

Evaluate and compare multiple classification algorithms: Main results and conclusion

Arnold Jiadong Yu

April 11, 2019

1 Task

The following classifiers are compared:

1. Naïve Bayesian
2. Decision tree classifier(entropy)
3. k-NN
4. Random Forest
5. Logistic regression
6. AdaBoost
7. SVM (Support Vector Machine)

As well as the following performance measures are computed based on 10 fold cross validation:

1. The confusion matrix.
2. Accuracy, precision, recall and F-1 measures.
3. ROC curves and AUC
4. Learning curves

2 Results

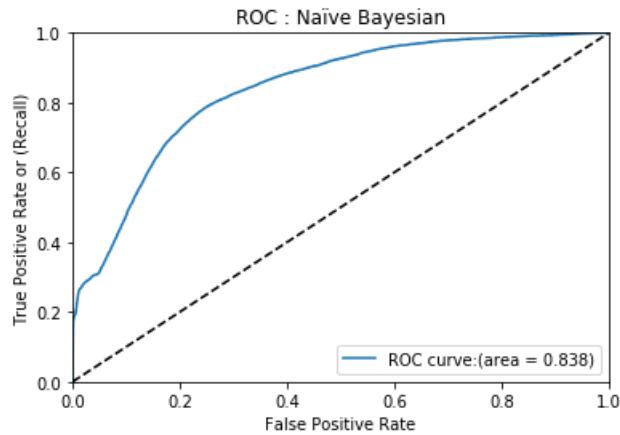
2.1 Census Income or Adult dataset

2.1.1 Naïve Bayesian

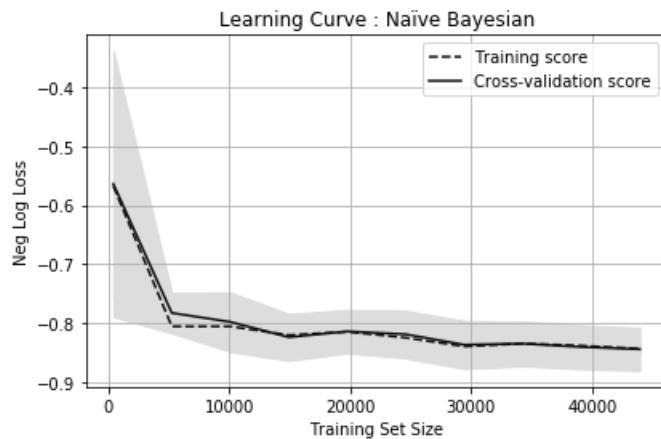
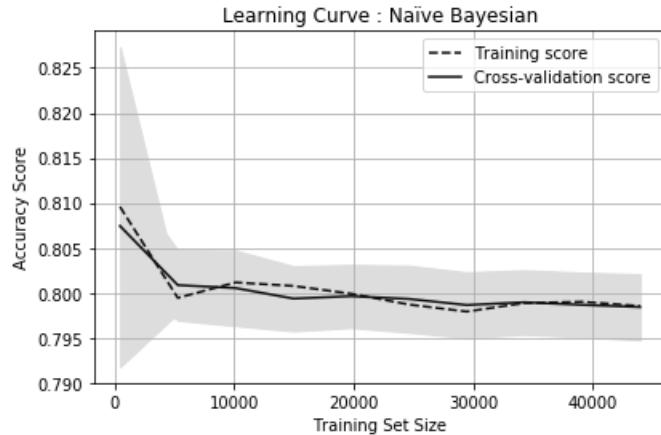
The confusion matrix, accuracy, precision, recall and F-1.

| | | | |
|-----------------------|-------|--------|--|
| Accuracy | float | 1 | 0.7984316776544778 |
| Classification_report | dict | 5 | {'Class 0':{'precision':0.8143126783905718, 'recall':0.952146413672453 ... |
| Confusion_matrix | int64 | (2, 2) | [[35377 1778] [8067 36201] |
| F1 | float | 1 | 0.8778520825320414 |
| Precision | float | 1 | 0.8143126783905718 |
| Recall | float | 1 | 0.9521464136724532 |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.

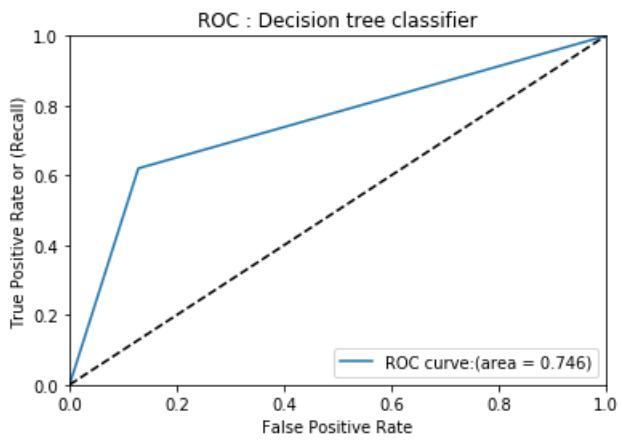


2.1.2 Decision tree classifier(entropy)

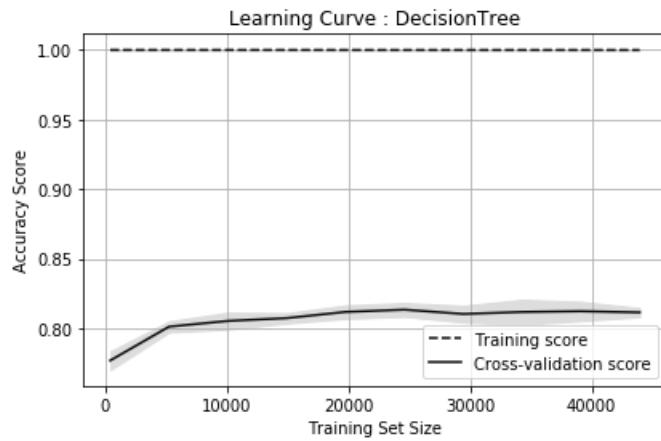
The confusion matrix, accuracy, precision, recall and F-1.

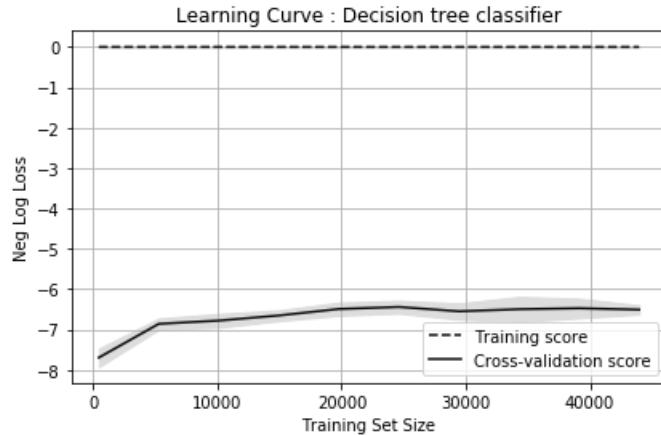
| | | | |
|-----------------------|-------|--------|--|
| Accuracy | float | 1 | 0.8115556283526473 |
| Classification_report | dict | 5 | {'Class 0': {'precision': 0.8794493768836035, 'recall': 0.871780379491320 ...} |
| Confusion_matrix | int64 | (2, 2) | [[32391 4764] [4440 7247]] |
| F1 | float | 1 | 0.8755980861244019 |
| Precision | float | 1 | 0.8794493768836035 |
| Recall | float | 1 | 0.8717803794913201 |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.



