proceedings of the 27th International Symposium Nuclear Electronics and Computing (NEC'2019) Budva, Becici, Montenegro, September 30 – October 4*, 2019.

[80] N. J. Vasant P. Gandhi. Decision-oriented information systems for farmers: A study of kisan call centres (kcc), kisan knowledge management system (kkms), farmers portal, and m-kisan portal. 2018.

[81] V. Veluchamy and Y. Hari. An expert system for automobile car diagnosis. In *[1988] Proceedings. The Twentieth Southeastern Symposium on System Theory*, pages 509–513, March 1988. doi: 10.1109/SSST.1988.17103.

Appendix A

Part 1: Questionnaire for the Educated Farmer

INDO-JAPAN PROJECT
EFFICIENT KNOWLEDGE TRANSFER

Hypothesis: Farmer is facing the problem in communicating the crop problem to agricultural expert through text/telephone/image [YES/NO]

1.Name	2.Age	3.Education.	4.Agri. Experience

5. Names of three crops you cultivated and have a good knowledge:

6. For each crop, list the crop the problems you have encountered in last 10 years from pre-sowing to harvesting.

a. Crop1.
b. Crop2
c.Crop3

Mode: Telephone/Cell Phone (Voice)

7. Are you confident of communicating each crop problem to scientist through telephone? (yes/no). If the answer is NO, give reasons. (Refer c3 of Table 1)
8. Which of the problems can not be communicated or difficult to communicate through telephone and why?

Mode: Images

9. Are you confident of communicating each of the problems to scientist through photos? If the answer is NO, give the reasons. (Refer Refer C3 of Table 1)
10. Which of the problems can not be communicated through photos and why?

Mode: Video

11. Are you confident of communicating each of the problem to scientist through videos? If the answer is NO, give reasons (Refer Refer C4 of Table 1)
12.

Mode: Text (SMS)

13. Are you confident of communicating each of the problem to scientist through text? (yes/no). If the answer to the above question is NO, explain the reason. (Refer C5 of Table 1).
14. Which of the above problems can not be communicated through text and why?

Mode: Combined (Voice, image, video, text)

15. Are you confident of communicating each of the problem by combining voice, image. Video and text (yes/no). If the answer to the above question is NO, explain the reason. (Refer C6 of Table 1). (Refer C6 of Table 1)

Common Questions

16. Suppose you have called/sent photos/text the scientist and the scientist about crop problem is not understanding the problem, what do you do?
17. In your view, what should be done (the list off steps the department/government should take) to enable the farmer to ask the question in an efficient manner through telephone?

18. Do you know any reference material available for agriculture? (Examples: Anna data, Vyavasaya Panchangam,...)
19. If you know reference material, did you refer any reference material and asked the question?
20. If you go through the reference material, will you able to ask a better question to the scientists, regarding the crop problem? What are the reasons?
21. Do you know any IT-based tools?
22. During your educational carrier, did you refer any guide/reference material?
23. Suppose we want to build IT-based tool to help you to put a better question to agricultural scientists. What kind of services should it provide to you to help you in forming a better question?

Table 1. <Crop, Problem> (1)	Confident of communicating through telephone (YES/NO), If NO, give the reasons. (2)	Confident of communicating through photos (WhatsApp) (YES/NO), If NO, give the reasons. (3)	Confident of communicating through video (YES/NO), If NO, give the reasons. (4)	Confident of communicating through text (YES/NO), If NO, give the reasons. (5)	Confident of communicating through combination of (text, image, video, voice) (YES/NO). IF NO, give the reasons. (6)

Part 2: Questionnaire for the Agricultural scientist

INDO-JAPAN PROJECT
EFFICIENT KNOWLEDGE TRANSFER

Hypothesis: Agricultural expert is facing the problem to resolve crop problem received through text/telephone/image [YES/NO]

1.Name	2.Age	3.Education.	4.Agri. Experience (providing crop advisory)

5. Names of three major crops you advice:

6. For each crop, list the crop the problems farmers face (from pre-sowing to harvesting).

a. Crop1.
b. Crop2
c.Crop3

Mode: Telephone/cell-phone/voice:

7. Are you confident that farmer will communicate crop problems clearly through telephone to you? (yes/no). If the answer to the preceding question is NO, explain the reasons. (Refer C2 of Table 1)
8. If the farmer is calling you, and you are not understanding the problem, what do you suggest the farmer to pose the question in a more efficient manner?
9. In your view, what should be done to enable the farmer to ask the question through telephone?
10. Which of the problems can not be communicated through telephone and why?

Mode: WhatsApp (photos)

11. Are you confident that farmer will communicate crop problems efficiently to you through photos? (yes/no). If NO, explain the reasons. (C3 of Table 2)
12. If the farmer has sent a photograph to you, and you are not understanding the problem, what do you suggest the farmer?
13. What should be done to enable the farmer to ask the question through images?
14. ich of the problems can not be communicated through photos and why?
15. What problems could only communicated through videos? Why? (Refer C4 of Table 1)

Mode: Text (SMS)

16. Are you confident that farmer will communicate each of the problem to you? (yes/no). If the answer to the preceding question is NO, explain the reasons. (Refer C5 of Table 1)
17. If the farmer has sent a text and you are not understanding the problem, what kind of suggestions you provide the farmer to ask the question in a better manner?
18. In your view, what should be done to enable the farmer to ask the question properly through text?
19. Which of the problems can not be communicated through text and why?

Mode. Combined media (text, telephone, images, video)

20. Explain the role of combined media in communicating each problem (C6 of the Table 1)

GENERAL

21. List the reference material available for agriculture. (Examples: Anna data, Vyavasaya Panchangam,...)
22. Do you feel that farmers are accessing reference material to ask the question?
23. If the farmers go through the reference material, do you feel that, farmers will ask better question to the scientists regarding the crop problem? What are the reasons?
24. Suppose, a farmer sees the crop. If we build a computer application on smart phone to help the farmer for asking better question to agricultural scientist, will it help? If yes, can you share your ideas/imagination/architecture on such tool.
25. What would be the role of text, images, voice, and video in building such tool?

Table 1. <Crop, Problem> (1)	Potential of telephone to communicate the problem iin a clear manner (YES/NO), If NO, give the reasons. (2)	Use of WhatsApp/photos to communicated the problem in a clear manner (YES/NO), If NO, give the reasons. (3)	Use of video to communicate the problem in a clear manner (YES/NO). If NO, give the reasons. (4)	Use of text to communicate the problem in a clear manner (YES/NO). If NO, give the reasons. (5)	Use of combination (text, phone, photos, video) to communicate the problem. (6)

Part 3: Crop Problem Information Form

INDO-JAPAN PROJECT
EFFICIENT KNOWLEDGE TRANSFER

1	Farmer Name		2	Date (DD/MM/YY)	
3	Year of Birth		4	Contact Number	
5	Village Name		6	Mandal	
7	District		8	State	
9	Education: (Up to 5^{th} / Up to 10^{th} / $>10^{th}$)		10	Acreage	
11	Caste : SC / ST / BC / OC, other_____		12	Work Style of Farmer	Serious / Not Serious
13	Farming Experience in Years: (Below 5 years / Between 5 years and 10 years / Above 10 years)				

Crop Information					
14	Crop Name		15	Sowing Date	
16	Water Source	Rain fed / Irrigated	17	Soil Type	Red / Black / Problematic soil
18	Weather Information	Sunny	Cloudy		Rainy
		Calm	Light Wind		Strong Wind

Previous Action Taken		
19	Fertilizers Applied	
20	Pesticides Applied	
21	Management Practices Taken	

22. Farmers Perception:

23. Photographs Attached:

Appendix B

Cotton diseases-symptoms data collected from domain experts (30 Diseases)

S.No	Disease	Symptoms	Symptom Question	Management
1	Spotted Boll Worm	1. Bored bolls	1. Do Bolls have holes?	**Chemical**: Spray Chlorantraniliprole @ 0.3 ml or Flubendiamide @ 0.25 ml or Indoxacarb 14.5%SC @1.0 ml or Profenofos @ 2.0 ml or Triazophos 40%EC @ 4.0 ml by mixing in 1 litre of water.
		2. Premature shedding of bolls	2. Do you find any Premature shedding of bolls?	
		3. Premature opening of bolls	3. Do you detect bolls with premature opening?	
		4. Presence of larvae/insects on bolls	4. Do you find any larvae/insects on bolls?	
		5. Drying of terminal buds	5. Do you detect any drying of terminal buds?	
		6. Presence of larvae/insects on bolls	6. Do you find any larvae/insects on bolls?	
		7. The larvae insert its head and ¾ of the body inside the boll	7. Is larvae's head inside the boll along with 3/4 th body?	
		8. The caterpillars have a number of black and brown spots on the body	8. Do you detect any caterpillar with number of black and brown spots on the body?	
	మచ్చల పురుగు / తలనత్త పురుగు	1. కాయలపై రంధ్రాలు ఉన్నాయి	1. కాయలపై రంధ్రాలు ఉన్నాయా ?	**రసాయన నిర్వహణ:** లీటరు నీటికి Chlorantraniliprole @ 0.3 ml లేదా Flubendiamide @ 0.25 ml లేదా Indoxacarb 14.5%SC @1.0 ml లేదా Profenofos @ 2.0 ml లేదా Triazophos 40%EC @ 4.0 ml చొప్పున
		2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయి	2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయా?	
		3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా ?	
		4. కాయలపై పురుగులున్నాయి	4. కాయలపై పురుగులున్నాయా?	
		5. చిగురు మొగ్గలు ఎండిపోతున్నాయి	5. చిగురు మొగ్గలు ఎండిపోతున్నాయా?	
		6. కాయలపై పురుగులున్నాయి	6. కాయలపై పురుగులున్నాయా?	

