Final Model

Table of Contents

Fragile States Index Category

The Fragile States Index Category Tells uses criteria on refugees, economy, brain drain to assert a different categories of fragility for countries around the world. These are given a corresponding accuracy score based on the model used. Here its seen that predicting sustainable countries yeilds much higher accuracy in all the models while for example, when prediciting if a country should be in the `Warning` category for fragility, ffnn performs the best. If the value closer to 1 then it is more likely that the model's assertion was correct.

Human Development Index Category (Code)

`hdr_hdicode` defining group which tells us how developed a country is based on a in-depth criteria used by the United Nations. This serves as a summary measure for assessing long-term progress in three basic dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living. Those three dimensions then are broken down into codes which refers the which category the index is coded towards (hence `hdicode`). These include:

- QUALITY OF HUMAN DEVELOPMENT
- LIFE-COURSE GENDER GAP
- WOMEN'S EMPOWERMENT
- SOCIOECONOMIC SUSTAINABILITY

Human Development Index Category (Region)

The Human Development Index Cateogry tells us how developed a country is based on a in-depth criteria used by the United Nations. This data set in particular focuses on the specific regions around the world that a country could be classified under. These include:

- EAP: East Asia and Pacific
- ECA: Europe and Central Asia
- LAC: Latin America and Caribbean
- SA: South Asia
- SSA: Sub-Saharan Africa

With their associated accuracy scores we can see how well our three models can predict the level of development in each region correctly. If it is closer to 1 then it is more likely that the model's assertion was correct.

Background on methods of comparison

Sensitivity

- Quantifies how well a test identifies a true positive
- E.g., a test with 95% sensitivity will return a positive result 95% of the time, and a negative result 5% of the time
- A higher rate in sensitivity tests results in fewer false negative tests and less cases are missed.
- However, sensitivity does not take into account false positives, meaning that results may be incorrect depending on the sensitivity range

Accuracy

- The degree of how close a calculated measure is to the actual measure.
- Accuracy refers to a set of measurements whereas preciseness refers to single measurements
- Although accuracy can help statisticians determine how correct their data is based on the actual value, accuracy considers all measures equally important
- For example if a dataset contains only 10% of positive instances, a majority baseline classifier which always assigns the negative label would reach 90% accuracy since it would correctly predict 90% instances.

Specificity

- Specificity is the percentage of true negatives in a test
- For example, a test with 95% specificity will generate a negative result for 95% of people who do not meet the criteria, but will return a positive (or false positive) for those who do
- A higher rate in specificity tests results in fewer false positive tests and less cases are missed.
- However, specificity does not take into account false negatives, meaning that results may be incorrect depending on the specificity range

Explaining performance metrics

- A classification metric is a number that measures the performance that your machine learning model when it comes to assigning observations to certain classes.
- Binary classification is a particular situation where you just have to classes: positive and negative.
- Typically the performance is presented on a range from 0 to 1 where a score of 1 is reserved for the perfect model.

Comparing relevant groupings of countries of interest

- We can compare a number of different groupings of countries based on different data sets. This is important because different data aims to capture different aspects of conflict that can affect certain parts of the world in unique ways. These include:

Fragile States Index Category: uses criteria on refugees, economy, brain drain, etc

We can also use:

- Confusion Matrix
- And the associated False Positive, True Negative, True Positive and False Negative

Multicollinearity definition

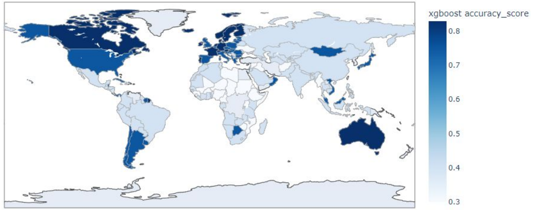
- Multicollinearity is something we are looking to avoid as it can lead to inflated standard errors of the regression coefficients. This makes it difficult to determine the precise contribution of each independent variable to the dependent variable.
- In numerical analysis, the condition number of a function measures how much the output value of the function can change for a small change in the input argument. This is used to measure how sensitive a function is to changes or errors in the input, and how much error in the output results from an error in the input.

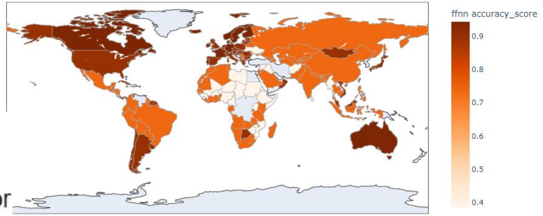
	yearmonth	fips	y_pred_transformer	y_pred_proba_transformer	y_true_transformer	y_pred_xgboost	y_pred_proba_xgboost
0	202211	FJ	False	0.183897	False	False	0.066500
1	202212	FJ	False	0.267831	False	False	0.099643
2	202211	TZ	False	0.482585	False	True	0.704086
3	202212	TZ	False	0.187792	False	True	0.638444
4	202301	TZ	True	0.539319	True	True	0.608380
...	...	...	...	...	...	...	...
359	202211	MJ	False	0.182196	False	False	0.079453
360	202212	MJ	False	0.203236	False	False	0.060189
361	202211	TD	True	0.527107	False	True	0.697625
362	202212	TD	True	0.555677	False	True	0.729246
363	202301	TD	True	0.565700	True	True	0.591722

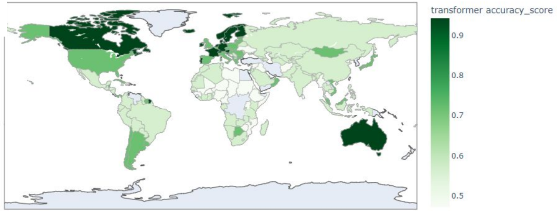
364 rows × 1343 columns

fsi_category

model	fsi_category	accuracy_score
xgboost	Alert	0.386
ffnn	Alert	0.300
transformer	Alert	0.471
xgboost	Stable	0.898
ffnn	Stable	0.750
transformer	Stable	0.716
xgboost	Sustainable	0.943
ffnn	Sustainable	0.829
transformer	Sustainable	0.943
xgboost	Warning	0.737
ffnn	Warning	0.398
transformer	Warning	0.561







Analysis

- The data shows that when the fsi_category shows the Stable and Sustainable category, the ffnn, xgboost, and transformer accuracy score is significantly higher in value.
  - The xgboost in these sections had the highest values of accuracy overall compared to the other models at 0.898 for Stable and 0.943 for Sustainable.
    - The ffnn shows 0.750 for Stable and 0.829 for Sustainable.
    - The transformer shows 0.716 for Stable and 0.948 for Sustainable.
- The category which shoes Alert and Warning have much lower accuracy for majority of the values.
  - The xgboost shows 0.386 for Alert and 0.737 for Warning.
    - The xgboost for Warning is significantly higher than the ffnn and transformer values.
  - The ffnn shows 0.300 for Alert and 0.398 for Warning.
  - The transformer shows 0.471 for Alert and 0.561 for Warning.