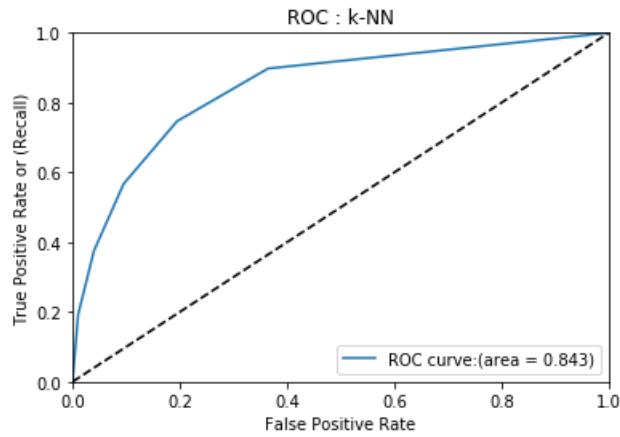


2.1.3 k-NN

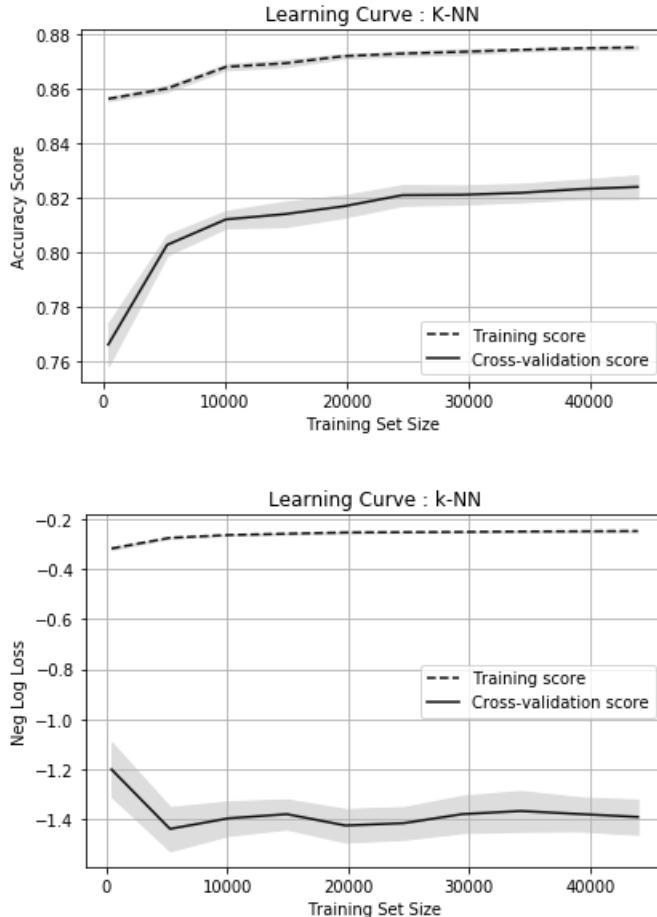
The confusion matrix, accuracy, precision, recall and F-1.

| | | | |
|-----------------------|-------|--------|---|
| Accuracy | float | 1 | <code>0.8239629826788419</code> |
| Classification_report | dict | 5 | <code>{'Class 0': {'precision': 0.8692301724806702, 'recall': 0.904696541515273 ...}</code> |
| Confusion_matrix | int64 | (2, 2) | <code>[[33614 3541] [5057 6630]]</code> |
| F1 | float | 1 | <code>0.8866088149183659</code> |
| Precision | float | 1 | <code>0.8692301724806702</code> |
| Recall | float | 1 | <code>0.9046965415152739</code> |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.

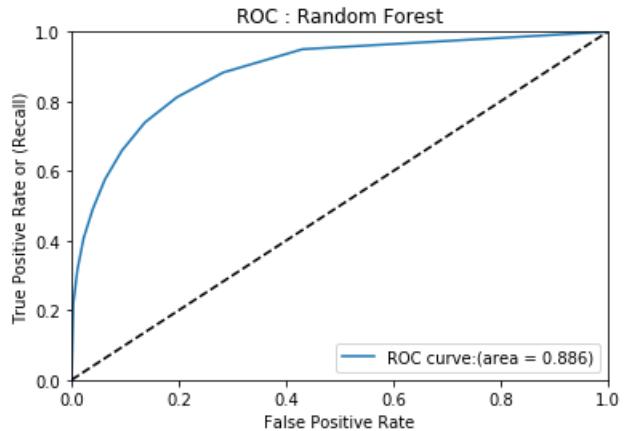


2.1.4 Random Forest

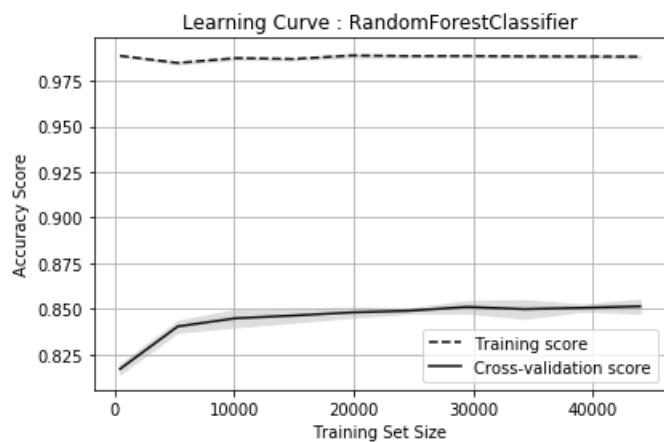
The confusion matrix, accuracy, precision, recall and F-1.

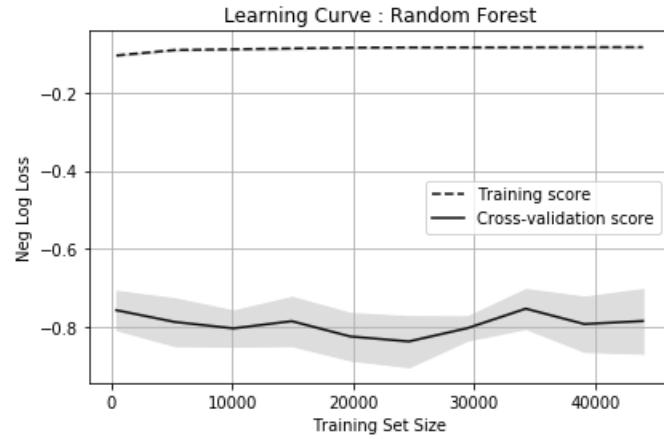
| | | | |
|-----------------------|-------|--------|---|
| Accuracy | float | 1 | <code>0.8469350149461529</code> |
| Classification_report | dict | 5 | <code>{'Class 0': {'precision': 0.8941748346481792, 'recall': 0.906015341138473 ...}</code> |
| Confusion_matrix | int64 | (2, 2) | <code>[[33663 3492] [3984 7703]]</code> |
| F1 | float | 1 | <code>0.9000561482313307</code> |
| Precision | float | 1 | <code>0.8941748346481792</code> |
| Recall | float | 1 | <code>0.9060153411384739</code> |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.



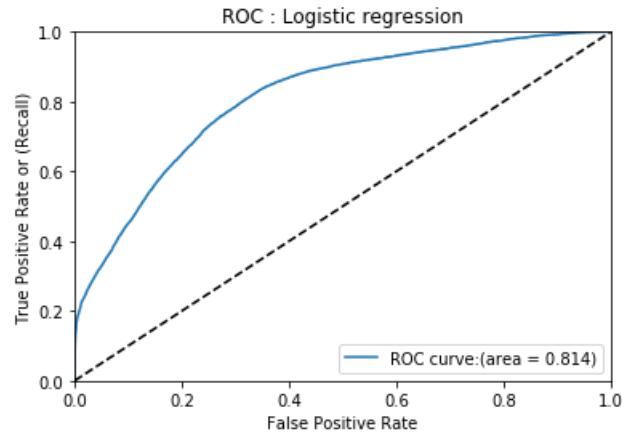


2.1.5 Logistic regression

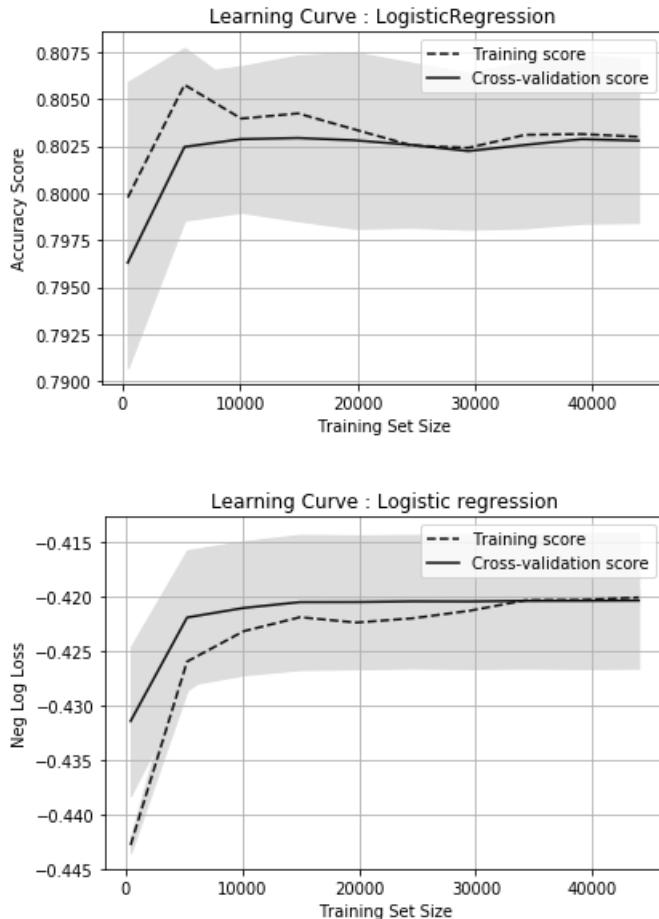
The confusion matrix, accuracy, precision, recall and F-1.

