

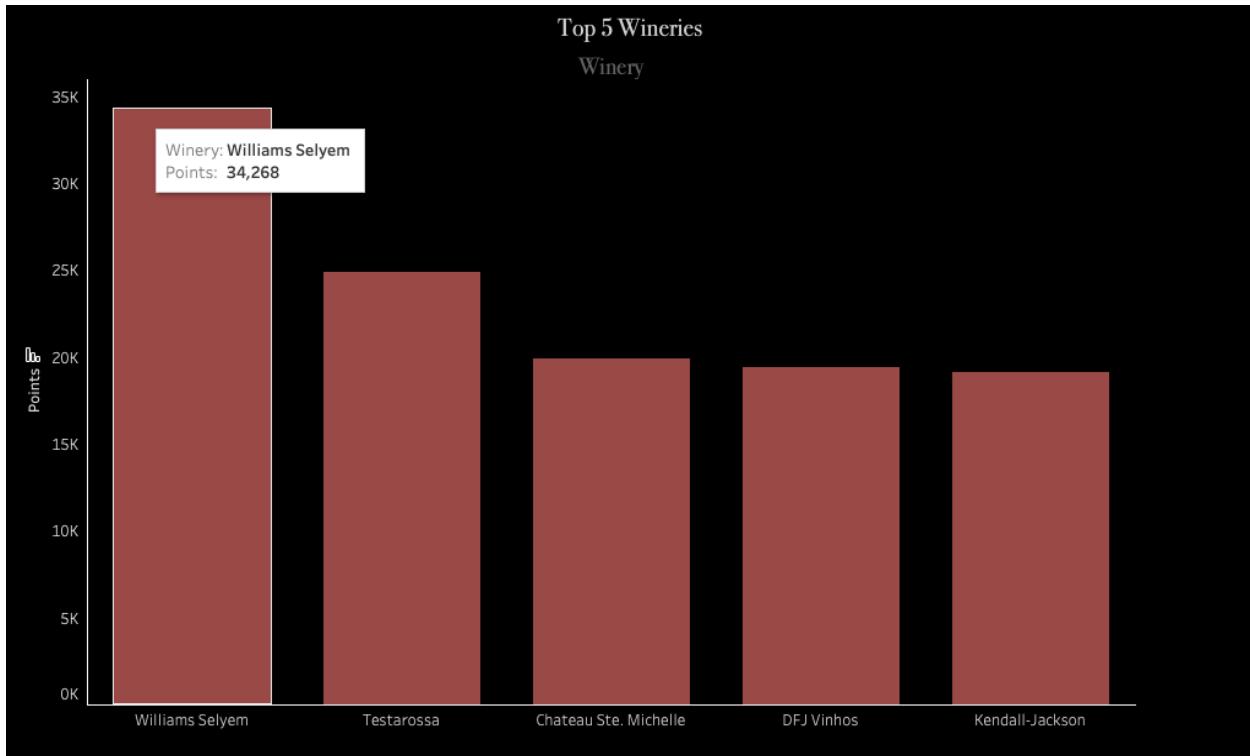
# The Wine Price is Right

## Appendix

## Table of Contents

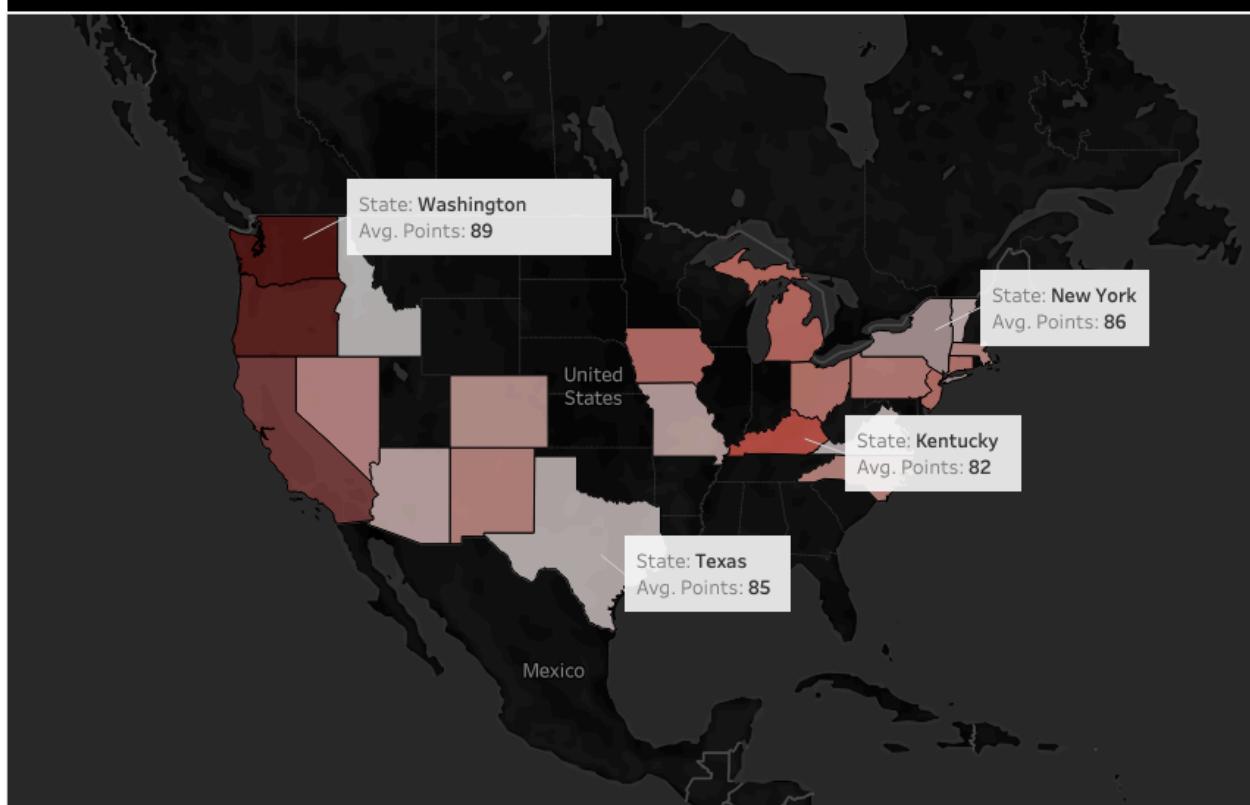
<b>Data Exploration</b> .....	<b>3</b>
<b>Machine Learning</b> .....	<b>5</b>
Round One:.....	5
Logistic Regression.....	5
KNN Model.....	6
Decision Tree Model .....	8
Bagging Classifier Model.....	9
Ada Boost Classifier .....	10
Gradient Boosting Classifier.....	11
Random Forest Classifier .....	12
Tuning .....	13
Neural Network .....	14
Round Two.....	15
Logistic Regression.....	15
KNN Model.....	16
Decision Tree Model .....	18
Bagging Classifier Model.....	19
Ada Boost Classifier .....	20
Gradient Boost Classifier.....	21
Random Forest Classifier .....	22
Tuning .....	23
Neural Network .....	24
Round Three .....	25
Logistic Regression.....	25
KNN Model.....	25
Ada Boost Classifier .....	28
Random Forest Classifier .....	29
Bagging Classifier .....	30
Tuning .....	30
Round Four .....	33
Logistic Regression.....	33
KNN Model.....	34
Decision Tree Classifier .....	35
Bagging Classifier .....	36
Ada Boost Classifier .....	38
Random Forest Classifier .....	39
Gradient Boosting Classifier.....	40
Tuning .....	41

## Data Exploration

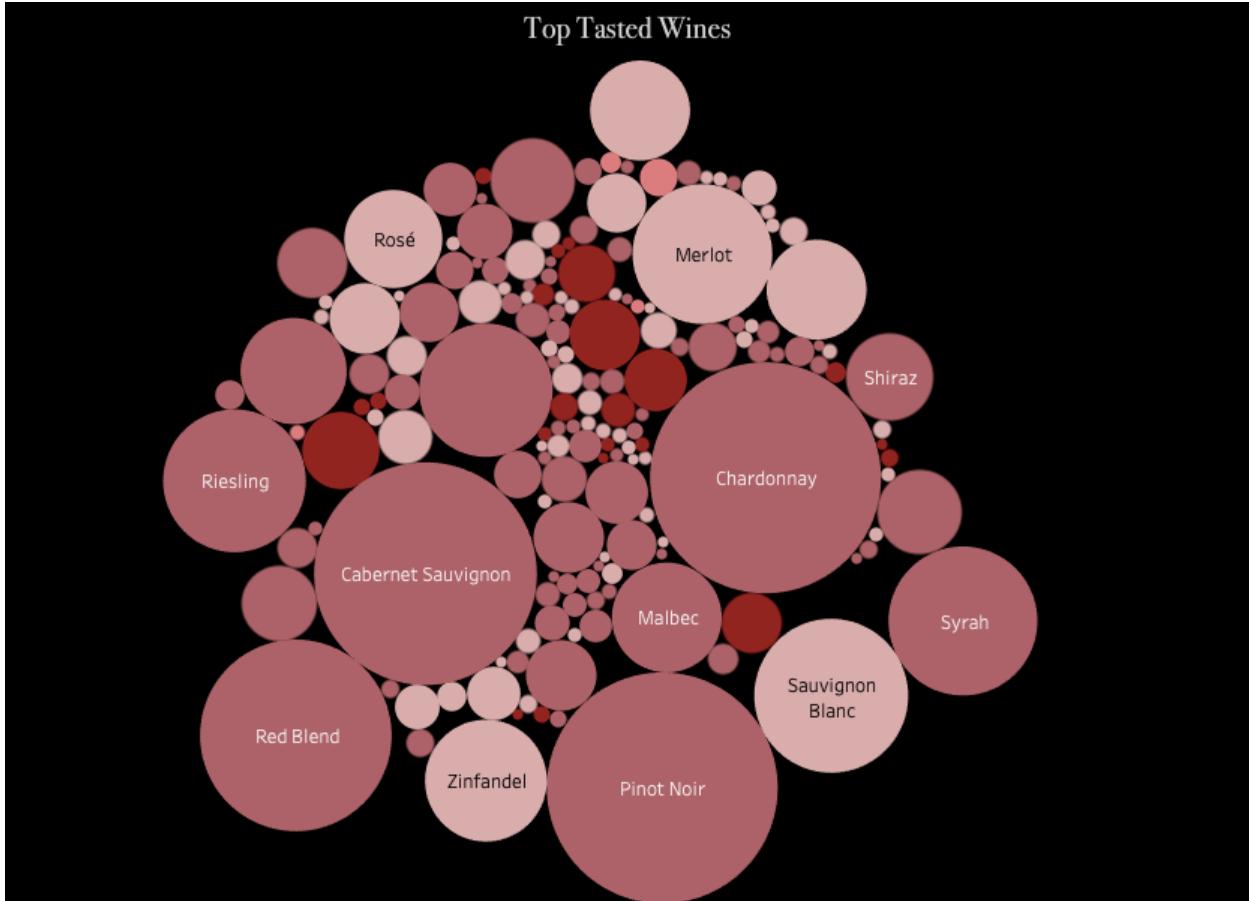


Top 5 Wineries classified by the total amount of points received.

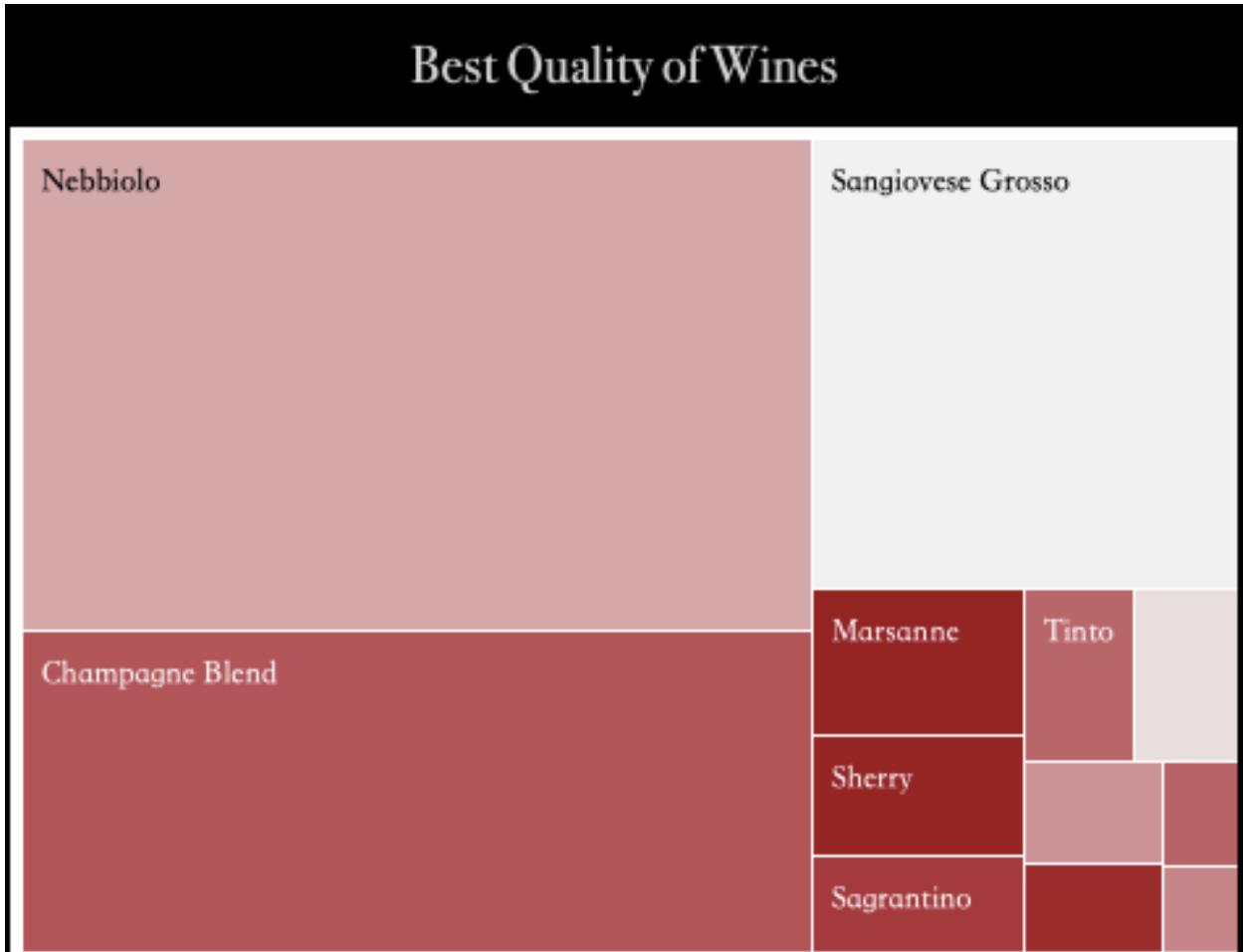
Where is the best wine in the US?



Map of which state has the highest quality of wine.

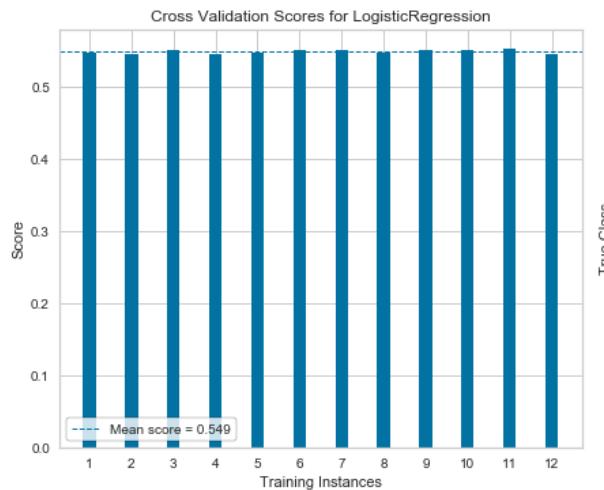


Bubble chart of the most tasted wines from Wine Enthusiast.



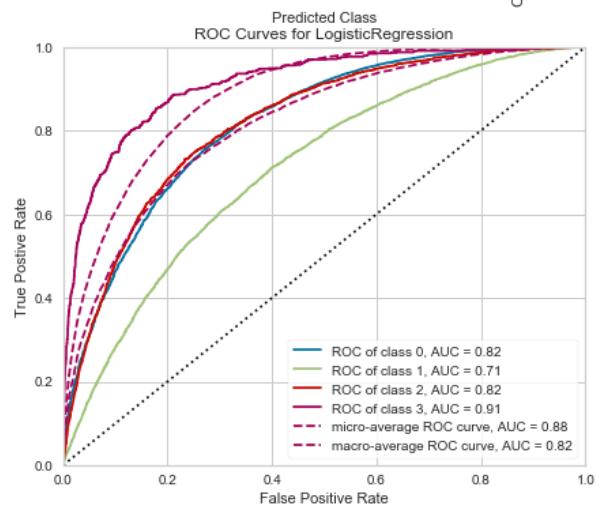
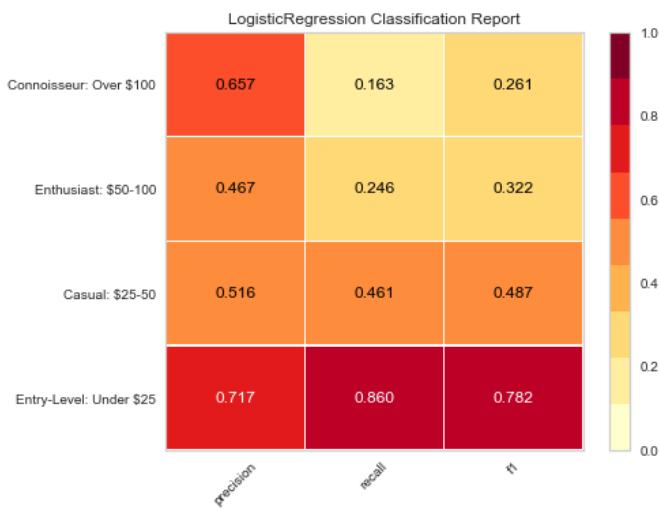
Graph of the best quality of wines by average points.

