Global Health Project

Core Clinical Fields

- **Age**
- **Gender**
- **Onset Time**
- **Medication Status**
- **Seizure Frequency**
- **Seizure Type**
- **Physical Examination**

Risk Factor Features

- risk_factors.febrile_seizures
- risk_factors.childhood_meningitis
- risk_factors.trauma_to_head
- risk_factors.prenatal_complication
- risk_factors.family_history_epilepsy
- risk_factors.neonatal_complication
- risk_factors.cerebral_malaria
- risk_factors.behavioral_problem
- risk_factors.psychiatric_problem
- risk_factors.cerebrovascular_disease
- risk_factors.developmental_delay

Feature-set we are using

Common

- **Age**
- **Gender**
- **Onset Time**
- **Seizure Frequency**
- **Seizure Type**

Data-driven (chatgpt suggested)

- **Female Adolescent**
- **Hepatic Impairment Flag**
- **Renal Impairment Flag**
- **Status Or Prolonged Flag**
- **Cognitive Priority**

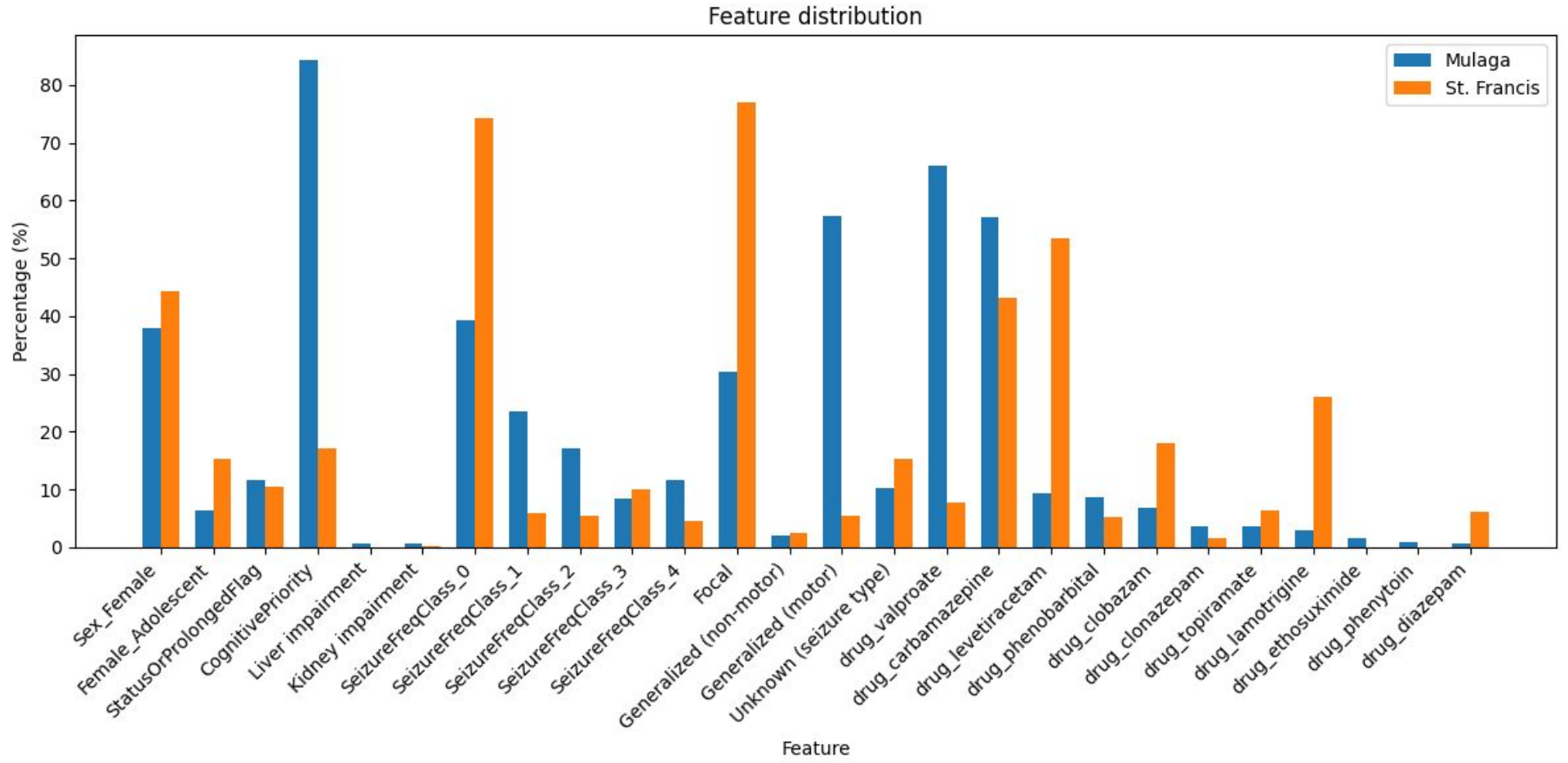
Feature	Importance / Citation	LLM Prompt / Questions / Context
Age **(Raj suggested)**	**Importance:** Fundamental demographic variable for dosing	**LLM Prompt:** Identify the patient's current age at the time of this note. Look for explicit mentions or calculate it from the Date of Birth relative to the Visit Date.
Gender **(Raj suggested)**	**Importance:** Biological determinant for hormonal influence.	**LLM Prompt:** Determine the biological sex of the patient from the text. **Context:** Assign 'Yes' (or 1) for Female. Assign 'No' (or 0) for Male.
Onset Timing Years **(Raj suggested)**	**Importance:** Helps estimate how long the patient has had epilepsy, which affects expected treatment response and outcomes. **Citation:** epilepsy duration, AAN	**LLM Prompt:** How long has the patient had epilepsy? Calculate the duration in years. **Questions:** (1) When did the seizures start (onset date or age at onset)? (2) What is the difference between the onset and the current date? **Context:** Look for "diagnosed in" or "seizures for [X] years". Return the duration as a number (Float).

