A Neural Network Model for Detecting Computer Networking Intrusions or Attacks

Adam Leonard Hubble

Problem Definition

As the study aim of the project undertaken, the development of a Neural Network (NN) model was anticipated for detecting and classifying computer-networking intrusions, as an Intrusion Detection System (IDS) [1]; for which the NN model configured, should be capable of differentiating between 'bad' connections, namely intrusions or attacks, and 'good', ordinary connections.

As the "most widely used data set for the evaluation of these systems" [2], the KDD Cup 1999 dataset [3] is applicated to both train and test the NN model configured for the problem domain bespoken. The dataset is comprised of a "standard set of data to be audited", which includes a wide variety of intrusions simulated in a military network environment. Instructed for the purposes of the IDS proposed, a subset of the database is applied to the model, representing ten-percent of the original quantity. Thereby containing '41' columns or features, and '494021' rows of network packet data; the columns are representative of each connection's features [4], where the last of which features is representative of their intrusion or attack type [5]. Therein, exists '23' different classes of intrusion or attack, where each class of attack is associated with a category, namely: Denial-of-Service (DOS), Remote-to-Local (R2L), User-to-Root (U2R) and Probe, whereas a good connection is associated with the category: 'normal'.

Dataset Pre-Processing

For implementation purposes, the Python [6] programming language was targeted as it "lets you work quickly and integrate systems more effectively", as well as being a predominant contender in the field of Machine Learning (ML). The dimensionality of the dataset is configured to be reduced using numerous approaches, dependant of the binary or multiclass classifier active for the IDS's configuration. Features that pose highly correlative natures with other features, that have missing field values (of 'null' type), or that can be identified by a single, constant value, are removed from the dataset. This aims to reduce the dimensional complexity of the dataset, and thus, the resultant amount of time required to train the model. For addressing feature removal, features can simply be 'dropped' from the dataset upon detecting missing values, value constancy or strong associations, or alternatively in compliance with a Singular Value Decomposition (SVD) technique, namely Principal Component Analysis (PCA); where the dataset when using said SVD, is projected into a lower dimensional space, whilst minimising information loss. For simplifying the multiclass classification capability of the relevant model, each of the '23' unique classes of intrusion or attack are mapped to their respective category, numerically, using a label encoding technique that enables each of said categories to be interpreted and processed by the model, as a one hot encoded vector; for binary classification, each of the categories are instead, manually assigned the following numerical values: DoS = 1, normal = 0, Probe = 1, R2L = 1 and U2R = 1. Notably, all categorical features are encoded to numerical values, as 'dummy' features, except the target category of each intrusion type or attack, within the multiclass configuration of the model. Proceeding from feature mapping, each feature's values are then relatively normalised for reducing their spatial complexities, to advance the efficiency of the models training procedure.

Experimental Design

Series of evaluative procedures were performed on both binary and multiclass configurations of the model derived. Said series concerned the trialling of feature removal techniques from the dataset, feature removal technique threshold value alteration, training and testing subset partition alteration, model architectural tuning concerning: active number of fully-connected layers (FCL's), number of neurons comprising each FCL and the active optimiser algorithm, as well as model hyperparameter tuning. Additional to all the experiments listed, K-fold Cross Validation (CV) is trialled for a range of dataset folds, to evaluate the performance of the model on a series of limited data samples, for better acknowledging and addressing the model's generalisation capabilities, when used to make predictions about samples of data that have not been used during training. For the exhaustive summary of model evaluation results, see *Appendix A* (multiclass classification model) and *Appendix B* (binary classification model). To note, cross-entropy loss functions were applicated to evaluate the performance of both binary and multiclass classification models explored.

Model Evaluation

Proceeding from the nature of the series of evaluative procedures conducted, it was found that the binary classification model configured, was capable of scoring a classification accuracy of '99.96%', which inevitably is adequate. Whereas the multiclass classification model configured, was capable of scoring a classification accuracy of '99.93%', that also is considered ample for purpose. For their performative visualisations across both training and validation procedures, line graphs were populated, which can be located within *Appendix C* for reference; each's performance was validated by the relevant testing subset partition assigned, via K-fold CV (for binary classification) and an independent dataset partitioning technique (for multiclass classification).

Conclusion

Summarily, the work submitted reflects upon the learning features of the study, such as the incorporation of data analysis techniques, data pre-processing techniques, dataset dimensionality reduction techniques, model configuration considerations and NN model performance evaluation methodologies.

Future Work

For future development purposes, it would be beneficial to explore hyperparameter tuning via optimisation algorithms, for acquiring parameter value optimality. As well, would trialling other NN's for classification and expanding model tuning.

References

[1] F. Sabahi and A. Movaghar, "Intrusion Detection: A Survey," *2008 Third International Conference on Systems and Networks Communications*, 2008, pp. 23-26, doi: 10.1109/ICSNC.2008.44.

[2] M. Tavallaee, E. Bagheri, W. Lu and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009, pp. 1-6, doi: 10.1109/CISDA.2009.5356528.

[3] KDD Cup 1999 (2021) *KDD Cup 1999 Data*. [Online] KDD Cup 1999. Available from: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html [Accessed: 27/05/21]

[4] KDD Cup 1999 (2021) *KDD Cup names*. [Online] KDD Cup 1999. Available from: http://kdd.ics.uci.edu/databases/kddcup99/kddcup.names [Accessed: 27/05/21]

[5] KDD Cup 1999 (2021) *Training attack types*. [Online] KDD Cup 1999. Available from: http://kdd.ics.uci.edu/databases/kddcup99/training_attack_types [Accessed: 27/05/21]

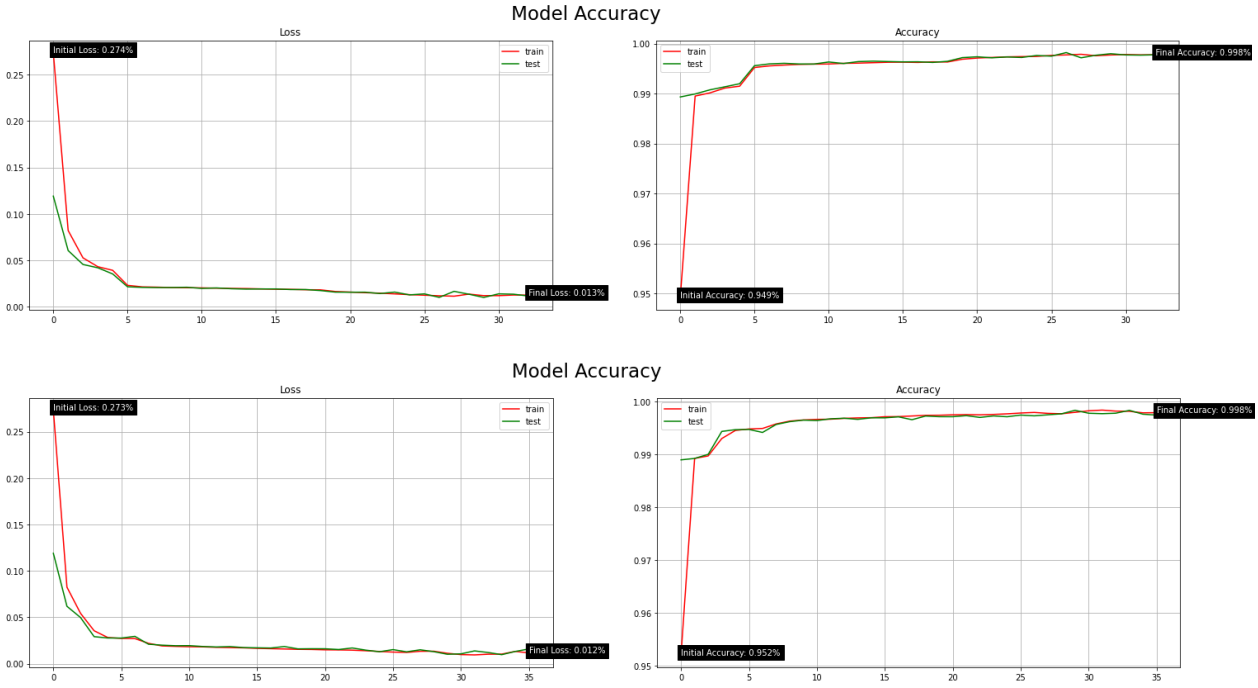
[6] Python (2001) *Welcome to Python*. [Online] Python Software Foundation. Available from: https://www.python.org/ [Accessed 27/05/21]

Appendices

Appendix A

Multiclass Classification Model Fine-tuning

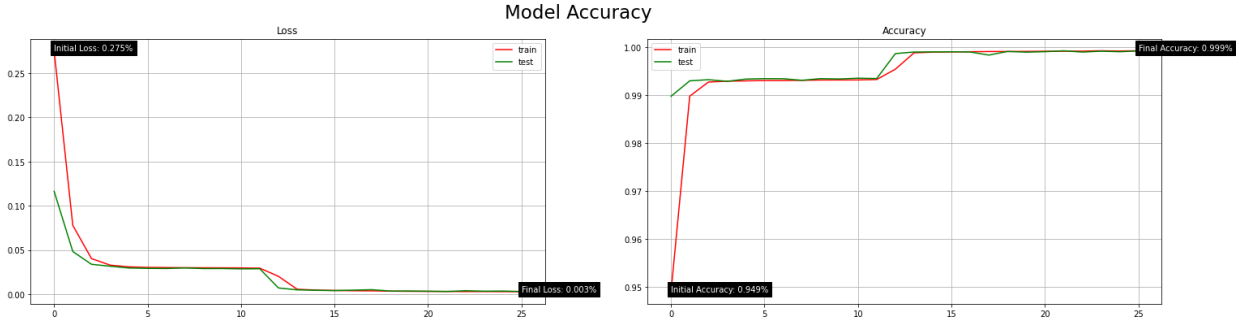
- Removing features of constant value from the dataset activeness alteration



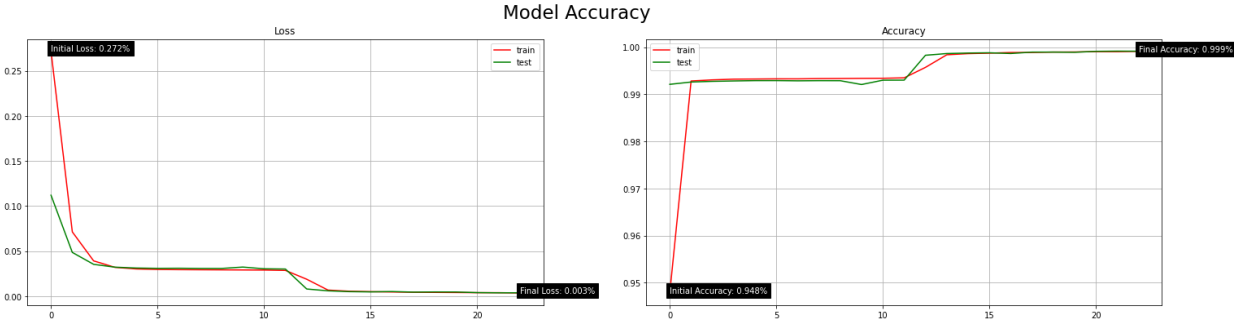
Removing Constant Features		
Removing Features?	Model Loss (Error)	Model Accuracy (%)
True	0.010	99.78%
False	0.016	99.72%

- Removing highly correlated features (Pearson Correlation Coefficient) threshold value alteration

