

First Phase Project report on

**Implementation of Six Sigma in Healthcare to reduce the errors
in Patient Identification**

Submitted to

Department of Industrial Engineering and Management

In partial fulfillment for the award of degree of

**Bachelor of Engineering in
Industrial Engineering and Management**

by

Alaparathi Charan Sai	1BM18IM002
Harini V	1BM18IM017
Sneha V	1BM18IM053



Under the Guidance of

Dr. RathanRaj K. J.

Professor



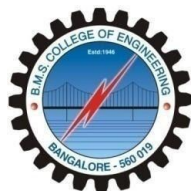
**DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT
B.M.S. COLLEGE OF ENGINEERING
BENGALURU - 560019
January 2022**

CERTIFICATE

DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

B.M.S. COLLEGE OF ENGINEERING

Bengaluru – 560019



This is to certify that the Project Work Phase 1 (21IM7DCPW1) entitled
**“Implementation of Six Sigma in Healthcare to reduce the errors
in Patient Identification”** has been successfully completed by

NAMES	USN
Alaparthi Charan Sai	1BM18IM002
Harini V	1BM18IM017
Sneha V	1BM18IM053

As a part of the 7th semester curriculum in Bachelor of Engineering in Industrial Engineering and Management, B.M.S. College of Engineering, an autonomous College under VTU, during the academic year 2021 - 2022

Under the guidance of:

Dr. RathanRaj K. J.
Professor and Guide
Dept. of IEM, BMSCE

Dr. B. Ramesh Nayak
Head of the Department
Dept. of IEM, BMSCE

Semester End Examinations:

	Name	Signature with date
Examiner 1:		
Examiner 2:		

Abstract

In the healthcare sector, quality should be given utmost importance in order to achieve good patient care. Quality in healthcare includes reducing the patient waiting time, avoiding billing errors, maintaining customer records properly and many others. An extensive literature survey was performed to understand the problems. In this project, problems existing in the Bangalore Baptist Hospital were identified, out of which the Patient Identification problem results in the majority of the risk. In order to reduce such errors caused, Six Sigma is the appropriate tool as it focuses on reducing defects per million. DMAIC approach has been implemented in this project.

Acknowledgement

The successful completion of the Phase 1 project work titled “**Implementation of Six Sigma in Healthcare to reduce the errors in Patient Identification**” gives us immense pleasure. We take this opportunity to express our deep sense of gratitude to all those who have helped us directly towards the completion of this project.

Members of the Team are indebted to and thank Dr. RathanRaj K. J, Professor, Department of Industrial Engineering and Management, B M S College of Engineering profusely for his constant guidance and mentorship throughout this project.

We would like to convey our sincere thanks to Dr. B Ramesh Nayak, Associate Professor and Head, Department of Industrial Engineering and Management, B M S College of Engineering.

We would also like to acknowledge the support and encouragement extended for this project by our coordinator, Dr. V. N. Shailaja, Associate Professor, Department of Industrial Engineering and Management, B M S College of Engineering.

We are grateful to Dr. Muralidhara S, Principal, B.M.S. College of Engineering for his continuous support.

We would like to express our gratitude toward the Bangalore Baptist Hospital for having given us an opportunity to undertake this project.

We would like to thank Dr. Girish.T.S. and Dr. Badri Datta H.C. for giving us the opportunity to work in the Bangalore Baptist Hospital.

We also would like to thank Dr. Manju Chacko for guiding us throughout the project and helping us realise the importance of the project’s objectives.

Lastly, we would like to thank all the non-technical and non-teaching staff of both B.M.S.College of Engineering and the Bangalore Baptist Hospital for all the support we have received.

Alaparathi Charan Sai

Harini V

Sneha V

Table of contents

Table of contents	1
List of Figures and Tables	2
List of Symbols, Abbreviations, Nomenclature	3
Chapter 1. Introduction	4
Chapter 2. Literature Review	6
Chapter 3. Six Sigma	10
3.1. What is six sigma?	10
3.2. Why six sigma?	10
3.3. Advantages of six sigma	10
3.4. Disadvantages of Six sigma	11
3.5 Methodologies of Six sigma	12
3.5.1 DMAIC	12
3.5.2 DMADV	13
Chapter 4. Bangalore Baptist Hospital	14
4.1. Problems observed at the BBH	14
Chapter 5. Problem Formulation	15
5.1. Problem Definition	15
5.2. Patient Identification Problem	15
5.3. Goals and objectives	16
5.4 Methodology	16
Chapter 6. Conclusion	20
Timeline	21
References	23

List of Figures and Tables

Fig. 3.5.1	DMAIC	12
Fig. 3.52	DMADV	13
Fig. 5.4.1	Methodology flowchart	18
Fig. 5.4.2	SIPOC diagram	20

Table 1.	Gantt chart	22
----------	-------------	----

List of Symbols, Abbreviations, Nomenclature

ANOVA	Analysis of Variance
BBH	Bangalore Baptist Hospital
CQI	Continuous Quality Improvement
CSF	Critical Success Factors
CTQ	Critical To Quality
DMADV	Define, Measure, Analyze, Design, Verify
DMAIC	Define, measure, Analyze, Improve, Control
KPI	Key Performance Indicator
NABH	National Accreditation Board for Hospitals & Healthcare Providers
NABL	National Accreditation Board for Testing and Calibration Laboratories
OP	Out-Patient
QCI	Quality Council of India
QFD	Quality Function Deployment
SIPOC	Suppliers, Inputs, Process, Outputs, Customers
SS	Six Sigma
VOC	Voice Of Customer

Chapter 1. Introduction

Healthcare is a huge sector which provides a lot of services pertaining to health. Managing any health institution requires many doctors, skilled and trained workers, finance and resources. Managing them properly makes the institution successful. Anyhow, in the healthcare sector maintaining quality is difficult when compared to other sectors, because in the healthcare sector almost all the work is done manually and is prone to have defects. The defects made do not only indicate loss of money, it can also be in terms of life since the errors made could be fatal sometimes. Therefore it is important to maintain and manage the institution with systematic techniques.

Hospitals have many functional divisions like Cardio, Ortho, Paediatrics, Surgery and others. All these functional divisions largely need laboratories to function properly and treat the patients according to the procedure. Laboratory is a hub to all the functional divisions. A small mistake made in laboratories can cause a huge impact on the treatment of the patient in the functional divisions. Few common problems seen in the laboratories are as follows

- Sample rejections
- Missing reports
- Breakdown time
- Long waiting time
- Redos of sample collection
- Wrong patient identification
- Delay in reports
- Disunion of safety measures
- Technical issues like no correlation between the report and patients' problem and many others.

Out of these, the wrong patient identification is a complex problem as identifying the patient wrongly can lead to exchange of reports, wrong supply of medicines, providing wrong treatments and wrong billings too. This is one of the problems that can be controlled by implementing systematic techniques since the errors are made manually.

This project aims to reduce the number of errors occurring by the patient identification problem at the Bangalore Baptist Hospital in an Out-patient laboratory, where blood samples and urine samples are being collected. This way, patient safety and quality of the laboratories can be significantly increased. In order to do the same, Six Sigma is being used as a quality tool.

Chapter 2. Literature Review

A methodical literature review of various publications on Six Sigma, extending from 2006-2021, was adopted to understand the methodology, statistical tools available and learn the advantages and disadvantages of Six Sigma to reduce the errors in any process. The key findings from these publications are summarized as follows:

Six Sigma has a wide variety of applications in many sectors other than manufacturing. The paper written by Cvetan Redzic and Jongmoon Baik focuses on application of Six Sigma to improve the quality of a Software product. Here, before applying product development life cycle, methodical study of the product was done. To improve the product process and increase the quality, Six Sigma was used along with many tools like Kano model, Voice of the customer, Pareto analysis to get the root cause of the problem in the process and solve it[1]. In the paper written by Celep Oguz *et al.*, a combination of Lean tools and Six Sigma tools were used to improve the process of production of concrete. This paper focuses on CTQ being a process indicator to a Six Sigma project. By eliminating the waste after applying lean, 25% of the lag in the schedule was compensated during the production. Process capability was the parameter used to do the same. Value stream mapping and control charts were used to measure the process capability [2]. The paper written by Ayon Chakraborty *et al.*, focuses on the empirical findings of the case studies where Six Sigma is implemented in a hospital, a hotel, a public sector and consultancy sector. Implementing Six Sigma was a successful project. This paper inference is that Six Sigma usage in the service sector also brings in huge improvements in the process. The methodology used to solve a project problem is explained. Using CTQs, KPIs and CSFs as the framework, the studies were made and compiled[3]. The paper written by Zhiyi Zhuo focuses on implementing Six Sigma in a bank using ANOVA to increase the service offered. Here, emphasis is given on the fact that Six Sigma works purely on data. DMAIC approach was being used because the process already existed and just needed to be improved. The goal of this paper was to increase the Value of the customer by reducing waiting time and providing effective servicing. This was done by reducing the errors occurring in the queue keeping time period as the CTQ parameter[4]. K. Prabu, J *et al.* talks about implementation of DMAIC in

production of submersible pumps. The root cause was identified by design of Experiments and it was identified in a component called stage casing. The calculation of sigma level before and after implementing six sigma had increased from 3.9 to 3.97. The implementation of six sigma to manufacturing pumps proved effective through this project[5].

The application of Six Sigma in the Healthcare sector has impacted a lot of lives by increasing the quality of treatment received by the patients. The paper written by Henk de Koning et al. focuses on the basics of lean six sigma and its application to Red cross hospital where the projects were not methodic and project tracking was neglected. Complexity reduction in hiring personnel and reducing operating theater starting times were solved using Lean Six Sigma[6]. Similar lean technologies and tools to increase the productivity of hospitals is discussed in the paper by Ross W. Simon and Elena G. Cancari[7]. Mehmet Tolga et al. discusses the attributes that contribute to the quality like safety, patient-centered, timely, efficient, equitable. They used Six sigma as a tool in various cases considering these quality aspects. The steps and methods to achieve maximum resource utilisation, reduce redundancies in the process and improve working conditions for healthcare personnel have been discussed [8]. Prathibha Varkey et al. gives a basic introduction to the different tools that can be used to achieve quality in healthcare. The CQI model emphasizes the view of healthcare as a process and gives importance to the focus on the system rather than the individual when considering improvement opportunities[9]. Hseng-Long Yeh et al. utilize the value stream map to seek out non-value-added activities, the cause and effect diagram to analyze the root causes of waste and generate the improvement project by brainstorming. They discussed how SS eliminates variation and Lean attempts to develop a theoretical or practical link between quality and speed[10]. The paper written by Giovanni Celano et al. defines the approach of integrating six sigma with discrete event simulation during the development of a quality improvement project. This was achieved by merging the two SS methodologies, to achieve the simulation model. The roadmap was tested and simulated to focus on the re-organisation of the flow of emergency patients affected by vertigo symptoms[11]. The paper written by Jami L.DelliFraine et al. gives a summary of various research papers that have attempted to achieve Six sigma in healthcare. It was found that out of 55 articles reviewed, 30 reported the correct use of SS; and Three of these articles

actually achieved near SS-level quality of care[12]. Matthew J. Liberatore discusses about a similar survey that conducted where Six sigma was applied to 88 hospitals which covered areas such as Administration, Ambulatory care, Emergency care, Inpatient care and Therapeutic support in healthcare[13]. Serkan Deniza and Mesut Çimenb talk about the barriers that exist in applying six sigma to the healthcare sector. Few of the barriers identified by this paper are: Belief in organizations that six sigma applications are complex, No sufficient knowledge of six sigma, Barriers related to organizational culture, Top management support is not sufficient for six sigma projects and Employees resistance to change[14]. The paper written by Deachen Angmo and Dr.Suman Kant deals with the obstacles in the healthcare sector and the tools used to prevent them. In India healthcare sectors are still following traditional systems ignoring the factors like patient safety, poor scheduling, competition, malpractice and rising costs. DMAIC is the methodology to be applied to implement six sigma. This methodology questions each and every step of the process(input) and hence delivers the best quality(output), fulfilling the customer's needs. The various tools used in each step are defined clearly in this paper[15].In this paper written by K. B. N.Raja Mohamed et al., DMAIC was applied to maximize the efficiency of treating the patients according to the priority of patients. The SIPOC diagram was used for suppliers, inputs, process flows, outputs and customers. CTQ analysis was done based on the ranking. QFD was also used where “the whats”, “the hows” were answered. Multiple regression model was used in the Improve stage. Later Monte Carlo simulation was used to control the process. Patient satisfaction was achieved and factors influencing it were identified[16].

Our project's main interest is to reduce the errors occurring in the laboratory pertaining to the patient identification. Errors in the laboratory of a healthcare center is a common problem that can actually be solved by taking proper measures. The study written by Naadira Vanker et al., focuses on the errors that occur during the pre-analytical phase of the laboratory processes in the hospital. Reduction of such errors was achieved by the application of Six sigma methodologies. The error rates and the sigma value before applying six sigma was calculated was 4.46 which had a greater scope for improvement and reducing errors[17]. The paper written by Susan Chu Walley et al., mainly focuses to reduce the errors caused by patient identification bands. This paper aimed at reducing such errors by 50% in a span of 12 months in a

selected pediatric institution. With the application of the DMAIC process and standardizing the patient identification band, there was a significant decrease in the errors[18]. In this paper written by Ahmed Alkuwaiti the focus is on reducing the medication error in the outpatient pharmacy department. With the help of DMAIC and other quality tools, the goal was to reduce the errors by 20%. It was seen that after applying Six sigma, the no. of defects per million reduced from 56000 to 5000. And the sigma level was also improved from 3.09 to 4.08[19]. In this paper written by Tamer C. Inal et al., it is seen that organizing workflow is a major part of achieving the laboratory processes without errors. This project uses lean six sigma to ease the laboratory process and reduce its turnaround time. The final results of this study are that the process time of 3 hours and 22.5 minutes of non-value added time near the reception area was reduced. This paper gives the confidence that errors in the clinical laboratories can be reduced by applying lean Six sigma and using quality tools[20].

3.1. What is six sigma?

Six sigma is the procedure adopted to improve and optimize the business processes. Six Sigma was first introduced in Motorola. Six sigma aims in achieving the customer satisfaction and requirements by reducing any variations in achieving the same. Statistically, six sigma means that the defects produced in a process must be less than 3.4 per million opportunities. Six sigma strives for achieving nearly zero defects in any process and improve the quality of the organization.

3.2. Why six sigma?

For process improvement, Six sigma is employed. We generally employ the DMAIC process in order to increase the quality, efficiency and accuracy of our work. This allows us to perform root cause analysis of a given problem, which allows us to define a problem, measure it using quantitative analysis. Perform analysis using control charts, pareto analysis, and histogram charts, and to improve and control the process, which allows us to reduce errors, lower costs, and increase profit, as well as improve customer satisfaction.

3.3. Advantages of six sigma

The advantages of Six Sigma are as follows[21]:

- ➔ Improved quality- Since Six Sigma is a scientific tool that is used for quality enhancement in a process or a system. By implementing Six Sigma there will be no or less rework and hence the rework costs are reduced significantly.
- ➔ Customer satisfaction- Customer satisfaction depends on the variations in the product, process or a system. Lesser the variation, more delighted the customers. By implementing Six Sigma, customer satisfaction also increases.

- Huge profits- Six Sigma implementation cuts down the usage of non-value added costs and eliminates waste in the process which is non-value added. By doing this, Days Sales Outstanding is increased. When DOSs are increased, huge profits are generated.
- Process efficiency- Implementation of Six Sigma, generates only 3.4 defects per million. We can say that for the input provided, output is 99.999666%. There is a significant increase in the process efficiency.
- Eliminates waste- Any kind of waste in the system, like talent, transportation, non- value added activity and other resources, is analysed and removed with the help of Six Sigma tools
- Decreases defects- The defects removal in a system can be performed by using empirical and statistical management techniques. And Six Sigma is one of the most powerful tools to perform the same.

3.4. Disadvantages of Six sigma

Drawbacks to any technique are inevitable. Likewise Six Sigma implementation too has few disadvantages. They are as follows[22]:

- Six Sigma level is an ideal technique. Currently, It is impossible to be met even though the Six Sigma steps are properly followed.
- Although the Six Sigma solution can be standardized and is being used for larger time periods, the initial investment of time is huge.
- Sometimes, the organization might forget that it is a process improvement technique and not the process itself and give more priority to Six Sigma projects than the actual project itself.
- Six sigma requires a lot of financial and human resources. It will be very difficult for small organizations to implement Six Sigma as it is expensive.

- Six Sigma uses a large range of tools at every stage. Selecting an appropriate tool according to the business processes is a real challenge.

3.5 Methodologies of Six sigma

The two approaches that are being used to accomplish Six Sigma are DMAIC and DMADV. Define, Measure, Analyze, Improve, and Control is the abbreviation for the DMAIC process. This is a method of enhancing current processes in order to decrease variance. DMADV stands for Define, Measure, Analyze, Design, and Validate/Verify. When a new process or product needs to be developed, or when a current process needs to be improved, the DMADV technique is utilized.

3.5.1 DMAIC [23]



Fig. 3.5.1 DMAIC

- Define: In this phase, the defects or errors existing in the organisation are described. This phase has two main steps: where the first step is the selection and scope of the project and the identification of defects.
- Measure: The objective of this phase is to define the defect and identify the root cause based on the progress of the process. The CTQ variables are identified which are used to express a measurable result of the process.
- Analyze: In this phase, the effect of process factors(inputs) on process outcomes(outputs) is understood.
- Improve: In this stage, the identification of the best solution and applying it to the process is the main task. The factors on which this can be decided are cost, risks, implementation requirements and expected process performance.