Hypothesis Testing and Confidence Intervals

- Sample is all of the data

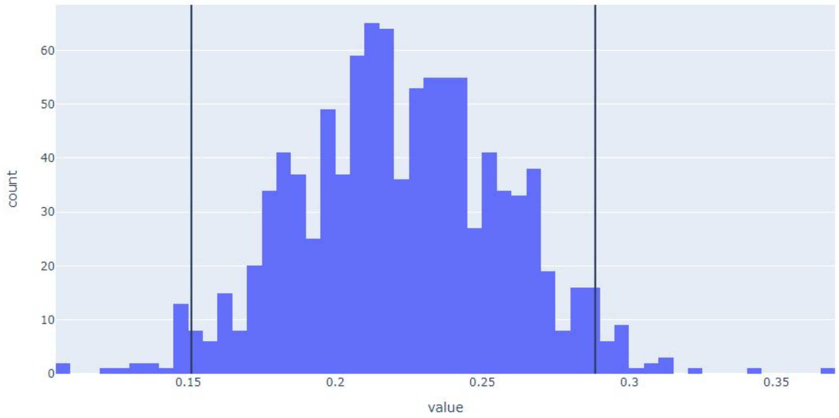
null hypothesis:

- H_0: $P_{\text{xgboost}} = P_{\text{ffnn}}$
- proportion of correct predictions for xgboost and ffnn are equal to one another

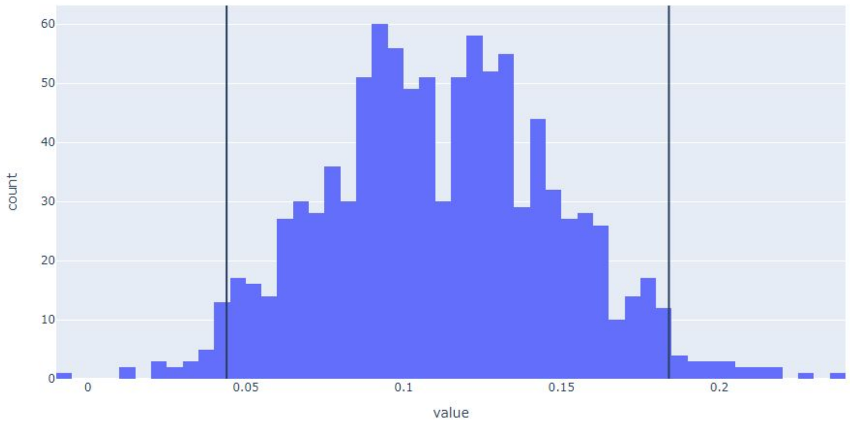
alternate hypothesis:

- H_1: $P_{\text{xgboost}} > P_{\text{ffnn}}$
- The proportion for xgboost is less than ffnn.

Confidence interval for xgboost vs ffnn



Confidence interval for xgboost vs transformer



Conclusion

Since our entire confidence interval is positive, we can conclude that we are 95% confident the true difference in proportions between xgboost model, ffnn model and transformer is positive. Therefore, we can conclude that we are 95% confident that the proportion of correct predictions for the xgboost model is greater than the proportion of correct predictions for the ffnn model and the transformer model.

Forward Selection Process

		coef	std err	t	P>|t|	[0.025	0.975]
0	const	0.3636	0.047	7.665	0.000	0.271	0.457
54	wbi_lending_category_Blend X xgboost	-0.3604	0.119	-3.032	0.002	-0.594	-0.127
88	hdr_region_ECA X xgboost X predicts1	0.1683	0.061	2.753	0.006	0.048	0.288
62	predicts1	0.2237	0.081	2.755	0.006	0.064	0.383
36	fsi_category_Warning X xgboost	0.1117	0.044	2.539	0.011	0.025	0.198
...	...	...	...	...	...	...	...
55	wbi_income_group_Low income X xgboost	0.0075	0.101	0.074	0.941	-0.190	0.205
24	wbi_other_(emu_or_hipc)_EMU X transformer	0.0024	0.038	0.062	0.951	-0.073	0.077
85	wbi_lending_category_IDA X xgboost X predicts1	0.0037	0.101	0.037	0.971	-0.195	0.202
5	wbi_lending_category_IDA X predicts1	0.0028	0.084	0.034	0.973	-0.162	0.168
12	wbi_income_group_Upper middle income X predicts1	0.0012	0.039	0.031	0.975	-0.075	0.077

```
model_2_variables=[]
model_2_variables.append('fsi_category_Alert X xgboost X predicts1')
model_2_variables.append('fsi_category_Warning X xgboost X predicts1')
model_2_variables.append('hdr_region_LAC X xgboost')
model_2_variables.append('fsi_category_Alert X transformer X predicts1')
model_2_variables.append('predicts1')
model_2_variables.append('fsi_category_Warning X xgboost')
model_2_variables.append('wbi_income_group_Low income X predicts1')
model_2_variables.append('fsi_category_Alert X predicts1')
model_2_variables.append('xgboost')
model_2_variables.append('hdr_region_ECA X xgboost')
model_2_variables.append('fsi_category_Alert X xgboost')
model_2_variables.append('fsi_category_Warning X predicts1')
model_2_variables.append('hdr_region_EAP X xgboost')
model_2_variables.append('wbi_lending_category_IDA X predicts1')
model_2_variables.append('hdr_region_ECA X transformer')
```

