

Building Socially Impactful Solutions using AI

Public Health & Agriculture

Jigar Doshi - June 2025

ML Journey

- Georgia Tech - Atlanta
- IBM Research - New York
- CrowdAI - San Francisco
- Wadhwani AI - Mumbai
- Artpark, IISC - Bangalore

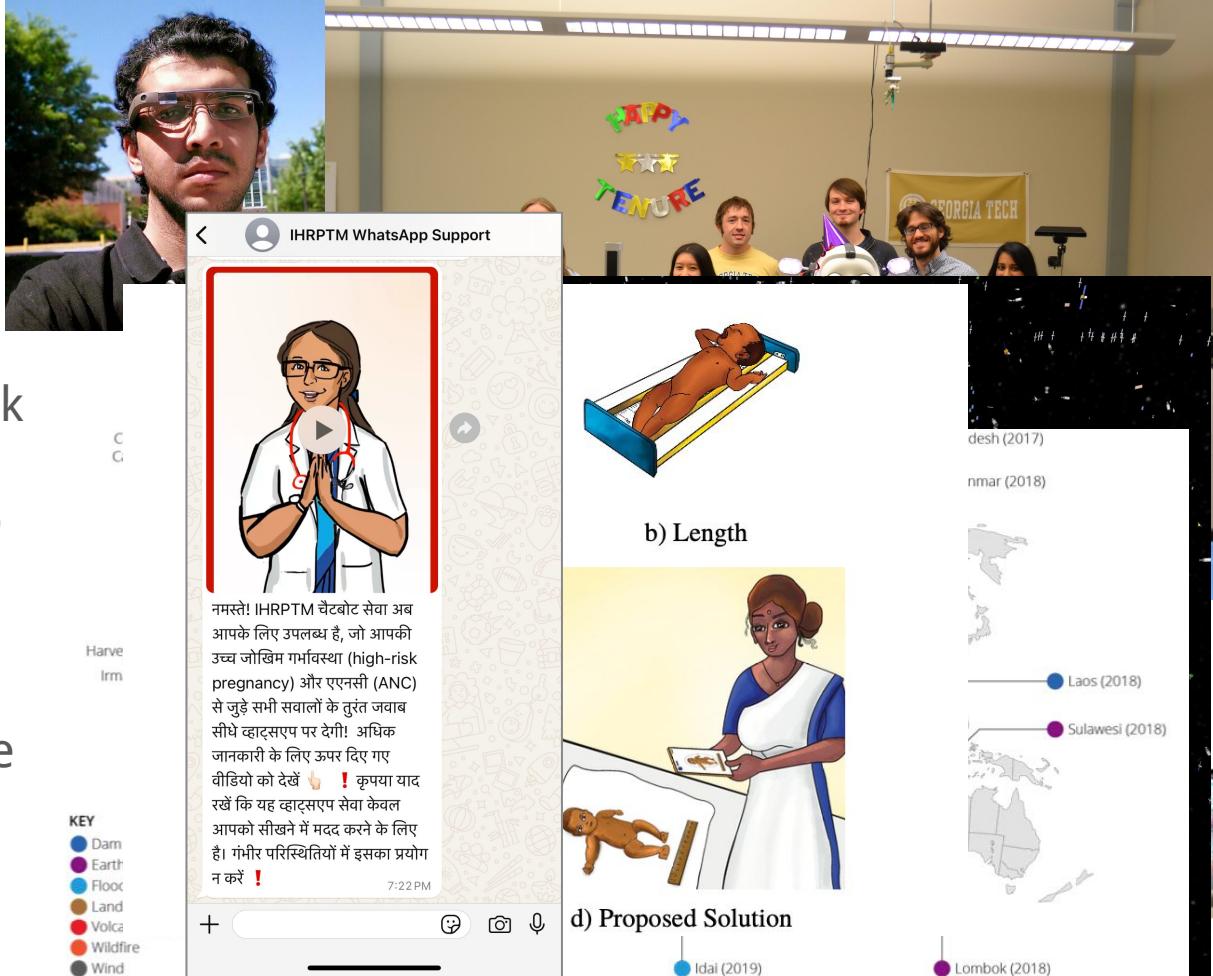


Figure 2: Disaster types and disasters represented in xBD over the world.

Talk Outline

- Day 1
 - LLMs for Mothers and Frontline health workers
 - TB & Covid Screening using Cough samples
 - AI Solutioning
- Day 2
 - Pest Management
 - Farmer Chat / Remote Sensing
 - AI Agents for Public Health
 - How to pick a problem
 - To build an impactful fulfilling ML career



LLM Copilot for Front Line Workers for High-Risk Pregnancy Management

May 2024

Poor Maternal and Child Health in India

A woman dies in childbirth every 20 minutes; for every woman who dies, 20 more suffer lifelong ailments.

2 children under 5 die every minute; 4 of 10 children don't realize their full potential due to chronic undernutrition/stunting.

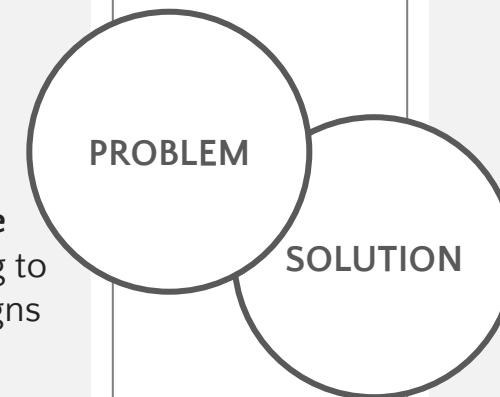
Factors influencing this include:



Lack of access to preventive care information and services, leading to poor understanding of danger signs and delayed care seeking



Inadequately trained & supported health workers who are unable to detect and manage high-risk conditions in time



ARMMAN leverages deep mobile penetration with existing health worker network and infrastructure to:



Provide preventive care information through pregnancy & infancy enabling women to seek care in time



Train and support health workers for timely detection & management of high-risk conditions

Our programs



PREGNANT WOMEN, MOTHERS AND CHILDREN

mMitra



- ▶ Free voice-call service providing critical preventive care information during pregnancy and infancy
- ▶ Largest mobile-based maternal messaging programme in the world in collaboration with MOHFW
- ▶ Live telephonic counselling for caregivers of moderately underweight infants to prevent their decline into severe malnutrition

60 MILLION+

WOMEN AND CHILDREN

HEALTH WORKERS AND SYSTEMS



- ▶ Mobile-based refresher training course for frontline health workers (ASHAs) in partnership with Ministry of Health and Family Welfare

- ▶ Implementation of high-risk management protocols for sustained reduction in delayed referrals and high-risk referrals to tertiary facilities

500,000+

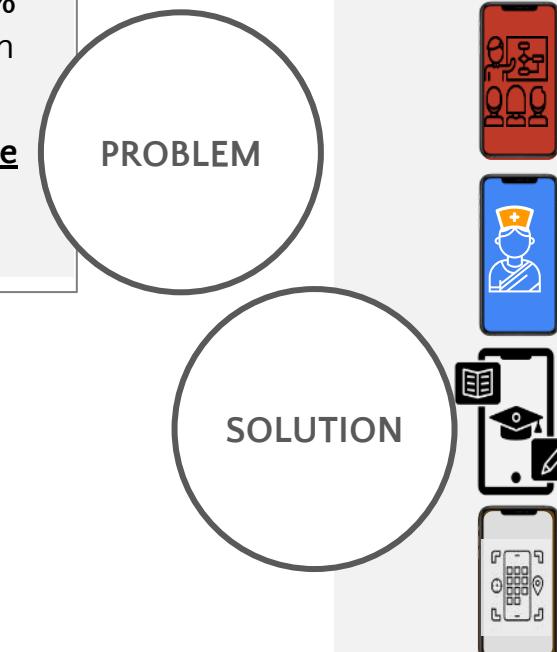
HEALTH WORKERS

ARMMAN's Integrated High Risk Pregnancy Tracking and Management (IHRPTM) program trains and supports 20,000+ health workers across 3 states to detect and manage high-risk pregnancies.

High Risk Pregnancy Vs Perinatal Morbidity & Mortality

20-30%¹ pregnancies belong to high-risk category, which are responsible for 75% of perinatal morbidity and mortality in India.

However, only 14% of the pregnancies are currently being recorded as 'High Risk'²



IHRPTM, an integrated, comprehensive, multi-step, systemic approach for improved identification, tracking and end-to-end management of high-risk pregnancies.

- Extensive Hybrid *Training on color coded HRP protocols for better role clarity* to avoid delayed & irrational referrals,
- *Tech+ Touch approach* using *Learning management app helps in access to self-paced, practice based multimedia learning content.*
- *Tracking of HRP through digital solution*
- *Data driven Regular Advocacy to ensure Availability, functionality and utility of logistics,*

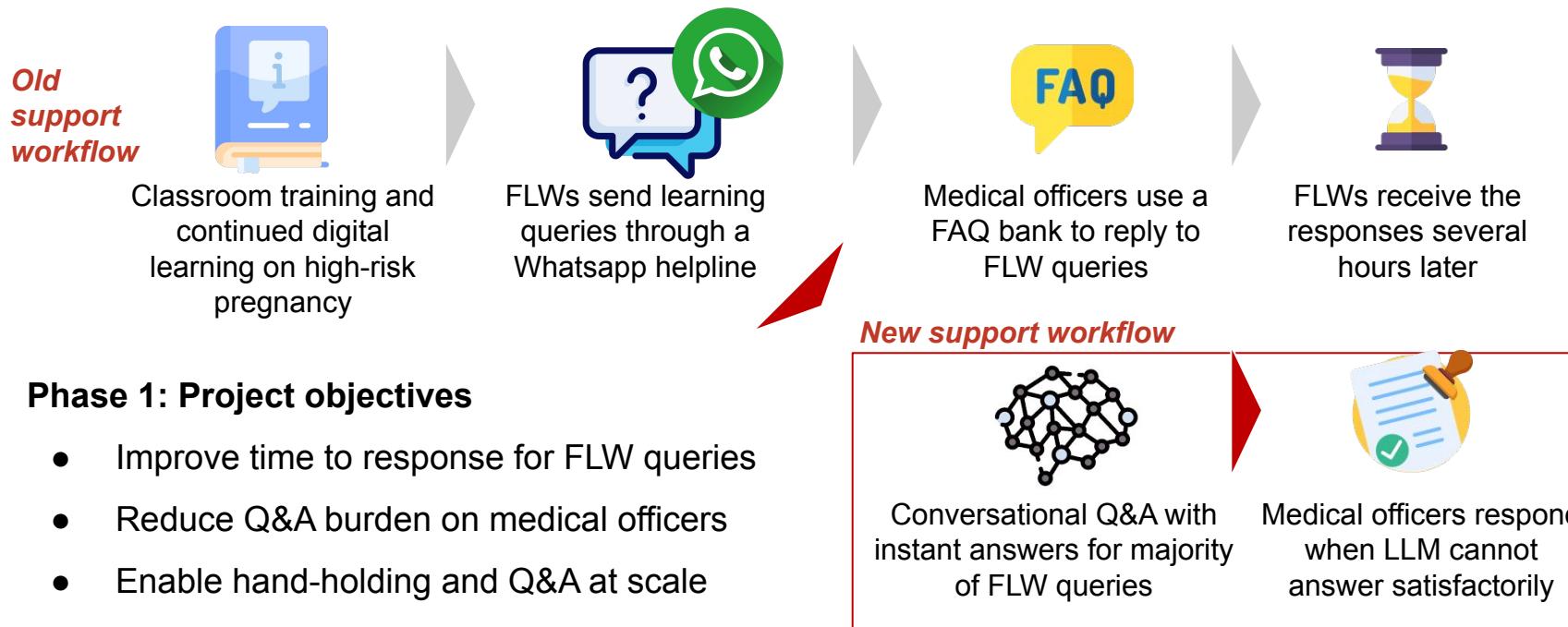
Source: 1. Research article

2. Guidance Note for "Extended PMSMA for tracking High Risk Pregnancies (HRPs)"

3. SRS Bulletin

Chatbot project introduction

ARMMAN and ARTPARK are building a learning copilot in the IHRPTM digital learning system, starting with automated ANM support for learning queries

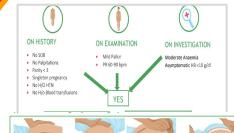


Project Journey: Iterative, human-centered approach

1. Project scoping:
Picking the use case



2. Data curation & creation:
Knowledge corpus, FAQs,
evaluation sets



4. Development &
“lab” evaluation of
chatbot

Rationale	FLW		FLW + Chatbot		QFA		QFA + Chatbot	
	Pre-Training Success Rate	Post-Training Success Rate	Pre-Training Success Rate	Post-Training Success Rate	Success Rate	Completion Rate	Success Rate	Completion Rate
Asymptomatic	66.3%	77.0%	72%	72%	74%	72%	90%	90%
Abnormal	66.3%	77.0%	72%	72%	74%	72%	90%	90%
All	66.3%	77.0%	72%	72%	74%	72%	90%	90%
All 3 evaluations	20%	25%	82%	82%	55%	55%	86%	86%
At least 2 of 3 evaluations agreed	4%	9%	62%	62%	4%	4%	86%	86%
All 2 evaluations agreed	1%	1%	98%	98%	98%	98%	98%	98%
No evaluations agreed	1%	2%	2%	2%	1%	1%	1%	1%

3. Wizard of Oz
testing for user
insights



5. Prototype
testing with 12
FLWs



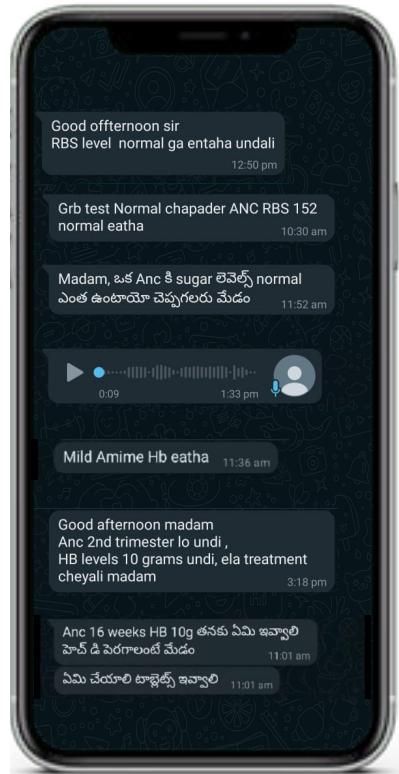
6. Deployment to 100
FLWs for ~3 months

NEXT: Scale up
in phases



User insights through Wizard of Oz experiments

We conducted 'Wizard of Oz' experiments to understand how an LLM bot will be used, perceived and interpreted by ANMs