## Machine Learning Round One: *Logistic Regression*

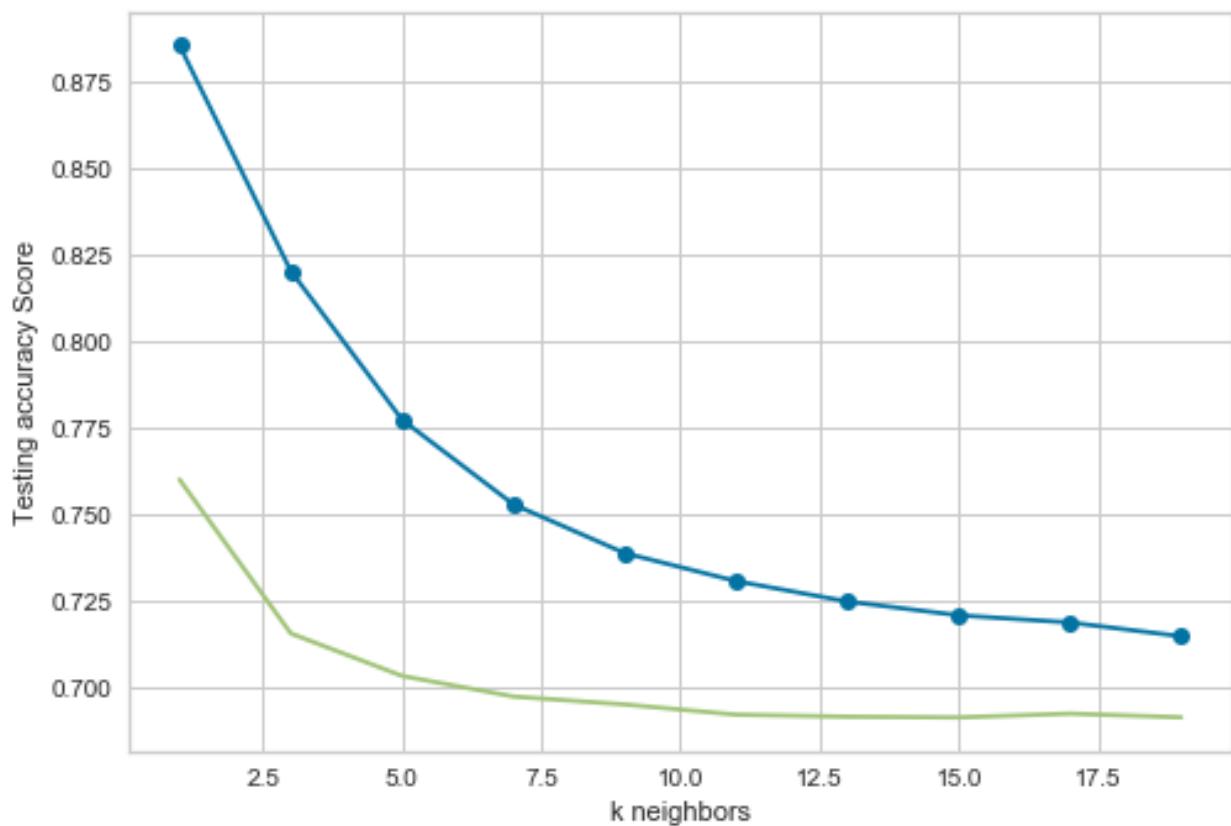


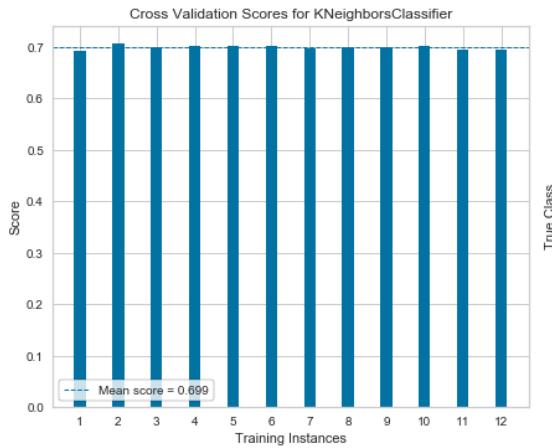
**LogisticRegression Confusion Matrix**

		True Class			
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Predicted Class	Entry-Level: Under \$25	15711	2430	127	0
	Casual: \$25-50	5051	4838	590	15
Enthusiast: \$50-100	1038	1886	971	58	
Connoisseur: Over \$100	103	222	393	140	



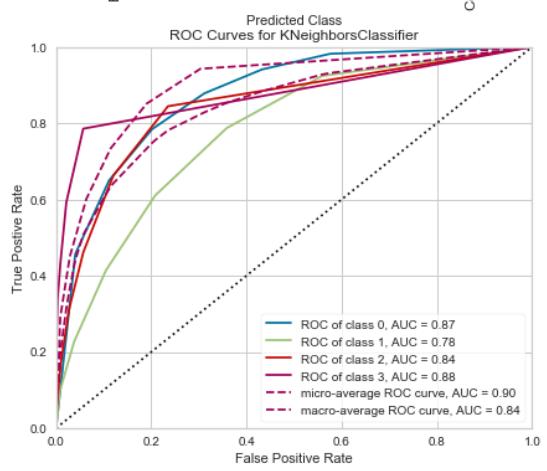
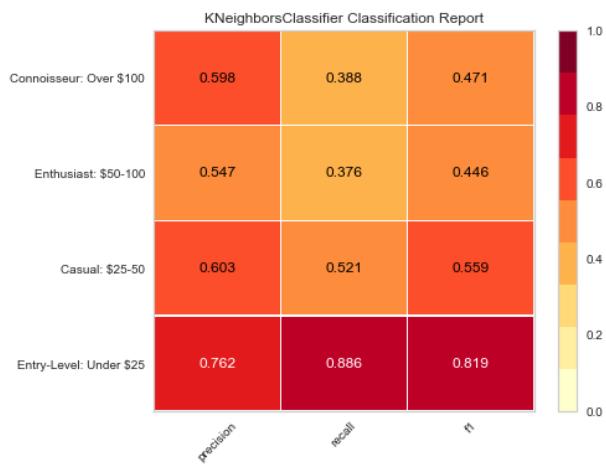
*KNN Model*



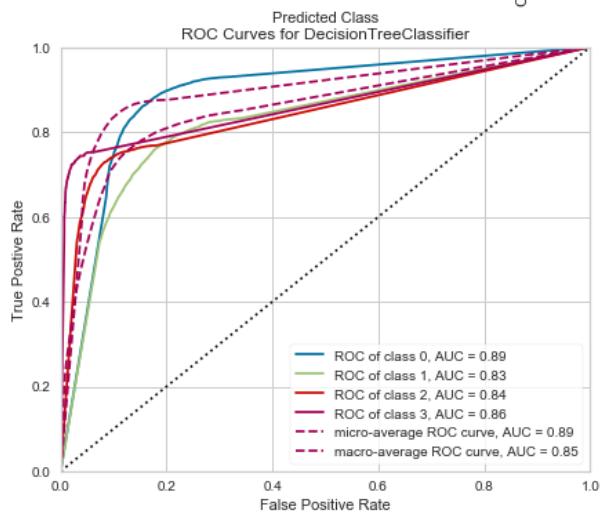
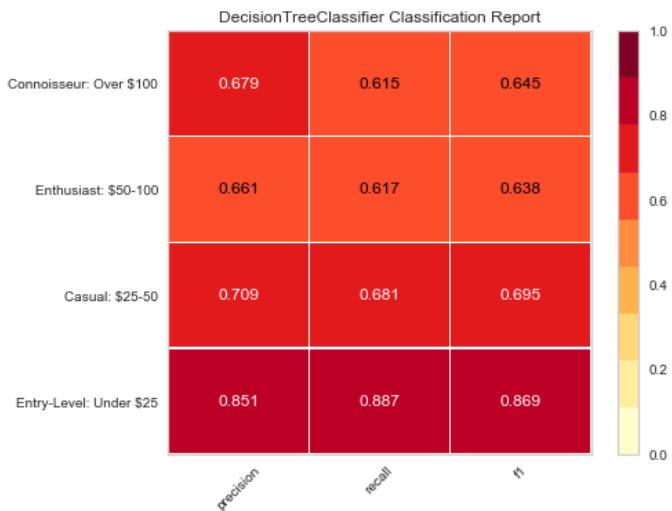
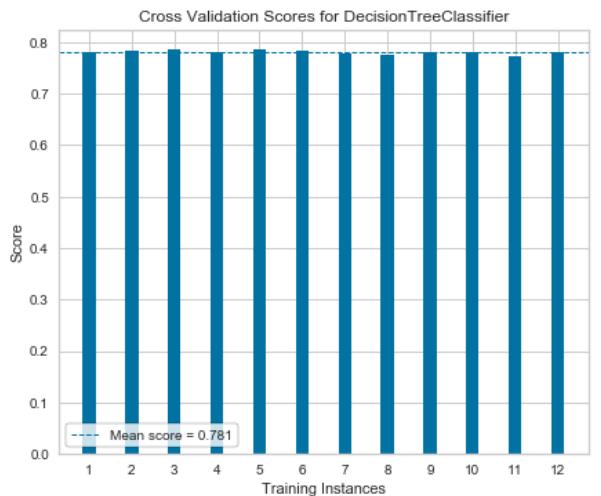


KNeighborsClassifier Confusion Matrix

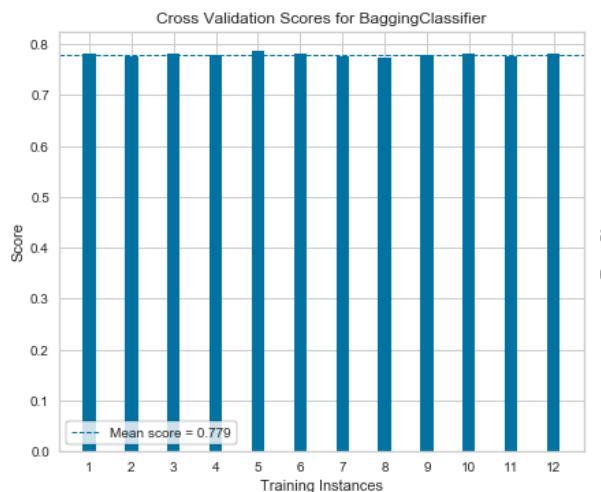
Entry-Level: Under \$25	16182	1867	199	20
Casual: \$25-50	4197	5468	762	67
Enthusiast: \$50-100	777	1553	1486	137
Connoisseur: Over \$100	79	176	270	333



## Decision Tree Model

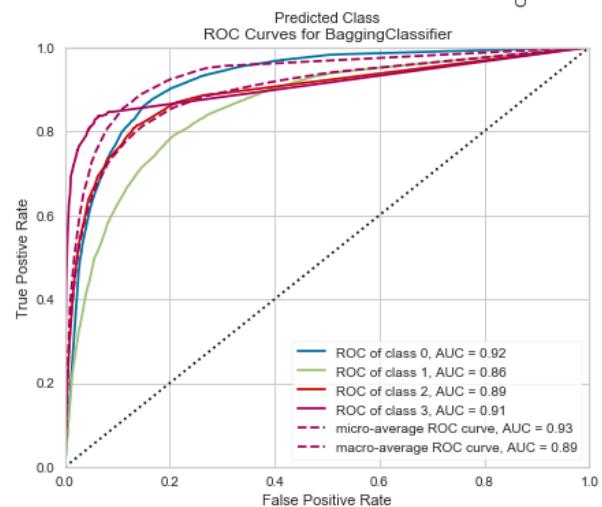
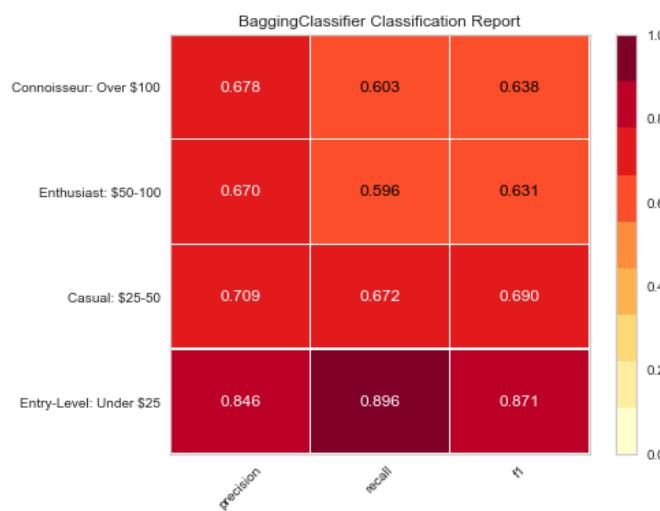


## Bagging Classifier Model

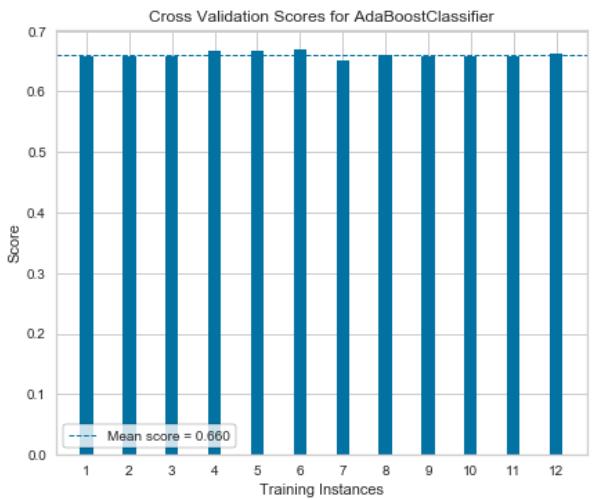


BaggingClassifier Confusion Matrix

	Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Entry-Level: Under \$25	16375	1694	182	17
Casual: \$25-50	2567	7052	781	94
Enthusiast: \$50-100	372	1091	2355	135
Connoisseur: Over \$100	35	110	196	517

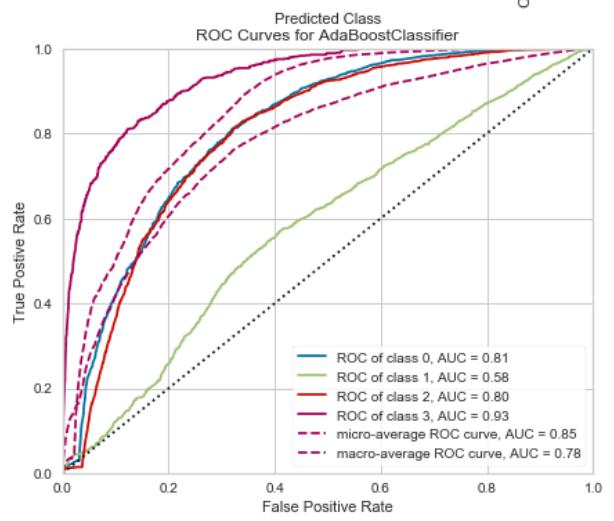
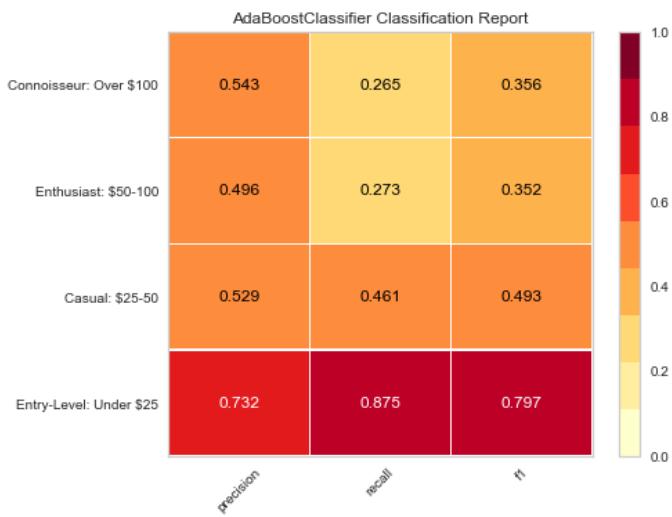