		coef	std err	t	P>|t|	[0.025	0.975]
0	const	0.3636	0.047	7.665	0.000	0.271	0.457
54	wbi_lending_category_Blend X xgboost	-0.3604	0.119	-3.032	0.002	-0.594	-0.127
88	hdr_region_ECA X xgboost X predicts1	0.1683	0.061	2.753	0.006	0.048	0.288
62	predicts1	0.2237	0.081	2.755	0.006	0.064	0.383
36	fsi_category_Warning X xgboost	0.1117	0.044	2.539	0.011	0.025	0.198
23	wbi_lending_category_Blend X xgboost X predicts1	0.3758	0.151	2.492	0.013	0.080	0.672
20	xgboost	-0.2279	0.093	-2.454	0.014	-0.410	-0.046
61	hdr_region_LAC X xgboost	0.1025	0.042	2.436	0.015	0.020	0.185
79	hdr_region_LAC X predicts1	-0.0504	0.023	-2.207	0.028	-0.095	-0.006
76	wbi_lending_category_Blend X transformer	-0.1963	0.091	-2.157	0.031	-0.375	-0.018
59	fsi_category_Alert X xgboost	0.1884	0.088	2.149	0.032	0.016	0.360
38	xgboost X predicts1	0.1750	0.090	1.947	0.052	-0.001	0.351
51	hdr_region_ECA	-0.0550	0.030	-1.849	0.065	-0.113	0.003
91	fsi_category_Alert	0.0486	0.031	1.574	0.116	-0.012	0.109
44	hdr_region_ECA X xgboost	-0.0909	0.058	-1.563	0.118	-0.205	0.023
16	fsi_category_Alert X xgboost X predicts1	-0.1264	0.081	-1.559	0.119	-0.285	0.033
81	wbi_lending_category_Blend X transformer X pre...	0.1690	0.131	1.287	0.198	-0.089	0.427
65	fsi_category_Warning	0.0299	0.024	1.271	0.204	-0.016	0.076
39	hdr_region_SA X transformer	0.0578	0.049	1.176	0.240	-0.039	0.154
35	fsi_category_Warning X predicts1	-0.0536	0.049	-1.093	0.275	-0.150	0.043

```
model_3_variables=[]
model_3_variables.append('fsi_category_Alert X xgboost X predicts1')
model_3_variables.append('fsi_category_Warning X xgboost X predicts1')
model_3_variables.append('hdr_region_LAC X xgboost')
model_3_variables.append('fsi_category_Alert')
model_3_variables.append('wbi_income_group_Lower middle income X xgboost')
model_3_variables.append('predicts1')
model_3_variables.append('fsi_category_Warning X xgboost')
model_3_variables.append('wbi_income_group_Low income X predicts1')
model_3_variables.append('fsi_category_Alert X predicts1')
model_3_variables.append('xgboost')
model_3_variables.append('hdr_region_ECA X xgboost')
model_3_variables.append('fsi_category_Alert X xgboost')
model_3_variables.append('wbi_income_group_Upper middle income X transformer X predicts1')
model_3_variables.append('wbi_income_group_Upper middle income X predicts1')
model_3_variables.append('fsi_category_Warning X predicts1')
model_3_variables.append('hdr_hdicode_Medium')
model_3_variables.append('wbi_lending_category_Blend X predicts1')
model_3_variables.append('hdr_region_EAP X xgboost')
model_3_variables.append('wbi_lending_category_IDA X predicts1')
model_3_variables.append('hdr_region_LAC')
model_3_variables.append('fsi_category_Warning')
```