| | | | |
|-----------------------|-------|--------|--|
| Accuracy | float | 1 | <code>0.8027926784324966</code> |
| Classification_report | dict | 5 | <code>{'Class 0': {'precision': 0.8155945419103314, 'recall': 0.957179383663033 ...</code> |
| Confusion_matrix | int64 | (2, 2) | <code>[[35564 1591] [8041 36461]]</code> |
| F1 | float | 1 | <code>0.8807330361565132</code> |
| Precision | float | 1 | <code>0.8155945419103314</code> |
| Recall | float | 1 | <code>0.9571793836630332</code> |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.

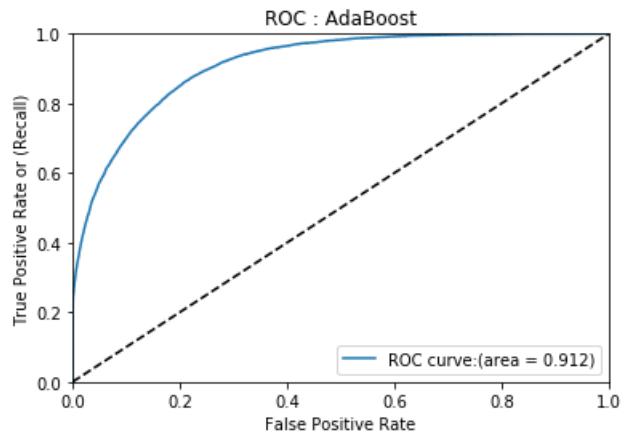


2.1.6 AdaBoost

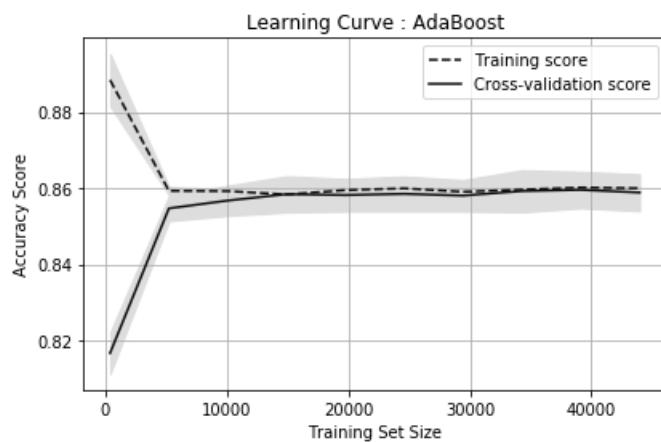
The confusion matrix, accuracy, precision, recall and F-1.

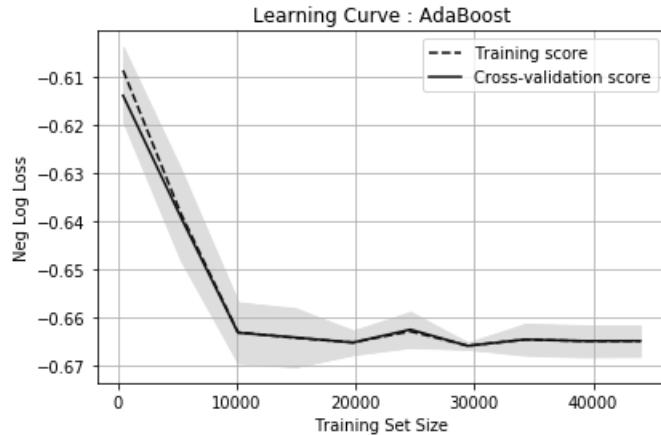
| | | | |
|-----------------------|-------|--------|--|
| Accuracy | float | 1 | <code>0.8588714630850498</code> |
| Classification_report | dict | 5 | <code>{'Class 0': {'precision': 0.8818838019282217, 'recall': 0.940438702731799 ...</code> |
| Confusion_matrix | int64 | (2, 2) | <code>[[34942 2213] [4680 7007]]</code> |
| F1 | float | 1 | <code>0.9102205087461088</code> |
| Precision | float | 1 | <code>0.8818838019282217</code> |
| Recall | float | 1 | <code>0.9404387027317992</code> |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.



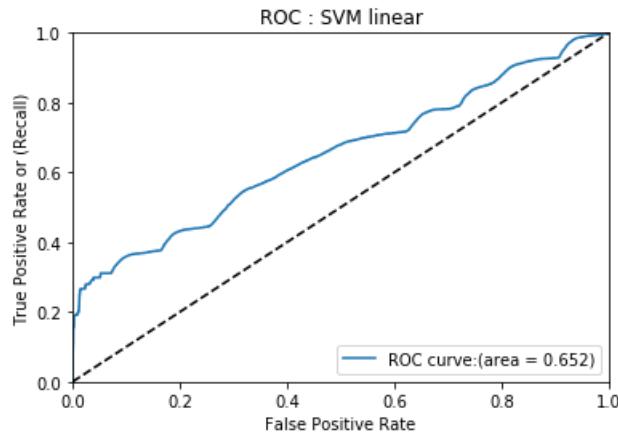


2.1.7 SVM (Support Vector Machine)

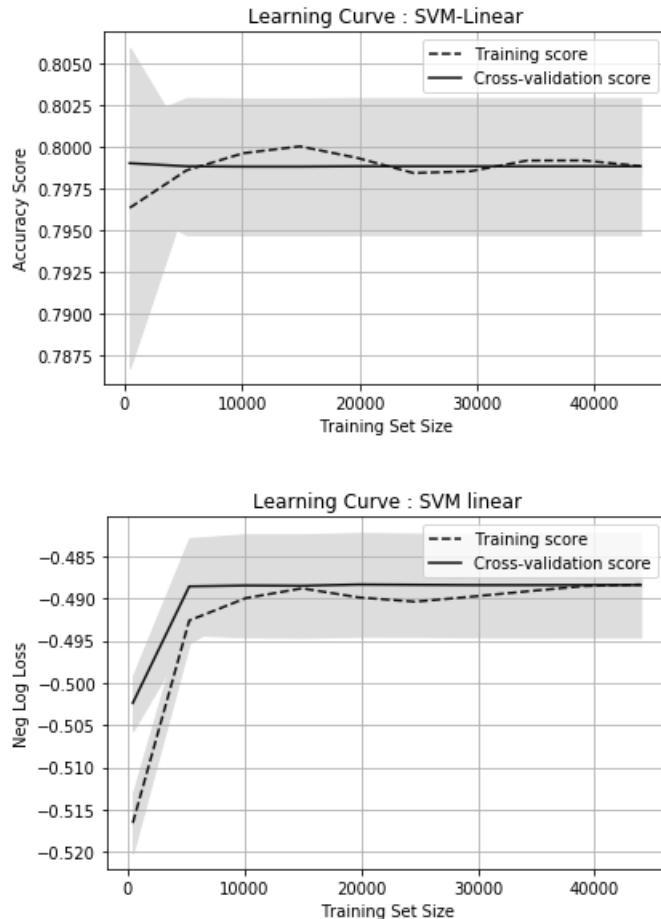
The confusion matrix, accuracy, precision, recall and F-1.

| | | | |
|-----------------------|-------|--------|---|
| Accuracy | float | 1 | <code>0.8037140166250358</code> |
| Classification_report | dict | 5 | <code>{'Class 0': {'precision': 0.8118128760801702, 'recall': 0.965872695464944 ...}</code> |
| Confusion_matrix | int64 | (2, 2) | <code>[[35887 1268] [8319 3368]]</code> |
| F1 | float | 1 | <code>0.8821671316724229</code> |
| Precision | float | 1 | <code>0.8118128760801702</code> |
| Recall | float | 1 | <code>0.9658726954649441</code> |

The ROC curve and AUC.