## Ada Boost Classifier

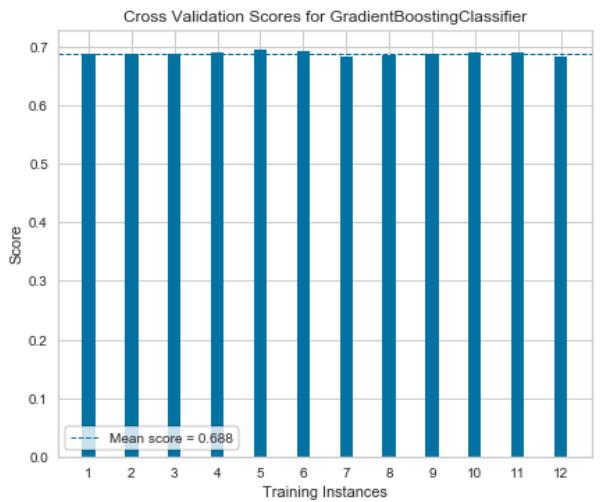


**AdaBoostClassifier Confusion Matrix**

		True Class			
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Predicted Class	Entry-Level: Under \$25	15979	2212	68	9
	Casual: \$25-50	4946	4842	663	43
Enthusiast: \$50-100	828	1906	1080	139	
Connoisseur: Over \$100	68	195	368	227	

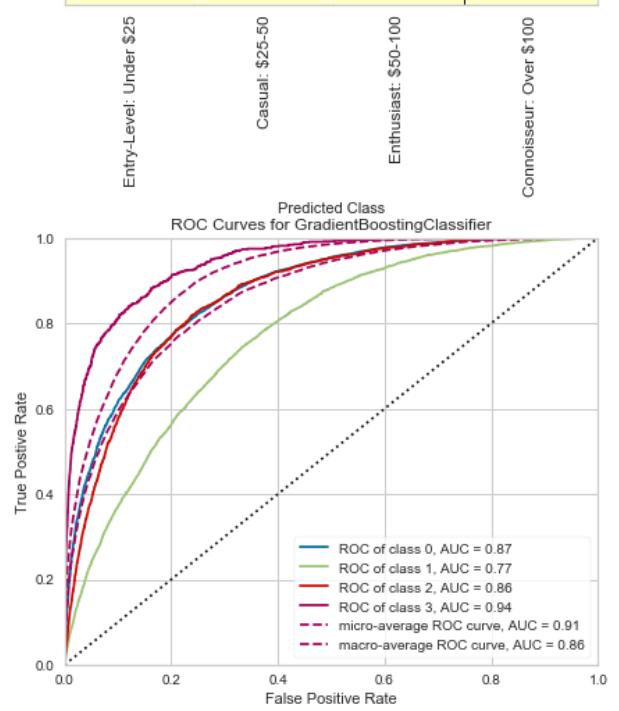
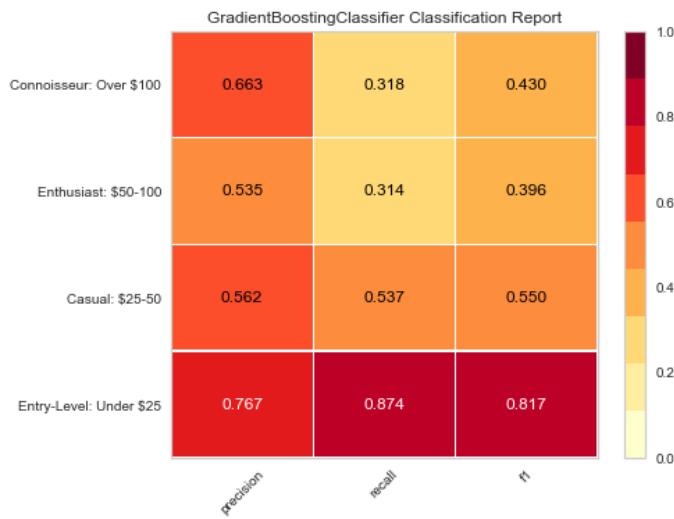


## Gradient Boosting Classifier

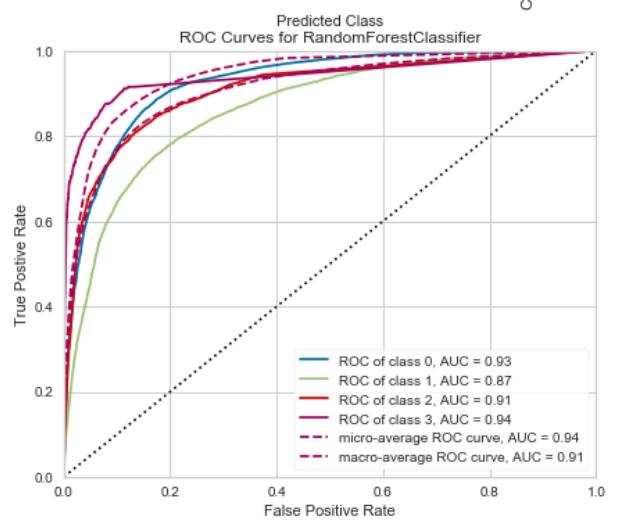
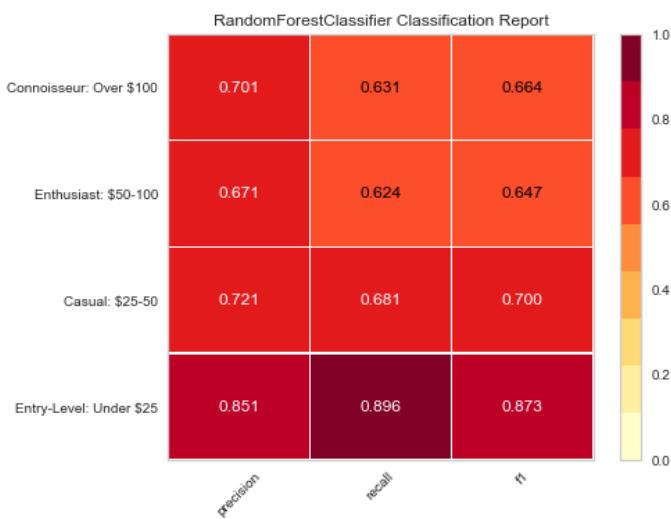
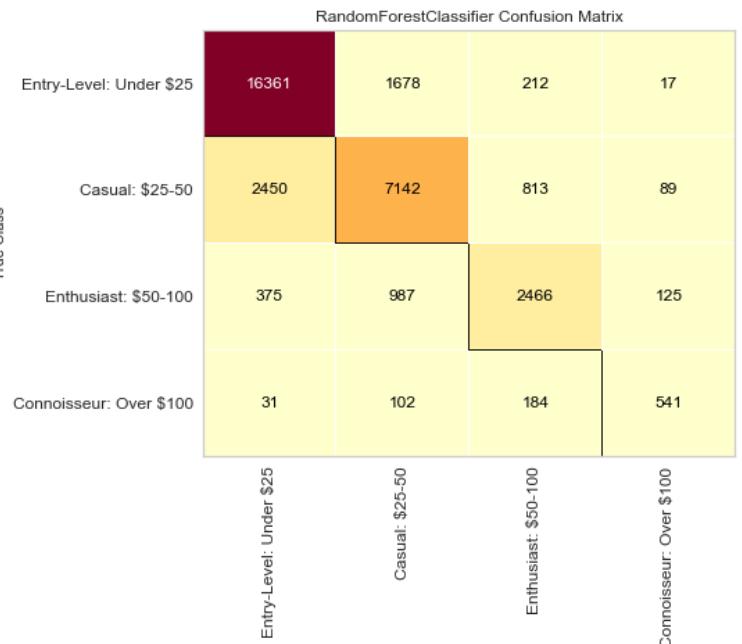
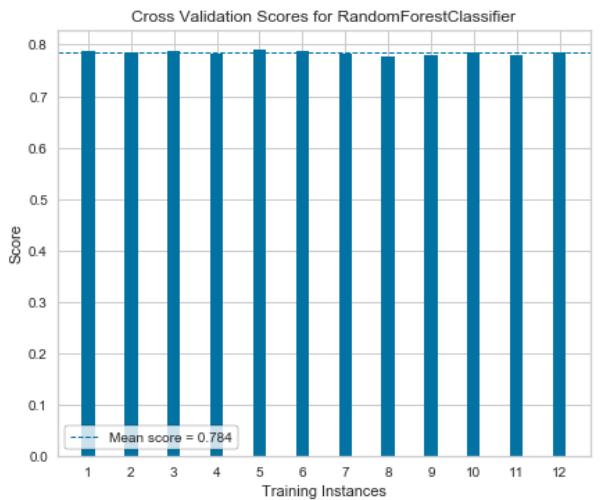


GradientBoostingClassifier Confusion Matrix

Entry-Level: Under \$25	15960	2241	65	2
Casual: \$25-50	4152	5638	671	33
Enthusiast: \$50-100	640	1968	1241	104
Connoisseur: Over \$100	65	178	342	273

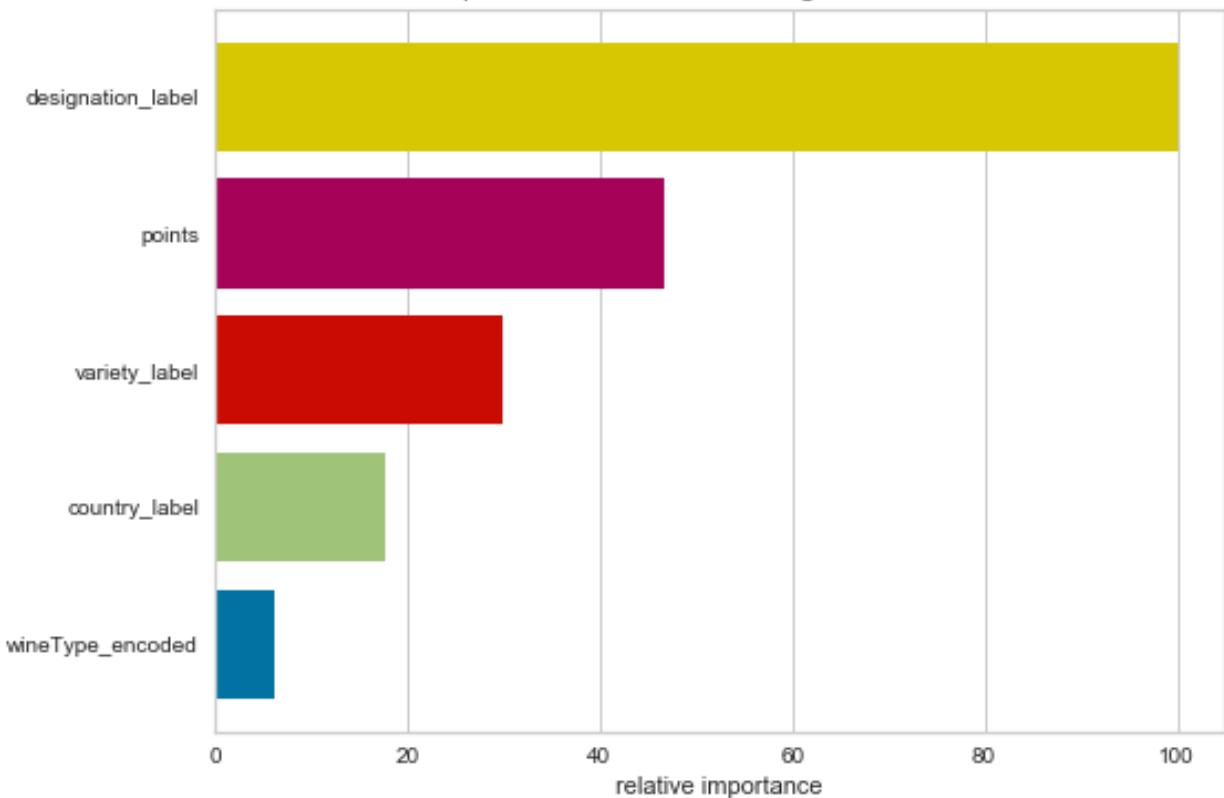


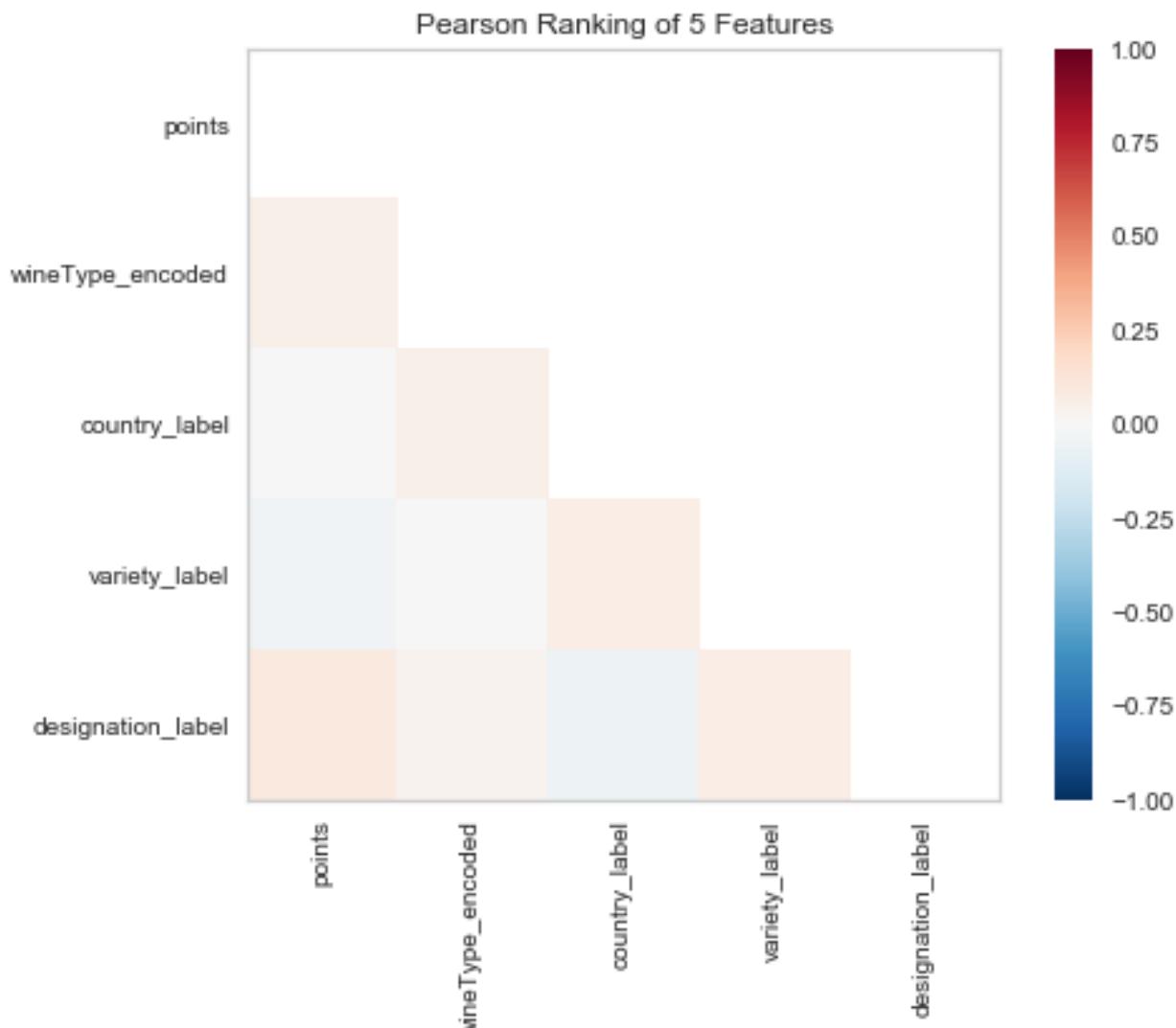
## Random Forest Classifier



*Tuning*

Feature Importances of 5 Features using RandomForestClassifier





### Neural Network

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
<hr/>		
dense_2 (Dense)	(None, 6)	36
<hr/>		
dense_3 (Dense)	(None, 4)	28
<hr/>		
Total params:	64	
Trainable params:	64	
Non-trainable params:	0	