Seizure Frequency **(Raj suggested)**	**Importance:** The primary measure of disease burden and treatment success. **Citation:** AAN Quality Measures	**Goal:** Classify seizure burden over the last 12 months into one of 5 strict categories. **Context:** Map the text to the closest class: 0 (Seizure Free): Absolutely no seizures for at least 1 year. 1 (Infrequent): Very rare events, occurring less than once a month (e.g., once or twice a year). 2 (Monthly): Occurring roughly once a month, but less than weekly. 3 (Weekly): Occurring at least once a week, or monthly events that happen in big clusters. 4 (Daily): Very severe burden, occurring every day or multiple times a day.

Seizure Type **(Raj suggested)**	**Importance:** Determines the choice of medication (Sodium channel blockers vs Broad spectrum). AAN Quality Measure. **Citation:** ILAE 2017 Operational Classification	**Goal:** Classify the dominant seizure type **Context:** 0 (Focal): Starts in one area (auras, one-sided movement, focal aware). 1 (Generalized Non-Motor): Staring spells or absence seizures without convulsions. 2 (Generalized Motor): Major convulsions (tonic-clonic/grand mal) dominating the picture. 3 (Mixed/Uncertain): Clinical picture is unclear or has features of both.
Female_Adolescent **(chatgpt suggested)**	**Importance:** Critical safety flag for "Women of Childbearing Potential" (WOCBP). Triggers avoidance of Valproate due to teratogenicity. Source: AES	**LLM Prompt:** Is this patient a female in the adolescent or early childbearing age range? **Questions:** (1) Is Sex_Female = Yes? (2) Is the age between 10 and 19? **Context:** Answer 'Yes' ONLY if both conditions are met.

Cognitive Priority risk_factors: - developmental_delay - behavioral_problem - psychiatric_problem	**Importance:** Important for understanding how epilepsy affects school, work, and day-to-day life. **Citation:** cognitive, psychological	**LLM Prompt:** Is there documented concern regarding the patient's cognitive function? **Questions:** (1) Are there mentions of intellectual disability, memory loss, school failure, or cognitive decline? (2) Are there "cognitive seizures"? **Context:** Answer 'Yes' if cognitive issues are a significant part of the clinical picture.
Hepatic Impairment Flag **(chatgpt suggested)**	**Importance:** Pharmacokinetic constraint. Prevents use of hepatotoxic or hepatically-metabolized drugs (Valproate, Phenytoin). **Citation:** data-driven, aes guide	**LLM Prompt:** Does the patient have liver disease that would affect medication choice? **Questions:** (1) Is there mention of "hepatic impairment", "liver failure", or "elevated liver enzymes"? **Context:** Answer 'Yes' if liver dysfunction is present. This is a safety constraint.