OLS Regression of Model 1

OLS Regression Results

Dep. Variable:	error	R-squared:	0.542
Model:	OLS	Adj. R-squared:	0.499
Method:	Least Squares	F-statistic:	12.83
Date:	Wed, 06 Dec 2023	Prob (F-statistic):	1.26e-115
Time:	23:40:37	Log-Likelihood:	759.54
No. Observations:	1092	AIC:	-1333.
Df Residuals:	999	BIC:	-868.5
Df Model:	92		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3146	0.064	4.885	0.000	0.188	0.441
predicts1	0.2594	0.104	2.494	0.013	0.055	0.464
fsi_category_Alert	0.0528	0.043	1.237	0.216	-0.031	0.136
fsi_category_Sustainable	-0.0096	0.028	-0.344	0.731	-0.064	0.045
fsi_category_Warning	0.0287	0.027	1.049	0.295	-0.025	0.082
hdr_hdicode_High	0.0328	0.054	0.608	0.543	-0.073	0.139
hdr_hdicode_Very High	0.0170	0.062	0.273	0.785	-0.105	0.139
hdr_hdicode_Medium	0.0739	0.053	1.392	0.164	-0.030	0.178
hdr_region_AS	0.0230	0.033	0.704	0.482	-0.041	0.087
hdr_region_EAP	-0.0018	0.034	-0.052	0.959	-0.069	0.065
hdr_region_ECA	-0.0643	0.033	-1.954	0.051	-0.129	0.000
hdr_region_LAC	0.0228	0.026	0.891	0.373	-0.027	0.073
hdr_region_SA	-0.0216	0.042	-0.520	0.603	-0.103	0.060
wbi_income_group_Low income	0.0947	0.146	0.649	0.516	-0.191	0.381
wbi_income_group_Lower middle income	0.0238	0.051	0.469	0.639	-0.076	0.123
wbi_income_group_Upper middle income	0.0202	0.040	0.504	0.614	-0.058	0.099
wbi_lending_category_Blend	0.0264	0.061	0.433	0.665	-0.093	0.146
wbi_lending_category_IDA	0.0090	0.032	0.277	0.782	-0.055	0.072
wbi_other_(emu_or_hipc)_EMU	0.0045	0.027	0.166	0.868	-0.049	0.058
fsi_category_Alert X predicts1	0.0060	0.065	0.092	0.926	-0.121	0.133
fsi_category_Warning X predicts1	-0.0237	0.061	-0.388	0.698	-0.144	0.096
hdr_hdicode_High X predicts1	-0.0005	0.037	-0.015	0.988	-0.074	0.073
hdr_hdicode_Very High X predicts1	0.0248	0.049	0.503	0.615	-0.072	0.122
hdr_hdicode_Medium X predicts1	-0.0923	0.066	-1.395	0.163	-0.222	0.038
hdr_region_AS X predicts1	-0.0344	0.061	-0.565	0.572	-0.154	0.085
hdr_region_EAP X predicts1	-0.0492	0.047	-1.052	0.293	-0.141	0.043
hdr_region_ECA X predicts1	0.0701	0.030	2.339	0.020	0.011	0.129
hdr_region_LAC X predicts1	-0.0513	0.023	-2.234	0.026	-0.096	-0.006
hdr_region_SA X predicts1	0.0314	0.050	0.625	0.532	-0.067	0.130
wbi_income_group_Low income X predicts1	-0.1653	0.211	-0.785	0.432	-0.578	0.248
wbi_income_group_Lower middle income X predicts1	-0.0816	0.154	-0.529	0.597	-0.384	0.221
wbi_income_group_Upper middle income X predicts1	0.0016	0.041	0.039	0.969	-0.078	0.082
wbi_lending_category_Blend X predicts1	-0.0574	0.104	-0.554	0.580	-0.261	0.146
wbi_lending_category_IDA X predicts1	-0.0088	0.092	-0.097	0.923	-0.189	0.171
transformer	0.1035	0.096	1.073	0.284	-0.086	0.293
xgboost	-0.1789	0.109	-1.648	0.100	-0.392	0.034
fsi_category_Alert X transformer	-0.0113	0.063	-0.178	0.859	-0.136	0.113
fsi_category_Alert X xgboost	0.1837	0.098	1.876	0.061	-0.008	0.376
fsi_category_Sustainable X transformer	-0.0262	0.040	-0.656	0.512	-0.105	0.052
fsi_category_Sustainable X xgboost	-0.0020	0.040	-0.051	0.959	-0.080	0.076
fsi_category_Warning X transformer	0.0193	0.041	0.467	0.641	-0.062	0.101
fsi_category_Warning X xgboost	0.1131	0.046	2.442	0.015	0.022	0.204
hdr_hdicode_High X transformer	-0.1039	0.073	-1.430	0.153	-0.247	0.039
hdr_hdicode_High X xgboost	0.0045	0.084	0.054	0.957	-0.160	0.169
hdr_hdicode_Very High X transformer	-0.1368	0.094	-1.454	0.146	-0.321	0.048
wbi_income_group_Upper middle income X transformer	-0.0414	0.071	-0.586	0.558	-0.180	0.097
wbi_income_group_Upper middle income X xgboost	0.0240	0.078	0.305	0.760	-0.130	0.178
wbi_lending_category_Blend X transformer	-0.1966	0.092	-2.146	0.032	-0.376	-0.017
wbi_lending_category_Blend X xgboost	-0.3637	0.119	-3.050	0.002	-0.598	-0.130
wbi_lending_category_IDA X transformer	-0.0287	0.051	-0.557	0.578	-0.130	0.072
wbi_lending_category_IDA X xgboost	0.0169	0.057	0.296	0.768	-0.095	0.129
wbi_other_(emu_or_hipc)_EMU X transformer	0.0022	0.039	0.057	0.955	-0.074	0.078
wbi_other_(emu_or_hipc)_EMU X xgboost	0.0283	0.039	0.732	0.465	-0.048	0.104
transformer X predicts1	-0.0901	0.138	-0.651	0.516	-0.362	0.182
xgboost X predicts1	0.1395	0.139	1.007	0.314	-0.132	0.411
fsi_category_Alert X transformer X predicts1	-0.0057	0.091	-0.062	0.951	-0.185	0.174
fsi_category_Alert X xgboost X predicts1	-0.1482	0.107	-1.384	0.167	-0.358	0.062
fsi_category_Warning X transformer X predicts1	-0.0128	0.079	-0.163	0.871	-0.167	0.142
fsi_category_Warning X xgboost X predicts1	-0.0611	0.074	-0.830	0.407	-0.206	0.083
hdr_hdicode_High X transformer X predicts1	0.0658	0.058	1.136	0.256	-0.048	0.180
hdr_hdicode_High X xgboost X predicts1	-0.0664	0.057	-1.164	0.245	-0.178	0.046
hdr_hdicode_Very High X transformer X predicts1	0.0862	0.077	1.124	0.261	-0.064	0.237
hdr_hdicode_Very High X xgboost X predicts1	-0.0613	0.075	-0.814	0.416	-0.209	0.087
hdr_hdicode_Medium X transformer X predicts1	0.0888	0.087	1.021	0.307	-0.082	0.259
hdr_hdicode_Medium X xgboost X predicts1	-0.0053	0.097	-0.055	0.956	-0.196	0.186
hdr_region_AS X transformer X predicts1	0.0036	0.081	0.044	0.965	-0.156	0.163
hdr_region_AS X xgboost X predicts1	0.0569	0.081	0.706	0.481	-0.101	0.215
hdr_region_ECA X transformer X predicts1	-0.0277	0.044	-0.626	0.531	-0.114	0.059
hdr_region_ECA X xgboost X predicts1	0.0978	0.045	2.169	0.030	0.009	0.186
hdr_region_LAC X transformer X predicts1	-0.0303	0.035	-0.862	0.389	-0.099	0.039
hdr_region_LAC X xgboost X predicts1	-0.0210	0.033	-0.629	0.529	-0.087	0.045
hdr_region_SA X xgboost X predicts1	0.0436	0.091	0.477	0.634	-0.136	0.223
wbi_income_group_Low income X transformer X predicts1	0.1742	0.242	0.720	0.472	-0.301	0.649
wbi_income_group_Low income X xgboost X predicts1	0.0922	0.254	0.363	0.717	-0.407	0.591
wbi_income_group_Lower middle income X transformer X predicts1	0.0995	0.185	0.539	0.590	-0.263	0.462
wbi_income_group_Lower middle income X xgboost X predicts1	0.0617	0.185	0.334	0.738	-0.301	0.424
wbi_income_group_Upper middle income X transformer X predicts1	0.0530	0.063	0.848	0.397	-0.070	0.176
wbi_income_group_Upper middle income X xgboost X predicts1	-0.0514	0.061	-0.838	0.402	-0.172	0.069
wbi_lending_category_Blend X transformer X predicts1	0.1776	0.134	1.320	0.187	-0.086	0.441
wbi_lending_category_Blend X xgboost X predicts1	0.3839	0.153	2.510	0.012	0.084	0.684
wbi_lending_category_IDA X transformer X predicts1	-0.0269	0.110	-0.243	0.808	-0.243	0.190
wbi_lending_category_IDA X xgboost X predicts1	0.0154	0.110	0.140	0.889	-0.200	0.231

Omnibus:	91.959	Durbin-Watson:	1.867
Prob(Omnibus):	0.000	Jarque-Bera (JB):	331.800
Skew:	0.341	Prob(JB):	8.92e-73
Kurtosis:	5.613	Cond. No.	9.78e+15