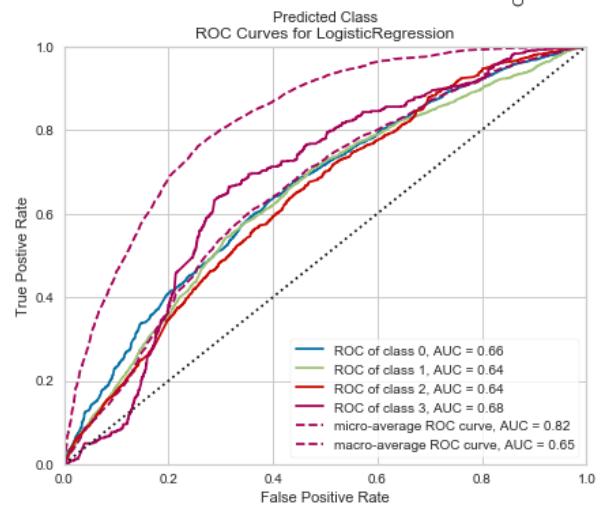
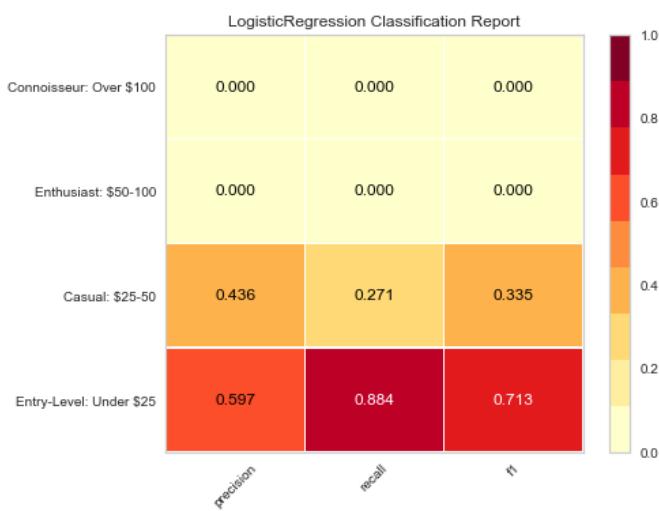
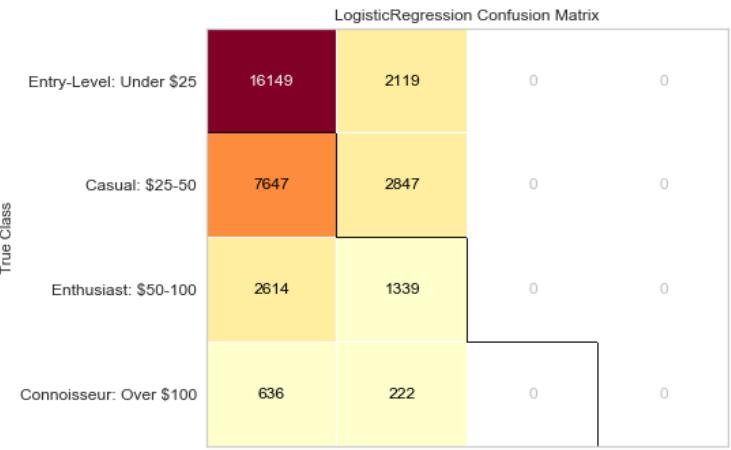
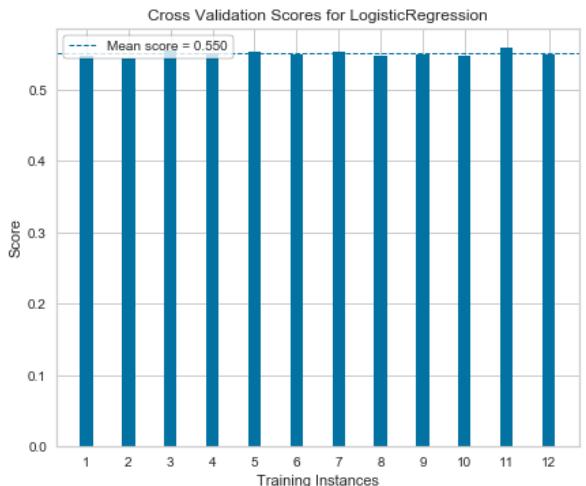
---

Epoch 47/100

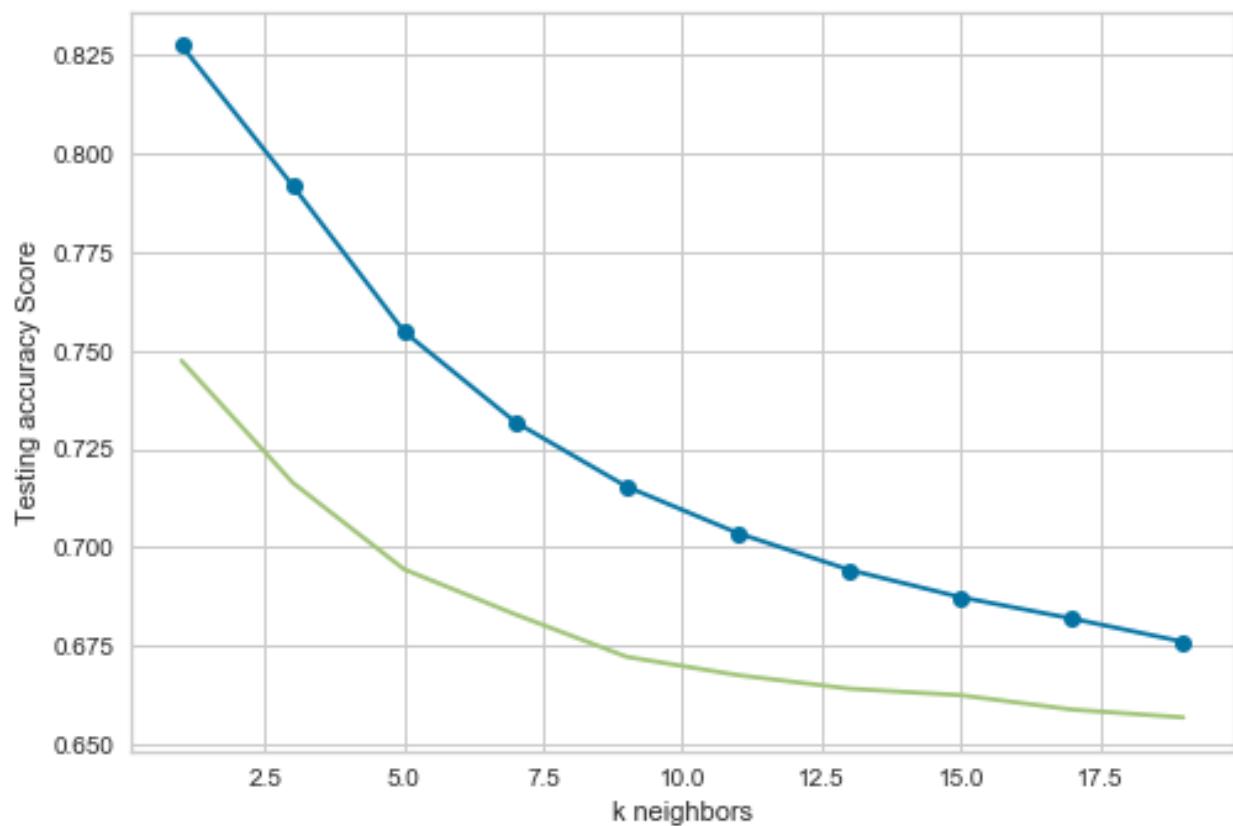
3148/3148 - 3s - loss: 0.7601 - accuracy: 0.6658

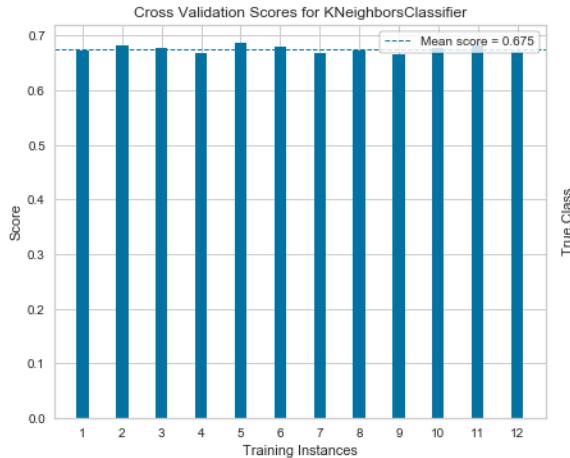
## Round Two

### Logistic Regression



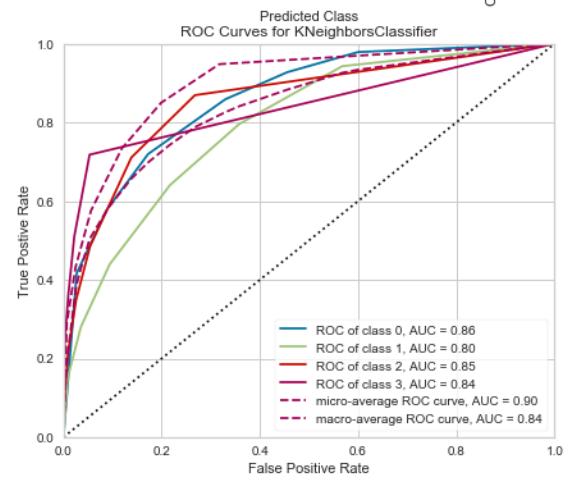
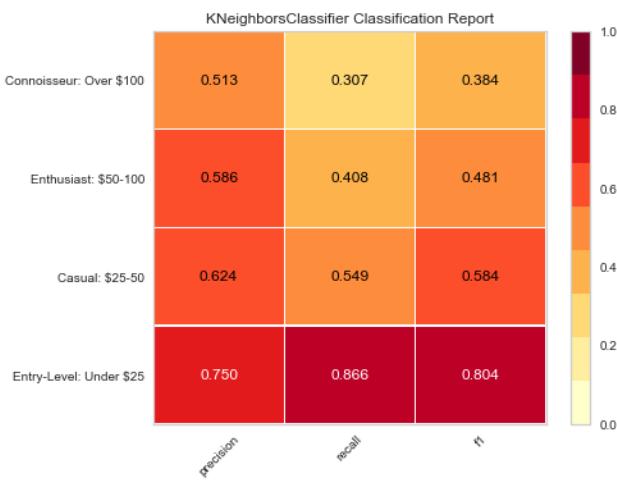
*KNN Model*



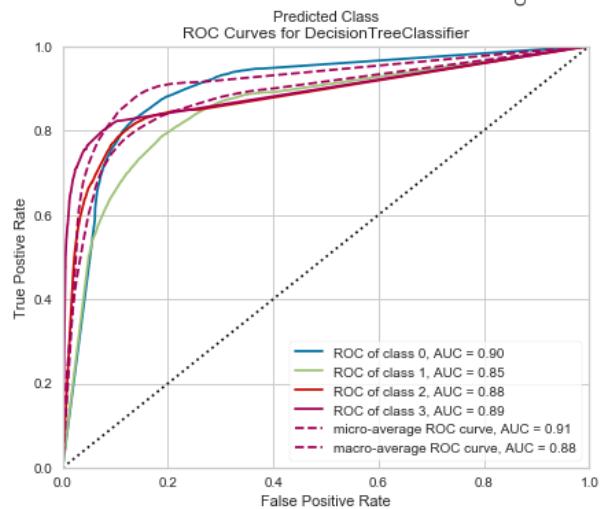
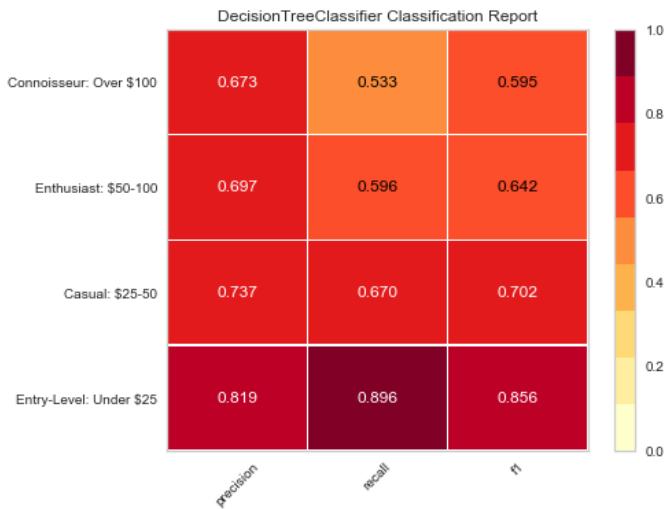
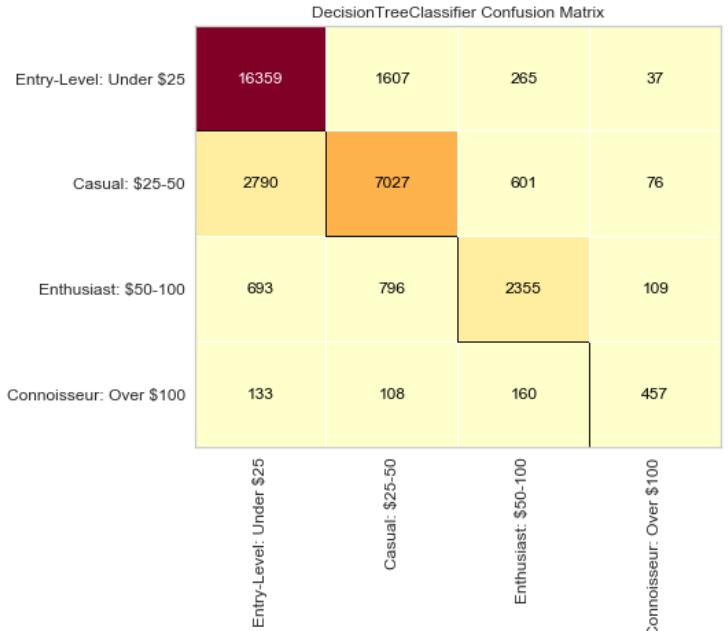
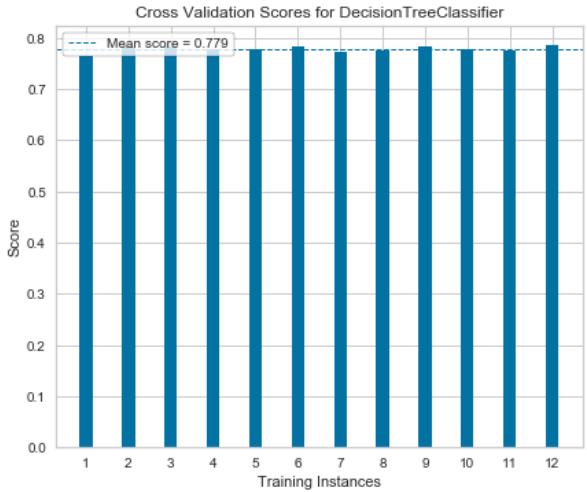


KNeighborsClassifier Confusion Matrix

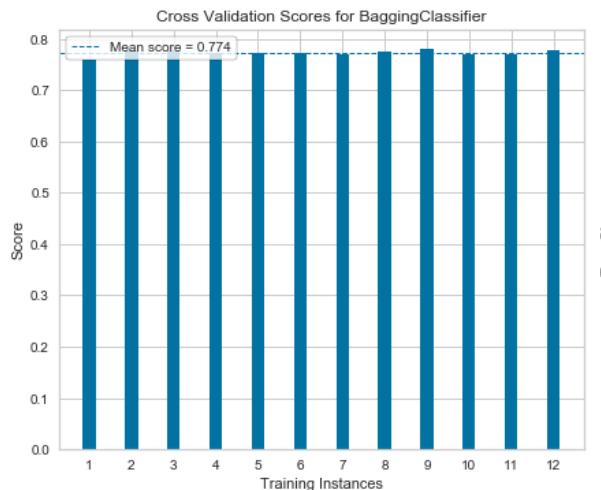
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100	
True Class	Predicted Class	15819	2055	315	79	
		3999	5763	648	84	
Entry-Level: Under \$25	Entry-Level: Under \$25	15819	2055	315	79	
Casual: \$25-50	Casual: \$25-50	3999	5763	648	84	
Enthusiast: \$50-100	Enthusiast: \$50-100	1081	1174	1611	87	
Connoisseur: Over \$100	Connoisseur: Over \$100	182	237	176	263	



## Decision Tree Model

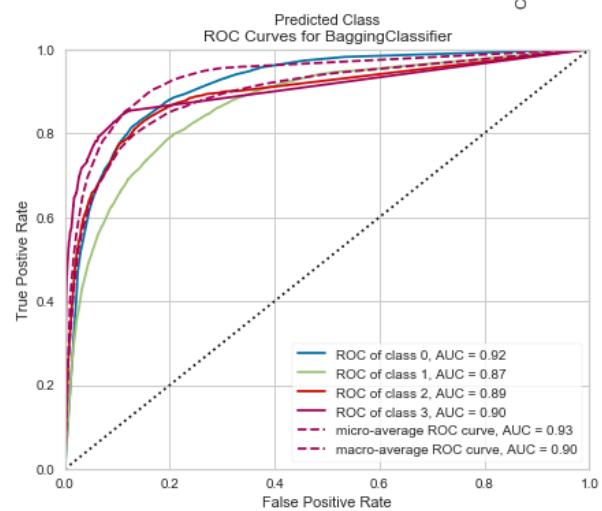
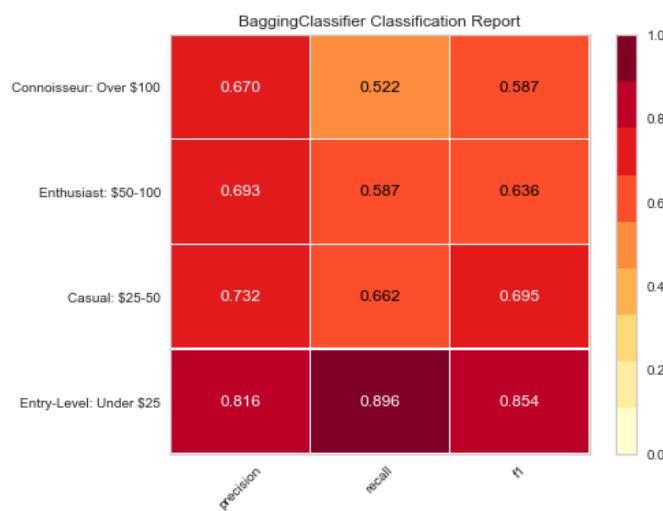


