National Children's Science Congress

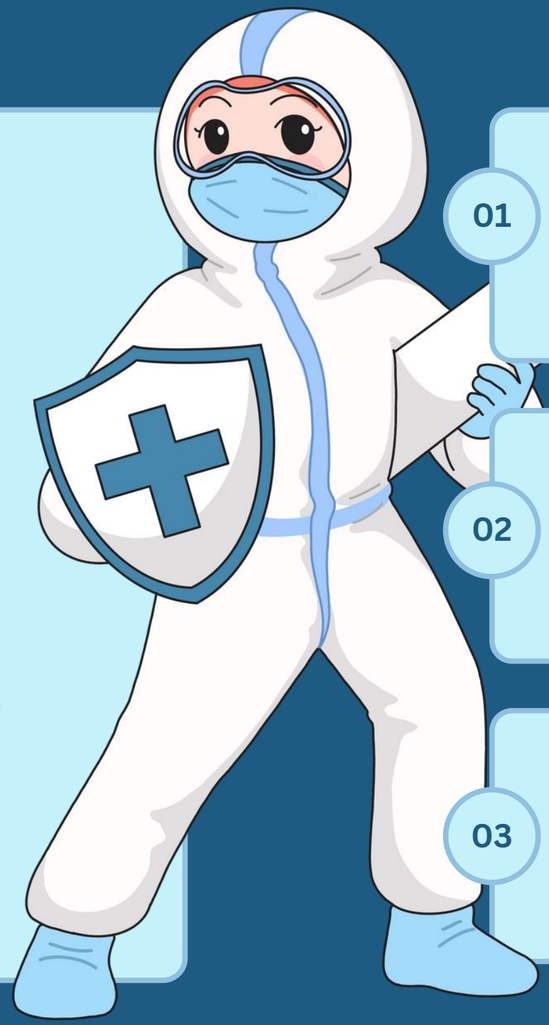
Early Detection of Pandemics using AI driven Geospatial Analysis of School Attendance Data

Sub Theme: Technological Innovation for Ecosystem and Health (V)

Guide Teacher: Varun P K (PGT Physics)
Team Leader: Krishnavyshak R
Team Member: Aditya Narayan Pandey

Project Objectives

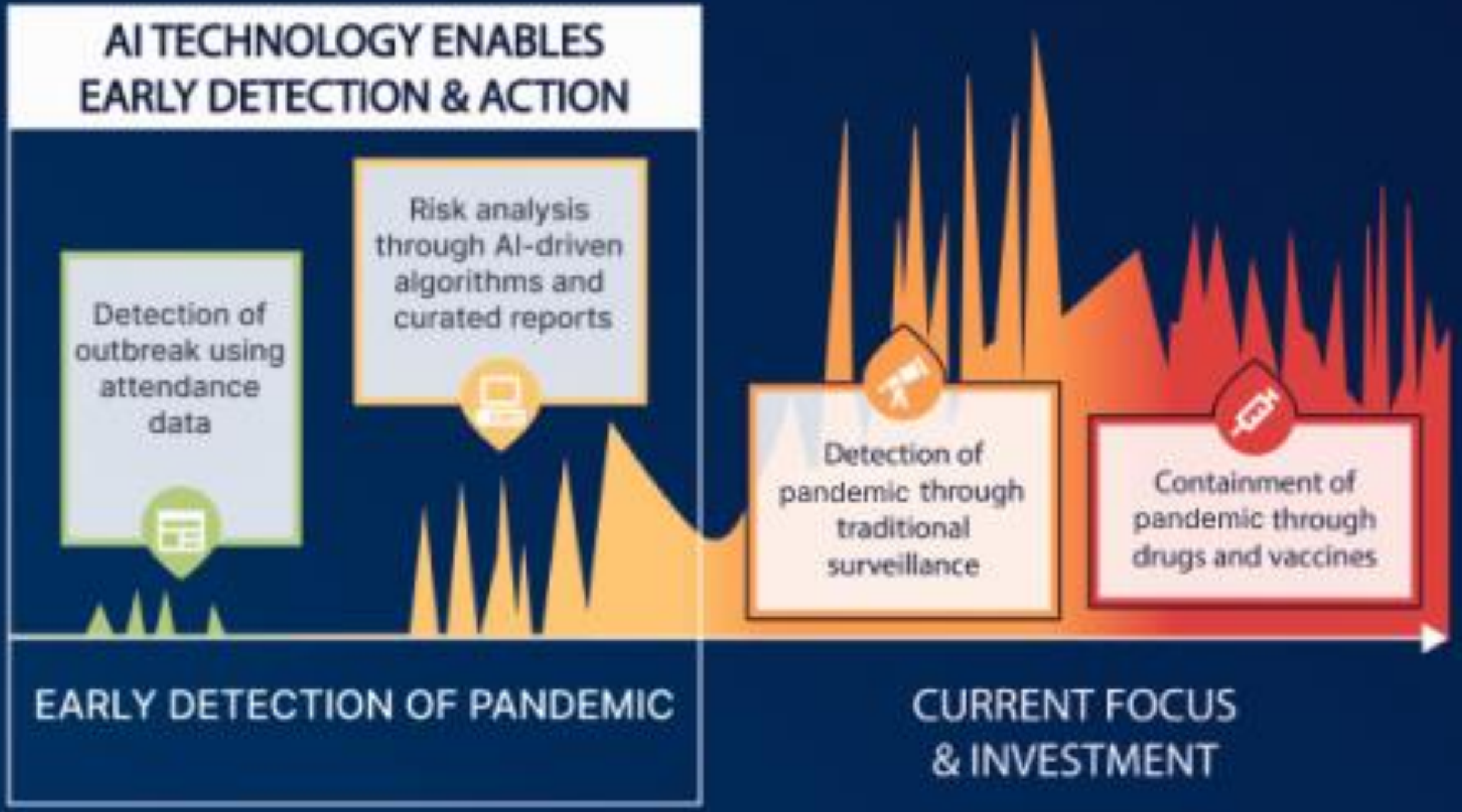
The project aims to utilize AI-driven geospatial analysis of school attendance data to detect potential pandemics early, allowing for swift response by health authorities and enhancing community safety.