Overall

For greater adoption, the intro to the bot must be facilitated by a medical officer

🤝 ANMs tend to interact with the platform respectfully ("OK Sir", "OK Madam") as they believe it to involve a credible authority figure

❓ Expect onboarding challenges: what to ask, how to ask, etc

Queries

💬 Queries likely to be phrased as search strings, with complex queries split into multiple messages

🗣️ Mix of English and Telugu phrases and fonts used, spelling mistakes common

🎙️ More detail in voice notes vs text queries. Large variations in dialect & pronunciation

Responses

ENG Preference for Telugu responses, with medical terminology in English

⌚ Quick response important; delays may trigger other medium of doubt clarification

↗️ Crisp language preferred. Utilize action-oriented verbs, replicating real communication

Phase-wise model evaluation strategy

Phase	Evaluations	Objective	Work done
I. Objective evaluation	<ul style="list-style-type: none">ANM pre/post quiz (<i>75 questions, English</i>)FAQ match (<i>-98 questions x 5 variations, English</i>)	<ul style="list-style-type: none">Validate “domain knowledge” of model	<ul style="list-style-type: none">Making program content machine-readable / RAG-friendlyFAQ match
II. Human evaluation by experts <i>(Indic Languages)</i>	<ul style="list-style-type: none">Evaluation of generated answers by clinical experts (<i>55 questions x 3 MTOs, English, text</i>) (<i>15 questions x 3 MTOs, Telugu, text</i>) (<i>65 questions x 3 MTOs, Hindi, text</i>)	<ul style="list-style-type: none">Measure correctness, completeness and conciseness of answers generated by modelMeasure model performance in Telugu & Hindi	<ul style="list-style-type: none">Generation of answers, with citationsMultilingual support (English, Telugu, Hindi)Tuning answers to be more “user-friendly”
III. Human evaluation by end users <i>(Indic Languages)</i>	<ul style="list-style-type: none">Field testing of Whatsapp chatbot by ANMs (<i>4 scenarios x 2 modes x 12 ANMs, Telugu, speech & text</i>)	<ul style="list-style-type: none">Understand user acceptability of responsesPredict likely failure points at deployment	<ul style="list-style-type: none">Multimodal support (text and voice)WhatsApp integration
IV. Pilot + deployment	<ul style="list-style-type: none">Limited deployment with 1000 ANMs	<ul style="list-style-type: none">Understand usability, correctness, usefulness of chatbot	<ul style="list-style-type: none">Integration with WhatsApp helplineHuman in the loop

I. Model performance on FLW quiz

Risks/Protocols	FLW Pre-Training Quiz (n= 6000+)	FLW Post-Training Quiz (n= 6000+)	GPT4 <u>without</u> Training manuals	GPT4 <u>with</u> Training manuals
Quality ANC	66.6%	77.8%	100%	100%
Anaemia	66.3%	74.6%	73%	93%
Antepartum Hemorrhage (APH)	67.2%	79%	80%	100%
Hypertension	65.9%	75.6%	93%	93%
Gestational Diabetes	46.8%	63.3%	73%	100%

I. Model performance on FAQ match

Evaluation	No. of questions	Mean accuracy	Standard deviation of accuracy
FAQ Set (Exact)	80	100%	-
FAQ Set (Paraphrased, 5-fold)	80x5	96%	1.37%
Expert-created question set (Exact)	18	94%	-
Expert-created question set (Paraphrased, 5-fold)	18x5	75%	7.96%

II. Model performance on expert evaluation

The model's response is...

	Good enough to send to FLW as is	Free of factual errors	Complete	Concise and easy to follow	Backed by satisfactory citations
All 3 evaluators agreed	71%	85%	82%	55%	49%
At least 2 of 3 evaluators agreed	84%	93%	93%	69%	80%
At least 1 of 3 evaluators agreed	98%	98%	98%	98%	89%
No evaluators agreed	<2%	<2%	<2%	<2%	11%

Chatbot overview

नमस्ते! IHRPTM चैटबोट सेवा अब आपके लिए उपलब्ध है, जो आपकी उच्च जोखिम गर्भवत्था (high-risk pregnancy) और एनसी (ANC) में जुड़े सभी सवालों के उरंत जवाब सीधे खाद्यारप पर देती। अधिक जानकारी के लिए ऊपर दिए गए वीडियो को दें! ! कृपया यद रखें कि यह खाद्यारप सेवा केवल आपको सीधे में मदद करने के लिए है। गंभीर परिस्थितियों में इसका प्रयोग न करें !

7:22 PM

Mild एनीमिया में anm को क्या करना चाहिए? 10:58 PM ✓

- माझ्य एनीमिया में ANM को गर्भवती महिला को आहार संबंधी सलाह देनी चाहिए और Iron और Folic Acid (IFA) की गोलियां प्रतिदिन एक बार देनी चाहिए, जब तक कि वे सहन की जा सकें।
- यदि गर्भवत्था के दूसरे या तीसरे तिमाही में हैं, तो IFA की गोलियां दिन में दो बार तब तक दी जानी चाहिए जब तक Hemoglobin (HB) स्तर 11 g/dL तक न पहुँच जाए।
- इसके बाद, IFA की गोलियां प्रतिदिन एक बार 180 दिनों तक जारी रखनी चाहिए।
- इसके अलावा, T. Albendazole 400 mg एक बार देना चाहिए और आहार परामर्श प्रदान करना चाहिए।

क्या आप दिए गए जवाब से संतुष्ट हैं? 10:58 PM

हाँ
नहीं

IHRPTM WhatsApp Support
क्या आप दिए गए जवाब से संतुष्ट हैं? 10:59 PM ✓

हाँ
नहीं

Multilingual, multimodal chat with user-friendly answers pulled from training protocols

- माझ्य एनीमिया में anm को क्या करना चाहिए?
- इसके बाद, IFA की बार 180 दिनों तक उपलब्ध रखनी चाहिए।
- इसके अलावा, T. Albendazole 400 mg एक बार देना चाहिए और आहार परामर्श प्रदान करना चाहिए।

Feedback prompt after every answer

क्या आप दिए गए जवाब से संतुष्ट हैं? 10:58 PM

हाँ
नहीं

IHRPTM WhatsApp Support
क्या आप दिए गए जवाब से संतुष्ट हैं? 10:59 PM ✓

माफ कीजिए, आपका सवाल का जवाब नहीं दे सकते। आपका सवाल MTO के साथ शेयर किया जा रहा है, जिसका जवाब आपको 24 घंटों के अंदर मिलेगा। 10:59 PM

No response to salutations or greetings

माफ कीजिए, आपका सवाल का जवाब नहीं दे सकते। आपका सवाल MTO के साथ शेयर किया जा रहा है, जिसका जवाब आपको 24 घंटों के अंदर मिलेगा। 11:01 PM

Baccha male hai ya female kaise pata chalega? 11:01 PM ✓

माफ कीजिए, आपका सवाल का जवाब नहीं दे सकते। आपका सवाल MTO के साथ शेयर किया जा रहा है, जिसका जवाब आपको 24 घंटों के अंदर मिलेगा। 11:01 PM

Thank you 11:01 PM ✓



Chatbot Demo

2:11

1 IHRPTM WhatsApp Support

पोषक आहार पर जोर दिया जा सके।

- एक महीने के बाद प्रगति की जांच के लिए एक follow-up Haemoglobin (Hb) test करना चाहिए।
- अगर Haemoglobin (Hb) स्तर में महत्वपूर्ण वृद्धि नहीं होती या कम से कम 1 g/dL तक नहीं बढ़ता, तो रोगी को Medical Officer (MO) या किसी उचित चिकित्सा केंद्र के लिए संदर्भित करना चाहिए।

3:34 PM

क्या आप दिए गए जवाब से संतुष्ट हैं?

3:34 PM

हाँ

नहीं

Today

IHRPTM WhatsApp Support

क्या आप दिए गए जवाब से संतुष्ट हैं?

हाँ

2:10 PM ✓

धन्यवाद! आप चाहें तो और सवाल पूछ सकतें हैं

2:10 PM

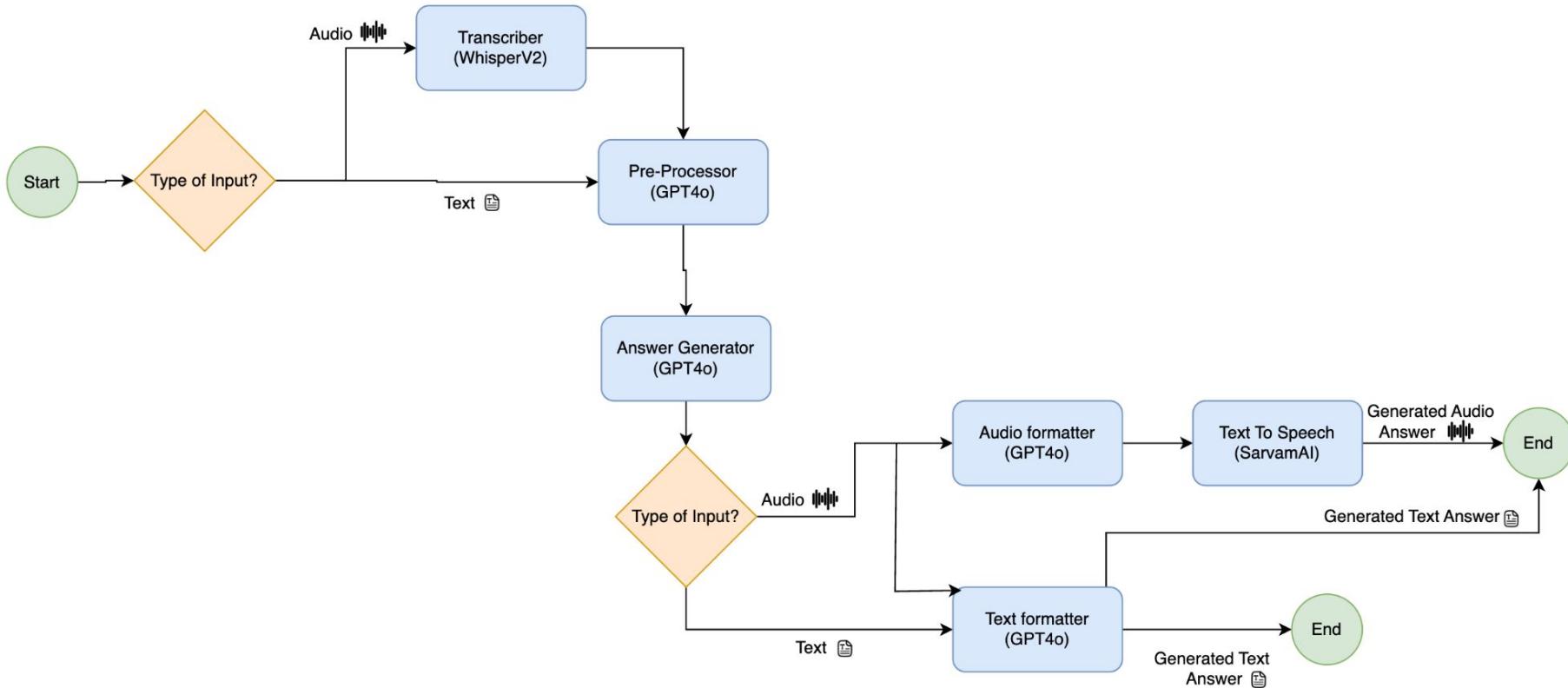
+

Info icon

Image icons: +, Camera, Download



System Architecture



Pilot: A Snapshot (March 2025)



100 FLWs enrolled

in stages across 2 districts in Uttar Pradesh

600+ queries

received from 72 FLWs

91% satisfactory responses

as rated by domain expert

92% positive feedback

by users on responses received

80% repeat users

who come back for second session

11% voice queries

vs 89% text queries

43% queries

escalated to human in the loop

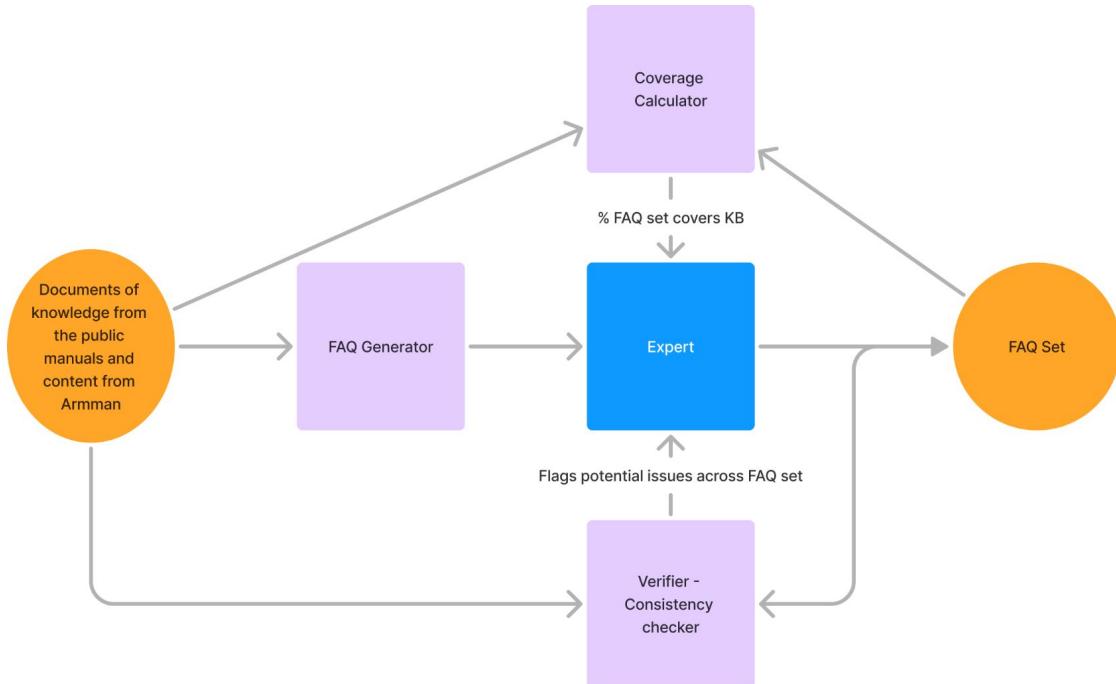


LLM assistant's for Mothers

- Kilkari National Program: ~5M Mothers each year
- Current program is canned messages 2 times a week
- Before moving to LLM based conversation, build a verifiable system
- Verifiable: All responses are vetted
- Solution:
 - Offline QA Bank generation + Retrieval System
 - Can be deployed locally. 33M Embedding Model + Similarity Search
 - Cheap and efficient