OLS Regression Results of Model 2

OLS Regression Results

Dep. Variable:	error	R-squared:	0.456
Model:	OLS	Adj. R-squared:	0.448
Method:	Least Squares	F-statistic:	60.07
Date:	Wed, 06 Dec 2023	Prob (F-statistic):	1.38e-130
Time:	18:37:59	Log-Likelihood:	665.90
No. Observations:	1092	AIC:	-1300.
Df Residuals:	1076	BIC:	-1220.
Df Model:	15		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3660	0.006	61.009	0.000	0.354	0.378
fsi_category_Alert X xgboost X predicts1	0.0164	0.047	0.350	0.726	-0.076	0.108
fsi_category_Warning X xgboost X predicts1	0.0457	0.026	1.755	0.080	-0.005	0.097
hdr_region_LAC X xgboost	0.0973	0.020	4.952	0.000	0.059	0.136
fsi_category_Alert X transformer X predicts1	-0.0054	0.028	-0.194	0.846	-0.060	0.050
predicts1	0.2648	0.020	13.099	0.000	0.225	0.304
fsi_category_Warning X xgboost	0.1112	0.023	4.801	0.000	0.066	0.157
wbi_income_group_Low income X predicts1	-0.0126	0.021	-0.587	0.557	-0.055	0.030
fsi_category_Alert X predicts1	-0.0557	0.031	-1.801	0.072	-0.116	0.005
xgboost	-0.1480	0.014	-10.322	0.000	-0.176	-0.120
hdr_region_ECA X xgboost	-0.0417	0.025	-1.701	0.089	-0.090	0.006
fsi_category_Alert X xgboost	0.1701	0.040	4.201	0.000	0.091	0.250
fsi_category_Warning X predicts1	-0.0669	0.025	-2.717	0.007	-0.115	-0.019
hdr_region_EAP X xgboost	0.0033	0.027	0.121	0.904	-0.050	0.057
wbi_lending_category_IDA X predicts1	-0.0133	0.019	-0.698	0.485	-0.051	0.024
hdr_region_ECA X transformer	-0.0344	0.023	-1.519	0.129	-0.079	0.010

Omnibus:	27.603	Durbin-Watson:	1.694
Prob(Omnibus):	0.000	Jarque-Bera (JB):	46.816
Skew:	0.193	Prob(JB):	6.82e-11
Kurtosis:	3.938	Cond. No.	19.0

OLS Regression Results of Model 3

OLS Regression Results

Dep. Variable:	error	R-squared:	0.492
Model:	OLS	Adj. R-squared:	0.482
Method:	Least Squares	F-statistic:	49.36
Date:	Wed, 06 Dec 2023	Prob (F-statistic):	3.96e-141
Time:	22:09:56	Log-Likelihood:	703.58
No. Observations:	1092	AIC:	-1363.
Df Residuals:	1070	BIC:	-1253.
Df Model:	21		
Covariance Type:	nonrobust		

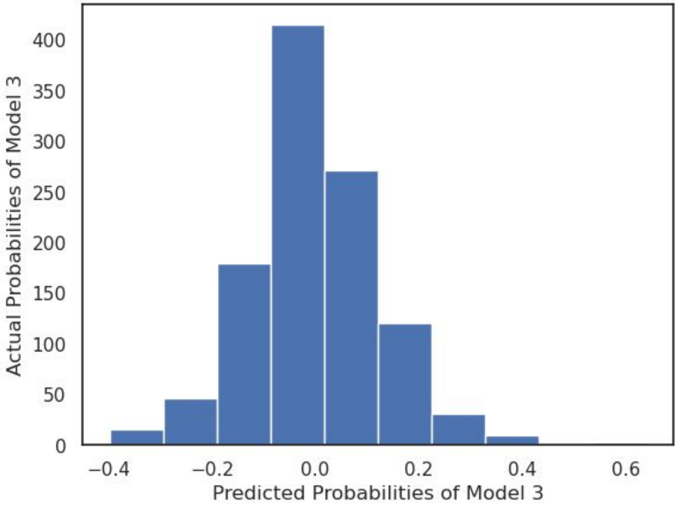
	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3122	0.009	35.921	0.000	0.295	0.329
fsi_category_Alert X xgboost X predicts1	0.0677	0.047	1.432	0.152	-0.025	0.160
fsi_category_Warning X xgboost X predicts1	0.0844	0.030	2.861	0.004	0.027	0.142
hdr_region_LAC X xgboost	0.0546	0.025	2.225	0.026	0.006	0.103
fsi_category_Alert	0.0973	0.020	4.752	0.000	0.057	0.137
wbi_income_group_Lower middle income X xgboost	-0.0136	0.022	-0.613	0.540	-0.057	0.030
predicts1	0.2823	0.021	13.608	0.000	0.242	0.323
fsi_category_Warning X xgboost	0.0456	0.026	1.741	0.082	-0.006	0.097
wbi_income_group_Low income X predicts1	-0.0072	0.023	-0.312	0.755	-0.052	0.038
fsi_category_Alert X predicts1	-0.1269	0.034	-3.697	0.000	-0.194	-0.060
xgboost	-0.0985	0.015	-6.521	0.000	-0.128	-0.069
hdr_region_ECA X xgboost	-0.0351	0.024	-1.436	0.151	-0.083	0.013
fsi_category_Alert X xgboost	0.0784	0.044	1.776	0.076	-0.008	0.165
wbi_income_group_Upper middle income X transformer X predicts1	0.0351	0.034	1.019	0.308	-0.032	0.103
wbi_income_group_Upper middle income X predicts1	-0.0005	0.029	-0.017	0.986	-0.058	0.057
fsi_category_Warning X predicts1	-0.1213	0.028	-4.350	0.000	-0.176	-0.067
hdr_hdicode_Medium	0.0342	0.011	3.231	0.001	0.013	0.055
wbi_lending_category_Blend X predicts1	-0.0343	0.025	-1.346	0.179	-0.084	0.016
hdr_region_EAP X xgboost	0.0043	0.028	0.155	0.877	-0.050	0.059
wbi_lending_category_IDA X predicts1	-0.0130	0.023	-0.575	0.565	-0.058	0.031
hdr_region_LAC	0.0382	0.013	2.871	0.004	0.012	0.064
fsi_category_Warning	0.0643	0.013	5.099	0.000	0.040	0.089

Omnibus:	63.222	Durbin-Watson:	1.764
Prob(Omnibus):	0.000	Jarque-Bera (JB):	149.968
Skew:	0.325	Prob(JB):	2.72e-33
Kurtosis:	4.695	Cond. No.	23.7