		7. లార్వా తలభాగం మరియు 3/4 వంతు శరీరం కాయలలో ఉన్నదా?	7. లార్వా తలభాగం మరియు 3/4 వంతు శరీరం కాయలలో ఉన్నదా?	కలిపి పిచికారి చేయండి.
		8. నల్లని లేదా ఉదా రంగు మచ్చలు గల లద్దె పురుగులు ఉన్నాయి	8. నల్లని లేదా ఉదా రంగు మచ్చలు గల లద్దె పురుగులు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
2	Helcoverpaor american boll worm	1. Bored squares/bored flower buds	1. Do flower buds have holes?	**Chemical**: If pest load is severe, spray indoxacarb @ 1 ml or spinosad @ 0.3 ml/ or emamectin benzoate @ 0.5 g by mixing in 1 litre of water. **Biological:**1. Spray HaNPV @ 200 LE by mixing with jaggery 1 kg + sandovit 100 ml or Robin Blue 50 g. This solution is sufficient for one acre. Spraying should be done thrice during evening time at 10-15 days interval on observing the eggs or first instar larvae. 2. Release Larvael parasitoid such as Campoletis chloridae (Ichneumonidae) or Eriborus angenteopilorus or Diadegma fenestalis or Bracon brevicornis or Peribaca orbata etc in the crop
		2. Premature shedding of bolls	2. Do you find any Premature shedding of bolls?	
		3. Premature opening of bolls	3. Do you detect bolls with premature opening?	
		4. Holes in leaves	4. Do you observe any holes in leaves?	
		5. Presence of larvae/insects on bolls	5. Do you find any larvae/insects on bolls?	
		6. Circular holes on leaves	6. Do you find circular holes on leaves?	
		7. The larvae insert only its head inside the boll	7. Is only larvae's head inside the boll?	
	శనగపచ్చ పురుగు	1. పూతమీద/మొగ్గ లమీద రంధ్రాలున్నాయి	1. పూతమీద/మొగ్గల మీద రంధ్రాలు ఉన్నాయా?	**రసాయన నిర్వహణ:** పురుగు తీవ్రంగా ఉంటే, లీటరు నీటికి Indoxacarb @ 1 ml లేదా Spinosad @ 0.3 ml/ లేదా Emamectin Benzoate @ 0.5 g చొప్పున కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** (1) గుడ్లు లేదా చిన్న లార్వాలు గమనించినచో HaNPV @ 200 LE తో 1 కిలో బెల్లం + Sandovit 100 ml లేదా Robin Blue 50 g కలిపి 1 ఎకరానికి సరిపడ ద్రావణాన్ని తయారు చేసుకొని పిచికారి చేయండి. ఈ ద్రావణాన్ని సుమారు 5 రొజులకొకసారి మూడు సార్లు సాయంత్రవేళలో పిచికారీ చేయండి. (2)పంటలో Larvael parasitoid వంటి Campoletis chloridae (Ichneumonidae) లేదా Eriborus angenteopilorus లేదా Diadegma fenestalis లేదా Bracon brevicornis లేదా Peribaca orbata మొదలైనవి వదలండి .
		2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయి	2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయా?	
		3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా?	
		4. ఆకులపై రంధ్రాలు ఉన్నాయి	4. ఆకులపై రంధ్రాలు ఉన్నాయా?	
		5. కాయలపై పురుగులున్నాయి	5. కాయలపై పురుగులున్నాయా?	
		6. ఆకులలో గుండ్రటి రంధ్రములు ఉన్నాయి	6. ఆకులలో గుండ్రటి రంధ్రములు ఉన్నాయా?	
		7. లార్వాలు, తమ తలభాగం కాయలో ఉంచి కాయలను తింటు ఉన్నాయి	7. లార్వాలు, తమ తలభాగం కాయలో ఉంచి కాయలను తింటు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
3	Pink bollworm	1. Unopened and twisted flowers in rose shape (rosette flowers)	1. Do you find any unopened and twisted flowers in rose shape?	**Chemical**: Use pheromone trap to monitor the adult moth activity. If adults are found in traps take up following control measures. (1) Spray Carbaryl 5%Dust @ 8 kg/acre or (2). Spray Phosalone 35%EC @ 4.0 ml or (3) Triazophos 40EC @ 5 ml by mixing in 1 litre of water. **Biological:**Use pheromone trap to monitor the adult moth activity. If adults are found in traps, release Trichogrammatoidea bacterae egg Parasitoid @ 40,000 per acre at three weeks interval.
		2. Shedding of flowers	2. Do you find any premature shedding of flowers?	
		3. Premature opening of bolls	3. Do you detect bolls with premature opening?	
		4. Fed bolls but holes not visible	4. Do you observe any boll damage (holes are not visible)?	
		5. Lint colour change	5. Do you detect any change in lint colour?	
		6. Presence of larvae/insects on bolls	6. Do you find any larvae/insects on bolls?	
		7. Pink colour lint	7. Do you find pink colour lint?	
		8. Presence of small and pink colour larvae with brown head inside the boll	8. Do you find any presence of small and pink colour larvae inside the boll?	
	గులాబి రంగు కాయ తొలుచు పురుగు	1. పువ్వులు సక్రమంగా విచ్చుకోవడం లేదు	1. పువ్వులు సక్రమంగా విచ్చుకోవడం లేదా?	**రసాయన నిర్వహణ:** లింగా ఆకర్షక బుట్టలను(Pheromone trap) పంటలో పెట్టండి. ఆ బట్టలలో పురుగులు గమనించినచో ఈ క్రింది విదంగా నియంత్రించండి (1) Carbaryl 5% Dust @ 8 kg/acre ని పిచికారీ చేయండి, లేదా (2).Phosalone 35%EC @ 4.0 ml లీటరు నీటిలో కలిపి పిచికారీ చేయండి, లేదా (3)Triazophos 40EC @ 5 ml లీటరు నీటిలో కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** లింగా ఆకర్షక బుట్టలను(
		2. పువ్వులు రాలుతున్నాయి	2. పువ్వులు రాలుతున్నాయా?	
		3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా?	
		4. కాయలపై రంద్రాలు కనపడకుండా పురుగు తిన్న నష్ట లక్షణాలు	4. కాయలపై రంద్రాలు కనపడకుండా పురుగు తిన్న నష్ట లక్షణాలు	

		ఉన్నాయి	ఉన్నాయా?	Pheromone trap) పంటలో పెట్టండి. ఆ బట్టలలో పురుగులను గమనించినచో Trichogrammatoidea bacterae egg Parasitoid @ 40,000 per acre మూడు వారాలకు ఒకసారి వదలండి
		5. దూది రంగు మారింది	5. దూది రంగు మారిందా?	
		6. కాయలపై పురుగులున్నాయి	6. కాయలపై పురుగులున్నాయా?	
		7. కాయ లోపలి పత్తి గులాబి రంగు లోకి మారుతున్నది	7. కాయ లోపలి పత్తి గులాబి రంగు లోకి మారుతున్నదా?	
		8. నల్లటి తల గల చిన్న గులాబి రంగు పురుగులు ఉన్నాయి	8. నల్లటి తల గల చిన్న గులాబి రంగు పురుగులు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
4	Tobacco caterpillar	1. Fed leaves	1. Do you observe any leaves eaten by insect?	**Chemical:** (1). To control small size larvae, spray Novaluron @ 1.0 ml or Lufenuron @ 1.25 ml or Thiodicarb @ 1.5 g or Imamectin Benzoate @ 0.5 g or Flubendamide @ 0.3 ml or Clorantraniliprol @ 0.3 ml by mixing in 1 litre of water. (2).If big size larvae observed, prepare poison bait with1.0 litre monochrotophos or 750 ml Chlorpyrifos or 300 g thiodicarb + 10 kg rice brawn +2.0 kg jaggery + sufficient water to make small grannules and broadcast in one acre field during evening time. **Biological:**(1) Arrange 4 pheramone traps per acre,
		2. Larvae/Insects on lower side of the leaf	2. Do you observe any larvae/insects on lower side of the leaves?	
		3. Skeletonization of leaves	3. Do you observe any Skeletonization of leaves?	
		4. Holes in leaves	4. Do you observe any holes in leaves?	
		5. Groups of small size larvae on lower side of the leaf	5. Do you observe any small size larvae on lower side of the leaves?	
		6. Irregular holes	6. Do you observe	

		on leaves	irregular holes on leaves?	(2).Grow 20 castor plant per acre as trap crop , (3) Spray 5% neem oil in ml by mixing in 1 litre of water.
	పొగాకు లద్దె పురుగు	1. ఆకులనూ పురుగులు తిన్నాయి	1. ఆకులనూ పురుగులు తిన్నాయా?	1.చిన్న సైజు లార్వేను నియంత్రిచడానికి novaluron @ 1.0 ml లేదా lufenuron @ 1.25 ml లేదా thiodicarb @ 1.5 g లేదా imamectin benzoate @ 0.5 g లేదా flubendamide @ 0.3 ml లేదా clorantraniliprol @ 0.3 ml లీటరు నీటిలో కలిపి పిచికారి చేయండి. (2). పెద్ద సైజు లార్వేను గమనించినచో, ఒక ఎకరాల పొలంలో, 1.0 లీటర్ Monochrotophos లేదా 750 మి.లీ Chlorpyriphos లేదా 300 గ్రా thiodicarb, 10 కిలోల బియ్యం + 2.0 కిలోల బెల్లతో చిన్న కణికలు తయారు చేయడానికి తగినంత నీటీలో కలిపి సాయంత్రం సమయంలో ఎరగా చల్లండి
		2. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయి	2. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయా?	
		3. ఆకులు జల్లెడగా మారినాయి	3. ఆకులు జల్లెడగా మారినాయా?	
		4. ఆకులపై రంద్రాలు ఉన్నాయి	4. ఆకులపై రంద్రాలు ఉన్నాయా?	
		5. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నది	5. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నదా?	
		6. ఆకులలో వివిధ పరిమాణం గల రంధ్రములు ఉన్నాయి	6. ఆకులలో వివిధ పరిమాణం గల రంధ్రములు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
5	Leaf hoppers	1. Stunted growth	1. Do you detect any stunted growth of a plant?	**Chemical:** (1).Seed treatment (after acid delinting) : Carbosulfan @ 40 g/kg seed or Imidacloprid 70 WS @ 5 g/kg seed or Thiamethoxam 70 WS@ 4 g/kg seed gives protection for 30 days against sap feeders sucking pests.(2).Soil application : Carbofuran 3G granules @ 13
		2. Curling of leaves	2. Do you observe any curling of leaves?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Spots on leaves	4. Do you observe any spots on leaves?	

		5. Upward curling of leaves	5. Are leaves curling upwards?	kg/acre at sowing.(3). Stem application : Monocrotophos @1:4 with water or Methyl Demeton @1:4 with water or Imidacloprid 200 SL @ 1: 20 with water three times at 20 – 25, 30 – 35, 40 – 45 Days after sowing using soft brush at tender green part of the stem.(4). If seed treatment and soil application not done take up foliar spray with monocrotophos @ 1.5 ml or imidacloprid @ 0.4 ml or methyldemeton @2 ml or acetamiprid @ 0.2 g or acephate @1.5 g or phosalone @ 2 ml or phosphamidon @ 0.5 ml by mixing in 1 litre of water at 15 & 30 days.
		6. Reddening of leaves from margins	6. Do you notice any reddening of leaves from margins?	
		7. Brown necrotic patches on leaves	7. Do you observe any brown necrotic patches on leaves?	
	పచ్చ దోమ	1. మొక్కల పెరుగుదల ఆగింది	1. మొక్కల పెరుగుదల ఆగిందా?	(1)విత్తన శుద్ధి (acid delinting తరువాత) : Carbosulfan @ 40 g/kg విత్తనానికి లేదా Imidacloprid 70 WS @ 5 g/kg విత్తనానికి లేదా Thiamethoxam 70 WS@ 4 g/kg, రసం పీల్చే పురుగులనుండి 30 రోజుల పాటు రక్షణ కల్పిస్తుంది.(2) Soil application : విత్తనం నాటేటప్పుడూ ఒక ఎకరాకు 13 కిలోల Carbofuran 3G కణజాల ఉపయోగించండి. (3). Stem application : Monocrotophos, నీళ్లు @1:4 లేదా Methyl Demeton, నీళ్లు @1:4 లేదా Imidacloprid 200 SL, నీళ్లు @ 1: 20., కాండం యొక్క లేత ఆకుపచ్చ భాగంలో మృదువైన బ్రష్ ఉపయోగించి 20 - 25, 30 - 35, 40 - 45 రోజుల తర్వాత మూడు సార్లు
		2. ఆకులు ముడత పడుతున్నాయి	2. ఆకులు ముడత పడుతున్నాయా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకులపై మచ్చలు ఉన్నాయి	4. ఆకులపై మచ్చలు ఉన్నాయా?	
		5. ఆకులు పైకి ముడుత పడుతున్నాయి	5. ఆకులు పైకి ముడుత పడుతున్నాయా?	
		6. అంచుల నుండి ఆకుల ఎర్రబడు	6. అంచుల నుండి ఆకుల ఎర్రబడుతున్నాయా?	