LLM + Human Exhaustive QA

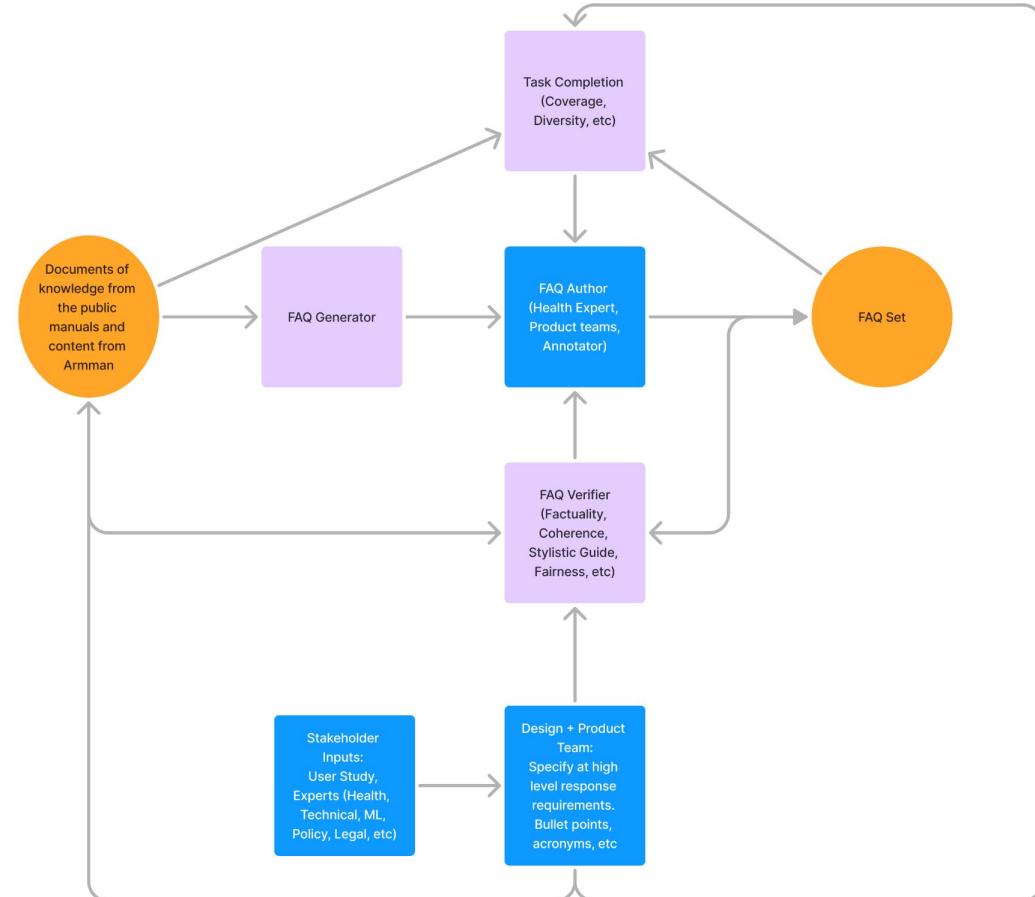


- Hallucination free
- Fast and Offline
- Converts 60K tokens knowledge base to 2K QA pairs



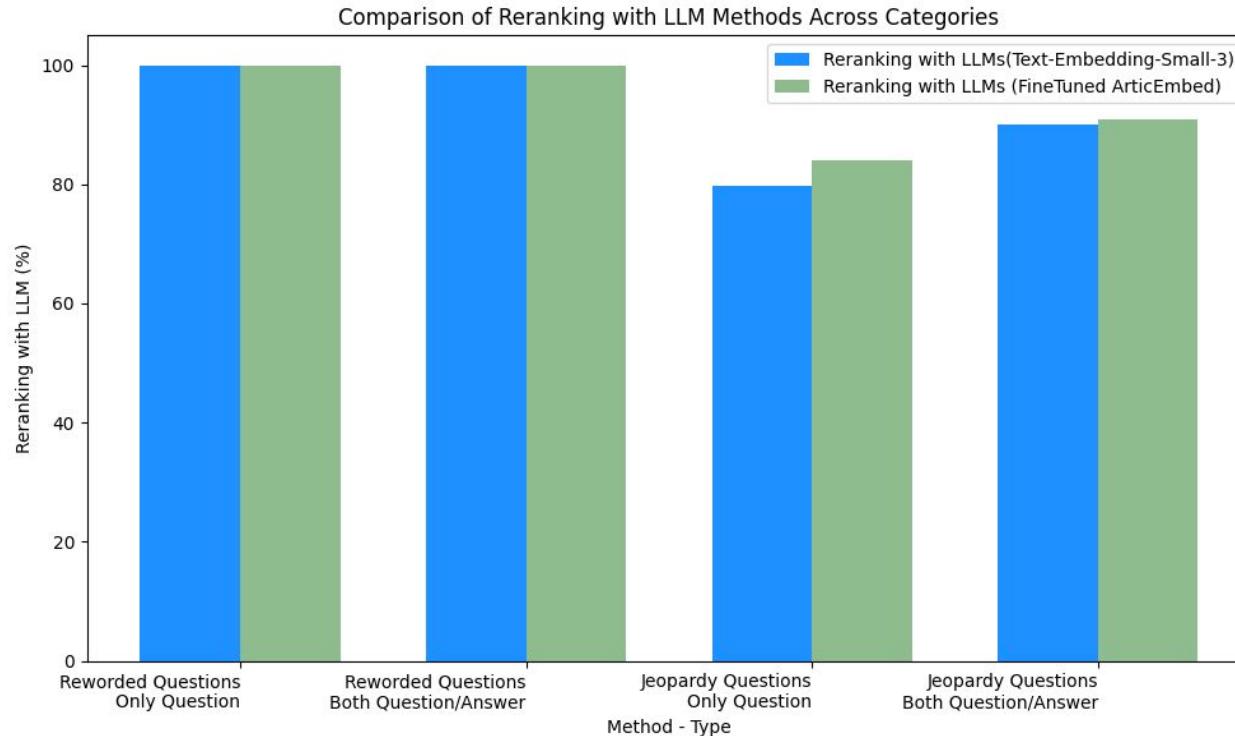
LLM + Human Exhaustive QA

- Hallucination free
- Converts 60K tokens knowledge base to 2K QA pairs
 - Product Insights
 - LLM Calls
 - Knowledge

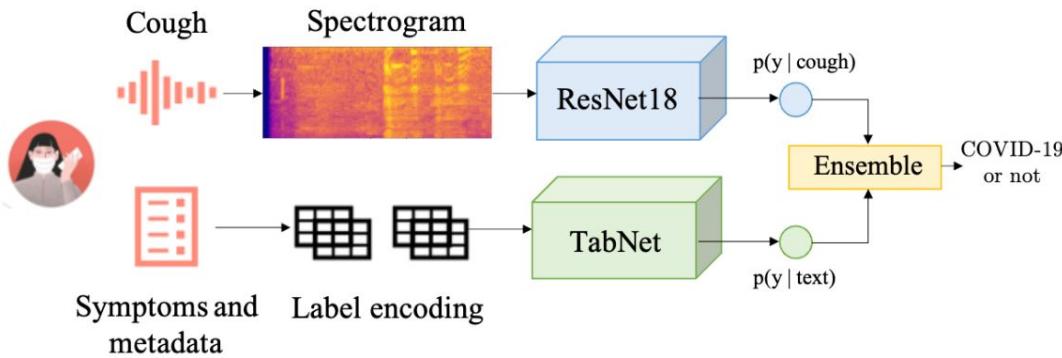




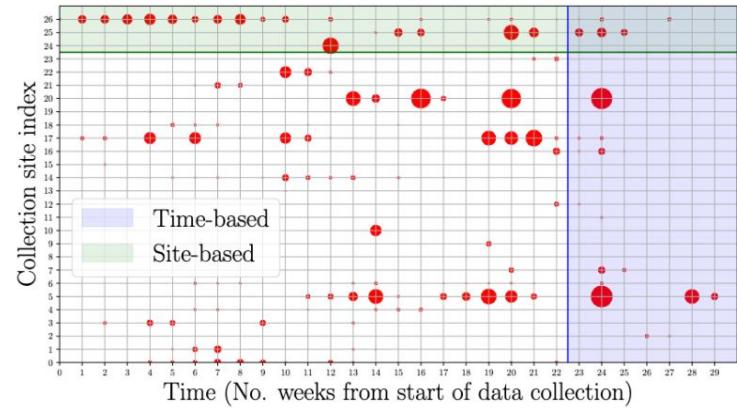
Results - Fine Tuned Embedder (33M)



Covid Screening using Cough



(a) *Schematic diagram of the proposed solution*



(b) *Data splitting strategy*

Covid Screening using Cough

Prevalence	Testing Capacity
1%	+44%
5%	+43%
10%	+41%
30%	+33%

Table 4: *Utility of our triaging tool.* We show the increase in the effective testing capacity of a system at different disease prevalence levels.

Screening TB based using Cough

Motivation:

- India accounts for ~26 % of global TB cases.
- Symptom-based screening misses partially symptomatic or asymptomatic cases.
- X-ray access is limited in many community settings.
- Need for a low-cost, deployable screening tool superior to symptom screening and more accessible than radiography.

Screening TB based using Cough

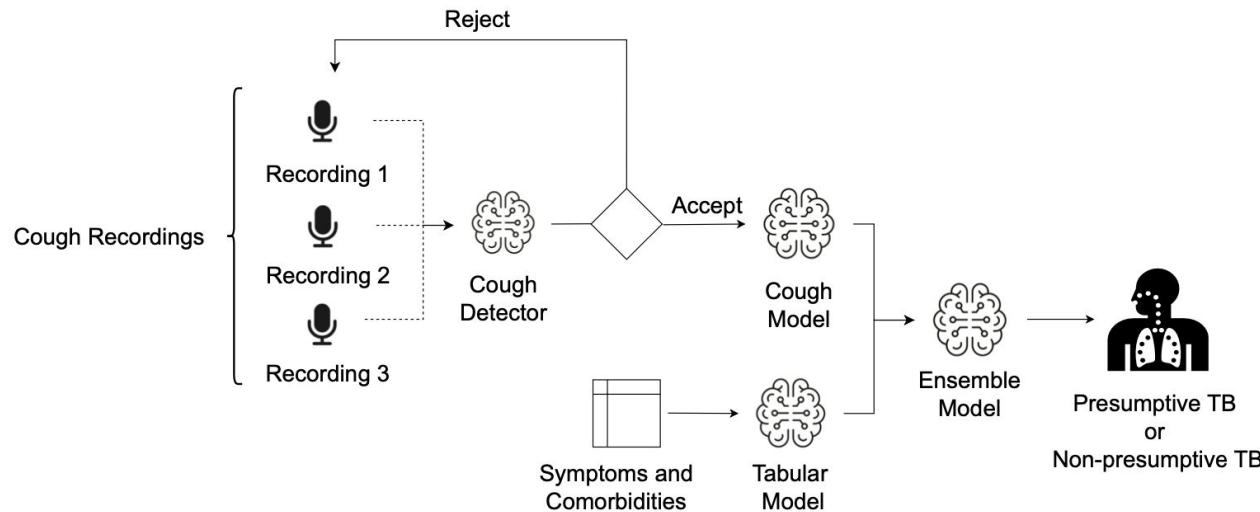


Fig. 4 An overview of the AI-based Pulmonary TB screening tool. The tool takes as input three solicited cough sound recordings which are validated using a cough detector and input to the cough TB model to obtain a likelihood score for the presence of TB. The symptoms, their duration, and comorbidities are input to the tabular model which gives another TB likelihood score. These two scores are combined and the resulting output thresholded to determine if the subject has presumptive TB or not.

Screening TB based using Cough

Key Results & Impact

Test-Set Performance (n = 1 551; TB prevalence = 53 %):

- Sensitivity: 91% Specificity: 69 %

Field Deployment:

- Screened 113063 subjects; flagged 17973 as presumptive TB.
- Of 5584 with diagnostic follow-up:
 - 359 confirmed TB cases (6.4 %), including **50 cases (14 %) missed by symptom screening.**

Implications:

- Detects partially/asymptomatic TB cases, reducing missed diagnoses.
- Scalable for low-resource settings without radiography, augmenting TB control efforts.

Day 2

Early pest warning and advisory system

CottonAce is an AI-powered early pest warning and instant advisory system developed by Wadhwani AI.

This pest management system helps cotton farmers to **protect their crops** by determining the right time to spray pesticides through **immediate and localized advice**, and helps extension program officers and administrators to monitor the solution.

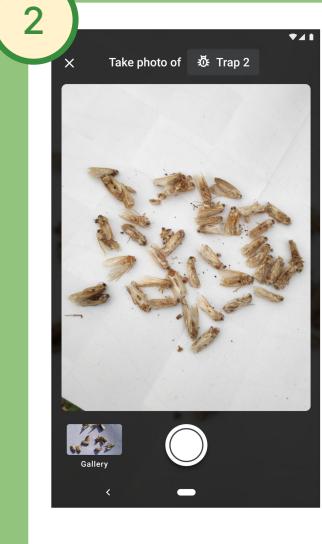


Early Pest Warning and Advisory System

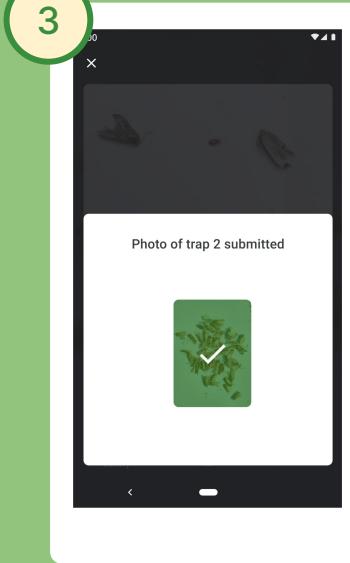
Using computer vision to count and identify pests infecting cotton crops and provide advisory to smallholder farmers.



