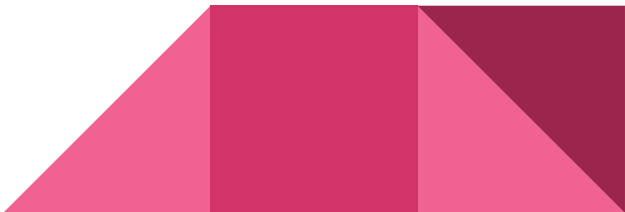# Crop data analytics using image and non-image features

Team members   :   Asish Varanasi
Dheeraj Raghavendra
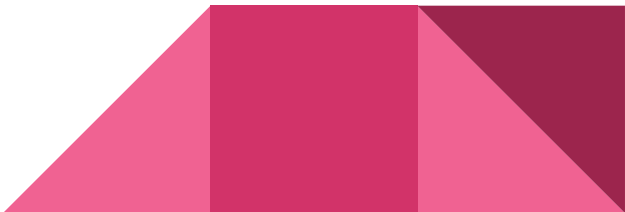Purnachand Jaddu

Mentors   :   Dr. Avinash Sharma
Dr. P. Krishna Reddy

Motivation

- Agriculture plays a vital role in economy for developing countries like India.

- Rapid technological advancement - but not much focus on agricultural domain.

- Several newly emerging problems in agriculture.

- In 2012, the NCRB of India reported 13,754 farmer suicides.

- Hence, a need for guidance to the farmers.

- Esagu is one such platform.

Key pain points

Recommendation of Diagnostics:

- Accuracy

- Timely advice

- Minimum required input

- Ease of communication

- Cost effectiveness

- Feedback system

Insights mining

Periodic behaviour of diseases with respect to :

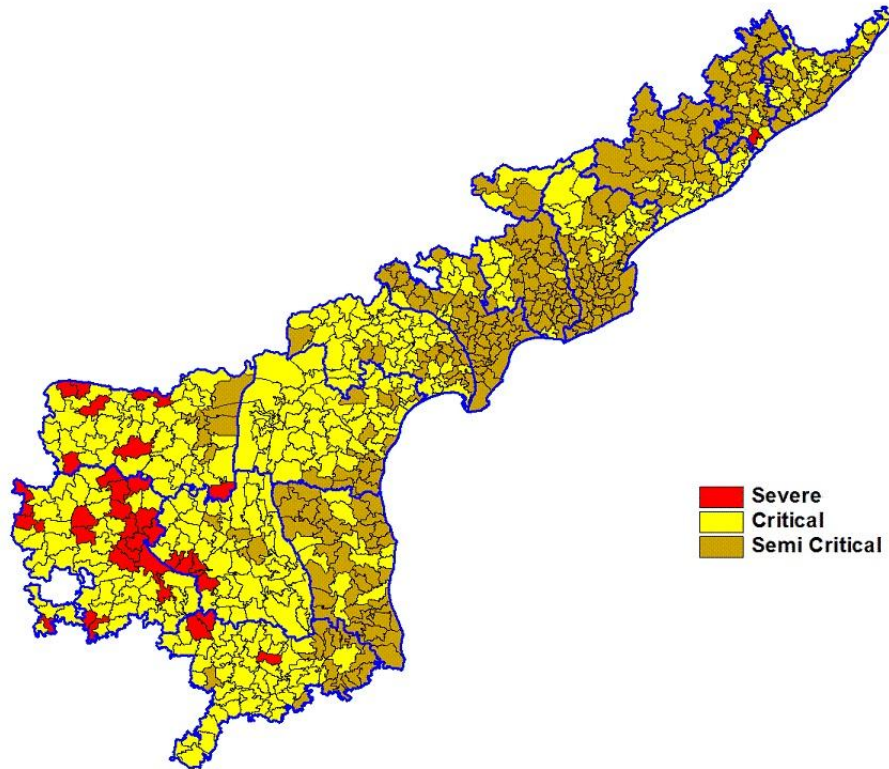- Season
- Crop
- Pesticides/Fertilizers
- Location
- Pests

Visualization

- Heatmaps showing the variation in the intensity of factors w.r.t:
  - ➤ Time
  - ➤ Location

- Plots like Bar charts, Pie diagrams etc.

- A dashboard for interactive visualization of the data.

Drought map

Severe
Critical
Semi Critical

Data

- Input data primarily contains non-image features.

- Weather details, soil conditions and farm report constitute the major part.

- Pre-processing involving data cleaning, integration and normalization.