- Control: The various steps are planned in this stage in order to keep the improved process in its current level using a quality control plan. Tools like statistical control charts, 5S and Poka-yoke can be used to keep the process in control.

3.5.2 DMADV [24]

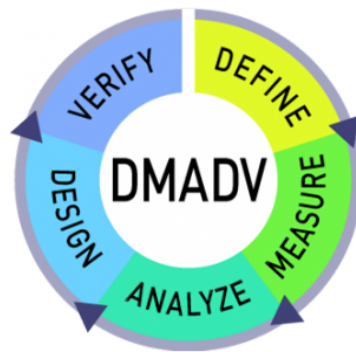


Fig. 3.5.2. DMADV

- Define: This phase is used to define the goals, objectives and deliverables needed by the customer. The customer can be internal or external.
- Measure: The customer needs and specifications are determined by the use of various statistical tools.
- Analyze: The various process and design options are analyzed by the team to meet the customer needs and customised requirements.
- Design: After the analysis, the processes and designs suitable and efficient designs are modeled. The best design model is chosen after considering different parameters.
- Verify: The implementation of the new design is tested and applied to the customer's requirement and is checked for any errors and its ability to meet customer requirement

Bangalore Baptist Hospital is a multi-speciality private hospital in Hebbal, Bengaluru. It is headed by Dr. Jasper McPhail and it was started in 1973. Since 1989, the hospital has been governed by Christian Medical College, Vellore. BBH is the first mission hospital which is NABH and NABL accredited. BBH is one of the hospitals which give priority to quality. Overall rating of the hospital is 4.2 out of 5 according to Google reviews given by the common people who have visited the hospital. The hospital has won the QCI DL Shah Quality award for two consecutive years for quality[25].

Bangalore Baptist Hospital was chosen for this project because the data is easily available from the primary source i.e the hospital quality manager.

4.1. Problems observed at the BBH

Bangalore Baptist hospital gives a lot of priority to quality. Even though too much time and work is put into it, it is difficult to make the system error free. There were few problems observed at the hospital. Although there is some quality controlled system, it is not error free. Six Sigma aims to achieve a system with only 3 defects per million. The general problems observed at the hospital are:

- Missing reports
- Wrong admission
- Wrong supply of medicine
- Patient-doctor miscommunication
- Wrong patient identification
- Equipment failure
- Ineffective waste management
- Medicine out of stock at the pharmacy

These are the common errors that happen frequently in the hospital. Most of these errors occur in the laboratory since the majority of the processing work happens there. Out of these 8 errors, wrong patient identification is very serious and quite frequent in the laboratories comparatively. Therefore Six Sigma is being applied and implemented to the 'wrong identification problem' between two patients.

Chapter 5. Problem Formulation

The healthcare sector is one of the sectors where errors should be minimized as much as possible. This is because the errors done could affect the stakeholders extremely and sometimes it could even be fatal. Since healthcare is a hands-on process, the variance is higher and there is less number of variables to work on. Therefore a systematic approach or an engineered approach is required to reduce the errors. This project aims to reduce the variance occurring at the Bangalore Baptist Hospital by implementing Six Sigma to one of its problem causing processes. Since the processes in the hospitals are already existing, the DMAIC (Define, Measure, Analyse, Improve and Control) approach is being applied.

5.1. Problem Definition

Implementation of Six Sigma in the Out-Patient laboratory processes to analyse and reduce the errors in identifying two patients distinctively using the DMAIC approach with the help of different Industrial Engineering tools at the Bangalore Baptist Hospital.

5.2. Patient Identification Problem

Patient Identification problem occurs when the hospital staff fail to distinguish between two different patients with the same name or in some cases different names too. This situation is very dangerous and it can sometimes be fatal too. This leads to exchange of medical reports, prescriptions and even blood samples. In the end, the staff will end up giving the wrong treatment to the patient. In one of the cases, an infant was supplied with the wrong medicine and the situation was almost fatal. This happened because the staff were unable to identify the infant's name properly. Even though it is a computerized system and the number of errors has reduced over the years, the system still has a significant number of such cases. Therefore it is important to apply Six Sigma to this problem and reduce the number of errors.

5.3. Goals and objectives

The goal of this project is to successfully implement Six Sigma in OP laboratory to reduce the number of errors in Patient Identification problem in the Bangalore Baptist Hospital.

The objectives of the project include the following:

- Understand the current process
- Understand the root cause of the problem and factors that can be improved
- Calculate the sigma value of the current process
- Improving the current process using DMAIC approach by reducing the defects
- Calculate sigma level of the improved process
- To achieve process improvement with minimum cost include

5.4 Methodology

The methodology that can be followed to obtain a solution to the problem defined can be understood from the figure. shown below.

In the first phase of the project work, the project team was formed and after a thorough literature review and discussion with the guide, the area of implementation of six sigma to the healthcare sector was finalized.

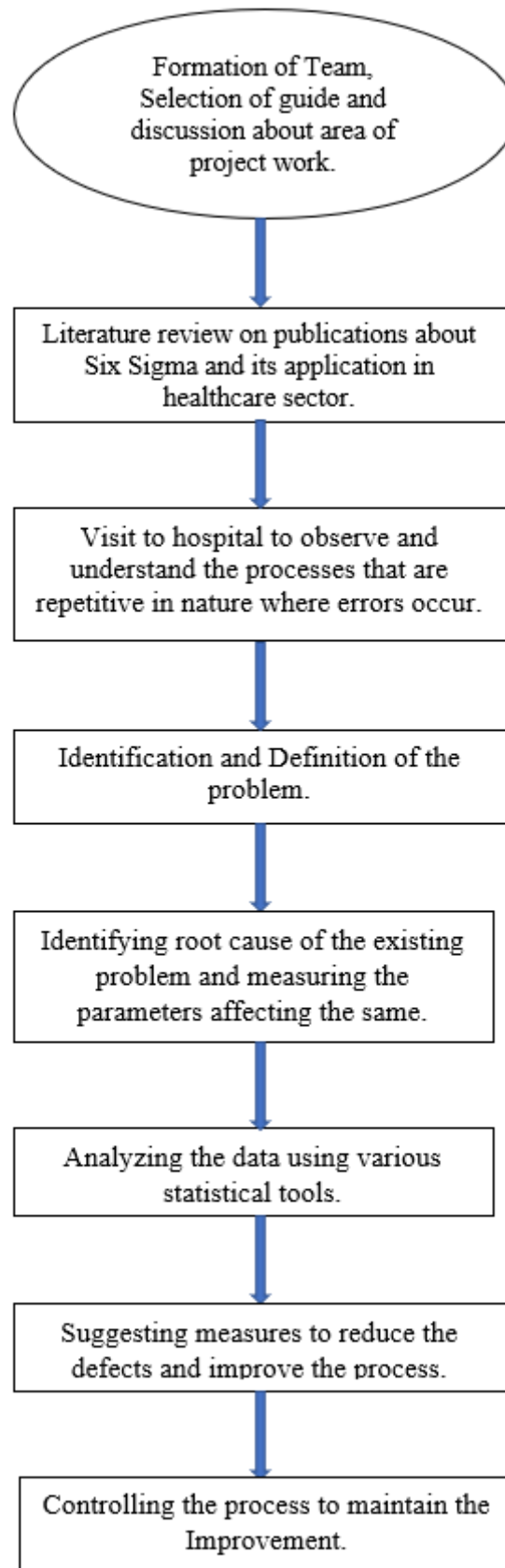


Fig.5.4.1. Methodology flowchart

To arrive at the problem statement, we approached the Bangalore Baptist Hospital, where we were given the permission to obtain previous data, measure the required data to carry out the project activities to eliminate errors existing in the processes. During the initial visits, different processes in the hospital were observed where we were exposed to the Key Indicators used to assess the level of performance of a process in the hospital. The data on the following were obtained:

- Type of errors
- No. of errors existing in different departments
- Frequency of occurrence and
- The measures taken to avoid such errors.

Analysing this, we recognized that “Patient identification” problems can be seen frequently in the laboratory processes and are significant.

The observation of the OP laboratory process was clearly understood by drawing SIPOC. A SIPOC diagram is the tool which is used for the process improvement techniques to identify every relevant element of the project before defining a problem which may be complex or may not be well scoped. SIPOC is an acronym which stands for supplier, input, process, output and customer. The aim is to understand the suppliers, inputs required and the outputs delivered to complete a process.[26]

Thus, a thorough observation of the out-patient laboratory process was carried out which is more specific than other processes in the BBH hospital. From the observations, we arrived at the SIPOC diagram to ease the process of defining the problem statement.

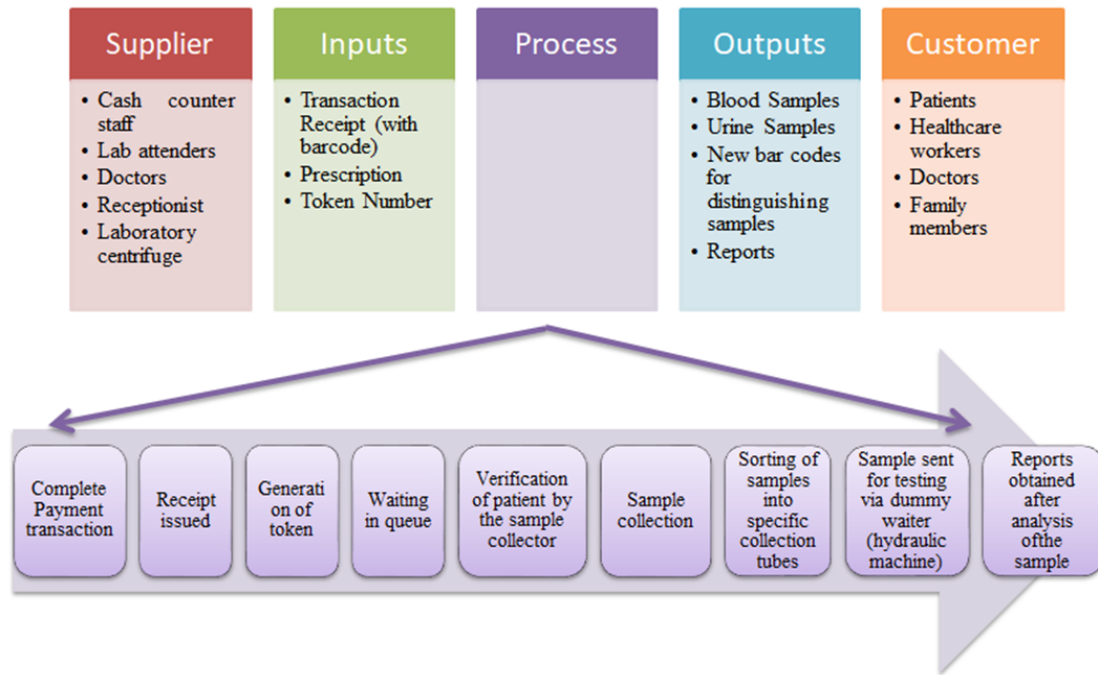


Fig.5.4.2. SIPOC Diagram

The details of the process have been briefly explained using the fig above. The definition of the problem statement completes the Define stage of DMAIC procedure.

In the second phase of the project, the remaining steps as described in the flowchart are to be achieved to obtain the solution of the problem which will improve the existing process and result in reduction of occurrence of errors in the process.

Chapter 6. Conclusion

This project work focuses on implementing Six Sigma for the OP laboratory process in the Bangalore Baptist Hospital. An extensive Literature Review was carried out to identify and analyse the problems existing in the healthcare sector. Out of many problems that were identified in the hospital, the patient identification problem was selected for this project as this problem is very common and threatening to the patient. Here, for the patient identification problem, the hospital does not identify the patients distinctively due to some technical and manual errors. This leads to errors that can be fatal sometimes. In order to avoid this, the DMAIC approach is followed. Currently, the Define phase is performed. In order to define the problem systematically, SIPOC was designed. This diagram gives information about the suppliers of the process, inputs and outputs of the process, the customers of the process and the detailed step wise procedure carried out in the process. Problem Statements and objectives are being made clear in the first phase of the project work. The other stages: Measure, Analyse, Improve and Control will be carried out with the help of its respective tools in the second stage of the project.

The timelines describe the list of steps to be completed in a chronological order. Gantt charts are used to understand the schedule of the different steps to complete the project within the given time frame. The Gantt chart for our project can be observed in the fig. shown below. For the first phase project work, the activities completed till now are highlighted in green and the plan for the second phase of the project work has been marked in red. The Gantt chart for our project is shown below in Table 1.

Table 1. Gantt chart

		OCTOBER				NOVEMBER				DECEMBER				JANUARY				FEBRUARY				MARCH				APRIL				MAY			
		W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4				
PROJECT INTRODUCTION																																	
Team and guide selection	PLAN																																
	ACTUAL																																
Discussion with guide	PLAN																																
	ACTUAL																																
Selection of the topic	PLAN																																
	ACTUAL																																
LITERATURE REVIEW																																	
What is Six Sigma	PLAN																																
	ACTUAL																																
Six Sigma Applications and Significance in Healthcare	PLAN																																
	ACTUAL																																
METHODOLOGY																																	
After Discussion with Guide, Approached Healthcare sectors to understand Processes	PLAN																																
	ACTUAL																																
Visiting Hospital	PLAN																																
	ACTUAL																																
DEFINE																																	
Understanding of Processes in Hospital	PLAN																																
	ACTUAL																																
SIPOC	PLAN																																
	ACTUAL																																
Problem definition	PLAN																																
	ACTUAL																																
MEASURE	PLAN																																
	ACTUAL																																
ANALYSIS	PLAN																																
	ACTUAL																																
IMPROVE	PLAN																																
	ACTUAL																																
CONTROL	PLAN																																
	ACTUAL																																

- [1] C. Redzic and J. Baik, "Six Sigma Approach in Software Quality Improvement", Fourth International Conference on Software Engineering Research, Management and Applications (SERA'06), pp. 396-406, Jan. 2006, doi: 10.1109/SERA.2006.61.
- [2] C. Oguz, Y.-W. Kim, J. Hutchison, and S. Han, "Implementing lean six sigma: A case study in concrete panel production". [Online]. Available: <https://iglcstorage.blob.core.windows.net/papers/attachment-28d001bd-e671-4a04-9da9-f811d3e02083.pdf>.
- [3] A. Chakraborty and K. Chuan Tan, "Case study analysis of Six Sigma implementation in service organisations", Business Process Management Journal, vol. 18, no. 6, pp. 992–1019, Nov. 2012, doi: 10.1108/14637151211283384.
- [4] Z. Zhuo, "Research on using Six Sigma management to improve bank customer satisfaction", International Journal of Quality Innovation, vol. 5, no. 1, Mar. 2019, doi: 10.1186/s40887-019-0028-6.
- [5] K. Prabu, J. Makesh, K. N. Raj, S. R. Devadasan, and R. Muruges, "Six Sigma implementation through DMAIC: a case study", International Journal of Process Management and Benchmarking, vol. 3, no. 3, pp. 386–400, 2013, Accessed: Jan. 07, 2022. [Online]. Available: <https://ideas.repec.org/a/ids/ijpmbe/v3y2013i3p386-400.html>.
- [6] H. de Koning, J. P. S. Verver, J. van den Heuvel, S. Bisgaard, and R. J. M. M. Does, "Lean Six Sigma in Healthcare", Journal For Healthcare Quality, vol.28, no. 2, pp. 4–11, Mar. 2006, doi: 10.1111/j.1945-1474.2006.tb00596.x.
- [7] R. W. Simon and E. G. Canacari, "A Practical Guide to Applying Lean Tools and Management Principles to Health Care Improvement Projects", AORN Journal, vol.95, no. 1, pp.85–103, Jan. 2012, doi: 10.1016/j.aorn.2011.05.021.
- [8] M. Tolga Taner, B. Sezen, and J. Antony, "An overview of six sigma applications in healthcare industry", International Journal of Health Care Quality Assurance, vol. 20, no. 4, pp. 329–340, Jun. 2007, doi:10.1108/09526860710754398.
- [9] P. Varkey, M. K. Reller, and R. K. Resar, "Basics of quality improvement in health care", Mayo Clinic proceedings, vol. 82, no. 6, pp. 735–9, 2007, doi:10.4065/82.6.735.
- [10] Hseng-Long Yeh, "Applying lean six sigma to improve healthcare: An empirical study", African Journal Of Business Management, vol. 5, no. 31, Dec. 2011,

doi:10.5897/ajbm11.1654.

[11] G. Celano, A. Costa, S. Fichera, and G. Tringali, “Linking Six Sigma to simulation: a new roadmap to improve the quality of patient care”, *International Journal of Health Care Quality Assurance*, vol. 25, no. 4, pp. 254–273, Apr. 2012, doi:10.1108/09526861211221473.

[12] J. L. DelliFraine, Z. Wang, D. McCaughey, J. R. Langabeer, and C. O. Erwin, “The Use of Six Sigma in Health Care Management”, *Quality Management in Health Care*, vol. 22, no. 3, pp. 210–223, 2013, doi: 10.1097/qmh.0b013e31829a838e.

[13] M. J. Liberatore, “Six Sigma in healthcare delivery”, *International Journal of Health Care Quality Assurance*, vol. 26, no. 7, pp.601–626, Aug. 2013, doi:10.1108/ijhcqa-09-2011-0054.