## Bagging Classifier Model

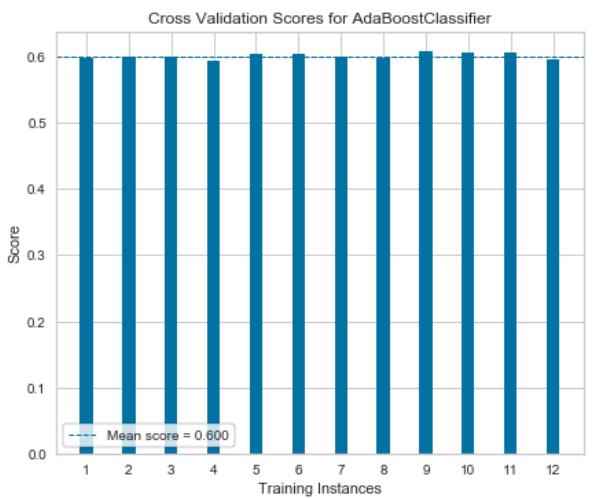


**BaggingClassifier Confusion Matrix**

		True Class			
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Predicted Class	Entry-Level: Under \$25	16368	1618	247	35
	Casual: \$25-50	2834	6951	625	84
Enthusiast: \$50-100	Enthusiast: \$50-100	720	811	2320	102
	Connoisseur: Over \$100	136	119	155	448

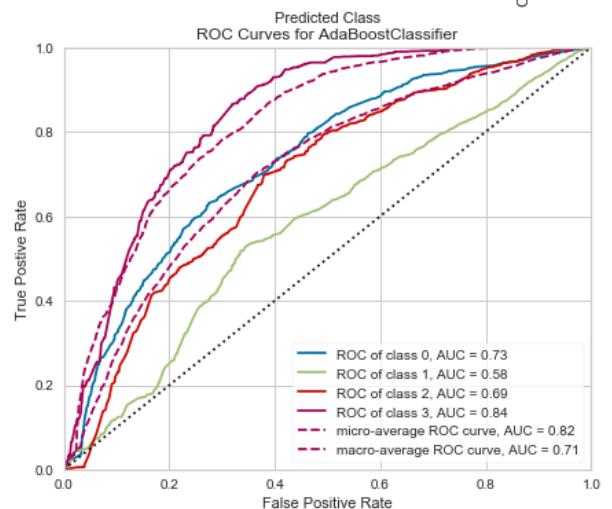
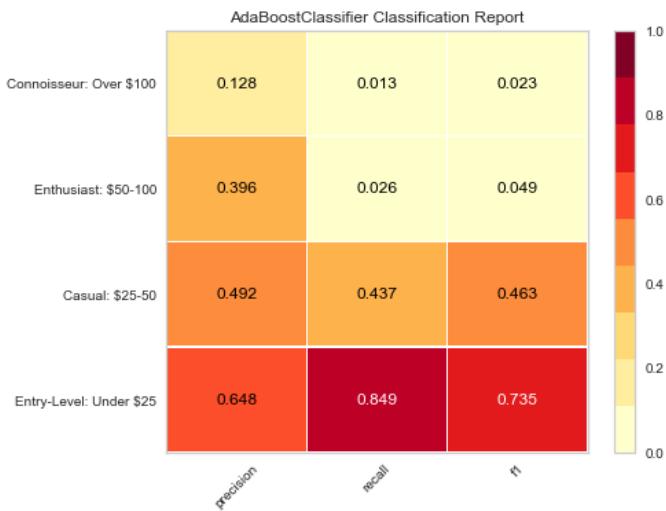


## Ada Boost Classifier

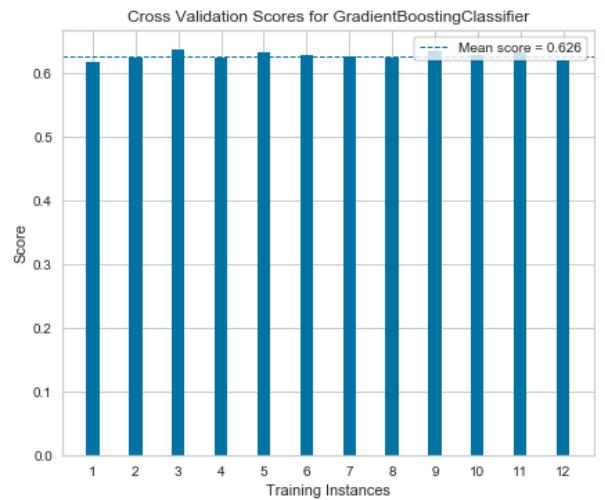


AdaBoostClassifier Confusion Matrix

		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100	
True Class	Predicted Class	Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100	
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100	
Entry-Level: Under \$25	Entry-Level: Under \$25	15504	2694	41	29	
Casual: \$25-50	Casual: \$25-50	5818	4584	73	19	
Enthusiast: \$50-100	Enthusiast: \$50-100	2074	1749	103	27	
Connoisseur: Over \$100	Connoisseur: Over \$100	515	289	43	11	

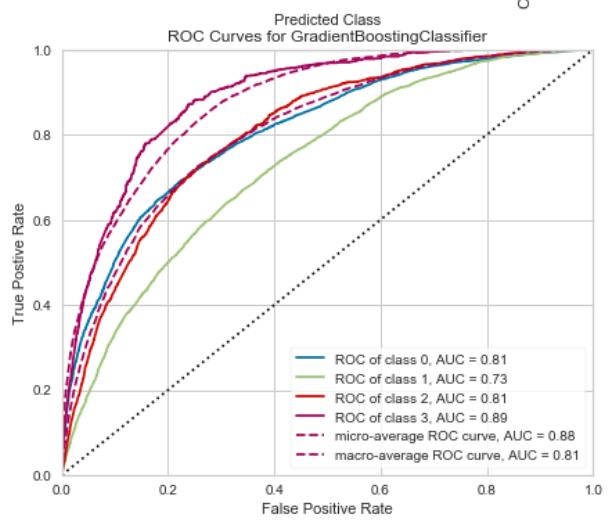
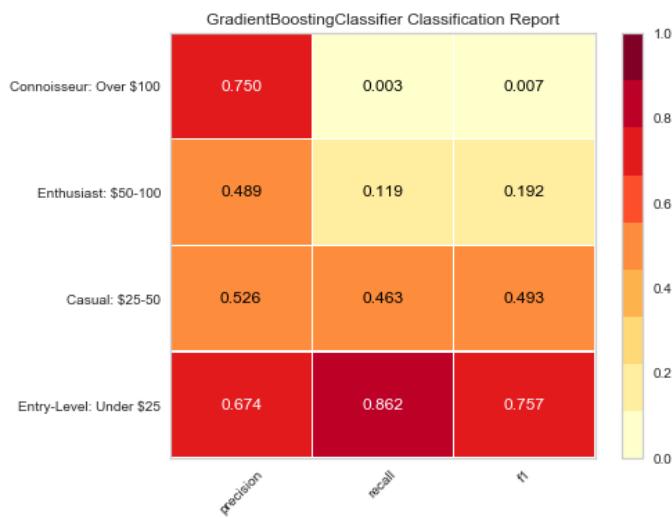


## Gradient Boost Classifier

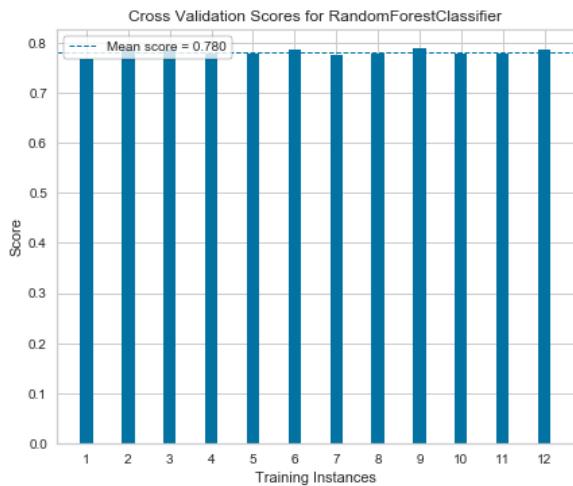


**GradientBoostingClassifier Confusion Matrix**

		True Class			
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Predicted Class	Entry-Level: Under \$25	15745	2449	74	0
	Casual: \$25-50	5363	4863	268	0
Enthusiast: \$50-100	1783	1697	472	1	
Connoisseur: Over \$100	464	239	152	3	

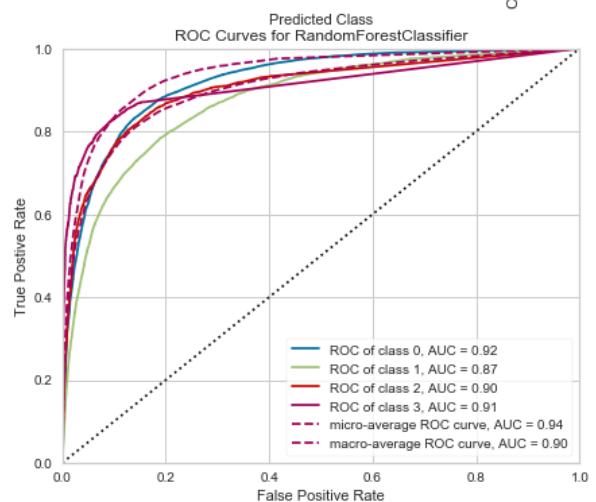
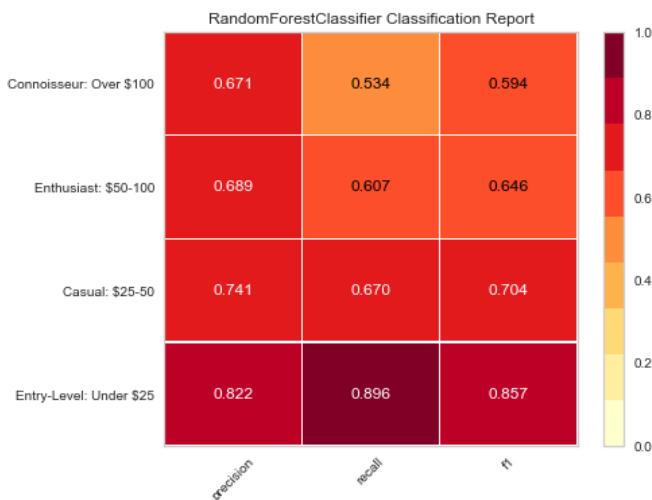


## Random Forest Classifier



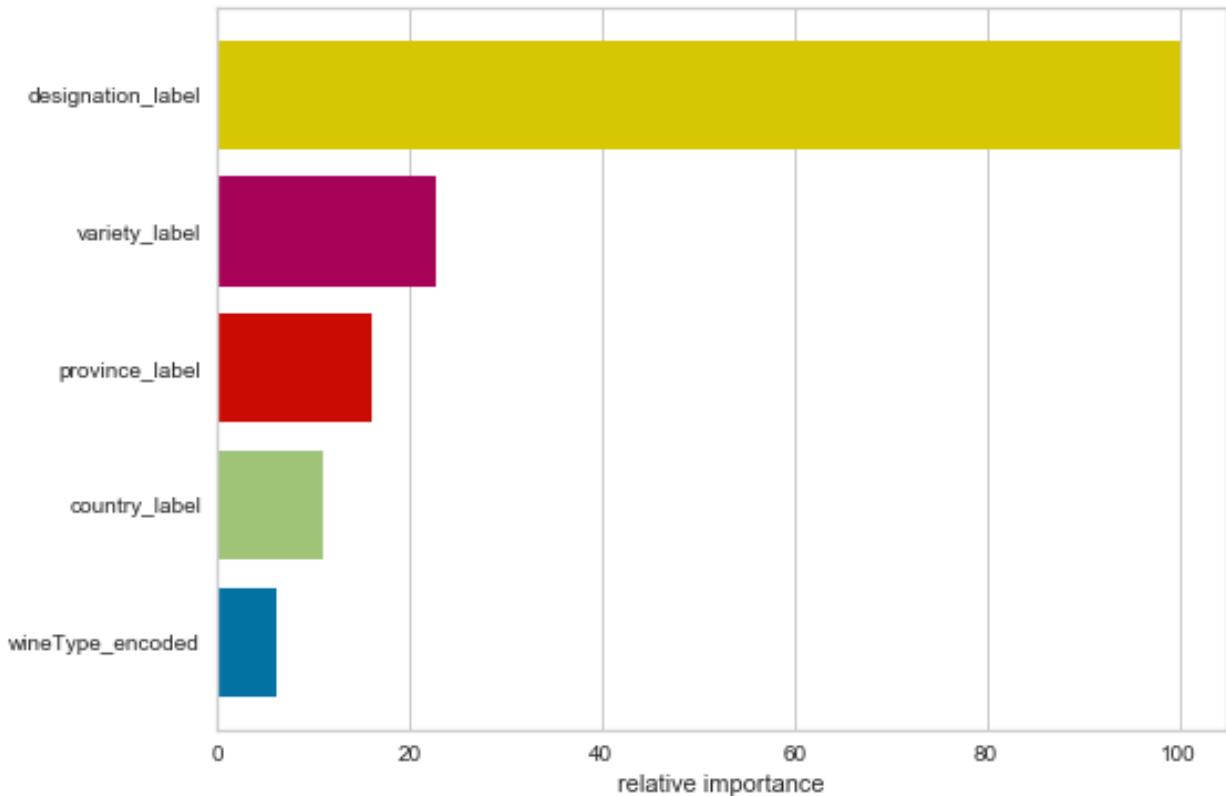
RandomForestClassifier Confusion Matrix

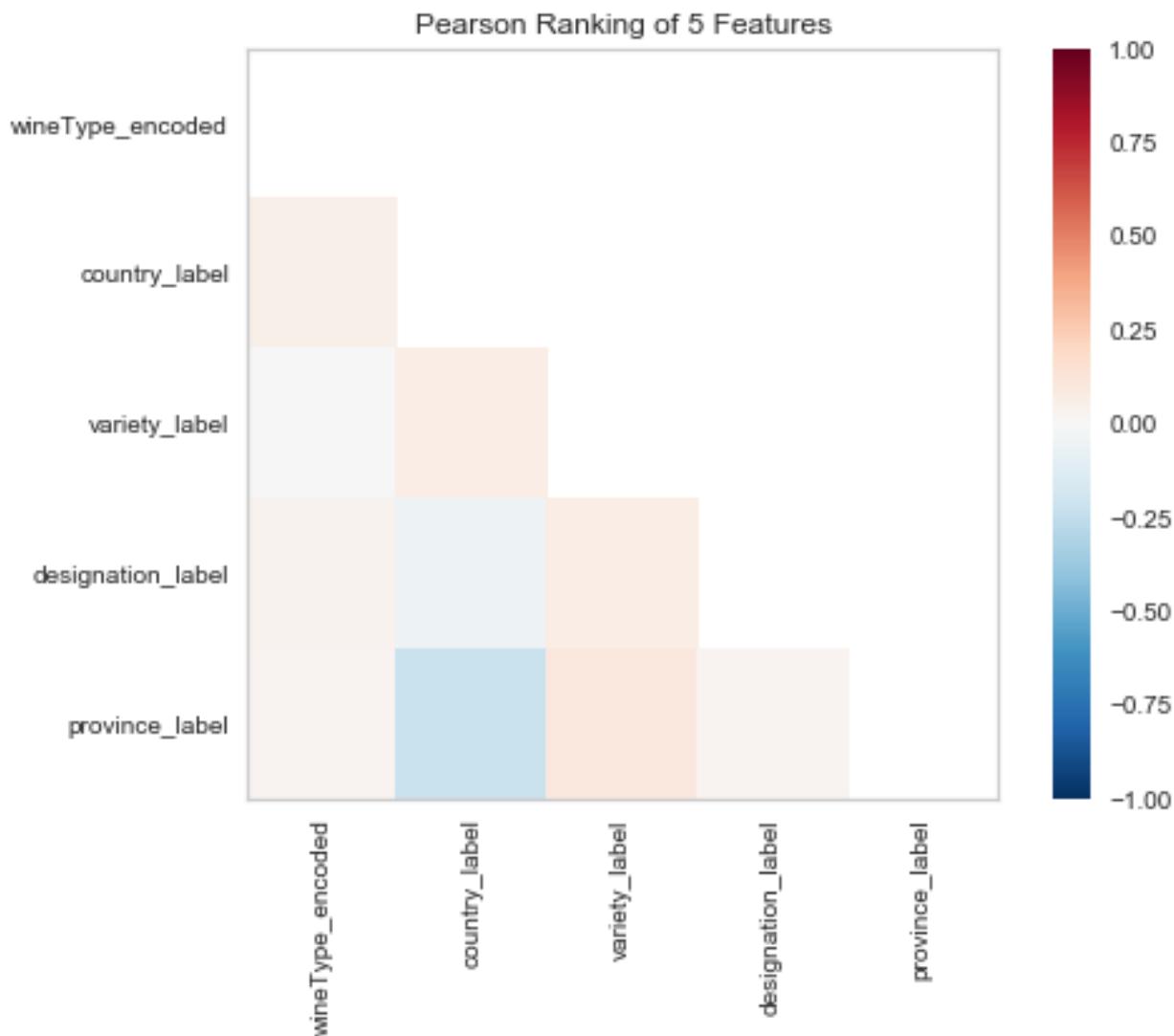
		True Class			
		Entry-Level: Under \$25	Casual: \$25-50	Enthusiast: \$50-100	Connoisseur: Over \$100
Predicted Class	Entry-Level: Under \$25	16363	1604	269	32
	Casual: \$25-50	2728	7035	653	78
Enthusiast: \$50-100	693	744	2401	115	
Connoisseur: Over \$100	132	107	161	458	



