Marmara University Faculty of Engineering

CSE4062 – Data Science, Spring2020

Group7

"DRIVER DROWSINESS DETECTION"

Delivery #4 - Report

CSE	Mahmut AKTAŞ	aktasmahmut97@gmail.com	150115010
CSE	Mustafa Abdullah HAKKOZ	mustafa.hakkoz@gmail.com	150117509
ME	Ozan Berke YABAR	ozanberkeyabar@gmail.com	150416822
MME	Ece HARPUTLU	harputlue@gmail.com	150515038
EE	Nurettin ABACI	abacinurettin@gmail.com	150715035

Submitted to: Assoc. Prof. Murat Can Ganiz

- 16.05.2020 -

Predictive Analytics

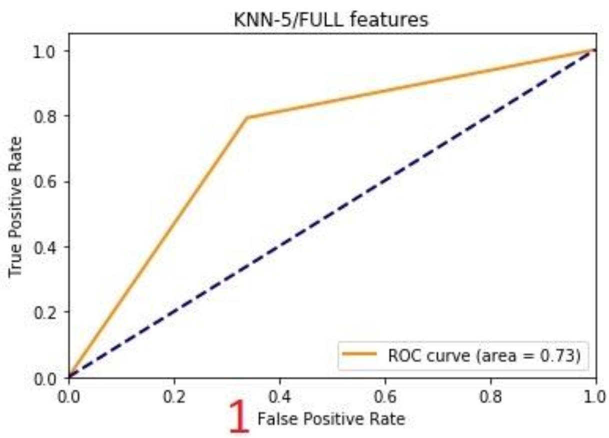
1- Feature Selection

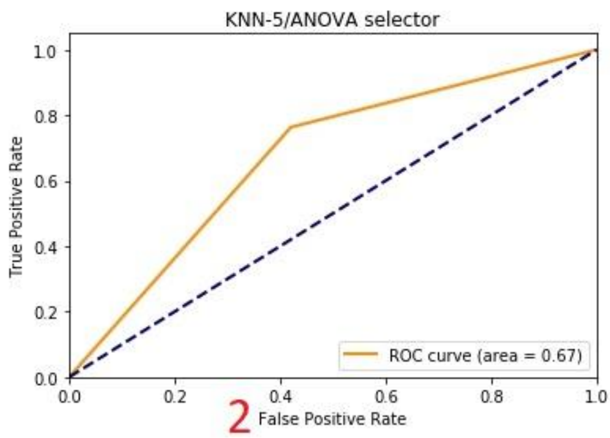
#	Feature Name	Description	ANOVA						Mutual Info						Decision Tree					
			Fold1	Fold2	Fold3	Fold4	Fold5	Avg	Fold1	Fold2	Fold3	Fold4	Fold5	Avg	Fold1	Fold2	Fold3	Fold4	Fold5	Avg
1	EAR	Eye Aspect Ratio	27790	27687	27644	27617	28057	27759	0.0543	0.0540	0.0539	0.0535	0.0552	0,0541	1	1	1	1	1	1
2	MAR	Mouth Aspect Ratio	3759	3792	3783	3765	3775	3775	0.1314	0.1313	0.1301	0.1309	0.1321	0,1311	1	1	1	1	1	1
3	MOE	Mouth Over Eye	4846	4873	4857	4830	4869	4855	0.0243	0.0249	0.0234	0.0235	0.0236	0,0239	1	1	1	1	1	1
4	EC	Eye Circularity	22696	22459	22432	22512	22879	22596	0.0361	0.0360	0.0360	0.0357	0.0360	0,0359	2	2	1	2	2	1.8
5	LEB	Level of Eyebrows	14953	14904	14869	15080	15043	14970	0.0646	0.0652	0.0637	0.0663	0.0648	0,0649	1	1	1	1	1	1
6	SOP	Size of Pupil	26841	26594	26570	26647	27051	26741	0.0892	0.0897	0.0883	0.0882	0.0896	0,089	3	3	1	3	3	2.6
7	PERCLOS	Percentage of Eye Closure	17311	17022	17106	16943	17031	17083	0.0359	0.0353	0.0359	0.0359	0.0360	0,0358	1	1	1	1	1	1
8	CLOSENESS	Eye Closure Status	8448	8250	8343	8229	8388	8332	0.0268	0.0281	0.0264	0.0279	0.0278	0,0274	4	4	1	4	4	3.4