[14] S. Deniz and M. Çimen, “Barriers of six sigma in healthcare organizations”, *Management Science Letters*, pp. 885–890, 2018, doi: 10.5267/j.msl.2018.6.009.

[15] Deachen Angmo and Dr. Suman Kant, “Six Sigma Implementation in Healthcare Industry: Past, Present and Future”, *International Journal of Engineering Research & Technology*, vol. 4, no. 6, 2015, Accessed: Jan. 29, 2020.[Online]. Available:<https://www.ijert.org/six-sigma-implementation-in-healthcare-industry-past-present-and-future>.

[16]K. B. N.Raja Mohamed, P. R. M., S. S., J. R. A., and R. Anderson, “Six sigma in health-care service: a case study on COVID 19 patients’ satisfaction”, *International Journal of Lean Six Sigma*, May2021, doi: 10.1108/ijlss-11-2020-0189.

[17] N. Vanker, J. van Wyk, A. E. Zemlin, and R. T. Erasmus, “A Six Sigma approach to the rate and clinical effect of registration errors in a laboratory”, *Journal of Clinical Pathology*, vol. 63, no. 5, pp. 434–437, May 2010, doi:10.1136/jcp.2009.072058.

[18] S. C. Walley, S. Berger, Y. Harris, G. Gallizzi, and L. Hayes, “Decreasing Patient Identification Band Errors by Standardizing Processes”, *HospitalPediatrics*, vol. 3, no.2, pp.108–117, Apr. 2013, doi: 10.1542/hpeds.2012-0075.

[19] A. A. Kuwaiti, “Application Of Six Sigma Methodology To Reduce Medication Errors In The Outpatient Pharmacy Unit: A Case Study From The King Fahd University Hospital, Saudi Arabia”, *International Journal for Quality Research*, vol. 10, no. 2, pp.267–278, Jun. 2016, doi: 10.18421/IJQR10.02-03.

[20] T. C. Inal et al.,“Lean six sigma methodologies improve clinical laboratory efficiency and reduce turnaround times”, *Journal of Clinical Laboratory Analysis*, vol. 32, no. 1, p. e22180, Feb. 2017, doi:10.1002/jcla.22180.

[21] “The six amazing benefits of employing six sigma in your organization”. <https://www.greycampus.com/blog/quality-management/6-benefits-of-6-sigma> (accessed: Jan. 06, 2022).

[22]“Six sigma- Accounting For Management”. <https://www.accountingformanagement.org/six-sigma/> (accessed: Jan. 07, 2022).

[23]“DMAICVersus DMADV”. Isixsigma. <https://www.isixsigma.com/new-to-six-sigma/design-for-six-sigma-dfss/dmaic-versus-dmadv/> (accessed: Jan. 05, 2022).

[24] DMAIC Process: Define, Measure, Analyze, Improve, Control”. ASQ. <https://asq.org/quality-resources/dmaic> (accessed: Jan. 05, 2022).

[25] Bangalore Baptist Hospital”. <https://www.bbh.org.in/> (accessed: Jan. 07, 2022).

[26]“Guide to High-Level Process Mapping (SIPOC)”. Juran. <https://www.juran.com/blog/guide-to-high-level-process-mapping-sipoc/> (accessed: Dec. 23, 2021).

ORIGINALITY REPORT

11%

SIMILARITY INDEX

9%

INTERNET SOURCES

6%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Kingston University Student Paper	1%
2	International Journal of Health Care Quality Assurance, Volume 25, Issue 4 (2012-04-28) Publication	1%
3	worldwidescience.org Internet Source	1%
4	academicjournals.org Internet Source	1%
5	docshare.tips Internet Source	1%
6	Submitted to University of Warwick Student Paper	1%
7	Submitted to William Carey University Student Paper	<1%
8	Submitted to The University of Manchester Student Paper	<1%
9	www.researchgate.net Internet Source	<1%

10	DelliFraine, Jami L., Zheng Wang, Deirdre McCaughey, James R. Langabeer, and Cathleen O. Erwin. "The Use of Six Sigma in Health Care Management : Are We Using It to Its Full Potential?", Quality Management in Health Care, 2013. Publication	<1 %
----	--	------

11	Submitted to SUNY, Binghamton Student Paper	<1 %
----	--	------

12	studydaddy.com Internet Source	<1 %
----	-----------------------------------	------

13	www.juran.com Internet Source	<1 %
----	----------------------------------	------

14	Submitted to Universidad San Francisco de Quito Student Paper	<1 %
----	--	------

15	meu.edu.jo Internet Source	<1 %
----	-------------------------------	------

16	docplayer.net Internet Source	<1 %
----	----------------------------------	------

17	keoxtech.com Internet Source	<1 %
----	---------------------------------	------

18	Submitted to Siddaganga Institute of Technology Student Paper	<1 %
----	--	------

hdl.handle.net

19

Internet Source

<1 %

20

ijqr.net

Internet Source

<1 %

21

www.coursehero.com

Internet Source

<1 %

22

www.emerald.com

Internet Source

<1 %

23

www.emeraldinsight.com

Internet Source

<1 %

24

www.iiis.org

Internet Source

<1 %

25

[International Journal of Health Care Quality Assurance, Volume 26, Issue 7 \(2013-09-14\)](#)

Publication

<1 %

26

Scholar.sun.ac.za

Internet Source

<1 %

27

asq.org

Internet Source

<1 %

28

downloads.hindawi.com

Internet Source

<1 %

29

ikee.lib.auth.gr

Internet Source

<1 %

30

tutlance.com

Internet Source

<1 %

31

www.econstor.eu

Internet Source

<1 %

32

Angels Niñerola, María-Victoria Sánchez-Rebull, Ana-Beatriz Hernández-Lara. "Quality improvement in healthcare: Six Sigma systematic review", Health Policy, 2020

Publication

<1 %

33

Prabu, K., J. Makesh, K. Naveen Raj, S.R. Devadasan, and R. Murugesh. "Six Sigma implementation through DMAIC: a case study", International Journal of Process Management and Benchmarking, 2013.

Publication

<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On

First Phase Project Report on

**A MACHINE LEARNING MODEL TO PREDICT CROP PRICES OF
SEASONAL COMMODITIES IN THE MARKETS OF KARNATAKA**

Submitted to the Department of Industrial Engineering and Management

In partial fulfillment for the award of Degree of

Bachelor of Engineering

in

Industrial Engineering and Management

By

DEEPA M PATIL

1BM18IM013

SAI PRANAV P

1BM18IM044

VAIDEHI BHASKARA

1BM18IM066



Under the Guidance of

Dr. K. T. Ramesh

Assistant Professor



DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT

B.M.S COLLEGE OF ENGINEERING

BENGALURU - 560019

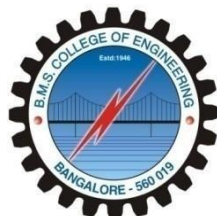
JANUARY -2022

CERTIFICATE

Department Of Industrial Engineering and Management

B.M.S. COLLEGE OF ENGINEERING

Bengaluru – 560019



This is to certify that the Final Year Project Phase-I (21IM7DCPW1) entitled

A MACHINE LEARNING MODEL TO PREDICT CROP PRICES OF SEASONAL COMMODITIES IN THE MARKETS OF KARNATAKA

has been successfully completed by

DEEPA M PATIL

1BM18IM013

SAI PRANAV P

1BM18IM044

VAIDEHI BHASKARA

1BM18IM066

as a part of the 7th semester curriculum in Bachelor of Engineering in

Industrial Engineering and Management,

B.M.S. College of Engineering, an autonomous College under VTU,

during the academic year 2021 - 2022

Guided by:

Dr. K. T. Ramesh

Assistant Professor

Dept. of IEM, BMSCE

Dr. B. Ramesh Nayak

Head of the Department

Dept. of IEM, BMSCE

Semester End Examinations:

	Name	Signature with date
Examiner 1:		
Examiner 2:		

અન્નદાતા સુખી ભવ

Abstract

Agriculture is the backbone of India. However, marginal farmers especially, earn infinitesimal amounts from their harvested yields partly because of their relatively smaller land holdings, and partly because of inadequate resources to obtain the required prices. In order to alleviate the stress caused by the lack of sound financial planning, this project proposes the utilization of machine learning to predict commodity prices. The solution obtained through such a model would assist farmers in predicting the price and associated yield of their crops, sowing patterns and suitable recommendations for sales. As a first step, this project thoroughly explores the literature in this domain, identifies potential sources for data collection, and proposes a methodology to design a machine learning model that predicts prices for seasonal crops specific to the markets of Karnataka.

Acknowledgement

The successful completion of this project titled “A Machine Learning Model to Predict Crop Prices of Seasonal Commodities in the Markets of Karnataka” gives us immense pleasure.

We are indebted to, and sincerely thank our project guide, Dr. Ramesh K. T., assistant professor, Department of Industrial Engineering and Management, B. M. S. College of Engineering, immensely for his time, patience, knowledge, and for guiding us throughout this project.

We would like to convey our special thanks to Dr. B. Ramesh Nayak, associate professor and HOD, Department of Industrial Engineering and Management, B. M. S. College of Engineering, for his support.

We would also like to acknowledge the support and encouragement extended for this project by our project coordinator, Dr. V. N. Shailaja, associate professor, Department of Industrial Engineering and Management, B. M. S. College of Engineering, and thank her for her leadership.

We also extend our gratitude to our families and friends, and thank all the teaching and nonteaching staff of our department for their bolstering presence.

Deepa M Patil
Sai Pranav P
Vaidehi Bhaskara

Table of Contents

List of Tables	2
List of Figures	2
List of Symbols, Abbreviations, Nomenclature	3
Chapter 1. Introduction	5
1.1. Agriculture in the Global Context	5
1.2. Agriculture and India	6
1.3. Indian Policies: NITI Aayog and India Farm Laws	7
1.4. Integrating Technology with Agriculture	9
1.5. Typical Problems faced by marginal farmers	13
1.6. Developing Price Prediction Models to assist Marginal Farmers	15
Chapter 2. Literature Review	16
Chapter 3. Problem Definition	37
3.1. Objectives of this Study	38
Chapter 4. Methodology	39
Chapter 5. Data Collection	42
Chapter 6. Project Roadmap	43
Chapter 7. Conclusion	44
Appendix A: Data Sources	45
References	47

List of Tables

Table 2.3.1.	Literature Review	19
--------------	-------------------	----

List of Figures

Figure 1.1.1.	UNSDG #2 focuses on zero hunger	5
Figure 1.3.1.	Multidimensional poverty index score in Karnataka	8
Figure 1.3.2.	NITI Aayog's blueprint for reforms in the farming sector	9
Figure 1.4.1.	Stages in Agriculture: Opportunity for technology	10
Figure 1.4.2.	AI and Robotics companies in Agriculture	12
Figure 1.4.3.	Artificial Intelligence in Agriculture	12
Figure 2.1.1.	Literature Review Overview	16
Figure 2.1.2.	Integrative Literature Review and Conceptual Framework Development Approach	17
Figure 2.2.1.	Frequency distribution of papers and articles addressing a given agricultural problem using different ML techniques	18
Figure 4.1.	Methodology Flowchart of this project	39
Figure 6.1.	This project's Gantt Chart	43
Figure 5.1.	Agmarknet	45
Figure 5.2.	Open Government Data Platform of India	45
Figure 5.3.	Krishi Marata Vahini	46
Figure 5.4.	Coffee Board of India	46
Figure 5.5.	Karnataka State Natural Disaster Monitoring Center	46

List of Symbols, Abbreviations, Nomenclature

UN SDG	United Nations Sustainable Development Goals
GDP	Gross Domestic Product
NITI	National Institution for Transforming India
SFAC	Small Farmers Agribusiness Consortium
VCA	Venture Capital Assistance Scheme
GIS	Geographic Information System
EGCGS	Equity Grant and Credit Guarantee Fund Scheme
ML	Machine Learning
AI	Artificial Intelligence
NDVI	Normalised Difference Vegetation Index
CCCI	Canopy Chlorophyll Content Index
NDRE	Normalised Difference RedEdge
MSAVI	Modified Soil Adjusted Vegetation Index
IOT	Internet of Things
MSP	Minimum Support Prices
y-o-y	Year-Over-Year
IoT	Internet of Things
ARIMA	Auto- Regressive Moving Average
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Square Error
STL	Seasonal-trend decomposition procedures based on loess
ELM	Extreme learning machines
AIC	Akaike Information Criterion
BIC	Bayesian Information Criteria
CORR	Correlation

VAR	Vector AutoRegression
LSTNet	Long- and Short-term Time-series network
NCDEX	National Commodity & Derivatives Exchange
GWO-KELM	Grey Wolf Optimization evolving Kernel Extreme Learning Machine
ANN	Artificial Neural Network
SFFS	Sequential Floating Forward Selection
RFE	Recursive Feature Elimination
EDA	Exploratory Data Analysis
API	Application Programming Interface
WPI	Wholesale Price Index
SVR	Support Vector Regression

Chapter 1. Introduction

1.1. Agriculture in the Global Context

The United Nations Sustainable Development Goal #2 focuses on “End hunger, achieve food security and improved nutrition and promote sustainable agriculture.” To achieve this goal, effort is needed to sustainably raise agricultural production, improve global supply chains, reduce food wastage and ensure that everyone who is suffering from hunger and malnutrition has access to nutritious food to survive.

It is seen that in many parts of the world, there's an increase in scarcity of water, fertile soil, and land, which are the major key inputs for crop production. Hence there is always a need to manage them sustainably. Increasing crop yields and restoring degraded lands would relieve the pressure of clearing all the forest area for agricultural production. Practices like management of scarce water by improving irrigation, storage technologies and producing new varieties of drought resistant crops will help in sustaining dry lands.

With the change in climatic conditions, there is an associated change in temperature and rainfall, because of which the global community is making an effort to increase investment for research and development of new technologies especially in agriculture, thus improving the food systems' sustainability globally.



Figure 1.1.1: UNSDG #2 focuses on zero hunger [1]

1.2. Agriculture and India

India is the second most populous country in the world with a population of around 1.38 billion. Agriculture has been the source of livelihood. Around 70% of rural households depend on agriculture. India is the largest producer of pulses, milk, and spices. Along with these, rice, cotton, tea, sugarcane, wheat and various fruits, vegetables are also majorly produced. It is seen that around 195mha is used for cultivation of which 63% is rainfed and the remaining 37% is irrigated [2].

Over time, as the Indian economy has diversified, the country's overall GDP has been declining indicating growing concerns. Even though India houses leading producers, many people in India are suffering from hunger: around 190 million people are undernourished. As per the Global Hunger Index- 2017, it was seen that India ranked 100 among 119 countries. Indian agriculture faces many challenges in the overall development towards improving welfare of its rural poor like increasing agriculture productivity per unit of land, reducing rural poverty through social strategy comprising both agriculture and non-farming employment, and making sure that agricultural growth responds to food security needs [3].

1.3. Indian Policies: NITI Aayog and India Farm Laws

The agriculture and allied sectors of the NITI Aayog aim to transform Indian agriculture fuelled by technology and innovation for the overall well being of Indian farmers through sustainable and inclusive growth.

The core functions of this vertical are as follows:

1. Policy Design and Analysis
2. Research and Knowledge Exchange
3. Policy Advocacy
4. Stakeholder Consultations
5. Working with States
6. Working with Line Ministries

India is expanding rapidly, and so is the utilization of technology in the country's developing sectors. A significant chunk of the population is still reliant on agriculture as their primary source of income. Since independence, India has been at odds with agricultural infrastructure, techniques, and associated people [2]. With the agri-sector still contributing to around 15-20% of the country's national GDP over the last few decades, and its diverse and ever-changing needs across regions, India has been driving necessary and timely interventions at the industrial, organisational, and individualistic farmer layers to ensure its consistent transformation.

Given the different needs of farming in India, strategies or interventions to serve the greater farming community is essential. One constructive amendment to an existing policy or the introduction of innovative clauses to assist farmer's developing goals or aspirations may harm crop yields in the long run. Increased subsidies on fertilizer purchases by farmer groups, for example, may degrade soil health in the long run, impacting the cash flow a farmer may have to budget y-o-y. In this connection, the introduction of technology-driven aids such as soil testing facilities around the country has had a significant impact on farmers' crop selection, cropping pattern, degree of automation and irrigation approaches [4].

Small Farmers Agribusiness Consortium (SFAC) is an Autonomous Society promoted by the Ministry of Agriculture, Cooperation and Farmers' Welfare, Government of India. It was registered under Societies Registration Act XXI of 1860 on 18th January,

1994. SFAC is implementing the central schemes of the Government of India namely VCA and EGCGS for economic inclusion of small and marginal farmers in agribusiness activities. Society is a trailblazer in organising small and marginal farmers into Farmers Interest Groups, Farmers Producers Organizations, and Farmers Producers Companies to provide them bargaining strength and economies of scale. They serve as platforms for improving accessibility and reducing price of agricultural inputs for small and marginal farmers, as well as for developing forward and backward supply chain linkages. This effort has prompted farmer engagement for aggregation across the country, with the ultimate goal of a viable business model and augmented incomes [5].