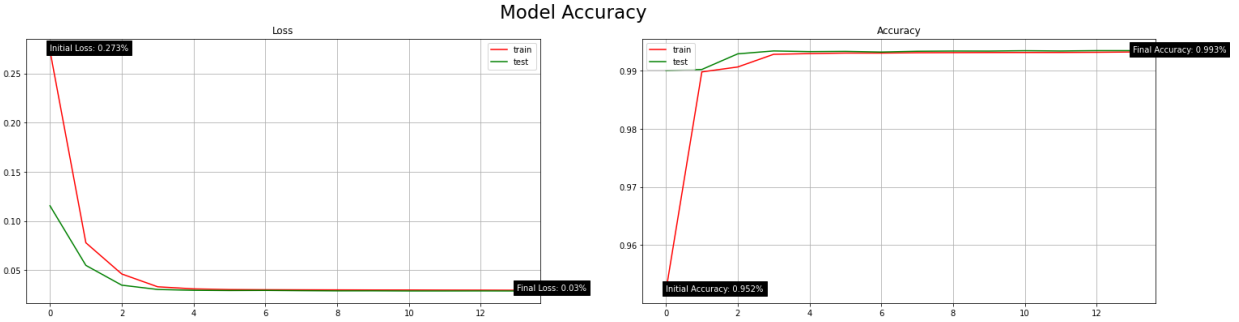
Model Accuracy



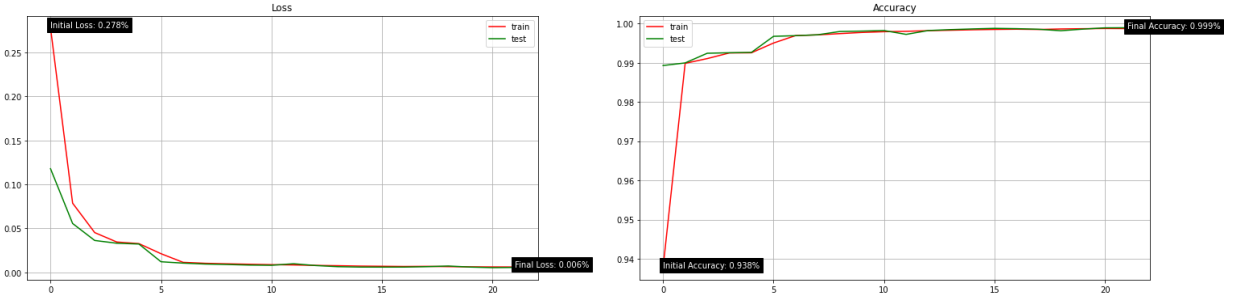
Model Accuracy



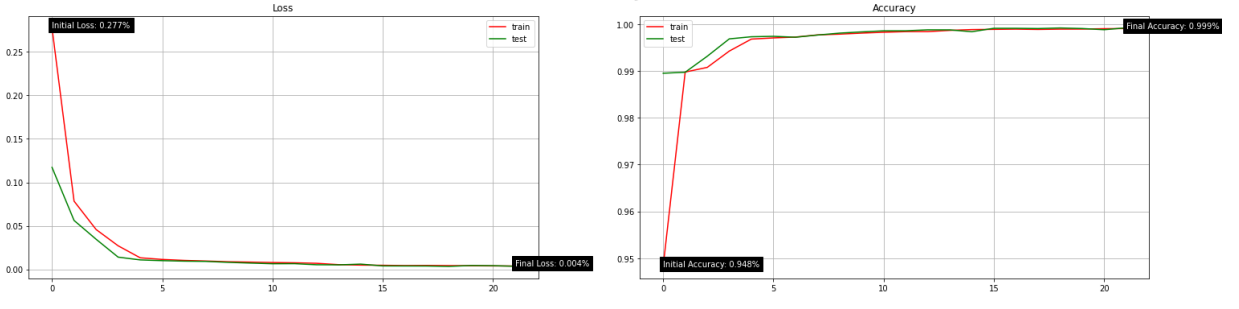
Model Accuracy

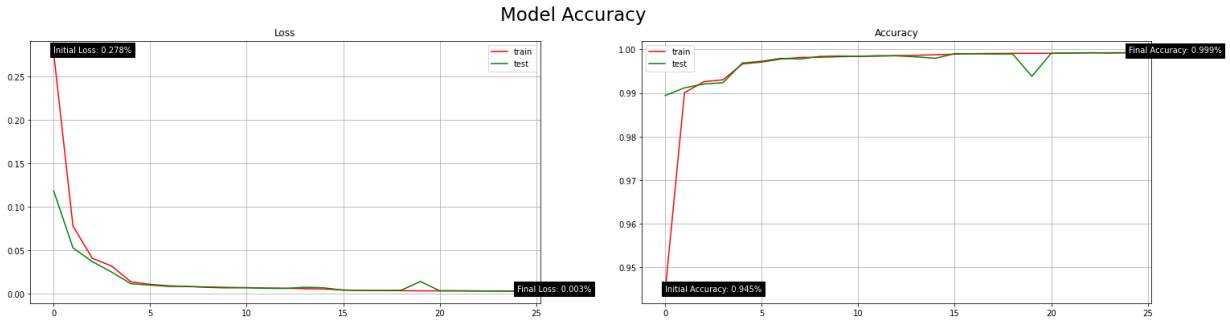


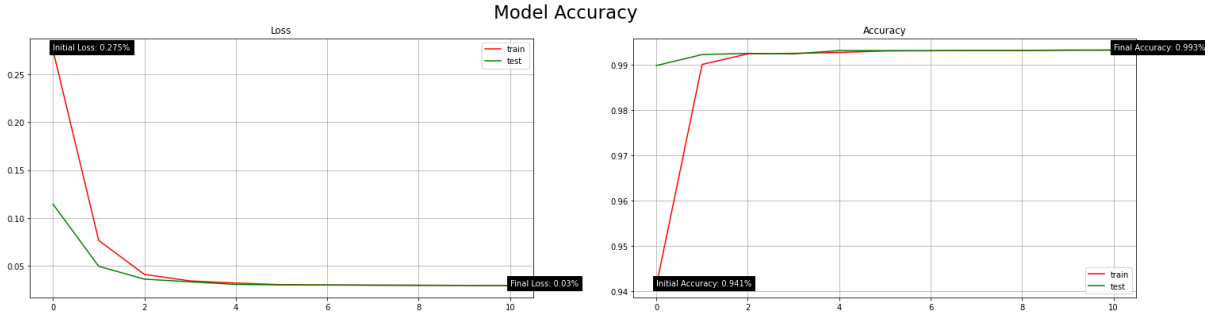
Model Accuracy

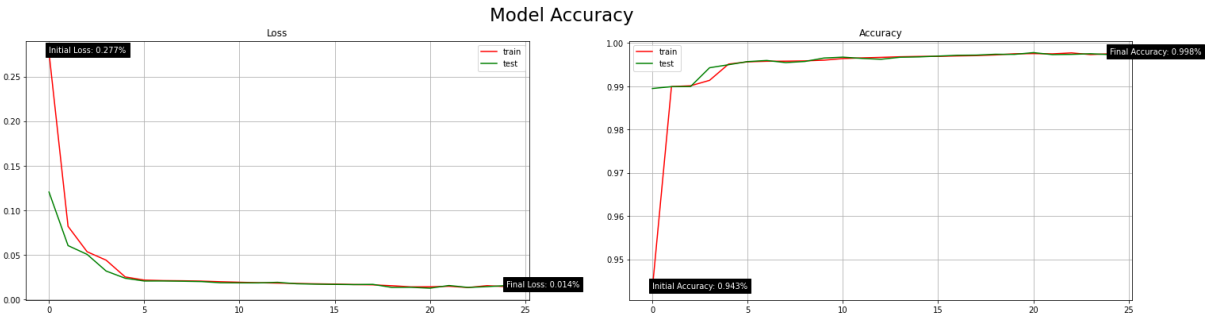


Model Accuracy



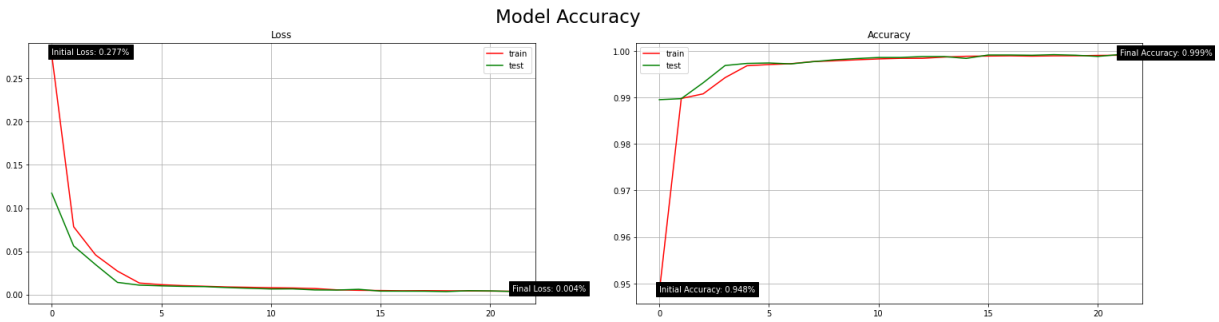


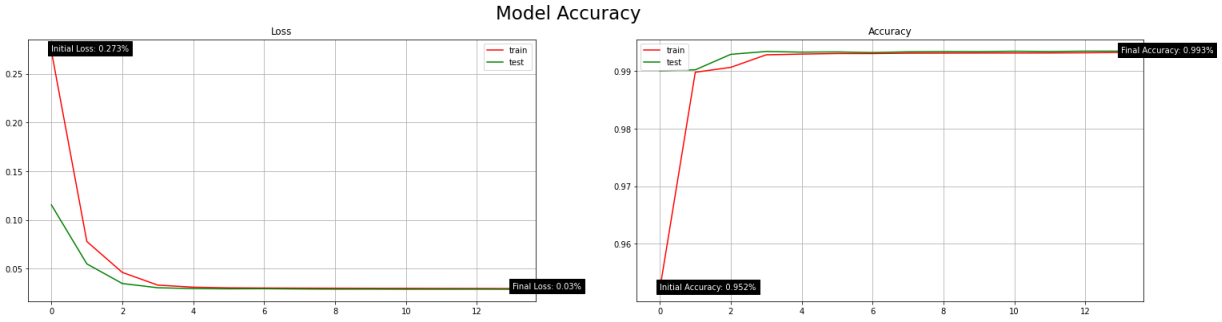




Removing Highly-Correlated Features		
Correlative Threshold	Model Loss (Error)	Model Accuracy (%)
0.95	0.003	99.90%
0.9	0.004	99.89%
0.85	0.029	99.34%
0.8	0.006	99.87%
0.75	0.004	99.91%
0.7	0.014	99.38%
0.65	0.030	99.31%
0.6	0.013	99.74%

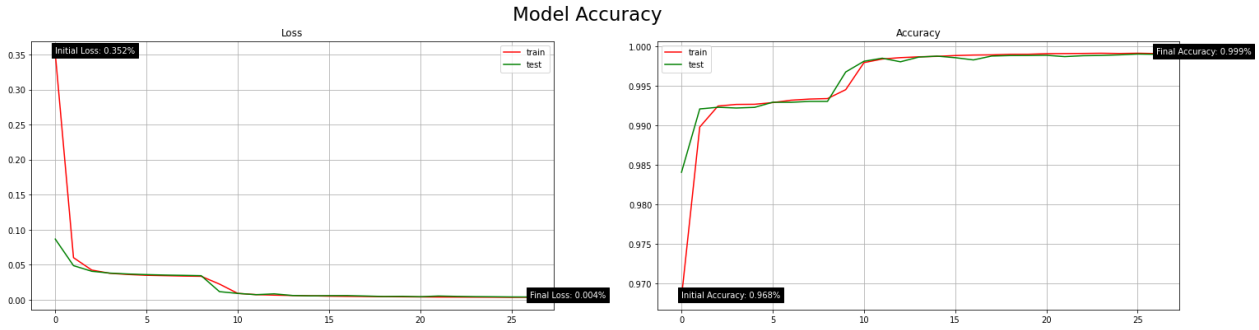
- Removing highly correlated features (Pearson Correlation Coefficient) activeness alteration

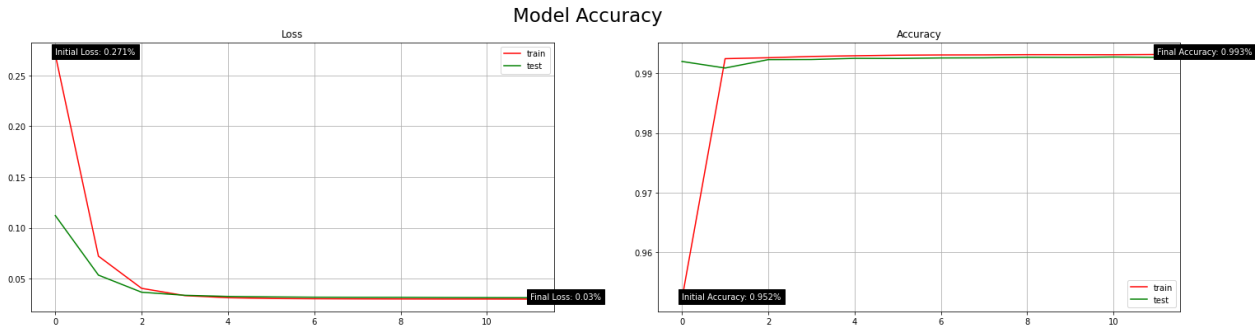


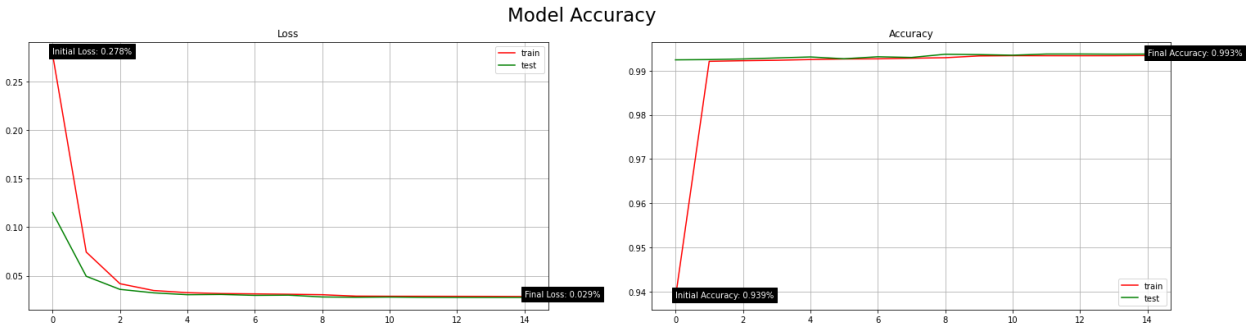


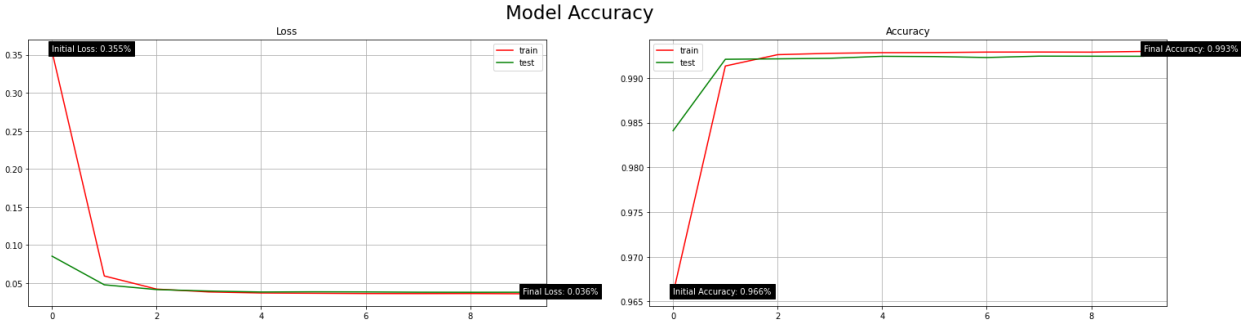
Removing Highly-Correlated Features		
Removing Features?	Model Loss (Error)	Model Accuracy (%)
True	0.004	99.91%
False	0.029	99.34%

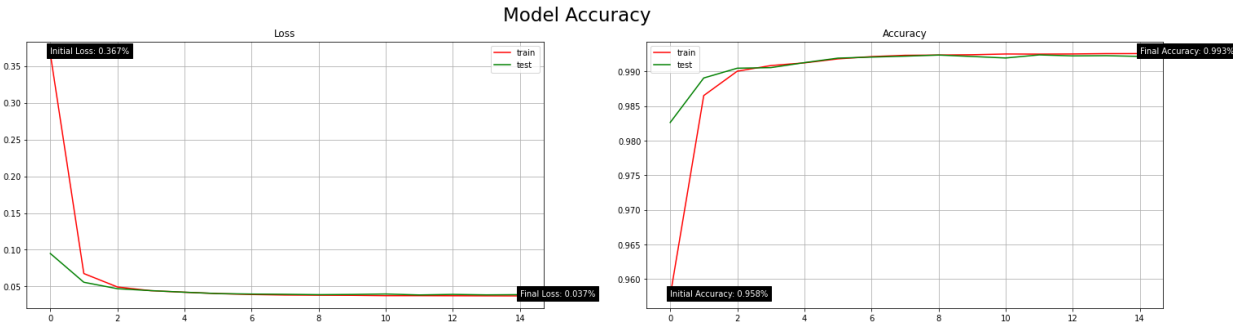
- Principal Component Analysis (PCA) number of components dataset reduced to (when correlation features not used) alteration

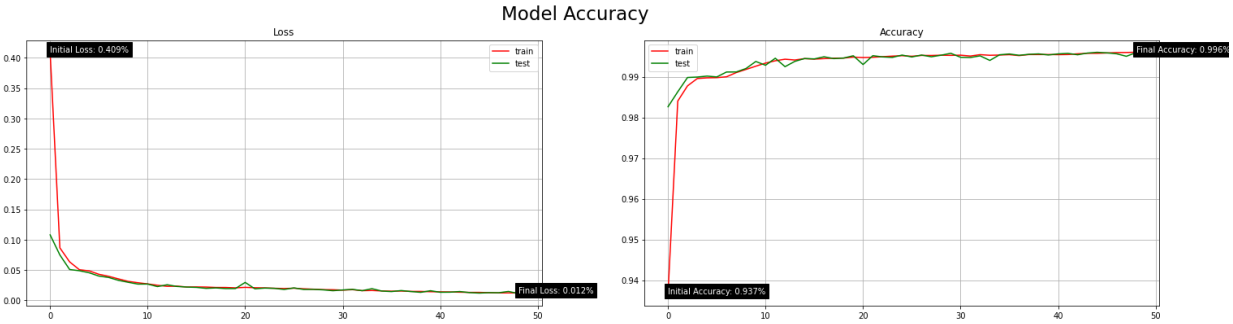






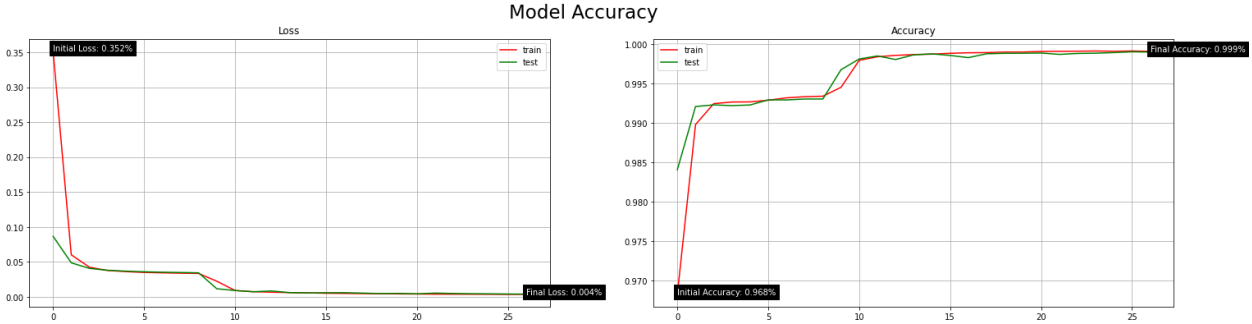


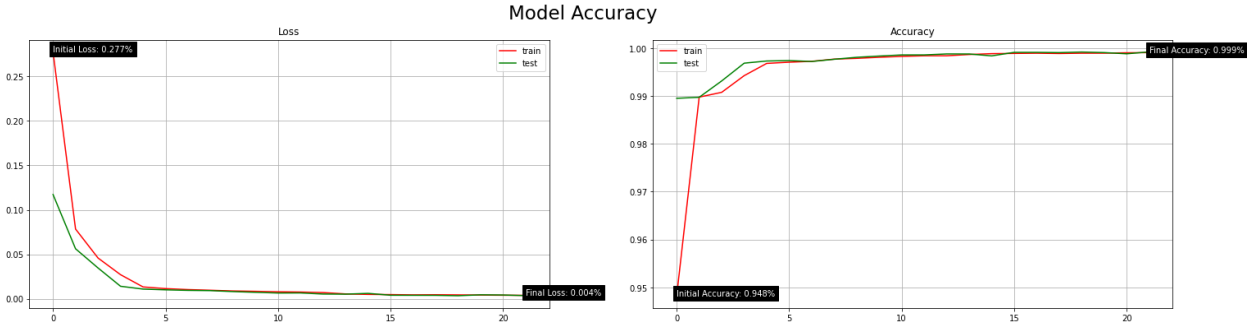




Principal Component Analysis (PCA)		
Dimensionality Reduction	Model Loss (Error)	Model Accuracy (%)
30	0.005	99.87%
25	0.031	99.26%
20	0.027	99.36%
15	0.038	99.24%
10	0.039	99.22%
5	0.012	99.58%

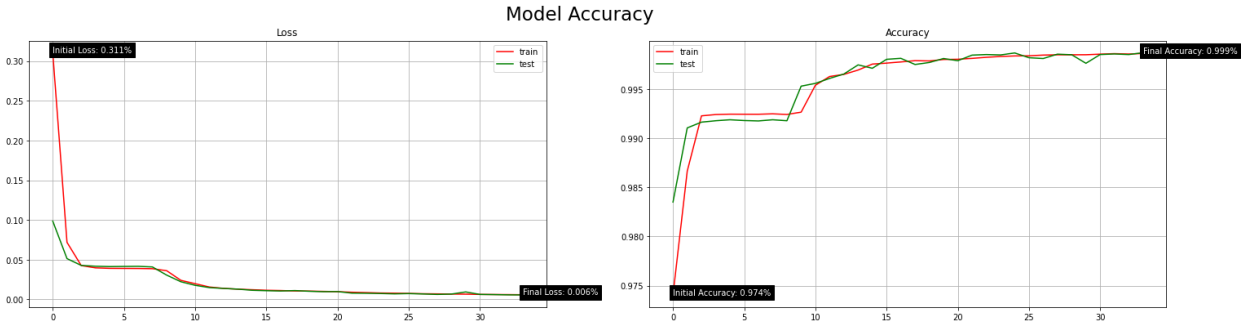
- Principal Component Analysis (PCA) activeness (when correlation features not used) alteration

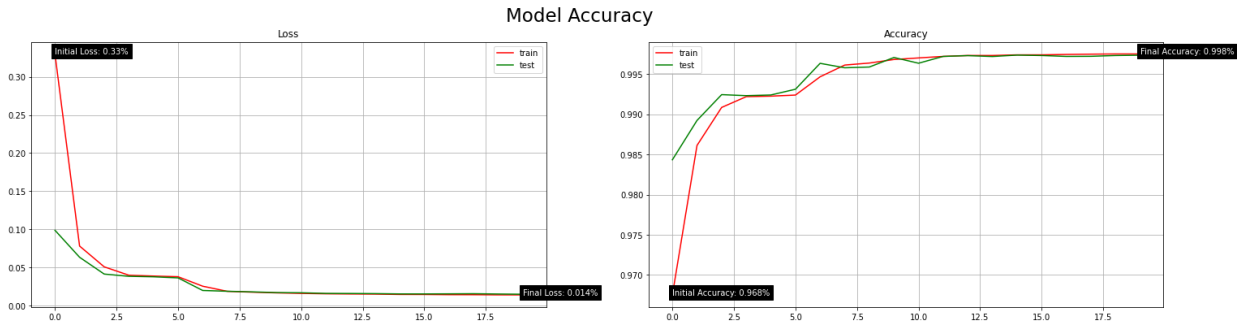


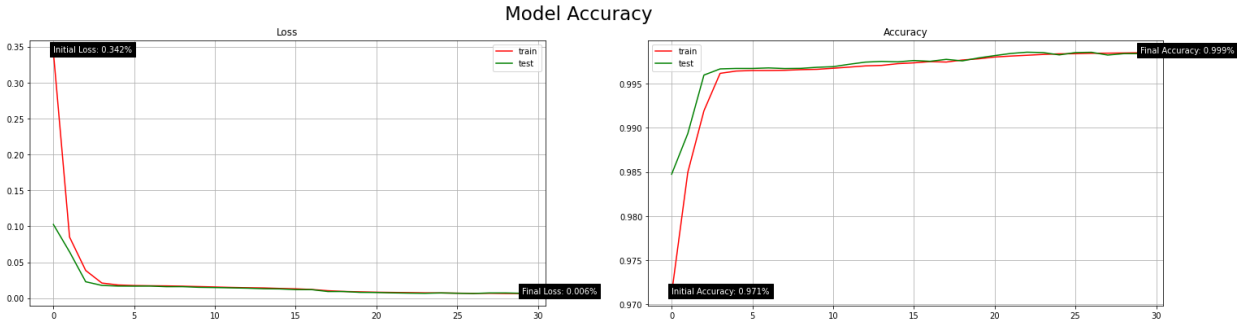


Principal Component Analysis (PCA)		
Active?	Model Loss (Error)	Model Accuracy (%)
True	0.005	99.87%
False	0.004	99.91%