01 **Early Detection:** Use AI to analyze school attendance data for early identification of potential pandemics.

02 **Geospatial Visualization:** Map student attendance, generate heat maps, and highlight areas with high absenteeism.

03 **Rapid Communication:** Enable swift alerts to health authorities for quick and targeted responses to emerging health threats.


AI TECHNOLOGY ENABLES EARLY DETECTION & ACTION
Detection of outbreak using attendance data
Risk analysis through AI-driven algorithms and curated reports
Detection of pandemic through traditional surveillance
Containment of pandemic through drugs and vaccines
EARLY DETECTION OF PANDEMIC
CURRENT FOCUS & INVESTMENT

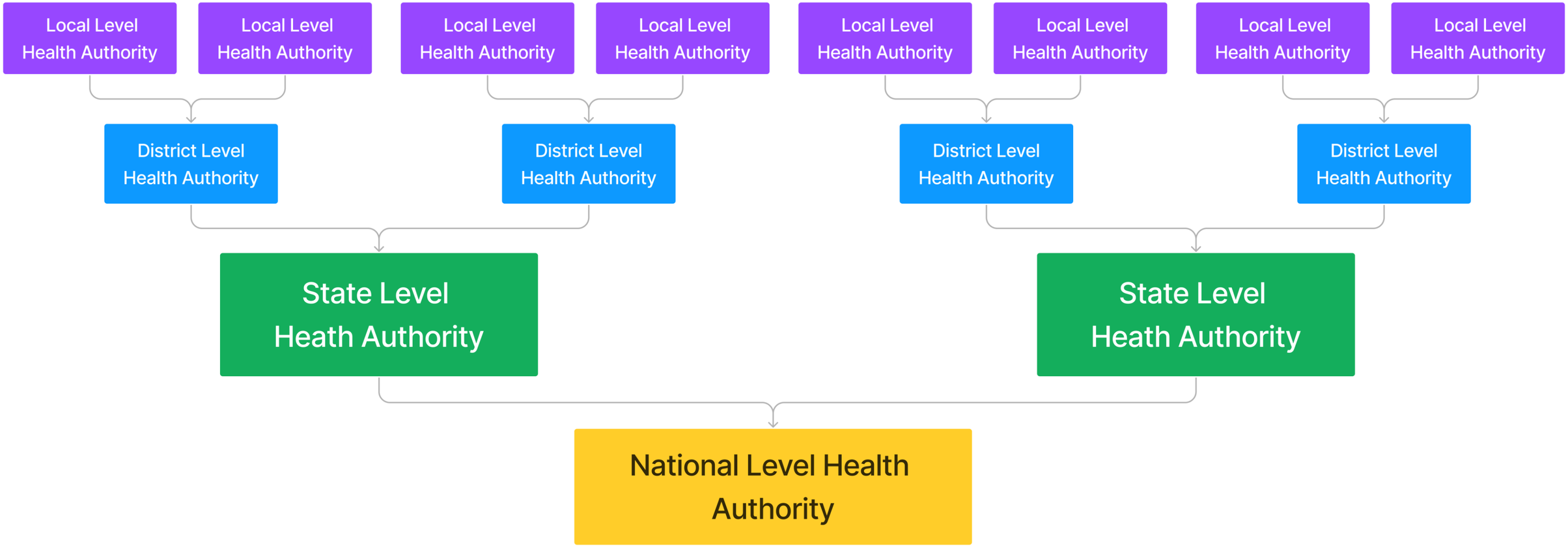
Significance of the project

Prevention

The project's preventive significance lies in its ability to proactively detect and alert health authorities to potential outbreaks, enabling timely interventions and safeguarding communities from the escalation of pandemics.

Significance

The project's significance lies in its advanced AI algorithms for early pandemic detection through school attendance analysis, geospatial visualization pinpointing high-risk areas, and swift communication with health authority for quick response.


Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
Local Level Health Authority
District Level Health Authority
District Level Health Authority
District Level Health Authority
District Level Health Authority
State Level Heath Authority
State Level Heath Authority
National Level Health Authority

Pandemic Real-time Observation and Absentees Clustering Technique [PROACT] Algorithm

The PROACT Algorithm is the powerhouse behind our project, designed for keeping a close eye on potential pandemics within schools.

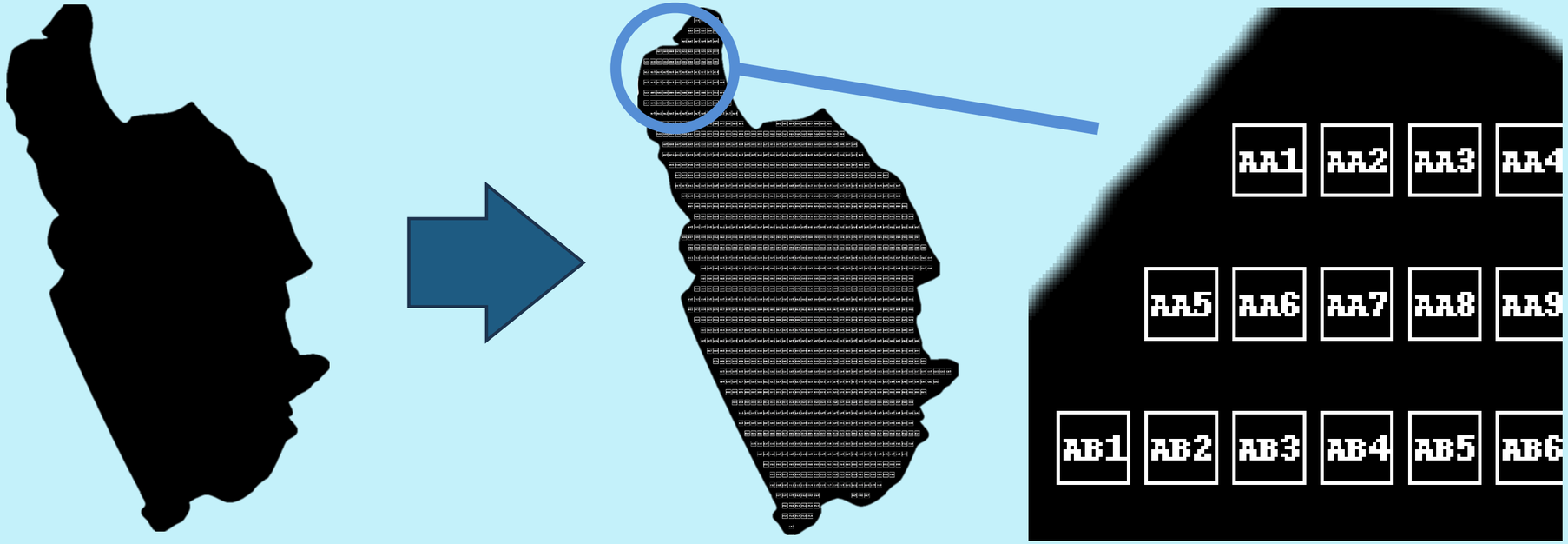
Key features:

1. Realtime monitoring of absentees' data for detecting unusual pattern related to illness among students.
2. Clustering addresses of absent students.
3. Immediately alerting health authorities on detecting any anomalies.
4. Generating easy to understand heatmaps.

Procedure of the Algorithm

STEP 1

- Making a plot of the area with girds dividing the whole area into several small pieces.
- Naming each grid with a unique ID.



STEP 2

- Collect the satellite and terrain view of the area using Google Maps API

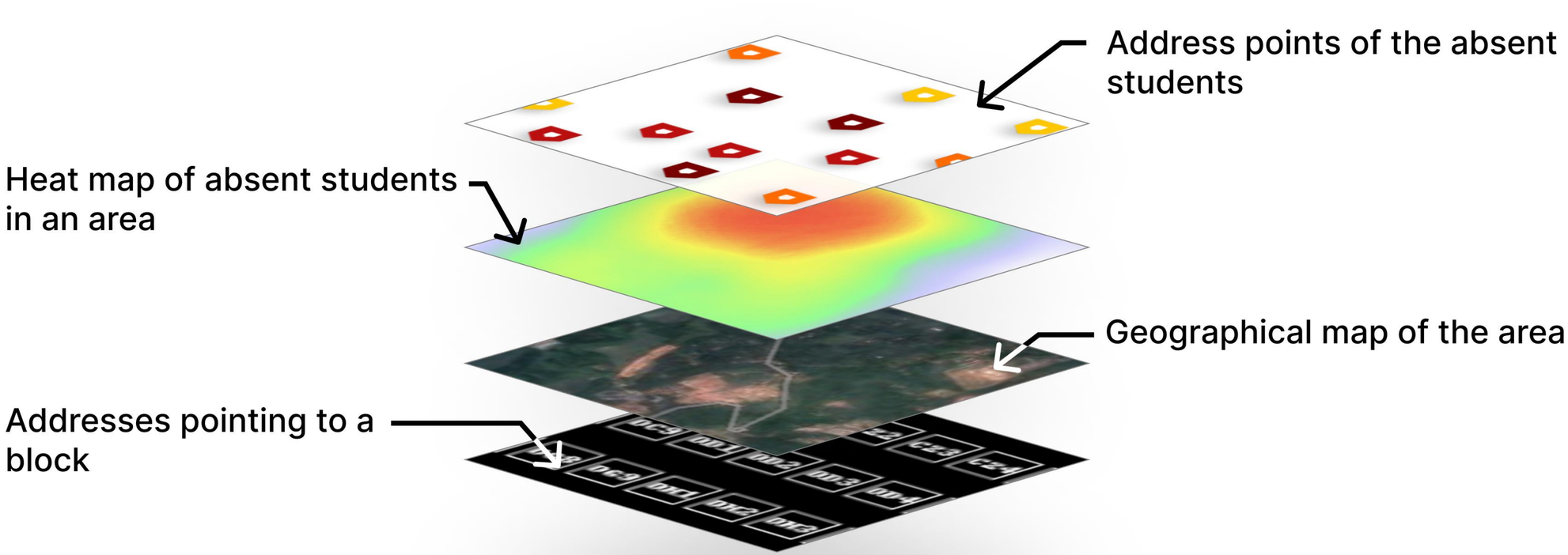
STEP 3

- Arrange all the layers of the map

STEP 4

- Generating heatmap with address points of using real-time data of absent students.

Different Layers of the Intractable Map



Working of the Web Interface

The web interface was built using React.js framework using Firebase as the database for storing the data of absent students.

The web interface provides:

- Intractable heatmaps
- Address highlighting
- Absentees' health records and contact information
- Cluster & hotspot detection

ANALYTICAL PANEL

http://localhost:5500

ANALYTICAL PANEL

STATISTICS

● NORMAL

✓ No unusual increase in student absenteeism. No signs of a potential outbreak.