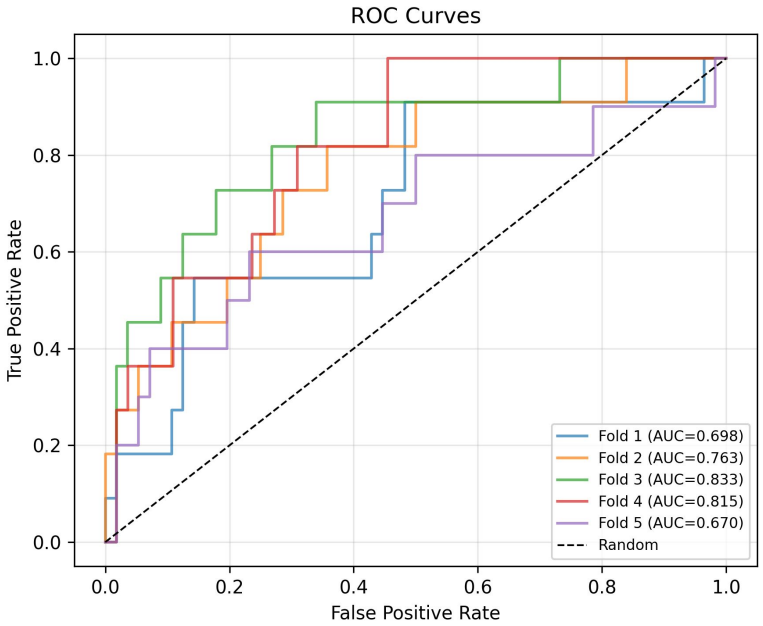
Renal Impairment Flag **(chatgpt suggested)**	**Importance:** Requires dose adjustment for renally-cleared drugs (Levetiracetam). **Citation:** data-driven, renal	**LLM Prompt:** Does the patient have kidney disease that would require drug dose adjustment? **Questions:** (1) Is there mention of "renal failure", "kidney disease", "CKD", or "dialysis"? **Context:** Answer 'Yes' if renal dysfunction is present. This is a safety constraint.
Status Or Prolonged Flag **(chatgpt suggested)**	**Importance:** "Red flag" for medical emergencies and high mortality risk. Differentiates self-limited seizures from those requiring rescue intervention. **Citation:** ILAE Status Epilepticus Definition	**LLM Prompt:** Does the patient have a history of Status Epilepticus or prolonged seizures? **Questions:** Is there a description of a single seizure lasting longer than 5 minutes? **Context:** Answer 'Yes' if either is true. This identifies dangerous/prolonged events requiring rescue plans.

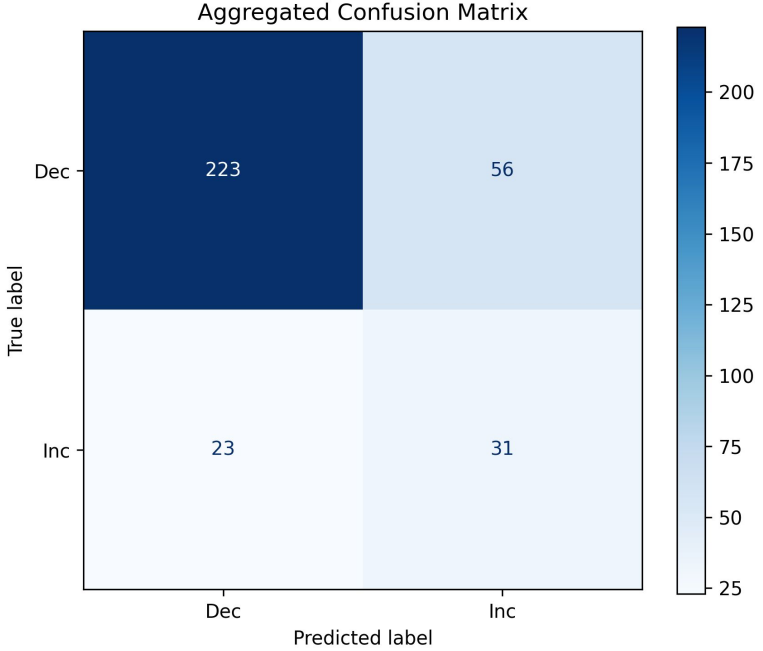

Feature distribution
Mulaga
St. Francis
Percentage (%)
0
10
20
30
40
50
60
70
80
Sex_Female
Female_Adolescent
StatusOrProlongedFlag
CognitivePriority
Liver impairment
Kidney impairment
SeizureFreqClass_0
SeizureFreqClass_1
SeizureFreqClass_2
SeizureFreqClass_3
SeizureFreqClass_4
Focal
Generalized (non-motor)
Generalized (motor)
Unknown (seizure type)
drug_valproate
drug_carbamazepine
drug_levetiracetam
drug_phenobarbital
drug_clobazam
drug_clonazepam
drug_topiramate
drug_lamotrigine
drug_ethosuximide
drug_phenytoin
drug_diazepam
Feature

Key differences

- Mulaga has considerably more cases with cognitive priority, generalized seizures and higher seizure frequency.
- St. Francis has considerably more patients with focal seizures and lower seizure frequencies.
- Mulaga relies more on valproate and carbamazepine, whereas St. Francis uses levetiracetam, lamotrigine, and phenobarbital far more frequently.