*Tuning*

Feature Importances of 5 Features using RandomForestClassifier





### Neural Network

Model: "sequential"

Layer (type)	Output Shape	Param #
<hr/>		
dense (Dense)	(None, 6)	36
<hr/>		
dense_1 (Dense)	(None, 4)	28
<hr/>		
Total params:	64	
Trainable params:	64	
Non-trainable params:	0	

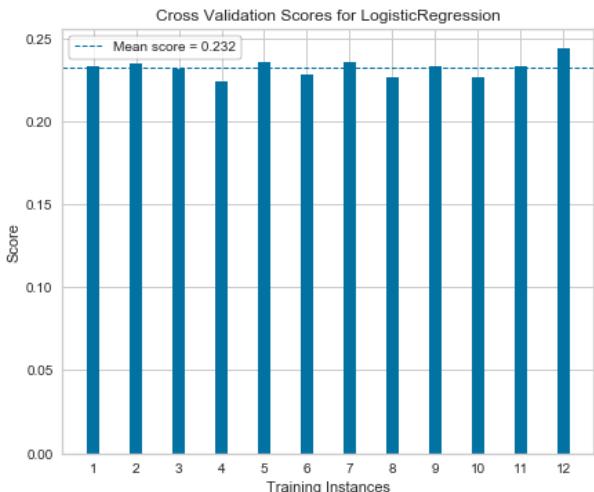
---

Epoch 48/100

3148/3148 - 3s - loss: 0.9217 - accuracy: 0.5894

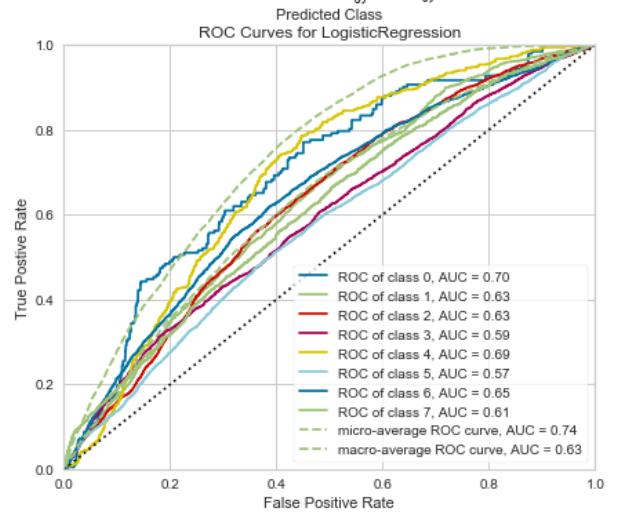
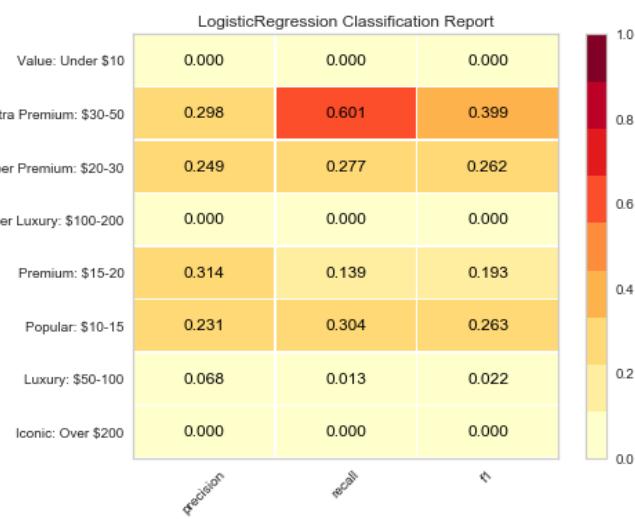
### Round Three

#### Logistic Regression

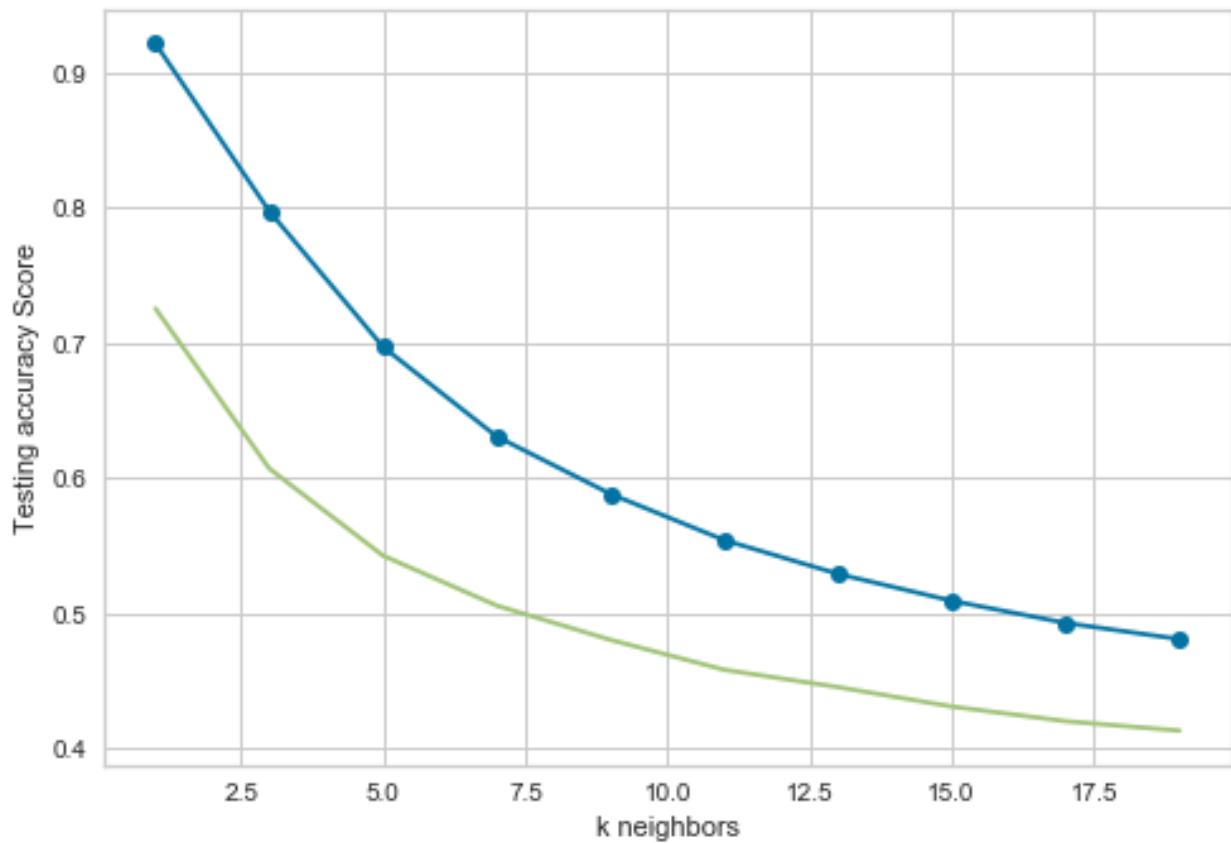


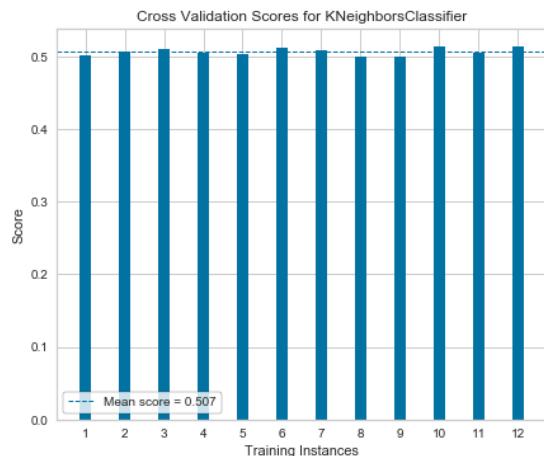
**LogisticRegression Confusion Matrix**

True Class	Iconic: Over \$200	Luxury: \$50-100	Popular: \$10-15	Premium: \$15-20	Super Luxury: \$100-200	Super Premium: \$20-30	Ultra Premium: \$30-50	Value: Under \$10	
Iconic: Over \$200	0	2	84	1	0	42	63	0	
Luxury: \$50-100	0	52	781	29	0	892	2199	0	
Popular: \$10-15	0	263	1727	867	0	1259	1574	0	
Premium: \$15-20	0	174	1471	846	0	1607	1972	0	
Super Luxury: \$100-200	0	7	198	2	0	129	329	0	
Super Premium: \$20-30	0	108	1316	517	0	2015	3315	0	
Ultra Premium: \$30-50	0	89	1057	132	0	1601	4336	0	
Value: Under \$10	0	72	831	304	0	556	754	0	



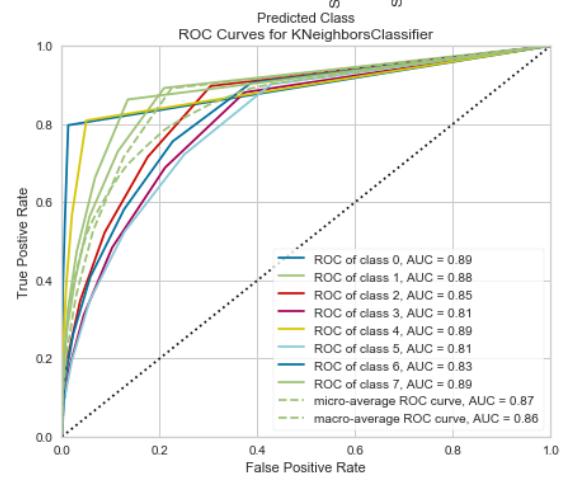
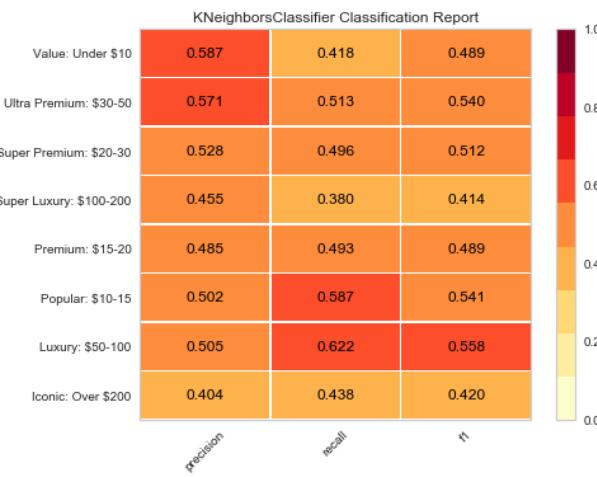
### KNN Model



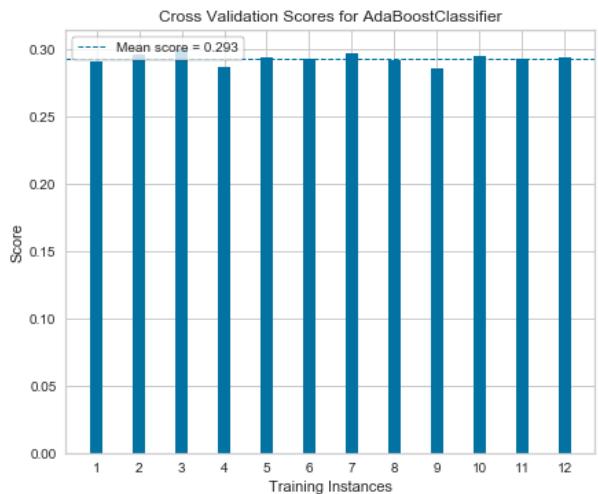


KNeighborsClassifier Confusion Matrix

Iconic: Over \$200	84	41	5	6	13	12	29	2
Luxury: \$50-100	33	2458	166	160	106	328	671	31
Popular: \$10-15	8	249	3340	830	14	547	364	338
Premium: \$15-20	8	316	1175	2993	24	889	517	148
Super Luxury: \$100-200	25	199	31	39	253	35	75	8
Super Premium: \$20-30	21	525	748	1209	52	3610	973	133
Ultra Premium: \$30-50	28	1014	482	633	83	1196	3698	81
Value: Under \$10	1	62	712	306	11	217	155	1053

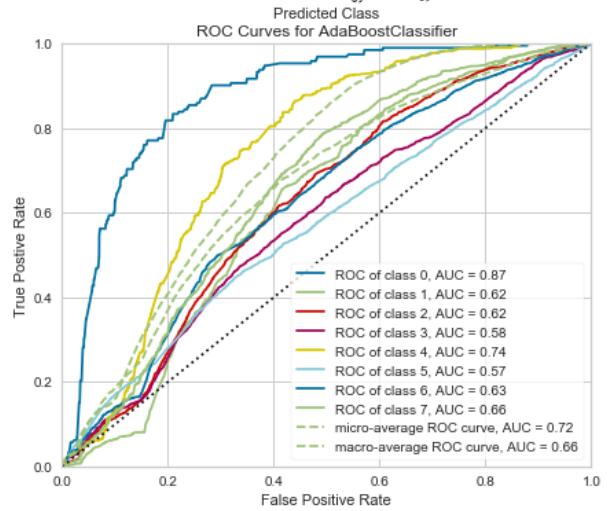
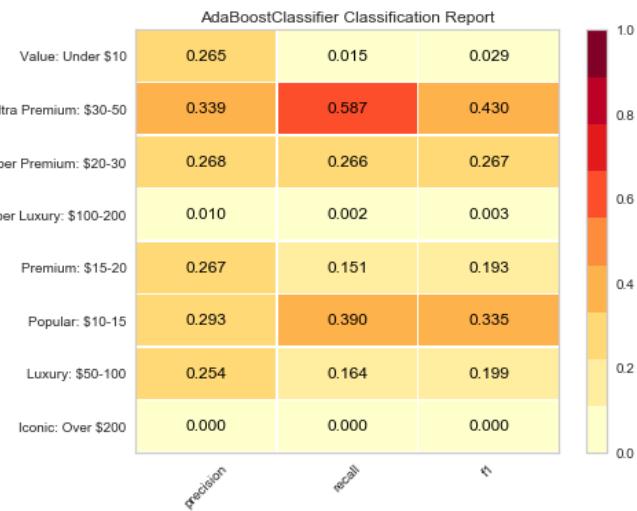


## Ada Boost Classifier

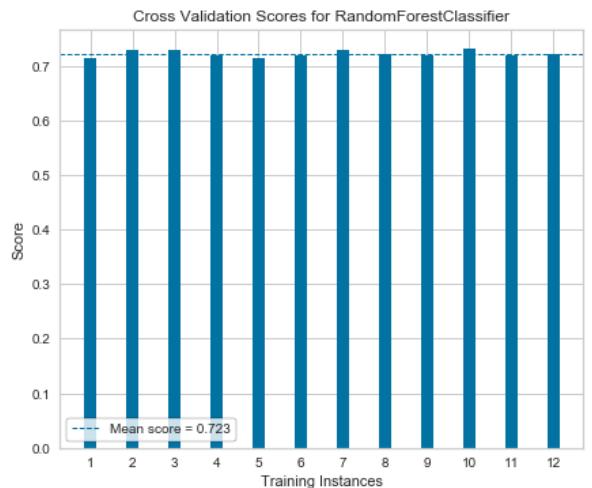


**AdaBoostClassifier Confusion Matrix**

True Class	Iconic: Over \$200	0	80	28	8	0	34	42	0
Luxury: \$50-100	5	648	358	269	13	487	2172	1	
Popular: \$10-15	11	220	2221	811	17	1357	1000	53	
Premium: \$15-20	8	334	1656	918	17	1692	1421	24	
Super Luxury: \$100-200	0	188	76	38	1	76	286	0	
Super Premium: \$20-30	14	437	1255	641	25	1934	2943	22	
Ultra Premium: \$30-50	8	581	722	505	18	1140	4233	8	
Value: Under \$10	2	59	1267	248	7	504	391	39	