TOTAL ABSENT CASES

3

HOTSPOTS	CLUSTERS
0	2

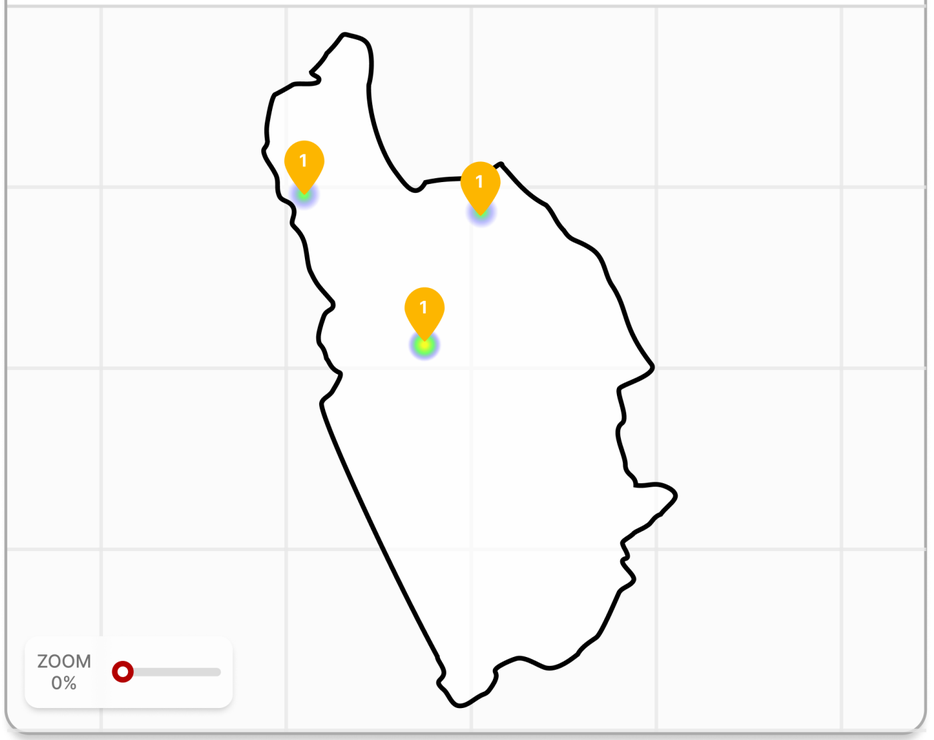
OUTBREAK PROBABILITY

No chance for an outbreak

0%

Intractable Map

RAMANTHALI



ANALYTICAL PANEL

http://localhost:5500

ANALYTICAL PANEL

STATISTICS

- SEVERE

Unusual increase in student absenteeism. Possible outbreak. Immediate investigation necessary.

TOTAL ABSENT CASES

50

- 43 Conjunctivitis cases

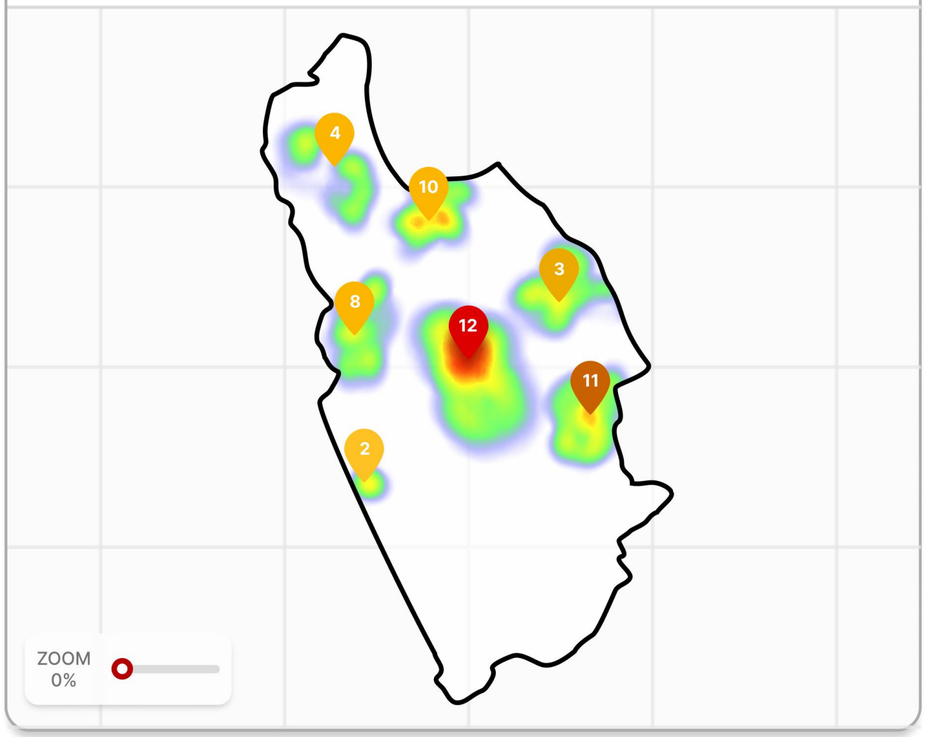
HOTSPOTS	CLUSTERS
2	7

OUTBREAK PROBABILITY

48%

Intractable Map

RAMANTHALI



ANALYTICAL PANEL

http://localhost:5500

ANALYTICAL PANEL

STATISTICS

● SEVERE

⚠ Unusual increase in student absenteeism. Possible outbreak. Immediate investigation necessary.

TOTAL ABSENT CASES

50

● 43 Conjunctivitis cases

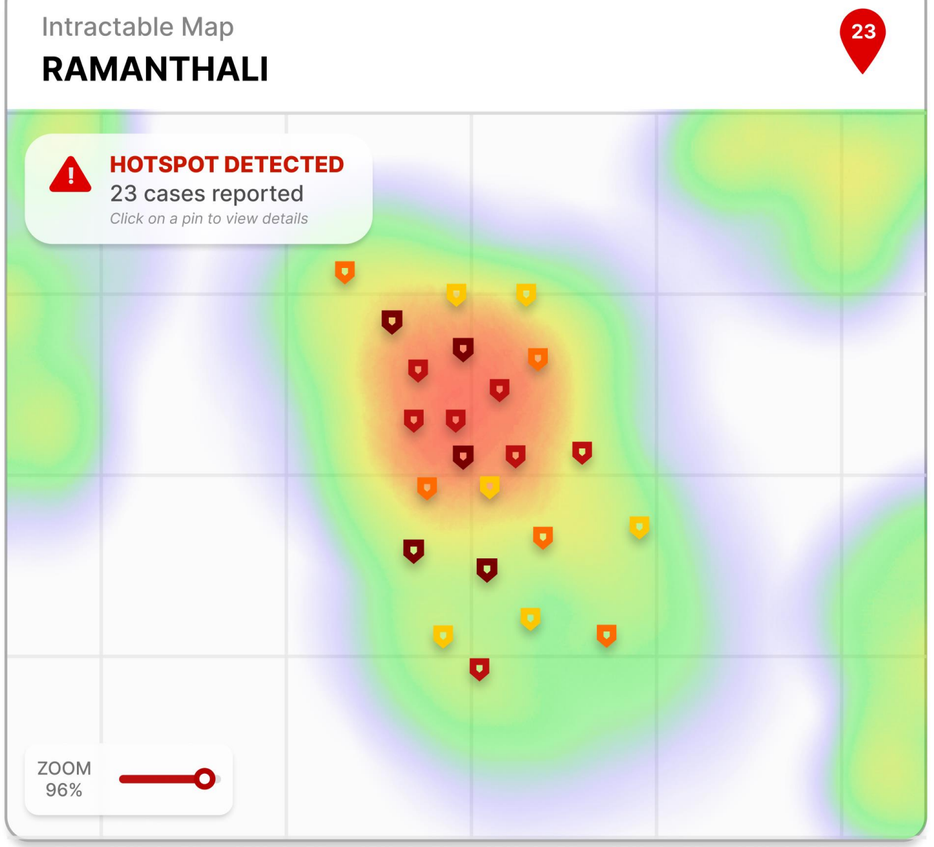
HOTSPOTS	CLUSTERS
2	7

OUTBREAK PROBABILITY

48%

Intractable Map

RAMANTHALI



ANALYTICAL PANEL

http://localhost:5500

ANALYTICAL PANEL

STATISTICS

- SEVERE

Unusual increase in student absenteeism. Possible outbreak. Immediate investigation necessary.