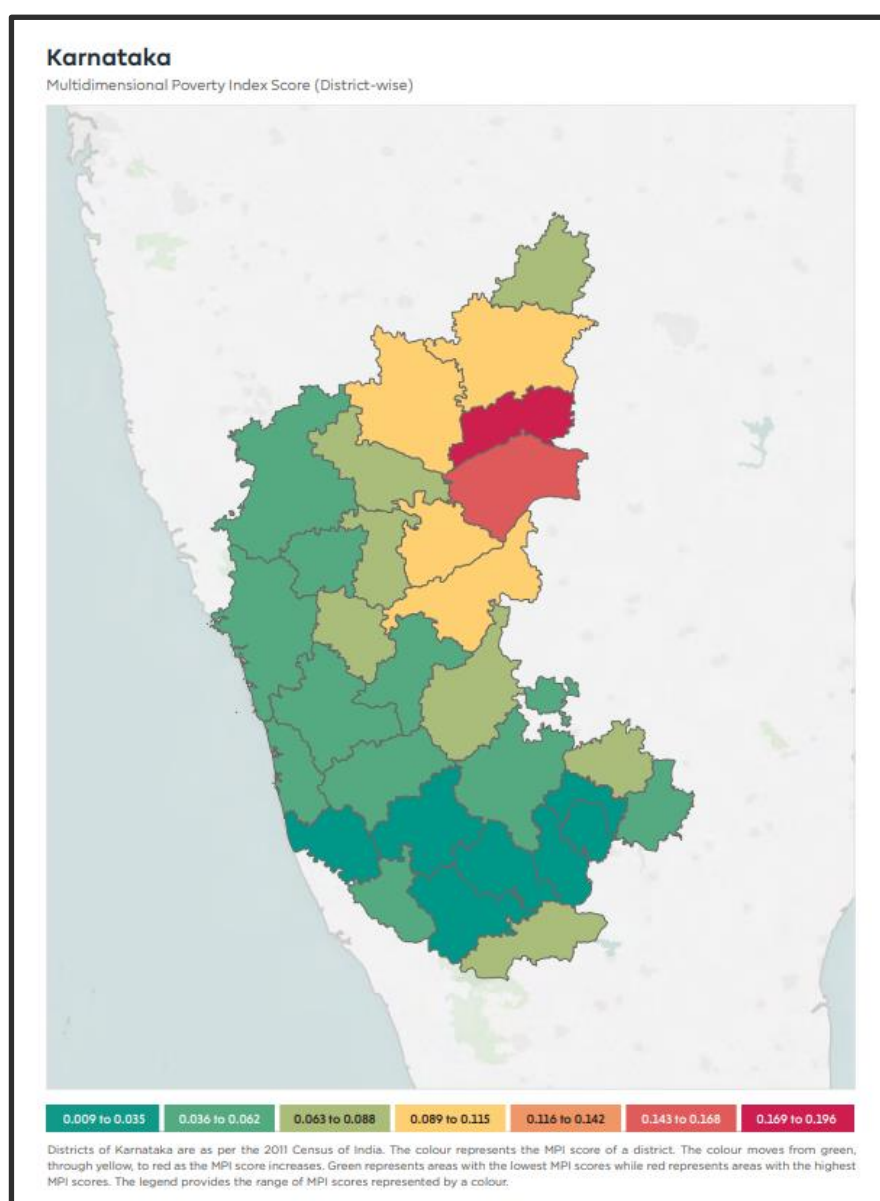


Figure 1.3.1. Multidimensional poverty index score in Karnataka

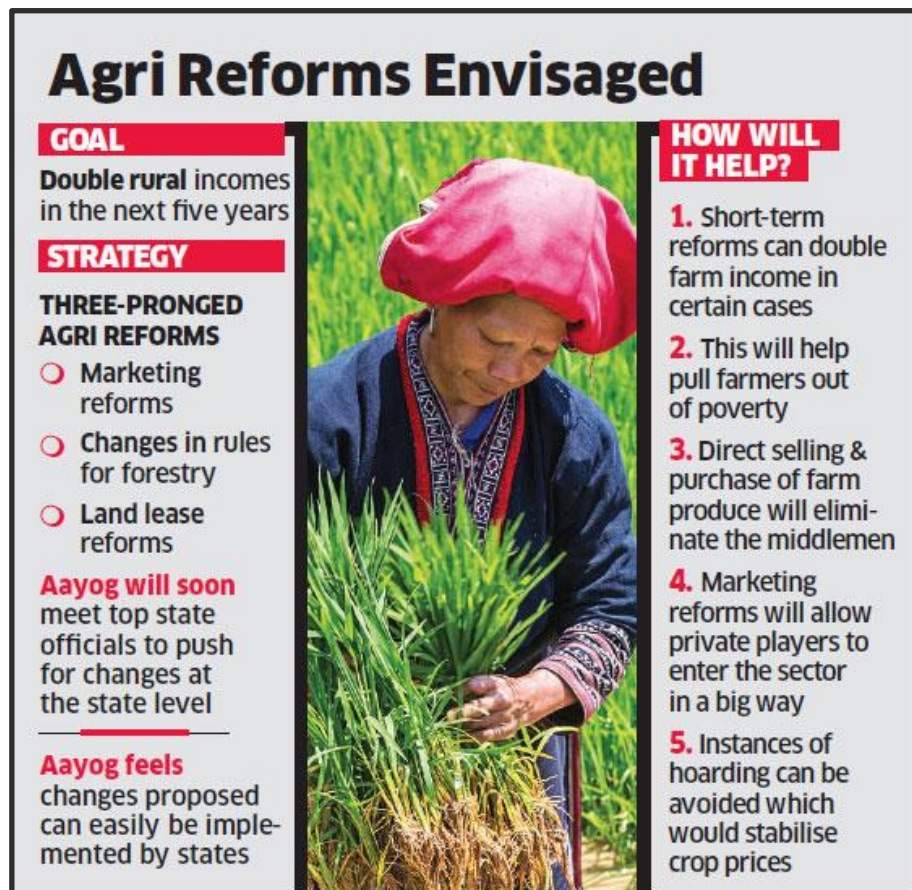


Figure 1.3.2. NITI Aayog's blueprint for reforms in the farming sector [4]

1.4. Integrating Technology with Agriculture

Farm mechanization is a strategy that has aided farmers all over the world, and it will do so in India as well. Over the years, the government has implemented a number of farm mechanization initiatives, which have contributed to certain advances in output and welfare standards. The Umang and mKisan programmes have also assisted with sharing crucial information to rural areas. Platforms like these have also supported farmers in receiving government benefits, subscribing to seed-fertilizer-credit incentives, and learning about other essential topics. However, in order to achieve exponential growth in agricultural output, a concerted effort encompassing the entire ecosystem is required, and the sector has seen the introduction of entrepreneurs into the agriculture industry.

Entrepreneurs' service-based models have encouraged farmers to boost their output and overcome a few weaknesses through consultations. This has also strengthened farmers' incomes, contributing to an improvement in their socioeconomic conditions. These disruptive models, which are transforming the face of the farmers' ecology, have been clearly acknowledged by the authorities and have demanded periodic changes in farming regulations. Startups are now present throughout the farming value chain, from sourcing

excellent seeds to providing mechanised farming tools, logistics help for crop movement, and improved sales channels [4].

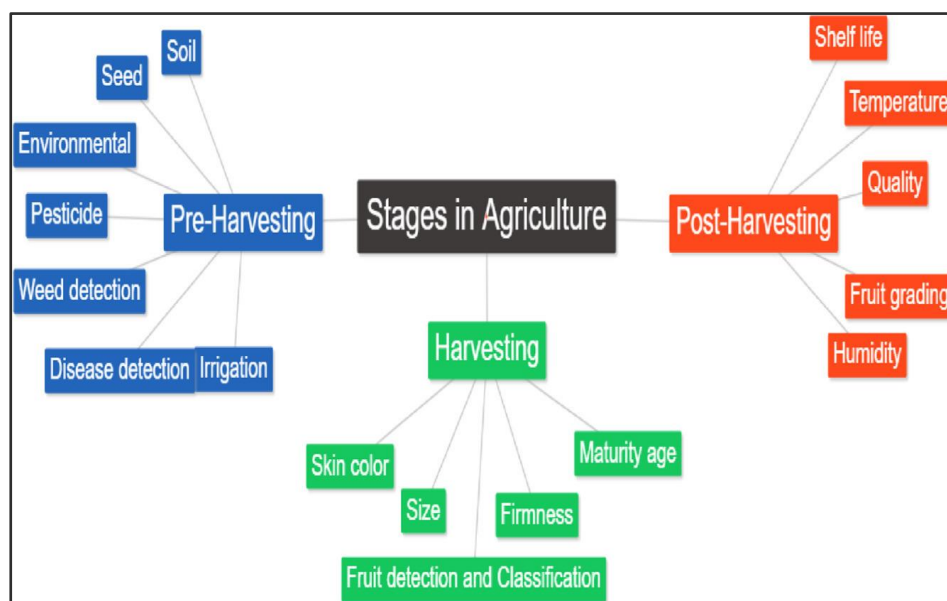


Figure 1.4.1. Stages in Agriculture: Opportunity for technology

In recent years, entrepreneurs have been able to offer farmers with integration of advanced solutions to boost their productivity by implementing new age technology like blockchain and artificial intelligence in the farming ecosystem. The use of such modern techniques has also presented entrepreneurs with new funding possibilities as they aim to focus on and solve the significant difficulties that the farming community will face as a result of climate change.

The ease of access to financing and the corresponding operational policies which would propel the entire value chain are significant variables that lead to growth in any sector. In recent years, there has been an undeniable shift in access to agricultural loans from non-institutional means to formal institutions, as compared to pre-2000 levels. Recognition of agricultural loans as a priority sector lending was a game changer in and of itself. This has had a significant impact on the overall loan outreach, and new age techniques with particular cum exclusive actions at the organizational level will boost access even further.

The use of digital technologies in smart farming has changed the way farmers cultivate crops and operate fields in recent years. It is not necessary to be an expert to

recognise how technology has altered the concept of farming, making it more economical, productive, secure, and accessible.

1. GIS-Based Agriculture

The use of digital technologies in smart farming has changed the way farmers cultivate crops and operate fields in recent years. It is not necessary to be an expert to recognise how technology has altered the concept of farming, making it more economical, productive, secure, and accessible.

The other significant advantage of GIS-based agriculture is the use of satellites and drones to obtain critical information from a bird's-eye view on plants, soil conditions, temperature, and topography. This type of data considerably increases decision-making accuracy.

2. Satellite-Derived Data

Predicting yields and doing near-real-time field monitoring with satellite data in operation to detect a number of risks has never been easier.

The sensors are capable of detecting information for indices like

- Normalised Difference Vegetation Index (NDVI)
- Canopy Chlorophyll Content Index (CCCI)
- Normalised Difference RedEdge (NDRE)
- Modified Soil Adjusted Vegetation Index (MSAVI)

3. Drone Data

Farmers may define harvest biomass, plant height, weed presence and water saturation on specific field regions with great precision with the help of drones. In comparison to satellites, they provide slightly more accurate information with high granularity. They deliver vital information considerably quicker than rangers when they are handled locally. Drones are also regarded as unequalled aides in the struggle against pests; the infestation is stopped by applying insecticide to prone regions with drones, all while minimizing the possibility of direct exposure resulting in chemical poisoning.

4. Online Data

A smart, digital monitoring system which enhances the decision making process and prevents missing crucial farming procedures, can be developed. NDVI which monitors the amount of chlorophyll in crops can be used to analyse the quality of crops. Farmers can accurately administer irrigation and avoid frost or heat loss by using meteorological data in conjunction with plant condition data collected from satellite vision. Drip irrigation with auto valve control, for example, is one of the greatest strategies for avoiding drought difficulties since the farmer can apply the needed amount of water to dry areas.

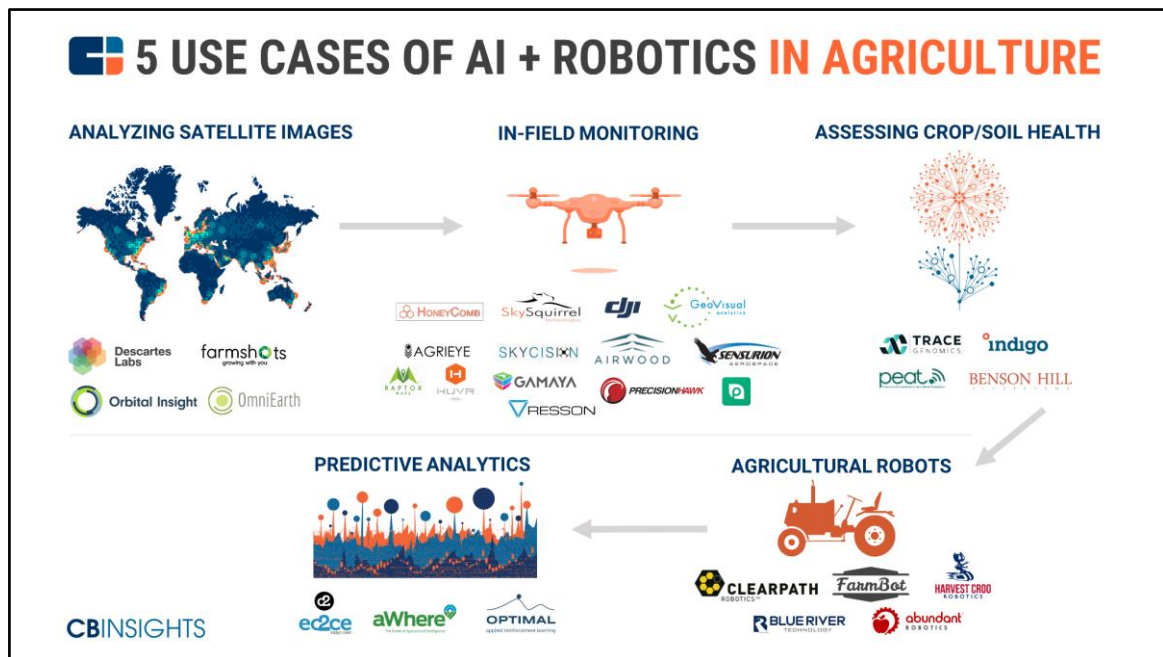


Figure 1.4.2: AI and Robotics companies in Agriculture

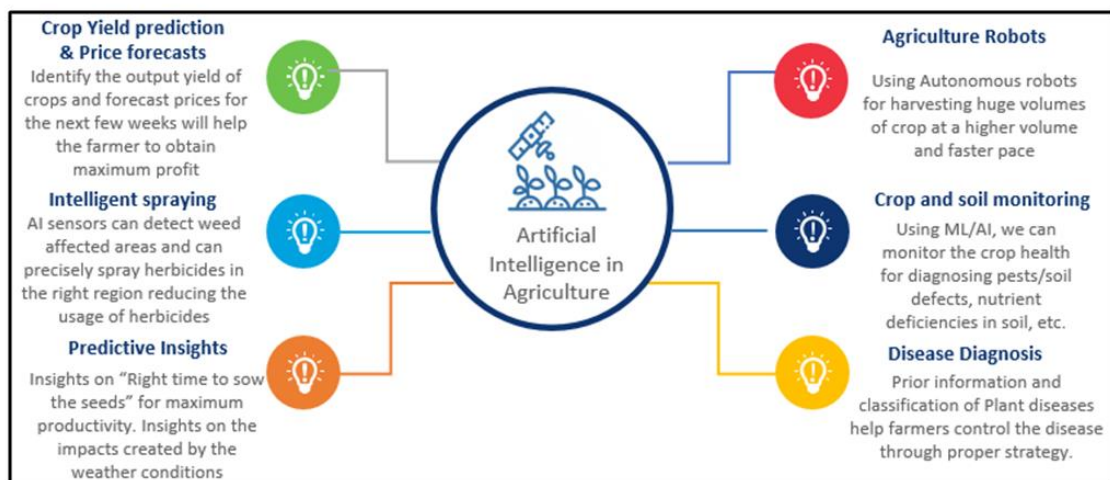


Figure 1.4.3: Artificial Intelligence in Agriculture

1.5. Typical Problems faced by marginal farmers

Marginal farmers are those farmers who own upto one hectare (~2.5 acres) of cultivable land [5]. The state of small and marginal farmers in India is very shocking. While 82% of Indian farmers fall into either one of these two categories, their land holdings account for only 44% of the total. In fact, one-third of India's farming community owns 0.4 hectares of cultivable land or less [6]. It is unfortunately ironic that while these farmers account for a larger proportion of working Indians, their incomes are meager and their contribution to the GDP is only around 18% [7]. The objective of this study is to identify the problems faced by marginal farmers of India, and contribute towards alleviating their struggles using forecasting techniques. This section discusses some typical problems faced by marginal farmers in India.

The most obvious adversity is that of debt and financial standing. Marginal farmers tend to lack the services that aid in marketability of their produce, eventually leading to lower margins earned per yield. This contributes to an accumulation of debt and inability to repay it, driving them to lower standards of living and agonizingly worse off though common- ending their lives. Lack of sufficient income also prohibits farmers from growing in the agricultural space: it prevents them from getting access to quality products like seeds, farming equipment, fertilizers, adequate market forecasts, etc.[8].

Middlemen play a major role in exploiting farmers. Due to misery, when farmers are left with no choice but to approach middlemen, it seems like even the slightest of their earnings are robbed away. In a similar extension, the government policies and farmer-benefit programs do not reach all farmers, particularly the marginal farmers who require the most assistance. Various schemes, loan and credit card policies introduced by the Indian government have unfortunately failed to reach the grassroot levels [9].

These financial factors have unfortunately led small and marginal farmers to switch to daily-wage labor as it provides them better and stable income than agriculture. In fact, as the land holding gradually increases from small and marginal farms to larger farms, the dependence of agriculture also increases. This in turn emphasizes that the Indian agriculture policies should address ways to increase farmer income and look at an adjacency expansion framework that adds additional value to the farmers' households [10].

