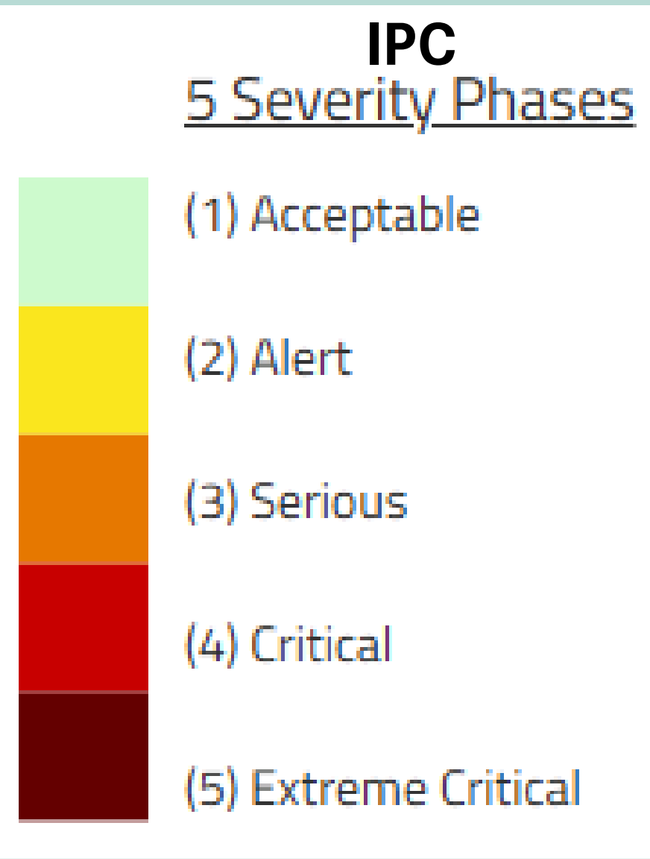
Predicting Food Insecurity In Ghana Using Machine learning

Integrated Food Security Phase Classification **(IPC) data modelling**

David Mathea Mithiga
Adm: 034915
Principles of Data Science

Predict **IPC Class** given [Region, Population]

86% Accuracy



The analysis of food insecurity in Ghana was conducted using the IPC dataset, with a focus on 'current' conditions and crisis levels corresponding to IPC phases 3, 4, and 5. After rigorous preprocessing and cleaning of the data, multiple machine learning models—Logistic Regression, Decision Tree, Random Forest, SVM, and XGBoost—were trained and evaluated for their predictive capabilities. The XGBoost model demonstrated superior performance, achieving an accuracy of 0.45, precision of 0.46, recall of 0.45, and F1-score of 0.45.

To further enhance prediction accuracy, IPC phases were reclassified into two distinct categories: "acceptable" (IPC 1-2) and "requiring urgent attention" (IPC 3-5). This strategic reclassification significantly improved model performance by minimizing penalties for classification errors between closely related classes. As a result, the overall accuracy increased to 0.86, underscoring the effectiveness of this approach.

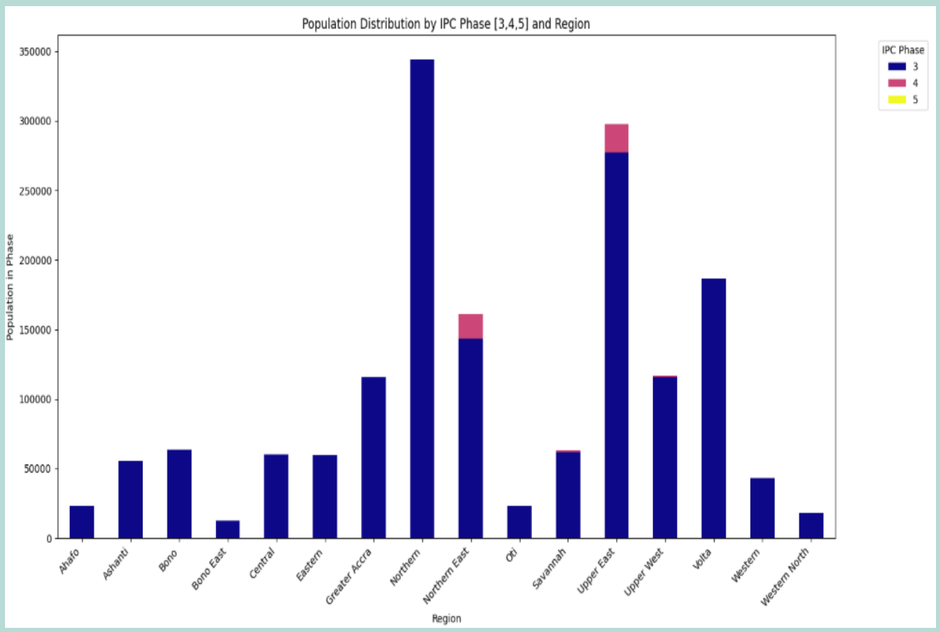
These findings provide robust and reliable insights, enabling targeted interventions in regions experiencing severe food insecurity. The refined model offers a powerful tool for policymakers and humanitarian organizations to address food insecurity with greater precision and efficiency.