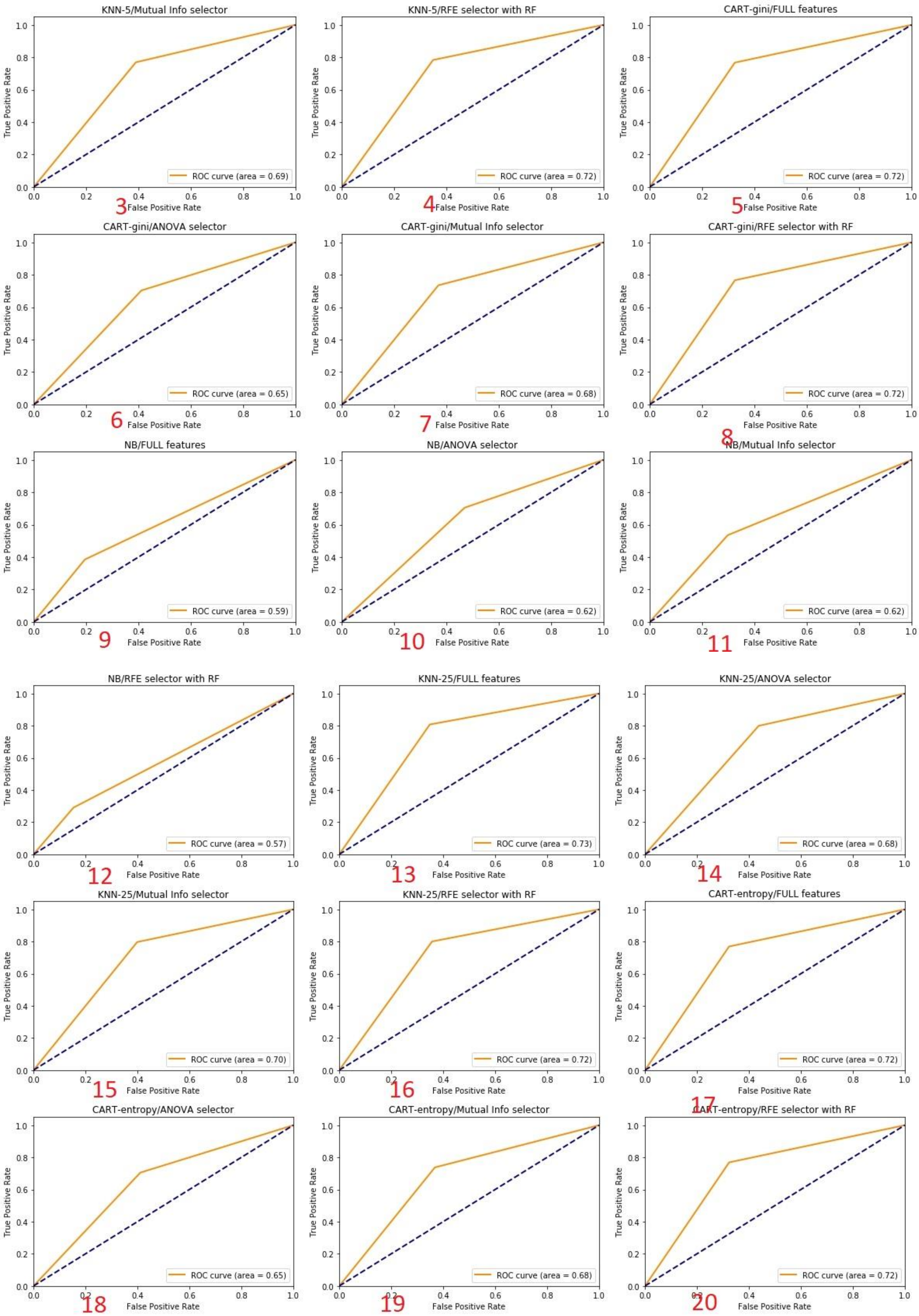
2- Classification Experiments

#	Experiment	Accuracy	Precision	Recall	AUC	F1
		Avg	Avg	Avg	Avg	Avg
1	**KNN5 FULL**	0,7378	0,7698	0,7931	0,7284	0,7813
2	**KNN5 ANOVA**	0,6879	0,7197	0,7636	0,6709	0,741
3	**KNN5 MI**	0,7029	0,7108	0,7542	0,6595	0,7319
4	**KNN5 RFE**	0,7286	0,7592	0,7813	0,7149	0,7701
5	**CART-GINI FULL**	0,7288	0,7698	0,7678	0,7211	0,7688
6	**CART-GINI ANOV**	0,6554	0,7074	0,7042	0,6455	0,7058
7	**CART MI**	0,6918	0,7132	0,7129	0,6531	0,713
8	**CART RFE**	0,7289	0,7693	0,767	0,7204	0,7681
9	**NB FULL**	0,5585	0,7341	0,3854	0,5937	0,5055
10	**NB ANOVA**	0,6326	0,6794	0,7031	0,6163	0,6911
11	**NB MI**	0,6039	0,6984	0,5635	0,6092	0,6237
12	**NB RFE**	0,5202	0,7222	0,267	0,5606	0,3899
13	**KNN25 FULL**	0,7429	0,7668	0,8071	0,7295	0,7864
14	**KNN 25 ANOVA**	0,7012	0,7209	0,7986	0,6801	0,7578
15	**KNN25 MI**	0,7159	0,7148	0,7901	0,6715	0,7506
16	**KNN25 RFE**	0,7351	0,7569	0,7963	0,7168	0,7761
17	**CART-ENT FULL**	0,7315	0,7711	0,7706	0,7231	0,7708
18	**CART-ENT ANOVA**	0,6568	0,7072	0,7046	0,6454	0,7059
19	**CART-ENT MI**	0,6939	0,7147	0,7142	0,6549	0,7144
20	**CART-ENT RFE**	0,7305	0,7712	0,7696	0,7229	0,7704

3- ROC Curves