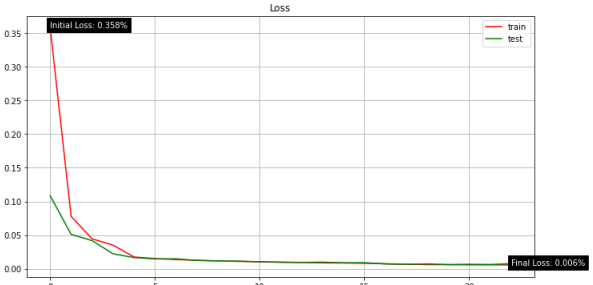
- Training and testing data split alteration

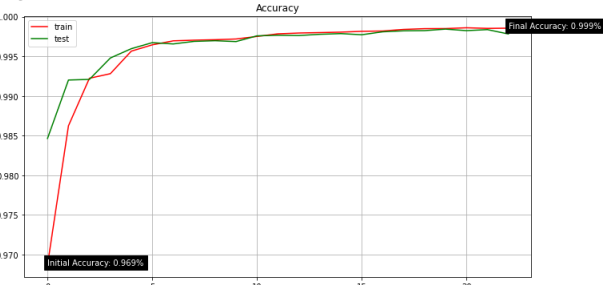




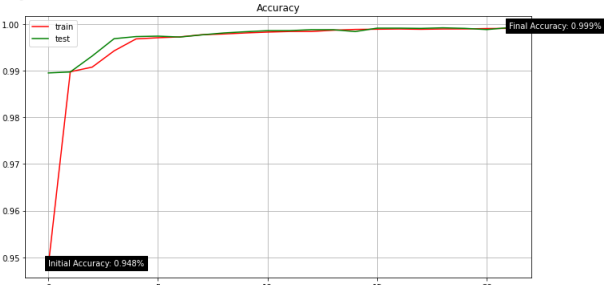


Model Accuracy

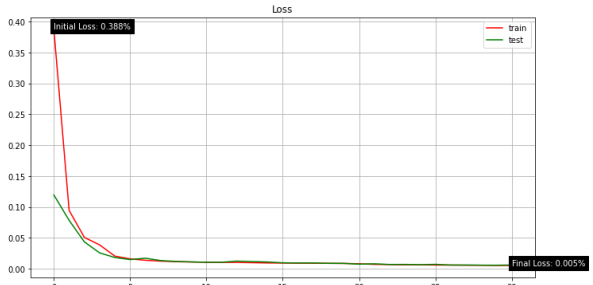


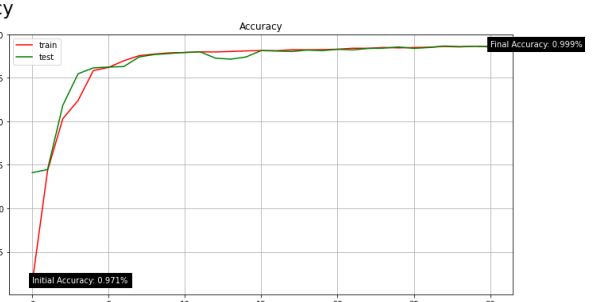


Model Accuracy

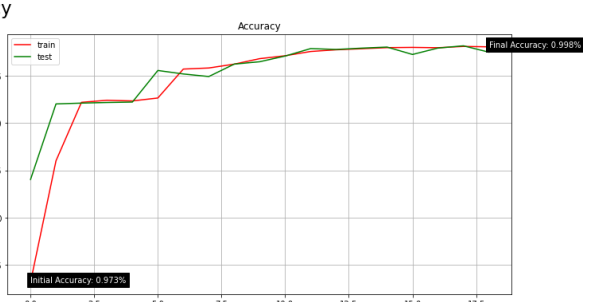


Model Accuracy

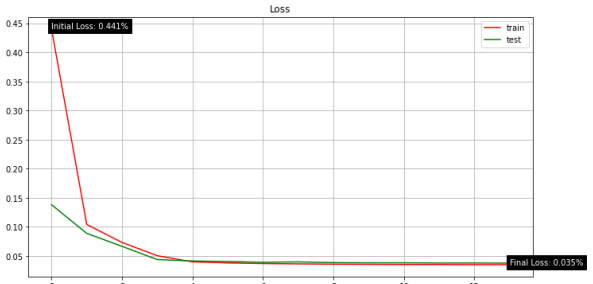


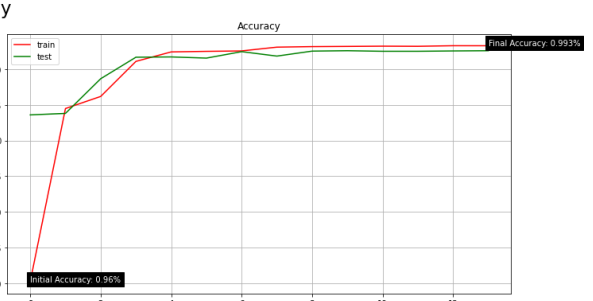


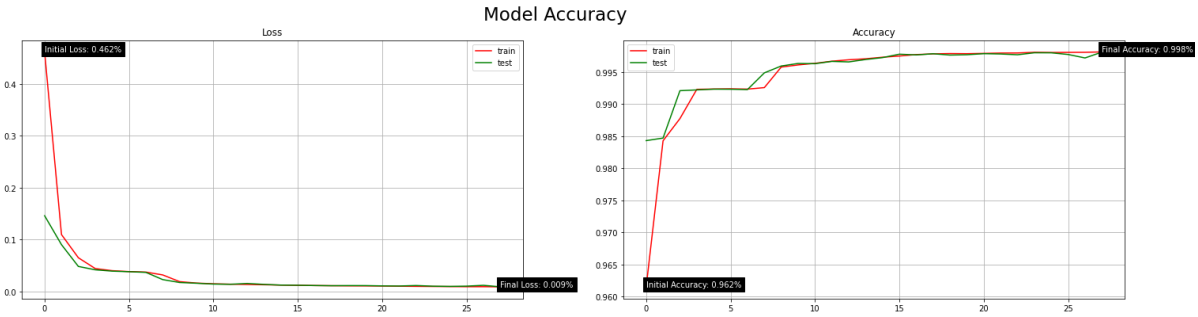
Model Accuracy

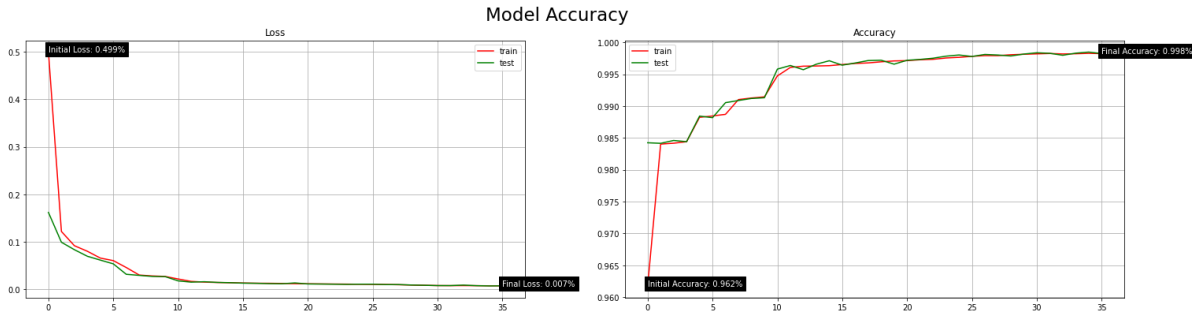


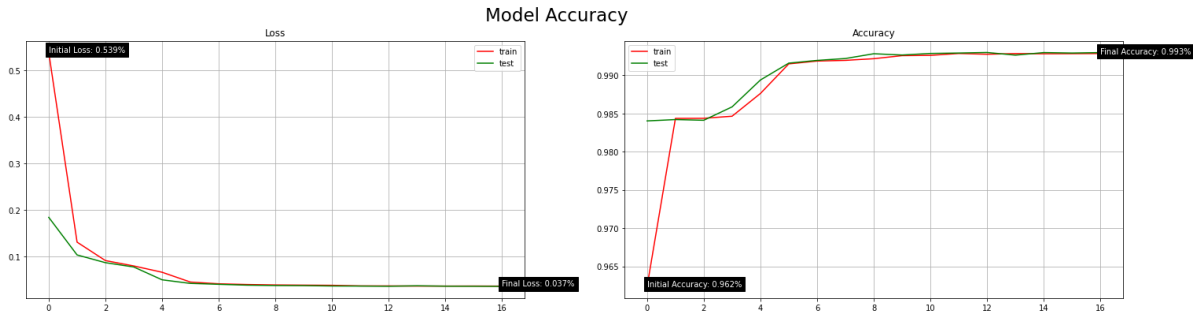
Model Accuracy







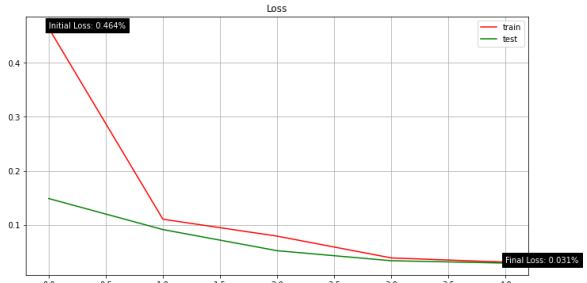


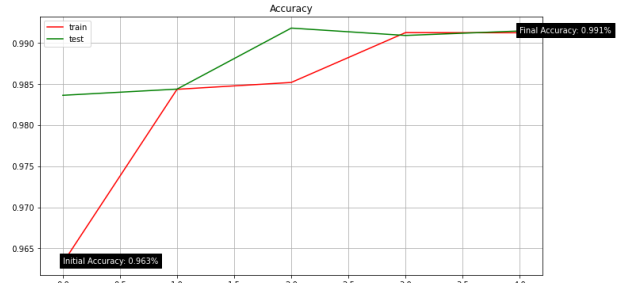


Training and Testing Dataset Partition		
Dataset Partition (%)	Model Loss (Error)	Model Accuracy (%)
0.9 | 0.1	0.006	99.85%
0.85 | 0.15	0.015	99.74%
0.8 | 0.2	0.007	99.83%
0.75 | 0.25	0.006	99.82%
0.7 | 0.3	0.004	99.91%
0.65 | 0.35	0.007	99.84%
0.6 | 0.4	0.010	99.79%
0.55 | 0.45	0.038	99.25%
0.5 | 0.5	0.011	99.77%
0.45 | 0.55	0.008	99.84%
0.4 | 0.6	0.036	99.29%

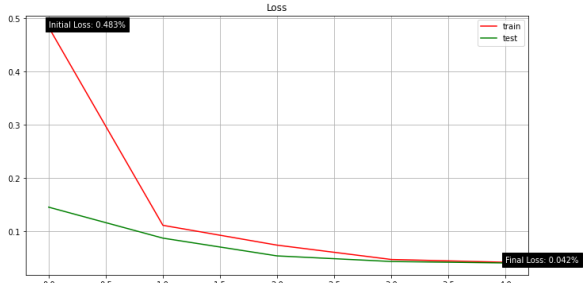
- K-fold Cross Validation (CV) dataset partitioning alteration