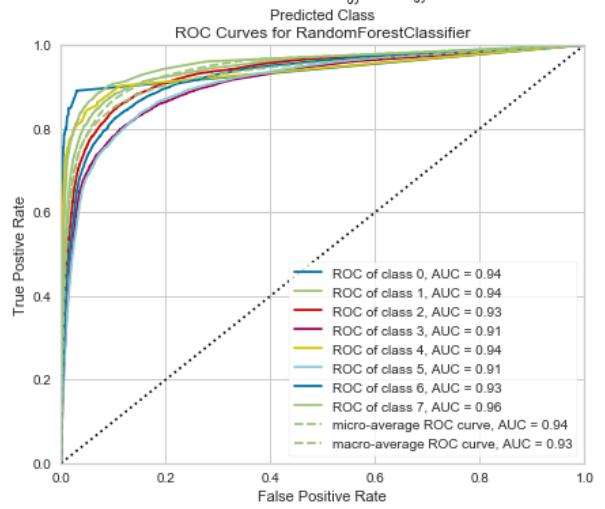


## Random Forest Classifier

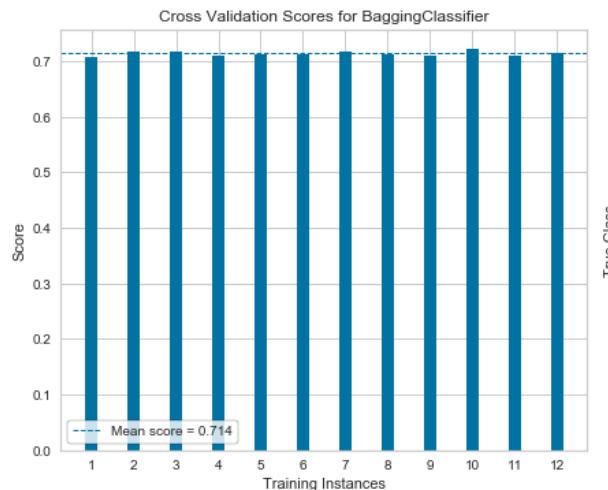


**RandomForestClassifier Confusion Matrix**

Iconic: Over \$200	116	12	0	4	31	18	11	0
Luxury: \$50-100	12	2974	52	84	117	165	530	19
Popular: \$10-15	2	63	4326	566	9	273	129	322
Premium: \$15-20	1	70	611	4311	4	712	266	95
Super Luxury: \$100-200	23	118	5	13	434	23	47	2
Super Premium: \$20-30	8	153	272	711	19	5310	728	70
Ultra Premium: \$30-50	12	530	132	218	31	745	5502	45
Value: Under \$10	0	11	403	92	0	73	50	1888

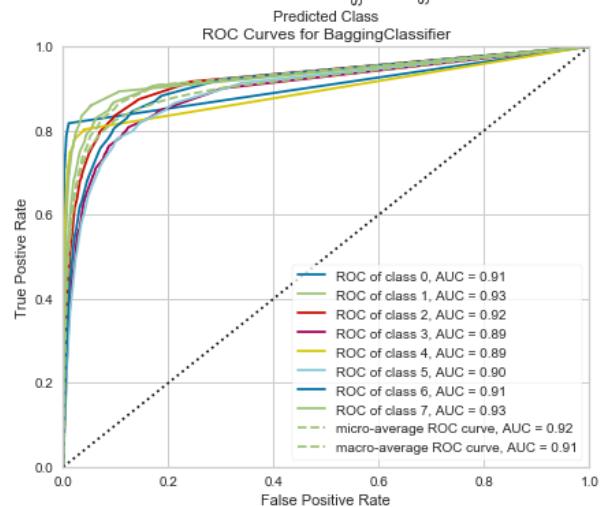


## Bagging Classifier



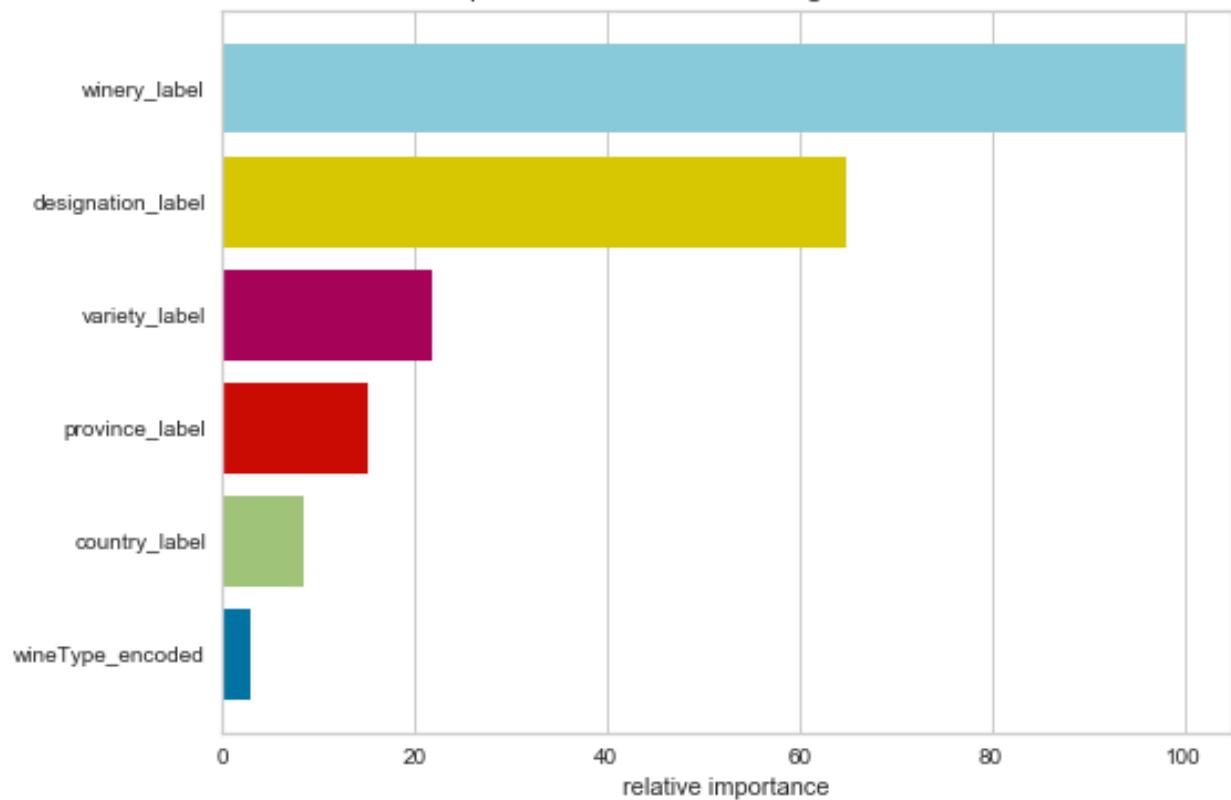
BaggingClassifier Confusion Matrix

True Class	Iconic: Over \$200	125	16	2	5	25	13	6	0
Luxury: \$50-100	17	2971	60	81	121	141	542	20	
Popular: \$10-15	1	79	4328	567	12	252	131	320	
Premium: \$15-20	1	83	632	4265	4	708	294	83	
Super Luxury: \$100-200	23	135	5	12	427	20	37	6	
Super Premium: \$20-30	8	162	290	754	29	5198	757	73	
Ultra Premium: \$30-50	16	592	151	251	38	769	5358	40	
Value: Under \$10	1	12	412	96	1	81	47	1867	

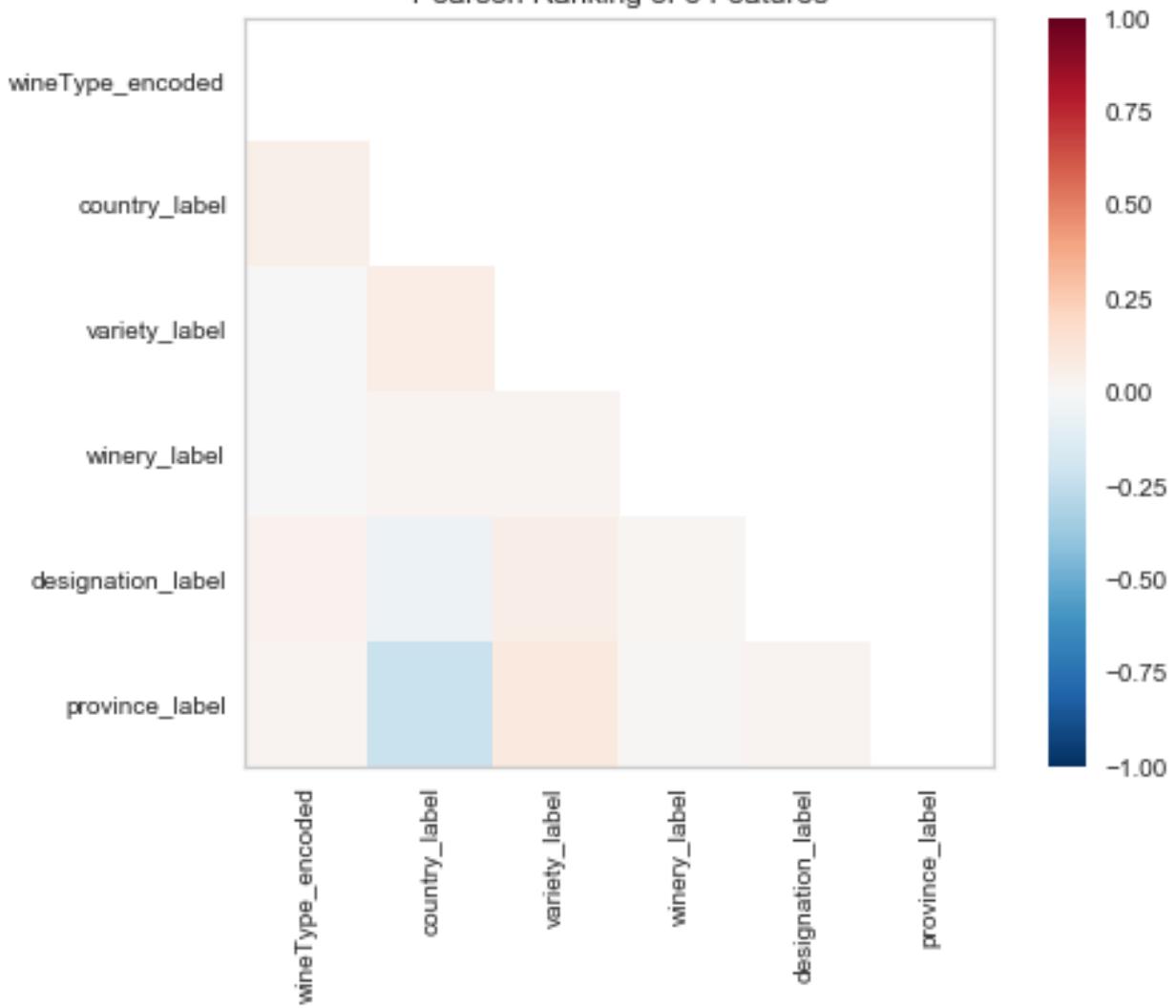


## Tuning

Feature Importances of 6 Features using RandomForestClassifier

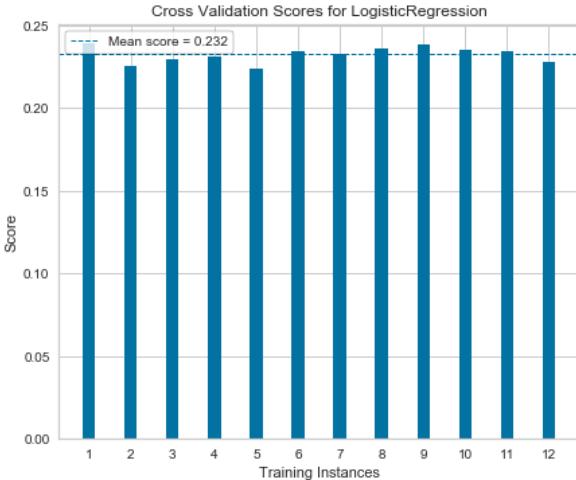


Pearson Ranking of 6 Features



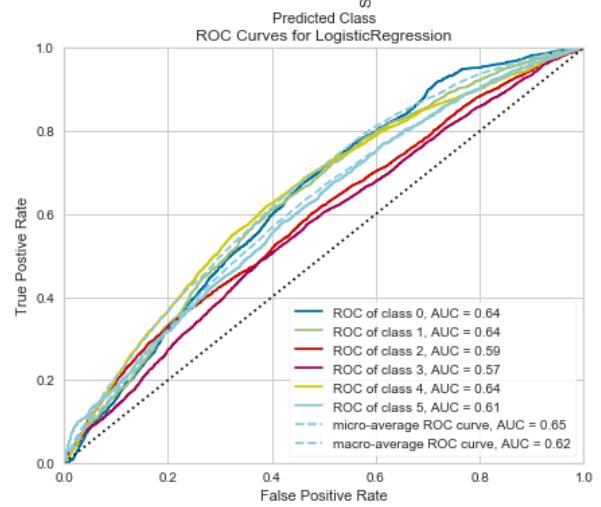
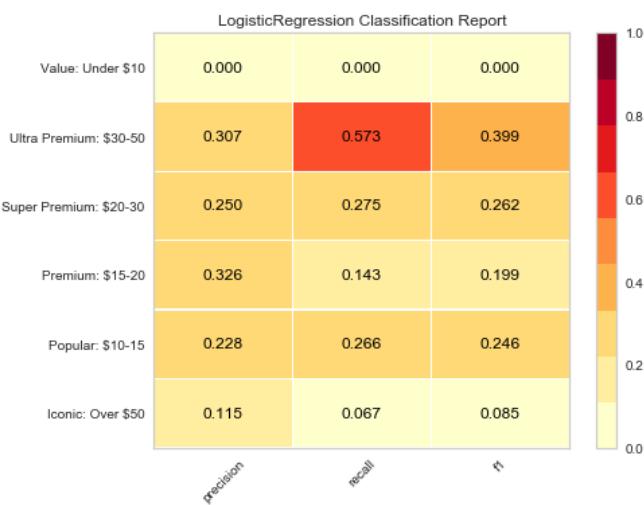
## Round Four

### Logistic Regression

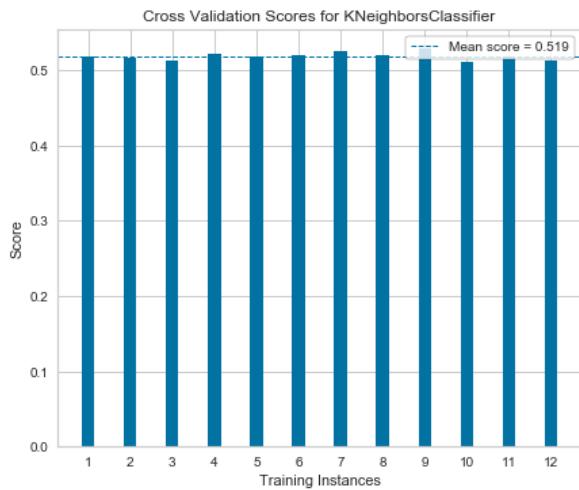


**LogisticRegression Confusion Matrix**

True Class	Iconic: Over \$50	Popular: \$10-15	Premium: \$15-20	Super Premium: \$20-30	Ultra Premium: \$30-50	Value: Under \$10	
Iconic: Over \$50	322	972	25	1050	2442	0	
Popular: \$10-15	652	1513	873	1279	1373	0	
Premium: \$15-20	607	1280	870	1554	1759	0	
Super Premium: \$20-30	542	1195	486	1999	3049	0	
Ultra Premium: \$30-50	404	971	128	1579	4133	0	
Value: Under \$10	262	695	290	545	724	0	

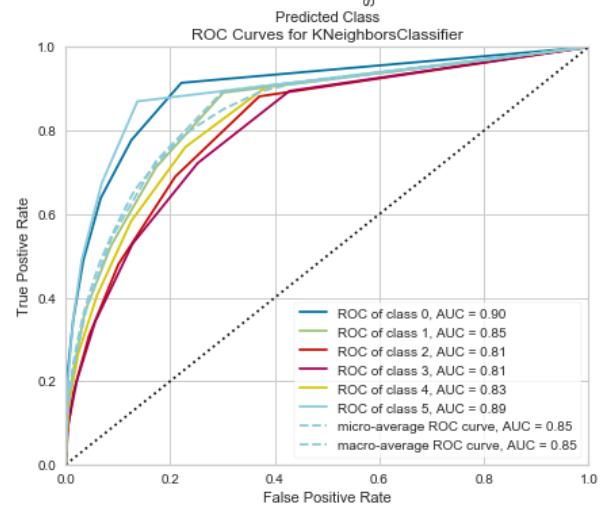


## KNN Model



**KNeighborsClassifier Confusion Matrix**

True Class						
	Iconic: Over \$50	Popular: \$10-15	Premium: \$15-20	Super Premium: \$20-30	Ultra Premium: \$30-50	Value: Under \$10
Iconic: Over \$50	3326	168	196	389	697	35
Popular: \$10-15	301	3312	843	545	325	364
Premium: \$15-20	362	1196	2963	890	504	155
Super Premium: \$20-30	648	723	1185	3612	964	139
Ultra Premium: \$30-50	1197	483	609	1223	3601	102
Value: Under \$10	87	686	277	216	151	1099