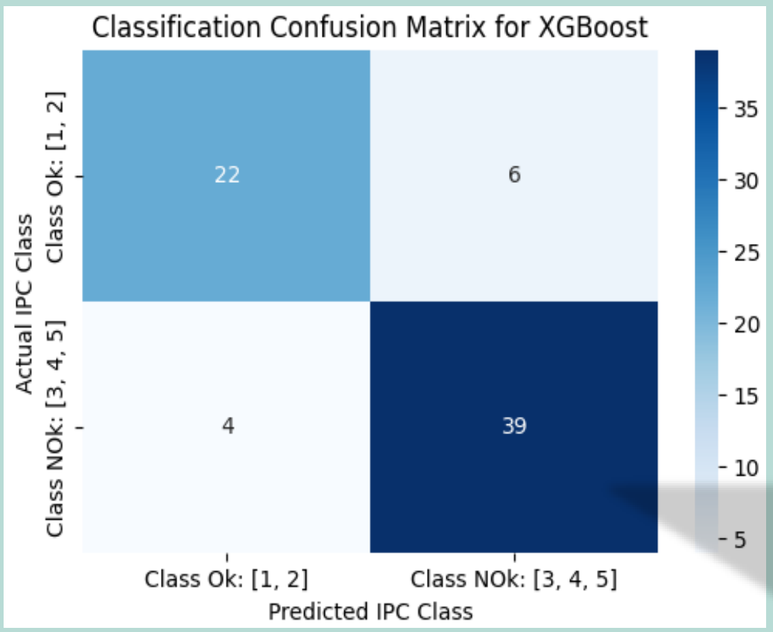
The analysis evaluated five machine learning models—Logistic Regression, Decision Tree, Random Forest, SVM, and XGBoost—to predict food insecurity in Ghana using the IPC dataset.

The results show that XGBoost achieved the highest performance with an accuracy of 86%, precision of 86%, recall of 86%, and F1-score of 86%. This indicates that XGBoost is the most effective model for predicting food insecurity phases in this context.

Other models also performed well, with Decision Tree achieving an accuracy of 85% and Random Forest reaching 80%. These findings highlight the potential of machine learning techniques to support targeted interventions in regions facing severe food insecurity.

Model	Precision	Recall	F1-Score	Accuracy
Logistic Regression	0.81	0.73	0.69	0.73
Decision Tree	0.84	0.85	0.84	0.85
Random Forest	0.80	0.80	0.80	0.80
SVM	0.80	0.69	0.62	0.69
XGBoost	0.86	0.86	0.86	0.86





The analysis of food insecurity in Ghana used the IPC dataset, focusing on crisis levels (phases 3-5).

After preprocessing, multiple models were evaluated, with XGBoost achieving an initial accuracy of 0.45.

Reclassifying IPC phases into "acceptable" (1-2) and "requiring urgent attention" (3-5) boosted accuracy to 0.86.

This model empowers policymakers and humanitarian organizations to target interventions effectively by predicting food insecurity with 86% accuracy on critical IPC phases 3,4 and 5.