The learning curve based on accuracy and neg log loss.



2.2 MoCap Hand Postures Data Set

2.2.1 Naïve Bayesian

The confusion matrix, accuracy, precision, recall and F-1 of each class.

| Naïve Bayesian - Dictionary (3 elements) | | | | |
|--|---------|--------|--------------------------------|--|
| Key | Type | Size | Value | |
| Accuracy | float64 | 1 | 0.9981945066905692 | |
| Classification_report | dict | 8 | {'Class 1': {'precision': 1.0, | |
| Confusion_matrix | int64 | (5, 5) | [[16214 0 0 0 0 14978 0 0 0 | |

| Confusion_matrix - NumPy array | | | | | |
|--------------------------------|-------|-------|-------|-------|-------|
| | 0 | 1 | 2 | 3 | 4 |
| 0 | 16214 | 0 | 0 | 0 | 51 |
| 1 | 0 | 14978 | 0 | 0 | 0 |
| 2 | 0 | 0 | 16267 | 77 | 0 |
| 3 | 0 | 0 | 0 | 14762 | 13 |
| 4 | 0 | 0 | 0 | 0 | 15733 |

| macro avg - Dictionary (4 elements) | | | | |
|-------------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9981999344259155 | |
| precision | float | 1 | 0.9981519137849457 | |
| recall | float | 1 | 0.998254671837732 | |
| support | int | 1 | 78095 | |

| micro avg - Dictionary (4 elements) | | | | |
|-------------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9981945066905692 | |
| precision | float | 1 | 0.9981945066905692 | |
| recall | float | 1 | 0.9981945066905692 | |
| support | int | 1 | 78095 | |

| weighted avg - Dictionary (4 elements) | | | | |
|--|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.998194906961333 | |
| precision | float | 1 | 0.998202079359611 | |
| recall | float | 1 | 0.9981945066905692 | |
| support | int | 1 | 78095 | |

| Class 1 - Dictionary (4 elements) | | | | |
|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9984297546106715 | |
| precision | float | 1 | 1.0 | |
| recall | float | 1 | 0.9968644328312327 | |
| support | int | 1 | 16265 | |

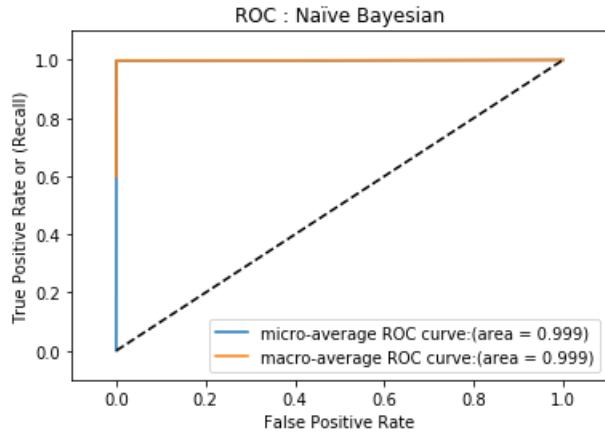
| Class 2 - Dictionary (4 elements) | | | | |
|-----------------------------------|-------|------|-------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 1.0 | |
| precision | float | 1 | 1.0 | |
| recall | float | 1 | 1.0 | |
| support | int | 1 | 14978 | |

| Class 3 - Dictionary (4 elements) | | | | |
|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9976388335224311 | |
| precision | float | 1 | 1.0 | |
| recall | float | 1 | 0.9952887909936368 | |
| support | int | 1 | 16344 | |

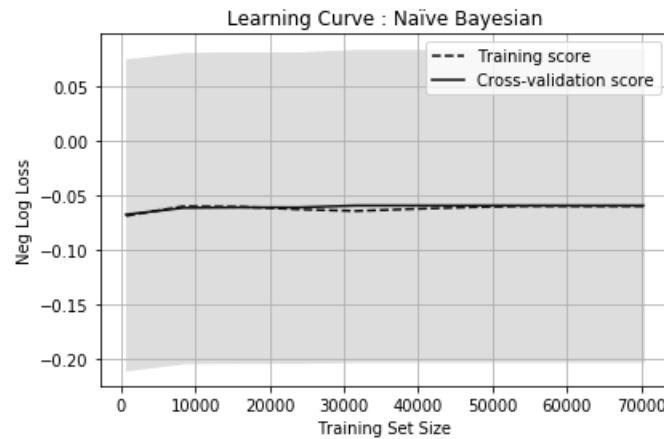
| Class 4 - Dictionary (4 elements) | | | | |
|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9969608968731005 | |
| precision | float | 1 | 0.9948109710896961 | |
| recall | float | 1 | 0.9991201353637902 | |
| support | int | 1 | 14775 | |

| Class 5 - Dictionary (4 elements) | | | | |
|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9979701871233746 | |
| precision | float | 1 | 0.995948597835032 | |
| recall | float | 1 | 1.0 | |
| support | int | 1 | 15733 | |

The ROC curve and AUC.



The learning curve based on neg log loss.



2.2.2 Decision tree classifier(entropy)

The confusion matrix, accuracy, precision, recall and F-1 of each class.

| Decision tree classifier - Dictionary (3 elements) | | | | Class 1 - Dictionary (4 elements) | | | | | | | |
|--|---------|--------|---|-----------------------------------|-------|--------|---|-----------|-------|---|---------------------------|
| Key | Type | Size | Value | Key | Type | Size | Value | | | | |
| Accuracy | float64 | 1 | 0.9975286510019847 | f1-score | float | 1 | 0.9948954489544896 | | | | |
| Classification_report | dict | 8 | {'Class 1': {'precision': 0.995204990591146, 'recall': 0.994589609591146, 'f1-score': 0.9948954489544896}, 'Class 2': {'precision': 0.9952014764687789, 'recall': 0.994589609591146, 'f1-score': 0.9948954489544896}, 'Class 3': {'precision': 0.9993990384615384, 'recall': 1.0, 'f1-score': 0.9993990384615384}, 'Class 4': {'precision': 0.999937711577522, 'recall': 1.0, 'f1-score': 0.999937711577522}, 'Class 5': {'precision': 0.9932724041634933, 'recall': 1.0, 'f1-score': 0.9932724041634933}, 'macro avg': {'precision': 0.9975711142546684, 'recall': 0.9975670325555605, 'f1-score': 0.9975670325555605}, 'micro avg': {'precision': 0.9975286510019847, 'recall': 0.9975286510019847, 'f1-score': 0.9975286510019847}, 'weighted avg': {'precision': 0.9975299860870034, 'recall': 0.9975286510019847, 'f1-score': 0.9975286510019847}} | Confusion_matrix | int64 | (5, 5) | [[16177, 0, 0, 0, 88], [2, 14967, 1, 1, 7], [0, 0, 16344, 0, 0], [0, 0, 0, 14764, 11], [76, 7, 0, 0, 15650]] | precision | float | 1 | 0.9952014764687789 |
| Confusion_matrix | int64 | (5, 5) | [[16177, 0, 0, 0, 88], [2, 14967, 1, 1, 7], [0, 0, 16344, 0, 0], [0, 0, 0, 14764, 11], [76, 7, 0, 0, 15650]] | precision | float | 1 | 0.9952014764687789 | | | | |

