

CMPE 256 - Advanced Data Mining Hackathon - Food Security

Team Data Farmers

Devi Priya Ravi
Dhanasree Rajamani
Sravani Thota

Colab URL :

https://colab.research.google.com/drive/1ffnhsIzDCCQPhM7O03__7NUjEKIvMFU6?usp=sharing

As a team, we have performed the following steps

- Business Understanding
- Data Understanding
- Data Collection
- Data Integration
- Data Selection
- Data Preprocessing
- Data Cleaning
- Data Imputation
- Feature Engineering
- Correlation Matrix
- Scaling and Normalization
- Model
- EDA
- Trend Analysis

- Statistical analysis of yield
- Conclusion - Results and Findings

Results and Findings

In light of the above performed data analysis, we conclude that there could be some miscellaneous data that could contribute more for the yield factor of the potatoes. In addition to that, understanding more about the pH and seed technology can give different dimensions to the goal that we plan to achieve here. Inverse correlations between the features (for eg Yield and temperature) strongly suggest the presence of certain implicit factors that can contribute to the growth of the potatoes. Taking the time span (1961-2022) under consideration, we also reflect the changes in technology and events (2020 pandemic that witnessed production/trade shutdown) which can highly influence the yield factor in our problem statement.