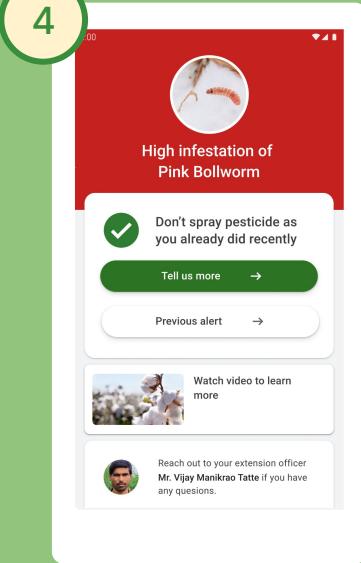
1



2



3



4

Users empty the pests caught in the pheromone trap installed at their farm.

Next, they capture two photos of the trap catch using the CottonAce app.

The AI technology detects and counts the number of pests found in the trap.

The CottonAce app generates a real-time advisory, based on scientific action thresholds.

Cotton in India...

... is important

- Leads the world in*
 - Production: ~25% of world total
 - Area under cultivation ~41% of world total
- Has an estimate 6 million cotton farmers
 - More than 75% are smallholder
 - And an additional 40-50 million people engaged in related activities

... is not easy

- Pest attacks responsible for ~30% crop yield loss
 - 70% estimated proportion of pest damage caused by bollworms
- Half of India's total pesticide usage is on cotton

Pink bollworm may eat up half of Maharashtra's cotton crop

Warning of unprecedented

Maharashtra cuts cotton forecast on worm infestation

November 17, 2017 12:34 am | Updated on

Maharashtra has cut its forecast for output of the fibre by 37 percent from its September outlook as a pest infestation has reduced yields.



Published: 24th January 2018 04:41 PM | Last Updated: 24th January 2018 04:41 PM

In Maharashtra, vicious cycle of cotton farmers' woes continues

Hindustan Times | By Ketaki Ghoge and Surendra P Gangan, Mumbai

Feb 18, 2018 12:18 AM IST



worry is that the pink bollworm present in the soil can attack the crop again in the upcoming kharif season.

The cotton lifecycle



The cotton lifecycle



Seeds Seedlings Buds Blossoms Bolls



Pest traps

- Traps that can be placed in the field to monitor pest populations
- A recent push by the government to get more farmers using them

Traps are mounted throughout the field



- Lure emits a pheromone that attracts males
- Different lures attract different bollworms



Bottom opens to release trap catch



Data

Modelling

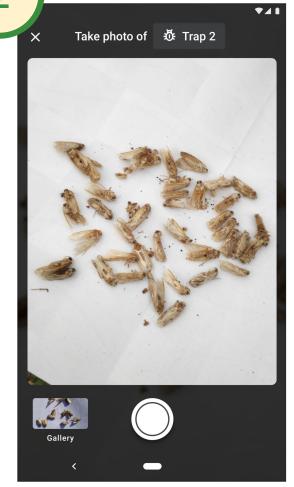
Evaluation

1



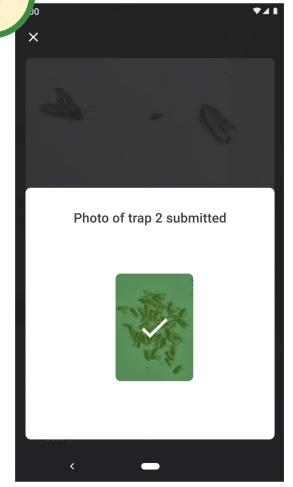
Users empty the pests caught in the pheromone trap installed at their farm.

2



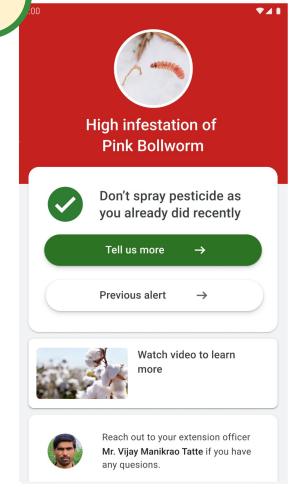
Next, they capture two photos of the trap catch using the CottonAce app.

3



The AI technology detects and counts the number of pests found in the trap.

4



The CottonAce app generates a real-time advisory, based on scientific action thresholds.

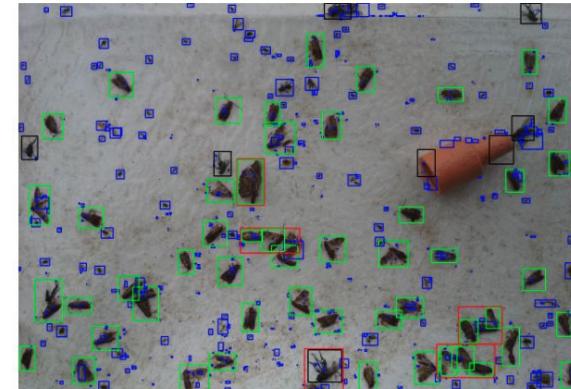
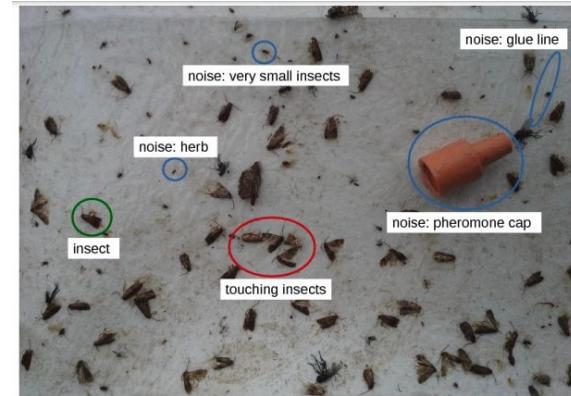
Existing datasets



Ding and Taylor, 2016

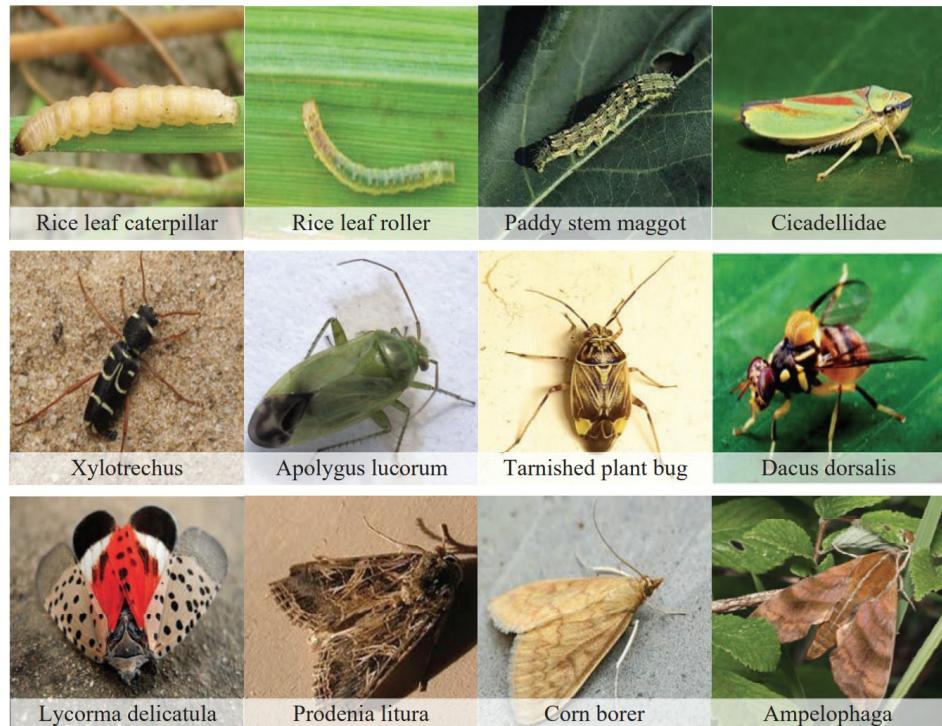


Bakkay et al., 2018



IP102 [Wu et al. 2019]

- 75K images across 102 categories
- Images extracted from internet search engine results



Data collection: process



BCI | RV-1 / P-2 / 01-10
INMH 8808017

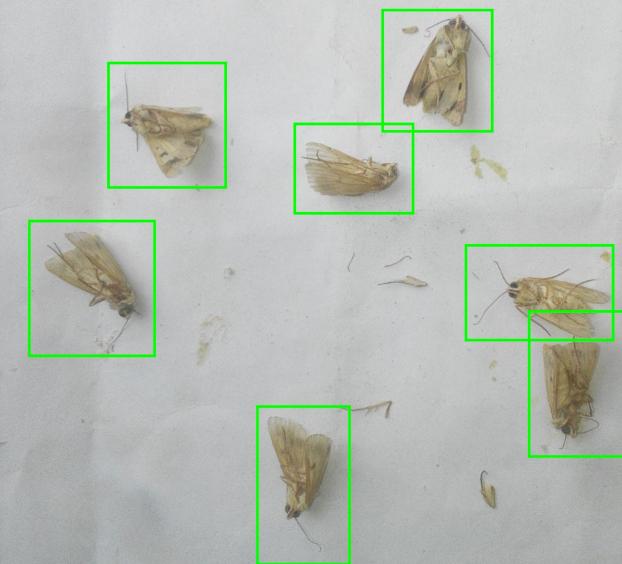


Pink bollworm



American bollworm

N.BCI/SG/P-1/B-12



Data

Modelling

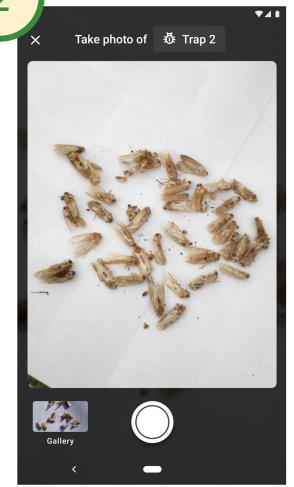
Evaluation

1



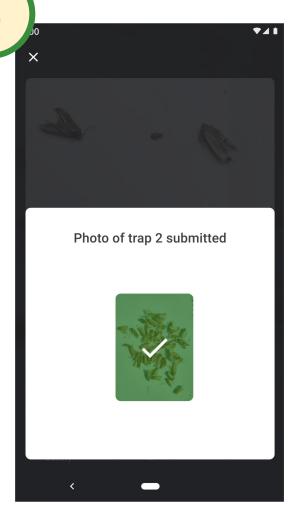
Users empty the pests caught in the pheromone trap installed at their farm.

2



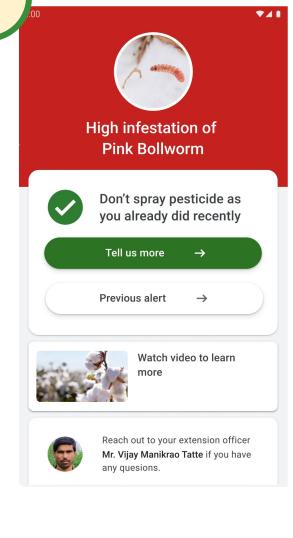
Next, they capture two photos of the trap catch using the CottonAce app.

3



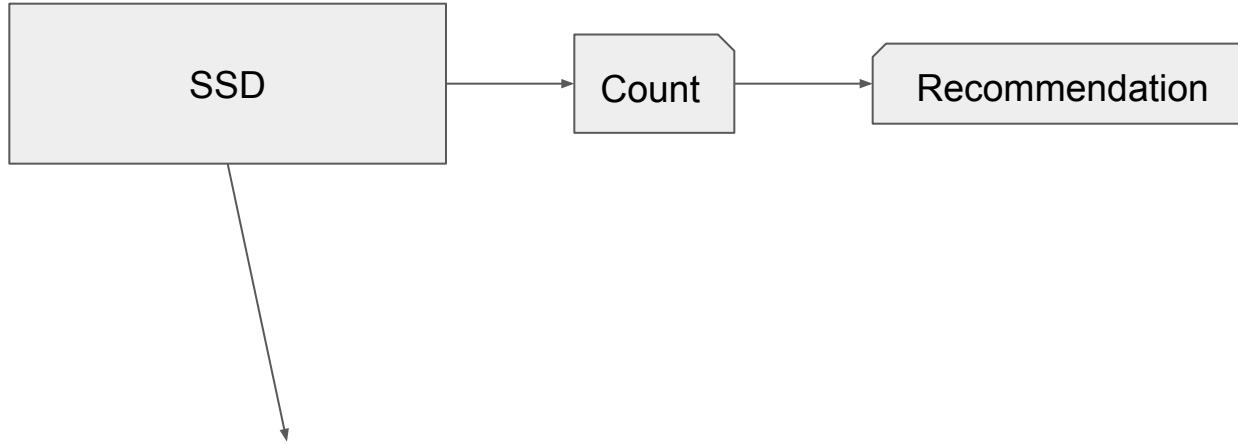
The AI technology detects and counts the number of pests found in the trap.

4



The CottonAce app generates a real-time advisory, based on scientific action thresholds.