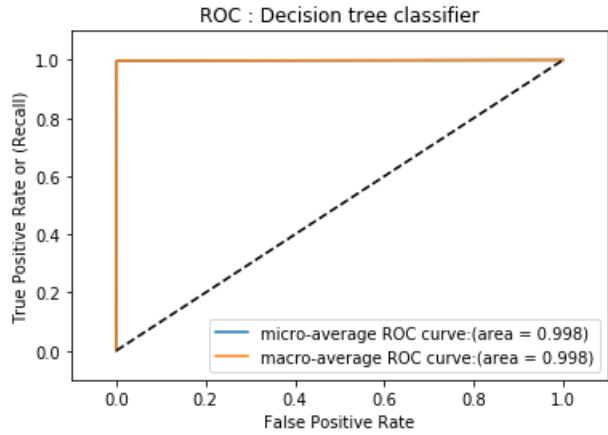
| Confusion_matrix - NumPy array | | | | | Class 2 - Dictionary (4 elements) | | | |
|--------------------------------|-------|-------|-------|-------|-----------------------------------|-------|------|---------------------------|
| | | | | | Key | Type | Size | Value |
| 0 | 1 | 2 | 3 | 4 | f1-score | float | 1 | 0.9948954489544896 |
| 0 | 16177 | 0 | 0 | 88 | precision | float | 1 | 0.9952014764687789 |
| 1 | 2 | 14967 | 1 | 7 | recall | float | 1 | 0.994589609591146 |
| 2 | 0 | 0 | 16344 | 0 | support | int | 1 | 16265 |
| 3 | 0 | 0 | 0 | 14764 | | | | |
| 4 | 76 | 7 | 0 | 15650 | | | | |

| macro avg - Dictionary (4 elements) | | | | | Class 3 - Dictionary (4 elements) | | | |
|-------------------------------------|-------|---|---------------------------|-----------|-----------------------------------|------|---------------------------|-------|
| | | | | | Key | Type | Size | Value |
| f1-score | float | 1 | 0.9975711142546684 | f1-score | float | 1 | 0.9993990384615384 | |
| precision | float | 1 | 0.9975754990296937 | precision | float | 1 | 0.9995325230399359 | |
| recall | float | 1 | 0.9975670325555605 | recall | float | 1 | 0.9992655895313126 | |
| support | int | 1 | 78095 | support | int | 1 | 14978 | |

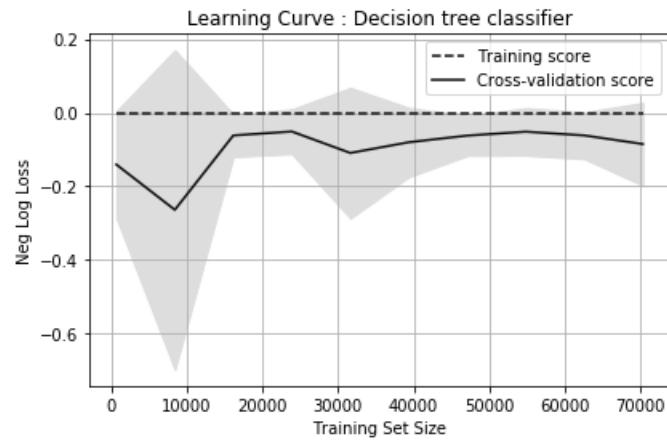
| micro avg - Dictionary (4 elements) | | | | | Class 4 - Dictionary (4 elements) | | | |
|-------------------------------------|-------|---|---------------------------|-----------|-----------------------------------|------|---------------------------|-------|
| | | | | | Key | Type | Size | Value |
| f1-score | float | 1 | 0.9975286510019847 | f1-score | float | 1 | 0.999969408669583 | |
| precision | float | 1 | 0.9975286510019847 | precision | float | 1 | 0.9999388192107678 | |
| recall | float | 1 | 0.9975286510019847 | recall | float | 1 | 1.0 | |
| support | int | 1 | 78095 | support | int | 1 | 16344 | |

| weighted avg - Dictionary (4 elements) | | | | | Class 5 - Dictionary (4 elements) | | | |
|--|-------|---|---------------------------|-----------|-----------------------------------|------|---------------------------|-------|
| | | | | | Key | Type | Size | Value |
| f1-score | float | 1 | 0.9975291668298457 | f1-score | float | 1 | 0.9995937711577522 | |
| precision | float | 1 | 0.9975299860870034 | precision | float | 1 | 0.9999322722654927 | |
| recall | float | 1 | 0.9975286510019847 | recall | float | 1 | 0.9992554991539763 | |
| support | int | 1 | 78095 | support | int | 1 | 14775 | |

The ROC curve and AUC.



The learning curve based on neg log loss.

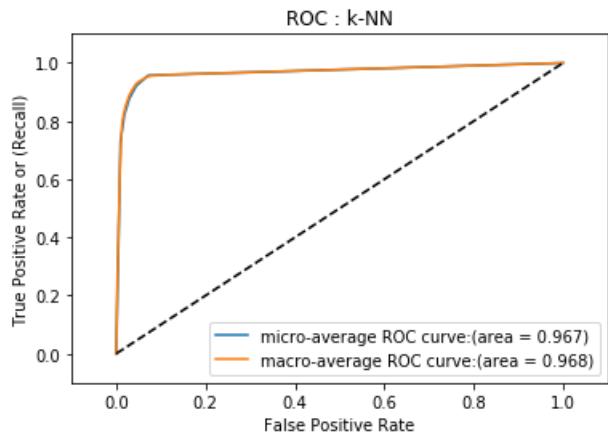


2.2.3 k-NN

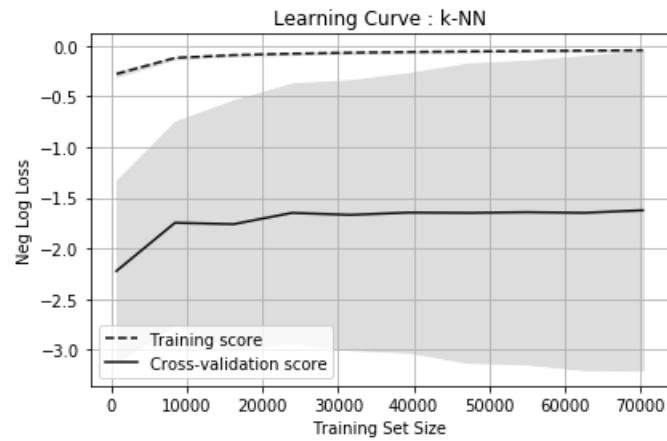
The confusion matrix, accuracy, precision, recall and F-1 of each class.

| k-NN - Dictionary (3 elements) | | | | | Class 1 - Dictionary (4 elements) | | | | |
|--|---------|--------|------------------------------------|--|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| Accuracy | float64 | 1 | 0.8786734105896664 | | f1-score | float | 1 | 0.9462638172439205 | |
| Classification_report | dict | 8 | {'class 1': {'precision': 0.9 | | precision | float | 1 | 0.9088901472253681 | |
| Confusion_matrix | int64 | (5, 5) | [[16051 48 166 0 [34 14346 9 7 | | recall | float | 1 | 0.9868429142330157 | |
| Confusion_matrix - NumPy array | | | | | support | int | 1 | 16265 | |
| | | | | | Class 2 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.8818539463978362 | | f1-score | float | 1 | 0.8170634468618294 | |
| precision | float | 1 | 0.8817221739649505 | | precision | float | 1 | 0.9578047803445052 | |
| recall | float | 1 | 0.8793969665779983 | | recall | float | 1 | 0.7897087616250612 | |
| support | int | 1 | 78095 | | support | int | 1 | 14978 | |
| macro avg - Dictionary (4 elements) | | | | | Class 3 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.8770204181782191 | | f1-score | float | 1 | 0.8308336015448985 | |
| precision | float | 1 | 0.8817221739649505 | | precision | float | 1 | 0.8764769794920548 | |
| recall | float | 1 | 0.8793969665779983 | | recall | float | 1 | 0.7897087616250612 | |
| support | int | 1 | 78095 | | support | int | 1 | 16344 | |
| micro avg - Dictionary (4 elements) | | | | | Class 4 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.8786734105896664 | | f1-score | float | 1 | 0.8669069550126819 | |
| precision | float | 1 | 0.8786734105896664 | | precision | float | 1 | 0.8550925011521496 | |
| recall | float | 1 | 0.8786734105896664 | | recall | float | 1 | 0.8790524534686971 | |
| support | int | 1 | 78095 | | support | int | 1 | 14775 | |
| weighted avg - Dictionary (4 elements) | | | | | Class 5 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.8772080963553975 | | f1-score | float | 1 | 0.8592437706917582 | |
| precision | float | 1 | 0.8828179818075667 | | precision | float | 1 | 0.9510877950933498 | |
| recall | float | 1 | 0.8786734105896664 | | recall | float | 1 | 0.7835759232187123 | |
| support | int | 1 | 78095 | | support | int | 1 | 15733 | |

The ROC curve and AUC.



The learning curve based on neg log loss.