Mulago referral hospital





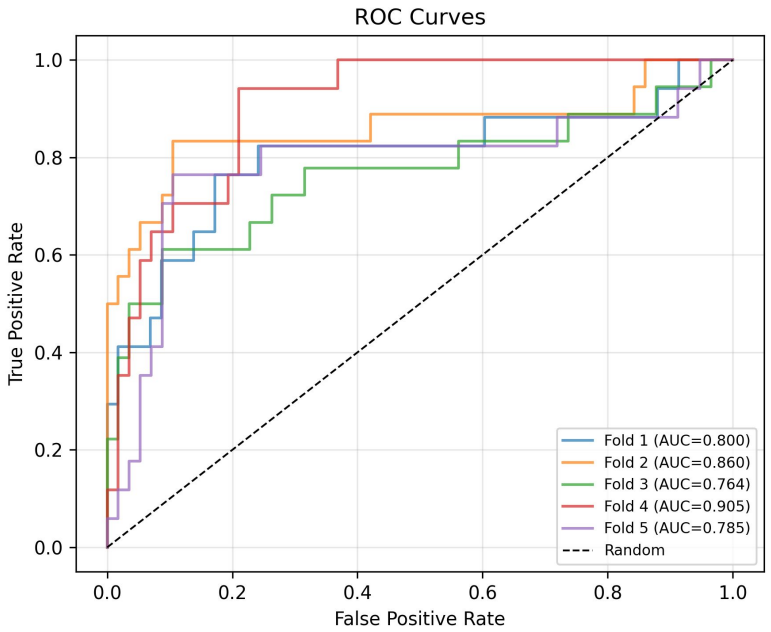
=== K-FOLD STRATIFIED RESULTS ===
Acc=0.806, F1=0.480, AUC=0.698, PR-AUC=0.397
Acc=0.746, F1=0.414, AUC=0.763, PR-AUC=0.505
Acc=0.746, F1=0.485, AUC=0.833, PR-AUC=0.532
Acc=0.773, F1=0.444, AUC=0.815, PR-AUC=0.470
Acc=0.742, F1=0.370, AUC=0.670, PR-AUC=0.356

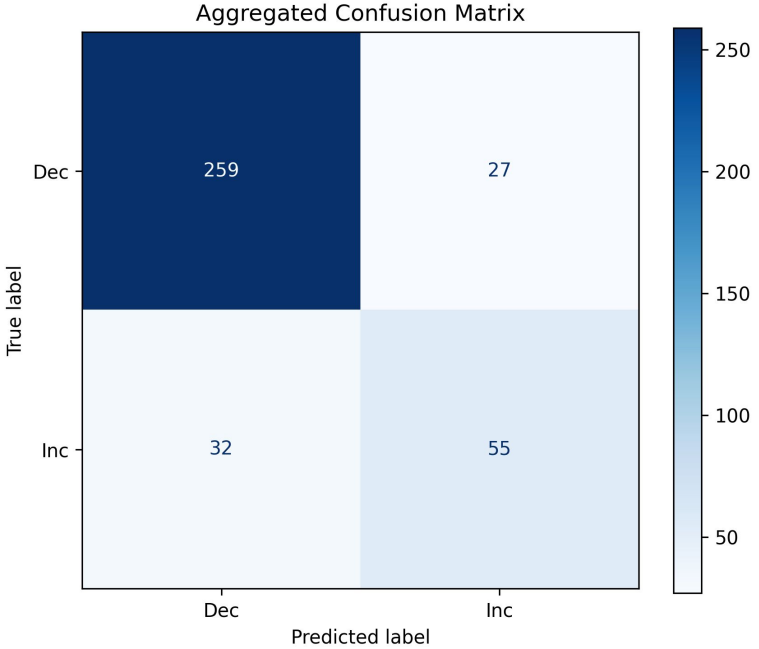
Total size: 333
Class balance:
Decrease 279
Increase 54

Mulago referral hospital

Rank	Feature Name	Importance	direction
1	SeizureFreqClass_0	4.823079	decrease
2	SeizureFreqClass_4	3.059332	increase
3	SeizureFreqClass_3	2.009308	increase
4	drug_phenobarbital	1.804581	increase
5	drug_clobazam	1.577393	increase
6	drug_phenytoin	1.504262	increase
7	Status Or Prolonged Flag	1.413082	decrease
8	drug_valproate	1.362898	increase
9	drug_carbamazepine	1.353787	increase
10	SeizureType_3 (Mixed/uncertain)	1.342673	increase