System workflow V1 (~2018)

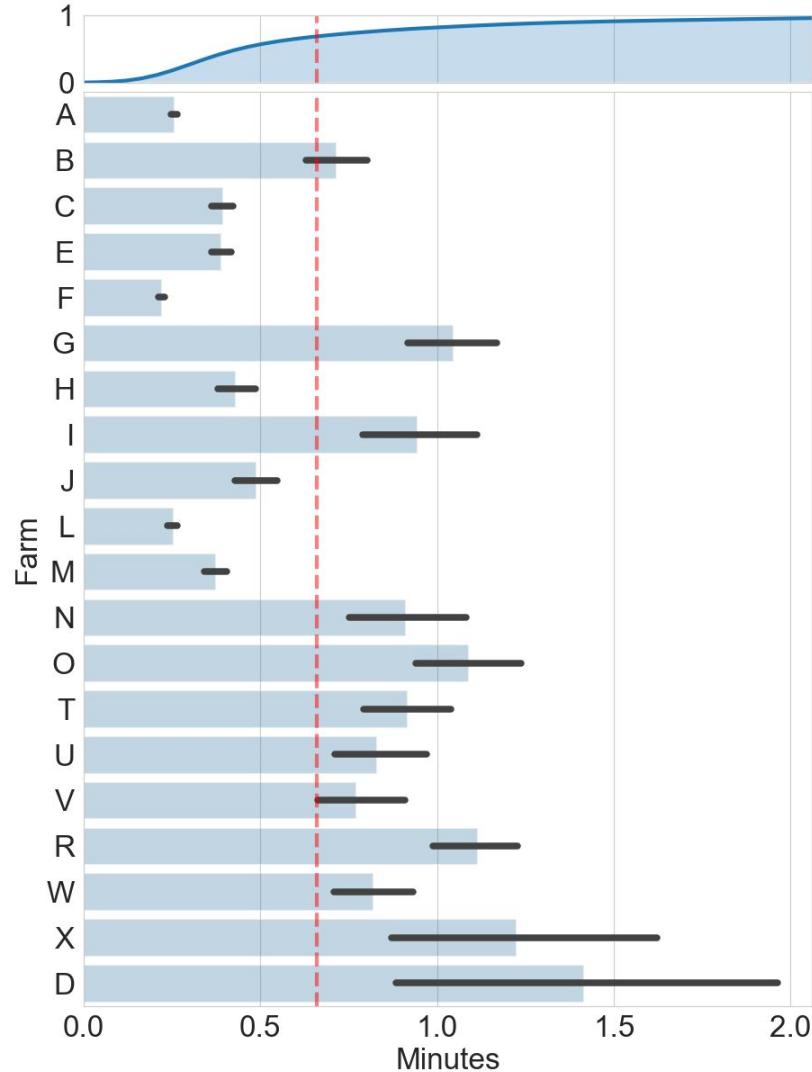


Single Shot MultiBox Detector [Lui et al 2016]

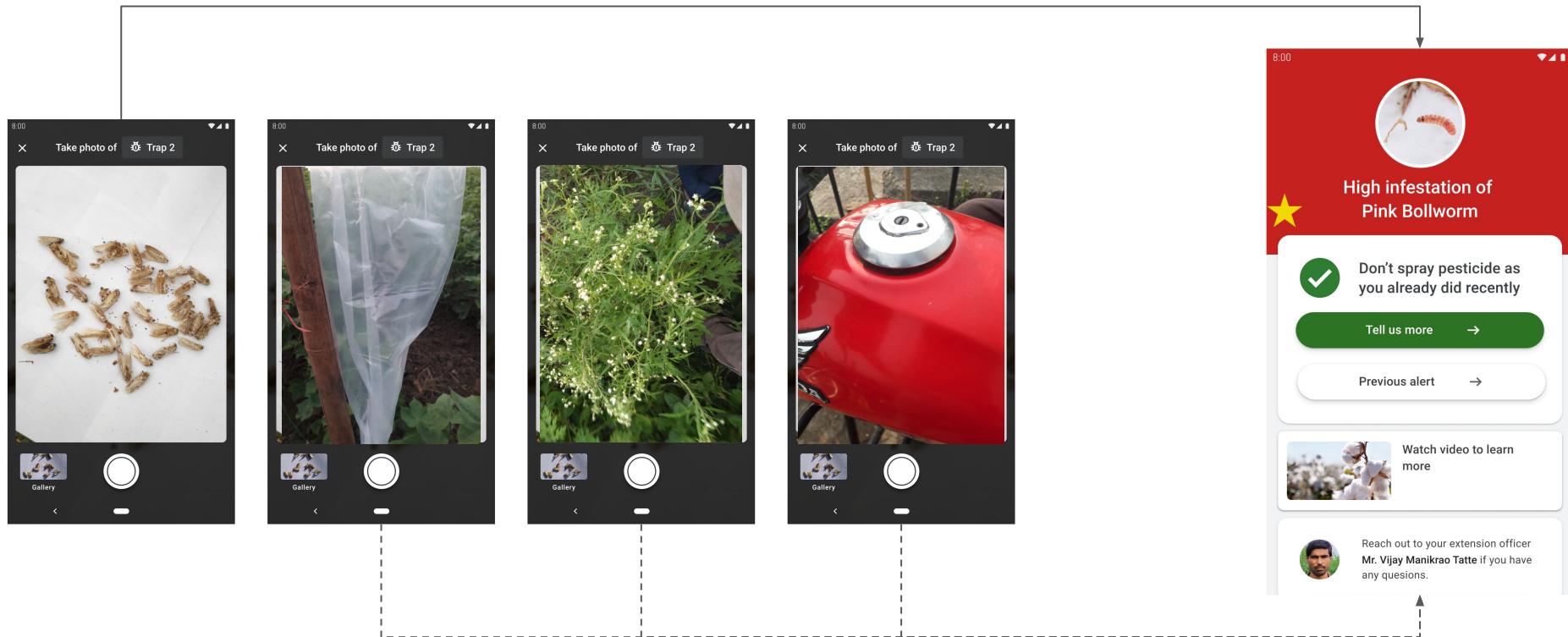
- Seminal architecture in object detection timeline
- Stable, established open source implementation in PyTorch
- Found to perform well on insect detection [Nam and Hung 2018]

Early lesson: upload time

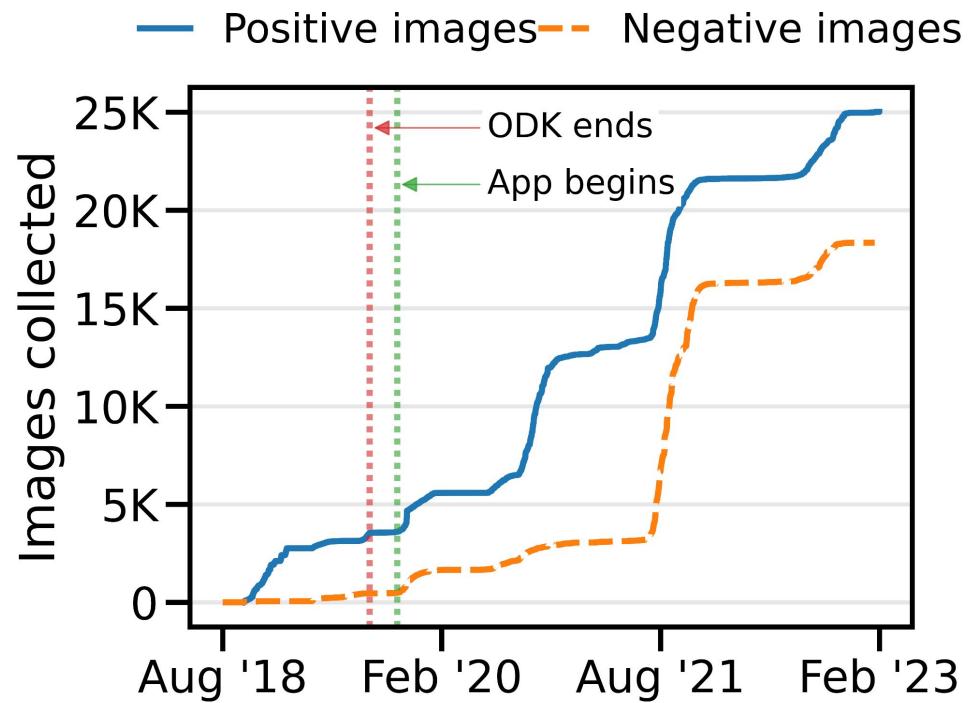
- Small study in early 2019
- Very early version of Cotton Ace
- Figure: Average upload time per field
- *Should not rely on internet!*



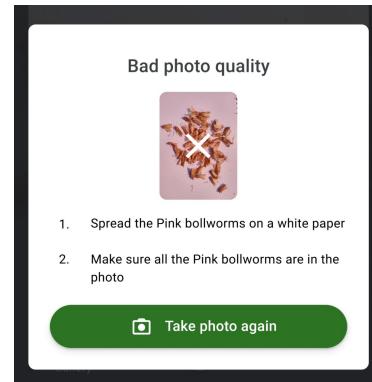
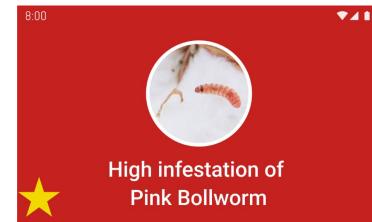
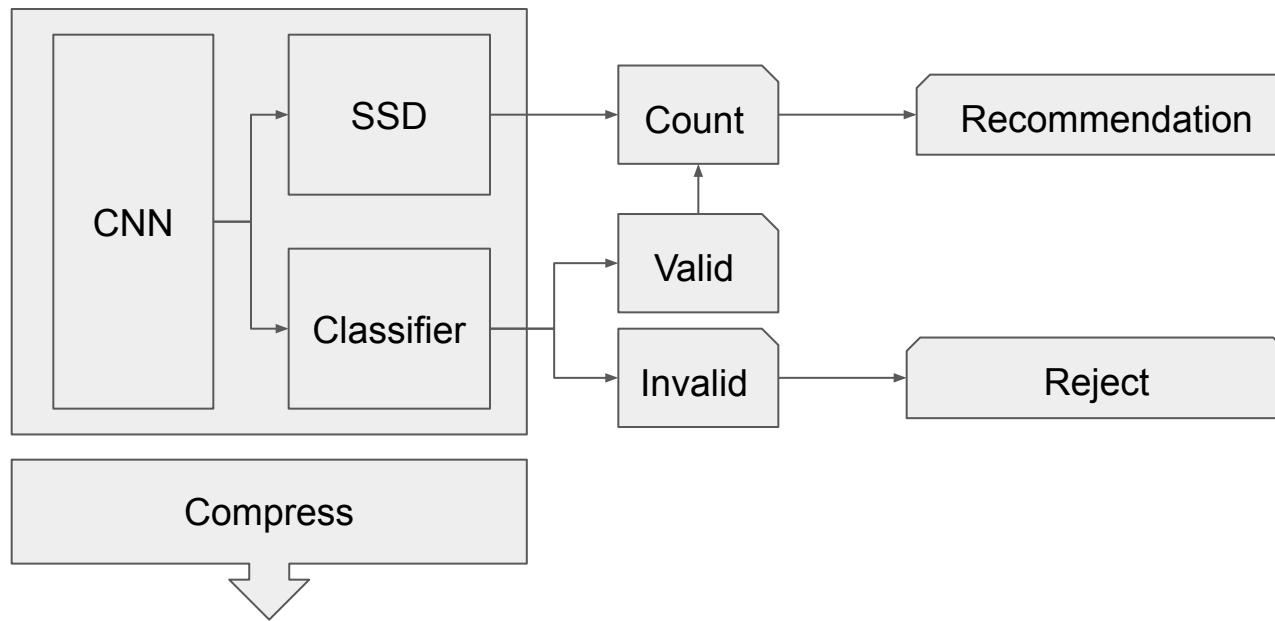
Early lesson: interaction varies



Scale of the negatives



System workflow V2 (~2019)

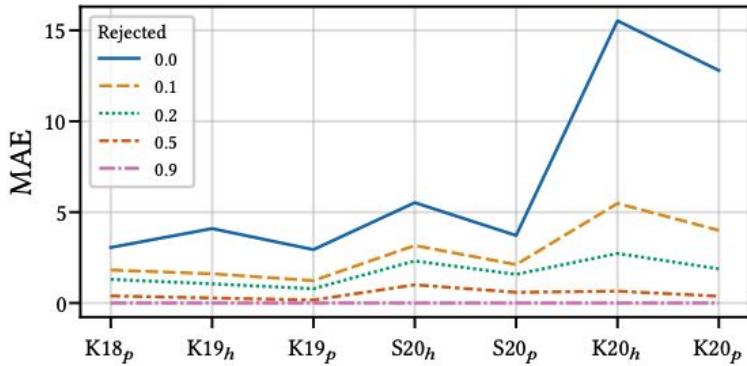


2022: A combination of ideas

Our current model (2023) is a result of several ideas coming together

1. Alert abstention
2. Alert delay
3. Model development using the crowd

OOD Rejection - Saw Tooth

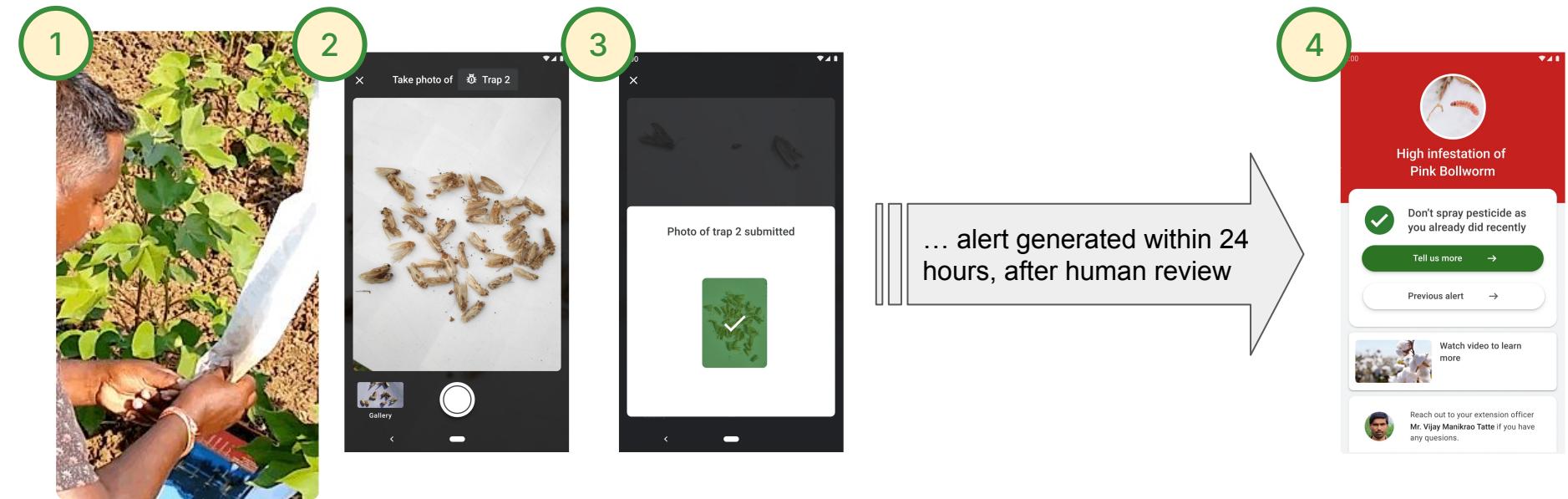


	Train	Test
S20 _h	<ul style="list-style-type: none">• Data up to summer 2020 start	Data from summer 2020
S20 _p	<ul style="list-style-type: none">• S20_h + data from that season	Data from summer 2020

Figure 1: Model performance across seasonal (2018–2020) development and test sets. With each season performance decreases (increasing X_p 's), but can be improved by training on that seasons data (decrease in immediate X_h 's). Performance can improve further by rejecting some fraction of worst-performing samples.