		తున్నాయి		పూయండి.(4).విత్తన చికిత్స మరియు నేల శుద్ధి చేయకపోతే Monocrotophos @ 1.5 ml లేదా Imidacloprid @ 0.4 ml లేదా Methyldemeton @2 ml లేదా Acetamiprid @ 0.2 g లేదా Acephate @1.5 g లేదా Phosalone @ 2 ml లేదా Phosphamidon @ 0.5 ml, 1 లీటరు నీటిలో కలిపి 15 మరియు 30 రోజులకు పిచికారి చేయండి.
		7. ఆకులు పైన గజ్జిలాగా ఉదా రంగు మచ్చలు ఉన్నాయి	7. ఆకులు పైన గజ్జిలాగా ఉదా రంగు మచ్చలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
6	White flies	1. Stunted growth	1. Do you detect any stunted growth of a plant?	**Chemical:** (1).Stem application: If the crop is at 30-45 days stage, go for stem application of Monochrotophos + water @ 1:4 ratio using soft brush at tender green part of the stem. (2). If crop age crosses 60 days, go for foliar spraying with Monochrotophos @ 1.6 ml or Acetamiprid @ 0.2 g or Fepronil @ 2.0 ml or Imidacloprid @ 0.25 ml or Acephate 75% S.P @ 1.5 g or Thiamethoxam @ 0.2 g or Difenthiuron @ 1.25 or Triazophos @ 2.0 ml or Spyromesifen @1.0 ml by mixing in 1 litre of water.(3). If crop is at 45- 60 days stage apply Imidacloprid +water @ 1:20 using soft brush at tender green part of the stem.(*Do not use synthetic Pyrithroids or organo phosphorus chemicals.). **Biological :** Spray neem oil @ 5 ml by mixing in 1 litre of water or Prepare 5% Neem Cake Kernal Extract(Incubate10 kg Neem seed powder in10 litres for one day and filter the extract and makeup to 200 litres with water which is sufficient for one acre) and spray.
		2. Premature shedding of bolls	2. Do you find any Premature shedding of bolls?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Larvae/Insects on lower side of the leaf	4. Do you observe any larvae/insects on lower side of the leaves?	
		5. Irregular yellowing of leaves extends from veins to outer edges of leaves	5. Do you observe irregular yellowing of leaves which extends from veins to outer edges?	
		6. White colour flies on lower surface of the leaves	6. Do you observe any white colour flies on lower side of the leaves?	
	తెల్ల దోమ	1. మొక్కల పెరుగుదల ఆగింది	1. మొక్కల పెరుగుదల ఆగిందా?	రసాయన నిర్వహణ: (1).పంట 30 - 45 రోజుల దశలో ఉంటే, Monochrotophos + నీరు @ 1: 4 నిష్పత్తితో కలిపి, కాండం యొక్క లేత ఆకుపచ్చ భాగంలో మృదువైన బ్రష్ ఉపయోగించి పూయండి. (2).పంట 45 - 60 రోజుల దశలో ఉంటే, Imidacloprid + నీరు @ 1:20 నిష్పత్తితో కలిపి, కాండం యొక్క లేత ఆకుపచ్చ భాగంలో మృదువైన బ్రష్ ఉపయోగించి పూయండి. (3).పంట వయస్సు 60 రోజులు దాటితే, Monochrotophos @ 1.6 ml లేదా Acetamiprid @ 0.2 g లేదా Fepronil @ 2.0 ml లేదా Imidacloprid @ 0.25 ml లేదా Acephate 75% S.P @ 1.5 g లేదా Thiamethoxam @ 0.2 g లేదా difenthiuron @ 1.25 లేదా Triazophos @ 2.0 ml లేదా
		2. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయి	2. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయి	4. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయా?	
		5. ఆకులలో ఈనెల నుండి అంచుల వరకు పసుపు రంగు చారలు లేదా మచ్చలు ఉన్నాయి	5. ఆకులలో ఈనెల నుండి అంచుల వరకు పసుపు రంగు చారలు లేదా మచ్చలు ఉన్నాయా?	

		6. తెలుపు రంగు దోమలు ఆకుల దిగువ ఉపరితలంపై ఉన్నాయి	6. తెలుపు రంగు దోమలు ఆకుల దిగువ ఉపరితలంపై ఉన్నాయా?	Spyromesifen @1.0 ml , 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.(* కృత్రిమ Pyrithroids లేదా Organo Phosphorus రసాయనాలను వాడకండి.). **సేంద్రియ వ్యవహణ:** 1 లీటరు నీటిలో 5 ml వేప నూనె కలిపి పిచికారి చేయండి లేదా 5% నీమ్ కేక్ కెర్నల్ ఎక్స్ట్రాక్ట్ (10 లీటర్ల నీమ్ సీడ్ పౌడర్ 10 లీటర్ల నీటిలో కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.) తయారు చేసుకొని పిచికారీ చేయండి.

S.No	Disease	Symptoms	Symptom Question	Management
7	Aphids	1. Stunted growth	1. Do you detect any stunted growth of a plant?	**Chemical:** (1). Stem application: If the crop is at 30-45 days stage, go for stem application of Monochrotophos + water @ 1:4 ratio using soft brush at tender green part of the stem. (2).If crop is at 45- 60 days stage apply Imidacloprid +water @ 1:20 using soft brush at tender green part of the stem. (3). If crop age crosses 60 days, go for foliar spraying with acetamiprid @ 0.2 g or fepronil @ 2.0 ml or Imidacloprid @ 0.25 ml or Thiamethoxam @ 0.2 g by mixing in 1 litre of water. **Biological:** Spray neem oil @ 5 ml by mixing in 1 litre of water or Spray Prepare 5% Neem Cake Kernal Extract (Incubate10 kg Neem seed powder in10 litres for one day and filter the extract and makeup to 200 litres with water which is sufficient for one acre).
		2. Black Sooty mould on stem	2. Do you observe any black sooty mould on stem?	
		3. Coating on leaves	3. Do you notice any coating on leaves?	
		4. Curling of leaves	4. Do you observe any curling of leaves?	
		5. Ants on leaves	5. Do you notice any Ants on leaves?	
		6. Ants on stems and branches	6. Do you find anta on stems and branches?	
		7. Black Sooty mould coating on leaves	7. Do you notice any black sooty mould coating on the leaves?	
		8. Crinkling of leaves along with curling	8. Do you observe any Crinkling of leaves along with curling?	
	పేనుబంక	1. మొక్కల పెరుగుదల ఆగింది	1. మొక్కల పెరుగుదల ఆగిందా?	**రసాయన నిర్వహణ:** (1).పంట 30 - 45 రోజుల దశలో ఉంటే, Monochrotophos + నీరు @ 1: 4 నిష్పత్తితో కలిపి, కాండం యొక్క లేత ఆకుపచ్చ భాగంలో మృదువైన బ్రష్ ఉపయోగించి పూయండి. (2).పంట 45 - 60 రోజుల దశలో ఉంటే, Imidacloprid + నీరు @ 1:20 నిష్పత్తితో కలిపి, కాండం యొక్క లేత ఆకుపచ్చ భాగంలో మృదువైన బ్రష్ ఉపయోగించి పూయండి. (3).పంట వయస్సు 60 రోజులు దాటితే, Acetamiprid @ 0.2 g లేదా Fepronil @ 2.0 ml లేదా Imidacloprid @ 0.25 ml లేదా Thiamethoxam @ 0.2 g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** 1 లీటరు నీటిలో 5 ml వేప నూనె కలిపి పిచికారి చయండి లేదా 5% నీమ్ కేక్ కెర్నల్ ఎక్స్ట్రాక్ట్ (10 లీటర్ల నీమ్ సీడ్ పౌడర్ 10 లీటర్ల నీటిలో 10 లీటర్ల నీటిలో కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.)
		2. కాండముపై నల్లటి మసి పూత ఉన్నది	2. కాండముపై నల్లటి మసి పూత ఉన్నదా?	
		3. ఆకులపై రంగు పూత ఉన్నది	3. ఆకులపై రంగు పూత ఉన్నదా ?	
		4. ఆకులు ముడత పడుతున్నాయి	4. ఆకులు ముడత పడుతున్నాయా?	
		5. ఆకులపై చీమలున్నాయి	5. ఆకులపై చీమలున్నాయా?	
		6. కాండం మరియు కొమ్మలపై చీమలున్నాయి	6. కాండం మరియు కొమ్మలపై చీమలున్నాయా?	
		7. ఆకులపై నల్లటి మసి పూత ఉన్నది	7. ఆకులపై నల్లటి మసి పూత ఉన్నదా?	
		8. ఆకులు ముడతలు కలిగి చుట్టుకుంటున్నాయి	8. ఆకులు ముడతలు కలిగి చుట్టుకుంటున్నాయా ?	

				తయారు చేసుకొని పిచికారీ చేయండి. కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.) తయారు చేసుకొని పిచికారీ చేయండి.

S.No	Disease	Symptoms	Symptom Question	Management
8	Mites	1. Brittle leaves	1. Do you discover any brittle leaves?	**Chemical:** Foliar sprays with Sulphur 50% WP @ 3 g/ or Dicofol @ 5 ml by mixing in 1 litre of water.
		2. Premature shedding of bolls	2. Do you find any Premature shedding of bolls?	
		3. Premature opening of bolls	3. Do you detect bolls with premature opening?	
		4. Larvae/Insects on lower side of the leaf	4. Do you observe any larvae/insects on lower side of the leaves?	
		5. Curling of leaves	5. Do you observe any curling of leaves?	
		6. Leaves colour change	6. Do you notice any change in leaf colour?	
		7. Spots on leaves	7. Do you observe any spots on leaves?	
		8. Shedding of leaves	8. Do you observe any shedding of leaves?	
		9. Groups of small size larvae on lower side of the leaf	9. Do you observe any small size larve on lower side of the leaves?	
		10. Downward curling of leaves	10. Are leaves curling downwards?	
		11. Reddening of leaf from middle portion and drying	11. Do you notice reddening of leaf from middle portion and drying?	
		12. Yellowish spots on leaves	12. Do you find yellowish spots on leaves?	
	నల్లి	1. ఆకులు పెళుసుగా ఉన్నాయి	1. ఆకులు పెళుసుగా ఉన్నాయా?	**రసాయన నిర్వహణ:** Sulphur 50% WP @ 3 g/ లేదా Dicofol @ 5 ml,1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయి	2. కాయలు పూర్తిగా పెరగక ముందే రాలిపోతున్నాయా?	
		3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా?	
		4. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయి	4. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయా?	
		5. ఆకులు ముడత పడుతున్నాయి	5. ఆకులు ముడత పడుతున్నాయా?	
		6. ఆకుల రంగు మారుతున్నది	6. ఆకుల రంగు మారుతున్నదా?	