St. Francis Hospital





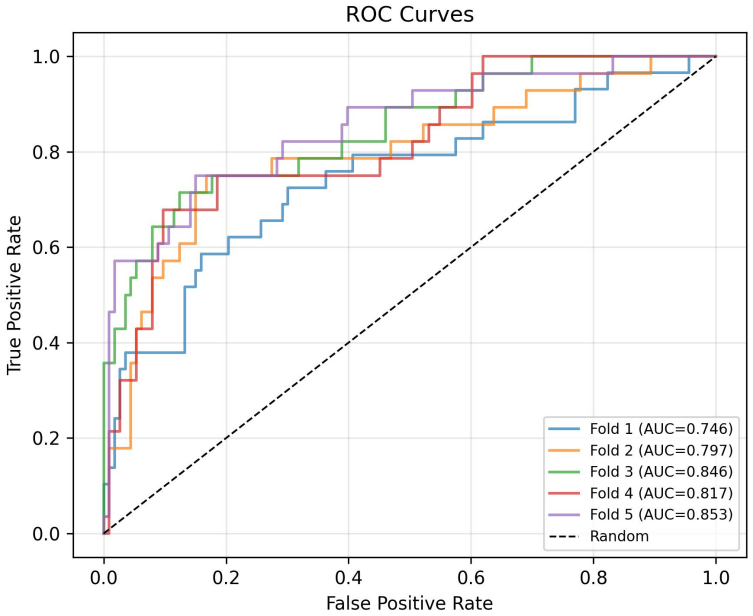
=== K-FOLD STRATIFIED RESULTS ===
Acc=0.813, F1=0.611, AUC=0.800, PR-AUC=0.687
Acc=0.880, F1=0.727, AUC=0.860, PR-AUC=0.812
Acc=0.813, F1=0.611, AUC=0.764, PR-AUC=0.667
Acc=0.838, F1=0.571, AUC=0.905, PR-AUC=0.729
Acc=0.865, F1=0.722, AUC=0.785, PR-AUC=0.589

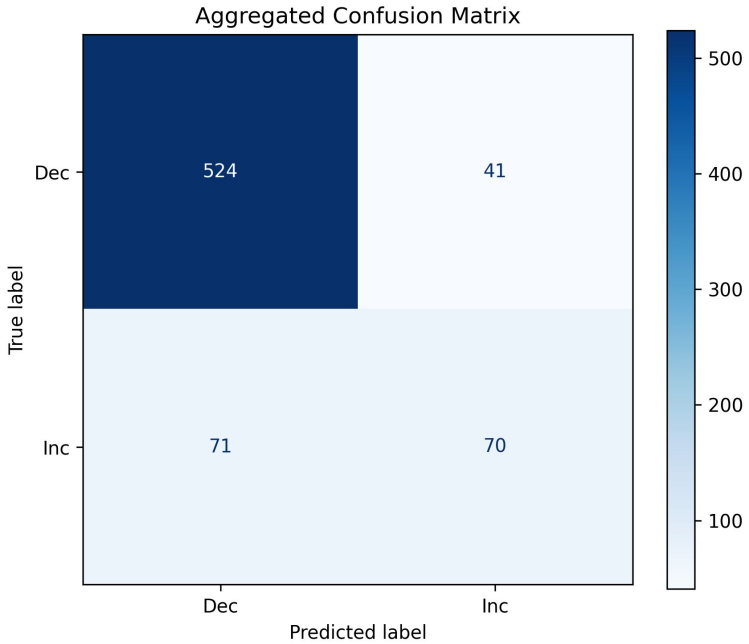
Total size: 373
Class balance:
Decrease 286
Increase 87

St. Francis hospital

Rank	Feature Name	Importance	direction
1	SeizureFreqClass_0	5.292486	decrease
2	SeizureFreqClass_3	4.122270	increase
3	SeizureType_0 (Focal)	2.071053	increase
4	SeizureFreqClass_4	1.769583	increase
5	drug_levetiracetam	1.584700	decrease
6	drug_valproate	1.575095	decrease
7	SeizureType_1(Gen non-motor)	1.493777	decrease
8	drug_phenobarbital	1.431862	increase
9	Sex_Female	1.412159	decrease
10	Cognitive Priority	1.384790	decrease