2022: A combination of ideas → *Alert delay*

- During the 2022 season, a group of users did not receive alerts immediately
- Instead...



2022: A combination of ideas → *Using the crowd*

ZindiWeekendz
Virtual hackathons

WEBINAR

Wadhwani AI Bollworm Classification Challenge

Can you improve a pest control app by classifying if an image contains a bollworm moth or not?

Fri, 07.10.22 | 11AM – 12PM GMT

Register today!

\$300 USD | 500 points

Harvard John A. Paulson School of Engineering and Applied Sciences

Wednesdays @ 12:45pm - 3:00pm
SEC LL2.223 (Allston Campus)

HOME PROJECTS PARTNER WITH US MATERIALS

Capstone Research Project Course

AC297R, FALL 2022
Weiwei Pan

WADHWANI AI
german cooperation
Implemented by giz

FAIR FORWARD Artificial Intelligence for all.

ZINDI

Wadhwani AI Bollworm Counting Challenge

Can you improve a pest control app by counting the number of bollworm moths per image?

€15 000 EUR | 15 000 points

Overall performance

Objective:

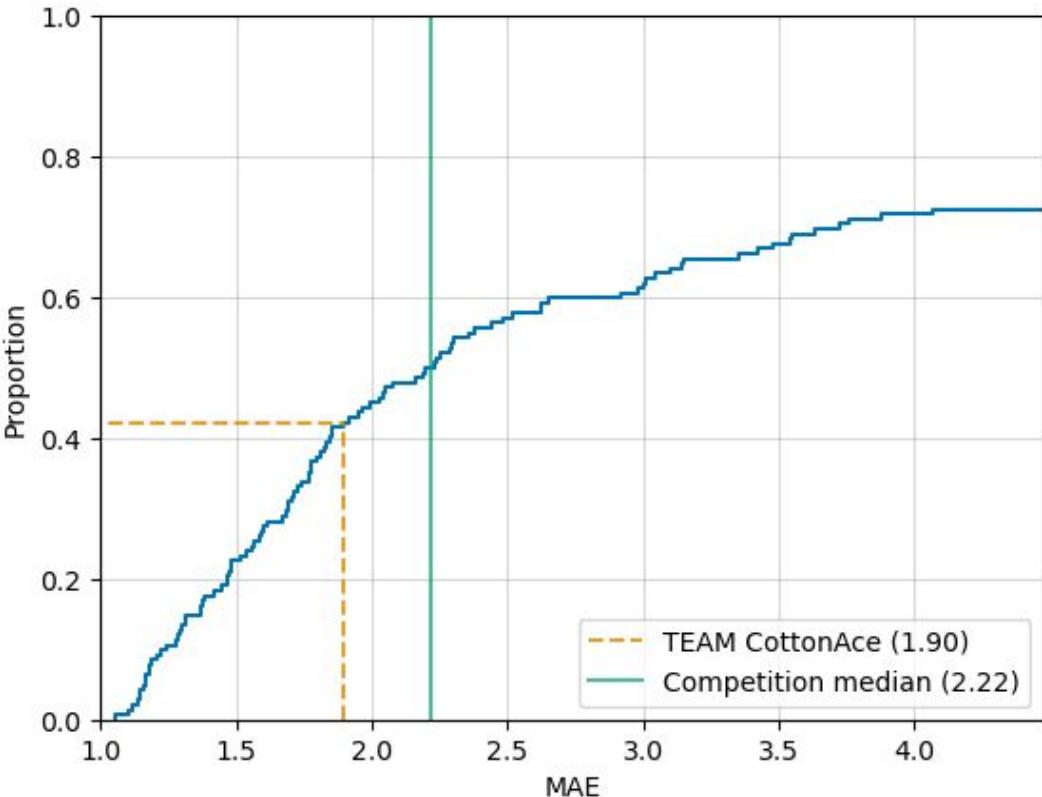
- Classify pests in an image and count them
- Correctly provide zero for images with no pests

Figure

- Proportion of teams with errors less than a given value
- MAE based on complete (public and private) test set

Takeaways

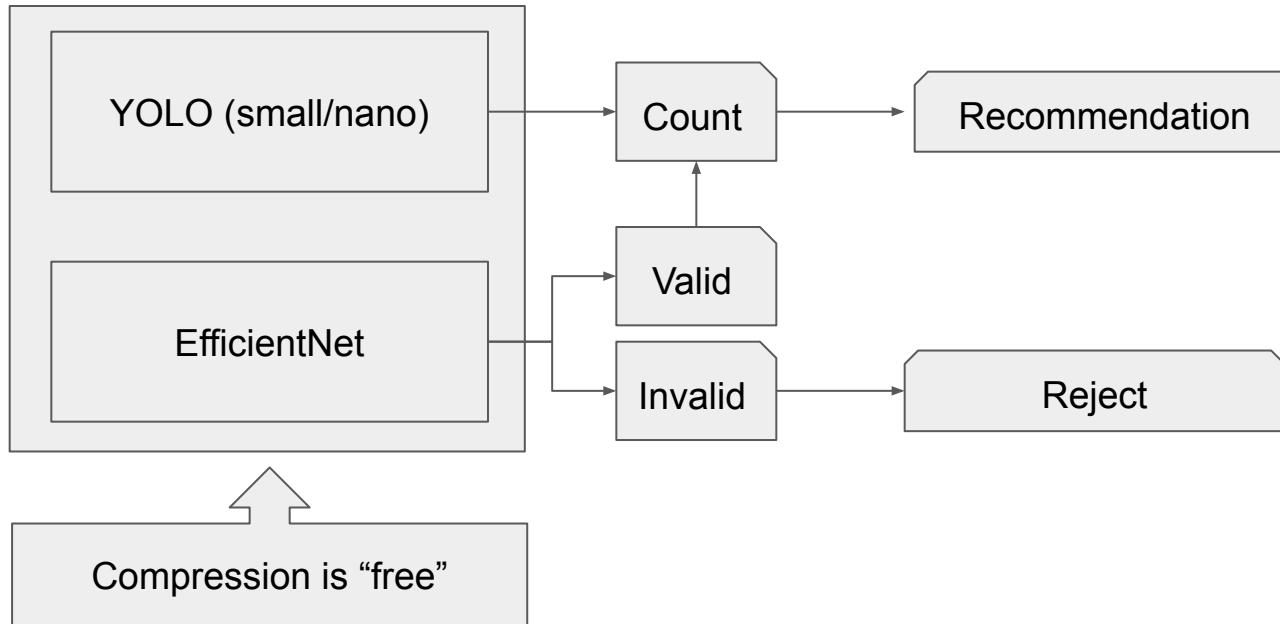
- Very strong numbers from best-performing teams
- Long tail is to be expected
- Our models performed better than most
 - Lots to learn from the crowd!



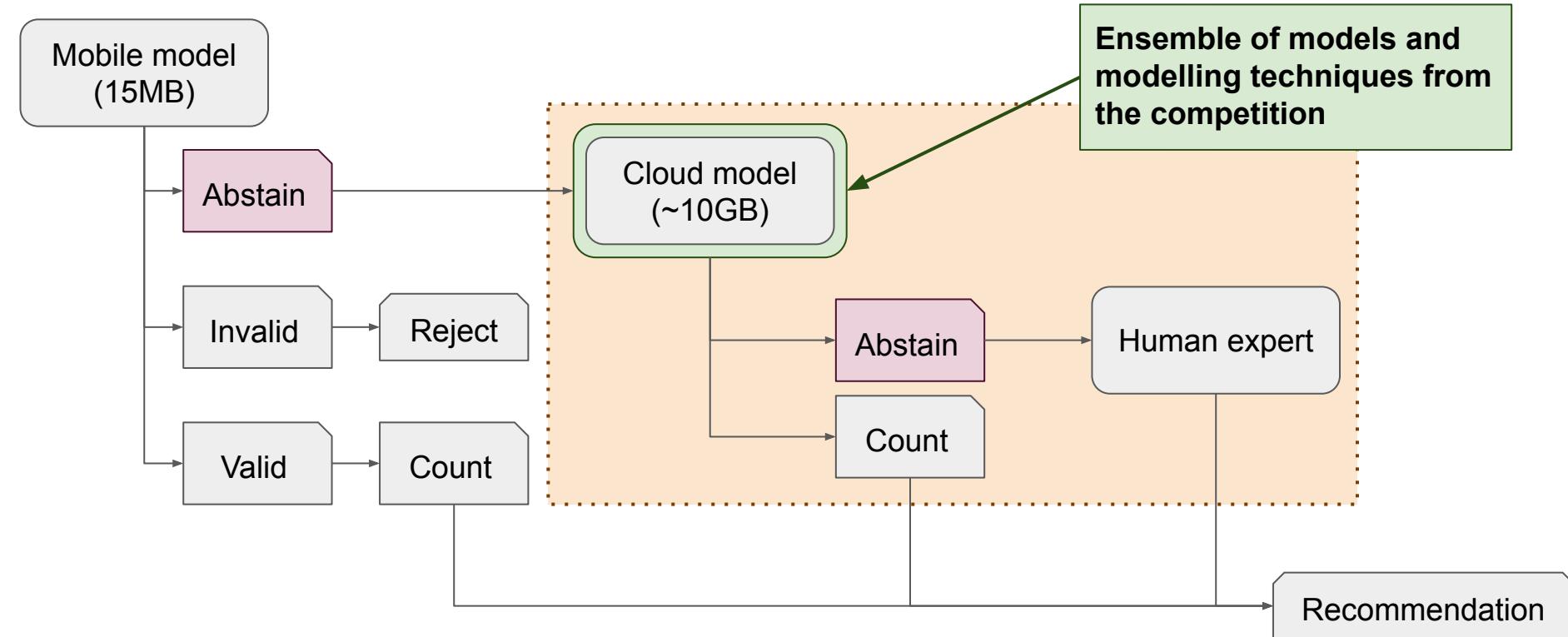
Notable winning techniques

YOLO	<ul style="list-style-type: none">• Four of the top six models used YOLO somewhere in their pipeline• ... including first and second place• ... transformers made a middling appearance (within the six)
Large images	<ul style="list-style-type: none">• Train resolution ranged from 1024 to over 2500!
Augmentations	<ul style="list-style-type: none">• Mosaic• Mixup
NMS magic	<ul style="list-style-type: none">• NMS tuning is key (we knew this, but good get validation)• <u>Weighted Boxes Fusion</u> (WBF)
Ensembling	<ul style="list-style-type: none">• Some used test time augmentation (TTA; we'd been experimenting with this, good to see examples)• Cross validation to improve confidence

System workflow (2023)



System workflow extended (2023)



Data

Modelling

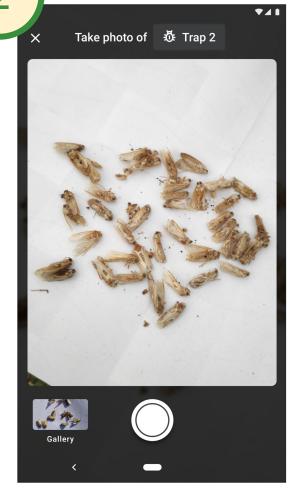
Evaluation

1



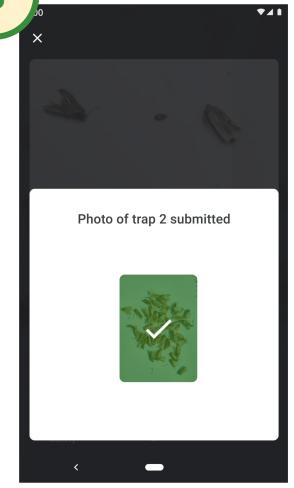
Users empty the pests caught in the pheromone trap installed at their farm.

2



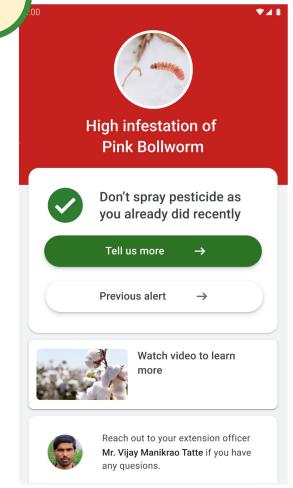
Next, they capture two photos of the trap catch using the CottonAce app.

3



The AI technology detects and counts the number of pests found in the trap.