		7. ఆకులపై మచ్చలు ఉన్నాయి	7. ఆకులపై మచ్చలు ఉన్నాయా?	
		8. ఆకులు రాలుతున్నాయి	8. ఆకులు రాలుతున్నాయా?	
		9. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నది	9. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నదా?	
		10. ఆకులు కిందకు ముడుత పడుతున్నాయి	10. ఆకులు కిందకు ముడుత పడుతున్నాయా?	
		11. ఆకులు మధ్య భాగం నుండి ఎర్రబడి చనిపోతున్నాయి	11. ఆకులు మధ్య భాగం నుండి ఎర్రబడి చనిపోతున్నాయా?	
		12. ఆకులపై పసుపు రంగు మచ్చలు ఉన్నాయి	12. ఆకులపై పసుపు రంగు మచ్చలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
9	Thrips	1. Larvae/Insects on lower side of the leaf	1. Do you observe any larvae/insects on lower side of the leaves?	**Chemical:** Spray Monochrotophos @ 1.6 ml or Fepronil @ 2.0 ml or Imidacloprid @ 0.25 ml or Acephate 75% S.P @ 1.5 g or Difenthiuron @ 1.25 g by mixing in 1 litre of water. **Biological:** Spray neem oil @ 5 ml by mixing in 1 litre of water or Spray Prepare 5% Neem Cake Kernal Extract (Incubate10 kg Neem seed powder in10 litres for one day and filter the extract and makeup to 200 litres with water which is sufficient for one acre)
		2. Curling of leaves	2. Do you observe any curling of leaves?	
		3. Coating on leaves	3. Do you notice any coating on leaves?	
		4. Groups of small size larvae on lower side of the leaf	4. Do you observe any small size larvae on lower side of the leaves?	
		5. Upward curling of leaves	5. Are leaves curling upwards?	
		6. Shiny silver colour coating on the lower side of the leaves along veins	6. Do you notice any shiny colour coating on the lower side of the leaves along veins?	
	తామర పురుగు	1. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయి	1. ఆకు అడుగు భాగంలో పురుగులు ఉన్నాయా?	**రసాయన నిర్వహణ:** Monochrotophos @ 1.6 ml లేదా Fepronil @ 2.0 ml లేదా Imidacloprid @ 0.25 ml లేదా Acephate 75% S.P @ 1.5 g లేదా Difenthiuron @ 1.25 g 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** 1 లీటరు నీటిలో 5 ml వేప నూనె కలిపి పిచికారి చేయండి లేదా 5% నీమ్ కేక్ కెర్నల్ ఎక్స్ట్రాక్ట్ (10 లీటర్ల నీమ్ సీడ్ పౌడర్ 10 లీటర్ల నీటిలో కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.) తయారు చేసుకొని పిచికారీ చేయండి.
		2. ఆకులు ముడత పడుతున్నాయి	2. ఆకులు ముడత పడుతున్నాయా?	
		3. ఆకులపై రంగు పూత ఉన్నది	3. ఆకులపై రంగు పూత ఉన్నదా ?	
		4. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నది	4. ఆకు అడుగు భాగములో చిన్న పురుగుల సముదాయము ఉన్నదా?	
		5. ఆకులు పైకి ముడుత పడుతున్నాయి	5. ఆకులు పైకి ముడుత పడుతున్నాయా?	
		6. ఆకు సిరల వెంట దిగువ భాగంలో మెరిసే వెండి రంగు పూత ఉన్నది	6. ఆకు సిరల వెంట దిగువ భాగంలో మెరిసే వెండి రంగు పూత ఉన్నదా ?	

S.No	Disease	Symptoms	Symptom Question	Management
10	Red cotton bug	1. Premature opening of bolls	1. Do you detect bolls with premature opening?	**Chemical:** Dust Malathion powder around each plant.
		2. Spots on bolls	2. Do you find spots on the bolls?	
		3. Colour of bolls changing	3. Do you detect any change in colour of the bolls?	
		4. Lint colour change	4. Do you detect any change in lint colour?	
		5. Yellow or white spots on bolls	5. Do you find any Yellow or White spots on bolls	
		6. Young bolls turns into black colour	6. Do you observe any young bolls turning into black colour?	
		7. Lint turns into red colour	7. Do you observe lint turns into red colour?	
	మరక నల్లి	1. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	1. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా ?	**రసాయన నిర్వహణ:** ప్రతి మొక్క చుట్టూ Malathion పౌడర్ చల్లండి.
		2. కాయలపై మచ్చలున్నాయి	2. కాయలపై మచ్చలు ఉన్నాయా?	
		3. కాయల రంగు మారుతున్నది	3. కాయల రంగు మారుతున్నదా?	
		4. దూది రంగు మారింది	4. దూది రంగు మారిందా?	
		5. కాయలు పసుపు లేదా తెలుపు రంగులోకి మారుతున్నాయి	5. కాయలు పసుపు లేదా తెలుపు రంగులోకి మారుతున్నాయా?	
		6. కాయలు/పిందెలు నలుపు రంగులోకి మారుతున్నాయి	6. కాయలు/పిందెలు నలుపు రంగులోకి మారుతున్నాయా?	
		7. కాయ లోపలి పత్తి ఎరుపు రంగు లోకి మారుతున్నది	7. కాయ లోపలి పత్తి ఎరుపు రంగు లోకి మారుతున్నదా?	

S.No	Disease	Symptoms	Symptom Question	Management
11	Dusky cotton bug	1. Lint colour change	1. Do you detect any change in lint colour?	**Chemical:** Spray phosphamidon 100 EC @1.25 ml or Triazophos p @ 2.0 ml or Imidacloprid @0.25 ml or Profenofos @ 2.0 ml or Dimethoate @ 2.0 ml by mixing in 1 litre of water.
		2. Presence of larvae/insects on bolls	2. Do you find any larve/insects on bolls?	
		3. Black colour of lint	3. Do you notice black colour lint?	
		4. Greyish brown bug with hyaline wings and roughly rectangular in shape	4. Do you detect any bugs with greyish brown bug with hyaline wings and roughly rectangular in shape?	
	మురికి నల్లి	1. దూది రంగు మారింది	1. దూది రంగు మారిందా?	**రసాయన నిర్వహణ:** Phosphamidon 100 EC @1.25 ml లేదా Triazophos p @ 2.0 ml లేదా Imidacloprid @0.25 ml లేదా Profenofos @ 2.0 ml లేదా Dimethoate @ 2.0 ml 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		2. కాయలపై పురుగులున్నాయి	2. కాయలపై పురుగులున్నాయా?	
		3. కాయ లోపలి పత్తి నల్లగా మారుతున్నది	3. కాయ లోపలి పత్తి నల్లగా మారుతున్నదా?	
		4. తూనీగ లాంటి రెక్కలతో , ఇంచుమించు దీర్ఘచతురస్రాకార ఆకారము గల గోధుమ రంగు రెక్కల పురుగులు ఉన్నాయి	4. తూనీగ లాంటి రెక్కలు కలిగి, ఇంచుమించు దీర్ఘచతురస్రాకార ఆకారము గల గోధుమ రంగు రెక్కల పురుగులు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
12	Leaf roller	1. Rolling of leaves and holes	1. Do you observe any rolling of leaves and holes?	**Chemical:** (1). Collection and destruction of sheded plant parts. (2). Spray chlorpyriphos 20 EC @ 2.5 ml + Dichlorovas @ 1.0 ml by mixing in 1 litre of water or fenitrothion 50 EC @1.0 ml by mixing in 1 litre of water.
	ఆకు చుట్టు పురుగు	1. రంద్రలతో ఉన్నా ఆకులు ముడుచుకుపోతున్నాయి	1. రంద్రలతో ఉన్నా ఆకులు ముడుచుకుపోతున్నాయా?	**రసాయన నిర్వహణ:** (1).వాడిపోయిన మొక్క భాగాలను సకరించి నాశనం చేయండి. (2).Chlorpyriphos 20 EC @ 2.5 ml + Dichlorovas @ 1.0 ml 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి లేదా Fenitrothion 50 EC @1.0 ml, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.

S.No	Disease	Symptoms	Symptom Question	Management
13	Stem weevil	1. Bored stem	1. Do you observe any bores on stem?	**Chemical:** Spray Chlorpyriphos @ 2.5 ml by mixing in 1 litre of water.
		2. Swelling (Galls) of stem just above the ground	2. Do you find any Swelling(galls) on stem just above the ground?	
	కాండం తొలుచు ముక్కు పురుగు	1. కాండముపై రంధ్రాలు ఉన్నాయి	1. కాండముపై రంధ్రాలు ఉన్నాయా?	**రసాయన నిర్వహణ:** Chlorpyriphos @ 2.5 ml 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి
		2. చెట్టు మొదలు భాగంపై మొటిమలు ఉన్నాయి	2. చెట్టు మొదలు భాగంపై మొటిమలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
14	Mealy bugs	1. Powdery white scale appearance on whole plant	1. Do you observe powdery white scale appearance on whole plant?	**Chemical:** Spray Profenofos 50% EC @ 3.0 ml or Methyl Parathion @3.0 ml or Acephate @ 1.5 g + Triton or Sandovit gum @ 1 ml by mixing in 1 litre of water. **Biological:** Spray neem oil @ 5 ml by mixing in 1 litre of water or Spray Prepare 5% Neem Cake Kernal Extract (Incubate10 kg Neem seed powder in10 litres for one day and filter the extract and makeup to 200 litres with water which is sufficient for one acre).
		2. Wilting of plant	2. Do you find any wilting of a plant?	
	పిండినల్లి	1. కాండముపైన తెల్లటి జిగురు లాంటి పదార్థం ఉన్నది	1. కాండముపైన తెల్లటి జిగురు లాంటి పదార్థం ఉన్నదా?	**రసాయన నిర్వహణ:** Profenofos 50% EC @ 3.0 ml లేదా methyl Parathion @3.0 ml లేదా Acephate @ 1.5 g + Triton లేదా Sandovit Gum @ 1 ml, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** 1 లీటరు నీటిలో 5 ml వేప నూనె కలిపి పిచికారి చేయండి లేదా 5% నీమ్ కేక్ కెర్నల్ ఎక్స్ట్రాక్ట్ (10 లీటర్ల నీమ్ సీడ్ పౌడర్ 10 లీటర్ల నీటిలో కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.) తయారు చేసుకొని పిచికారీ చేయండి.
		2. మొక్కలు మాడిపోతున్నాయి	2. మొక్కలు మాడిపోతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
15	Rust	1. Shedding of leaves	1. Do you observe any shedding of leaves?	**Chemical:** Spray Calixine @ 1.0 ml or Wettable Sulfur @ 3.0 g by mixing in 1 litre of water.
		2. Spots on leaves	2. Do you observe any spots on leaves?	
		3. Spots on stems	3. Do you find any spots on stem?	
		4. Colour of bolls changing	4. Do you detect any change in colour of the bolls?	
		5. Bright yellow/orange/brown spots on lower surface of the older leaves surrounded by purple borders	5. Do you observe spots with bright yellow/orange/brown in colour usually on under surface of the older(lower) leaves surrounded by purple border?	
		6. Bright yellow/orange/brown spots usually on stems surrounded by purple borders	6. Do you observe spots with bright yellow/orange/brown in colour usually on stems surrounded by purple border?	
		7. Bright yellow/orange/brown bolls	7. Do you observe any spots with bright yellow/orange/brown coloured bolls?	
	తుప్పు తెగులు	1. ఆకులు రాలుతున్నాయి	1. ఆకులు రాలుతున్నాయా?	**రసాయన నిర్వహణ:** Calixine @ 1.0 ml లేదా Wettable Sulfur @ 3.0 g 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		2. ఆకులపై మచ్చలు ఉన్నాయి	2. ఆకులపై మచ్చలు ఉన్నాయా?	
		3. కాండముపై మచ్చలు ఉన్నాయి	3. కాండముపై మచ్చలు ఉన్నాయా?	
		4. కాయల రంగు మారుతున్నది	4. కాయల రంగు మారుతున్నదా?	
		5. ముదురు ఆకుల అడుగు భాగంలో ఉదా రంగు వలయంతో పసుపు లేదా నారింజ లేదా ముదురు గోధుమ రంగు మచ్చలు ఉన్నాయి	5. ముదురు ఆకుల అడుగు భాగంలో ఉదా రంగు వలయంతో పసుపు లేదా నారింజ లేదా ముదురు గోధుమ రంగు మచ్చలు ఉన్నాయా?	
		6. కాండముపై ఉదా రంగు వలయము గల పసుపు లేదా నారింజ రంగు మచ్చలు ఉన్నాయి	6. కాండముపై ఉదా రంగు వలయము గల పసుపు లేదా నారింజ రంగు మచ్చలు ఉన్నాయా?	
		7. కాయలు ముదురు పసుపు/ ముదురు నారింజ/ ముదురు గోధుమ రంగులోకి మారుతున్నాయి	7. కాయలు ముదురు పసుపు/ ముదురు నారింజ/ ముదురు గోధుమ రంగులోకి మారుతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
16	Helmenthosporium spots	1. Dried leaves	1. Do you notice dried leaves?	**Chemical:** Spray Mancozeb @ 2.5g or Copper Oxychloride @ 3.0g by mixing in 1 litre of water.
		2. Spots on leaves	2. Do you observe any spots on leaves?	
		3. The spots are greyish white in colour with deep purple margin	3. Do you observe greyish white colour spots with green purple margin?	
		4. Circular spots	4. Do you find circular spots on leaves?	
	హెల్మెంతోస్పోరియం ఆకు మచ్చ తెగులు	1. ఆకులు ఎండుతున్నాయి	1. ఆకులు ఎండుతున్నాయా?	**రసాయన నిర్వహణ:** Mancozeb @ 2.5g లేదా Copper Oxychloride @ 3.0g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		2. ఆకులపై మచ్చలు ఉన్నాయి	2. ఆకులపై మచ్చలు ఉన్నాయా?	
		3. ఆకులపై తెలుపు లేదా బూడిద రంగు మచ్చలు మరియు అంచుల వెంబడి వంకాయ రంగు వలయాలు ఉన్నాయి	3. ఆకులపై తెలుపు లేదా బూడిద రంగు మచ్చలు మరియు అంచుల వెంబడి వంకాయ రంగు వలయాలు ఉన్నాయా?	
		4. ఆకులపై గుండ్రటి మచ్చలు ఉన్నాయి	4. ఆకులపై గుండ్రటి మచ్చలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
17	Cercospora leaf spot	1. Shedding of leaves	1. Do you observe any shedding of leaves?	**Chemical:** Spray Mancozeb @ 2.5g or Copper Oxychloride@0.3% or Zineb 75% WP @2.5g or Copper Oxychloride @ 3.0 g or Carboxin 37.5% + Thiram 37.5% (Vitavax) @ 2.0 g by mixing in 1 litre of water 3- 4 times by changing the chemical at 15 days interval.
		2. Spots on leaves	2. Do you observe any spots on leaves?	
		3. Circular or irregular spots with greyish white centre surrounded by brown margin	3. Do you find lesions with circular or irregular spots with greyish white centre surrounded by brown margin?	
		4. Water soaked spots (lesions) appear on upper surface of leaf	4. Do you find water soaked lesions on upper surface of the leaf?	
	సెర్కొస్పారా ఆకుమచ్చ తెగులు	1. ఆకులు రాలుతున్నాయి	1. ఆకులు రాలుతున్నాయా?	**రసాయన నిర్వహణ:** Mancozeb @ 2.5g లేదా Copper Oxychloride@0.3% లేదా Zineb 75% WP @2.5g లేదా Copper Oxychloride @ 3.0 g లేదా Carboxin 37.5% + Thiram 37.5% (Vitavax) @ 2.0 g, 1 లీటరు నీటిలో కలిపి, 15 రోజుల వ్యవధిలో,రసాయనం
		2. ఆకులపై మచ్చలు ఉన్నాయి	2. ఆకులపై మచ్చలు ఉన్నాయా?	
		3. ఆకులపై వివిధ పరిమాణం గల గుండ్రటి బూడిద మరియు ఉదా రంగు గల మచ్చలు ఉన్నాయి	3. ఆకులపై వివిధ పరిమాణం గల గుండ్రటి బూడిద మరియు ఉదా రంగు గల మచ్చలు ఉన్నాయా?	