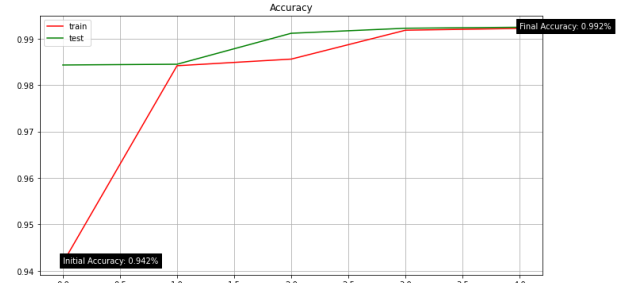
Dataset Fold 1



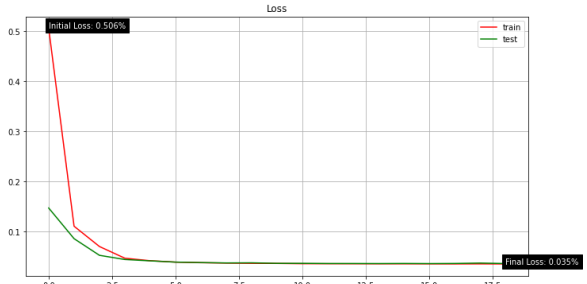


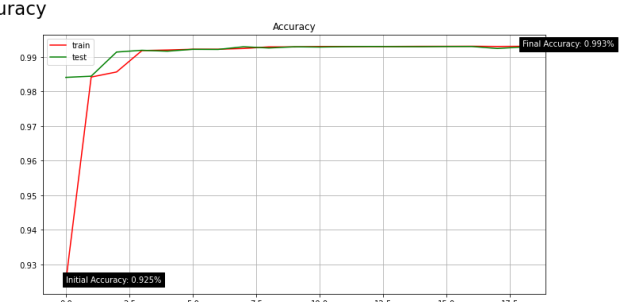
Dataset Fold 2



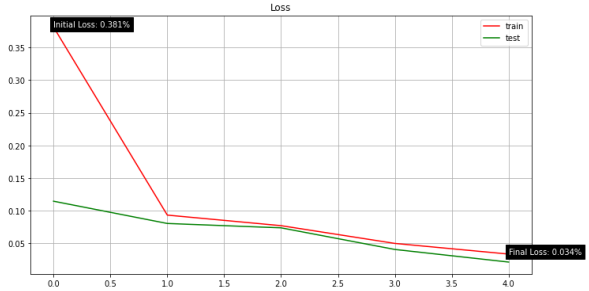


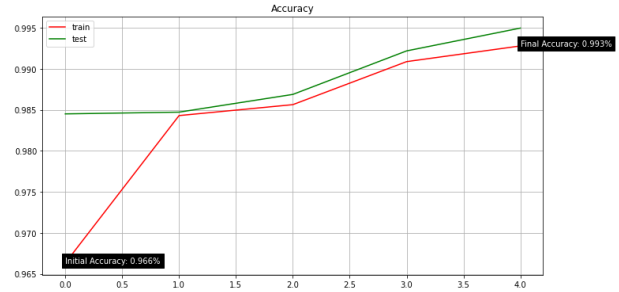
Model Accuracy



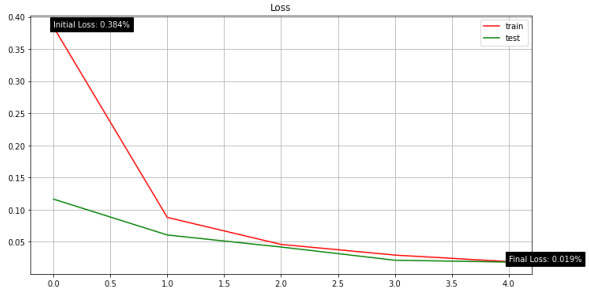


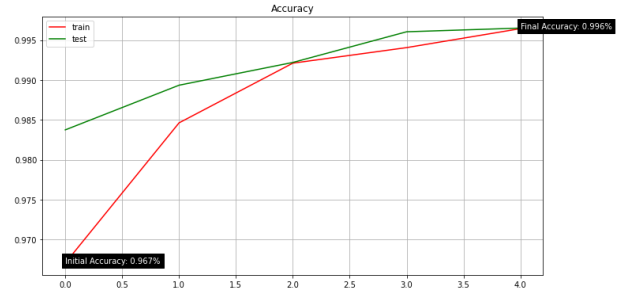
Dataset Fold 1



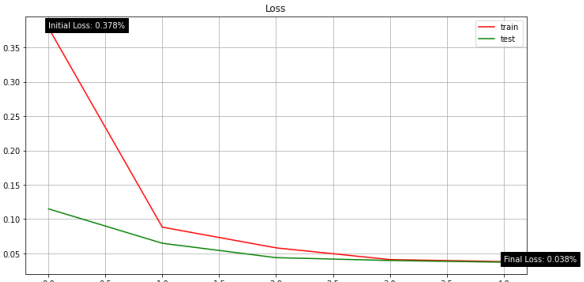


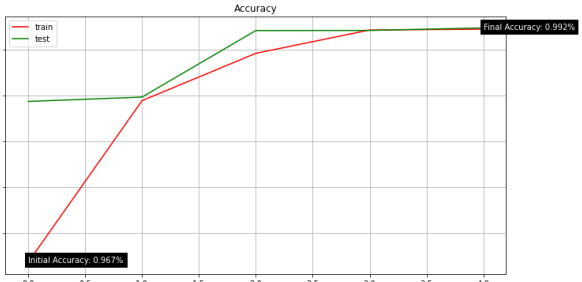
Dataset Fold 2



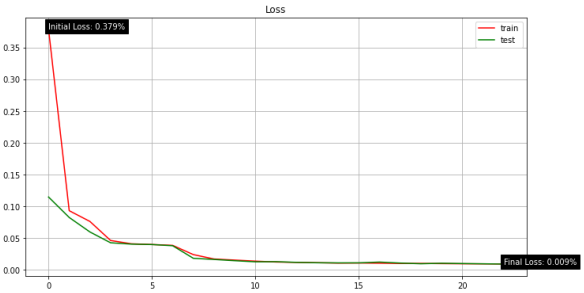


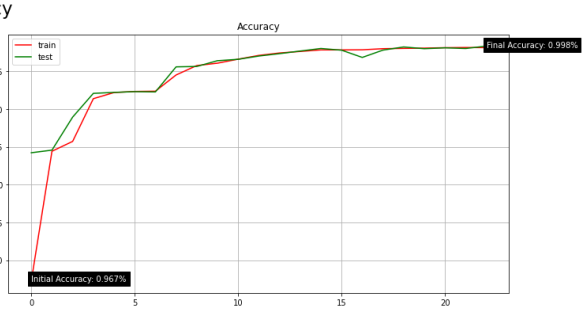
Dataset Fold 3



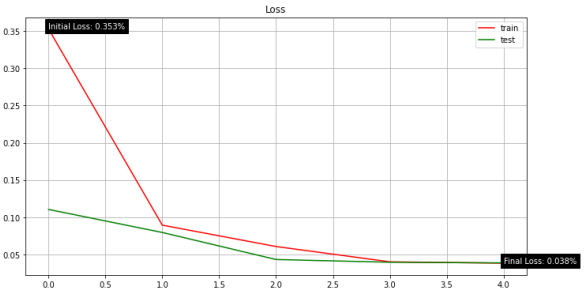


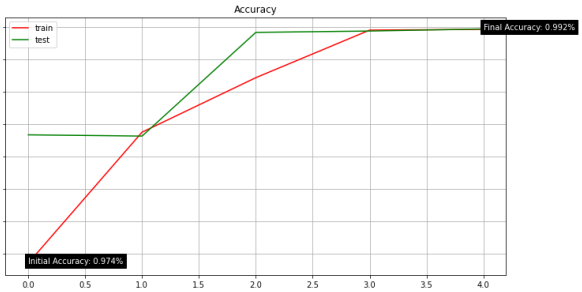
Model Accuracy



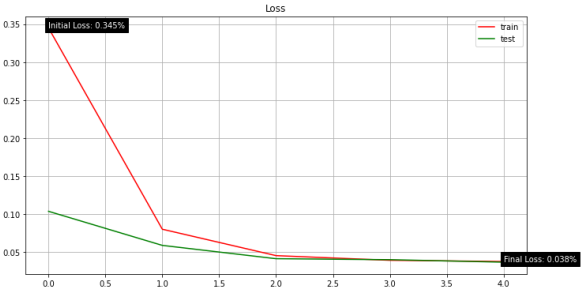


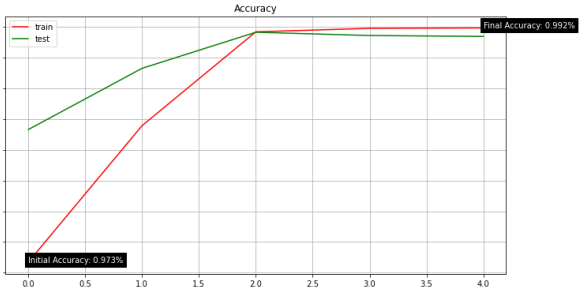
Dataset Fold 1



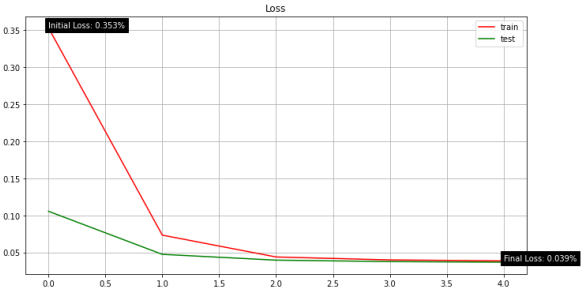


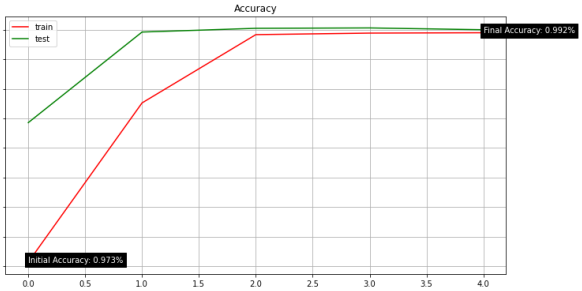
Dataset Fold 2



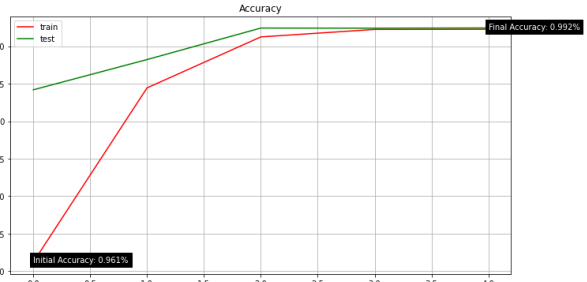


Dataset Fold 3

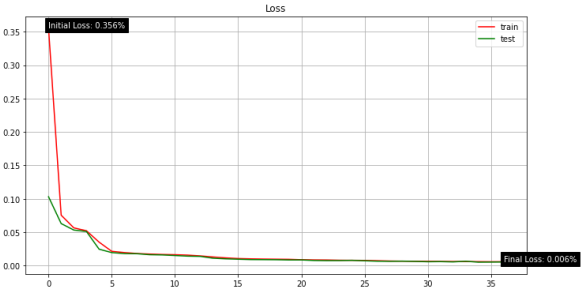


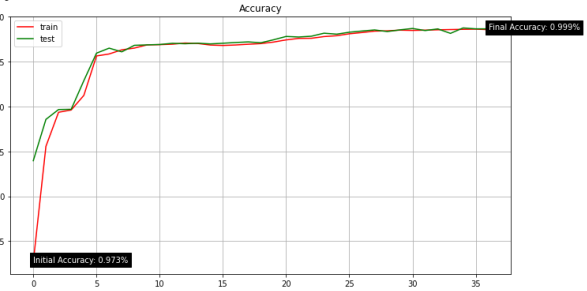


Dataset Fold 4

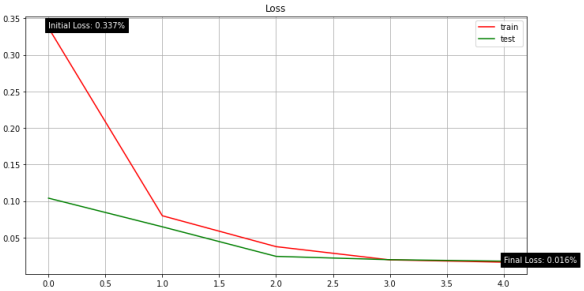


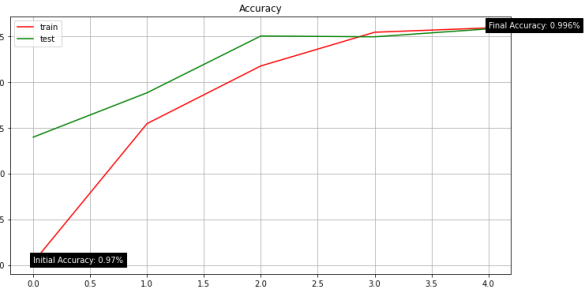
Model Accuracy



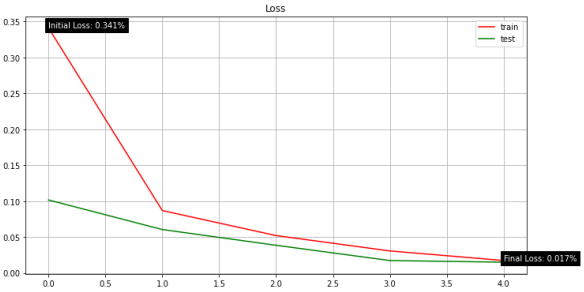


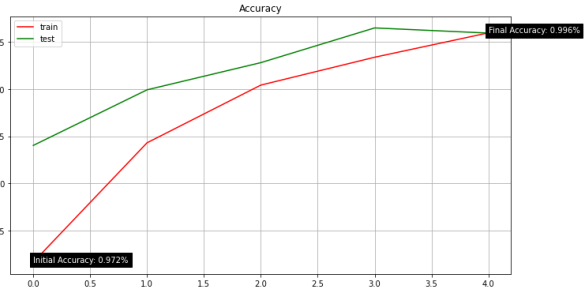
Dataset Fold 1



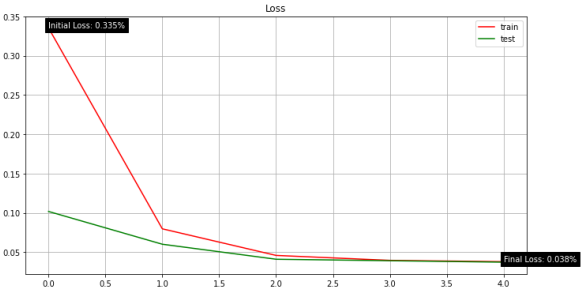


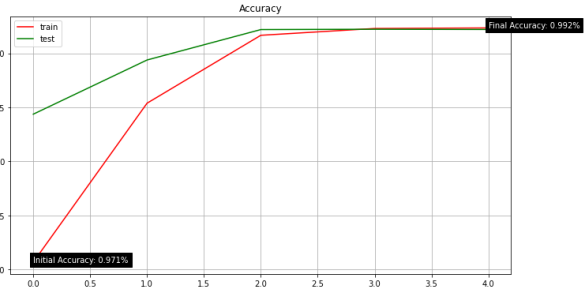
Dataset Fold 2



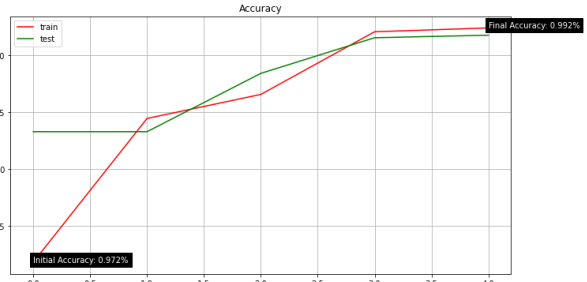


Dataset Fold 3

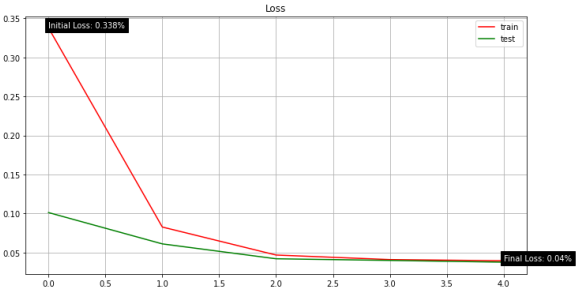


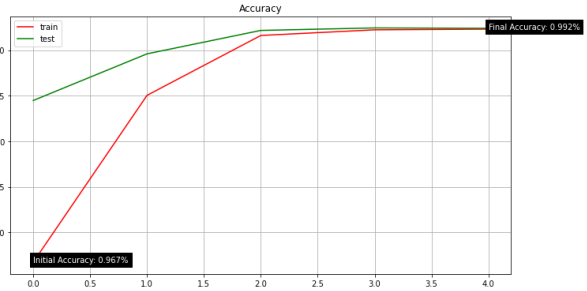


Dataset Fold 4

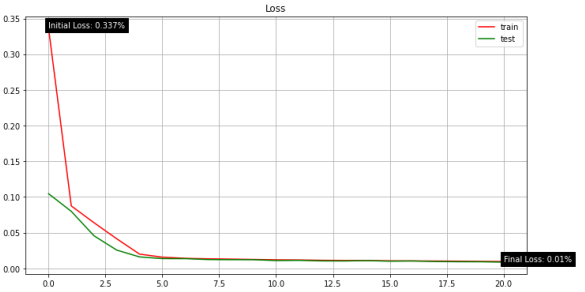


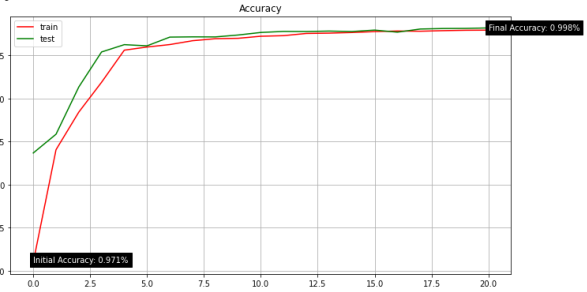
Dataset Fold 5



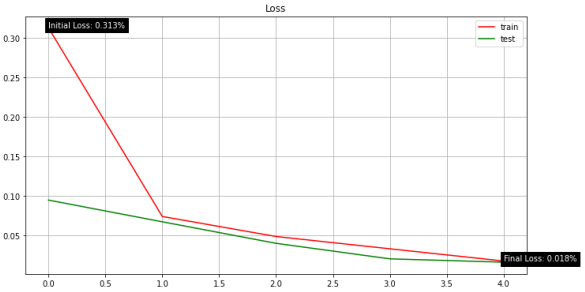


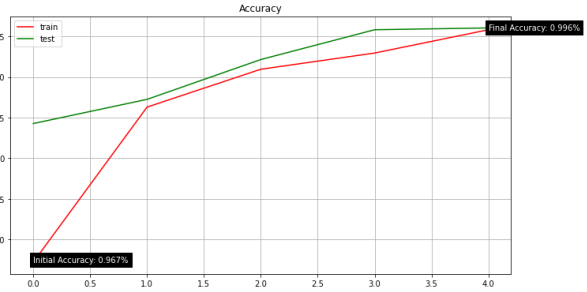
Model Accuracy



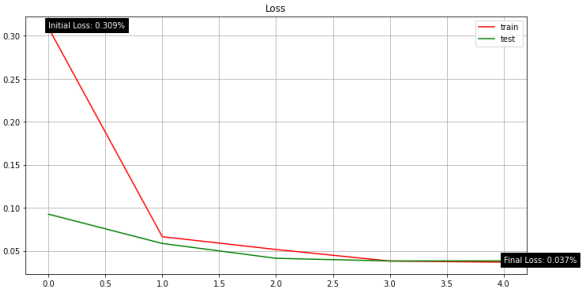


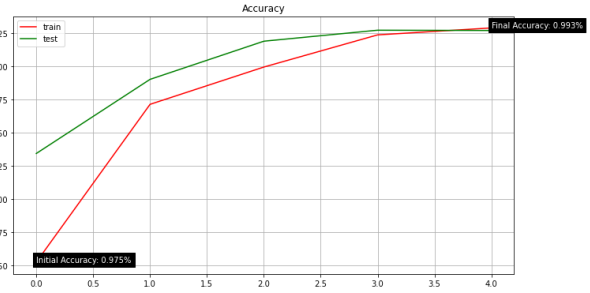
Dataset Fold 1



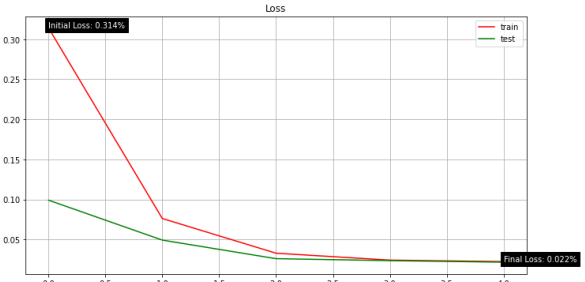


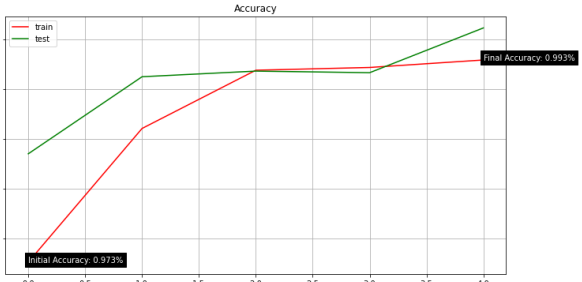