- Finally, analyzation after data transformation.

Representation

PGM (Probabilistic Graphical Model)

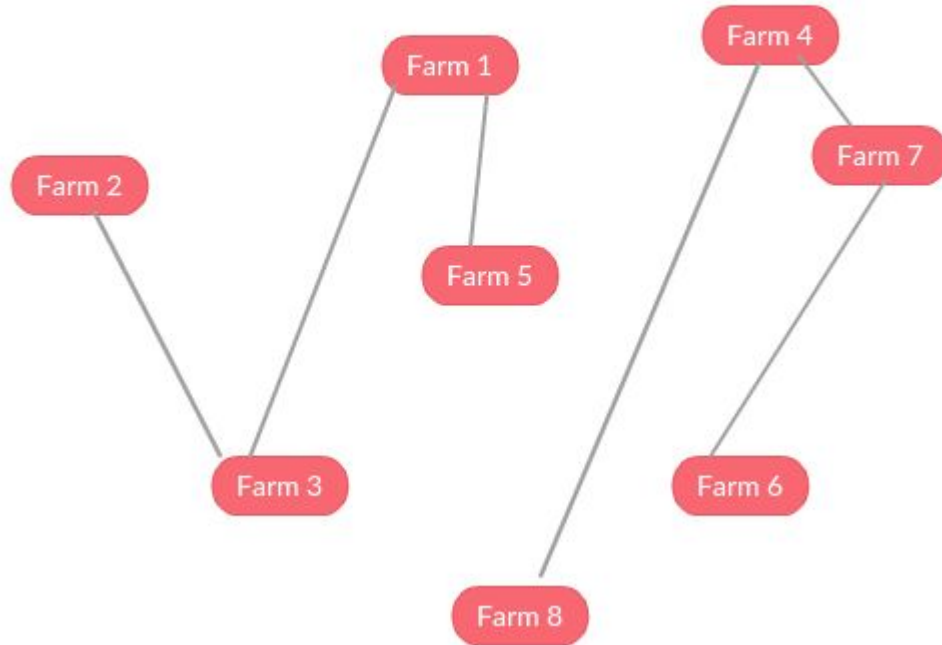- Farms as vertices and edges indicating the dependency on various factors.