		4. ఆకుల పై భాగంలో నీటి చుక్కల లాంటి మచ్చలు ఉన్నాయి	4. ఆకుల పై భాగంలో నీటి చుక్కల లాంటి మచ్చలు ఉన్నాయా?	మారుస్తూ, 3-4 సార్లు పిచికారీ చేయండి.

S.No	Disease	Symptoms	Symptom Question	Management
18	Alternaria leaf spot	1. Coating on leaves	1. Do you notice any coating on leaves?	**Chemical:** Seed treatment with Pseudomonas Fluorescens @ 10 g/ kg seed. Spray Mancozeb@ 2.5g or Copper Oxychloride @ 3 g + Poshamycin or Plantomycin @ 1.0 g by mixing in 1 litre of water four to five sprays at 15 days interval
		2. Shedding of leaves	2. Do you observe any shedding of leaves?	
		3. Spots on leaves	3. Do you observe any spots on leaves?	
		4. Spots on stems	4. Do you find any spots on stem?	
		5. Spots on bolls	5. Do you find spots on the bolls?	
		6. Black Sooty mould coating on leaves	6. Do you notice any black sooty mould coating on the leaves?	
		7. Small brown, round spots surrounded by a purple margin appear on leaves	7. Do you find any small brown, round spots surrounded by a purple margin?	
		8. On older leaves, the necrotic centre of the spots may be marked by a pattern of concentric zonation	8. Do you find any older leaves with the necrotic centre of the spots (marked by a pattern of concentric zonation)?	
		9. Spherical or elliptical purple spots on stems	9. Do you observe spherical or elliptical spots with purple colour on stems?	
		10. Spherical or elliptical purple spots on bolls	10. Do you observe any spherical or elliptical spots with purple colour on bolls?	
	ఆల్టెర్నేరియా ఆకుమచ్చ తెగులు	1. ఆకులపై రంగు పూత ఉన్నది	1. ఆకులపై రంగు పూత ఉన్నదా ?	**రసాయన నిర్వహణ:** 1 కిలో విత్తనాలలో 10g Pseudomonas Fluorescens కలిపి విత్తన శుద్ధి చెయాలి. Mancozeb@ 2.5g లేదా Copper Oxychloride @ 3 g + Poshamycin లేదా Plantomycin @ 1.0 g, 1 లీటరు నీటిలో
		2. ఆకులు రాలుతున్నాయి	2. ఆకులు రాలుతున్నాయా?	
		3. ఆకులపై మచ్చలు ఉన్నాయి	3. ఆకులపై మచ్చలు ఉన్నాయా?	
		4. కాండముపై మచ్చలు ఉన్నాయి	4. కాండముపై మచ్చలు ఉన్నాయా?	
		5. కాయలపై మచ్చలున్నాయి	5. కాయలపై మచ్చలు ఉన్నాయా?	

		6. ఆకులపై నల్లటి మసి పూత ఉన్నది	6. ఆకులపై నల్లటి మసి పూత ఉన్నదా?	కలిపి, 15 రోజుల వ్యవధిలో 4 సార్లు పిచికారీ చేయండి.
		7. ఆకులపై ఉదా రంగు వలయము గల చిన్నచిన్న ముదురు గోధుమ రంగు మచ్చలు ఉన్నాయి	7. ఆకులపై ఉదా రంగు వలయము గల చిన్నచిన్న ముదురు గోధుమ రంగు మచ్చలు ఉన్నాయా?	
		8. పండుటాకులలో గజ్జిలాగా ఉదా రంగు మచ్చలు ఏకీకృత కేంద్రం గల గుండ్రటి వృత్తాల లాగా మచ్చల సముదాయాలు ఉన్నాయి	8. పండుటాకులలో గజ్జిలాగా ఉదా రంగు మచ్చలు ఏకీకృత కేంద్రం గల గుండ్రటి వృత్తాల లాగా మచ్చల సముదాయాలు ఉన్నాయా?	
		9. కాండముపై గుండ్రటి లేదా దీర్ఘ వృత్తకారం గల మచ్చలు ఉన్నాయి	9. కాండముపై గుండ్రటి లేదా దీర్ఘ వృత్తకారం గల మచ్చలు ఉన్నాయా?	
		10. కాయలకు గుండ్రటి లేదా దీర్ఘ వృత్తకారంలో ఉదా రంగు మచ్చలు ఉన్నాయి	10. కాయలకు గుండ్రటి లేదా దీర్ఘ వృత్తకారంలో ఉదా రంగు మచ్చలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
19	Anthracnose	1. Splitting of stem and Shredding of bark	1. Do you discover any splitting of stem and shredding of bark?	**Chemical:** (1). Treat the delinted seeds with Carbendazim or Carboxin @ 2g/kg seed or Thiram or Captan at 4g/kg seed. (2). Remove and burn the infected plant debris and bolls in the soil. (3). Rogue out the reservoir weed hosts.(4).Spray the crop at boll formation stage with Mancozeb@ 2.5g or Copper Oxychloride@ 3.0g or Ziram2.5g or Carbendazim@ 1.0g by mixing in 1 litre of water.
		2. Lint colour change	2. Do you detect any change in Lint colour?	
		3. Premature opening of bolls	3. Do you detect bolls with premature opening?	
		4. Spots on leaves	4. Do you observe any spots on leaves?	
		5. Spots on bolls	5. Do you find spots on the bolls?	
		6. Black colour of lint	6. Do you notice black colour lint?	
		7. Reddish circular spots on leaves	7. Do you observe circular reddish colour spots on leaves?	
		8. Small water soaked, circular, reddish brown depressed spots appear on the bolls	8. Do you observe any circular spots on bolls with reddish brown in colour, small water soaked and depressed?	
		9. Yellow or brown colour of lint	9. Do you observe yellow or brown colour lint?	
	ఆంత్రాక్నోస్	1. కాండంపై ఉన్న బెరడు ముక్కలు రాలిపోతున్నాయి మరియు కాండము పగులుతున్నది.	1. కాండంపై ఉన్న బెరడు ముక్కలు రాలిపోయి మరియు కాండము పగులుతున్నదా ?	**రసాయన నిర్వహణ:** (1).Delinted విత్తనాలను, Carbendazim లేదా Carboxin @ 2g/kg విత్తనాకు లేదా Thiram లేదా Captan @ 4g/kg విత్తనాకు, కలిపి శుద్ధి చేయండి. (2).వ్యాధి సోకినా మొక్కలను, కాయలను ఒక్కదగ్గరికి సేరించి తగలబెట్టండి. (3). రిజర్వాయర్ కలుపు మొక్కల సముదాయాన్ని తొలగించండి. (4) కాయ / పిందే ఏర్పడే సమయంలో
		2. దూది రంగు మారింది	2. దూది రంగు మారిందా?	
		3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	3. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా?	
		4. ఆకులపై మచ్చలు ఉన్నాయి	4. ఆకులపై మచ్చలు ఉన్నాయా?	
		5. కాయలపై మచ్చలున్నాయి	5. కాయలపై మచ్చలు ఉన్నాయా?	
		6. కాయ లోపలి పత్తి నల్లగా మారుతున్నది	6. కాయ లోపలి పత్తి నల్లగా మారుతున్నదా?	
		7. ఆకులపై గుండ్రని ఎరుపు రంగు మచ్చలు ఉన్నాయి	7. ఆకులపై గుండ్రని ఎరుపు రంగు మచ్చలు ఉన్నాయా?	

		8. కాయలుపై ముదురు ఎరుపు లేదా ఉదా రంగు గల చిన్న గుంతల లాగ మచ్చలు ఉన్నాయి	8. కాయలుపై ముదురు ఎరుపు లేదా ఉదా రంగు గల చిన్న గుంతల లాగ మచ్చలు ఉన్నాయా?	Mancozeb@ 2.5g లేదా Copper Oxychloride@ 3.0g లేదా Ziram2.5g లేదా Carbendazim@ 1.0g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		9. కాయ లోపలి పత్తి పసుపు లేదా ముదురు గోధుమ రంగు లోకి మారుతున్నది	9. కాయ లోపలి పత్తి పసుపు లేదా ముదురు గోధుమ రంగు లోకి మారుతున్నదా?	