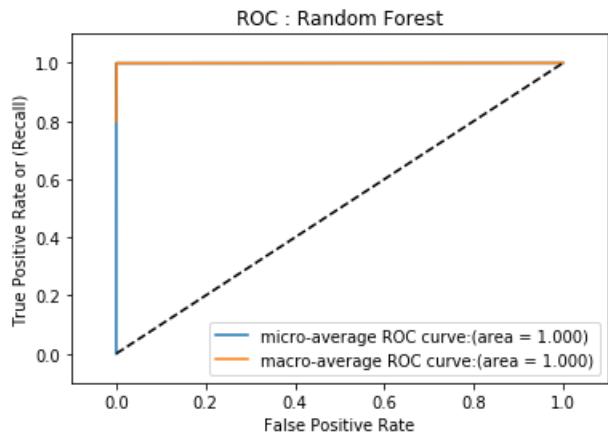


2.2.4 Random Forest

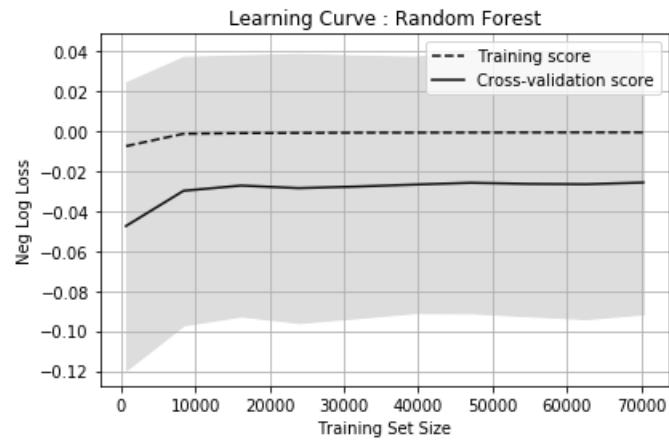
The confusion matrix, accuracy, precision, recall and F-1 of each class.

| Random Forest - Dictionary (3 elements) | | | | Class 1 - Dictionary (4 elements) | | | | |
|---|---------|--------|---|-----------------------------------|-----------------------------------|-------|--------------------|--------------------|
| Key | Type | Size | Value | Key | Type | Size | Value | |
| Accuracy | float64 | 1 | 0.9990908508867405 | f1-score | float | 1 | 0.998245668031147 | |
| Classification_report | dict | 8 | {'class 1': {'precision': 0.9990908508867405, 'recall': 0.9970488779588073, 'f1-score': 0.998245668031147, 'support': 16265}, 'class 0': {'precision': 0.999453346480956, 'recall': 0.9997328703085347, 'f1-score': 0.9997328703085347, 'support': 14978}, 'macro avg': {'precision': 0.9991052097648645, 'recall': 0.9991073247731649, 'f1-score': 0.9991052097648645, 'support': 78095}, 'weighted avg': {'precision': 0.9990908508867405, 'recall': 0.9990908508867405, 'f1-score': 0.9990908508867405, 'support': 78095}}, 'Confusion_matrix': int64 (5, 5) [[16217 0 0 0 48] [0 14970 3 0 0]]} | precision | float | 1 | 0.999453346480956 | |
| Confusion_matrix | int64 | (5, 5) | [[16217 0 0 0 48] [0 14970 3 0 0]] | recall | float | 1 | 0.9970488779588073 | |
| | | | Confusion_matrix - NumPy array | support | int | 1 | 16265 | |
| Confusion_matrix - NumPy array | | | | | | | | |
| 0 | 1 | 2 | 3 | 4 | | | | |
| 0 | 16217 | 0 | 0 | 48 | Class 2 - Dictionary (4 elements) | | | |
| 1 | 0 | 14970 | 3 | 5 | f1-score | float | 1 | 0.9997328703085347 |
| 2 | 0 | 0 | 16344 | 0 | precision | float | 1 | 1.0 |
| 3 | 0 | 0 | 0 | 14769 | recall | float | 1 | 0.9994658832955 |
| 4 | 9 | 0 | 0 | 15724 | support | int | 1 | 14978 |
| macro avg - Dictionary (4 elements) | | | | | | | | |
| Key | Type | Size | Value | Class 3 - Dictionary (4 elements) | | | | |
| f1-score | float | 1 | 0.9991052097648645 | f1-score | float | 1 | 0.9999082316233826 | |
| precision | float | 1 | 0.9991047230815587 | precision | float | 1 | 0.9998164800880895 | |
| recall | float | 1 | 0.9991073247731649 | recall | float | 1 | 1.0 | |
| support | int | 1 | 78095 | support | int | 1 | 16344 | |
| micro avg - Dictionary (4 elements) | | | | | | | | |
| Key | Type | Size | Value | Class 4 - Dictionary (4 elements) | | | | |
| f1-score | float | 1 | 0.9990908508867405 | f1-score | float | 1 | 0.9997969130787977 | |
| precision | float | 1 | 0.9990908508867405 | precision | float | 1 | 1.0 | |
| recall | float | 1 | 0.9990908508867405 | recall | float | 1 | 0.9995939086294416 | |
| support | int | 1 | 78095 | support | int | 1 | 14775 | |
| weighted avg - Dictionary (4 elements) | | | | | | | | |
| Key | Type | Size | Value | Class 5 - Dictionary (4 elements) | | | | |
| f1-score | float | 1 | 0.9990910838242535 | f1-score | float | 1 | 0.9978423657824598 | |
| precision | float | 1 | 0.9990929743079253 | precision | float | 1 | 0.9962618006716087 | |
| recall | float | 1 | 0.9990908508867405 | recall | float | 1 | 0.9994279539820758 | |
| support | int | 1 | 78095 | support | int | 1 | 15733 | |

The ROC curve and AUC.



The learning curve based on neg log loss.

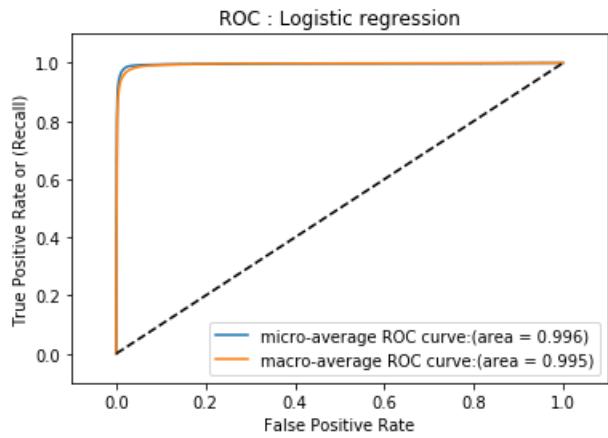


2.2.5 Logistic regression

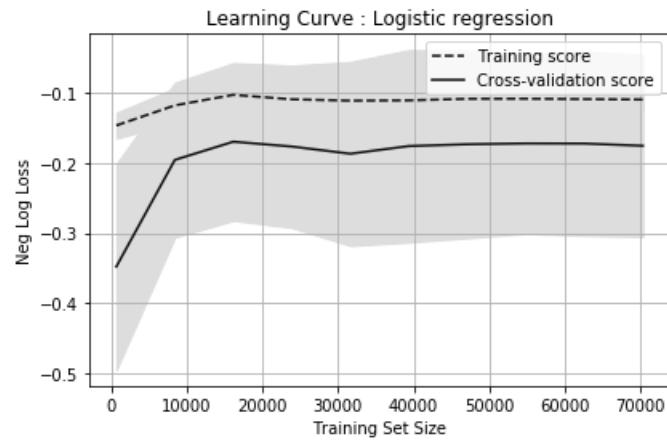
The confusion matrix, accuracy, precision, recall and F-1 of each class.

