Joshua Chang
Professor Coggeshall
DSO 562
February 9, 2023

Homework 5: Build Preliminary Models

Model	Parameters							Number of Variables			
Logistic Regression	**Number of Variables**	**max_iter**						**Train**	**Test**	**OOT**	**Fit**
1	5	20						0.477	0.484	0.466	underfit
2	10	20						0.487	0.491	0.474	
3	15	20						0.484	0.476	0.467	
Decision Tree	**Number of Variables**	**max_depth**	**min_samples_leaf**	**min_samples_split**	**max_features**			**Train**	**Test**	**OOT**	**Fit**
1	10	2	500	1000	none			0.461	0.461	0.444	underfit
2	10	50	250	500	none			0.531	0.523	0.503	
3	10	50	300	550	none			0.526	0.521	0.497	
4	10	100	2	5	none			0.543	0.512	0.495	overfit
5	15	50	250	500	5			0.523	0.514	0.496	
6	15	50	300	550	5			0.520	0.518	0.496	
7	20	50	250	500	8			0.526	0.521	0.498	
8	20	100	300	700	10			0.526	0.519	0.498	
Random Forest	**Number of Variables**	**max_depth**	**min_samples_leaf**	**min_samples_split**	**max_features**	**n_estimator**		**Train**	**Test**	**OOT**	**Fit**
1	10	2	500	1000	8	3		0.473	0.473	0.459	underfit
2	10	5	30	50	3	5		0.519	0.517	0.495	
3	10	10	25	45	6	10		0.529	0.527	0.505	
4	10	15	20	40	10	15		0.535	0.525	0.502	
5	15	10	25	45	6	10		0.528	0.525	0.503	
6	15	30	10	30	10	100		0.542	0.523	0.501	
7	20	10	25	45	6	10		0.530	0.523	0.505	
8	20	30	5	50	10	100		0.544	0.515	0.501	overfit
LGBM	**Number of Variables**	**n_estimators**	**max_depth**	**num_leaves**				**Train**	**Test**	**OOT**	**Fit**
1	10	2	2	2				0.464	0.452	0.444	underfit
3	10	20	2	2				0.512	0.513	0.489	
4	10	50	3	4				0.516	0.518	0.494	
5	10	50	6	10				0.528	0.527	0.504	
6	15	300	4	5				0.526	0.532	0.507	
7	15	500	6	10				0.533	0.520	0.507	
8	20	100	4	8				0.528	0.526	0.503	
9	20	75	6	10				0.527	0.528	0.505	
10	20	500	100	50				0.536	0.516	0.505	overfit
Neural Networks	**Number of Variables**	**hidden_layer_sizes**	**activation**	**alpha**	**learning_rate**	**solver**	**learning_rate_init**	**Train**	**Test**	**OOT**	**Fit**
1	10	(5)	logistic	0.1	constant	adam	0.01	0.494	0.496	0.478	underfit
3	10	(20,20,20)	relu	0.01	adaptive	lbfgs	0.01	**0.528**	**0.527**	0.505	
4	15	(5)	relu	0.01	adaptive	lbfgs	0.01	0.521	0.523	0.500	
5	15	(10,10)	relu	0.1	adaptive	lbfgs	0.0001	0.526	0.529	0.505	
6	20	(10,10)	logistic	0.01	adaptive	lbfgs	0.0001	0.516	0.516	0.496	
7	20	(20,20,20)	relu	0.01	constant	lbfgs	0.01	0.529	0.524	0.507	
GBC	**Number of Variables**	**n_estimators**	**max_depth**					**Train**	**Test**	**OOT**	**Fit**
1	10	2	2					0.487	0.483	0.469	underfit
2	10	10	4					0.524	0.513	0.499	
3	10	20	6					0.529	0.524	0.505	
4	10	1000	6					0.544	0.513	0.497	overfit
5	15	30	4					0.523	0.523	0.502	
6	15	50	6					0.527	0.529	0.504	
7	15	100	6					0.532	0.525	0.507	
8	20	10	6					0.526	0.520	0.501	
9	20	40	6					0.526	0.531	0.503	
XGB	**Number of Variables**	**n_estimators**	**max_depth**					**Train**	**Test**	**OOT**	**Fit**
1	10	2	2					0.492	0.490	0.474	underfit
2	10	10	4					0.519	0.519	0.497	
3	10	20	6					0.529	0.525	0.507	
4	15	30	4					0.528	0.525	0.505	
5	15	60	6					0.534	0.524	0.505	
6	20	50	6					0.531	0.528	0.506	
7	20	2000	5					0.544	0.518	0.498	overfit