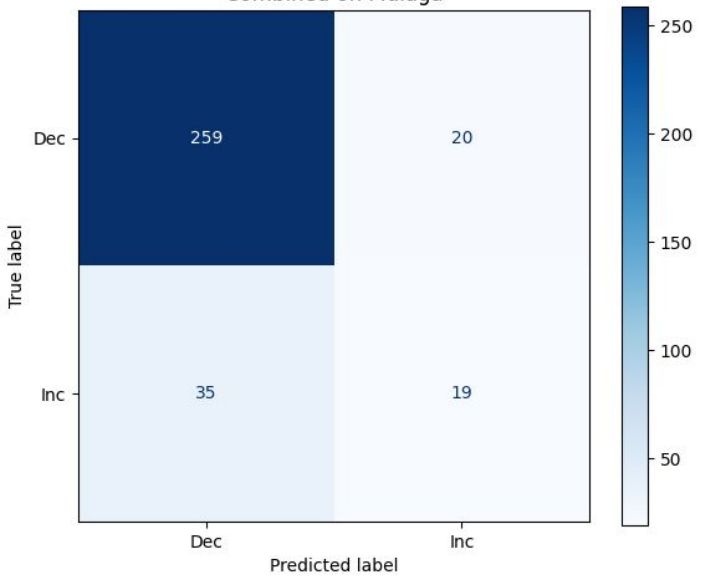
Combined hospitals

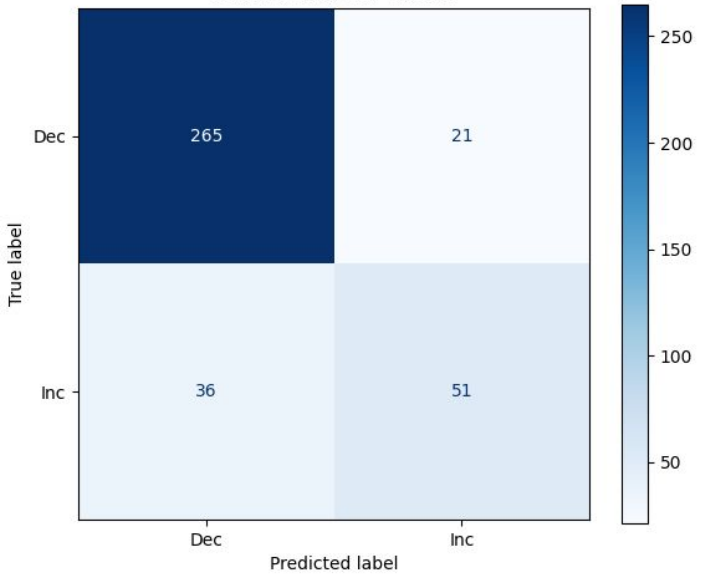


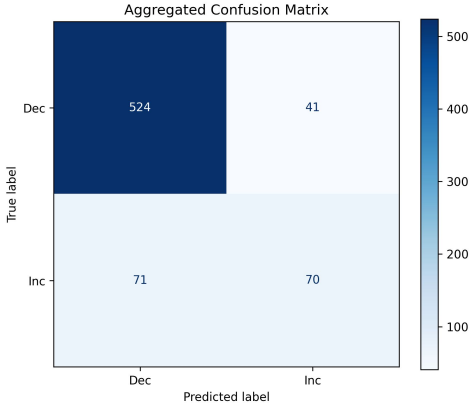


=== K-FOLD STRATIFIED RESULTS ===
Acc=0.739, F1=0.493, AUC=0.746, PR-AUC=0.538
Acc=0.752, F1=0.545, AUC=0.797, PR-AUC=0.554
Acc=0.702, F1=0.512, AUC=0.846, PR-AUC=0.712
Acc=0.794, F1=0.580, AUC=0.817, PR-AUC=0.567
Acc=0.716, F1=0.535, AUC=0.853, PR-AUC=0.683