With these on hand, there are multiple areas that may be addressed in order to contribute an infinitesimal but impactful value addition to the marginal farmer

communities. A lot of research has been carried out to identify several areas where technological advancements and intelligence may be used to develop solutions that assist farmers [11]:

1. Surveillance alert systems: Video surveillance systems that use AI-ML algorithms can distinguish between agents like people, vehicles, and animals. Video analytics helps detect any foreign agent entering the farmland and triggers alerts to the farmer regarding a potential breach. This assists marginal farmers with protecting yield and physical mishap.
2. Crop yield prediction: Using sensors and IoT devices provides agronomists and farmers with real time information about soil parameters, climatic factors and nutrition to name a few, in order to provide constraint-flexible intelligence to improve and increase crop yield, as ML algorithms have the capability to consolidate and work with large datasets.
3. Crop planning: Yield mapping using supervised algorithms to assimilate orthogonality and detect features, patterns and trends in the data. This thus informs the farmer about yield forecasts even prior to the sowing and harvest seasons.
4. Pest management: Infrared data from underground sensors and drone imagery assists with plant health monitoring and detecting pest infections and manifestations like that of locust plagues, even before they occur. Heat maps identify potential risk locations of the insect/ pest plagues.
5. Crop price prediction: This is critical in order to justify minimum support prices and pricing strategies based on the crop yield, to provide farmers with an estimate of the worth of their crops. By factoring the potential yield and demand for any given crop, farmers can manage their finances and plan suitable growing recommendations to maximize returns.
6. Irrigation management: AI-ML techniques can be suitably used to monitor irrigation rates according to the crop type and expected yields. Using various optimization techniques like linear programming, the optimum water-levels can be calibrated for utilization in environments that have scarce water availability.
7. Livestock management: Various parameters like livestock health, fodder quantities, daily activity and monitoring vitals can be measured using advanced AI-ML algorithms. Particularly in farms that use livestock, it is essential to understand the patterns of milk and manure production for farming purposes.

1.6. Developing Price Prediction Models to assist Marginal Farmers

Agricultural price prediction has been the focus of multiple research efforts to aid the marginal farmer communities globally. With the ever-increasing discovery and implementation of up and coming technologies and innovative solutions, price prediction models are constantly evolving. This study also aims to contribute to these efforts and develop a price prediction model that addresses seasonal agricultural commodities particularly in the markets of Karnataka. Various machine learning techniques have been used in the past in order to carry out a holistic predictive analysis for accurate price forecasting.

Price prediction is required to anticipate and cushion the impacts of fluctuating prices. Agriculture 4.0 marked the interconnection with Industry 4.0 which included the compilation of AI, ML, Big Data, Data Mining, etc. A survey paper [12] provides a thorough summary of agriculture price prediction, and from the various techniques surveyed, results show that more than 80% of the research uses time-series forecasting in AI-ML models. The specific types of models can be classified into the following categories based on the settlement methods:

1. Artificial Intelligence and Machine Learning
2. Data Mining
3. Regression
4. Stochastic Modeling
5. Dynamic Modeling

Additionally, based on the data type, the models can be categorized into the following forecasting techniques [13]:

1. Time Series
2. Qualitative
3. Causal Models

Chapter 2. Literature Review

2.1. Methodology

A systematic literature review of various publications obtained from search engines, extending from 2006-2021, was adopted to learn the applications of decision sciences, data analytics, forecasting and machine learning in the specific context of agriculture. An overview of our literature review is shown in Figure 2.1.1. A wide spectrum of papers and articles from journals of international reputation, government policy archives and white papers published by certain organizations, in order to understand the intensity of agricultural solutions being developed as an effect of the farmer crises. In order to ensure that validated research was studied, various criteria for selecting the papers were considered and are displayed in Figure 2.1.1.

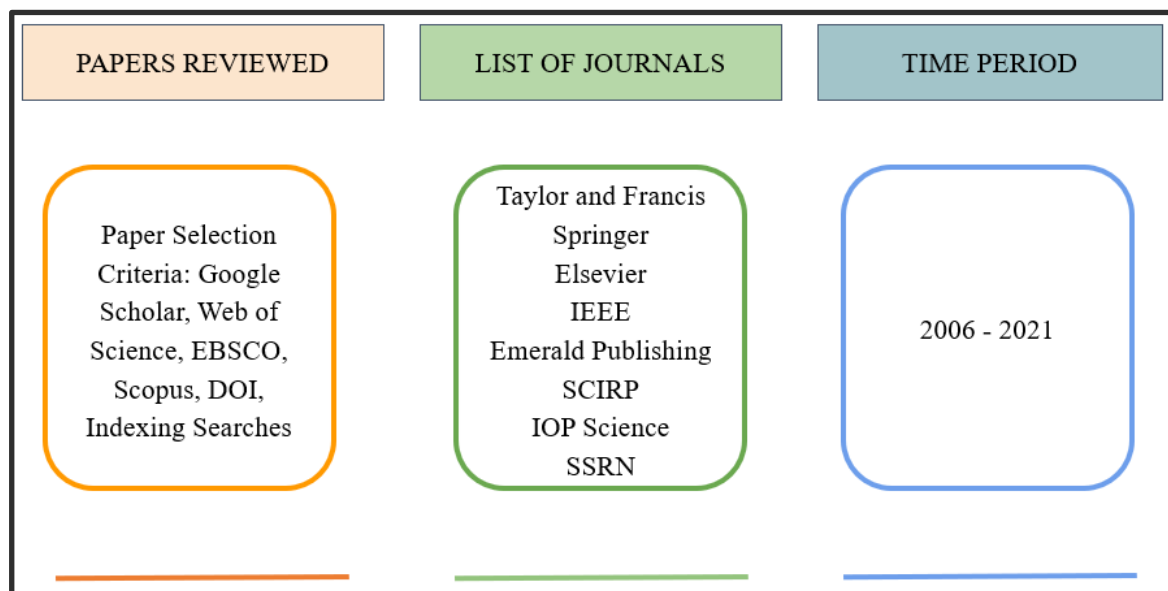


Figure 2.1.1. Literature Review Overview

The methodology followed to come out with an exhaustive literature review is shown in Figure 2.1.2. This methodology was adopted from a thesis on Conceptual Framework Analysis [14]. This integrative literature review and conceptual framework development approach is divided into a sequence of three interconnected phases:

1. Knowledge Acquisition: This phase consists of two iterations. In the first iteration, the initial screening of papers and articles was carried out on search engines like Google Scholar and Web of Science using generic keywords like ‘agriculture’, ‘farmers’, ‘technology for agriculture’, etc. The second iteration of knowledge acquisition followed

the first iteration of knowledge abstraction. Here, the screening was more specific with respect to this project's focus. Secondary literature search was carried out using more specific keywords like 'crop price prediction', 'ARIMA models for agriculture', 'seasonal commodities', 'price trends in mandis', etc.

2. Knowledge Abstraction: This phase also consists of two iterations. The first iteration of knowledge abstraction follows the first iteration of knowledge acquisition. A basic understanding of appropriate attributes with respect to the Indian agricultural scenario was completed and used as a basis for the second iteration of knowledge acquisition. The sequence of knowledge acquisition and knowledge abstraction continued even in the second iteration. The results of this phase included a specific assimilation of key pointers relevant to this study. A set of research papers and articles were thus compiled to carry out an exhaustive literature review.

3. Representation: In this final phase of Conceptual Framework, the findings from each of the papers and articles surveyed is consolidated into a form that is explainable and understandable by the reader. The representation of the findings from this study's thorough literature review is shown in Table 2.3.1 The various parameters that were considered to come up with these key findings is described in Section 2.2.

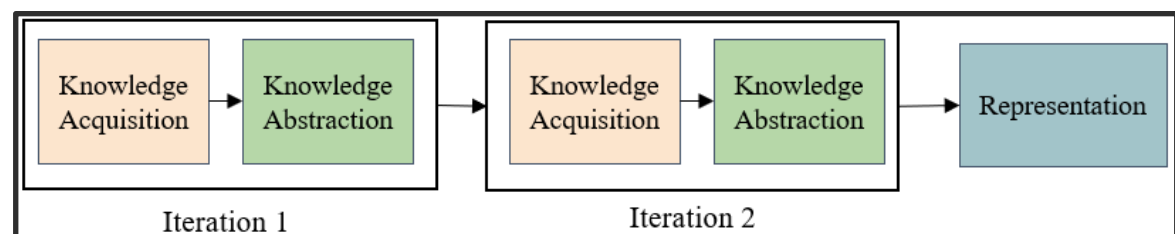


Figure 2.1.2. Integrative Literature Review and Conceptual Framework Development Approach

2.2. Parameters considered for Key Findings

The various parameters considered for research and analysis while developing the key findings from each research paper and article are listed below. A summary of the frequency of papers and articles using a given ML model to solve a given agricultural problem is pictorially represented in Figure 2.2.1.

1. Type of agricultural problem being addressed: price prediction, yield prediction, recommendation systems, marketplace design, etc.
2. Tools used for data analysis: Python, R, MATLAB, PowerBI, etc.
3. Technique employed and Justification: Machine learning models used for prediction like Linear Regression, ARIMA, Random Forests, etc.
4. Data drivers used as input for the study's analysis like historical price, yield, rainfall, temperature, etc.
5. Data preprocessing techniques used for data cleansing and curation for missing values and outliers like eliminating empty rows, collaborative filtering, normalization, etc.
6. Applicability to India and Crop type: The type of crops and datasets used for analysis and the geographical and logical applicability to the Indian context
7. Results: The metrics used to conclude the results of the algorithm like r-squared value, RMSE, MAPE, MSE, etc., and the performance evaluation of the results thus generated.

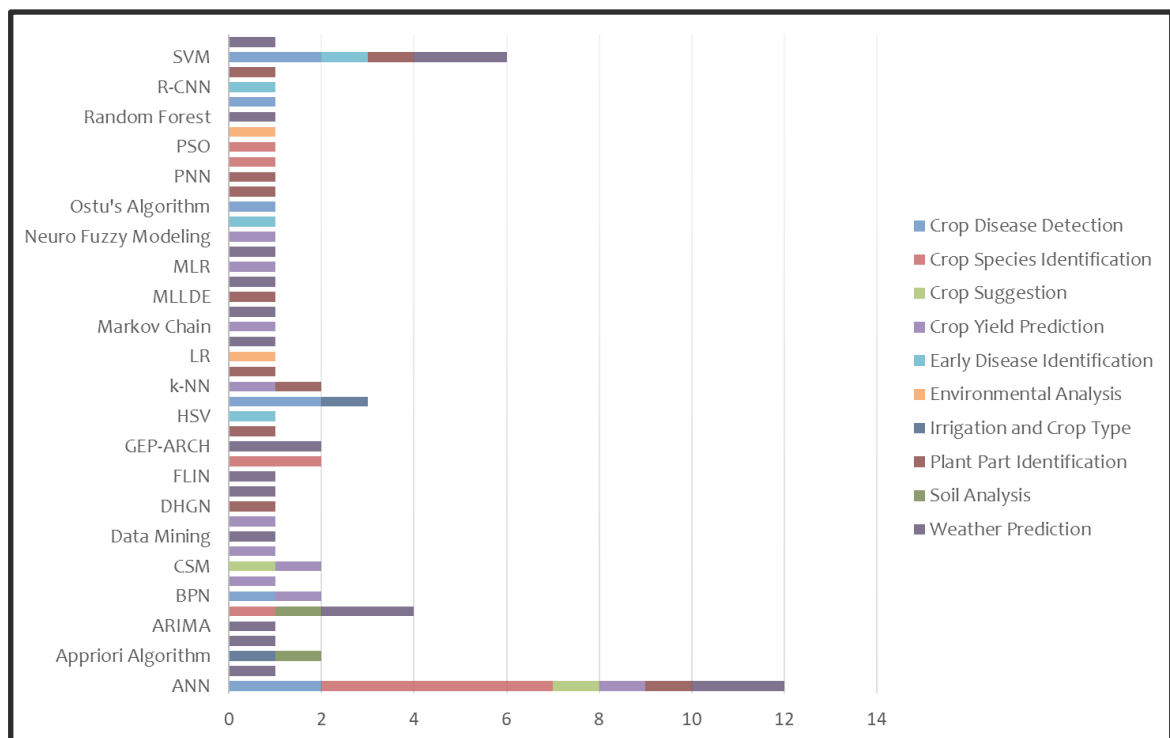


Figure 2.2.1. Frequency distribution of papers and articles addressing a given agricultural problem using different ML techniques

2.3. Literature Review

Using the methodology described in Section 2.1. and the parameters elucidated in Section 2.2., an exhaustive literature review was carried out. These papers have been analyzed across a 16-year period, and it is very interesting to infer the shift in tools and techniques from rudimentary technology to more intelligent systems. Key findings observed from these publications have been summarized in Table 2.3.1.

Table 2.3.1. Literature Review

Sl. No.	Title	Year	Author(s)/ Organizations	Key Findings
1.	ARIMA model and forecasting with three types of pulse prices in Bangladesh: a case study <i>Emerald: International Journal of Social Sciences</i> [15]	2006	Md Zakir Hossain, Quazi Abdus Samad et al.	At a very rudimentary level, this paper addresses the issue of crop price prediction for pulses in Bangladesh. SHAZAM was the software used to carry out this forecasting for three sets of pulses: motor, mash and mung, and ARIMA models were selected as the best models. Box-Jenkins methodology was adopted using six criteria for evaluation of the model: R^2 , $R\text{-bar}^2$, AIC, MAPE, Standard error, and BIC. Additionally, secondary time series data for a one-year period was used, which helped generate the best fit that followed the historical forecast, followed by ex-post and ex-ante. The monthly data used for this prediction did not contain any missing values, however, the scale of measure was converted from taka per maund to taka per quintal.

2.	Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications <i>International Journal of Computer Applications</i> [16]	2014	Manpreet Kaur, Heena Gulati et al.	This paper uses data mining for crop price prediction. MATLAB was used to develop the prediction model using naive and regression forecasting techniques. For a three year period, the prices of oil and tomato were subject to prediction. This paper specifies the type of preprocessing technique used. Data cleaning was carried out by linear extrapolation. The applications that use K-means need to be refined in order to be effectively used as a price predictor algorithm. This study was carried out for the markets of Coimbatore where a back-propagation neural network algorithm was also designed. This study acknowledged the role played by supplemental data sources like seasonal rainfall and economic indicators.
3.	Extension of MSP: Fiscal and Welfare Implications <i>NITI Aayog</i> [17]	2016	Jyoti Parikh, Chandrashekh ar Singh	This paper studies the impact of policy decisions, specifically, providing the Minimum Support Price (MSP) for crops on three critical societal goals: a) ensuring that the producers are protected against price fluctuations, b) influencing farmers to grow the produce of customers' choice and c) protecting end customers against surge prices. The study was conducted on two crops - rice and wheat and in two states – Uttar Pradesh and Madhya Pradesh. A

				simulation modelling was developed for this study. The findings point to clear benefits with the MSP offered by the Food Corporation of India to procure crop produce for supplying to consumers through public distribution system channels. Both the producers and consumers stand to gain in that, the MSP offer was more than what mandis offered and consumers had the benefit of sourcing the produce through public distribution system channels at the rationed price.
4.	Forecast of Agricultural Crop Price using Time Series and Kalman Filter Method <i>Asia Pacific Journal of Multidisciplinary Research</i> [18]	2017	Paulo V. Cenas	Crop price prediction is carried out in this research, using a combination of ARIMA and Kalman Filter algorithms. The crop of focus in this study was rice, and the data drivers included the four-year historic monthly prices of rice in the Philippines. In order to bring out a holistic view of the historic and current prices, ARIMA was used where the trend and seasonal indices were factored. Following this, the price prediction was carried out using the ARIMA and Kalman Filter ML models. The combination of both these models was designed in order to forecast future prices. The results were analyzed based on the RMSE values, which showed that the combination model had a lower RMSE score than the only- ARIMA

				model. An in- sample period was used for forecasts and an out-of-sample period was used to test the accuracy.
5.	Smart Farming System: Crop Yield Prediction Using Regression Techniques <i>Springer: Proceedings of International Conference on Wireless Communication</i> [19]	2018	Ayush Shah, Akash Dubey, et al.	Crop yield prediction is the central focus of this paper. Regression is the major technique employed here, using models like multivariate polynomial regression, support vector machine regression, and random forest models. In addition to historic yield data, humidity, rainfall, and temperature data were also considered to predict the yield of corn in Iowa, USA. This paper justifies the need for regression because of the usage of historical data. Given that yield is dependent on other multiple environmental parameters, multivariate polynomial regression has been factored. RMSE and r-squared values were used as the evaluation metrics, and the highest r-squared value was obtained for the support vector machine regression model.
6.	An interpretable produce price forecasting system for small and marginal farmers in India using collaborative filtering and	2018	Wei Ma, Kendall Nowocin et al.	This paper provides an extensive application of machine learning algorithms for crop price prediction especially addressed towards the small and marginal farmers in India. Decision tree based classifiers and ANN were employed to carry out this analysis based on the modal prices of commodities in Odisha from India's

	adaptive nearest neighbors <i>ACM: Proceedings of the tenth International Conference on Information and Communication Technologies and Development</i> [20]			Agmarknet portal. ANN was selected because of its ability to provide evidence for the given forecast, which in this case corresponds to predictors justifying the change in direction in a forecast. The data was analyzed over a six month period for which the random forests method was found to be the most accurate for prediction. This research also created a web application that is convenient and explainable to farmers. Crops considered in this study include mango, brinjal, tomato, green chilli, pointed gourd and cauliflower. Collaborative filtering was used as the data imputation technique.
7.	Development & performance evaluation of hybrid KELM models for forecasting of agro-commodity price <i>Springer: Evolutionary Intelligence</i> [21]	2019	Nirjharinee Parida, Debahuti Mishra et al.	The aim of this paper is to predict crop prices. This paper highlights the implementation of GWO-KELM machine learning technique on MATLAB using spot pricing as a major data driver. The price prediction was carried out for Chana and Barley based on the opening and closing prices, and the minimum and maximum prices for a six year period from the datasets available on India's NCDEX portal. The hybrid GWO-KELM model was chosen in order to result in an improved prediction. The mean squared error was used to test the results of the model. Statistical parameters used in this simulation

				include the measures of central tendency and normal distribution.
8.	Agricultural commodity futures prices prediction via LSTNet <i>Taylor and Francis: Journal of Applied Economics</i> [22]	2019	Hongbing Ouyanga, Xiaolu Weia and Qiufeng Wu	With the objective of predicting crop futures prices, this paper assesses the utilization of LSTNet neural networks. The key data drivers in this study are solely historic prices of daily, weekly and annual closing prices. LSTNet was selected as the ML technique because of the pattern of the futures prices comprising both long and short term information and linear and non-linear trends where ARIMA, VAR is not optimum. This was proved by the results obtained where LSTNet showed better forecast performance mainly because it factors interdependencies among all variables. The metrics for analyzing this forecast performance included using RSE, CORR, and Fluctuation Tests. This paper mentions the data preprocessing techniques employed- dropping null values, and splitting 60% of the dataset for training 20% for validation and 20% for testing.
9.	Seasonal crop yield forecast: Methods, applications, and accuracies <i>Elsevier:</i>	2019	Bruno Basso, Lin Liu	This paper focuses on crop yield forecasting. Remote sensing data from IOT and sensors was extracted for primary data drivers being temperature, precipitation, humidity and soil maps. This data was used to carry out a linear regression analysis.