| Logistic regression - Dictionary (3 elements) | | | | Class 1 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---------|--------|---|-----------------------------------|-------|------|--------------------|-----|------|------|-------|----------|-------|---|--------------------|-----------|-------|---|--------------------|--------|-------|-------|--------------------|---------|-----|---|-------|---|-------|----|---|---|---|---|-----|-------|----|---|----|------|----|----|-------|
| Key | Type | Size | Value | Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Accuracy | float64 | 1 | 0.9615980536526026 | f1-score | float | 1 | 0.9967416697405631 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Classification_report | dict | 8 | {'class 1': {'precision': 0.99100938036728763, 'recall': 0.9968029511220412, 'f1-score': 0.9967416697405631, 'support': 16265}, 'class 0': {'precision': 0.9615980536526026, 'recall': 0.9615980536526026, 'f1-score': 0.9615980536526026, 'support': 78095}, 'macro avg': {'precision': 0.9609575763156517, 'recall': 0.9615980536526026, 'f1-score': 0.9609575763156517, 'support': 78095}, 'weighted avg': {'precision': 0.9615980536526026, 'recall': 0.9615980536526026, 'f1-score': 0.9615980536526026, 'support': 78095}}, 'Confusion_matrix': int64 (5, 5) [[16213 48 4 0 0], [21 14206 5 0 0], [6 0 16264 70 4], [8 0 115 14636 16], [19 1866 10 61 13777]}] | precision | float | 1 | 0.996803958935267 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Confusion_matrix | int64 | (5, 5) | [[16213 48 4 0 0], [21 14206 5 0 0], [6 0 16264 70 4], [8 0 115 14636 16], [19 1866 10 61 13777]] | recall | float | 1 | 0.9968029511220412 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | support | int | 1 | 16265 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Confusion_matrix - NumPy array | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th></th> <th>0</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>16213</td> <td>48</td> <td>4</td> <td>0</td> <td>0</td> </tr> <tr> <td>1</td> <td>21</td> <td>14206</td> <td>5</td> <td>0</td> <td>746</td> </tr> <tr> <td>2</td> <td>6</td> <td>0</td> <td>16264</td> <td>70</td> <td>4</td> </tr> <tr> <td>3</td> <td>8</td> <td>0</td> <td>115</td> <td>14636</td> <td>16</td> </tr> <tr> <td>4</td> <td>19</td> <td>1866</td> <td>10</td> <td>61</td> <td>13777</td> </tr> </tbody> </table> | | | | | | | | | 0 | 1 | 2 | 3 | 4 | 0 | 16213 | 48 | 4 | 0 | 0 | 1 | 21 | 14206 | 5 | 0 | 746 | 2 | 6 | 0 | 16264 | 70 | 4 | 3 | 8 | 0 | 115 | 14636 | 16 | 4 | 19 | 1866 | 10 | 61 | 13777 |
| | 0 | 1 | 2 | 3 | 4 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 0 | 16213 | 48 | 4 | 0 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 21 | 14206 | 5 | 0 | 746 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | 6 | 0 | 16264 | 70 | 4 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 8 | 0 | 115 | 14636 | 16 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | 19 | 1866 | 10 | 61 | 13777 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| macro avg - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9609575763156517</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9616463312889814</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9613266950440684</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>78095</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9609575763156517 | precision | float | 1 | 0.9616463312889814 | recall | float | 1 | 0.9613266950440684 | support | int | 1 | 78095 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9609575763156517 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9616463312889814 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9613266950440684 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 78095 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| micro avg - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9615980536526026</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9615980536526026</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9615980536526026</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>78095</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9615980536526026 | precision | float | 1 | 0.9615980536526026 | recall | float | 1 | 0.9615980536526026 | support | int | 1 | 78095 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9615980536526026 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9615980536526026 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9615980536526026 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 78095 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| weighted avg - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9615464173257661</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9625365616299422</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9615980536526026</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>78095</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9615464173257661 | precision | float | 1 | 0.9625365616299422 | recall | float | 1 | 0.9615980536526026 | support | int | 1 | 78095 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9615464173257661 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9625365616299422 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9615980536526026 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 78095 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Class 1 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9967416697405631</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.996803958935267</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9968029511220412</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>16265</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9967416697405631 | precision | float | 1 | 0.996803958935267 | recall | float | 1 | 0.9968029511220412 | support | int | 1 | 16265 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9967416697405631 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.996803958935267 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9968029511220412 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 16265 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Class 2 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9136278860376873</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.8812655086848635</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9484577380157564</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>14978</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9136278860376873 | precision | float | 1 | 0.8812655086848635 | recall | float | 1 | 0.9484577380157564 | support | int | 1 | 14978 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9136278860376873 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.8812655086848635 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9484577380157564 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 14978 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Class 3 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9934640522875818</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9918282717404562</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9951052373959863</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>16344</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9934640522875818 | precision | float | 1 | 0.9918282717404562 | recall | float | 1 | 0.9951052373959863 | support | int | 1 | 16344 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9934640522875818 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9918282717404562 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9951052373959863 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 16344 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Class 4 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9908604698395505</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9911288684228347</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.9905922165820643</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>14775</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9908604698395505 | precision | float | 1 | 0.9911288684228347 | recall | float | 1 | 0.9905922165820643 | support | int | 1 | 14775 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9908604698395505 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9911288684228347 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.9905922165820643 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 14775 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Class 5 - Dictionary (4 elements) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <thead> <tr> <th>Key</th> <th>Type</th> <th>Size</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td>f1-score</td> <td>float</td> <td>1</td> <td>0.9100938036728763</td> </tr> <tr> <td>precision</td> <td>float</td> <td>1</td> <td>0.9473286117032249</td> </tr> <tr> <td>recall</td> <td>float</td> <td>1</td> <td>0.8756753321044938</td> </tr> <tr> <td>support</td> <td>int</td> <td>1</td> <td>15733</td> </tr> </tbody> </table> | | | | | | | | Key | Type | Size | Value | f1-score | float | 1 | 0.9100938036728763 | precision | float | 1 | 0.9473286117032249 | recall | float | 1 | 0.8756753321044938 | support | int | 1 | 15733 | | | | | | | | | | | | | | | | |
| Key | Type | Size | Value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| f1-score | float | 1 | 0.9100938036728763 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| precision | float | 1 | 0.9473286117032249 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| recall | float | 1 | 0.8756753321044938 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| support | int | 1 | 15733 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

The ROC curve and AUC.



The learning curve based on neg log loss.



2.2.6 AdaBoost

The confusion matrix, accuracy, precision, recall and F-1 of each class.

AdaBoost - Dictionary (3 elements)

| Key | Type | Size | Value |
|-----------------------|---------|--------|--|
| Accuracy | float64 | 1 | 0.46064408732953455 |
| Classification_report | dict | 8 | {'Class 1': {'precision': 0.358995, ...} |
| Confusion_matrix | int64 | (5, 5) | [[16265, 0, 0, 0, 0], [1, 14975, 0, 0, 2]] |

Confusion_matrix - NumPy array

| | 0 | 1 | 2 | 3 | 4 |
|---|-------|-------|------|------|-------|
| 0 | 16265 | 0 | 0 | 0 | 0 |
| 1 | 1 | 14975 | 0 | 0 | 2 |
| 2 | 0 | 0 | 3268 | 0 | 13076 |
| 3 | 13309 | 0 | 0 | 1466 | 0 |
| 4 | 15732 | 0 | 1 | 0 | 0 |

macro avg - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|---------------------------|
| f1-score | float | 1 | 0.4084007034756828 |
| precision | float | 1 | 0.6717378789587096 |
| recall | float | 1 | 0.4597944833632404 |
| support | int | 1 | 78095 |

micro avg - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|----------------------------|
| f1-score | float | 1 | 0.46064408732953455 |
| precision | float | 1 | 0.46064408732953455 |
| recall | float | 1 | 0.46064408732953455 |
| support | int | 1 | 78095 |

weighted avg - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|----------------------------|
| f1-score | float | 1 | 0.40570653571313303 |
| precision | float | 1 | 0.6649729027454423 |
| recall | float | 1 | 0.46064408732953455 |
| support | int | 1 | 78095 |

Class 1 - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|---------------------------|
| f1-score | float | 1 | 0.5283245631131034 |
| precision | float | 1 | 0.3589952987397091 |
| recall | float | 1 | 1.0 |
| support | int | 1 | 16265 |

Class 2 - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|---------------------------|
| f1-score | float | 1 | 0.9998998430875038 |
| precision | float | 1 | 1.0 |
| recall | float | 1 | 0.9997997062358125 |
| support | int | 1 | 14978 |

Class 3 - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|----------------------------|
| f1-score | float | 1 | 0.3332483556824555 |
| precision | float | 1 | 0.9996940960538391 |
| recall | float | 1 | 0.19995105237395985 |
| support | int | 1 | 16344 |

