

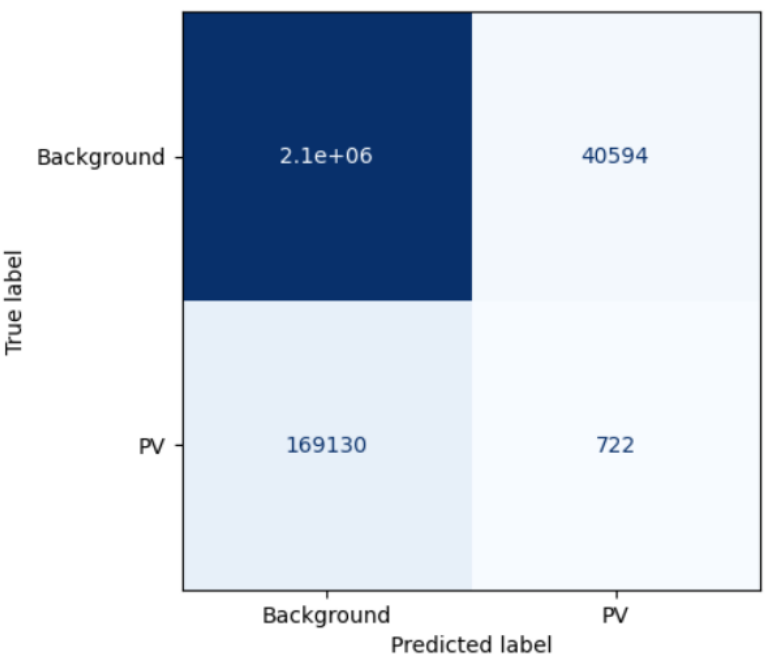
# Compare the Baseline Models' performances:

## Random Model:

TP:722 TN:2148850 FP:40594 FN:169130

Accuracy: 0.9111  $(TP + TN) / (TP + TN + FP + FN)$   
Recall: 0.0043  $TP / (TP + FN)$   
Specificity: 0.9815  $TN / (TN + FP)$   
Precision: 0.0175  $TP / (TP + FP)$   
F1-Score: 0.0068  $(2 * precision * recall) / (precision + recall)$   
IoU: 0.0034  $TP / (TP + FN + FP)$

Confusion Matrix

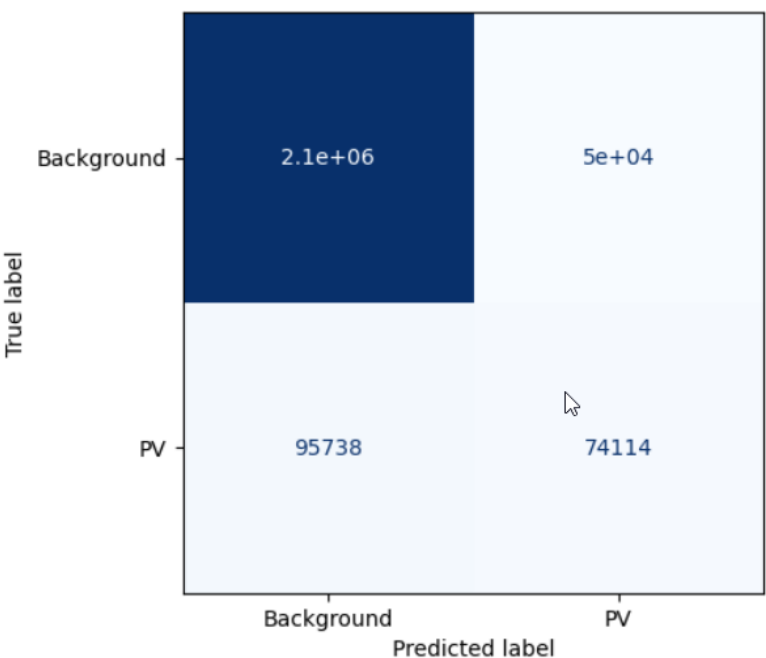


## Threshold Classifier:

TP:74114 TN:2139055 FP:50389 FN:95738

Accuracy: 0.9381  $(TP + TN) / (TP + TN + FP + FN)$   
Recall: 0.4363  $TP / (TP + FN)$   
Specificity: 0.9770  $TN / (TN + FP)$   
Precision: 0.5953  $TP / (TP + FP)$   
F1-Score: 0.5036  $(2 * precision * recall) / (precision + recall)$   
IoU: 0.3365  $TP / (TP + FN + FP)$

Confusion Matrix



## Logistic Regression Model:

TP:104260 TN:2088443 FP:101001 FN:65592

Accuracy: 0.9294  $(TP + TN) / (TP + TN + FP + FN)$   
Recall: 0.6138  $TP / (TP + FN)$   
Specificity: 0.9539  $TN / (TN + FP)$   
Precision: 0.5079  $TP / (TP + FP)$   
F1-Score: 0.5559  $(2 * precision * recall) / (precision + recall)$   
IoU: 0.3849  $TP / (TP + FN + FP)$

Confusion Matrix