Graph signalling

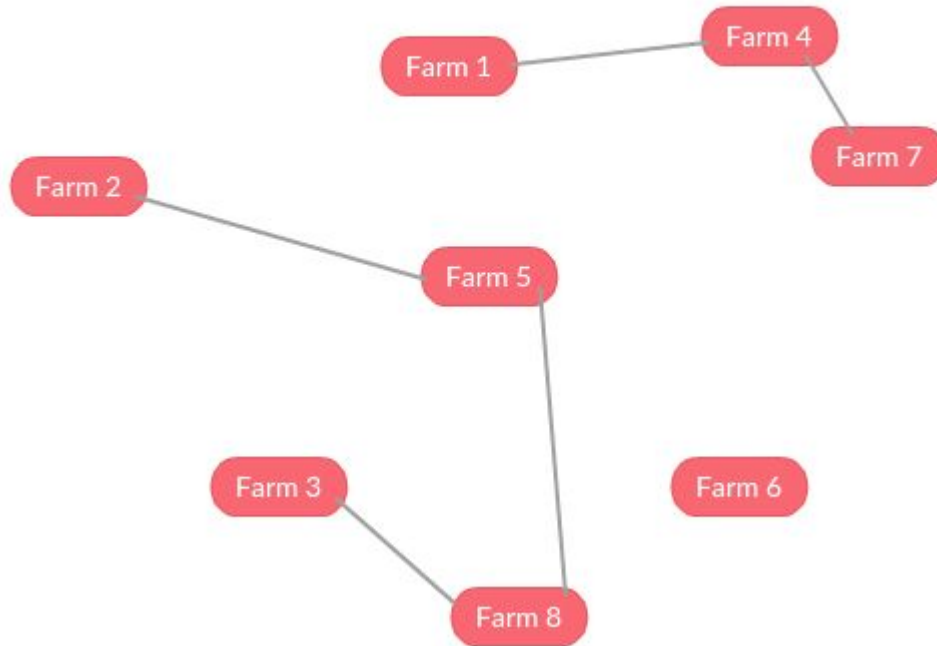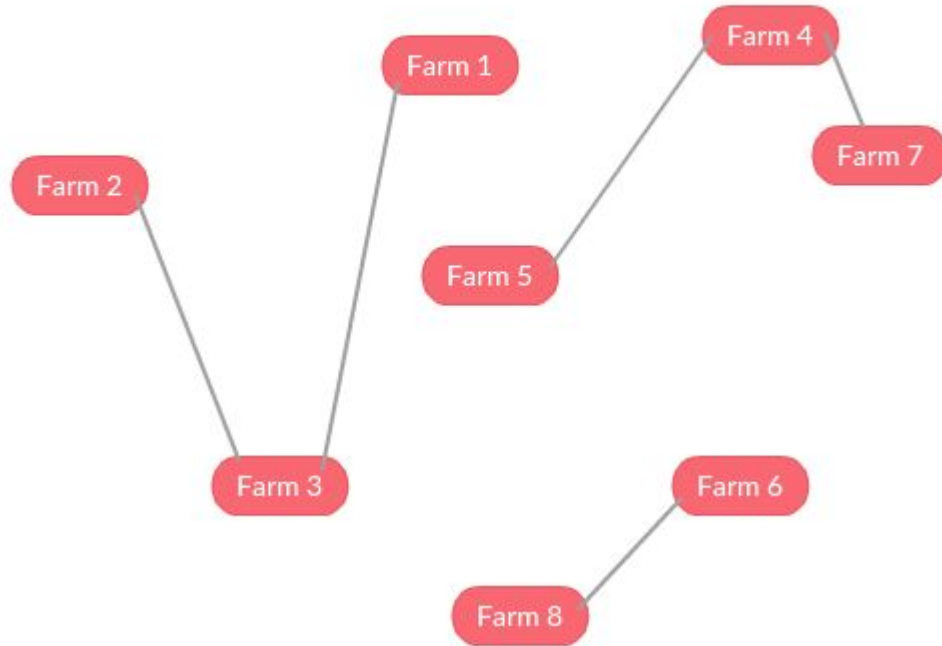- A plot of the condition of each farm with respect to time/season.
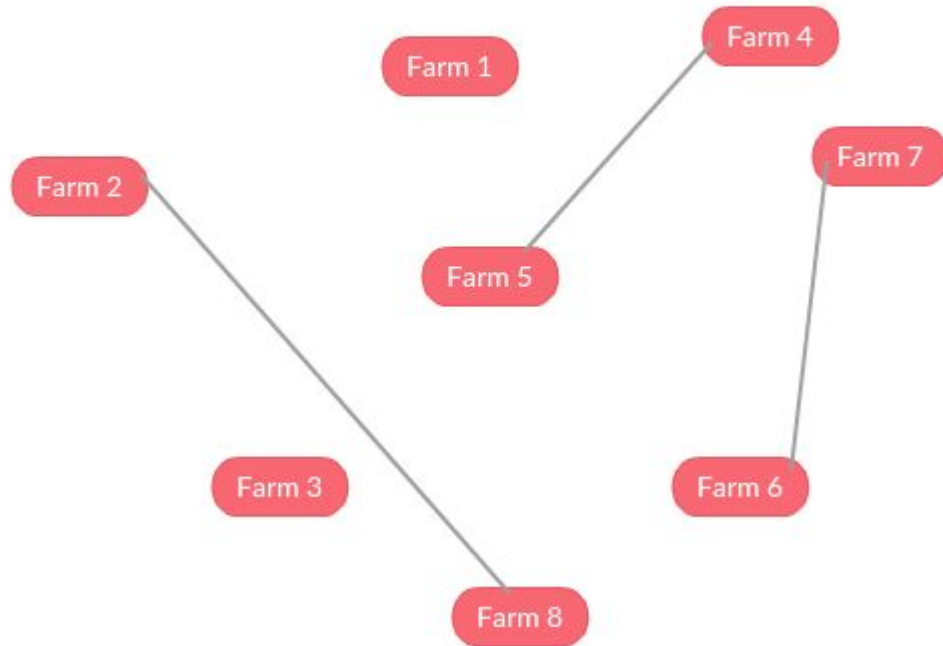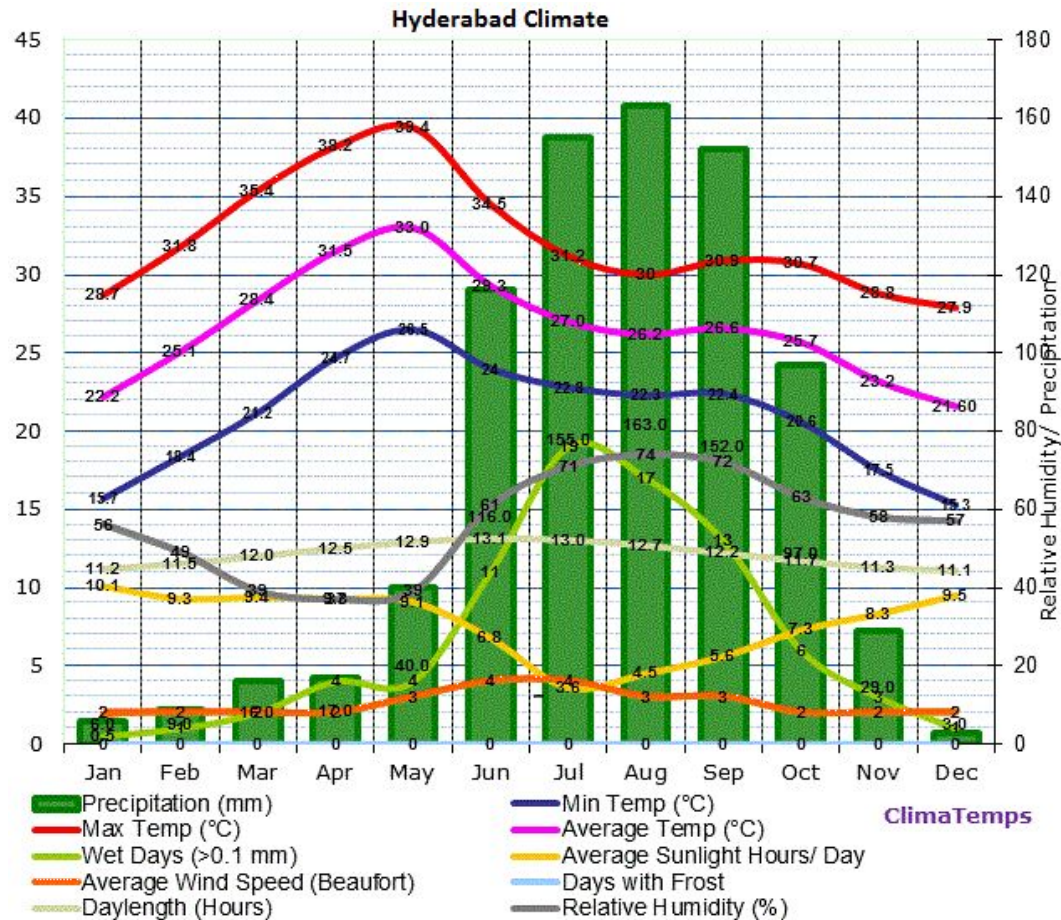
Graph based on neighbourhood