4



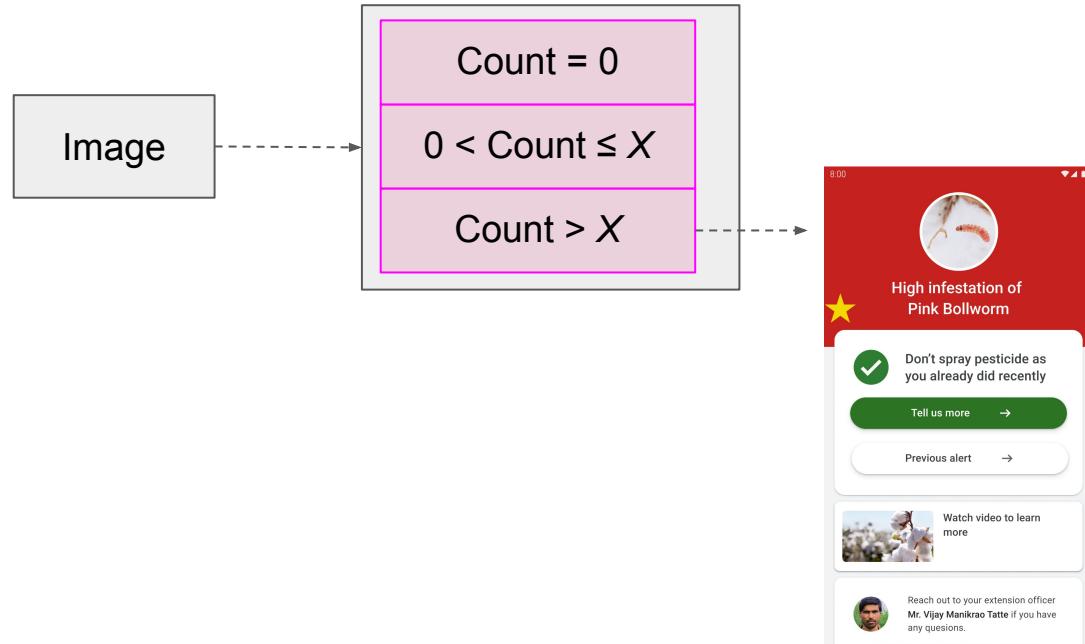
The CottonAce app generates a real-time advisory, based on scientific action thresholds.

User-centric evaluation

- Modelling produces counts
- Business logic produces categorical RED/Yellow/Green alerts

Business logic can vary, and can be difficult to model directly

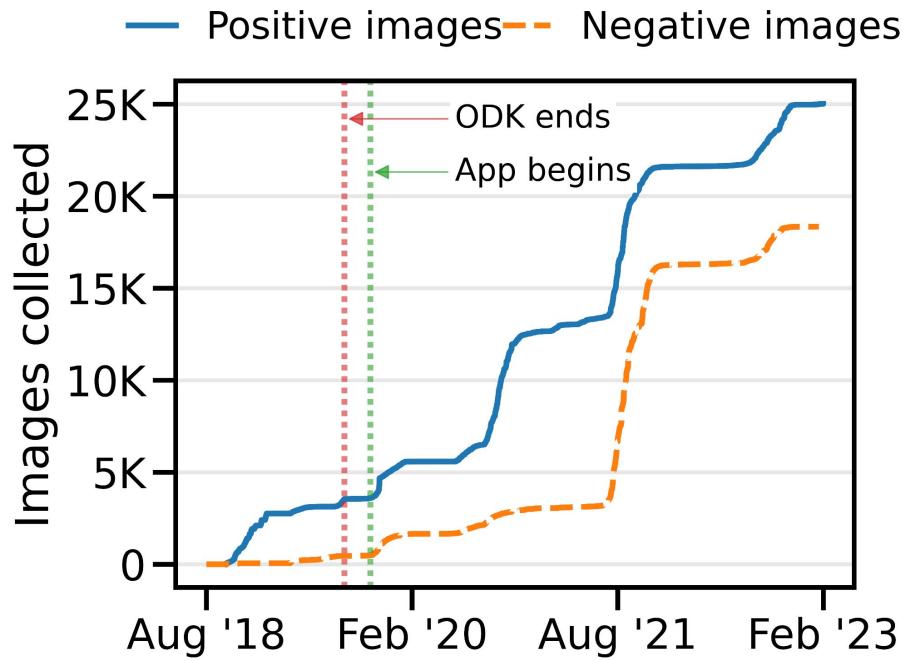
... but it's prudent to keep in mind during tuning and evaluation



Open source data: facilitating collaboration

Overview

- 36K+ images
- Captured during formal data collection and app deployment
- Mix of
 - pink and American bollworms
 - Images without bollworms



Low barrier to entry

```
$> git clone https://github.com/WadhwaniAI/pest-management-opendata.git  
$> cd pest-management-opendata  
$> aws s3 sync --no-progress s3://wadhwani-ai-agri-opendata/ data/  
$> ./bin/to-ultralytics.sh -d data -o /desired/output/location  
$> git clone https://github.com/ultralytics/yolov5.git  
$> cd yolov5  
$> python train.py --data /desired/output/location/config.yaml ...
```

```
from datasets import load_dataset # HuggingFace!  
dataset = load_dataset('wadhwani-ai/pest-management-opendata', streaming=True)
```

Scaled Social Impact

- Reduction in pesticide usage by **20%**
- Translates to increase in profit by **25%**
- Reached 500K farms in 2023
- Adopted by National Pest Surveillance System (NPSS)
- Awards and Recognition
 - \$3.3M grant from Google.org 2023
 - Global Change Award Winner - H&M Foundation 2021
 - Fast Company World Changing Ideas 2021
 - \$120K grant from GIZ for kaggle-like competition

Building AI Solutions

Beginning AI for Social Impact - Story (~2017)

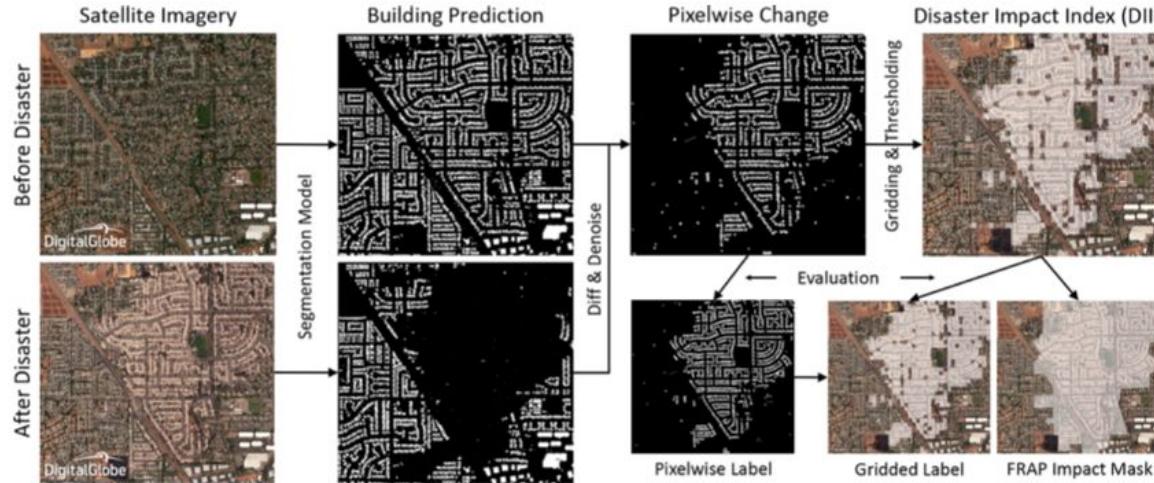


Figure 1: Flow diagram of our proposed approach for disaster impact analysis. We run pre-trained CNN on satellite imagery before and after disaster, compare the change in extracted man-made feature, then compute Disaster Impact Index (DII) to understand impact of each area

Doshi, J., Basu, S., & Pang, G. (2018). From Satellite Imagery to Disaster Insights. *AI for Social Good, Neural Information Processing Systems (NeurIPS) Workshop*.

Boin, J.-B., Roth, N., Doshi, J., Llueca, P., & Borensztein, N. (2020). Multi-class segmentation under severe class imbalance: A case study in roof damage assessment. *Humanitarian Assistance and Disaster Response, NeurIPS Workshop*.

Wadhwani Institute for Artificial Intelligence



Non-profit dedicated to making positive social impact

- By developing and deploying AI/ML based solutions
- Focusing on underserved communities in India and other developing countries

Inaugurated in February 2018



P. Anandan

Born Madras, India
Alma mater Indian Institute of Technology Madras (B.Tech, Electrical engineering)
University of Nebraska (Masters in Computer Science)
University of Massachusetts, Amherst (PhD in Computer Science)

How to pick a problem? 8 Questions

Impact potential: Is it a big problem? What is the specific problem that AI is expected to solve, and in what settings? Will solving that specific problem have large enough impact?

AI feasibility: Is an AI-based solution feasible, technically? Are there logical and scientific reasons to believe the envisioned AI model is possible? Does the data exist or acquirable?

AI necessity: Will AI be sufficiently better than other (potentially easier) approaches to warrant the (additional) effort?

Full solution + System design: Will AI make enough difference, given other issues? What else needs work (workflows/processes upstream, downstream, in parallel) for the full system-integrated solution to have expected outcomes?

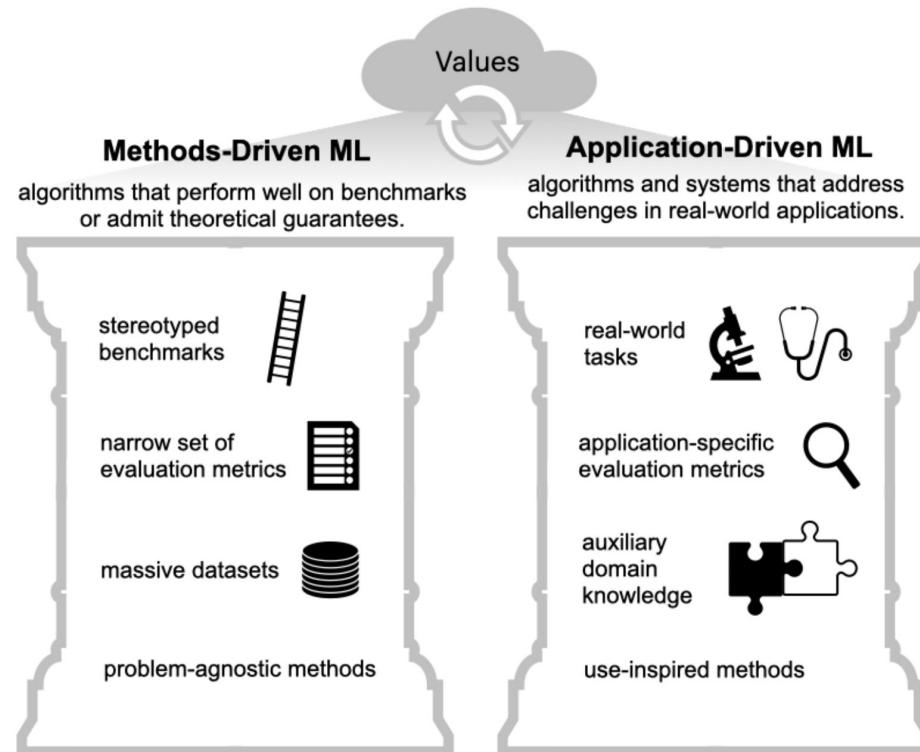
Acceptance & Usage: Will stakeholders (chooser, payer, user, beneficiary) accept it and use it the way it's envisioned?

Partners: Who are the partner orgs to co-innovate, co-deploy, iterate, pilot with? Do they have the complementary competencies required?

Scaling: Who are the owners, programs and pathways for scaling? Can the all the (technical and non-technical) components of the solution scale to ensure financial and operational feasibility?

Why Us: What makes us uniquely qualified to work on this? Is this the most highest leverage work I could be working on?

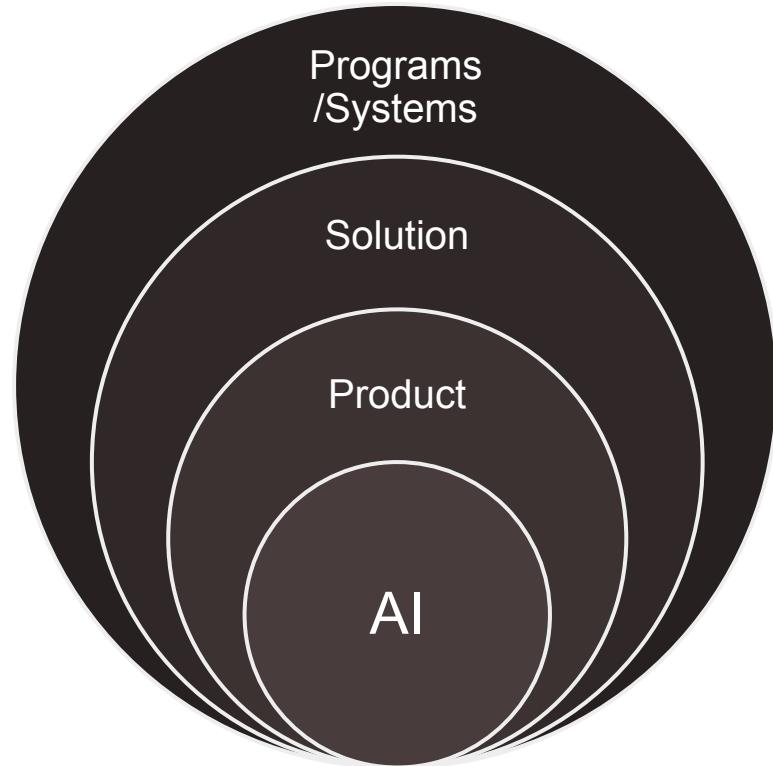
Methods Driven ML vs Application Driven ML