S.No	Disease	Symptoms	Symptom Question	Management
20	Grey or aerolate mildew	1. Shedding of leaves	1. Do you observe any shedding of leaves?	**Chemical:** Spray the crop with Carbendazim @1.0g or Bordeaux mixture @1.0 g or Wettable Sulphur @ 3.0 g by mixing in 1 litre of water at two times at weekly interval
		2. Spots on leaves	2. Do you observe any spots on leaves?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Frosty irregular shape powdery spots on the lower surface of leaf	4. Do you find any Frosty irregular shape powdery spots on the lower surface of leaf?	
		5. Leaf turns into yellow colour	5. Are leaves turn into yellow colour?	
	బూజు తెగులు	1. ఆకులు రాలుతున్నాయి	1. ఆకులు రాలుతున్నాయా?	**రసాయన నిర్వహణ:** Carbendazim @1.0g లేదా Bordeaux mixture @1.0 g లేదా Wettable Sulphur @ 3.0 g, 1 లీటరు నీటిలో కలిపి 7 రోజులులకు ఒక సారి, 2 మార్లు పిచికారీ చేయండి.
		2. ఆకులపై మచ్చలు ఉన్నాయి	2. ఆకులపై మచ్చలు ఉన్నాయా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకు క్రింది భాగంలో నీటి బిందువులతో కూడిన అడ్డదిడ్డమైన బూడిద మచ్చలు ఉన్నాయి	4. ఆకు క్రింది భాగంలో నీటి బిందువులతో కూడిన అడ్డదిడ్డమైన బూడిద మచ్చలు ఉన్నాయా?	
		5. ఆకులు పసుపు రంగులోకి మారుతున్నాయి	5. ఆకులు పసుపు రంగులోకి మారుతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
21	Root rot	1. Dried leaves	1. Do you notice Dried leaves?	**Chemical:** (1). Seed treatment with Carbendazim @2.0g / kg seed or Trychoderma Viridi @ 4 g / kg seed. (2).Spray carbendazim @2.0g or copper oxychloride @ 3.0g by mixing in 1 litre of water
		2. Shedding of leaves	2. Do you observe any shedding of leaves?	
		3. Easily uprootable from the soil	3. Do you find plant is easily uprootable from soil?	
		4. Wilting of plants in concentric circular patches in the crop	4. Do you find wilting of plants in concentric circular patches in the crop?	
		5. Rotted root or few roots	5. Do you find rotted root or few roots?	
	వేరు కుల్లు తెగులు	1. ఆకులు ఎండుతున్నాయి	1. ఆకులు ఎండుతున్నాయా?	**రసాయన నిర్వహణ:** (1). 1 కిలో విత్తనాలకు Carbendazim @2.0g లేదా Trychoderma Viridi @ 4 g కలిపి శుద్ధి చేయండి.(2).Carbendazim @2.0g లేదా Copper Oxychloride @ 3.0g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.
		2. ఆకులు రాలుతున్నాయి	2. ఆకులు రాలుతున్నాయా?	
		3. నేల నుండి సులభంగా వేరుచేయవచ్చు	3. నేల నుండి సులభంగా చేయవచ్చా?	
		4. పంటలో గుంపులు గుంపులుగా మొక్కలు ఎండిపోతున్నాయి	4. పంటలో గుంపులు గుంపులుగా మొక్కలు ఎండిపోతున్నాయా?	
		5. కుళ్ళిన వేర్లు ఉన్నాయి	5. కుళ్ళిన వేర్లు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
22	Verticelliu m wilt	1. Dried leaves	1. Do you notice Dried leaves?	**Chemical:** 1. Seed treatment with Carbendazim @2.0g / kg seed or Trychoderma Viridi @ 4 g / kg seed.(2).Apply heavy doses of farm yard manure or Compost at 4 t/acre. (3).Follow crop rotation by growing Paddy or Lucerne or Chrysanthemum for 2-3 years. (4).Spot drench with 0.05 per cent Benomyl or Carbendazim.
		2. Pinkish to pinkish brown discolouration of stem	2. Do you observe discolouration of stem from pinkish to pinkish brown in colour?	
		3. Shedding of leaves	3. Do you observe any shedding of leaves?	
	ఎండు తెగులు	1. ఆకులు ఎండుతున్నాయి	1. ఆకులు ఎండుతున్నాయా?	**రసాయన నిర్వహణ:** (1).1 కిలో విత్తనాలకు Carbendazim @2.0g లేదా Trychoderma Viridi @ 4 g కలిపి శుద్ధి చేయండి.(2).Carbendazim @2.0g లేదా Copper Oxychloride @ 3.0g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.(3).2-3 సంవత్సరాలవరకు పంట మార్పిడి పద్ధతిలో వరి లేదా Lucerne లేదా Chrysanthemum పండించండి. (4). Spot drench with 0.05 per cent Benomyl or Carbendazim.
		2. కాండము చీల్చి చుస్తే గులాబీ నుండి ముదురు గులాబీ గోధుమ లేదా నలుపు రంగులో మారియున్న లోపలి కణజాలం ఉన్నది	2. కాండము చీల్చి చుస్తే గులాబీ నుండి ముదురు గులాబీ గోధుమ లేదా నలుపు రంగులో మారియున్న లోపలి కణజాలం ఉన్నదా ?	
		3. ఆకులు రాలుతున్నాయి	3. ఆకులు రాలుతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
23	Fusarium wilt	1. Stunted growth	1. Do you detect any stunted growth of a plant?	**Chemical:** (1). Crop rotation with non-host species to reduce presence of Fusarium Wilt in the soil. Long rotation periods (several years) are more effective than short rotations. (2). Resistant cotton varieties are the main method of managing Fusarium Wilt in cotton crops.
		2. Leaves colour change	2. Do you notice any change in leaf colour?	
		3. Browning or blackening of vascular tissues ie black streaks or stripes may be seen extending upwards to the branches	3. Do you find browning or blackening of vascular tissues of stem (black streaks or stripes may be seen extending upwards to the branches)?	
		4. Yellowing of edges of leaves ie discolouration starts from the margin and spreads towards the midrib Finally leaf turns brown	4. Are leaves edges turn into yellow (discolouration starts from the margin and spreads towards the midrib. Finally leaf turns brown)?	
		5. Shedding of leaves	5. Do you observe any shedding of leaves?	
	ఫ్యుజేరియం వడలు తెగులు	1. మొక్కల పెరుగుదల ఆగింది	1. మొక్కల పెరుగుదల ఆగిందా?	**రసాయన నిర్వహణ:** (1).మట్టిలో ఏండు తెగులు తగ్గడానికి నాన్-హోస్ట్ జాతుల పంటలతో పంట మార్పిడి చెయాలి.దీర్ఘకాలం(చాలా సంవత్సరాలు) పంట మార్పిడి చేస్తే మంచిది. (2).రోగ నిరోధక శక్తి గల పత్తి విత్తనాలను ఎంచుకోవడం వలన ప్రధానంగా ఏండు తెగులును అరికట్టవచ్చు.
		2. ఆకుల రంగు మారుతున్నది	2. ఆకుల రంగు మారుతున్నదా?	
		3. కాండము చీల్చి చుస్తే ముదురు గోధుమ లేదా నలుపు రంగులో మారియున్న లోపలి కణజాలం ఉన్నది	3. కాండము చీల్చి చుస్తే ముదురు గోధుమ లేదా నలుపు రంగులో మారియున్న లోపలి కణజాలం ఉన్నదా ?	
		4. ఆకులు కోణలు పసుపు రంగులో మారి క్రమముగా మధ్యభాగం వరకు వ్యాపించి ఉదా రంగులోకి మారుతున్నాయి	4. ఆకులు కోణలు పసుపు రంగులో మారి క్రమముగా మధ్యభాగం వరకు వ్యాపించి ఉదా రంగులోకి మారుతున్నాయా?	
		5. ఆకులు రాలుతున్నాయి	5. ఆకులు రాలుతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
24	Bacterial blight or angular leaf spot or black arm	1. Shedding of leaves	1. Do you observe any shedding of leaves?	**Chemical:** Spray quinolphos @ 2.0 ml or chlorpyrifos @ 2.5 ml or acephate @ 1.5 g or triazophos @ 2.0 ml or thiodicarb @ 1.5 g by mixing in 1 litre of water. **Biological:** Spray neem oil @ 5 ml by mixing in 1 litre of water or Spray Prepare 5% Neem Cake Kernal Extract (Incubate10 kg Neem seed powder in10 litres for one day and filter the extract and makeup to 200 litres with water which is sufficient for one acre).
		2. Leaves colour change	2. Do you notice any change in leaf colour?	
		3. Splitting of stem and Shredding of bark	3. Do you discover any splitting of stem and shredding of bark?	
		4. oozing of gum over the stems	4. Do you detect oozing of gum over the stems?	
		5. Lint colour change	5. Do you detect any change in lint colour?	
		6. Premature opening of bolls	6. Do you detect bolls with premature opening?	
		7. Premature shedding of bolls	7. Do you find any premature shedding of bolls?	
		8. Spots on leaves	8. Do you observe any spots on leaves?	
		9. Spots on bolls	9. Do you find spots on the bolls?	
		10. Blackening of veins in leaves	10. Do you observe any leaves with blackened veins?	
		11. Black colour of lint	11. Do you notice black colour lint?	
		12. Reddish brown colour angular spots on both sides of leaves	12. Do you find any angular spots with reddish brown in colour on both sides of leaves?	
		13. Dark black irregular spots	13. Do you find any irregular spots with dark black in colour?	
	బాక్టీరియా నల్ల మచ్చ తెగులు	1. ఆకులు రాలుతున్నాయి	1. ఆకులు రాలుతున్నాయా?	**రసాయన నిర్వహణ:** Quinolphos @ 2.0 ml లేదా Chlorpyrifos @ 2.5 ml లేదా Acephate @ 1.5 g లేదా Triazophos @ 2.0 ml లేదా Thiodicarb @ 1.5 g, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి. **సేంద్రియ నిర్వహణ:** 1 లీటరు నీటిలో 5 ml వేప
		2. ఆకుల రంగు మారుతున్నది	2. ఆకుల రంగు మారుతున్నదా?	
		3. కాండంపై ఉన్న బెరడు ముక్కలు రాలిపోతున్నాయి మరియు కాండము పగులుతున్నది.	3. కాండంపై ఉన్న బెరడు ముక్కలు రాలిపోయి మరియు కాండము పగులుతున్నదా ?	
		4. కాండముపై బంక కారుతున్నది	4. కాండముపై బంక కారుతున్నదా?	
		5. దూది రంగు మారింది	5. దూది రంగు మారిందా?	
		6. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయి	6. కాయలు పూర్తిగా పెరగక ముందే పగులుతున్నాయా?	