Total size: 706
Class balance:
Decrease 565
Increase 141

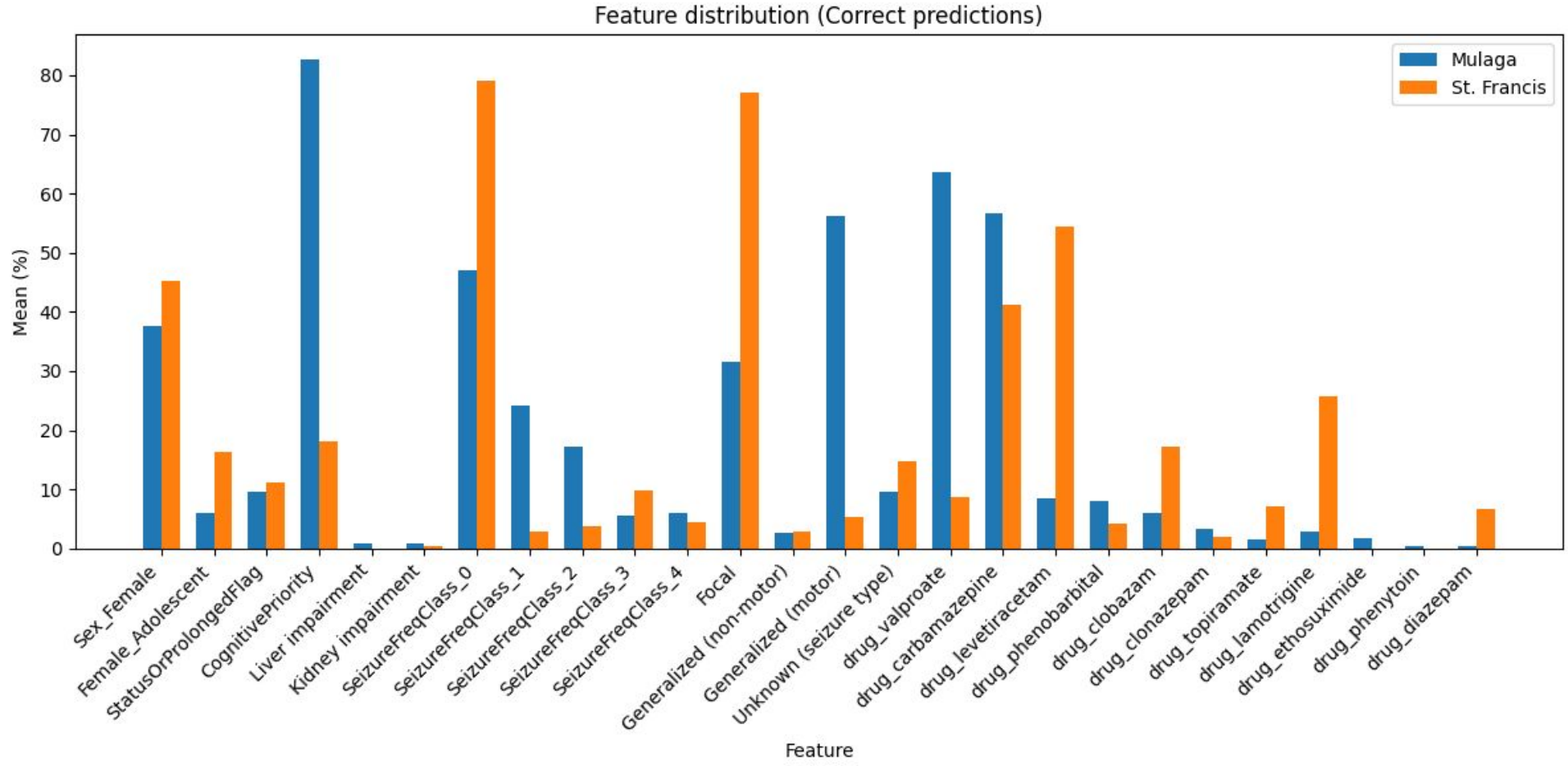

Combined on Mulaga
Dec
Inc
259
20
35
19
True label
Dec
Inc
Predicted label
250
200
150
100
50