Dataset Fold 2



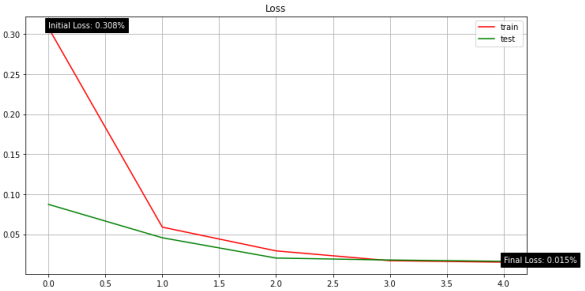


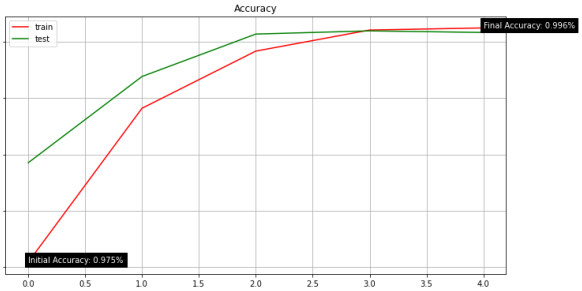
Dataset Fold 3



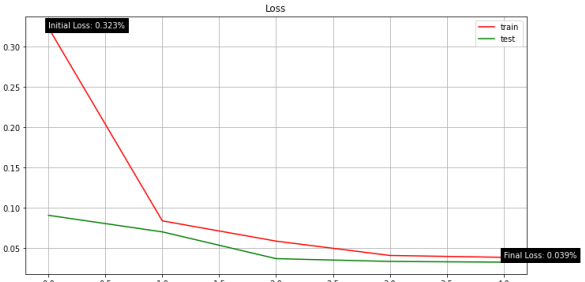


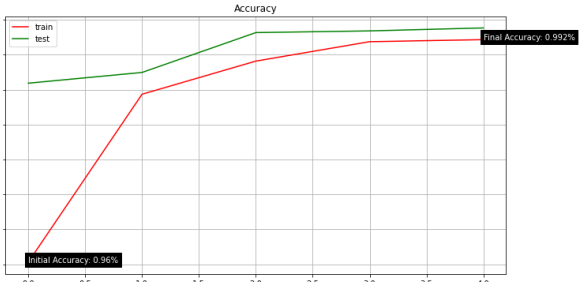
Dataset Fold 4



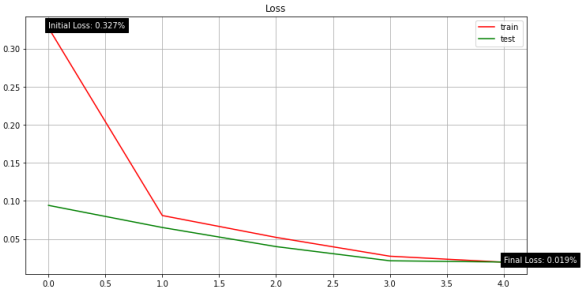


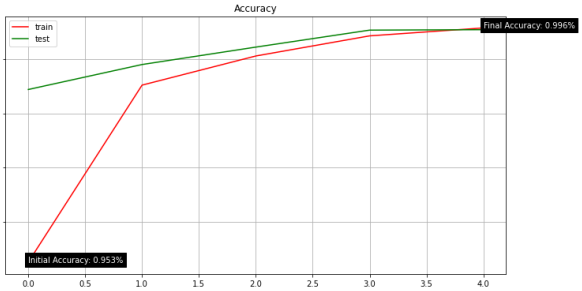
Dataset Fold 5



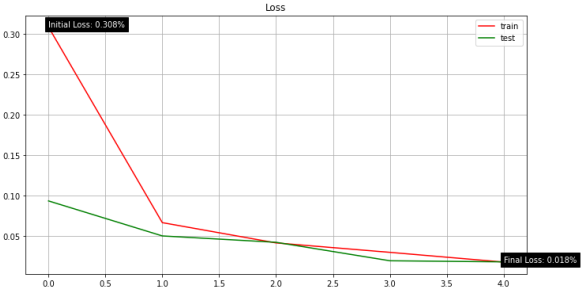


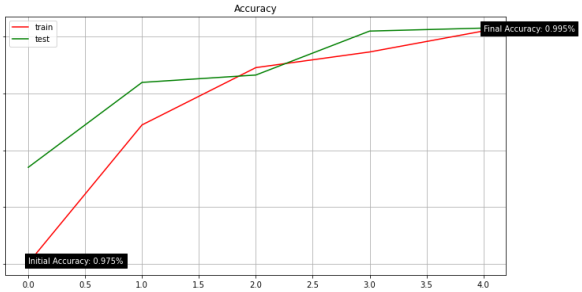
Dataset Fold 6

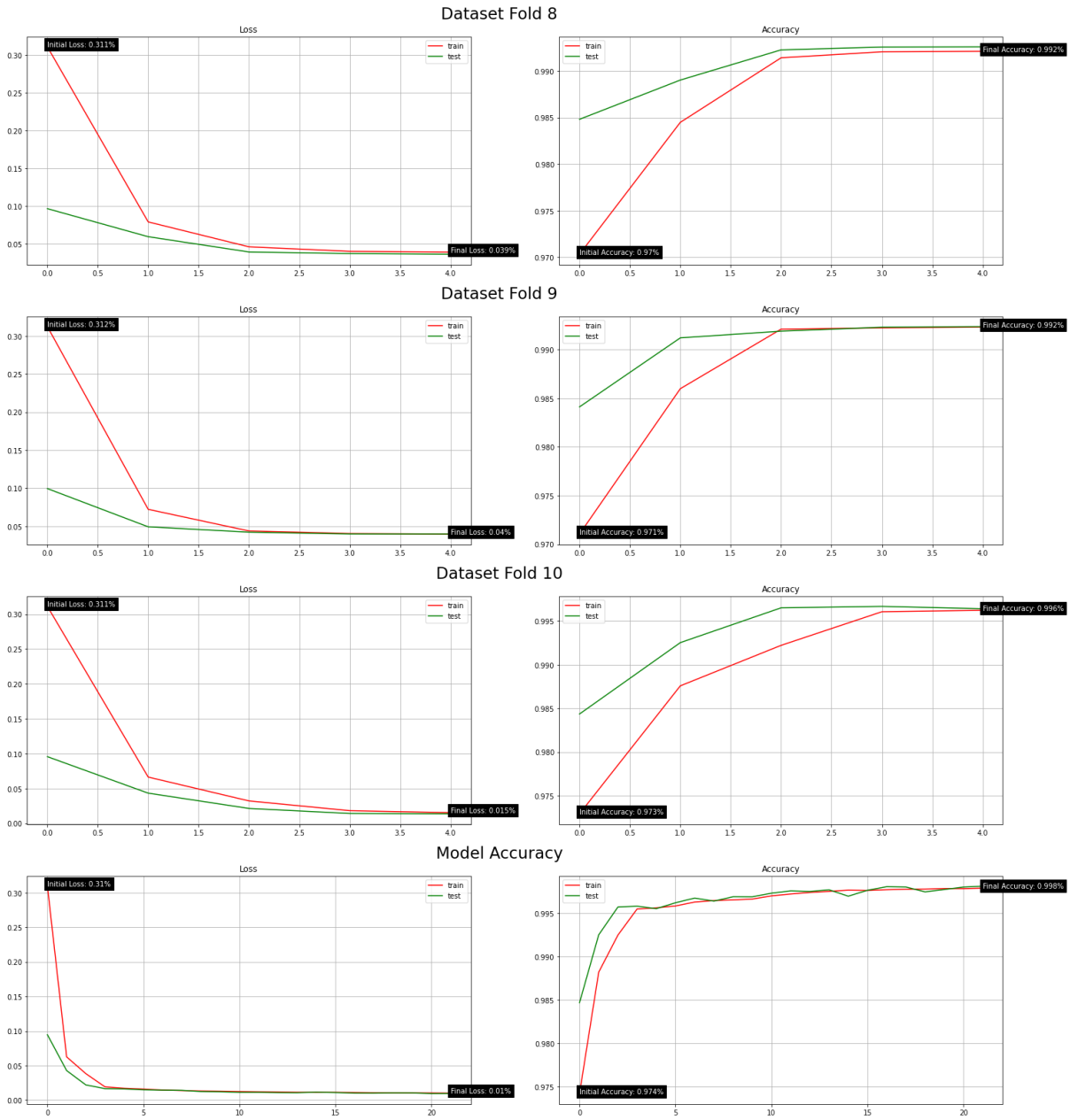




Dataset Fold 7

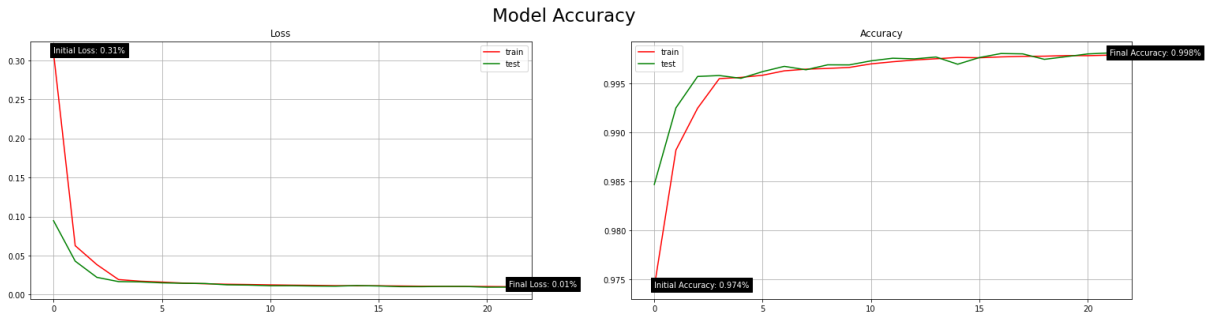






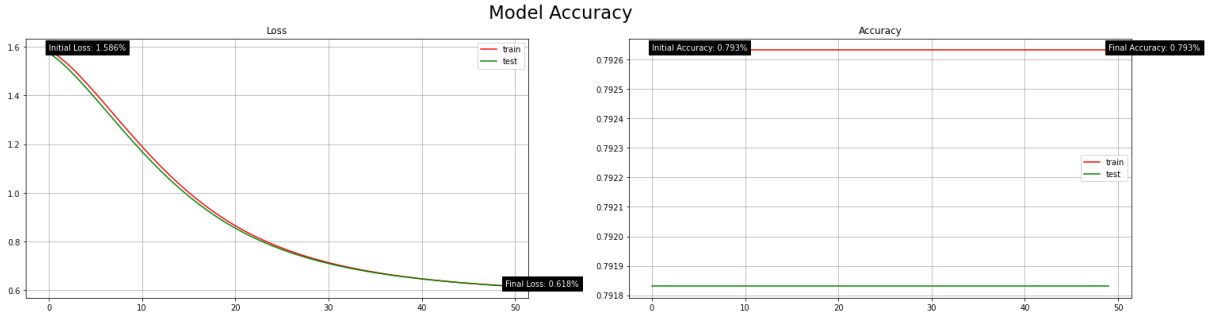
K-fold Cross Validation (CV)		
K-folds	Model Loss (Error)	Model Accuracy (%)
2	0.036	99.30%
3	0.010	99.78%
4	0.006	99.85%
5	0.010	99.80%
10	0.010	99.81%

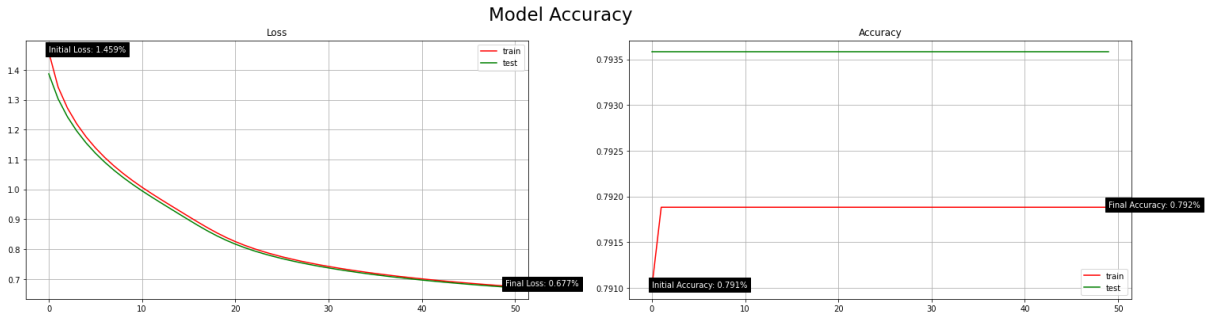
- K-fold Cross-validation (CV) activeness alteration



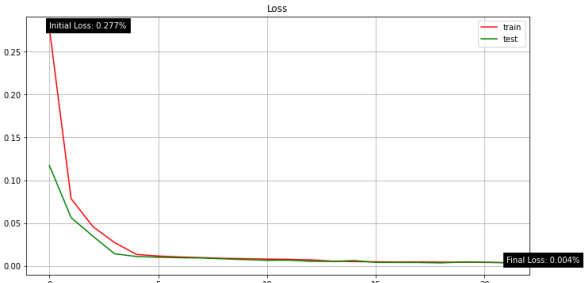
K-fold Cross Validation (CV)		
Active?	Model Loss (Error)	Model Accuracy (%)
True	0.006	99.85%
False	0.004	99.91%

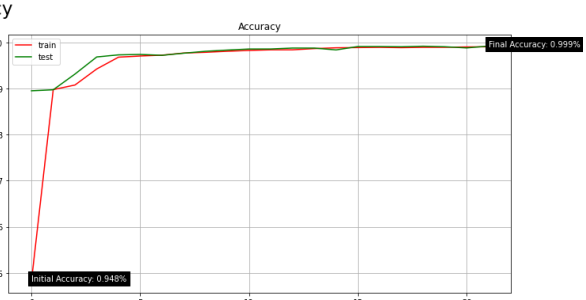
- Active model optimiser algorithm alteration



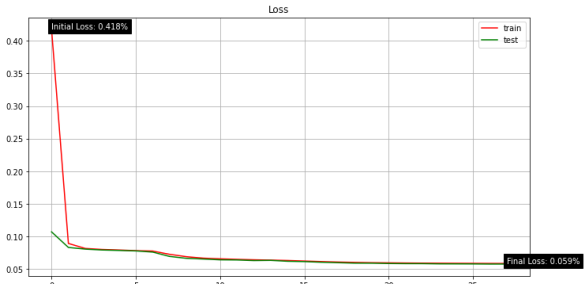


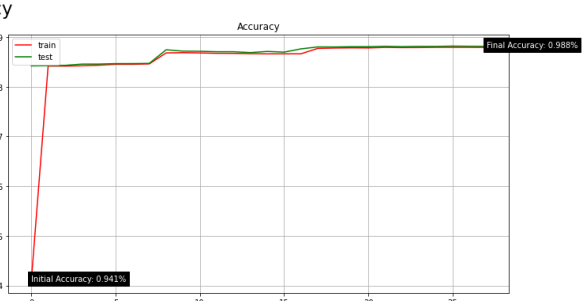
Model Accuracy



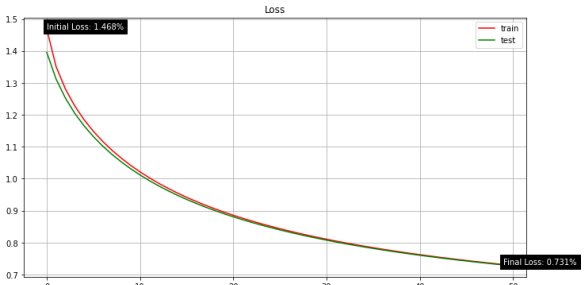


Model Accuracy

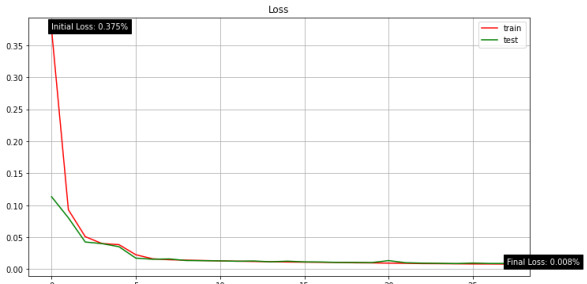


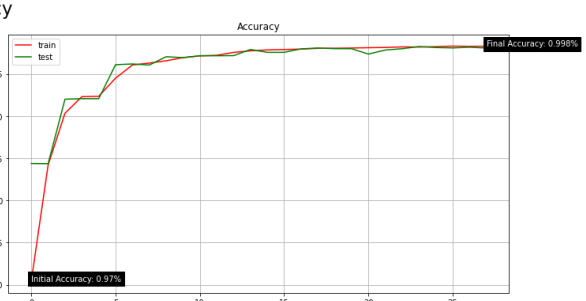


Model Accuracy

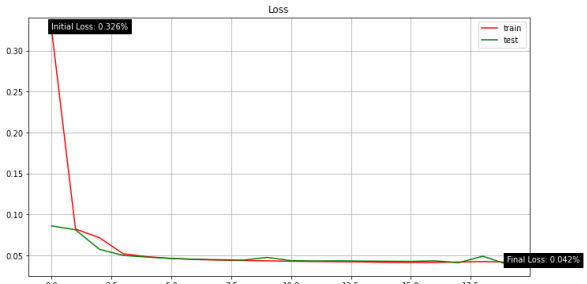


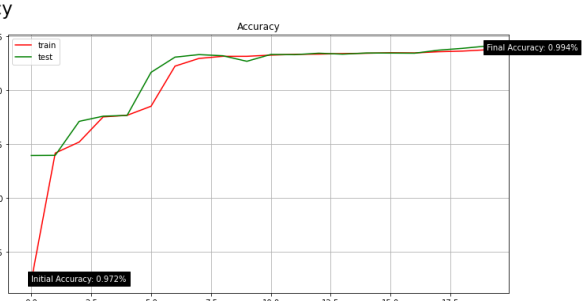
Model Accuracy

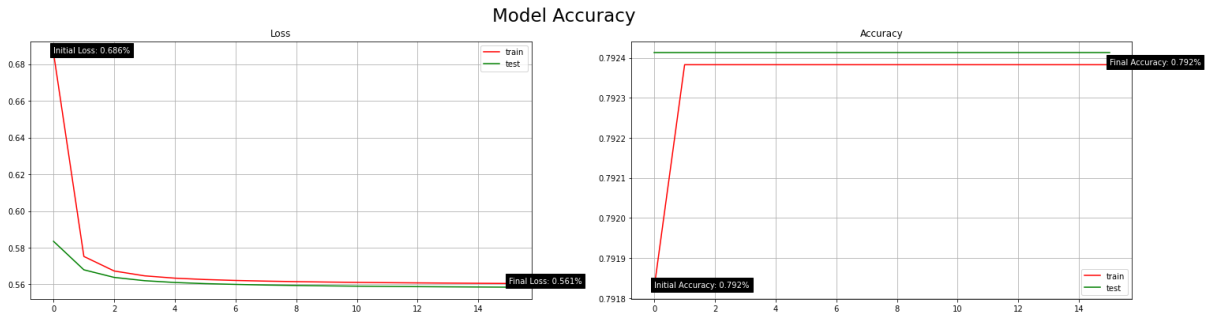




Model Accuracy

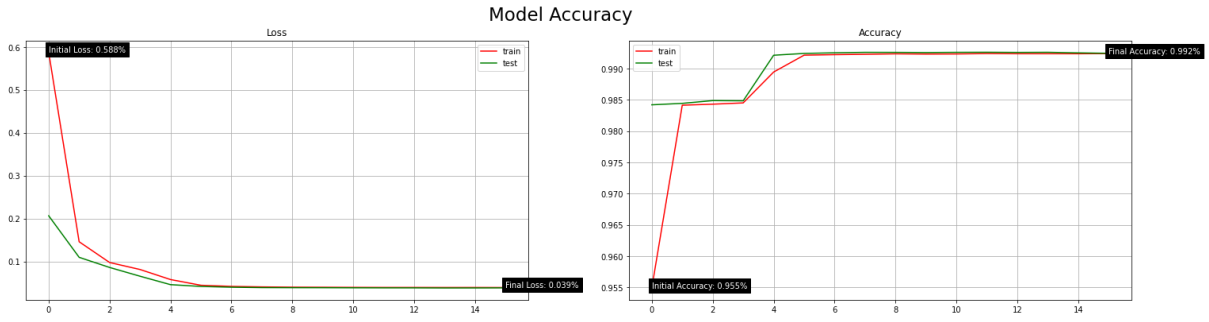


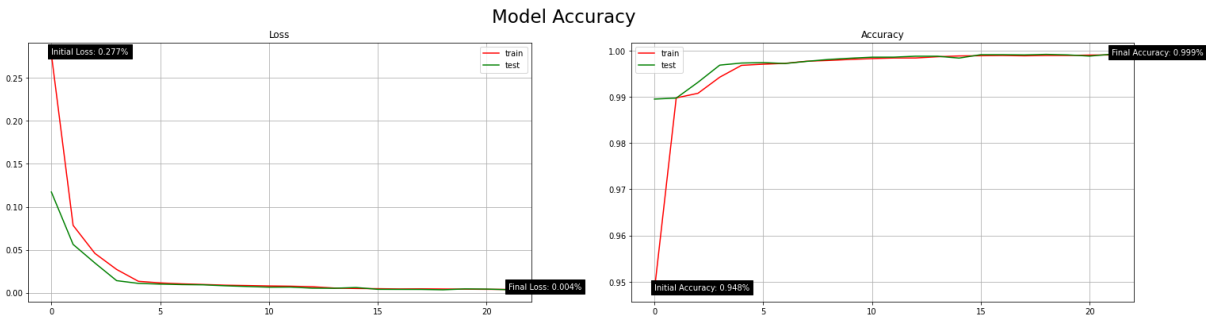


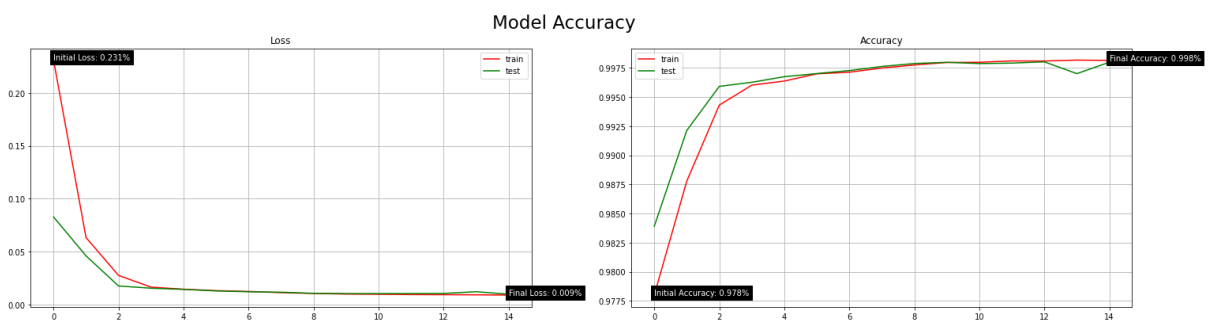


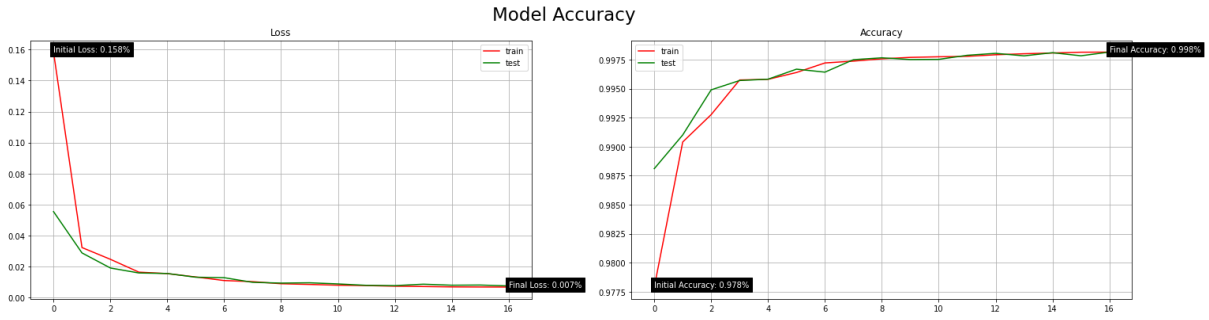
Model Optimiser Algorithm		
Optimiser Algorithm	Model Loss (Error)	Model Accuracy (%)
Adadelta	0.618	79.18%
Adagrad	0.673	79.36%
Adam	0.004	99.91%
Adamax	0.005	98.81%
Ftrl	0.729	79.27%
Nadam	0.009	99.80%
RMSprop	0.043	99.34%
SGD	0.559	79.24%