Residuals Graph of Model 3

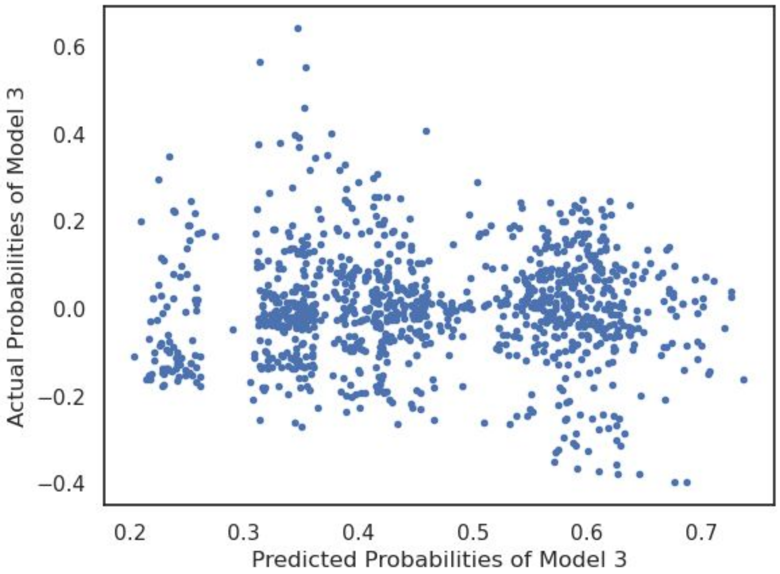
$$Residual = ObservedValue - PredictedValue$$

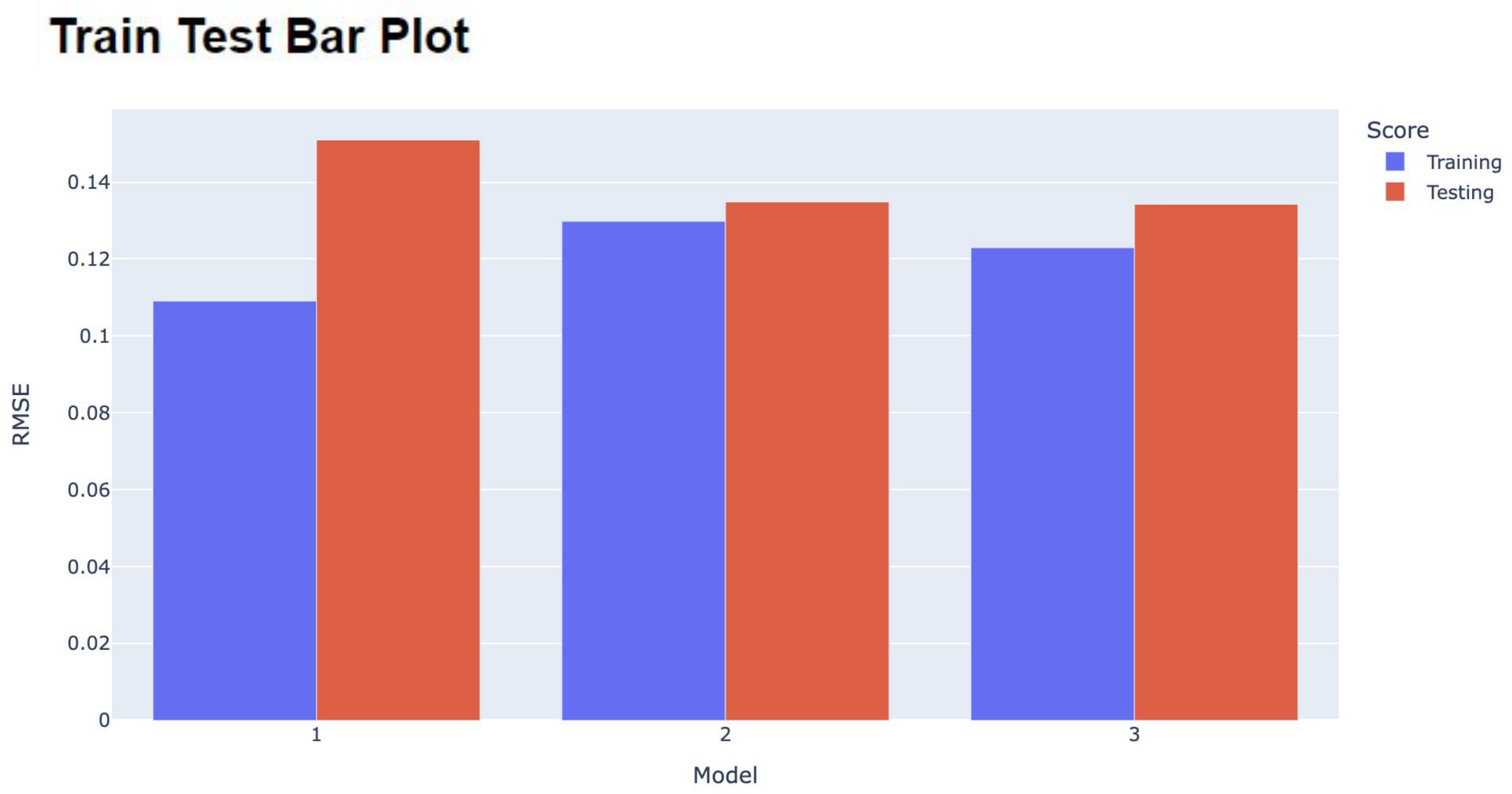
Our residual graphs shows rough normal distribution which is generally an ideal outcome as it allows us to make further statstical inferences. We also have a slight left skew meaning the mean is less than the median residual value. Additionally, since this is left skewed we can say that the plot is also homoskadistic. This means the variance in errors (within the groups being compared) is largely constant across all levels of the independent variable. Also, we can infer based on the normality that the model's predictions have a high reliability and accuracy.