		7. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయి	7. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయా?	నూనె కలిపి పిచికారి చేయండి లేదా 5% నీమ్ కేక్ కెర్నల్ ఎక్స్ట్రాక్ట్ (10 లీటర్ల నీమ్ సీడ్ పౌడర్ 10 లీటర్ల నీటిలో కలిపి 1 రోజు నానపెట్టాలి. దానినుండి తీసిన సారం ఒక ఎకరానికి సరిపడిన 200 లీటర్ నీటిలో కలపాలి.) తయారు చేసుకొని పిచికారీ చేయండి.
		8. ఆకులపై మచ్చలు ఉన్నాయి	8. ఆకులపై మచ్చలు ఉన్నాయా?	
		9. కాయలపై మచ్చలున్నాయి	9. కాయలపై మచ్చలు ఉన్నాయా?	
		10. ఆకు యొక్క సిరలు నల్లరంగులోకి మారుతున్నాయి	10. ఆకు యొక్క సిరలు నల్లరంగులోకి మారుతున్నాయా?	
		11. కాయ లోపలి పత్తి నల్లగా మారుతున్నది	11. కాయ లోపలి పత్తి నల్లగా మారుతున్నదా?	
		12. ఆకుల ఇరువైపులా ఈనెల మధ్యభాగంలో ముదురు గోధుమ రంగు గల మచ్చలు ఉన్నాయి	12. ఆకుల ఇరువైపులా ఈనెల మధ్యభాగంలో ముదురు గోధుమ రంగు గల మచ్చలు ఉన్నాయా?	
		13. కాయలపై నల్లటి మచ్చలు అడ్డదిడ్డంగా ఉన్నాయి	13. కాయలపై నల్లటి మచ్చలు అడ్డదిడ్డంగా ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
25	Nitrogen deficiency	1. Stunted growth	1. Do you detect any stunted growth of a plant?	**Chemical:** (1). Foliar spray with 10 to 20 g Urea in 1 litre water two to three times.(2). For varieties apply 50 per cent of Nitrogen and Potassium full dose of P2O5 as basal and remaining ½ Nitrogen and Potassium at 40 – 45 days after sowing. (3).For hybrids apply Nitrogen in three splits viz., basal, 45 and 65 days.
		2. Leaves colour change	2. Do you notice any change in leaf colour?	
		3. A pale yellowish green colour of leaf(older)	3. Are older leaves pale yellowish green in colour?	
		4. Reduction in leaf size	4. Do you notice any reduction in size of leaves?	
	నత్రజని లోపం	1. మొక్కల పెరుగుదల ఆగింది	1. మొక్కల పెరుగుదల ఆగిందా?	**రసాయన నిర్వహణ:** (1).10-20g యూరియా, 1 లీటర్ నీటిలో కలిపి 2-3 సార్లు పిచికారీ చేయండి. (2).దేశీయ రకాల పత్తి పంటకు 50% నత్రజని మరియు పొటాషియం + పూర్తిస్థాయి మోతాదులో P2O5 బేసల్ మరియు మిగిలిన 1/2 నైట్రోజెన్ మరియు పొటాషియంతో, విత్తనాలు నాటిన 40 - 45 రోజుల వ్యవధిలో వాడండి.
		2. ఆకుల రంగు మారుతున్నది	2. ఆకుల రంగు మారుతున్నదా?	
		3. ముదురు ఆకులు పసుపు రంగులోకి మారుతున్నాయి	3. ముదురు ఆకులు పసుపు రంగులోకి మారుతున్నాయా?	
		4. ఆకులు చిన్నవిగా ఉన్నాయి	4. ఆకులు చిన్నవిగా ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
26	Phosphorus deficiency	1. Delay in blooming	1. Is there a delay in blooming?	**Chemical:** (1).If the crop has Phosphorus deficiency spray with 10-20 g DAP in 1 litre water. (2).If the soil has Phosporous deficiency, apply 25% excess Phosphorus fertilizers than specified as basal or 15 days after sowing.
		2. Delay in maturity of the bolls	2. Do you find any delay in maturity of the bolls?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Leaf is dark green in colour	4. Are leaves dark green in colour?	
		5. Reduction in leaf size	5. Do you notice any reduction in size of leaves?	
	భాస్వరం లోపం	1. పూత రావటం ఆలస్యమౌతున్నాది	1. పూత రావటం ఆలస్యమౌతు ఉన్నదా?	**రసాయన నిర్వహణ:** (1).మొక్కలో Phosphorus లోపం ఉంటే 10-20g DAP, 1 లీటరు నీటిలో కలిపి పిచికారి చేయండి.(2).మట్టిలో Phosporous లోపం ఉంటే అవసమైనదానికంటే 25% ఎక్కువ Phosphorus ఎరువును విత్తనాలు నాటిన 15 రోజులతరువాత వాడండి.
		2. కాయలు పూర్తిగా అభివృద్ధి చెందడం ఆలస్యం అవుతున్నది	2. కాయలు పూర్తిగా అభివృద్ధి చెందడం ఆలస్యం అవుతున్నదా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకులు ముదురు ఆకుపచ్చ రంగు లోకి మారుతున్నాయి	4. ఆకులు ముదురు ఆకుపచ్చ రంగు లోకి మారుతున్నాయా?	
		5. ఆకులు చిన్నవిగా ఉన్నాయి	5. ఆకులు చిన్నవిగా ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
27	Potassium deficiency	1. Bolls failed to open	1. Are bolls not opening?	**Chemical:** (1).Spay with 10 g of Potassium Nitrate by mixing in 1 litre Water two to three times at weekly interval. (2).Drain out stagnated water from the field.
		2. Retarded growth of bolls	2. Is growth of bolls retarted?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Spots on leaves	4. Do you observe any spots on leaves?	
		5. The whole leaf become reddish brown	5. Do you observe whole leaves turn into reddish brown in colour?	
		6. Light yellowish green spots appear between the veins	6. Do you detect light yellowish green spots appear between the veins?	
	పొటాషియం లోపం	1. కాయలు పగలటం లేదు	1. కాయలు పగలటం లేదా?	**రసాయన నిర్వహణ:** (1).10 g Potassium Nitrate, 1 లీటర్ నీటిలో కలిపి, 7 రోజులకు ఒకసారి, 2-3 సార్లు పిచికారీ చేయండి. (2).పొలంలో నిలిచినా నీళ్లను తొలగించండి.
		2. కాయలు వృద్ధి చెందడం ఆగిపోయింది	2. కాయలు వృద్ధి చెందడం ఆగిపోయిందా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకులపై మచ్చలు ఉన్నాయి	4. ఆకులపై మచ్చలు ఉన్నాయా?	
		5. ఆకు మొత్తం ఎర్రటి గోధుమ రంగులోకి మారుతుంది	5. ఆకు మొత్తం ఎర్రటి గోధుమ రంగులోకి మారుతున్నదా?	
		6. ఆకుల ఈనెల మధ్యలో పసుపు లేదా పచ్చ రంగు మచ్చలు ఉన్నాయి	6. ఆకుల ఈనెల మధ్యలో పసుపు లేదా పచ్చ రంగు మచ్చలు ఉన్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
28	Magnesium deficiency	1. Shedding of leaves	1. Do you observe any shedding of leaves?	**Chemical:** (1). Spray 10g MgSO4 / litre water on 50th and 80th day. (2).If soil deficient of Mg, apply MgSO4 @ 8 kg/ha as basal before sowing.
		2. Leaves colour change	2. Do you notice any change in leaf colour?	
		3. Purplish red leaves with green veins (older leaves)	3. Are older leaves turns into purplish red in colour with green veins?	
		4. Drying of terminal buds	4. Do you detect any drying of terminal buds?	
	మెగ్నీషియం లోపం	1. ఆకులు రాలుతున్నాయి	1. ఆకులు రాలుతున్నాయా?	**రసాయన నిర్వహణ:** (1).50వ రోజు మరియు 80వ రోజు 10g MgSO4, 1 లీటరు నీటిలో కలిపి పిచికారీ చేయండి.(2).మట్టిలో Magnesium లోపం ఉంటే, విత్తనాలు నాటేముంది MgSO4 @ 8 kg/ha వాడండి.
		2. ఆకుల రంగు మారుతున్నది	2. ఆకుల రంగు మారుతున్నదా?	
		3. ఈనెలు మినహా ముదురు ఆకులు, ఎరుపు రంగులోకి మారుతున్నాయి	3. ఈనెలు మినహా ముదురు ఆకులు, ఎరుపు రంగులోకి మారుతున్నాయా?	
		4. చిగురు మొగ్గలు ఎండిపోతున్నాయి	4. చిగురు మొగ్గలు ఎండిపోతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
29	Boron deficiency	1. Drying of terminal buds	1. Do you detect any drying/ falling of terminal buds?	**Chemical:** Spray Borox @ 1-1.5 g / litre water twice at weekly interval between 60-90 days after sowing.
		2. Bushy appearance of plant	2. Do you observe any bushy appearance of a plant?	
		3. Premature shedding of bolls	3. Do you find any Premature shedding of bolls?	
		4. Shedding of flowers	4. Do you find any premature shedding of flowers?	
	బోరాన్ లోపం	1. చిగురు మొగ్గలు ఎండిపోతున్నాయి	1. చిగురు మొగ్గలు ఎండిపోతున్నాయా?	**రసాయన నిర్వహణ:** Borox @ 1-1.5 g, 1 లీటర్ నీటిలో కలిపి, 60-90 రోజుల పంట సమయంలో, 7 రోజుల వ్యవధితో, రెండు మార్లు పిచికారీ చేయండి.
		2. మొక్కలు గిడసబారినట్లు చిన్న పొదలుగా ఉన్నాయి	2. మొక్కలు గిడసబారినట్లు చిన్న పొదలుగా ఉన్నాయా?	
		3. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయి	3. కాయలు పూర్తిగా పెరగకముందే రాలిపోతున్నాయా?	
		4. పువ్వులు రాలుతున్నాయి	4. పువ్వులు రాలుతున్నాయా?	

S.No	Disease	Symptoms	Symptom Question	Management
30	Zinc deficiency	1. Brittle leaves	1. Do you discover any brittle leaves?	**Chemical:** Apply ZnSO4 @ 20 kg/acre as basal or spray ZnSO4 @ 1.0 g / litre of water at 45, 60 and 75 days after sowing.
		2. Shedding of flowers	2. Do you find any premature shedding of flowers?	
		3. Leaves colour change	3. Do you notice any change in leaf colour?	
		4. Bronzing and interveinal chlorosis appearance	4. Do you observe any appearance of bronzing and interveinal chlorosis?	
	జింక్ లోపం	1. ఆకులు పెళుసుగా ఉన్నాయి	1. ఆకులు పెళుసుగా ఉన్నాయా?	**రసాయన నిర్వహణ:** ZnSO4 @ 20 kg/acre ఏరువు వాడండి. లేదా ZnSO4 @ 1.0 g, 1 లీటర్ నీటిలో కలిపి, విత్తనాలు నాటిన 45, 60 మరియు 75 రోజులకు పిచికారీ చేయండి.
		2. పువ్వులు రాలుతున్నాయి	2. పువ్వులు రాలుతున్నాయా?	
		3. ఆకుల రంగు మారుతున్నది	3. ఆకుల రంగు మారుతున్నదా?	
		4. ఆకుల యొక్క ఈనెలల మధ్యభాగంలో పసుపు రంగు ఉన్నది	4. ఆకుల యొక్క ఈనెలల మధ్యభాగంలో పసుపు రంగు ఉన్నదా?	