TOTAL ABSENT CASES

50

- 43 Conjunctivitis cases

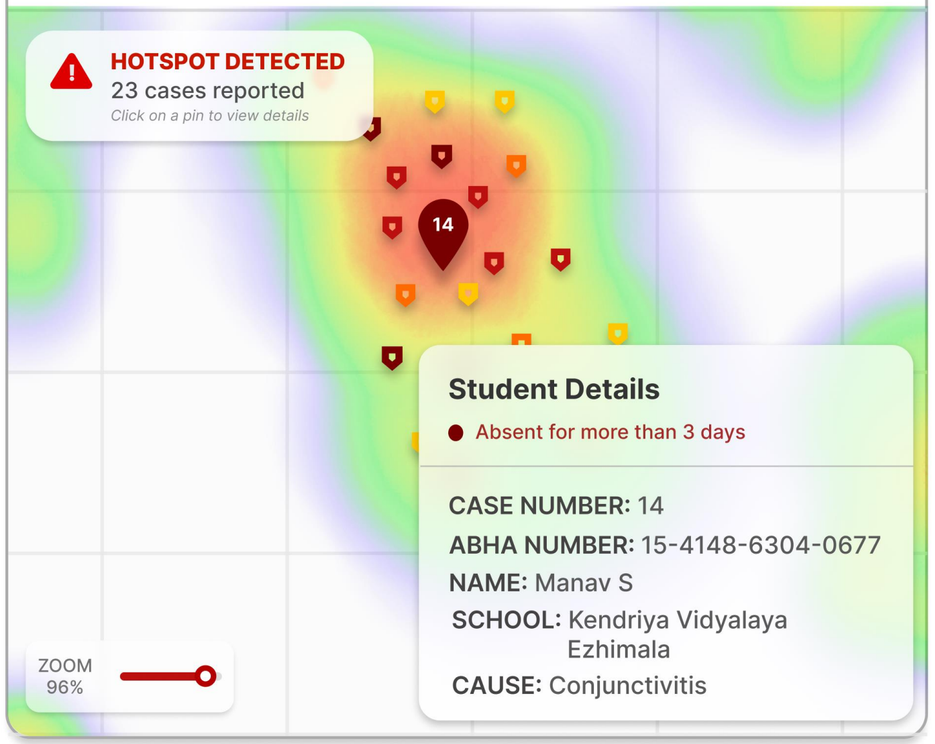
HOTSPOTS	CLUSTERS
2	7

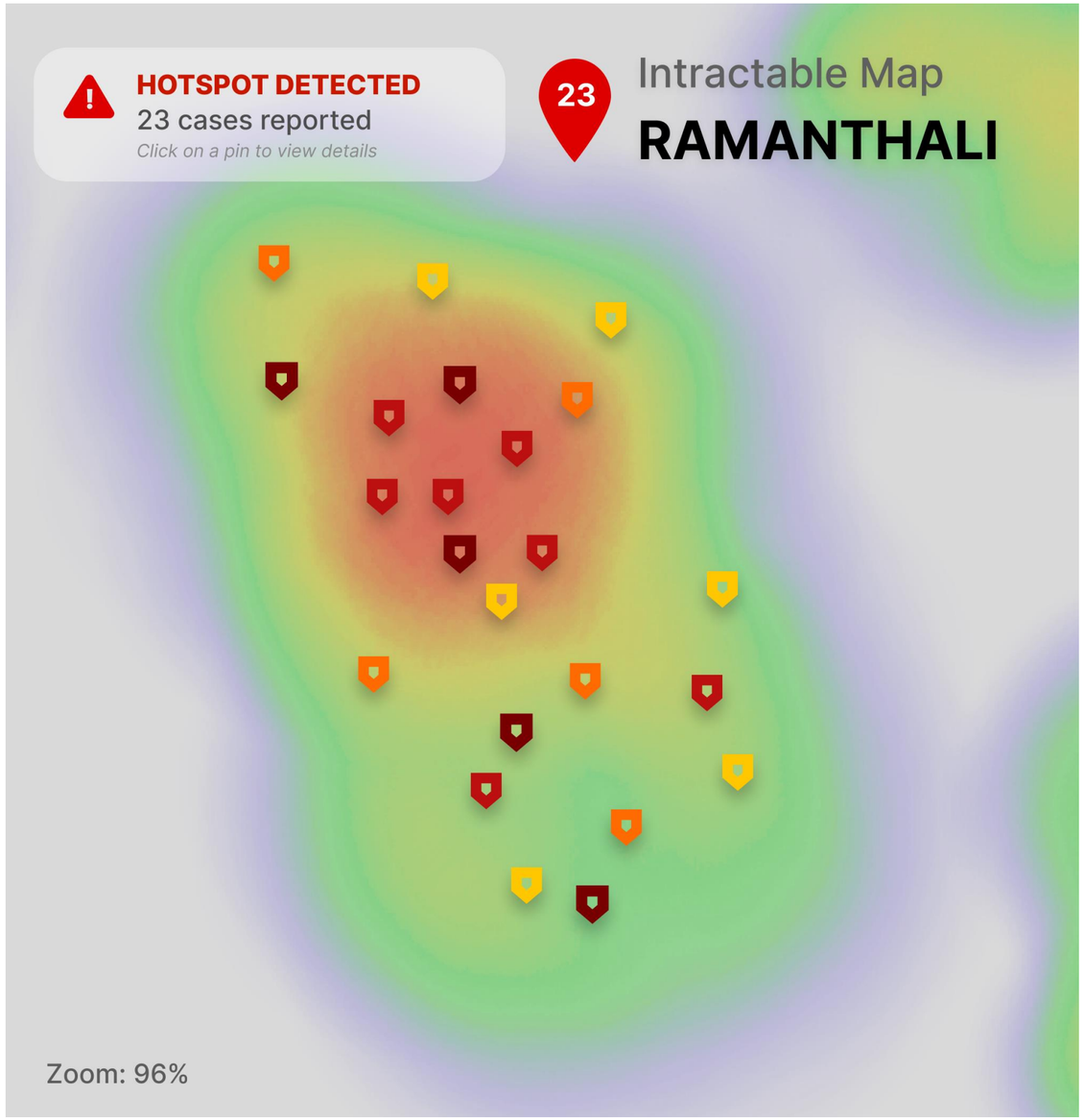
OUTBREAK PROBABILITY

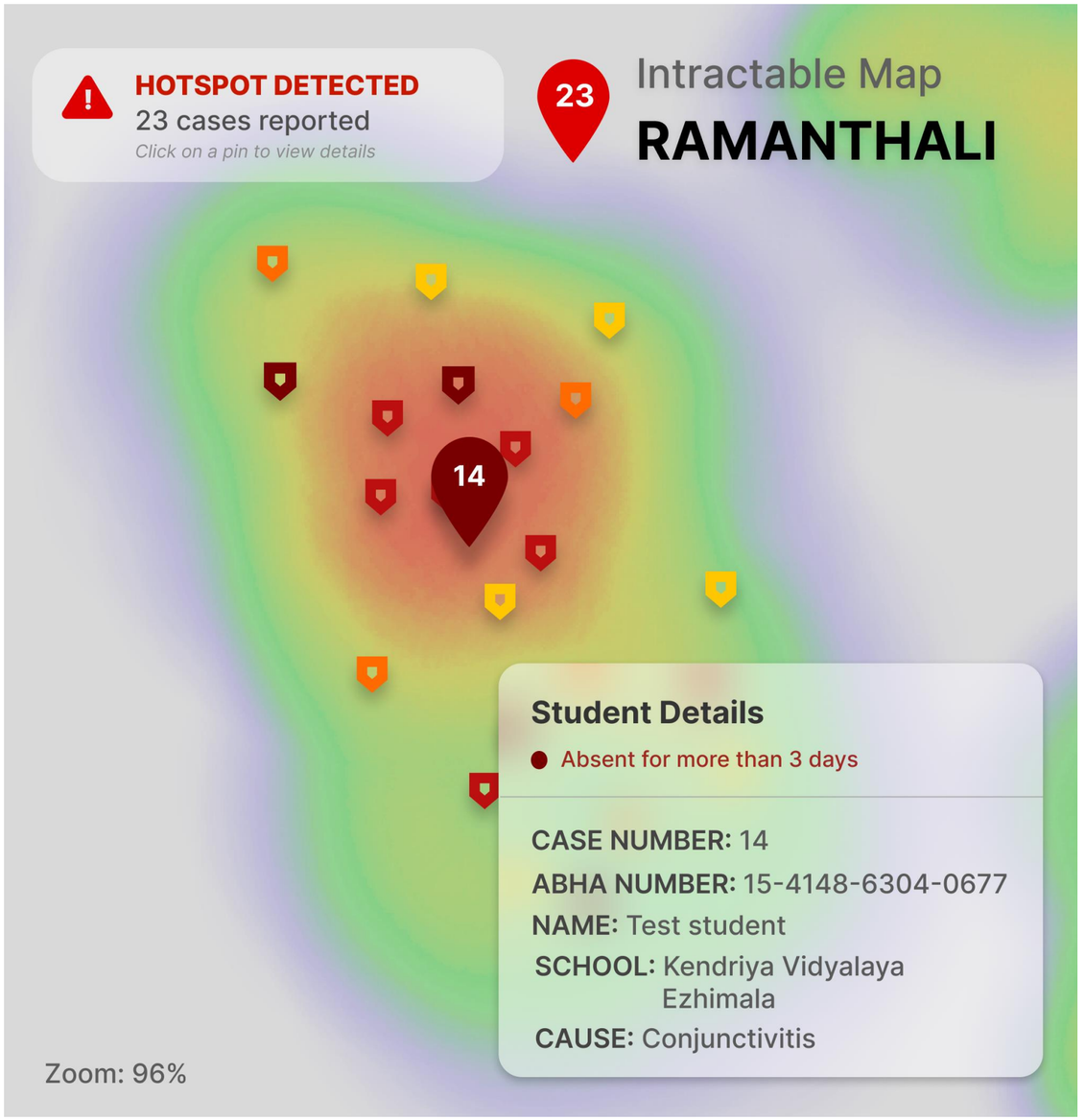
48%

Intractable Map

RAMANTHALI




HOTSPOT DETECTED
23 cases reported
Click on a pin to view details
23
Intractable Map
RAMANTHALI
Zoom: 96%


HOTSPOT DETECTED
23 cases reported
Click on a pin to view details
23
Intractable Map
RAMANTHALI
14
Student Details
Absent for more than 3 days
CASE NUMBER: 14
ABHA NUMBER: 15-4148-6304-0677
NAME: Test student
SCHOOL: Kendriya Vidyalaya Ezhimala
CAUSE: Conjunctivitis
Zoom: 96%

Working of the Mobile App

The Android app is built for the class teachers to feed the absentees' data. The data is then sent to the firebase database.

The app also shows statistics of the absentees and graph of trend in absenteeism of the class, making it easy for the teacher to visualize.