Residuals Scatter Plot of Model 3

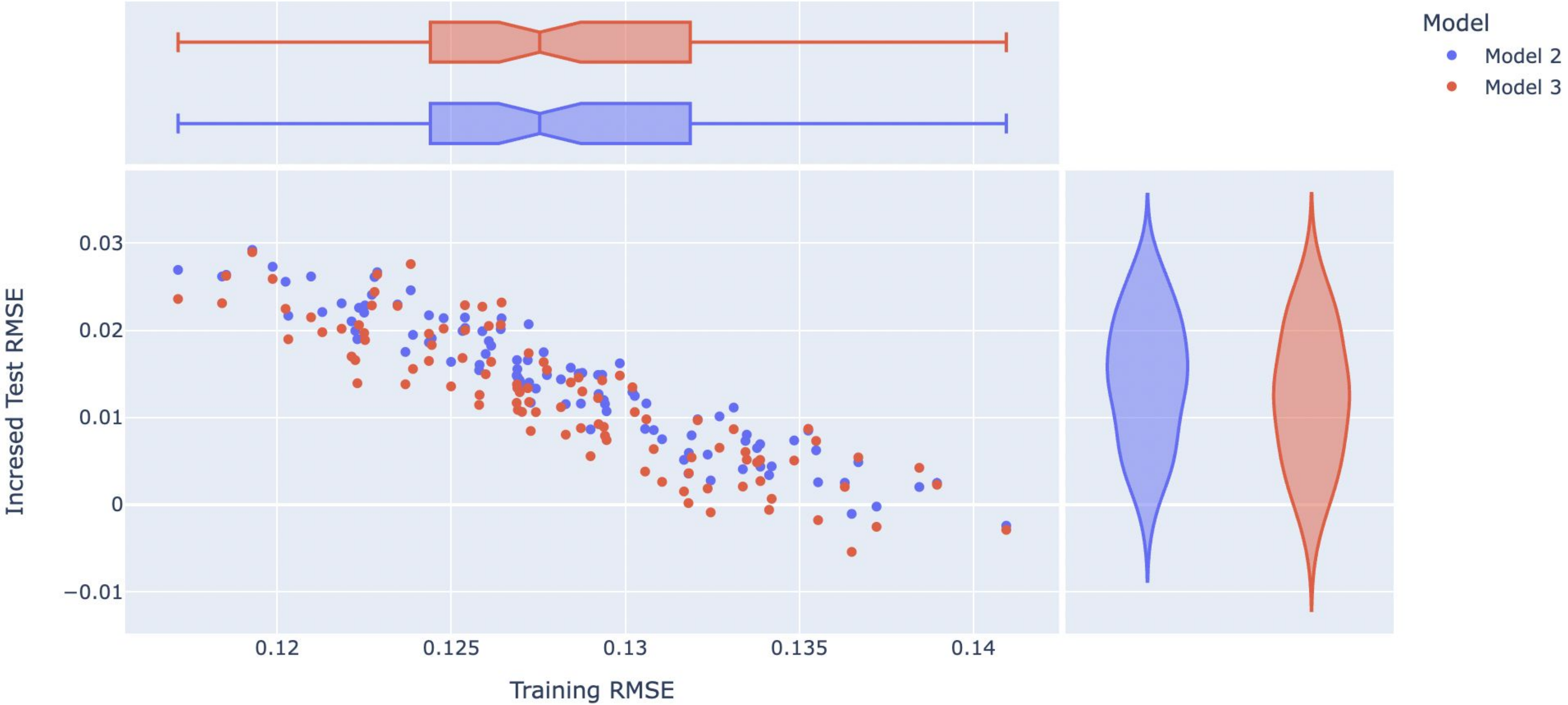
In general, we use a scatter plot of residuals to determine whether or not a linear model would be appropriate for modelling the data that we use. In our case our residuals scatter plot shows no correlation and could be classified as entirely random in the placement of the points. This type of residuals scatter plot is known as a null residual plot. Generally, this is seen as an ideal outcome in the context of comparing models as it indicates that the models have no obvious defects. As a result, we can infer that our data can be modelled by a linear model using an expected value function.





Train Test Bar Plot
Score
Training
Testing
RMSE
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1
2
3
Model

Train Test Scatter Plot



Train Test Observations

The root mean squared error for model 1 testing on our chosen variables is 4.18713% which is vastly different from the actual training model, which has an RMSE of 15.09523%. This signalfies overfitting as the RMSEs are vastly different with a change in error of 10.9081%. The reason why the difference in RMSE is so high is because model 1 contains multiple variables which have a significant impact on its accuracy. These variables are those with high p-values, which are included, resulting in this large difference.

In our second model, we ordered our variables by p-value and selected only those with low values, resulting in our model 2 having an RMSE of 13.5854%, compared to the testing RMSE of 13.12096%. This demonstrates an extremely narrow change in RMSE, which again, signalfies overfitting. This is because too few variables that have direct correlation with the training data were selected, resulting in model 2 memorizing the data instead of learning the patterns of the data to make predictions.

In model 3, we decided to add 11 more variables to ensure there would not be overfitting, resulting in an RMSE of our test data of 12.29667%, compared to the 13.4248%, resulting in a significant enough difference that is close enough to the training data that overfitting would not be a potential issue.

These observations of model 3 vs model 2 were echoed by the scatter plot created. It shows that despite the similaries between the differences in RMSE, model 3 still has a significant enough difference of ~2% between training and testing data to say potential overfitting is not an issue. By these observations, we can come to the conclusion that out of our models, model 3 is the best performing as it has a similar RMSE to the test data, yet is not too close such that overfitting is a concern. Based on this conclusion, we can use this model to create the expected value formula that can be expressed as:

$$
\begin{aligned}
E(x) = {} & 0.3278 + (0.0677 * \text{fsi\_category\_Alert X xgboost X predicts1}) + (0.0844 * \text{fsi\_category\_Warning X xgboost X predicts}) \\
& + (0.0546 * \text{hdr\_region\_LAC X xgboost}) + (0.1753 * fsi_category_Alert) + (-0.0028 * \text{wbi\_income\_group\_Lower middle income X xgboost}) \\
& + (0.2420 * \text{predicts1}) + (0.1385 * \text{fsi\_category\_Warning X xgboost}) + (0.0225 * \text{wbi\_income\_group\_Low income X predicts1}) \\
& + (-0.2846 * \text{fsi\_category\_Alert X predicts1}) + (-0.0632 * \text{xgboost}) + (-0.0716 * \text{hdr\_region\_ECA X xgboost}) \\
& + (0.1213 * \text{fsi\_category\_Alert X xgboost}) + (0.0656 * \text{wbi\_income\_group\_Upper middle income X transformer X predicts1}) \\
& + (-0.0325 * \text{wbi\_income\_group\_Upper middle income X predicts1}) + (0.0644 * \text{wbi\_lending\_category\_Blend}) \\
& + (-0.2619 * \text{fsi\_category\_Warning X predicts1}) + (0.0301 * \text{hdr\_hdicode\_Medium}) + (-0.1166 * \text{wbi\_lending\_category\_Blend X predicts1}) \\
& + (-0.0193 * \text{hdr\_region\_EAP X xgboost}) + (0.0274 * \text{wbi\_lending\_category\_IDA X predicts1}) + (0.0481 * \text{hdr\_region\_LAC}) \\
& + (-0.0074 * \text{hdr\_region\_ECA X transformer}) + (0.1025 * \text{fsi\_category\_Warning})
\end{aligned}
$$

Correlation Heatmap

Based on the heatmap, we can observe few dark areas, suggesting little correlation between variables which implies having low multicollinearity.