KNN-5/Mutual Info selector
ROC curve (area = 0.69)
3
KNN-5/RFE selector with RF
ROC curve (area = 0.72)
4
CART-gini/FULL features
ROC curve (area = 0.72)
5
CART-gini/ANOVA selector
ROC curve (area = 0.65)
6
CART-gini/Mutual Info selector
ROC curve (area = 0.68)
7
CART-gini/RFE selector with RF
ROC curve (area = 0.72)
8
NB/FULL features
ROC curve (area = 0.59)
9
NB/ANOVA selector
ROC curve (area = 0.62)
10
NB/Mutual Info selector
ROC curve (area = 0.62)
11
NB/RFE selector with RF
ROC curve (area = 0.57)
12
KNN-25/FULL features
ROC curve (area = 0.73)
13
KNN-25/ANOVA selector
ROC curve (area = 0.68)
14
KNN-25/Mutual Info selector
ROC curve (area = 0.70)
15
KNN-25/RFE selector with RF
ROC curve (area = 0.72)
16
CART-entropy/FULL features
ROC curve (area = 0.72)
17
CART-entropy/ANOVA selector
ROC curve (area = 0.65)
18
CART-entropy/Mutual Info selector
ROC curve (area = 0.68)
19
CART-entropy/RFE selector with RF
ROC curve (area = 0.72)
20
True Positive Rate
False Positive Rate

4- Confusion Matrices

5 - T-Test

- **Accuracy T_Test** between KNN-5 & KNN-25:

  T Value = [-6.9592847], P Value = [0.00018816]

  p-value<=0.05 so there's a significant difference between models.