10:00AM

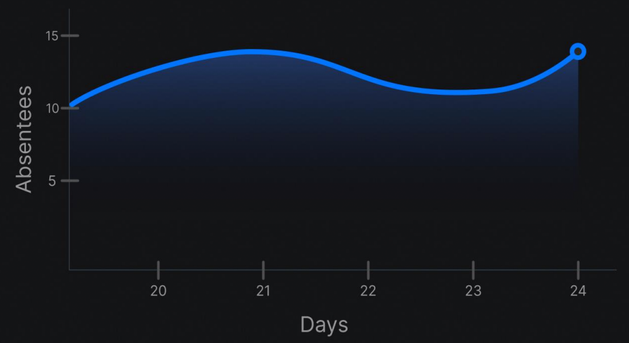
PROACT - ATTENDANCE MARKING

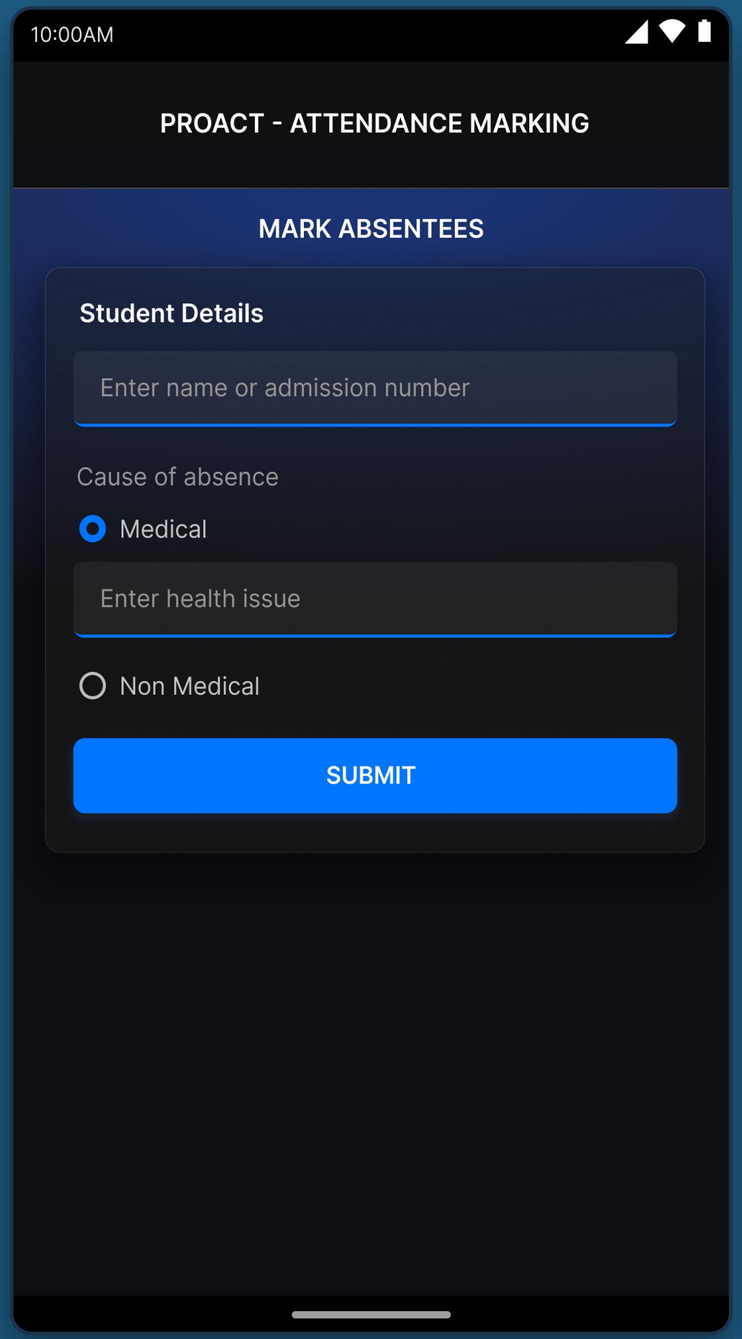
CLASS XI

ABSENTEES

Today	Last day	This week
Mark Absentees	13	16

TREND




10:00AM
PROACT - ATTENDANCE MARKING
MARK ABSENTEES
Student Details
Enter name or admission number
Cause of absence
Medical
Enter health issue
Non Medical
SUBMIT

AI model

The AI model is the core of our project, which is responsible for predicting the change of an outbreak to plotting the heatmap and identifying the clusters.

The AI model was built using the `RandomForest` model for outbreak detection.

Susceptible-Infectious-Susceptible model was also used to find the chance of a student to get affected with the viral disease.

The AI model was trained on a large dataset which also consists trend of absentees during the dengue situation of 2018 in India.

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
```

jupyter prediction_model Last Checkpoint: 8 days ago

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python 3 (ipykernel)

[2]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
```

[4]:
```
# Assuming your dataset is in a CSV file named 'absenteeism_data.csv'
data = pd.read_csv('absentees_dataset.csv')

# Check the first few rows to understand the structure of your data
data.head()
```

[4]:

	Day	Cause
0	1	Health Related Illness
1	1	Health Related Illness
2	1	Health Related Illness
3	1	Health Related Illness
4	1	Health Related Illness

[5]:
```
# Convert categorical data to numerical values
data['Cause'] = data['Cause'].astype('category')
data['Cause'] = data['Cause'].cat.codes

# Splitting the data into features and target variable
X = data.drop('Cause', axis=1)  # Features
y = data['Cause']  # Target variable

# Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```python
def calculate_outbreak_probability(absentee_data):
    # Grouping data by cause and counting occurrences per day
    grouped_data = absentee_data.groupby(['Day', 'Cause']).size().unstack(fill_value=0)
    cause_days = grouped_data.astype(bool).sum(axis=0)
    potential_outbreaks = cause_days[cause_days > 3]
    total_days = absentee_data['Day'].nunique()
    outbreak_probability = len(potential_outbreaks) / total_days * 100

    return outbreak_probability
absentee_data = pd.read_csv('test_dataset.csv')

outbreak_probability = calculate_outbreak_probability(absentee_data)
print(f"Probability of outbreak: {outbreak_probability:.2f}%")
```