	<i>Advances in Agronomy</i> [23]			The accuracy of this analysis was measured using the r-squared values between the predicted and actual yields. This paper's applicability to India is gauged as it addresses food security issues, and directly considers the forecast of rice, maize and wheat for weekly periods in various districts in India.
10.	Decision Making Support System for Prediction of Prices in Agricultural Commodity <i>IEEE: Amity International Conference of Artificial Intelligence</i> [24]	2019	Aman Vohra; Nitin Pandey; et al.	Agricultural commodity price prediction has long been a serious issue for farmers. Farmers are unable to obtain the desired price for their crops, which is why the suicide rate rises with each passing year. This research primarily proposes a decision-making assistance model for agricultural commodity price prediction. This article also offers agricultural data mining approaches that will assist farmers in forecasting agricultural commodity prices. The goal of this work is to create a system with enhanced efficiency of price prediction characteristics. The goal is to present a conceptual platform and build a system that will allow us to make some progress toward more accurate price prediction.
11.	A Framework for Crop Price Forecasting in Emerging	2019	Ayush Jain, Smit Marvaniya, et al. : IBM	Crop price forecasting accuracy is crucial because it allows distribution network strategists and government agencies to take appropriate measures

	Economies by Analyzing the Time-series Data <i>arXiv.org</i> [25]		Research- India	by anticipating market elements like demand and supply. Crop prices in marketplaces in emerging nations such as India are keyed every day, which can lead to human-induced errors such as the entry of false data or the submission with no data for many days. Taking into account the historical crop prices, historical weather data and data quality related features, this paper demonstrates the efficiency of Tomato and Maize in fourteen market places. The proposed models not only produce accurate results but end up with a more robust model.
12.	Crop yield prediction using machine learning: A systematic literature review <i>Elsevier: Computers and Electronics in Agriculture</i> [26]	2020	Thomasvan Klompenburg, Ayalew Kassahun, et al	Machine learning is a key guide that helps for forecasting crop prices, including advising on which crops to cultivate during the crop's growth season. A variety of machine learning techniques have been used to aid crop production prediction studies. Using data drivers like temperature, rainfall, soil type an Artificial Neural Network model is applied. An extensive study of around 50 electronic databases to identify deep learning algorithms was conducted. Convolutional Neural Network is by far the most used algorithm which forms the basis of a crop prediction deep learning model.
13.	Time series	2020	Koichi	As the prices of crops show a

	forecasting of agricultural product prices based on recurrent neural networks and its evaluation method <i>Springer: SN Applied Sciences</i> [27]		Kurumatani	seasonality trend over the years, there has been a wide acceptance of the fact that recurrent neural networks with machine learning capability, have a more accurate time series forecast than the conventional methods. Error rate calculations are as important as the criterion for conservation of the statistical characteristics of the probability distribution function from the original past time series to the future time series in the forecasted future. The time alignment of time point forecast (TATP) running on Long Short Term Memory (LSTM) has greater efficacy and conservation. On a concluding note, it can be confirmed that the forecasting models are in direct relation to the evaluation criteria and the combination of these models would be used in the practical application where continuous datasets are unavailable.
14.	Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala <i>Elsevier: Procedia Computer</i>	2020	Kiran M.Sabu, T.K. Manoj Kumar	This study begins by acknowledging the psychological trauma that farmers face due to fluctuating prices, both monetarily and mentally. By addressing risk mitigation in times of such price fluctuation, along with climatic uncertainty, this paper highlights the significance of predictive analytics for price prediction of Areca Nuts in Kerala. Recent trends in shifts of farmer

	<i>Science</i> [28]			proclivity because of the seasonal nature of areca and its associated price fluctuations forms the basis of this paper. Time series forecasting is used to develop a machine learning model using monthly historical prices. The RMSE is used as the evaluation metric for a variety of ML models: SARIMA, Holt-Winter's Seasonal Method and LSTM neural networks, for a ten-year period, of which the LSTM model displayed the best results.
15.	A Study On Agriculture Commodities Price Prediction and Forecasting <i>IEEE: International Conference on Smart Technologies in Computing, Electrical and Electronics</i> [29]	2020	Girish Hegde, Vishwanath R Hulipalled, et al.	This paper addresses the problem of price prediction using human computer interaction, specially in the area of speech recognition. By directly addressing the Indian farmers, this paper proposes a solution that would assist farmers in the regional language. The current policies of the Indian government are briefly discussed, highlighting how the current price forecasting is carried out at the central level using historical data on short-term yield intervals. But the lack of evidence-based justification prevents the farmers from getting valid justification for suitable recommendation and planning. A holistic survey is carried out in the price forecasting domain in this paper by underlining the advantages and disadvantages of various models currently used to address this

				agriculture problem.
16.	Kisan Mitra: A Data Analytic Based Prototype for Avoiding Crops Price Crash SSRN: <i>Proceedings of the International Conference on Innovative Computing and Communication</i> [30]	2020	Akshit Kumar, Binayak Chakrabarti, et al.	As per the recent by the government, there was a fall in price in the financial year 2018-2019 this adversely affected the farmers income the most in twenty years. The Machine learning model for prediction and analysis is deployed on the cloud for faster execution. This paper shows the various factors which will be used during the process of building the model and thus, leads us to the inference that the importance of both external factors (international stock price, exports, imports) and internal factors (production, yield) must be known to have a clear and accurate prediction.
17.	Crop Value Forecasting using Decision Tree Regressor and Models <i>European Journal of Molecular &</i>	2020	Akshay Prassanna S, B A Harshanand, B Srishti et al.	A ML is an emerging field in research which can be used for crop price prediction and providing solutions accurately. Comparative study has been between random forest and decision tree regressor and performance of each has been observed. Data was collected from data.gov.in which had a dataset of 23 crops. Each file contains rainfall and

	<i>Clinical Medicine</i> [31]			WPI from 2012 to 2018. When the accuracy of both models were compared it was seen that Decision Tree 92.6 percent and Random Forest Ensemble Learning 97.57 percent
18.	A Comparative Study of ARIMA and Holt-Winters Exponential Smoothing Models for Rice Price Forecasting in Tanzania <i>Open Access Library Journal</i> [32]	2021	Yohana James Mgale Yunxian Yan, et al.	High crop price fluctuations have a great influence on making production decisions. It becomes difficult for the producers to estimate the price or to adjust the timing of sale. This not only affects the producers but also the poor consumers as they have to reduce the food consumption because of increase in price. In addition, it has a great impact on economic growth and socially. ARIMA and Holt winter exponential smoothing models were used to forecast the price, showing time series data for both the trend and seasonal variation and a comparative study was made between them. The data used for this forecasting was the monthly rice prices in the market in the region Mbeya from 2004 to 2019. By comparing the predicted value with the last nine observations of nominal rice prices, it was seen that Holt- winters exponential model was closer to the actual value than ARIMA, with Mean Absolute percentage error of 5.3314. Thus, with this a conclusion was made saying that Holt- winters exponential model is more efficient in predicting the future

				price values for agriculture products.
19.	Crop price prediction using supervised machine learning algorithms <i>IOP: Journal of Physics: Conference Series</i> [33]	2021	Ranjani Dhanapal, AjanRaj et al.	Due to price fluctuations in the market, farmers are facing immense losses, this not only affects the farmers but also there is a reduction in the country's GDP. Variation in crop prices in the market leads to demand fluctuations and the market value of that crop. Reduction in the value of the crop makes farmers suffer from investment losses. Decision tree regressor, A machine learning technique was used to predict crop prices for the coming twelve months. Whole price index (WPI) and rainfall per month data was collected for prediction. It was found that decision tree regressor algorithms gave the most accurate and efficient predicted results. By testing it with different algorithms it was seen that regressor models' root mean square value error was 3.8 which was less when compared with others. From the obtained it was concluded that out of 20 different crops chosen for prediction, crops soybean, Niger, ragi, copra and arhar were top gainers and crops jute, barley, sunflower, moong and Bajra were top losers. A time series forecast graph of previous years and next year's prices was shown.
20.	Crop yield	2021	Pallavi	Estimation of crop production is a

	forecasting using data mining <i>KeAi: Global Transitions Proceedings</i> [34]		Kamath, Pallavi Patil, et al.	difficult task in our country mainly because of our population. Most of the farmers rely on their long-term experience to predict crop yield for the coming season, but this prediction isn't reliable. Hence there is always a need for crop yield forecasting. Crop production prediction helps farmers in advance to make market related decisions. A supervised classification technique called random forest was used to predict crop yield. Dataset on climate variables like humidity, temperature, and soil characteristics such as loamy, chalky, clay and so on and also the location or the area for growing was collected for the prediction. Random forest Is a combination of decision trees, where a decision tree was constructed for each individual set of training dataset and average of it was taken and further predictions were made based on it. A website was created containing four modules i) crop module ii) soil module iii) weather module and iv) predict module. With the help of these modules' farmers can easily get to the estimated crop yield.
21.	Onion price prediction for the market of Kayamkulam	2021	Anubha, Tripathi, Kshitiz Kumar, et al.	Due to variation in production and market arrival time there is always a fluctuation in the price of onions. Thus , there is always a need for forecasting onion prices. Here ARIMA model is

	<i>Springer: Data Analytics and Management</i> [35]			used for predictions of annual prices of onions in the market of Kayamkulam, Kerala. modal prices of big varieties of onions in this region was collected from data.gov.in website. ARIMA eliminates the non-stationarity of data used to give a prediction of an existing time series based on past values. Also shows the hidden trend of the moving average pattern. It was found that there was a Close relationship between actual and obtained values. MSE value was found to be around 64,394, which is considered to measure accuracy.
22.	Crop prediction based on soil and environmental characteristics using feature selection techniques <i>Taylor and Francis: Mathematical and Computer Modelling of Dynamical Systems</i> [36]	2021	A. Suruliandi, G. Mariammal et al.	Choosing the right crop for cultivation based on soil and environmental characteristics by manual by farmers always resulted in failure. Here a machine learning technique was used to make a comparative study on wrapper feature selection techniques combined with classification techniques to predict the best suited crop for cultivation on that land. Soil and environmental data were collected based on which the prediction was made. It was concluded that feature selection techniques Boruta, SFFS and RFE combined with bagging classifiers showed the best results, RFE With bagging classifier outperforms other combinations.

23.	Application of the decision tree method for predicting the yield of spring wheat <i>IOP: Earth and Environmental Science [37]</i>	2021	V K Kalichkin et al	Predictive decision tree model was developed to predict the yield of spring wheat. This study used previous years wheat yield data (2001 - 2019). Some qualitative factors like level of intensification, tillage system and quantitative factors like temperature and precipitation were taken into account. based on the obtained result accuracy was calculated for training and testing samples. average absolute error - 3.455 (training sample) and 4.446 (test sample); determination coefficient - 0.895 (training sample) and 0.811 (test sample).
24.	An intelligent Crop Price Prediction using suitable Machine Learning Algorithm <i>EDP Sciences: ITM Web of Conferences [38]</i>	2021	Ishita Ghutake, Ritesh Verma et al	Selection of crops has been a tedious task for farmers as it is difficult for them to predict the prices of the crop for that particular season which is always dependent on the weather. This problem can be solved using ML algorithms to predict the future crop prices. Thus, helping farmers to select the right crop and gain more than usual. This study considered wheat crops for prediction using two regression algorithm techniques - decision tree and random forest and result was analysed. With the successful prediction it was seen that the obtained result was 92 percent accurate and aiming to increase the accuracy to 97 percent by considering

				more features in addition to rainfall like temperature, humidity and soil fertility for better prediction.
25.	Analysis of minimum support price prediction for Indian crops using machine learning and numerical methods <i>Springer: International Conference on Innovative Computing and Communications</i> [39]	2021	Sarthak Gupta, Akshara Agarwal et al.	Ensuring that farmers receive the optimum threshold of minimum support prices is the key focus of this paper. Here, a combination of machine learning tools and statistical numerical methods have been considered, and a holistic comparison between predicted and actual values is evaluated. The methodology employed includes: linear regression, recurrent neural networks, Newton's forward and backward interpolation, Lagrange's technique and exponential smoothing, using data extracted from the Indian government's Ministry of Agriculture and Farmer welfare for 32 crops across 16 years. Results indicated that linear regression, Newton's interpolation and Lagrange's technique present the least MAPE values for the prediction of MSP.
26.	Farming made easy using machine learning <i>IEEE: International Conference for Emerging</i>	2021	Manasi Jadhav, Neha Kolambe et al.	Crop price prediction using Decision Tree Regressor algorithm is the central focus of this study. By clearly elucidating the primary and secondary data sources, this paper highlights the parameters that drive the forecast: Rainfall, wholesale price index, monthly and annual price and yield

	<i>Technology</i> [40]			data. This futuristic forecast aids the farmer with financial planning, and can incorporate additional segments like fertilizer and crop recommendations along with weather forecast. A web application was developed using Naive Bayes, k-nearest neighbor and decision trees. Python was employed as the tool to carry out the ML model. Decision trees provided the best results in terms of addressing the issue of overfitting, thereby increasing the accuracy of the model significantly. Three metrics were used for performance evaluation: MAE, MSE and r-square value. Paddy's data was analyzed, and a testing accuracy of greater than 90% was achieved.
27.	Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques <i>arXiv.org: Advances in Science, Technology and Engineering Systems</i> [41]	2021	Zhiyuan Chen, Howe Seng Goh et al.	This paper proposes a dynamic web application for commodity price forecasts in Malaysia using commonly known ML algorithms like ARIMA, SVR, Prophet, XGBoost and LSTM. Like many papers discussed, this study also reveals that LSTM produces a better result for long term predictions with a relatively low MSE of 0.304. Another aspect discussed here is the added benefit of a price prediction model for planning the plantations. The configuration of each of these models is described. The non linearity of these datasets is

				considered while using these ML techniques. The crops analyzed in this study include chillies and tomatoes. ARIMA shows the highest stability for short term prediction with an MSE of 0.25, and it has the ability to deal with small amounts of data for forecasting. However, when additional parameters are considered like temperature and humidity, LSTM performed better for long-term predictions.
--	--	--	--	---

As it can be observed from this thorough literature review, there has been a pressing focus on integrating smart technology with agriculture for more than a decade. Beginning with rudimentary functionalities on tools like SHAZAM, all the way to establishing neural networks in tools of today like Python, R and MATLAB, solutions to agricultural problems have been ubiquitous. The trends in machine learning algorithms also indicate a steady implementation of models employing ARIMA, regression, LSTM, decision trees and random forests, to name the most frequently used. Time series forecasting was the basis for most price prediction models, and crops like paddy, chilli, tomato, onion, brinjal were studied. Based on the findings from this literature review and the kind of data available from sources mentioned in Chapter 5, this study will identify the types of seasonal crops specific to Karnataka that may be used for testing and forecasting.

Chapter 3. Problem Definition

Agriculture contributes 17 percent of India's GDP. It is seen that almost around 80 percent of the farmers are small and marginal farmers, owning less than 2 hectares of land [2]. Income of such farmers are usually from the profit they make out of their produce, because of which they generally tend to sell their commodities at high prices. However, as these prices reach unreasonably high levels, it is seen that there is an increase in the risk like leftovers of commodities and unsold produce which are unfit for consumption. This brings farmers back to square one, of selling their commodities at meager prices to avoid cost of facilities like cold storage, market price forecasting, etc. In order to assist farmers with feasible price prediction solutions, this project proposes a machine learning model to predict crop prices of seasonal commodities in markets of Karnataka.