Multicollinearity is bad because it can lead to inflated standard errors of the regression coefficients. This makes it difficult to determine the precise contribution of each independent variable to the dependent variable.

Discussing Ethics

Understanding the Ethical Implications of Predictive Models in Conflict Resolution

In conflict resolution, the use of predictive models has opened new ways in mitigating crises. However, this comes with a complex ethical terrain that needs carefully considerated. The application of predictive analytics in conflict zones can promise more advanced foresight and humanitarian preparedness but also poses significant ethical challenges such as predictive privacy concerns, the need for human-centered AI, and the adherence to legal and moral responsibilities in humanitarian actions.

Addressing these issues is crucial for the responsible and effective use of predictive models in conflict resolution.

Transparency and Privacy

The foundations of ethical decision-making in AI powered conflict resolution centers on ensuring that the development and application of AI in conflict resolution are guided by principles of transparency, accountability, and respect for individual privacy.

Clarity in Model Development and Application:

It's essential to be transparent about how predictive models are built and used. When models assess conflict situations, we must be clear about their capabilities and limits to avoid misinterpretation and misuse. This point is aligned with the emphasis on transparency and accountability in conflict resolution (Reference: ICRC documents).

Respecting Predictive Privacy:

Using predictive analytics in conflict areas might violate predictive privacy by exposing sensitive information without consent. An example is when models predict conflict zones, potentially revealing private data about individuals in those areas, as discussed by Rainer Mühlhoff.

Human-Centric Approach in AI and Humanitarian Actions

This emphasizes the importance of maintaining human control in AI applications for humanitarian work. It advocates for AI as an aid, not a replacement, for human decision-making in critical areas like identifying missing persons and evaluating humanitarian needs.

Human-Centric AI in Humanitarian Efforts:

Ethical AI use in conflict resolution should prioritize human decision-making. A relevant example is the AI-based identification of missing persons by the ICRC, where AI assists but doesn't replace human judgment, reflecting UNICEF's guidelines on human-centered AI.

Legal and Moral Responsibilities in Conflict Scenarios:

International humanitarian law requires human judgment in conflict resolution. Predictive models used to identify humanitarian needs must not bypass the legal applications in conflict zones (Reference: ICRC documents).

Balancing Efficiency with Ethical Implications

This highlights the importance of considering the impact of AI decisions on local communities and individuals, especially in conflict zones, while maintaining operational efficiency.

Bias and Inclusion in Predictive Modeling:

Predictive models must be scrutinized for biases. For instance, pattern recognition in satellite imagery should not perpetuate existing inequalities. This is particularly relevant in the context of the "Sampling bias in climate–conflict research" by Courtland Adams et al., highlighting the risk of focusing only on conflict-prone regions while ignoring peaceful areas.

Balancing Efficiency and Ethical Considerations:

While predictive models are efficient in data analysis, their ethical impact is significant. For example, models predicting the viability of food crops must weigh their effects on local communities, not just focus on efficiency.

In summary, the use of predictive models in conflict resolution is complex and demands a careful ethical approach, considering transparency, predictive privacy, inclusion, human-centric AI, and legal and moral responsibilities.

In Conclusion

Considering ethics in AI-driven conflict resolution is crucial because it ensures that we can still have respect for individual rights and making sure we uphold our societal values. Ethical considerations is a reminder for us to maintain a responsible use of technology, especially AI, and to ensure that AI decisions are fair, transparent, and accountable. This approach not only safeguards privacy and human dignity but also strengthens the trust in AI applications, especially in a highly sensitive area like conflict resolution.

Conclusion and Evaluating Machine Learning Models

Beginning with our first observations, we looked qualitatively at the maps created my FSI and HDR variables and found that some machine leanring models preformed better or worse depending on the region or variable.

Then we moved forward in our investigation by looking at which models preformed better on different variables by using hypothesis testing and confidence intervals. This allowed us to conclude that certain variables that preformed well on certian models so we included them in our new models in forward selection for models 2 and 3. This found xgboost to preform well on many variables which we included in our model 3 in particular because of their high accuracy score, low multicolinearity because of a low condition number near 20. This was because in our confidence interval and hypothesis testing, we found that the absolute difference between the accuracy scores of xgboost and ffnn/transformer was always positive meaning that it always had a higher accuracy scores. As such it and its variables associated with it became a major part of the predictive power of our model 3 because of their high accuracy scores.

In order accurately maintain the correctless of proper statstical testing and analysis we utilized random sampling to preform forward selection for model 2 and 3. This is because in order to preform forward selection we must start by narrowing down to a smaller subset of the variables in order to ensure that the variables used have enough observations. As such we limited our sample to varibales that had >14 observations. Since we're using indicator variables, some of the situations being considered might not have actually happened very frequently in our data. In this case, we're trying to make predictions on the basis of very little data to actually inform these predicitons.

This culminated in us producing model 3 which combined a low condition number and a low difference in RMSE while also balancing a low level of overfitting to ensure that the model had maximum predictive power. This is not to say that the performanceof model three is entirely due to xgboost and its variables, as other variables and machine learning models were included so that it can be applied to other sets of data unseen to the model. Its performance in this case was echoed in its train test which showed good performance in comparison to our model 1 and 2. This is furhter suppported by our heatmap that indicates that suggests little correlation between variables which implies having low multicollinearity.

This allows us to conclude that our model 3 and its expected value formula is the best multiple linear regression model to predict occurances of error. This also tells us that in general an effective expected value formula shoould include values from xgboost in particular as well as other categories/models to ensure there is an appropriate balance between multicolinearity and RMSE.

The signficance of our model 3 is that there is enough statistical evidence to support that it can be used to accruately predict conflict as part of the enitiatives by UNICEF. Accurate predictions are important because these predicitons are what can alow organizations and nations to be proactive in the face of potentially dangerous conflicts and save civillian lives. At the same time, we must also consider the probablity of error as these models are not perfect, even model 3, as indicated by its RMSE. But, we can still be confident in its performance as a perfect model would be highly susceptible to overfitting which would cause futher complications. Therefore, our model can be counted on by UNICEF.