- Training dataset batch size alteration



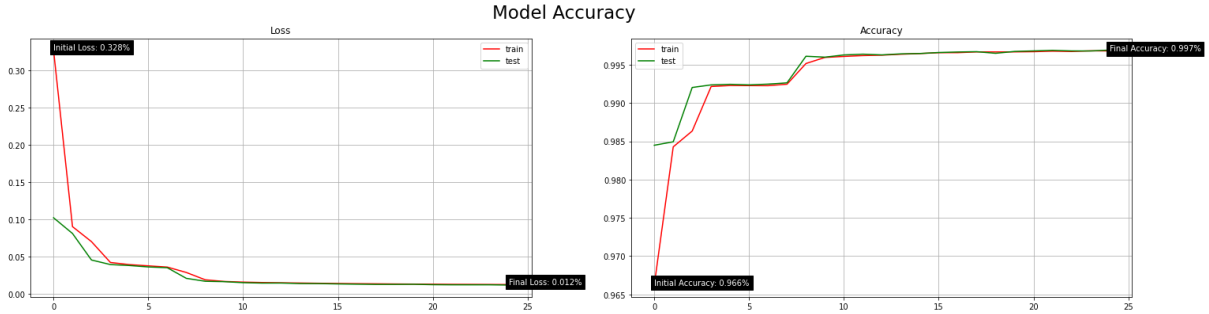


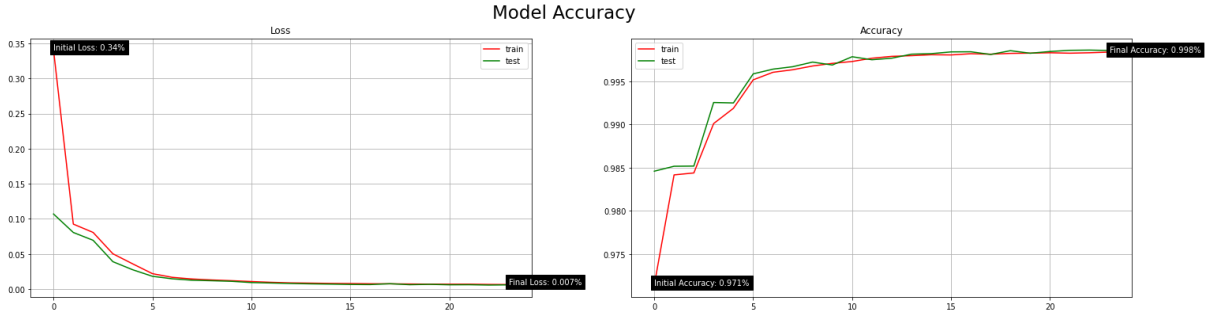


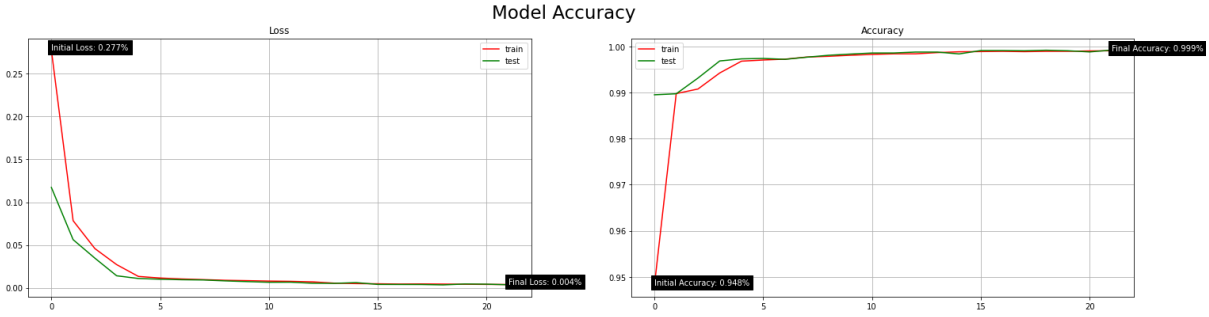


Training Dataset Sample (Batch) Size		
Batch Sample Size	Model Loss (Error)	Model Accuracy (%)
256	0.038	99.26%
128	0.004	99.91%
64	0.010	99.80%
32	0.008	99.79%

- Fully-connected layer (FCL) count in the model (when model topology is [32, 32, 32, 5, 5]) alteration



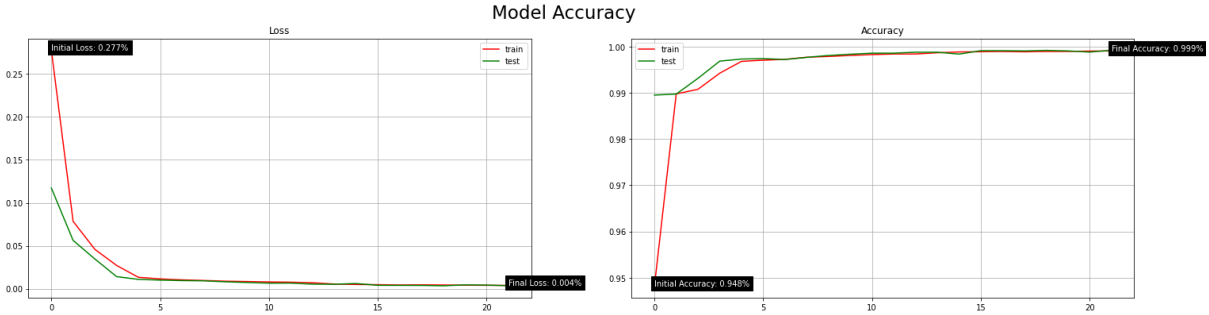


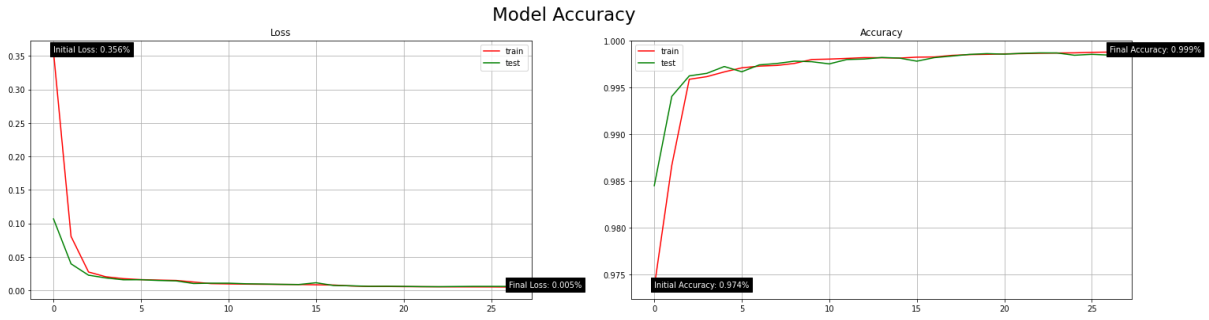


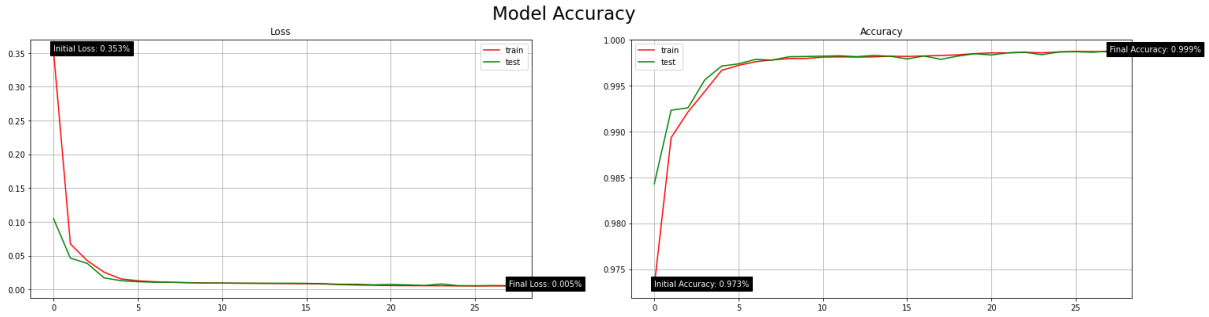
Model Fully-Connected Layer (FCL) Count

FCL Count	Model Loss (Error)	Model Accuracy (%)
3	0.013	99.67%
4	0.006	99.85%
5	0.004	99.91%

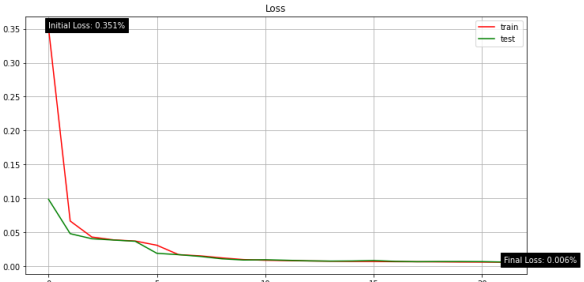
- Model topology (neuron counts) alteration