Graph based on temperature

Graph based on rainfall

Graph based on crop species

Hyderabad Climate

Models

Graphical method

- Taking into account all the dependencies in the graphs, we model the system.

Spectral analysis

- Considering timely behaviour of the respective feature.
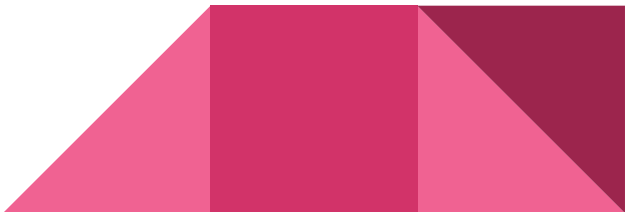
Topic modeling

- Discovering the abstract symptoms using keywords.

Predicting the solution using Multi View Learning techniques.

- Specific MVL solution will depend on the data.

- If the dimension is large, then we have to use Canonical Correlation Analysis.

- If the entire data is labeled, we can use multiview supervised learning.

- If we have both labeled and unlabeled data, then we have to use multiview semi supervised learning.

- Several such procedures like active learning, ensemble learning etc.

- Finally, a feedback to improve the accuracy of the prediction algorithm.

Expected deliverables

- Data visualization .

- Data aggregation/pre-processing.

- Model implementation.

- Developing dashboard for interactive visualization of the data and prescriptive

  analytics service.

- Recommendation system.

Timeline

- One month from now-union of data and representative model.

- Two months from now-POC of representation.

- By the end of this semester model implementation.

- Next semester recommender system.

References

- http://insait.in/AIPA2012/articles/009.pdf

- http://www.saravananraj.net/wp-content/uploads/2014/12/45_AFITA_ICT-for-Agricultural-extension_India.pdf

- http://insait.in/AIPA2012/articles/005.pdf

Thank you