

Problem Scoping

It is the first time we are exposed to problems. They may be easier or big, sometimes general or sometimes very critical. Many times we can become as part of a problem to becomes a part of the solution.

Schemer → To identify problem we'll follow schema.

List of Problems

Theme	Type
-------	------

Themes: Digital Literacy Topics: Online Learning platforms, Digital Awareness, e-books, etc.
Theme: Health Topics: Medical App, Mobile Medicines, Screening of diseases, etc.
Theme: Entertainment Topics: Media, Virtual/Gaming, Interactive AVs, Promotions etc.

Application → Yield rates
Sowing & harvesting → How might we increase the former decreasing the best time for sowing
Partners: & saving their input?

What Problem Convex



Who?

The 'Who?' block helps you in analyzing the various groups affected directly or indirectly due to the problem. You find out who is responsible for the problem and what do they know about the problem. This will help you to understand the people involved in the problem and the people who are affected by the problem.

Who are Stakeholders?

For Farmers, Fertilizer producers, labourers & Tractors companies.

What do we know about them?

E.g.: These are some of the people worst affected by the problem & loses their money & time.

What?

Under the 'What?' block, you need to look into what you have learned. At this stage, you need to determine the nature of the problem. What is the problem and how do you know about it? You can also list down the possible causes of the problem. If there are any more related details, newspaper articles, media announcements, etc. are some examples.



Where?

New that you know who is associated with the problem and what the problem actually is, you need to focus on the context/location of the problem. This block will help you to look into the situation in which the problem arises, the context of it, the location where it is present.



Why?

You have finally listed all the major elements that affect the problem directly. Now it's convenient to know who the people it would be beneficial by the solution are, what are the benefits of the solution, and how the solution would benefit the stakeholders at the base of any are end to solve this problem. This is the 'Why?' sense, this about the benefits which the stakeholders would get from the solution and how would it benefit them, as well as to whom.



problem statement template	
formally	informally
structured	unstructured
positions	positions
labels	labels

for problem	for solution
Describing what has happened	Describing what needs to happen
by the user and	by the user and
describing the cause	describing the cause
describing the effect	describing the effect

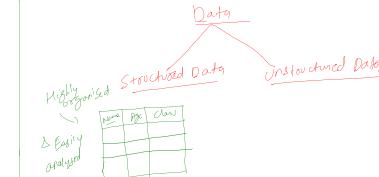
when?	where?
Time	Location
Time	Location

Data Acquisition
It is the stage in which we acquire data for the project.

What is Data?

- Data can be defined as a representation of facts or instructions about some entity (school, sports, business, animal, student etc) that can be processed or communicated by humans or machine.
- Data is a collection of facts, such as numbers, words, pictures, audios, videos, map, measurements, observations or even just description of things.

- Data may be represented with the help of characters such as alphabets (A-Z, a-z), digits (0-9) or special characters (+, -, /, *, <, >, = etc).



- Structured data is categorized as quantitative data.
- It's the type of data most of us work with every day.
- Structured data has predetermined data types & format so that it fits well in the column fields of database or spreadsheet.
- Highly organized & easily analyzed → Storing in form of tables.
- Data is stored in relational databases or spreadsheets (like Excel sheets).
- Examples of structured data are Name, age, address, etc.
- Unstructured data is difficult to deconstruct because it has no predefined model, meaning it can't be organized in relational databases.
- NOSQL databases for managing unstructured data.
- Example of unstructured data: Video, audio, satellite imagery, social media activity.

Data

- A Data set is just a set of certain objects.
- This set is normally represented in tabular pattern.

→ Every column describes a particular variable.
→ Each row corresponds to a given member of the data set.

Dishwasher	Predictive		
	Name	Age	Garden
A	Tom	36	No
B	Bob	62	Yes
C	Mona	63	Yes
D	John	43	No
E	Jane	36	No

Test Predictive

Data Features

- A measurable piece of data that can be used for analysis.
- In CSV & Excel files they could be known as columns.
- Features are also sometimes referred to as "Variables" or "Attributes".

→ Depending on what we're trying to analyse, the features we include in our dataset can vary widely.

Acquiring Data (Sources)



Surveys
A research method used for collecting data from a predefined group of respondents to gain information & insights into various topics of interest.