Source: <https://arxiv.org/abs/2403.17381>

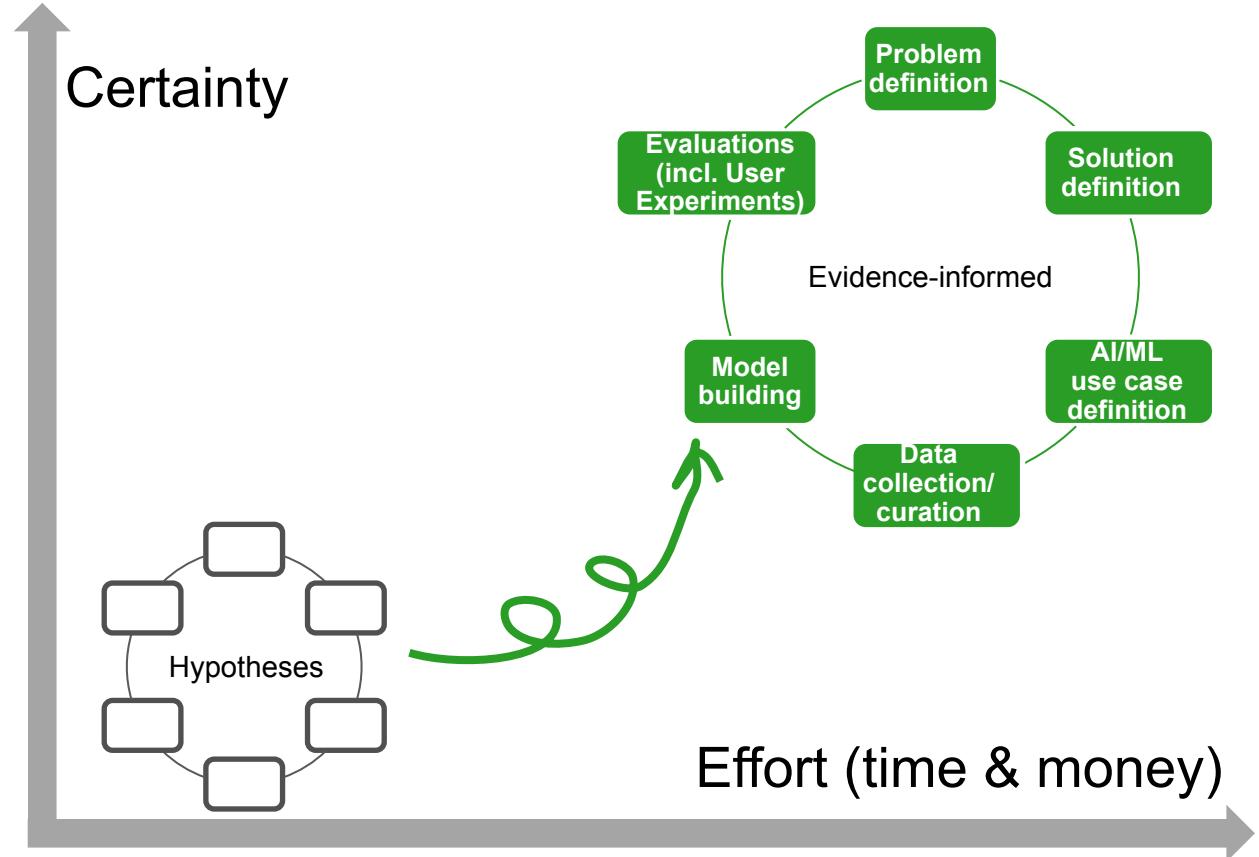
AI Solutioning

AI can play a potentially transformative role, but only **as part of complete** solutions integrated with systems for impact at scale



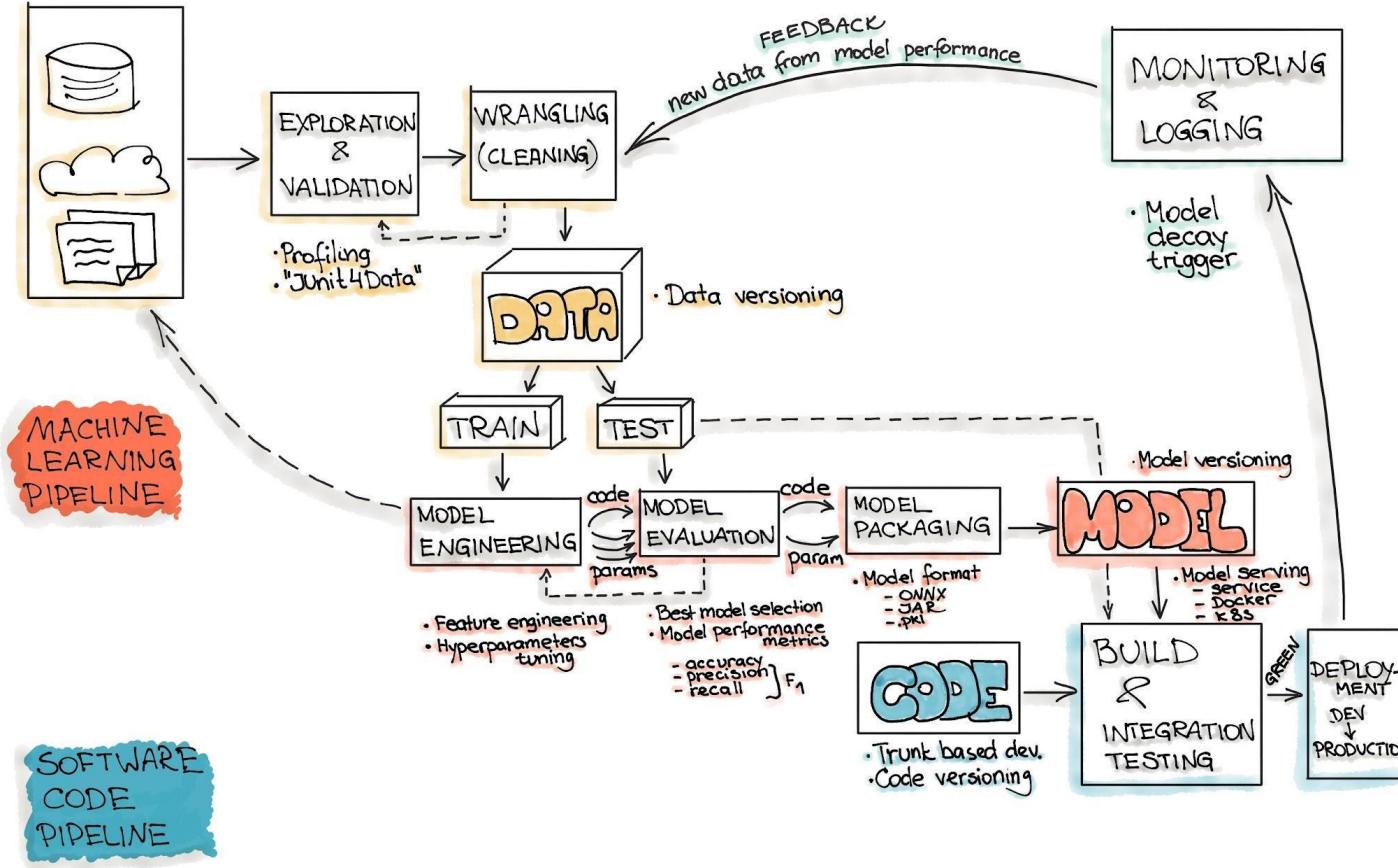
A mental model
to help imagine
the real process of
building ~~AI~~
AI-based solutions
responsibly

*The whole process is
about making key
interconnected choices
with increasingly greater
certainty, iteratively,
while figuring out
trade-offs and
generating evidence.*

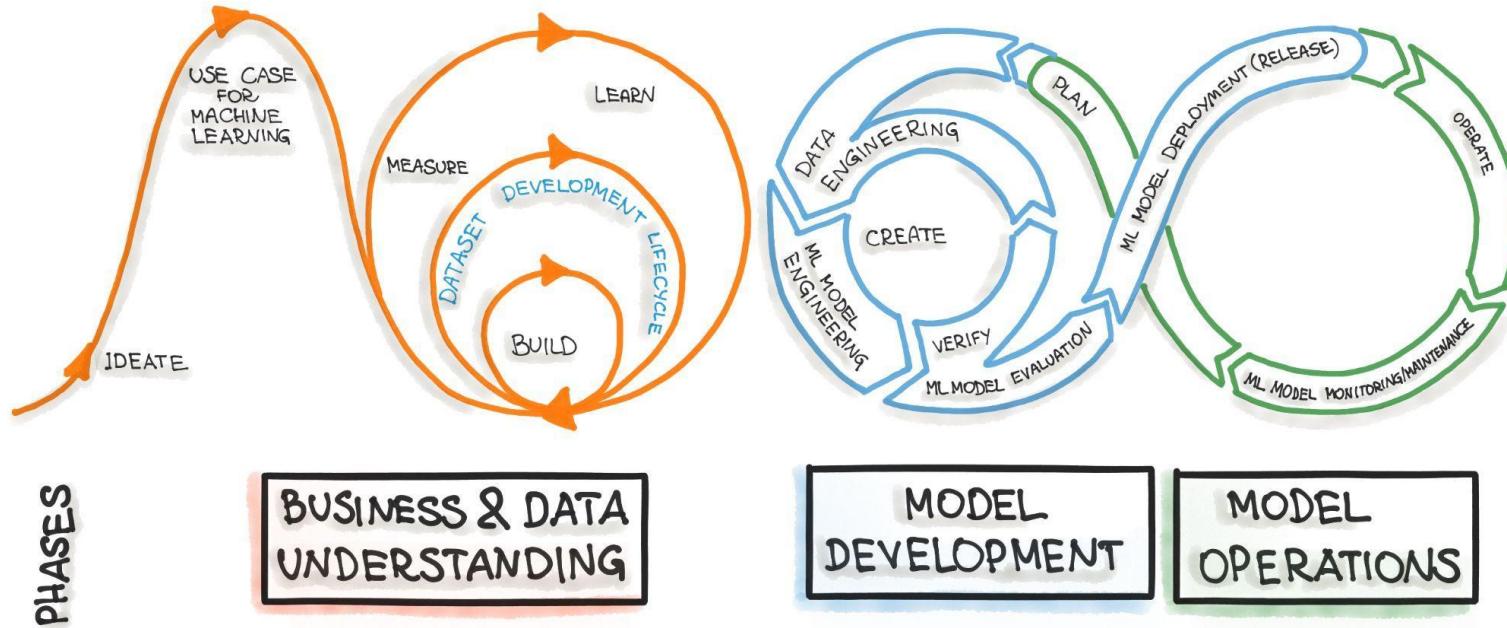


DATA PIPELINE

MACHINE LEARNING ENGINEERING.



CRISP-ML(Q)



@visenger

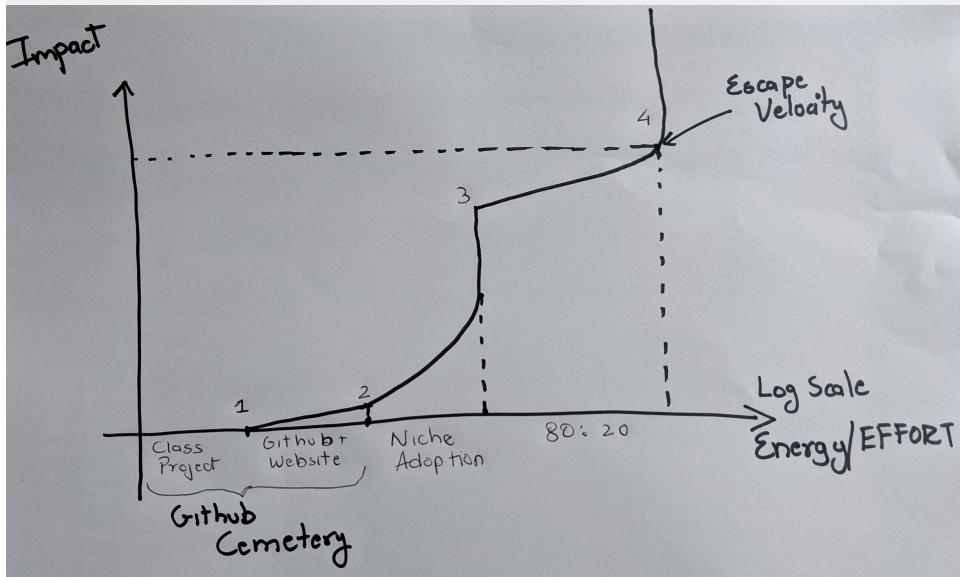
Principles and Heuristics

1. Imperfect solution to a right problem > perfect solution to a wrong problem
 - a. Breadth-first. Not depth-first
2. Premature optimization is the curse of ML (and any field)
 - a. Start simple.
 - b. Simple, low-tech debt solution to complex but highly performant
3. No perfect launch. It is a journey.
 - a. Iterate quicker. Fail faster.
 - b. Prioritise where to improve, not what you think is important
4. “So what” > Why > What > How (solution)
 - a. Insist on “So What?” seven times.
 - b. But we often start with a solution (or a technology)
5. Err on “data” side.
 - a. Collect more but purpose driven (even if the purpose is anticipatory).
6. Outputs will be wrong
 - a. Abstain when not sure (can the UX support it?)
 - i. Models need not make decision all the time
 - ii. Abstention is a lever, in addition to “scope (functionality)”, “resources (time, compute, human)”, “quality”.
 - iii. Design is about tuning these four degrees of freedom via negotiation.
7. Not all decisions are equal.
 - a. Cost-aware decision making.

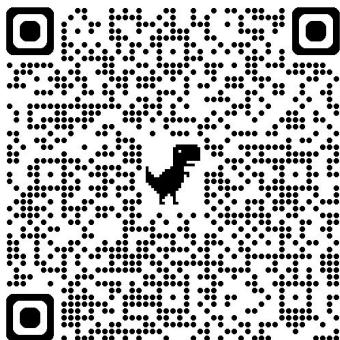
How develop taste for problems?

- Opinionated First Principled Framework
 - Find your **values**, what change do you wish to see in the world?
 - You must be willing to live that yourself first
 - Boil it down to 2 or 3 values
 - Given a set of values, find large '**unsolved**' domain
 - Hill you are willing to 'die' on
 - <https://encyclopedia.uia.org/problems/1>
 - Climate Change, Mental Health, Financial inclusion, Epidemics, Social Injustice, Drug Discovery Misinformation, etc
 - **Pursue it** with all your time, energy and attention
 - Be open minded on the 'how' to get there. Startup, Non-profit, Academia, Content Creator etc
 - Immerse yourself within these communities or build your own
 - Aim higher than your imagination. Though pursue it with integrity, humility and honesty
 - Don't chase status, awards, recognition. Make enough money to continue to pursue this.

Effort vs Impact Curve



Thanks!
jigarkdoshi.com



Outline

- Jigar's Journey into Social Good
 - Story time with Disaster Mapping and Joining Wadhwani AI
- Problem Selection
 - 8 + 1 Question Framework
 - Innovative Taste
- LLM for FLWs
- Anthropometry
- TB & Covid Cough
- Building Solution Journey
 - How to
- Satellite Imagery
- Pest Management
- Future: Agents for Public Health System