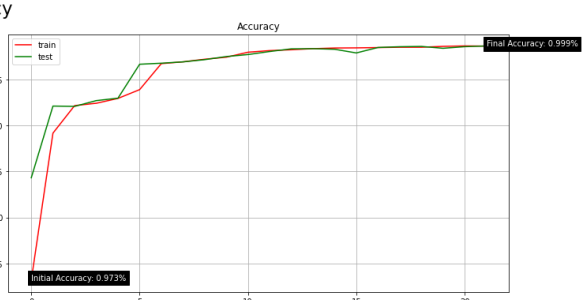




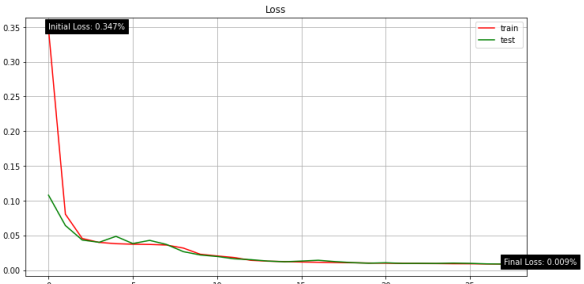


Model Accuracy

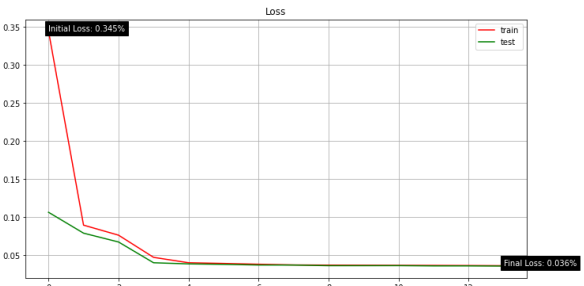


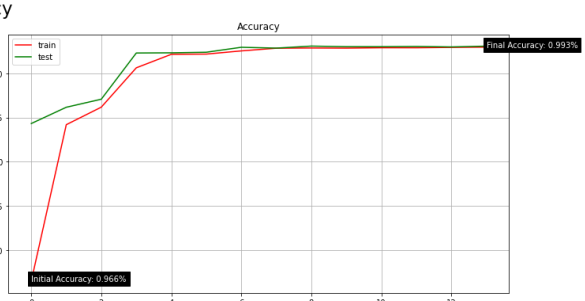


Model Accuracy

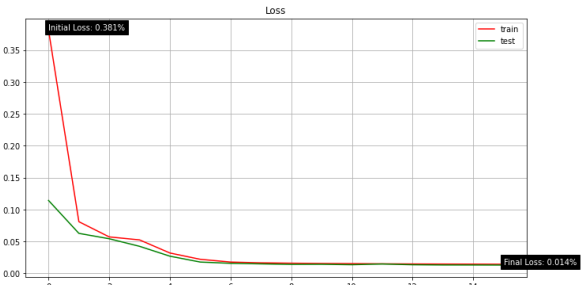


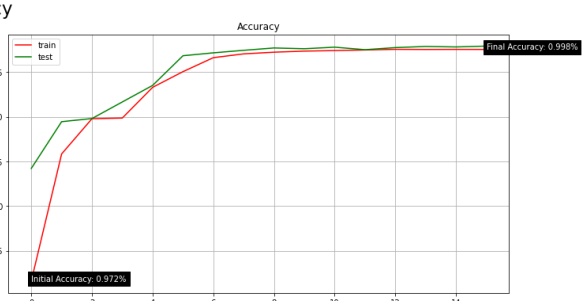
Model Accuracy



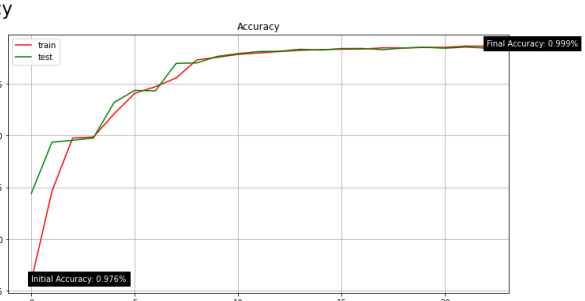


Model Accuracy

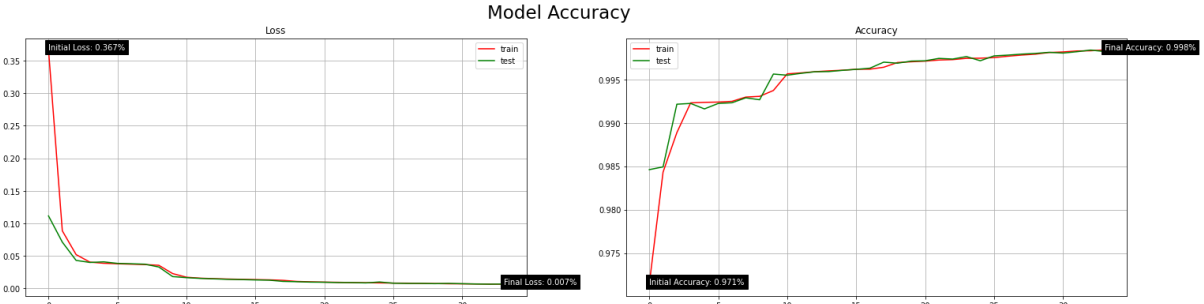




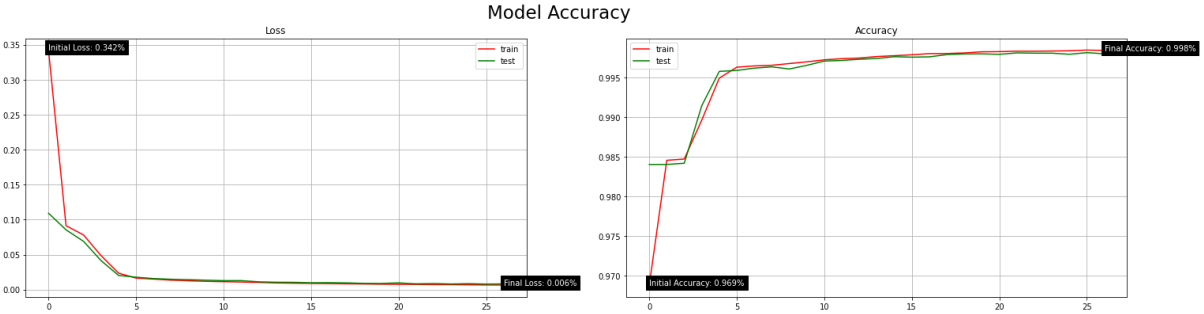
Model Accuracy



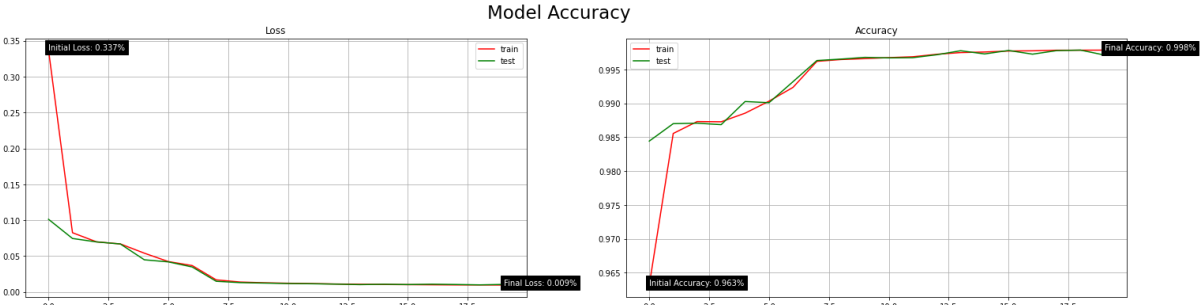
Model Accuracy



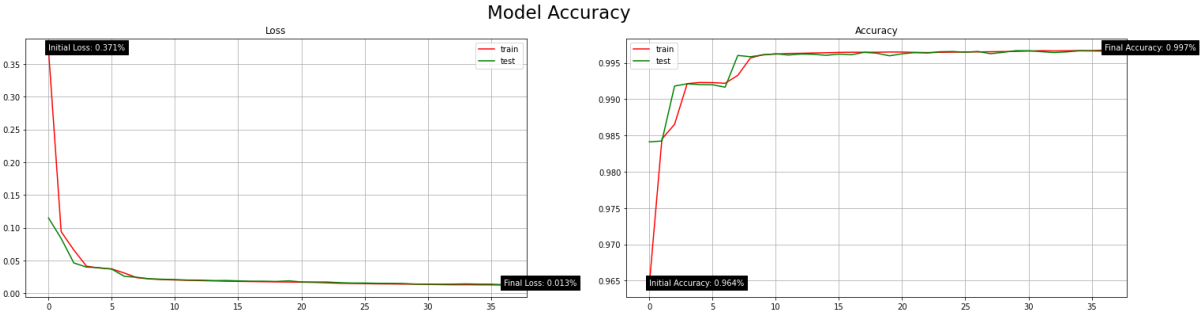
Model Accuracy



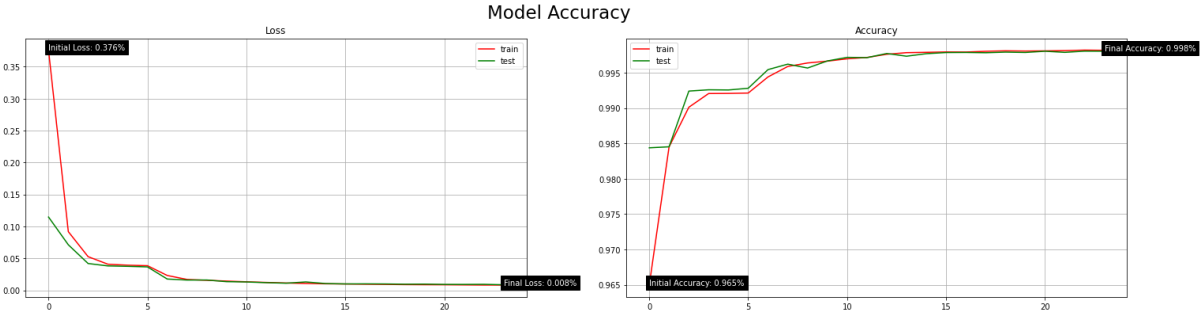
Model Accuracy

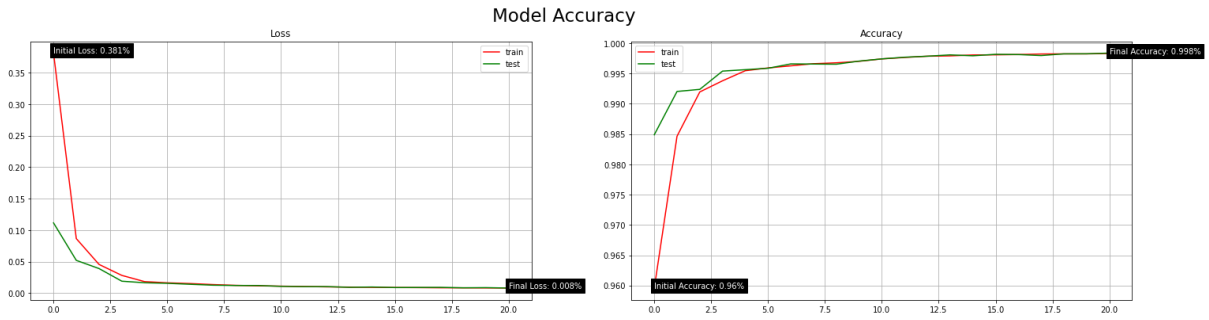


Model Accuracy



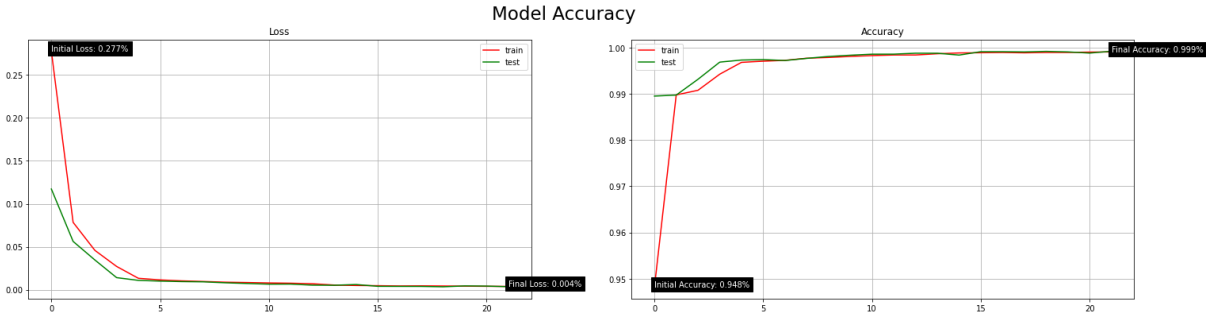
Model Accuracy

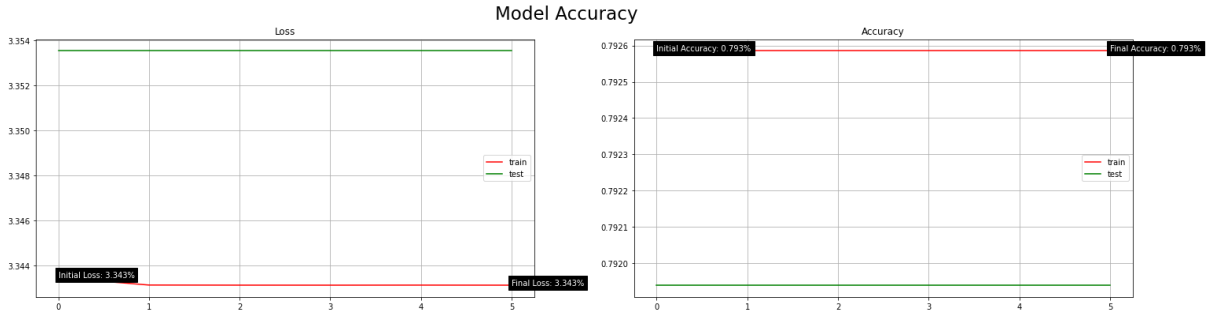




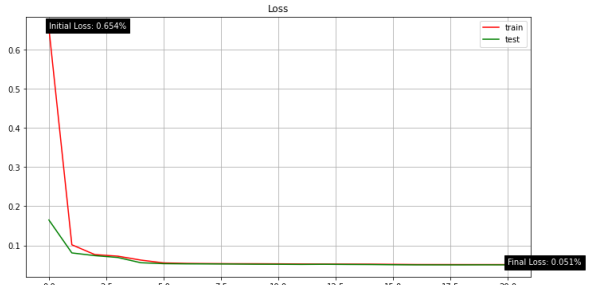
Model Topology		
Neurons Per Layer	Model Loss (Error)	Model Accuracy (%)
32, 32, 32, 5, 5	0.004	99.91%
80, 80, 80, 80, 5	0.005	99.87%
70, 70, 70, 70, 5	0.006	99.86%
60, 60, 60, 60, 5	0.006	99.86%
50, 50, 50, 50, 5	0.007	99.85%
40, 40, 40, 40, 5	0.010	99.81%
30, 30, 30, 30, 5	0.036	99.31%
80, 80, 80, 5, 5	0.013	99.78%
40, 40, 40, 5, 5	0.008	99.83%
30, 30, 30, 5, 5	0.007	99.81%
30, 30, 30, 30, 5	0.011	99.73%
20, 20, 20, 20, 5	0.008	99.81%
20, 20, 20, 5, 5	0.014	99.66%
35, 35, 35, 5, 5	0.009	99.82%

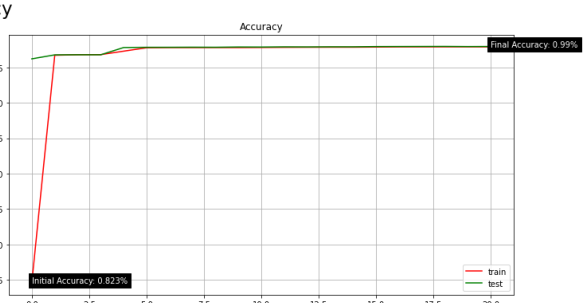
- Model layer transfer function alteration



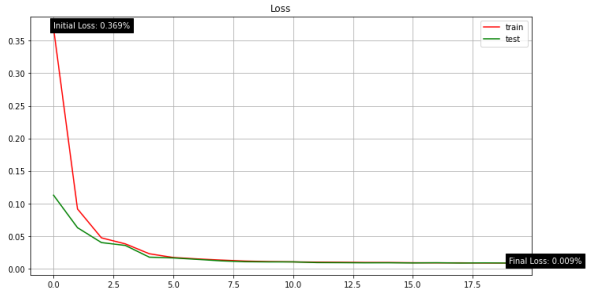


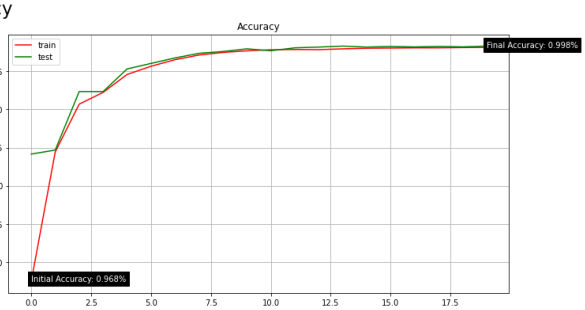
Model Accuracy



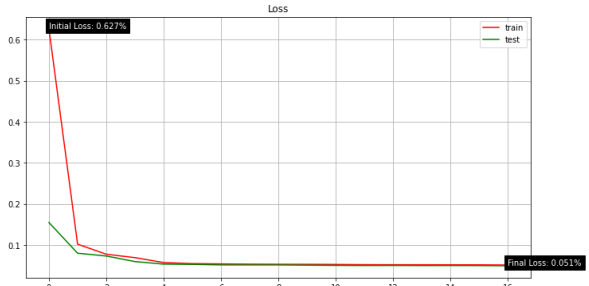


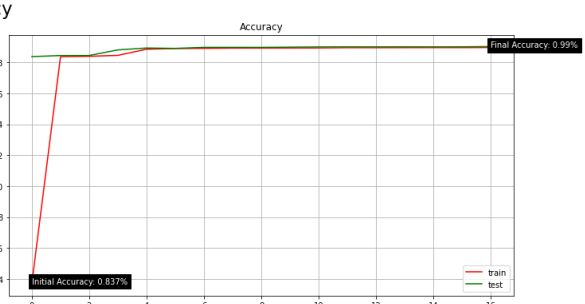
Model Accuracy



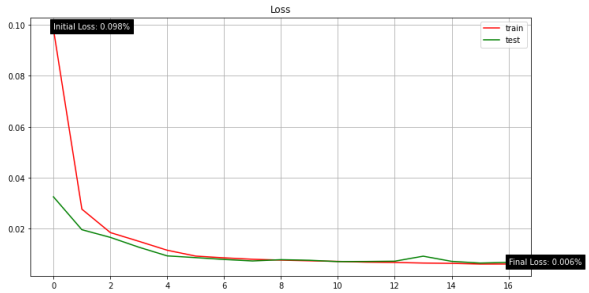


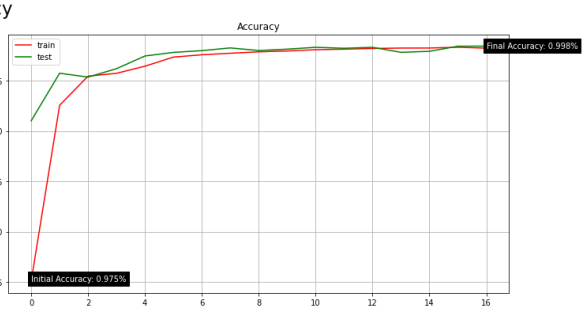
Model Accuracy



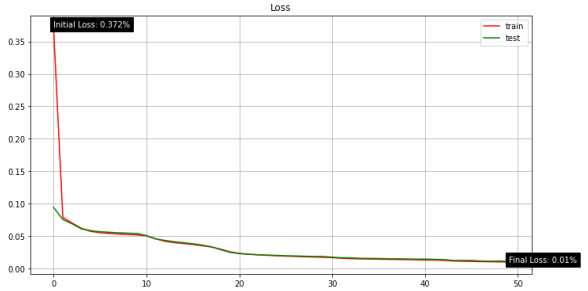


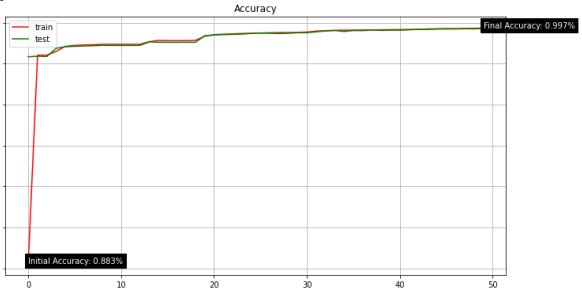
Model Accuracy

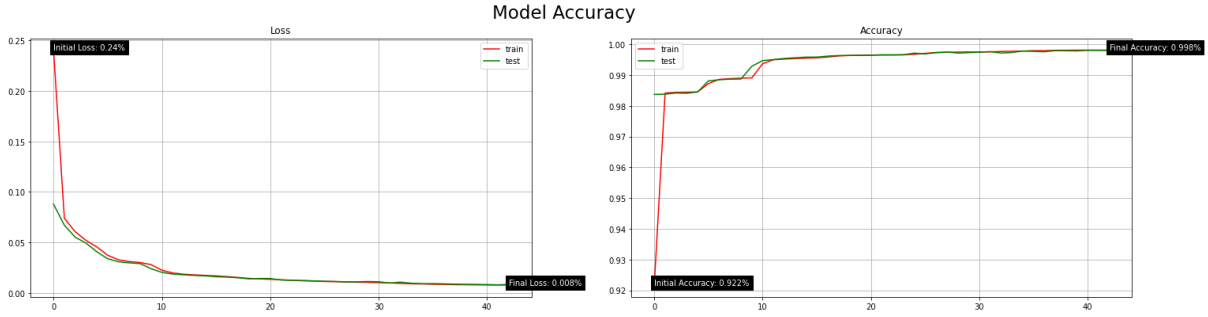


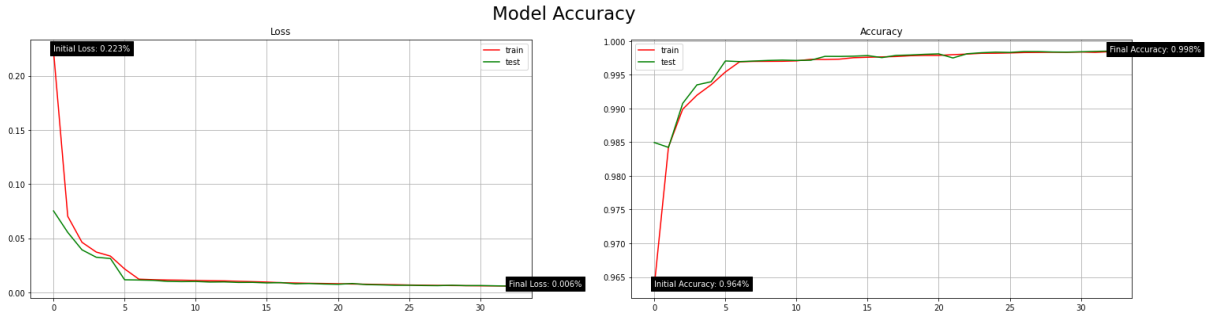


Model Accuracy







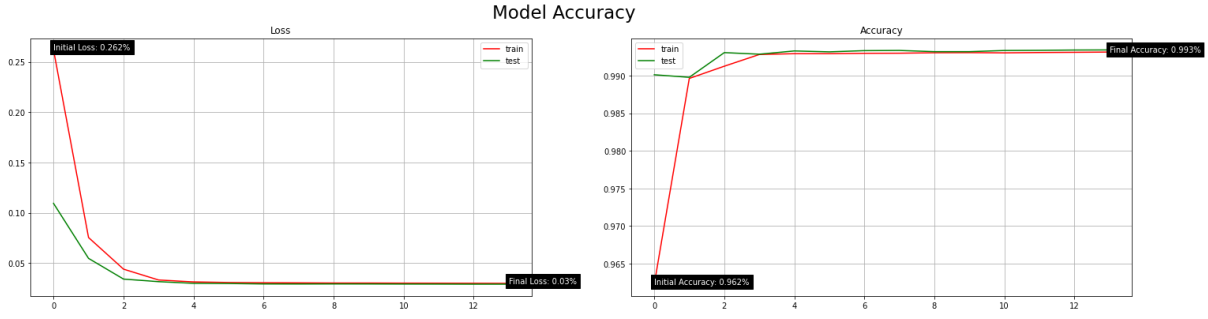


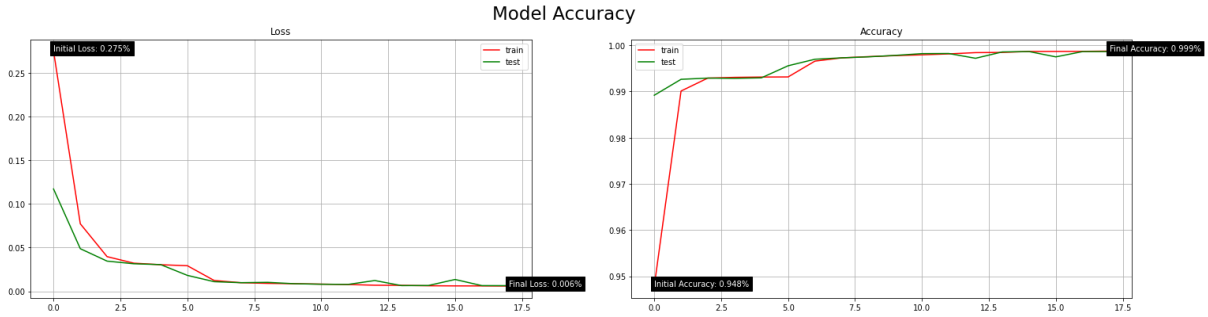
Model Layer Transfer (Activation) Functions		
Transfer Function Per layer	Model Loss (Error)	Model Accuracy (%)
relu, relu, relu, softmax, sigmoid	0.004	99.91%
relu, relu, relu, softmax, relu	3.354	79.19%
sigmoid, sigmoid, sigmoid, softmax, sigmoid	0.051	98.98%
relu, relu, relu, softmax, softmax	0.009	99.82%
sigmoid, sigmoid, sigmoid, softmax, softmax	0.050	99.00%
relu, relu, relu, relu, softmax	0.007	99.82%
sigmoid, sigmoid, sigmoid, sigmoid, softmax	0.011	99.72%
relu, sigmoid, sigmoid, relu, softmax	0.009	98.80%
relu, relu, softmax, relu, softmax	0.024	99.84%
relu, softmax, relu, softmax, sigmoid	0.007	99.85%