3.1. Objectives of this Study

The main aim of this study is to contribute towards research intended to promote sustainable agriculture which would aid in rural development and help the farmer-communities, especially in India. By designing and developing an intelligent model that would aid in the price prediction of seasonal crops in order to assist the agriculture supply chain in decision making towards minimizing risk caused by crop price fluctuations. This would also aid in financial planning, and make crop prices transparent to the farmer well in advance before the season, thereby assisting them with suitable crop planning and recommendations.

The central focus of this project is to closely predict crop prices for seasonal crops in Karnataka based on the data available with respect to previous trends in weather, crop yield, soil data, wholesale price index data, price, MSPs, consumption trends, etc. using suitable machine learning techniques in forecasting.

Chapter 4. Methodology

The methodology followed in this project is represented below in Figure 4.1., and is explained as follows. This methodology began with identifying the area of focus, followed by a specific literature review system to analyze and infer from the key findings.

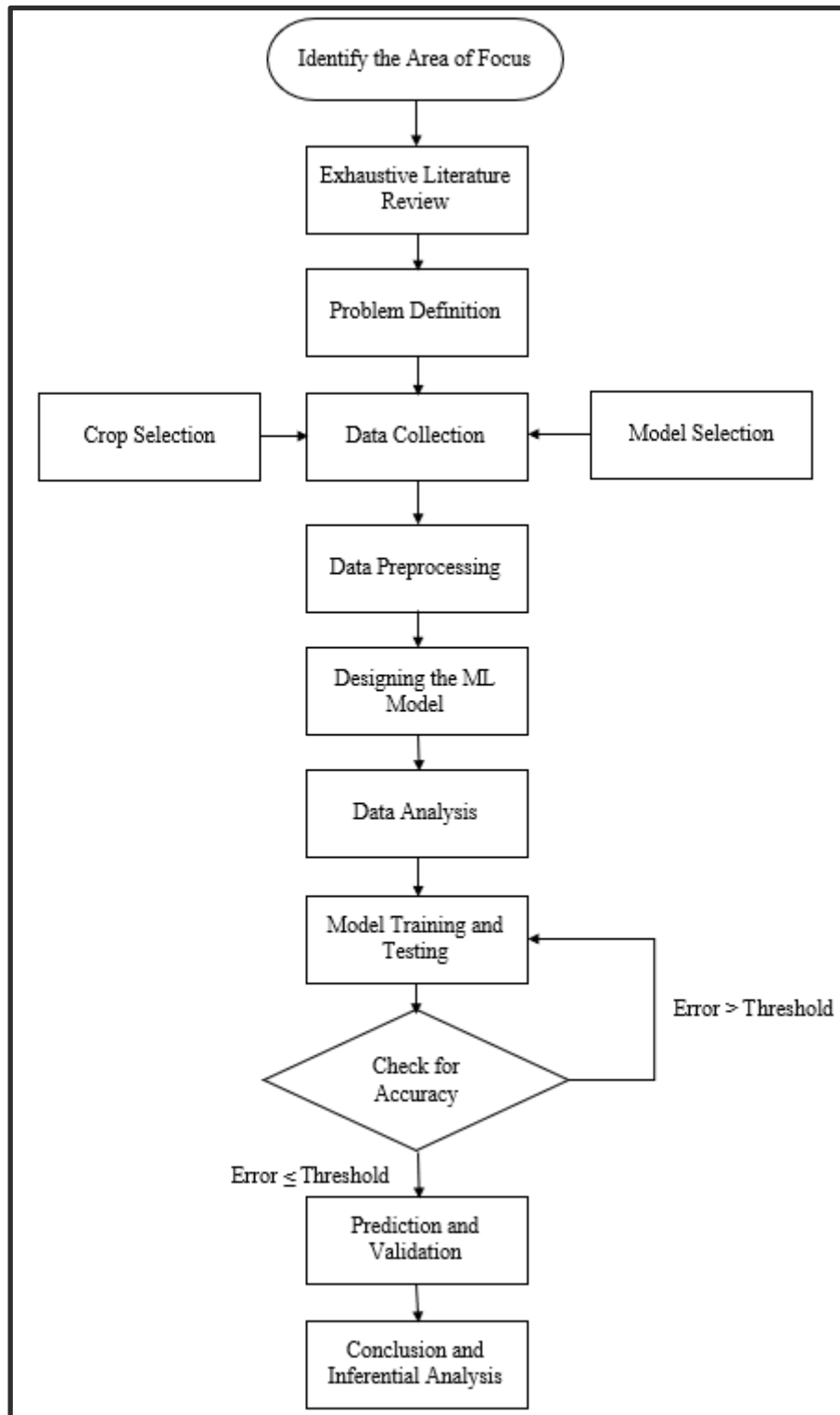


Figure 4.1. Methodology flowchart of this project

1. Identifying the Area of Focus

High crop price fluctuations have a great influence on making production decisions. It becomes difficult for producers to estimate the price or to adjust the timing of sale. Hence there is always a need for crop price forecasting. Crop price prediction helps farmers in advance to make market related decisions.

2. Literature Review

More than 25 literature papers and policy studies were collected from various conference venues and various reputed journals like Taylor and Francis, Emerald, ScienceDirect etc. A detailed study was done on these papers and a clear understanding of the problem was obtained.

3. Problem Definition

The problem was specifically defined in this step: As prices become unreasonably high, it is seen that there is an increase in the risk like leftovers of commodities and unsold produce, which are unfit for consumption. This brings farmers back to square one, of selling their commodities at meager prices to avoid cost of facilities like cold storage, market price forecasting, etc. In order to assist farmers with feasible price prediction solutions, this project proposes a machine learning model to predict crop prices of seasonal commodities in the markets of Karnataka.

4. Data Collection

Data is gathered from various sources like Krishi Marata Vahini and Agmarknet portals, based on which the type of crop may be selected along with the selection of ML techniques like ARIMA, Decision Tree Regressor etc.

5. Data Preprocessing

Data obtained from these data sources mentioned above is raw and includes many null values. Hence, there is always a need to filter these redundant data and fit the existing data. In data preprocessing, data is cleaned, normalised and checked if all the data types are valid.

6. Designing the ML Model

An ML model is to be designed based on which crop prices prediction is done for the selected crop.

7. Data Analysis

Data exploration also known as E.D.A (Exploratory Data Analysis) is carried out, where based on the data model designed, analysis is made on the trends or significant characteristics of the designed model and conclusion is drawn.

8. Model Training and Testing

Designed model is made to run to obtain the result, based on which the accuracy of the model is checked by various performance metrics like mean square error (MSE), mean absolute error (MAE) etc. If the error depicted is greater than the threshold value then the model is made to train again, else the obtained result is used further to draw conclusion

9. Prediction and Validation

Result that is obtained from the previous phase is tested before the conclusion is drawn. Validation is done by comparing the predicted crop price result with the actual market price data. If the comparison has less variations, then the result obtained is valid for further inferential analysis.

10. Conclusion and Inferential Analysis

Final phase of this process is to draw conclusions from the obtained results and infer is made on the findings from this model. Other ways to assist farmers may be calculated by enhancing this model to solve problems like recommendations and sowing patterns.

With this structured methodology, the project has been scheduled and the roadmap is described in Chapter 6. With respect to Phase-1 of this project, steps 1-4 were completed. Beginning with step 5, data needs to be cleaned and curated, followed by a thorough analysis and model design based on the type of crop and ML technique. The prediction results need to be evaluated against the actual values to establish accuracy.

Chapter 5. Data Collection

There are various data sources that this project will make use of, in order to extract data with respect to primary price and yield data, and secondary data with respect to temperature, rainfall, soil moisture, etc. These data sources are listed below:

1. **Agmarknet** [42]: This is the Government of India's official agriculture marketing portal that contains information about commodity and market profiles, price and arrival data for a wide range of agricultural commodities with district-wise, market-wise and period-wise granularity.
2. **Open Government Data Platform of India** [43]: This is India's open government data platform catering to various sectors with 100,000+ datasets solely addressing the agriculture space with various APIs, visualizations, services and schemes. In addition to the price and yield data, additional supplemental datasets like rainfall, soil moisture, etc.
3. **Krishi Marata Vahini** [44]: This is Karnataka's online agricultural marketing information system maintained by the Government of Karnataka's Department of Agricultural Marketing Board. This portal hosts data about prices and yields of a variety of commodities for multiple markets within Karnataka on daily, monthly, and yearly periods of granularity.
4. **Coffee Board of India** [45]: This portal is the Government of India's official coffee board managed by the Ministry of Commerce and Industry. Reports are generated on a seasonal basis with state and market-wise yields and production quantities.
5. **Karnataka State Natural Disaster Monitoring Center** [46]: This portal is the government of Karnataka's weather, wind and rainfall portal with annual, monthly, weekly state-wise and district-wise reports.

These sources serve as authentic data repositories for primary data points like historic prices and arrival quantities, and secondary data points like rainfall, temperature, humidity, soil moisture, etc. Given that these datasets contain a lot of missing values and outliers, suitable techniques for data imputation and preprocessing are to be employed to obtain feasible datasets for analysis.

Chapter 6. Project Roadmap

This project has been structured sequentially in order to achieve the goals required at every stage. The Gantt chart of the project roadmap is shown in Figure 6.1. Beginning from the second half of the month of August, 2021, this project timeline has been divided into half-monthly targets, with a tentative completion scheduled in H2- May, 2022. As the chart indicates, each member of the team has been assigned specific tasks, and tasks that require group synergy have also been identified.

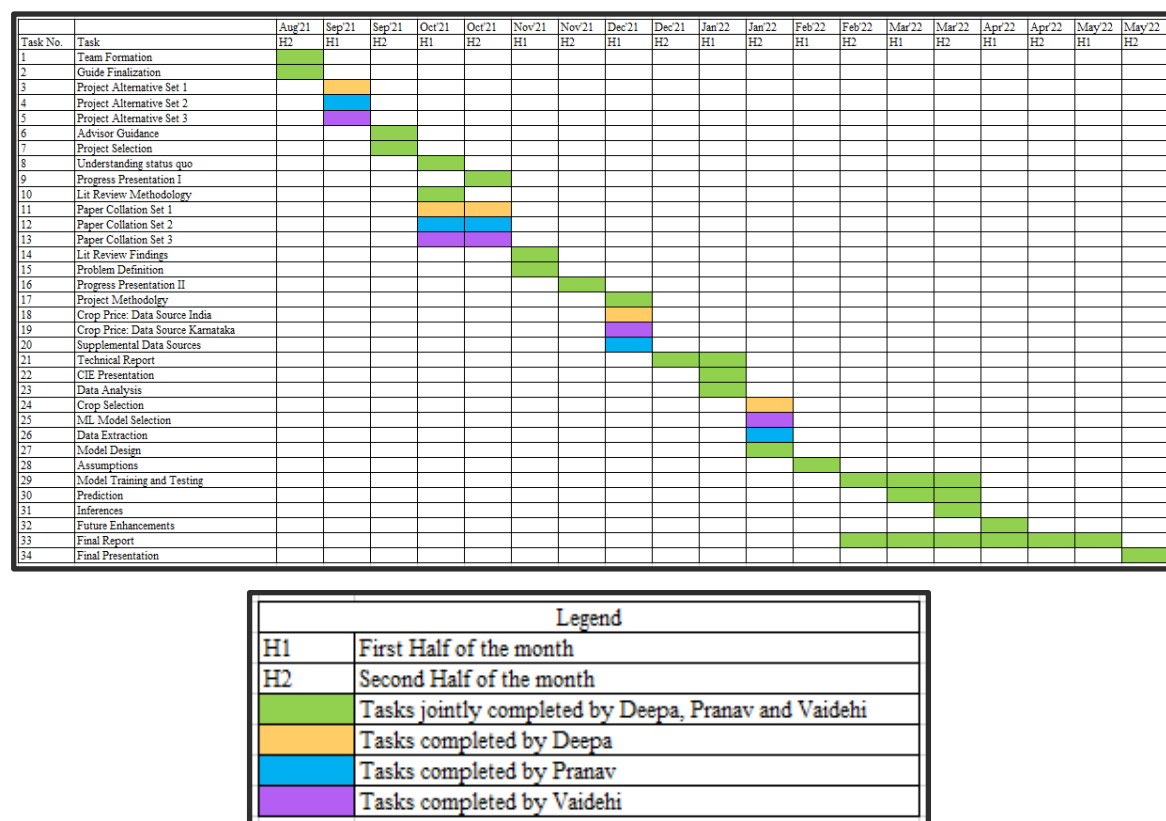


Figure 6.1. This project's Gantt Chart

The successful completion of Phase-I of the project comprised of team and guide selection, problem identification- its definition and objectives, researching a framework for literature review and carrying out an exhaustive literature review to understand the current state of the art applications of machine learning in agriculture, followed by identifying potential data sources from credible sources. Going forward towards Phase-II of the project, the analysis and extraction of data is critical. Based on the data available, a thorough analysis of the types of crops and suitable techniques to develop the intended machine learning model are to be evaluated. This would result in an extensive application of programming knowledge to develop a suitable ML model, categorizing datasets into training and testing portions, and ultimately resulting in near-accurate predictions.

Chapter 7. Conclusion

Machine learning has gained traction in a wide spectrum of fields, and its implementation in agriculture is not new. Finding its applications in crop price and yield predictions, crop recommendations, visual quality management, irrigation and soil health management, etc., machine learning continues to bring innovative solutions in this much-needed space. This project aims to use such machine learning techniques to address farmers, especially the marginal and small farmers of India, with intelligent solutions for price prediction, particularly for seasonal commodities in the markets of Karnataka. While the outcomes of phase-1 of this project has proposed the problem definition, thoroughly assimilated findings from an extensive literature review and framed a methodology for next steps, the next phase will consist of data curation, analysis, model design, testing and prediction aspects that this project intends to solve. Addressing seasonal commodities brings an interesting challenge to the model designers, and being able to predict and deal with such trends sufficiently accurately is the essence of this project. This model would use historic price data, along with supplemental data points like yield, temperature, rainfall, etc., to create a holistic model with adequate accuracy so that it would not only assist farmers with price prediction for a given seasonal crop, but also aid them with better financial and crop planning over their harvesting seasons.

Appendix A: Data Sources

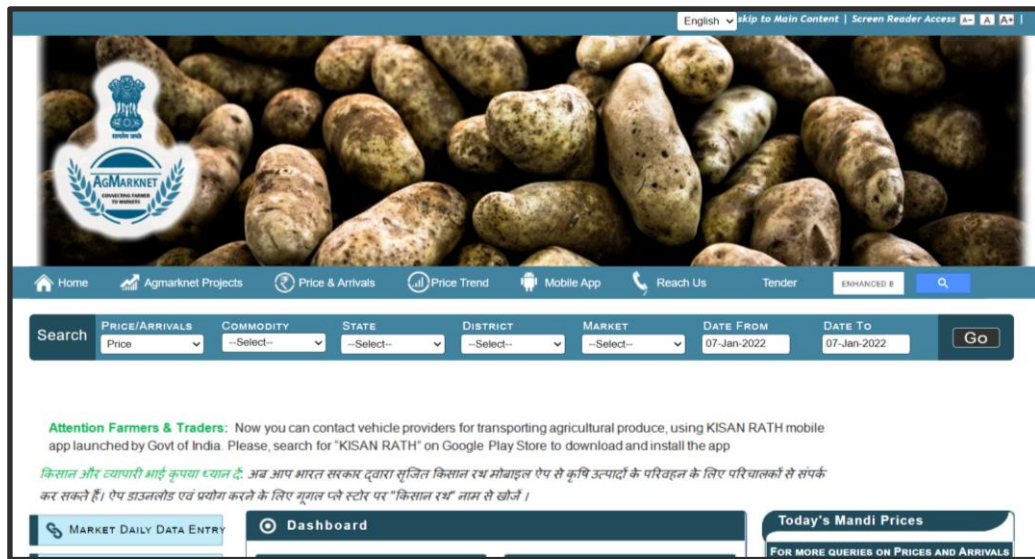


Figure 5.1. Agmarknet



Figure 5.2. Open Government Data Platform of India

Government of Karnataka
Department of Agricultural Marketing and Karnataka State Agricultural Marketing Board

Krishi Marata Vahini

Online Agricultural Marketing Information System

HOME ACT DEPARTMENT KSAMB LOGIN GALLERY CONTACT US APMC Resolutions Appeal Orders 07/01/2022 January 7, 2022

[Cereals](#)
[Vegetables](#)
[Plantation Crops](#)
[Fishes](#)
[Spices](#)
[Forest Products](#)
[Oil Seeds](#)
[Dry Fruits](#)
[Live Stock/Poultry](#)
[Fibre Crops](#)
[Flowers](#)
[Fruits](#)
[Other](#)
[All](#)

State Level Daily Report as on 07/01/2022

Group: Cereals (Prices in Rs)

