

1. The project before me was “Plant Disease Classification and Explanation”
 - a. **Problem:** is to identify the plant and the disease, if any and provide explanation for the prediction.
 - b. **User:** Government (for large scale plant monitoring) and individuals (farmers/cultivators)
 - c. **AI Methods:** Classifier (CNN, pretrained Inception V3) and LIME for generating explanations.
 - d. **Data:** PlantVillage (54303 leaf images divided into 38 categories by species and disease) and PlantDoc (2598 leaf images divided into 13 plant species and upto 17 disease classes). I am not sure of what preprocessing or transformation techniques (as it is images) have been used.
 - e. **Evaluation:** Recall, Precision, F1-Score, Accuracy, BCE Loss
 - f. **Trust/Human Values:** The fairness issues that would be relevant for the problem are: type of model (as CNN is a black box model), and group discrimination. There might be data imbalance too for the disease classes but the presentation does not talk about this.
 - g. **Human-AI:** The trust of the human in this approach is garnered by using LIME for providing explanations regarding the model predictions.
 - h. **Project’s Impact:** The idea is impactful as it would help individuals (farmers and also horticulturists or those who grow plants as a hobby) to monitor their crops/plants. However, there are many startups who have already deployed working apps that have features in addition to the one’s mentioned in this project.
2. The project after me was “Measuring the Spatio-Temporal Psychological Impact of Government Policies during Covid-19 Pandemic in different US states using Twitter Data”.
 - a. **Problem:** Identifying the trust compliance of the public towards the rules implemented during Covid19
 - b. **User:** Policy Makers and Public Health Specialists
 - c. **AI Methods:** Naive Bayes, Random Forest, Balanced Random Forest, SubSample Balanced Random Forest, Semantic Encoding and Decoding Optimization.
 - d. **Data:** The dataset used comprises the Tweet IDs collected during Covid19. The tweet IDs hydrated using a hydrator. In addition to this, the government policies data, diagnostic and statistical manual of mental disorders, ontology for drug abuse and geonames and the subreddit for depression, addiction and anxiety.
 - e. **Evaluation:** Precision, Recall and F1-score
 - f. **Trust/Human Values:** political and social values affecting the compliance of policy (trust issue of compliance)
 - g. **Human-AI:** There's no solution built, by the work can be used to gain insight for the policy makers which can be further used to reformat the tool to get policies and hence obtain more human centric policies
 - h. **Project’s Impact:** The project is useful for policy makers to better assess the sentiment of the people in a region.

3. The fairness dimensions are as follows:
 - a. **Data Modality:** The fairness issues that arise from the different types of data (structured/unstructured) fall under this category. For example, a dataset consisting of information regarding credit card usage might be having more entries for males compared to females.
 - b. **Type of Model:** There are two types of models - black box and white box and three types based on the type of training - supervised, unsupervised and semi-supervised. Different models might have different fairness issues, for example, IBM has found the model for facial detection to be unfair to black people.
 - c. **Type of Fairness:** Group and Individual fairness refers to the model having the same predictions for a similar group or an individual.
 - d. **Cause:** The fairness issues can be caused by two factors - "How it originated" and "Where it originated".
 - e. **Testing Verification:** Direct, Indirect and Hidden biases contribute to the testing verification. Direct bias occurs if we use protected attributes to create the models. And in indirect bias, a correlating feature contributes to the bias and in hidden bias, we are unsure of which feature is causing the fairness issue in the model.
 - f. **Analysis:** Data scientists try to maximise the model performance and select features accordingly which might lead to fairness issues in the model.
 - g. **Mitigation:** There are different approaches to mitigate the bias in a model and improve upon fairness, namely, in-processing, pre-processing, and post-processing.
4. The type of fairness that I would be concerned about is the type of fairness - group discrimination or individual discrimination as the ordering needs to be fair to both the individuals and the group as a whole.