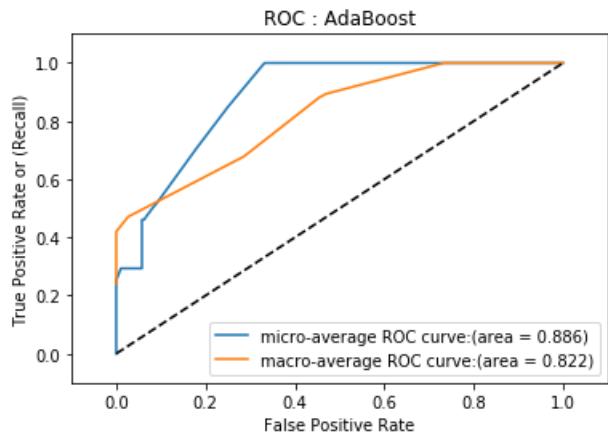
Class 4 - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|----------------------------|
| f1-score | float | 1 | 0.18053075549535127 |
| precision | float | 1 | 1.0 |
| recall | float | 1 | 0.09922165820642978 |
| support | int | 1 | 14775 |

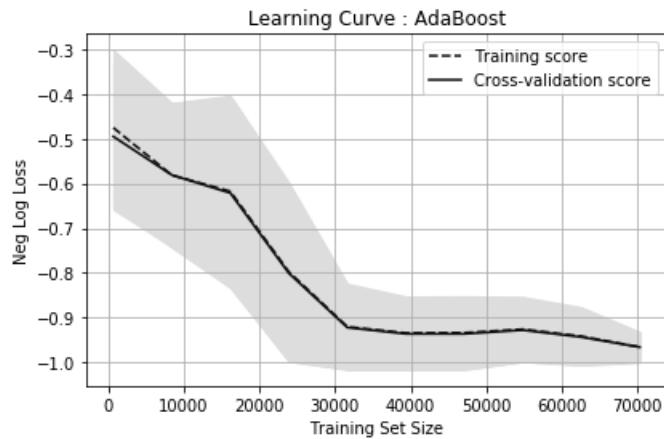
Class 5 - Dictionary (4 elements)

| Key | Type | Size | Value |
|-----------|-------|------|--------------|
| f1-score | float | 1 | 0.0 |
| precision | float | 1 | 0.0 |
| recall | float | 1 | 0.0 |
| support | int | 1 | 15733 |

The ROC curve and AUC.



The learning curve based on neg log loss.

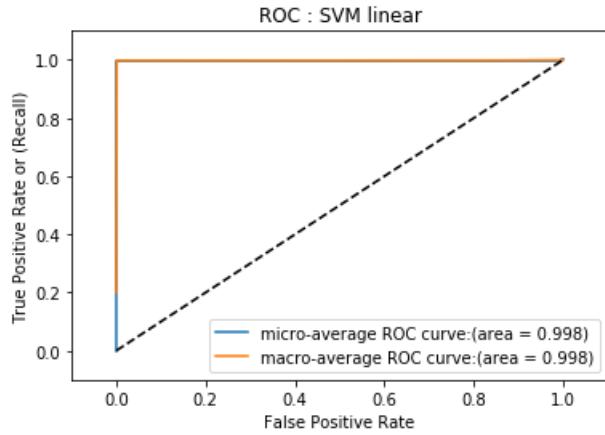


2.2.7 SVM (Support Vector Machine)

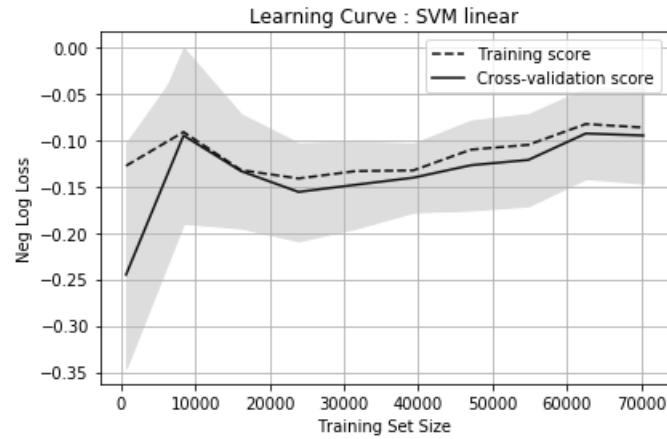
The confusion matrix, accuracy, precision, recall and F-1 of each class.

| SVM linear - Dictionary (3 elements) | | | | | Class 1 - Dictionary (4 elements) | | | | |
|--|---------|--------|--|--|-----------------------------------|-------|------|--------------------|--|
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| Accuracy | float64 | 1 | 0.9962737691273449 | | f1-score | float | 1 | 0.9972628017837921 | |
| Classification_report | dict | 8 | {'Class 1': {'precision': 0.9977230769230769, 'recall': 0.9968029511220412, 'f1-score': 0.9986634589681904, 'support': 16265}} | | precision | float | 1 | 0.9977230769230769 | |
| Confusion_matrix | int64 | (5, 5) | [[16213 0 4 0 4], [19 14944 8 0 4], [18 0 16256 70 0], [0 0 105 14664 6], [0 6 0 0 15727]] | | recall | float | 1 | 0.9968029511220412 | |
| Confusion_matrix - NumPy array | | | | | support | int | 1 | 16265 | |
| | | | | | Class 2 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9986634589681904 | | f1-score | float | 1 | 0.9986634589681904 | |
| precision | float | 1 | 0.9995986622073578 | | precision | float | 1 | 0.9995986622073578 | |
| recall | float | 1 | 0.9977300040058753 | | recall | float | 1 | 0.9977300040058753 | |
| support | int | 1 | 14978 | | support | int | 1 | 14978 | |
| macro avg - Dictionary (4 elements) | | | | | Class 3 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9962786119558393 | | f1-score | float | 1 | 0.9937341443286365 | |
| precision | float | 1 | 0.9963087084434422 | | precision | float | 1 | 0.9928540890490442 | |
| recall | float | 1 | 0.9962509323792444 | | recall | float | 1 | 0.9946157611355849 | |
| support | int | 1 | 78095 | | support | int | 1 | 16344 | |
| micro avg - Dictionary (4 elements) | | | | | Class 4 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9962737691273449 | | f1-score | float | 1 | 0.9937315759156982 | |
| precision | float | 1 | 0.9962737691273449 | | precision | float | 1 | 0.9949789659383905 | |
| recall | float | 1 | 0.9962737691273449 | | recall | float | 1 | 0.99248730964467 | |
| support | int | 1 | 78095 | | support | int | 1 | 14775 | |
| weighted avg - Dictionary (4 elements) | | | | | Class 5 - Dictionary (4 elements) | | | | |
| Key | Type | Size | Value | | Key | Type | Size | Value | |
| f1-score | float | 1 | 0.9962735978102523 | | f1-score | float | 1 | 0.9980010787828789 | |
| precision | float | 1 | 0.9962758209321964 | | precision | float | 1 | 0.9963887480993411 | |
| recall | float | 1 | 0.9962737691273449 | | recall | float | 1 | 0.9996186359880506 | |
| support | int | 1 | 78095 | | support | int | 1 | 15733 | |

The ROC curve and AUC.



The learning curve based on neg log loss.



3 Conclusion

3.1 Census Income or Adult dataset

Compare to all the performance measures, the top two picks are **Random Forest** and **Adaboost**.

Naïve Bayesian has a low accuracy and low auc compared to the picks. The main drawback is the learning curve divergence.

Decision tree classifier(entropy) has a little higher accuracy than Naïve Bayesian, but still lower than the picks. The auc is too low and the cross validation score of neg log loss of learning curve is too large to be considered.

k-NN has a little better accuracy and auc than Naïve Bayesian and Decision Tree classifier, but still lower than the picks. The auc value and the cross validation score of neg log loss are also lower than the picks.

Random Forest has the second highest score of each type.

Logistic regression has the second lowest score of accuracy. The auc score is better than Decision tree, but the variance is very high for the learning curve.

AdaBoost has the highest score of accuracy, auc. The score of learning curve is the second highest with low variance and it converges.

SVM linear has lower accuracy than the picks and the lowest auc score which is not included in consideration.

3.2 MoCap Hand Postures Data Set

Compare to all the performance measures, the top two picks are **Naïve Bayesian** and **Random Forest**.

Although, k-NN, logistic regression and SVM linear have very high scores of each type(not highest), their computation time is way longer than other classifiers. Therefore, these three classifiers are not in the consideration.

Adaboost has the lowest scores. Moreover, the learning curve divergence given the training size.

Decision tree classifier(entropy) has very high scores of each type, but since random forest is a lot of more decision trees. It has better scores than decision tree.