Appendix C

Part 1: Sample questionnaire filled by a domain expert

www.cropdarpan.in

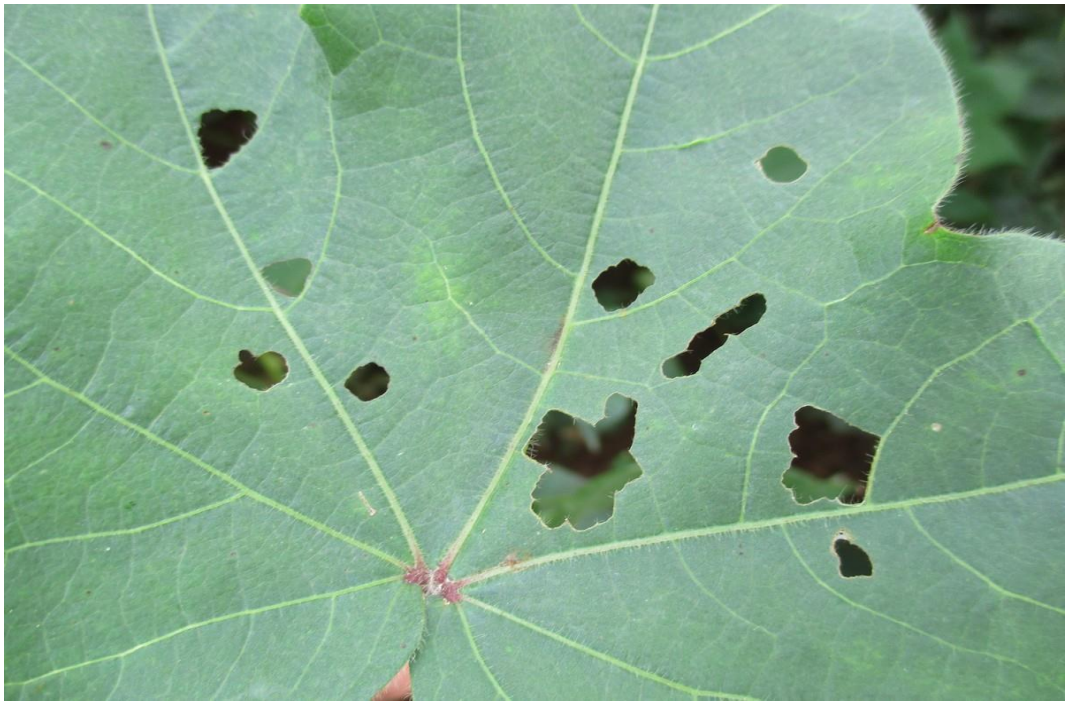
Expert Evaluation Form Questionnaire

1. Pls. try to detect four diseases using the Crop Darpan system using multiple paths. Out of four crop diseases, how many of the crop diseases are correctly identified? (a) 1 (b) 2 (c) 3 (d) 4	If you are unable to identify any disease, provide the reason:
2. After providing training, do you agree that proposed system will help the educated farmer to identify the crop problem using the Smartphone through Internet? (a) It will not help the educated farmer. (b) It will help the educated farmer to some extent. (c) It will definitely help the educated farmer to identify the crop problem.	Provide your comment, if any: Training programmes should be conducted to train uneducated faimers also as it is also available in Telegu.
3. The proposed system extends generalization / specialization framework (starting with generic symptoms to specialized symptoms) to build the crop diagnostic system to help the farmer. What is your view on the methodology? (a) The methodology does not make sense. (b) The methodology seems logical and it is as per the common sense. (c) Unable to judge.	Provide your comment, if any: Very good development. It should be developed in many crops as soon as possible.
4. You have used the system to identify few crop problems. How difficult is it for you to use? (a) Difficult to use. (b) Moderately difficult to use. (c) Easy to use.	Provide your comment, if any: Along with question, if pictures of symptoms are provided, it would help in correct diagnosis by faimers.
5. We have employed subject matter specialists to identify the crop disease symptom data. As a subject matter specialist, what is your opinion on the data? (a) The data is inaccurate. (b) The data is accurate. (c) The data is satisfactory and could be refined.	Provide your comment, if any:
6. The system is intended to resemble the real world diagnostic procedure followed by a typical agricultural scientist to identify the crop problem if he/she visits the field. What is your view on methodology? (a) Typical agricultural scientist follows a different methodology to identify the crop problem. (b) To some extent, the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem. (c) In general (overall), the system is capturing the methodology followed by typical agricultural scientist to identify the crop problem.	Provide your comment, if any: Try to provide trade names of chemicals along with the chemical names so that the faimer can easily get it.
7. Do you think that similar systems can be developed for other crops? (a) No, it is difficult. (b) Yes, definitely	Provide your comment, if any:
8. Do you think this kind of tool is needed in agriculture? (a) No (b) Yes	Provide your comment, if any: Very good tool & very helpful

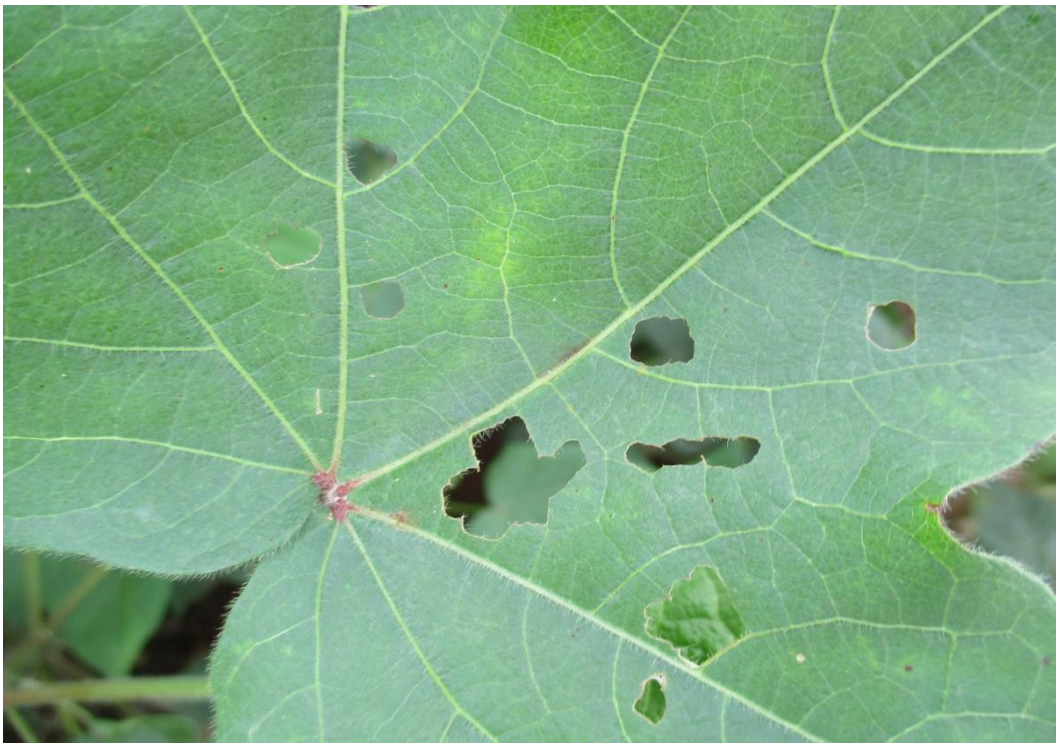
Part 2: Sample cotton crop data record created by educated farmer

Photograph 1

Photograph 2

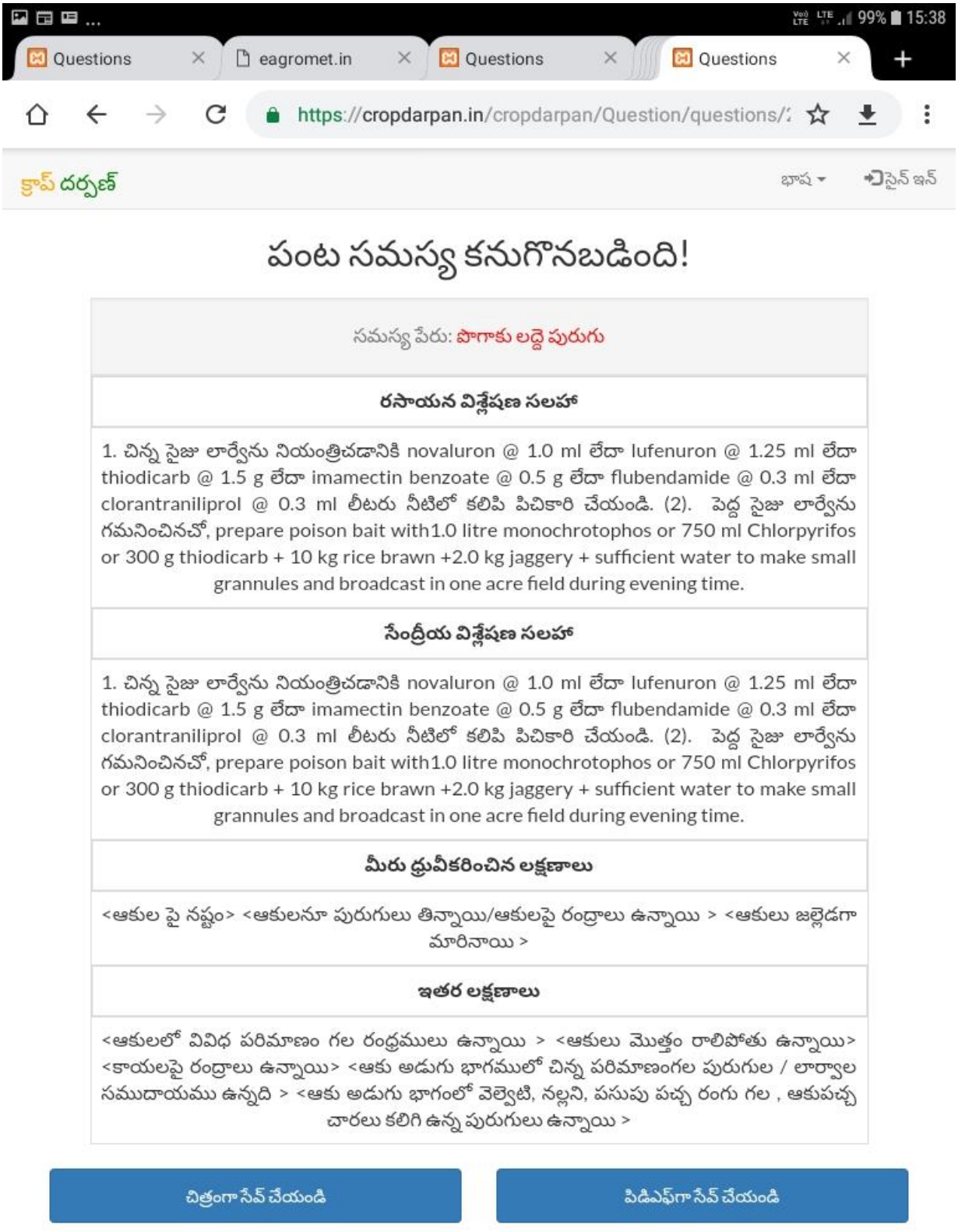


Photograph 3



Photograph 4

Snapshot of diagnosis found by the farmer



Part 3: Ground truth comparison study of 10 sample data records

ID	Sample File Name	Farmer Diagnosis	Expert Diagnosis	Reason for Expert Diagnosis	Correct/ Wrong
1	adivireddy0022date102019	Tobacco caterpillar	Tobacco caterpillar	Irregular holes on leafs and caterpillers presence and its texture	Correct
2	b_ramulu0015date201019	Aphids	Aphids	presence of group of black color insects	Correct
3	farmernamebhartkumar003	Leaf hoppers	Leaf hoppers	browing of leaves, presence of insects	Correct
4	farmernamepentaiah0029date15112019	Spotted Boll Worm	Helcoverpaor american boll worm	bolls holes and insects	Wrong
5	g_mallikarjan002322102019	White flies	White flies	presence of insects	Correct
6	m_d_khadeer0019date22102019	Spotted bollworm	Spotted bollworm	insect and fecal matter	Correct
7	m_gopalreddy0014date201019	Bacterial blight	Bacterial blight	black color irregular spots on bolls	Correct
8	k_veeresham001	Helcoverpaor american boll worm	Tobacco caterpillar	insect and leaf damage	Wrong
9	suresh0001date07112019farmername	whiteflies	whiteflies, Alternaria leaf spot	presence of insects and chlorotic spots, presence of spots on leaves	Correct
10	jangaaiha007	White flies	Leaf hoppers	presence of insects	Wrong