Commodity	Variety	Grade	Market	Arrival	Unit	Min	Max	Modal
Bajra	Other	Average	LINGASARGUR	538	Quintal	1650	2000	1700
Jowar	Jowar Hybrid	Average	SINCHANUR	50	Quintal	1900	1900	1900
Jowar	Jowar Hybrid	Average	HARAPANAHAL	36	Quintal	1500	1500	1500
Mazra	Hybrid Local	Average	HARAPANAHAL	311	Quintal	1650	1700	1650
Mazra	Hybrid Local	Average	SHIVAMOGGA	249	Quintal	700	1760	1240
Mazra	Local	Average	RAKODURGA	278	Quintal	1550	1650	1670
Paddy	Paddy Sona	Average	LINGASARGUR	525	Quintal	2400	2420	2410
Paddy	Paddy	Average	TNARSIPUR	600	Quintal	2000	2300	2250
Paddy	Sona Masuri Old	Average	SINCHANUR	161	Quintal	2100	2350	2300
Paddy	Sona Masuri New	Average	SINCHANUR	1872	Quintal	1600	1800	1724
Paddy	Paddy RNR Old	Average	SINCHANUR	291	Quintal	2000	2000	2000
Paddy	Paddy RNR Old	Average	HONNALI	452	Quintal	2000	2000	2000
Paddy	Paddy Medium Variety	Average	SHIVAMOGGA	5	Quintal	2500	3200	2850
Ragi	Red	Average	SHIVAMOGGA	138	Quintal	1900	2700	2250
Rice	Coarse	Average	SHIVAMOGGA	412	Quintal	3600	6800	5250
Rice	Fine	Average	SHIVAMOGGA	275	Quintal	2600	4800	3250
Rice	Medium	Average	SHIVAMOGGA	10	Quintal	2600	2800	2700
Wheat	Medium	Average	SHIVAMOGGA	10	Quintal	2600	2800	2700

[Report Menu](#)
[Print](#)

[Privacy Policy](#)
[Copyright Policy](#)
[Disclaimer Policy](#)
[Terms and Conditions](#)
[Help](#)
[Disclaimer](#)
[Screen Reader Access](#)
[Site Map](#)

Version 1.0 (Rev 0105)

Website Content Managed by Department of Agricultural Marketing, Govt. of Karnataka. Designed, Developed and Hosted by National Informatics Centre (NIC). Last Updated: January 7, 2022

india.gov.in

Figure 5.3. Krishi Marata Vahini

कॉफी बोर्ड
COFFEE BOARD
Government of India
Ministry of Commerce & Industry

ABOUT US CONTACT 07/01/2022

Exports Services News Tenders Trade Enquiry Statistics Schemes Media & Events Publications

COFFEE BOARD OF INDIA
Government of India
Ministry of Commerce & Industry

COFFEE NEWS

International Coffee Day on 1st October.
5th International Coffee Day Programme
www.internationalcoffeeday.org

List of Quality Testing Services-
Coffee Quality Division (TIES)

AIC - CCRI - Centre for Entrepreneurship Development

WELCOME TO COFFEES OF INDIA

IBSA Coffee Festival
August 4-5 2021

- The saga of Indian coffee began on a humble note, with planting of 'Seven seeds' of 'Mocha' during 1600 AD by the legendary holy saint Baba Budan, in the courtyard of his hermitage on 'Baba Budan Giri' in Karnataka. For quite a considerable period, the plants remained as a garden curiosity and spread slowly as back yard plantings.
- It was during 18th century that the commercial plantations of coffee were started, thanks to the success of British entrepreneurs in conquering the hostile forest terrain in south India.

Share this: Today's coffee industry has made great strides and earned a dynamic status in the coffee row of

Figure 5.4. Coffee Board of India

Karnataka State Natural Disaster Monitoring Centre
(A Regd. Society Autonomous Body Affiliated to Dept. of Revenue, GOK)

Home Reports Gallery Video Gallery FAQ Feedback Glossary Wd. of Sign In

KSNDMC Report As On Friday, January 7, 2022

*Note: Below Mentioned Daily Report has been uploaded at 09:30 AM and verified after 2:00 PM on Daily Basis.

[Earthquake Reports](#)
[District/Taluk/Hobli Wise Rainfall Pattern](#)

[Reservoir Level](#)
[Rainfall Graphs](#)

[Daily Rainfall Map 07/01/2022](#)
[Weather Maps](#)

[Forecast Maps](#)
[Weekly Rainfall Maps](#)

[BBMP Daily Map](#)
[Monthly Rainfall Maps](#)

[PreMonsoon Seasonal Maps](#)
[SouthWest Monsoon Seasonal Maps](#)

[High Wind Wave / Wave Alert](#)
[Annual Rainfall Maps](#)

[NorthEast Monsoon Seasonal Maps](#)
[Normalized Difference Vegetation Index\(NDVI\)](#)

[Minor Irrigation Tanks In State](#)
[Ground Water Level in State](#)

[Drought Affected Areas](#)

NDMA Guidelines

Tenders

Click for subscription/updates of mobile alerts

Rainfall Details

BAGALKOTE
RANGE: NO RAIN
TOTAL RAIN: NO RAIN

BALLARI
RANGE: (9mm) TO: (11.50mm)
TOTAL RAIN: 0.96(mm)

BELAGAVI
RANGE: NO RAIN
TOTAL RAIN: 0.00(mm)

Figure 5.5. Karnataka State Natural Disaster Monitoring Center

References

- [1] “Food Security and Nutrition and sustainable agriculture | Department of Economic and Social Affairs,” *United Nations*. [Online]. Available: <https://sdgs.un.org/topics/food-security-and-nutrition-and-sustainable-agriculture>. [Accessed: 07-Jan-2022].
- [2] “Fao.org,” *India at a glance | FAO in India | Food and Agriculture Organization of the United Nations*. [Online]. Available: <https://www.fao.org/india/fao-in-india/india-at-a-glance/en/>. [Accessed: 07-Jan-2022].
- [3] India Brand Equity Foundation, “Analysis about Indian agriculture industry, market size, Export & Investment Opportunity,” *IBEF*, 17-Jul-2017. [Online]. Available: <https://www.ibef.org/industry/agriculture-india.aspx>. [Accessed: 07-Jan-2022].
- [4] “[NITI Aayog,” *NITI Aayog*. [Online]. Available: <http://www.niti.gov.in/>. [Accessed: 07-Jan-2022].
- [5] *Agriculture Investment Company, Agricultural Finance Services*. [Online]. Available: <http://sfacindia.com/Aboutus.aspx#:~:text=SFAC%20is%20implementing%20the%20central,marginal%20farmers%20in%20agribusiness%20activities>. [Accessed: 07-Jan-2022].
- [6] “You are future civil servant (IAS/IPS/IRS/IFS/state civil servant officer) that is why you are reading this !!,” *You Are Future Civil Servant (IAS/IPS/IRS/IFS/State Civil Servant Officer) That is Why you are Reading This !!* [Online]. Available: <http://www.theexamsmadesimple.com/p/home.html>. [Accessed: 07-Jan-2022].
- [7] “What is marginal farmer,” *IGI Global*. [Online]. Available: <https://www.igi-global.com/dictionary/marginal-farmer/68592>. [Accessed: 07-Jan-2022].
- [8] G. R. Kumar, “Focusing on major problems of marginal farmers,” *The Hans India*, 23-Jan-2017. [Online]. Available: https://www.thehansindia.com/posts/index/Hans/2017-01-23/Focusing-on-major-problems-of-marginal-farmers/275349?infinite_scroll=1. [Accessed: 07-Jan-2022].
- [9] statistics can display more up-to-date data than referenced in the text. This text provides general information. Statista assumes no liability for the information given being complete or correct. Due to varying update cycles, “Topic: Agriculture in India,” *Statista*. [Online]. Available: <https://www.statista.com/topics/4868/agricultural-sector-in->

india/#dossierKeyfigures. [Accessed: 07-Jan-2022].

[10] “Small and marginal farmers in India - difficulties and solutions,” *India - Israel Trade & Economic Office, Embassy of Israel*, 11-May-2020. [Online]. Available: <https://itrade.gov.il/india/2016/08/29/small-and-marginal-farmers-in-india-difficulties-and-solutions/>. [Accessed: 07-Jan-2022].

[11] “Issues and problems of small and marginal farmer: A study ...” [Online]. Available: https://www.researchgate.net/profile/B-Suresh-Lal/publication/318492801_Issues_and_Problems_of_Small_and_Marginal_Farmer_A_Study_in_Karimnagar_District/links/596dbc04aca2728ade75cf8c/Issues-and-Problems-of-Small-and-Marginal-Farmer-A-Study-in-Karimnagar-District.pdf. [Accessed: 07-Jan-2022].

[12] “Issues and problems of small and marginal farmer: A study ...” [Online]. Available: https://www.researchgate.net/profile/B-Suresh-Lal/publication/318492801_Issues_and_Problems_of_Small_and_Marginal_Farmer_A_Study_in_Karimnagar_District/links/596dbc04aca2728ade75cf8c/Issues-and-Problems-of-Small-and-Marginal-Farmer-A-Study-in-Karimnagar-District.pdf. [Accessed: 07-Jan-2022].

[13] L. Columbus, “10 ways AI has the potential to improve agriculture in 2021,” *Forbes*, 17-Feb-2021. [Online]. Available: <https://www.forbes.com/sites/louiscolumbus/2021/02/17/10-ways-ai-has-the-potential-to-improve-agriculture-in-2021/?sh=3ffd1b007f3b>. [Accessed: 07-Jan-2022].

[14] Y. Jabareen, “Building a conceptual framework: Philosophy, definitions, and procedure,” *International Journal of Qualitative Methods*, vol. 8, no. 4, pp. 49–62, 2009.

[15] Z. Hossain, Q. Abdus Samad, and Z. Ali, “Arima model and forecasting with three types of pulse prices in Bangladesh: A case study,” *International Journal of Social Economics*, vol. 33, no. 4, pp. 344–353, 2006.

[16] M. Kaur, H. Gulati, and H. Kundra, “Data mining in agriculture on crop price prediction: Techniques and applications,” *International Journal of Computer Applications*, vol. 99, no. 12, pp. 1–3, 2014.

[17] J. Parikh and C. Singh, “Extension of MSP: Fiscal and welfare implications - NITI,” Oct-2007. [Online]. Available:

https://www.niti.gov.in/planningcommission.gov.in/docs/reports/sereport/ser/ser_msp.pdf. [Accessed: 07-Jan-2022].

[18] P. V. Cenas, “[PDF] forecast of agricultural crop price using time series and Kalman filter method: Semantic scholar,” *[PDF] Forecast of Agricultural Crop Price using Time Series and Kalman Filter Method | Semantic Scholar*, 01-Jan-2017. [Online]. Available: <https://www.semanticscholar.org/paper/Forecast-of-Agricultural-Crop-Price-using-Time-and-Cenas/7eca5c374b0fe4eb95688f6bc1dcdcbf4ebe3825>. [Accessed: 07-Jan-2022].

[19] A. Shah, A. Dubey, V. Hemnani, D. Gala, and D. R. Kalbande, “Smart farming system: Crop yield prediction using regression techniques,” *Lecture Notes on Data Engineering and Communications Technologies*, pp. 49–56, 2018.

[20] W. Ma, K. Nowocin, N. Marathe, and G. H. Chen, “An interpretable produce price forecasting system for small and marginal farmers in India using collaborative filtering and adaptive nearest neighbors,” *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*, pp. 11–11, 2019.

[21] N. Parida, D. Mishra, K. Das, and N. K. Rout, “Development and performance evaluation of Hybrid Kelm models for forecasting of agro-commodity price,” *Evolutionary Intelligence*, vol. 14, no. 2, pp. 529–544, 2019.

[22] H. Ouyang, X. Wei, and Q. Wu, “Agricultural Commodity Futures prices prediction via long- and short-term time series network,” *Journal of Applied Economics*, vol. 22, no. 1, pp. 468–483, 2019.

[23] B. Basso and L. Liu, “Seasonal crop yield forecast: Methods, applications, and accuracies,” *Advances in Agronomy*, vol. 154, pp. 201–255, 2019.

[24] A. Vohra, N. Pandey, and S. K. Khatri, “Decision making support system for prediction of prices in Agricultural Commodity,” *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 2019.

[25] A. Jain, S. Marvaniya, S. Godbole, and V. Munigala, “A framework for crop price forecasting in emerging economies by analyzing the quality of time-series data,” *arXiv.org*, 09-Sep-2020. [Online]. Available: <https://arxiv.org/abs/2009.04171>. [Accessed: 07-Jan-2022].

[26] T. van Klompenburg, A. Kassahun, and C. Catal, “Crop yield prediction using

Machine Learning: A Systematic Literature Review,” *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

[27] K. Kurumatani, “Time series forecasting of agricultural product prices based on recurrent neural networks and its evaluation method,” *SN Applied Sciences*, vol. 2, no. 8, 2020.

[28] K. M. Sabu and T. K. M. Kumar, “Predictive analytics in agriculture: Forecasting prices of arecanuts in Kerala,” *Procedia Computer Science*, vol. 171, pp. 699–708, 2020.

[29] G. Hegde, V. R. Hulipalled, and J. B. Simha, “A study on Agriculture Commodities Price Prediction and forecasting,” *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, 2020.

[30] A. Kumar, B. Chakrabarti, and A. A. Ansari, “Kisan Mitra: A data analytic based prototype for avoiding crops price crash,” *SSRN Electronic Journal*, 2020.

[31] A. P. S and H. B A, “Crop value forecasting using decision tree regressor and ...,” *European Journal of Molecular & Clinical Medicine*, 2020. [Online]. Available: https://ejmcm.com/article_2585_86ae450aad4837565e30767586a4d64b.pdf. [Accessed: 09-Jan-2022].

[32] Y. J. Mgale, Y. Yan, and S. Timothy, “A comparative study of Arima and holt-winters exponential smoothing models for Rice Price forecasting in Tanzania,” *OALib*, vol. 08, no. 05, pp. 1–9, 2021.

[33] R. Dhanapal, A. AjanRaj, S. Balavinayagapragathish, and J. Balaji, “Crop price prediction using supervised machine learning algorithms,” *Journal of Physics: Conference Series*, vol. 1916, no. 1, p. 012042, 2021.

[34] P. Kamath, P. Patil, S. S, Sushma, and S. S, “Crop yield forecasting using data mining,” *Global Transitions Proceedings*, vol. 2, no. 2, pp. 402–407, 2021.

[35] Anubha, K. Tripathi, K. Kumar, and G. Khandelwal, “Onion price prediction for the market of kayamkulam,” *Data Analytics and Management*, pp. 77–85, 2021.

[36] A. Suruliandi, G. Mariammal, and S. P. Raja, “Crop prediction based on soil and environmental characteristics using feature selection techniques,” *Mathematical and Computer Modelling of Dynamical Systems*, vol. 27, no. 1, pp. 117–140, 2021.

- [37] V. K. Kalichkin, O. K. Alsova, and K. Yu Maksimovich, "Application of the decision tree method for predicting the yield of spring wheat," *IOP Conference Series: Earth and Environmental Science*, vol. 839, no. 3, p. 032042, Sep. 2021, doi: 10.1088/1755-1315/839/3/032042.
- [38] I. Ghutake, R. Verma, R. Chaudhari, and V. Amarsinh, "An intelligent Crop Price Prediction using suitable Machine Learning Algorithm," *ITM Web of Conferences*, vol. 40, p. 03040, 2021, doi: 10.1051/itmconf/20214003040.
- [39] S. Gupta, A. Agarwal, P. Deep, S. Vaish, and A. Purwar, "Analysis of Minimum Support Price Prediction for Indian Crops Using Machine Learning and Numerical Methods," *Advances in Intelligent Systems and Computing*, pp. 267–277, Jul. 2020, doi: 10.1007/978-981-15-5148-2_24.
- [40] M. Jadhav, N. Kolambe, S. Jain, and S. Chaudhari, "Farming Made Easy using Machine Learning," *2021 2nd International Conference for Emerging Technology (INCET)*, May 2021, doi: 10.1109/incet51464.2021.9456351.
- [41] Z. Chen, H. S. Goh, K. L. Sin, K. Lim, N. K. H. Chung, and X. Y. Liew, "Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques," *arXiv.org*, 24-Jun-2021. [Online]. Available: <https://arxiv.org/abs/2106.12747>. [Accessed: 09-Jan-2022].
- [42] "Agriculture Marketing," *Dashboard*. [Online]. Available: <https://agmarknet.gov.in/>. [Accessed: 07-Jan-2022].
- [43] "Open Government Data (OGD) Platform India," *A Digital India Initiative*. [Online]. Available: <https://data.gov.in/>. [Accessed: 07-Jan-2022].
- [44] "Home Page ," *Online Agricultural Marketing Information System*. [Online]. Available: <http://krishimaratavahini.kar.nic.in/>. [Accessed: 07-Jan-2022].
- [45] "Coffee Board," *Welcome to Coffees of India*. [Online]. Available: <https://www.indiacoffee.org/>. [Accessed: 07-Jan-2022].
- [46] "Weather Dash Board: - ksndmc.org," *KSNDMC Report*. [Online]. Available: <https://ksndmc.org/KsndmcDashBoard/>. [Accessed: 07-Jan-2022].

10%

SIMILARITY INDEX

6%

INTERNET SOURCES

2%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1

[Www.Niti.Gov.In](http://www.niti.gov.in)

Internet Source

3%

2

Submitted to National Institute of Business Management Sri Lanka

Student Paper

1%

3

upscenvironment.wordpress.com

Internet Source

1%

4

Submitted to Technological Institute of the Philippines

Student Paper

<1%

5

Submitted to BMS College of Engineering

Student Paper

<1%

6

www.theexamsmadesimple.com

Internet Source

<1%

7

oecdinsights.org

Internet Source

<1%

8

"Agricultural Value Chains in India", Springer Science and Business Media LLC, 2022

Publication

<1%