## Decision Tree Classifier



DecisionTreeClassifier Confusion Matrix

Iconic: Over \$50	3902	69	81	201	530	28
Popular: \$10-15	91	4345	538	267	133	316
Premium: \$15-20	114	664	4301	625	258	108
Super Premium: \$20-30	235	251	778	5195	720	92
Ultra Premium: \$30-50	694	147	268	763	5301	42
Value: Under \$10	16	390	100	59	43	1908

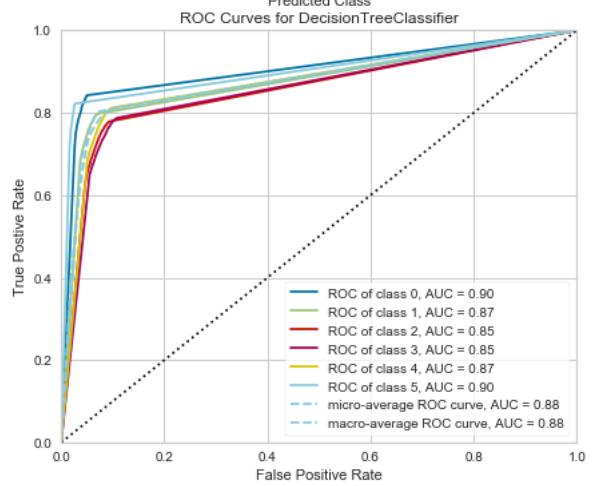
True Class

Iconic: Over \$50  
Popular: \$10-15  
Premium: \$15-20  
Super Premium: \$20-30  
Ultra Premium: \$30-50  
Value: Under \$10

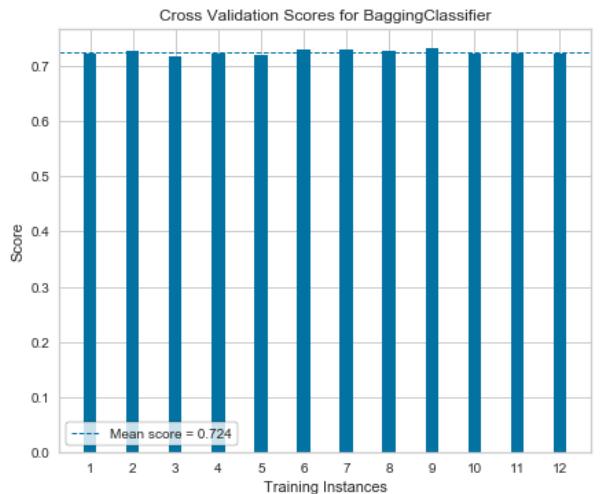
Predicted Class



ROC Curves for DecisionTreeClassifier



*Bagging Classifier*

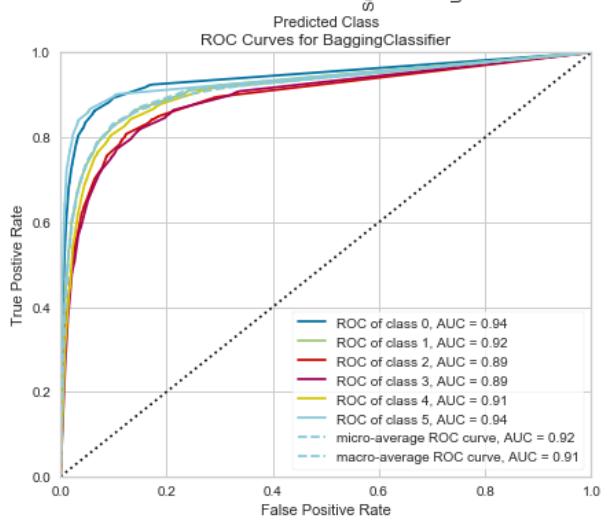


BaggingClassifier Confusion Matrix

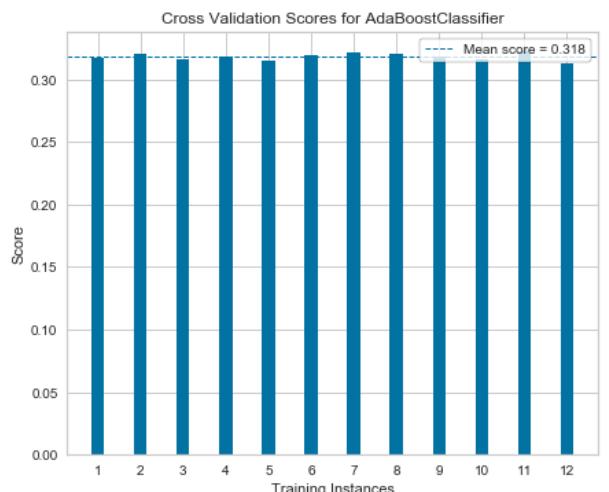
	Iconic: Over \$50	65	78	204	583	23
Iconic: Over \$50	3858					
Popular: \$10-15	107	4229	640	263	130	321
Premium: \$15-20	118	657	4234	706	274	81
Super Premium: \$20-30	219	288	758	5177	763	66
Ultra Premium: \$30-50	606	148	244	770	5409	38
Value: Under \$10	21	370	108	72	41	1904

True Class

Iconic: Over \$50  
Popular: \$10-15  
Premium: \$15-20  
Super Premium: \$20-30  
Ultra Premium: \$30-50  
Value: Under \$10

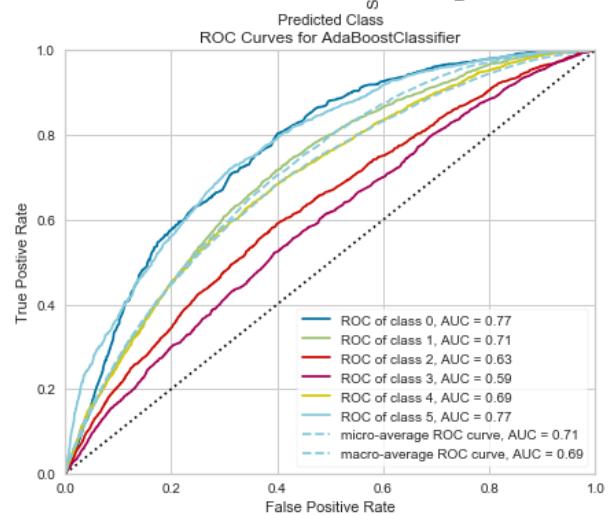
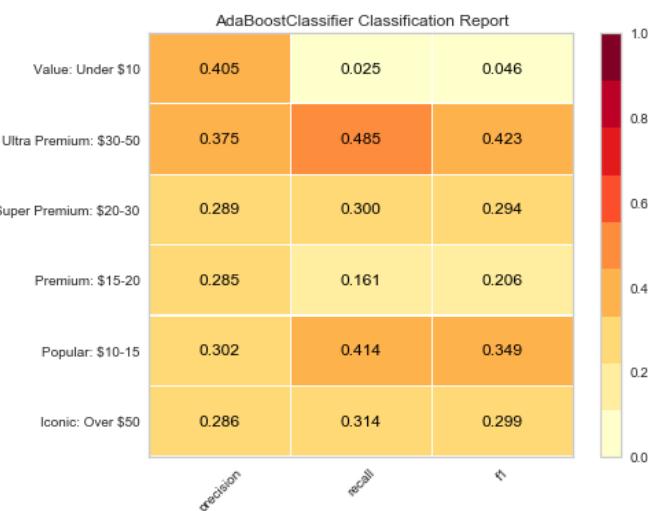


## Ada Boost Classifier



**AdaBoostClassifier Confusion Matrix**

Iconic: Over \$50	1510	471	210	571	2042	7
Popular: \$10-15	582	2358	853	1264	591	42
Premium: \$15-20	797	1680	976	1674	924	19
Super Premium: \$20-30	1069	1212	742	2184	2051	13
Ultra Premium: \$30-50	1145	731	421	1412	3496	10
Value: Under \$10	182	1358	222	463	229	62

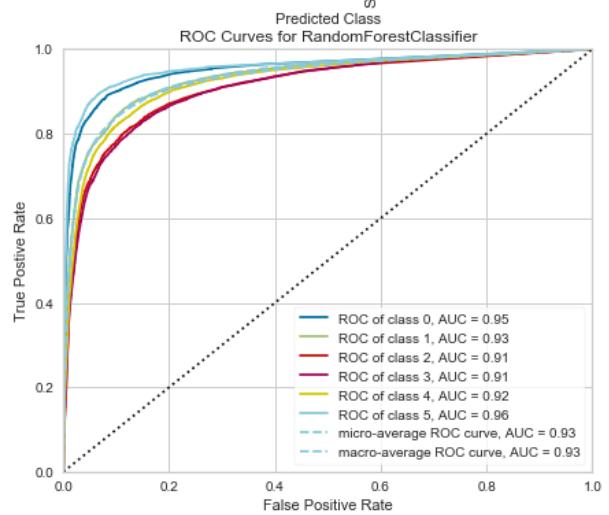


## Random Forest Classifier

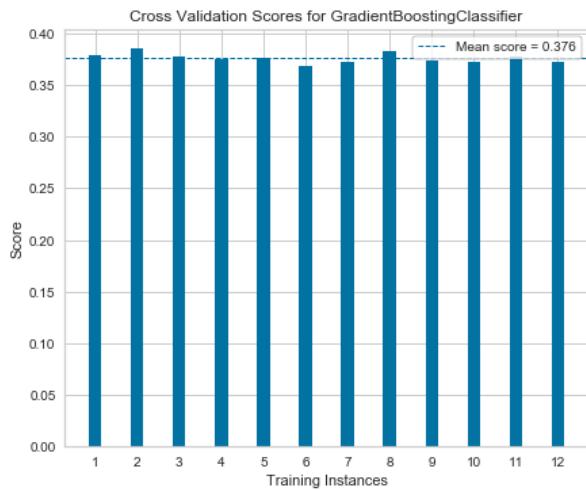


RandomForestClassifier Confusion Matrix

True Class	Iconic: Over \$50	Popular: \$10-15	Premium: \$15-20	Super Premium: \$20-30	Ultra Premium: \$30-50	Value: Under \$10
Iconic: Over \$50	3866	47	72	203	601	22
Popular: \$10-15	73	4283	620	269	128	317
Premium: \$15-20	105	591	4318	708	265	83
Super Premium: \$20-30	205	272	701	5266	763	64
Ultra Premium: \$30-50	601	129	225	749	5472	39
Value: Under \$10	9	398	94	61	42	1912

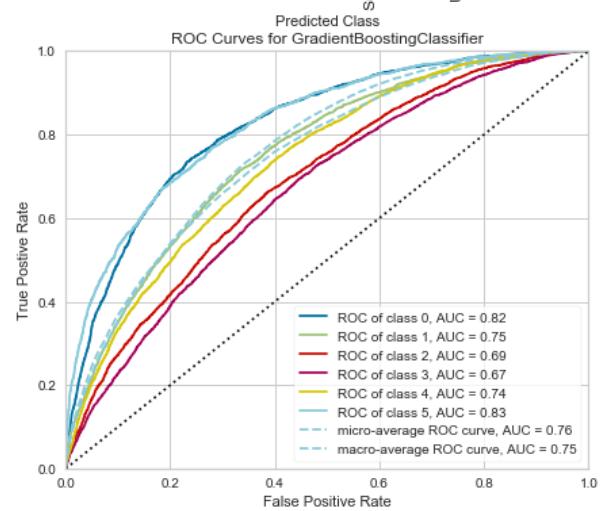
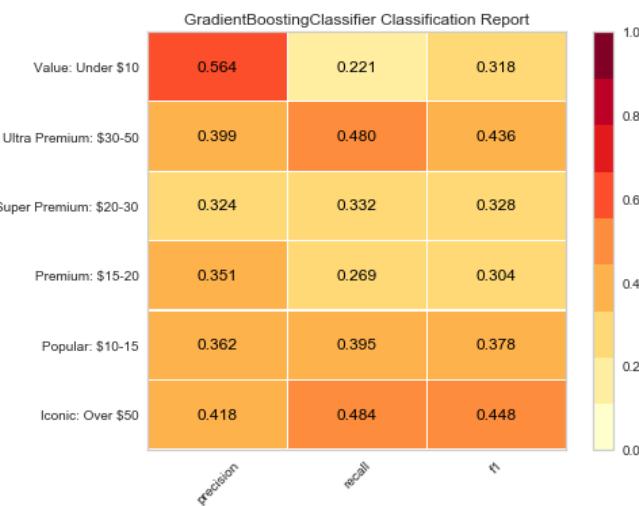


## Gradient Boosting Classifier



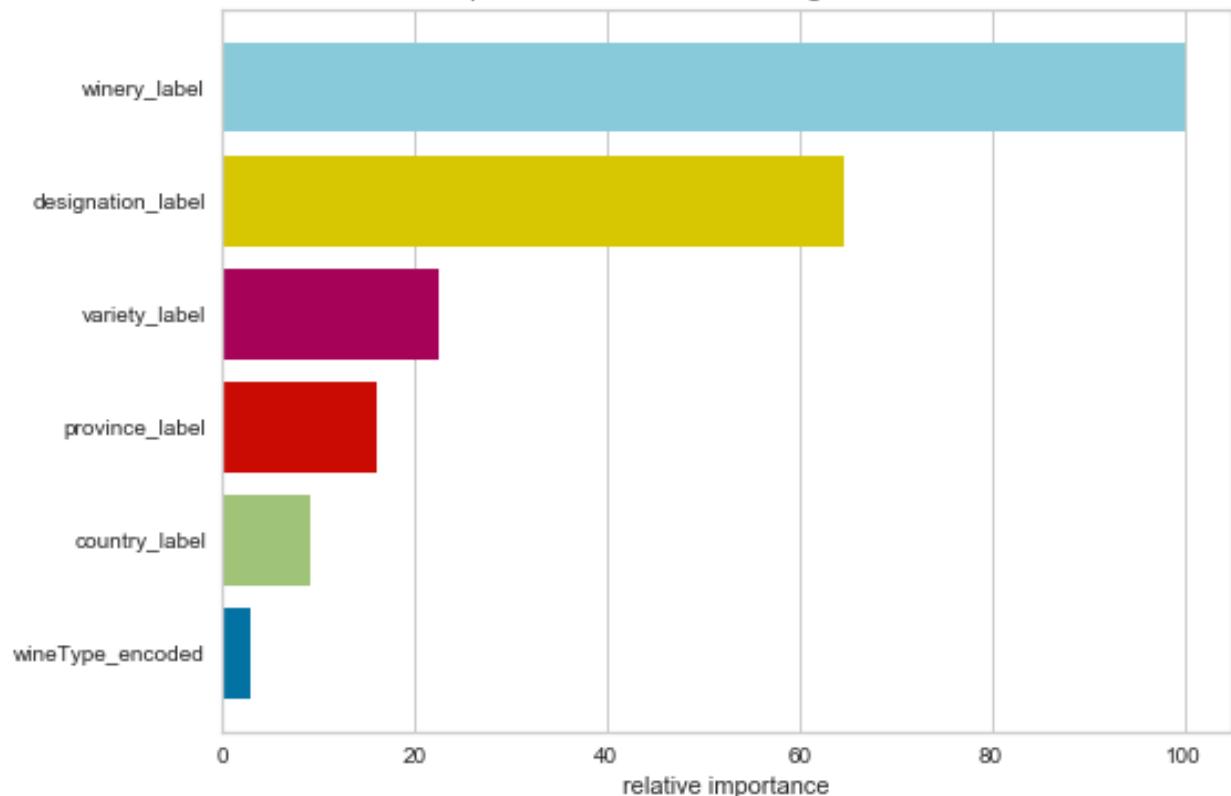
GradientBoostingClassifier Confusion Matrix

True Class							Value: Under \$10
	Iconic: Over \$50	Popular: \$10-15	Premium: \$15-20	Super Premium: \$20-30	Ultra Premium: \$30-50	Iconic: Over \$50	
Iconic: Over \$50	2327	343	153	589	1387	12	
Popular: \$10-15	346	2250	1207	1014	628	245	
Premium: \$15-20	507	1286	1633	1592	954	98	
Super Premium: \$20-30	887	924	987	2416	2007	50	
Ultra Premium: \$30-50	1388	518	391	1429	3464	25	
Value: Under \$10	112	903	287	425	233	556	



*Tuning*

Feature Importances of 6 Features using RandomForestClassifier



Pearson Ranking of 6 Features