- **Precision T_Test** between KNN-5 & KNN-25:

  T Value = [2.06812435], P Value = [0.07486675]

  p-value>0.05 so there's no significant difference between models.

- **Recall T_Test** between KNN-5 & KNN-25:

  T Value = [-34.54848005], P Value = [9.06573874e-08]

  p-value<=0.05 so there's a significant difference between models.

- **F1 T_Test** between KNN-5 & KNN-25:

  T Value = [-11.94455852], P Value = [6.83443097e-06]

  p-value<=0.05 so there's a significant difference between models.

- **AUC T_Test** between KNN-5 & KNN-25:

  T Value = [-3.56333437], P Value = [0.00841836]

  p-value<=0.05 so there's a significant difference between models.

6 – Methods

For this iteration of the project, we first determined 4 different feature subsets:

1. Full sets of features
2. Feature selection with SelectKBest method of SkLearn. As scoring function we choosed f_classif (ANOVA) .
3. Feature selection with SelectKBest method of SkLearn. As scoring function we choosed f_ mutual_info_classif (Mutual Information) .
4. Feature selection with RFECV (Recursive feature elimination on CV). As estimator parameter we chosoed DecisionTreeClassifier.

And we determined 5 different classifying models:

1. KNeighborsClassifier model of Sklearn with 5 neighbours
2. KNeighborsClassifier model of Sklearn with 25 neighbours
3. DecisionTreeClassifier model of Sklearn with gini
4. DecisionTreeClassifier model of Sklearn with entropy
5. GaussianNB model of Sklearn.

And finaly we determined 5 metrics to calculate t-test on:

1. accuracy_score metric of Sklearn.
2. precision_score metric of Sklearn.
3. recall_score metric of Sklearn.
4. f1_score metric of Sklearn.
5. roc_auc_score metric of Sklearn.

7 – Implementation

After researching on how to implement feature selection without causing bias on test-set, we concluded to use 5x2 cross-validation technique referenced on 1998 paper titled "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms" by Thomas Dietterich. [1]

It's recommended to use an outer CV (i.e. 5 folds) to repeat the whole pipeline to produce an array of results which will be necessary to run t-test on. In every fold, you can do feature elimination without effecting test-set. Also, you can use an inner cv to run your experiments.

In our case we used 5 folds for outer CV and we did train-test splitting there. Then in every fold we first run our 3 feature elimination methods. For third method (RFE), we needed another CV to run so we choose 2 folds for inner CV. We also transformed corresponding outer test-sets by using selected features. After that, we run our 5 different classifiers and run predictions on corresponding test-sets. And finally we evaluate models by using 5 metrics we determined above. In total, we run 5 x 4 = 20 different experiments in for 5 folds.

8 – Conclusion

For t-test, we chose our top two experiments according to ROC-AUC metric :

1. **KNN-25 with full features: 80.71**
2. **KNN-5 with full features: 79.31**

Then run scipy's ttest_ind() method on their CV scores. We compared them by using 5 different metrics on **topic 5: T-Test.**

According to precision metric there's no significant difference between them but for all other metrics, there is. So we can concluded that, **KNN-25 with full features** is our best experiment.

We also observed that, in general our feature elimination methods didn't provide better results than full features unless Naïve-Bayes classifier. In the case of NB, feature elimination methods based Mutual Information and ANOVA resulted in higher accuracy.

9 – References

[1] T. G. Dietterich, "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms," in Neural Computation, vol. 10, no. 7, pp. 1895-1923, 1 Oct. 1998, doi: 10.1162/089976698300017197.