Cameras
→ A camera captures a visual image.
→ Could be used to collect data for CV tasks.
→ Practice at measuring visual assets in the form of Photo, film or video.

Web Scraping

- ✓ Web Scraping is the process of collecting structured web data in an automated fashion. It's also called web extraction.
- ✓ Use cases of web scraping include price monitoring, price intelligence, news monitoring, lead generation & market research among many others.

Observations

- help us to find relationship b/w different elements of the system which we have seen!

Sensors

→ To understand complex events with multiple factors that affect each other.

- ✓ An device which collects or measures a physical property & records, indicates or otherwise responds to it.
- ✓ Eg - Temperature Sensors, Humidity Sensors, Pressure Sensors, Proximity Sensors, Level Sensors, Acceleration Sensors.

Capacitive → Measure Rate change
Infrared → Measure Distance

- Infrared Sensors - (i) Eye - Cash counting machine.
(ii) Infrared sensor
Count up of people
Count of people
Count of car

API (Application Program Interface)

- It's a software intermediary that allows two applications to talk to each other.
- ↳ External to feature

Acquiring Data from Reliable Sources

- ✓ Internet, Random websites
 - from which we collect is open sourced & not commercial property
- ✓ Open source websites hosted by the good
- ✓ Some other sources: Google Scholar, ResearchBazaar, YouTube, websites

System Map

→ help us to find relationship b/w different elements of the system which we have seen!

- To understand complex events with multiple factors that affect each other.
- Every element is interconnected.
- one can easily define a relationship among different elements which come under a system.

In a System Map



Data modeling



Deep learning



Let's imagine we're trying to teach machines to recognize faces. The machine needs to learn what this face looks like.

- give them always to machines
- to recognize human faces
- to move & manipulate objects
- to understand the voice commands by humans, and also do various tasks

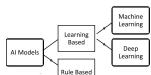
Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.

To basically a term used for the use of Machine learning

Modeling

- ML modeling refers to developing algorithms, also called models, which can be trained to get intelligent outputs
- That is, writing codes to make a machine intelligently respond

In the previous model of class experiments, we have seen various types of graphical representation which can be used for representation of data for further analysis. In this section, we will learn about decision trees. A decision tree is a classification or regression tree in which each internal node represents a test on an attribute (e.g., whether a pixel is red), and each branch represents an outcome of the test. The leaf nodes represent the class label (e.g., what is the heart disease for this patient).



Rule Based Approach

Rule based approach tries to fit in modeling with the information or pattern of data and directly to develop. The machine follows the rule or instruction mentioned by the developer, and performs task accordingly.



Decision Tree



Learning Based Approach

Machine learning is a machine learning where the machine learns by itself to learn and predict by the developer. The machine learns from the data provided to it. The machine is trained with the data (Figure out patterns and trends out of it). Generally it is followed when the data is large and complex. It is a process of training the machine with the data. The machine then uses the data to predict similar features of it and clusters some datasets together. That is, to say, the machine tells us about the trends which it observed while it was trained.

Decision Tree

→ It is the most powerful & popular tool for classification & prediction.

↳ A decision tree is a hierarchical tree structure, where each internal node shows a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.



Dataset	Temperature	Humidity	Wind	Classify
Sunny	High	Normal	Weak	Yes
Sunny	High	High	Strong	No
Overcast	High	High	Weak	Yes
Rainy	Normal	Normal	Weak	Yes
Rainy	Normal	High	Normal	No
Rainy	Normal	High	Strong	No
Cloudy	High	Normal	Weak	Yes
Cloudy	High	High	Normal	Yes
Cloudy	Normal	High	Weak	Yes
Sunny	High	High	Weak	No
Sunny	High	Normal	Normal	Yes
Overcast	Normal	Normal	Normal	Yes
Overcast	Normal	High	Strong	No
Rainy	Normal	Normal	Weak	No

Evaluation



80% - Training data

20% - Testing data

Once a model has been made and trained, it needs to go through proper testing so that one can calculate the efficiency & performance of the model.

∴ Hence, the model is tested with the help of Testing data (which was separated out of the acquired dataset at Data Acquisition stage) & the efficiency of the model is calculated on the basis of parameters mentioned below:

True positive	True Negative
False Positive	False Negative
False Negative	True Positive

The Confusion Matrix		Reality	
		Yes	No
Prediction	Yes	TP	FP
	No	FN	TN