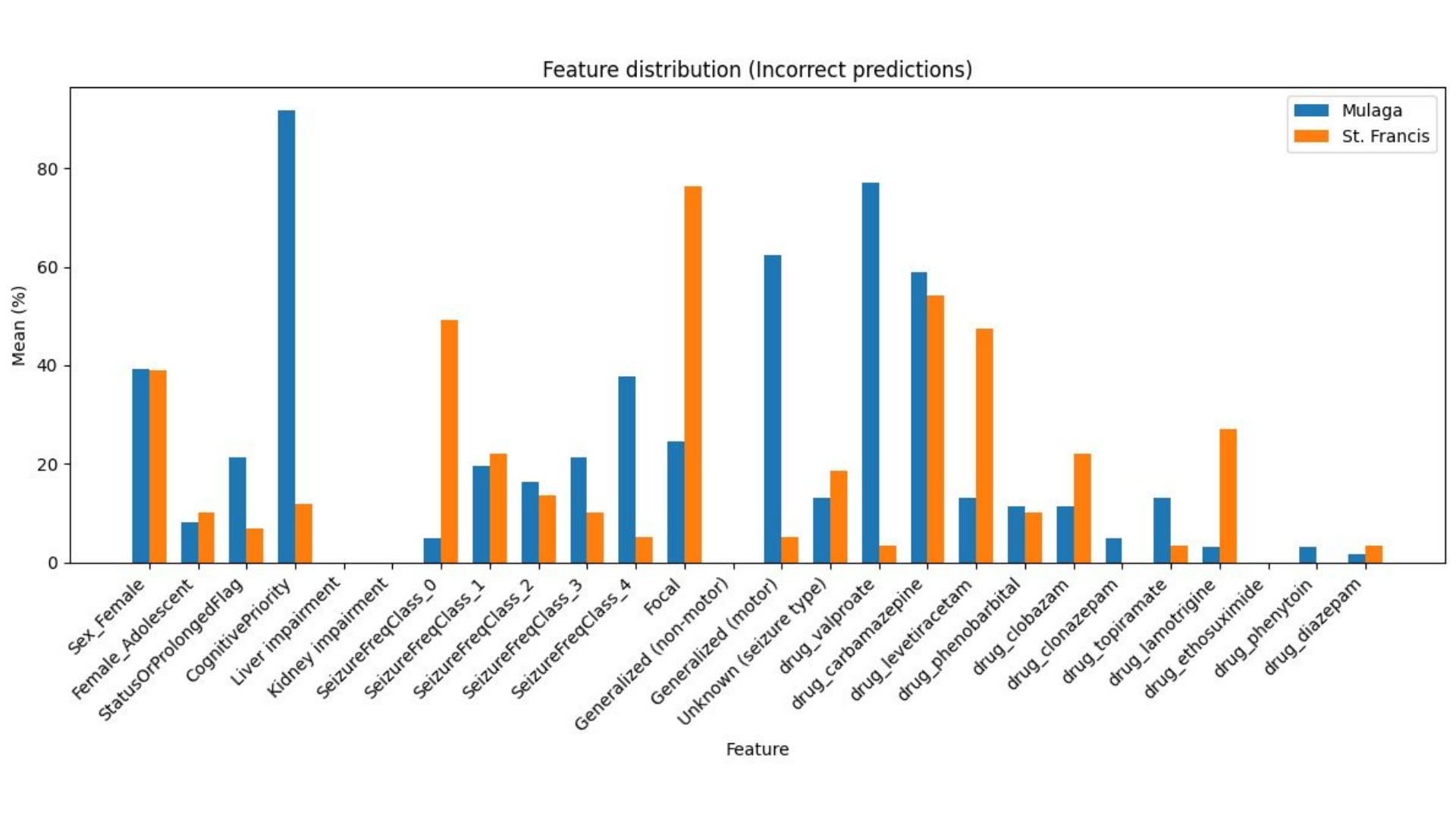

Combined on St. Francis
Dec
Inc
265
21
36
51
True label
Dec
Inc
Predicted label
250
200
150
100
50


Aggregated Confusion Matrix
Dec
Inc
524
41
71
70
True label
Dec
Inc
Predicted label
500
400
300
200
100

Combined hospitals

Rank	Feature Name	Importance	direction
1	SeizureFreqClass_0	4.051975	decrease
2	SeizureFreqClass_3	2.650415	increase
3	SeizureFreqClass_4	2.344753	increase
4	drug_phenobarbital	1.621822	increase
5	SeizureType_0 (Focal)	1.457884	increase
6	drug_clobazam	1.429757	increase
7	Cognitive Priority	1.398338	decrease
8	drug_lamotrigine	1.340478	increase
9	drug_diazepam	1.256786	increase
10	Status Or Prolonged Flag	1.215455	decrease


Feature distribution (Correct predictions)
Mulaga
St. Francis
Mean (%)
0
10
20
30
40
50
60
70
80
Sex_Female
Female_Adolescent
StatusOrPrologedFlag
CognitivePriority
Liver impairment
Kidney impairment
SeizureFreqClass_0
SeizureFreqClass_1
SeizureFreqClass_2
SeizureFreqClass_3
SeizureFreqClass_4
Focal
Generalized (non-motor)
Generalized (motor)
Unknown (seizure type)
drug_valproate
drug_carbamazepine
drug_levetiracetam
drug_phenobarbital
drug_clobazam
drug_clonazepam
drug_topiramate
drug_lamotrigine
drug_ethosuximide
drug_phenytoin
drug_diazepam
Feature


Feature distribution (Incorrect predictions)
Mulaga
St. Francis
Mean (%)
0
20
40
60
80
Sex_Female
Female_Adolescent
StatusOrProlongedFlag
CognitivePriority
Liver impairment
Kidney impairment
SeizureFreqClass_0
SeizureFreqClass_1
SeizureFreqClass_2
SeizureFreqClass_3
SeizureFreqClass_4
Focal
Generalized (non-motor)
Generalized (motor)
Unknown (seizure type)
drug_valproate
drug_carbamazepine
drug_levetiracetam
drug_phenobarbital
drug_clobazam
drug_clonazepam
drug_topiramate
drug_lamotrigine
drug_ethosuximide
drug_phenytoin
drug_diazepam
Feature