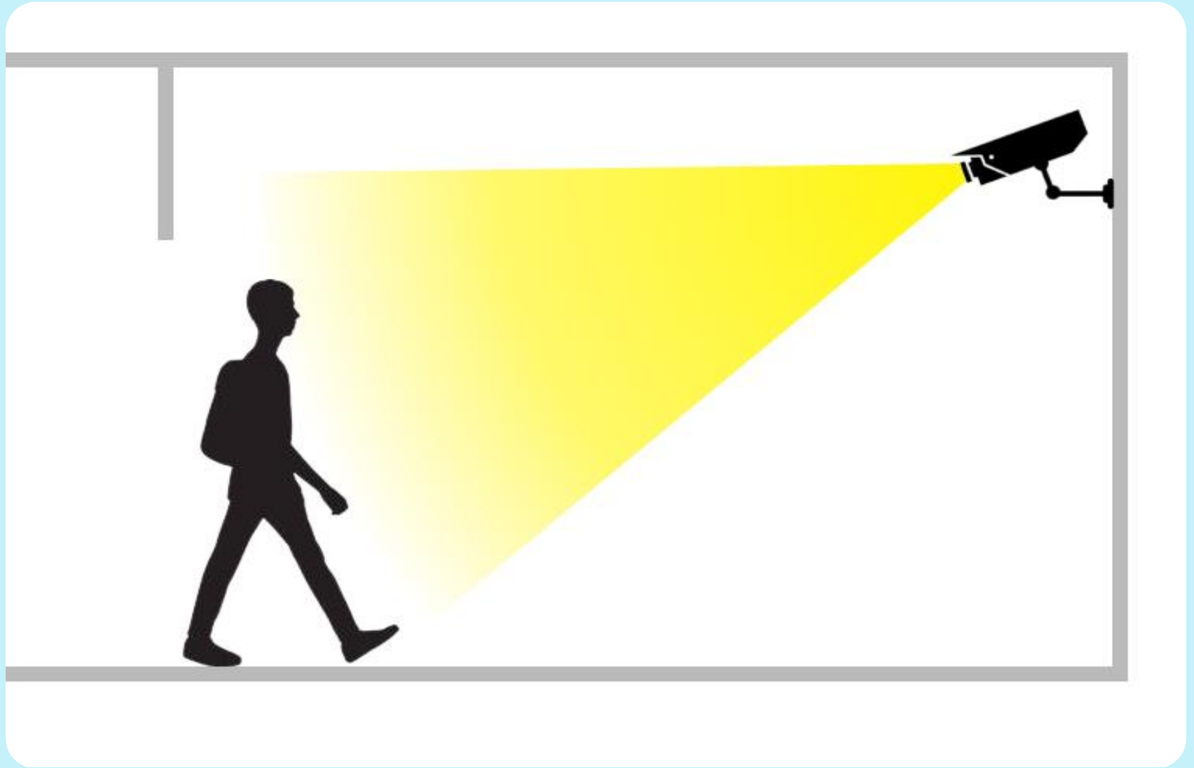
```
Probability of outbreak: 10.00%
```

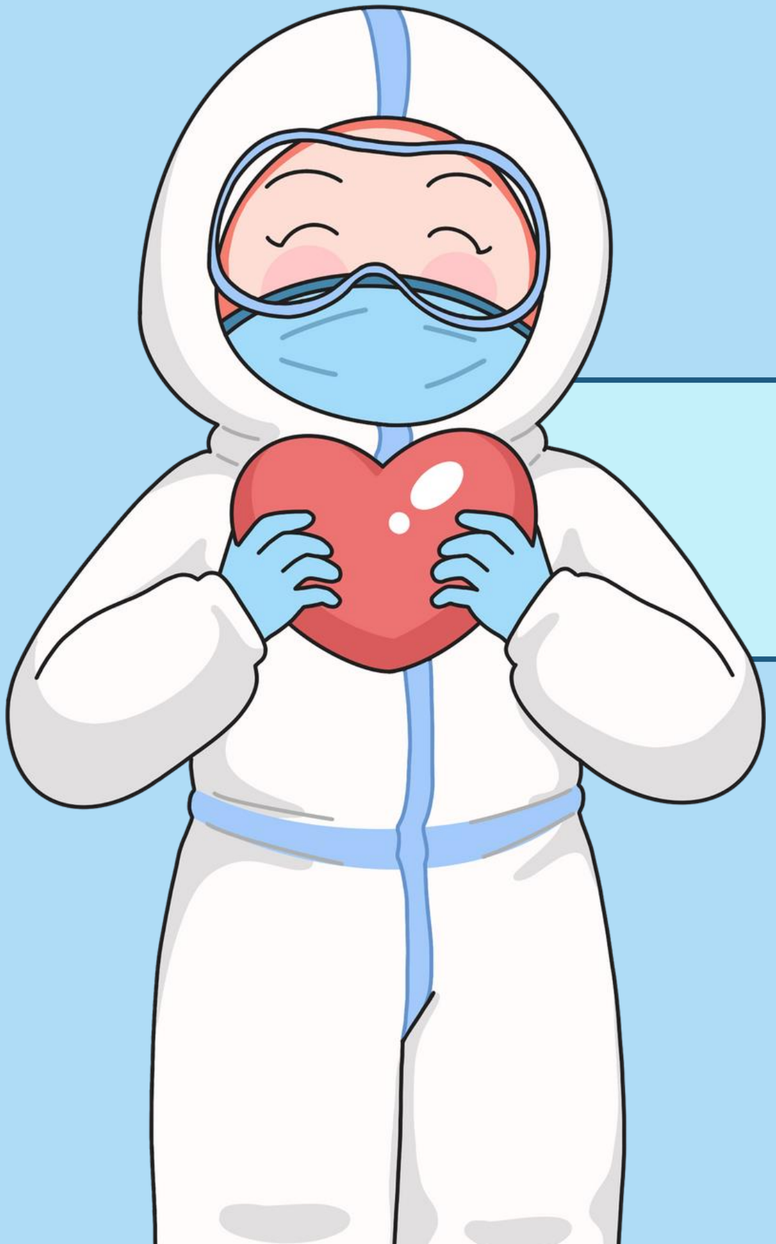
[31]:
```
predictions = model.predict(X_test)

accuracy = round(accuracy_score(y_test, predictions), 2)
report = classification_report(y_test, predictions)

print(f'Accuracy: {accuracy} ({accuracy*100}%)')
print(f'Classification Report:\n{report}')
```

```
Accuracy: 0.92 (92.0%)
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.84      0.91        19
           1       1.00      1.00      1.00         1
           2       0.86      1.00      0.93        19

    accuracy                           0.92        39
   macro avg       0.95      0.95      0.95        39
weighted avg       0.93      0.92      0.92        39
```

Future Scopes

Using AI cameras to automatically mark the attendance of students using facial recognition technology.





THANK YOU!