Appendix B

Binary Classification Model Fine-tuning

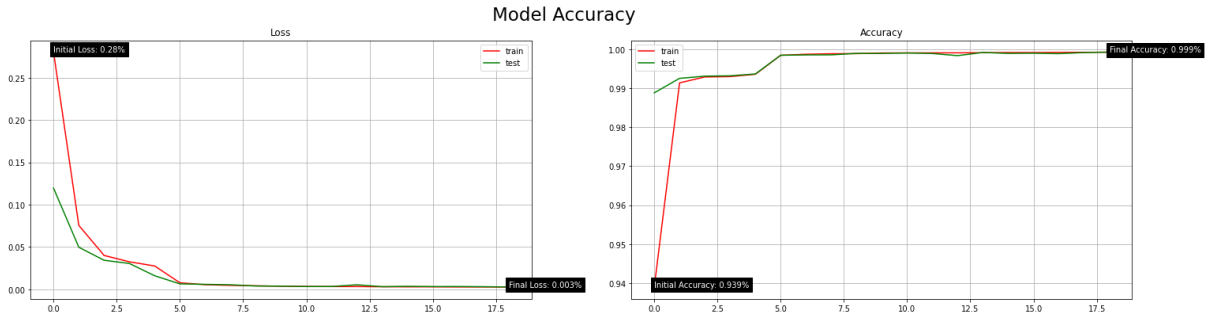
- Removing features of constant value from the dataset activeness alteration

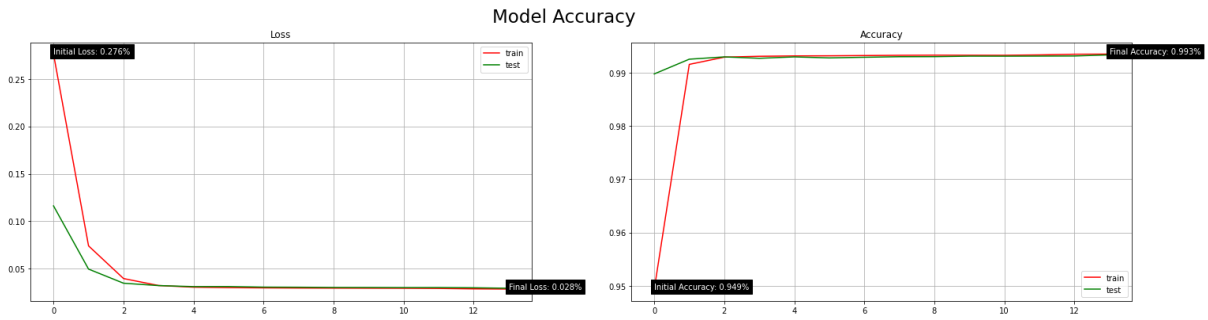


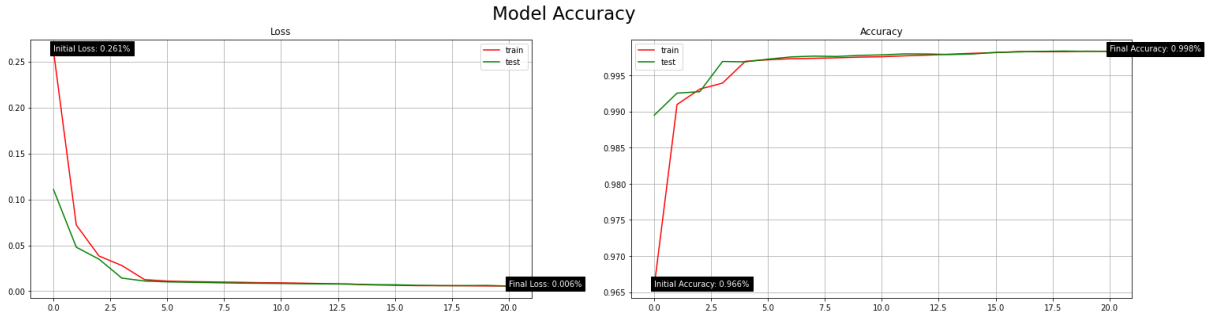


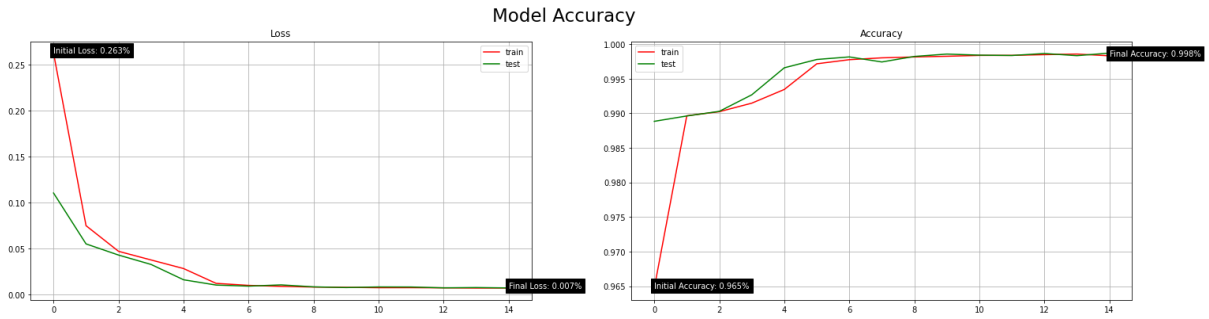
Removing Constant Features		
Removing Features?	Model Loss (Error)	Model Accuracy (%)
True	0.012	99.87%
False	0.029	99.32%

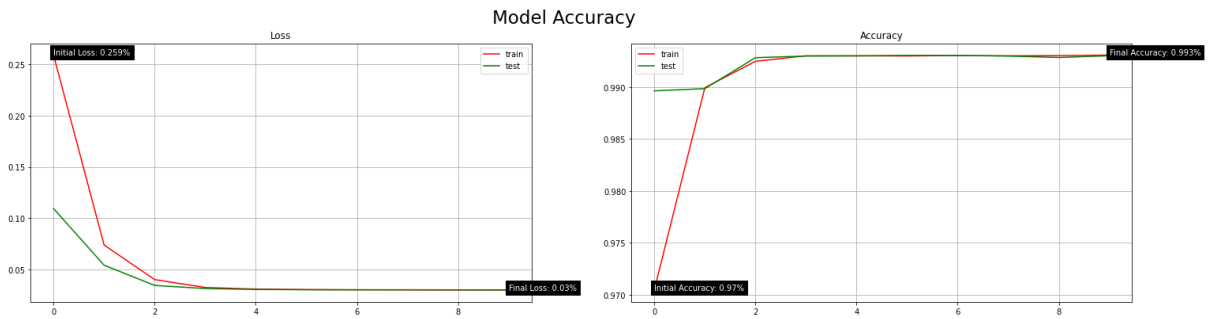
- Removing highly correlated features (Pearson Correlation Coefficient) threshold value alteration

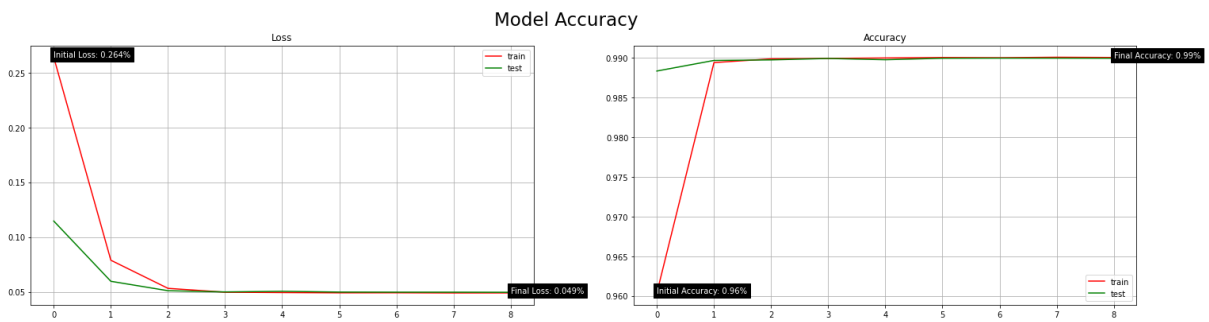








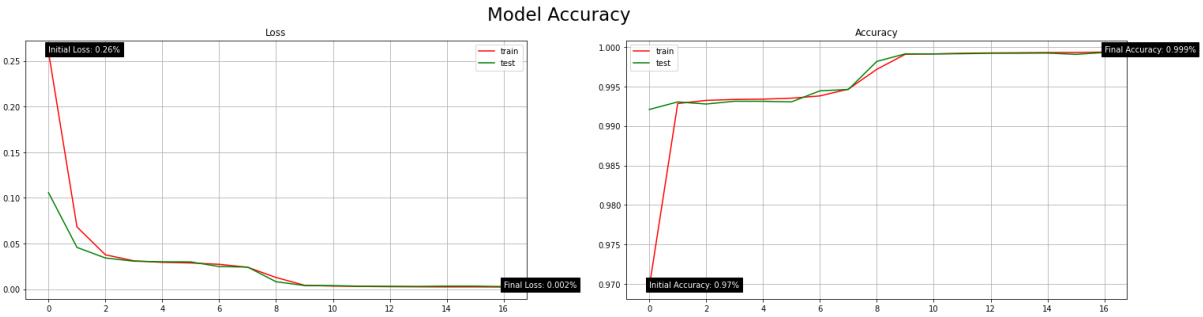


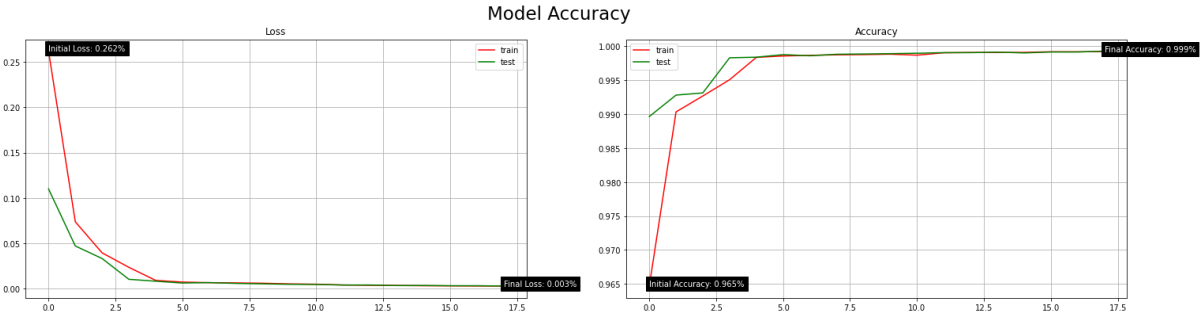


Removing Highly-Correlated Features		
Correlative Threshold	Model Loss (Error)	Model Accuracy (%)
0.95	0.003	99.92%

0.9	0.030	99.30%
0.85	0.030	99.31%
0.8	0.007	99.82%
0.75	0.012	99.87%
0.7	0.007	99.86%
0.65	0.031	99.30%
0.6	0.050	98.99%

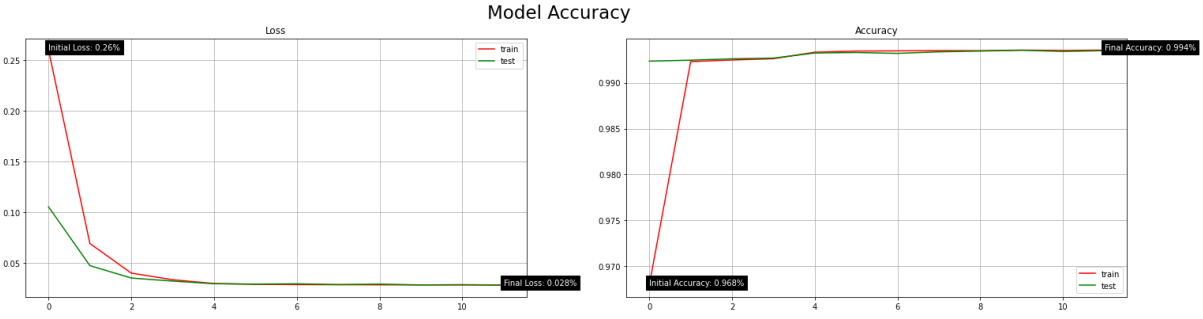
- Removing highly correlated features (Pearson Correlation Coefficient) activeness alteration



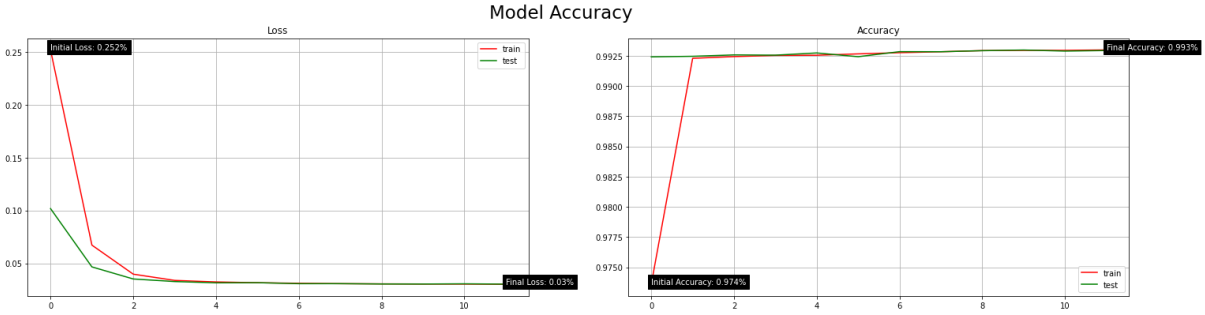


Removing Highly-Correlated Features		
Removing Features?	Model Loss (Error)	Model Accuracy (%)
True	0.004	99.91%
False	0.003	99.92%

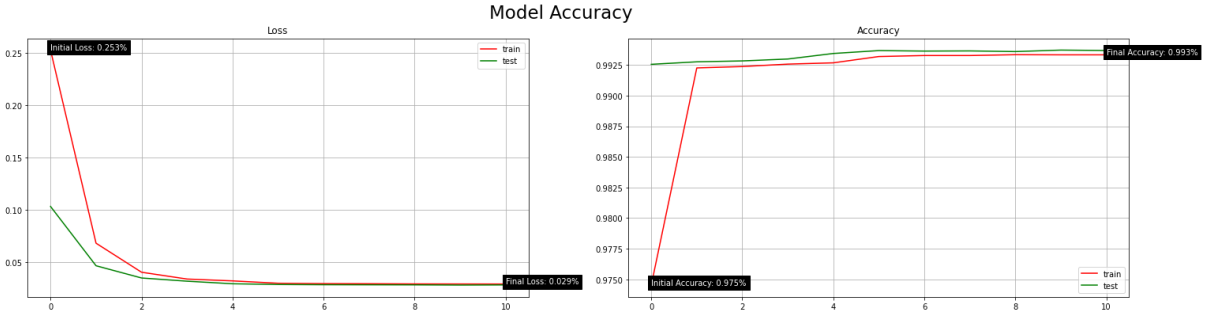
- Principal Component Analysis (PCA) number of components dataset reduced to (when correlation features not used) alteration



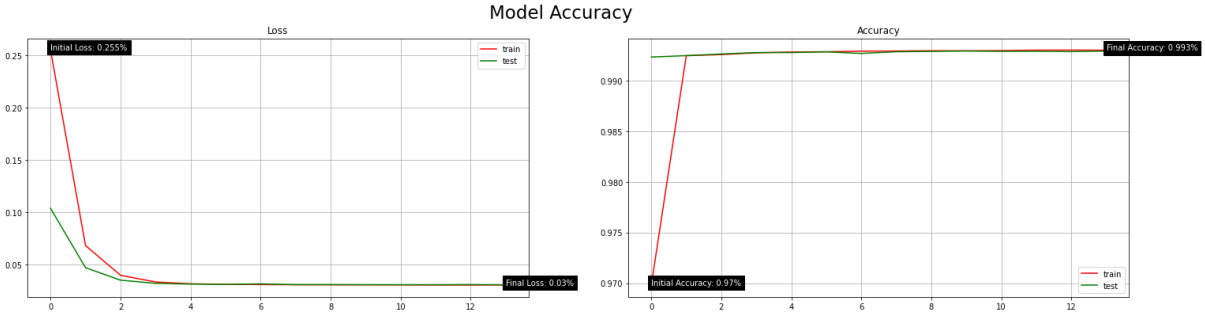
Model Accuracy



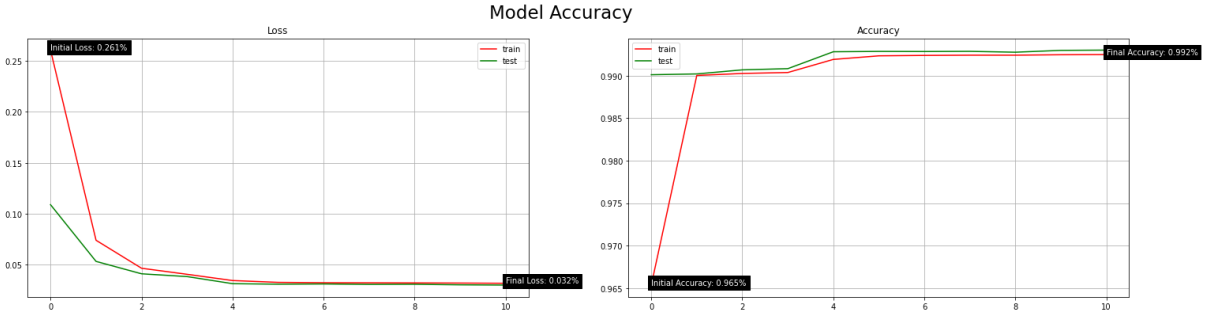
Model Accuracy



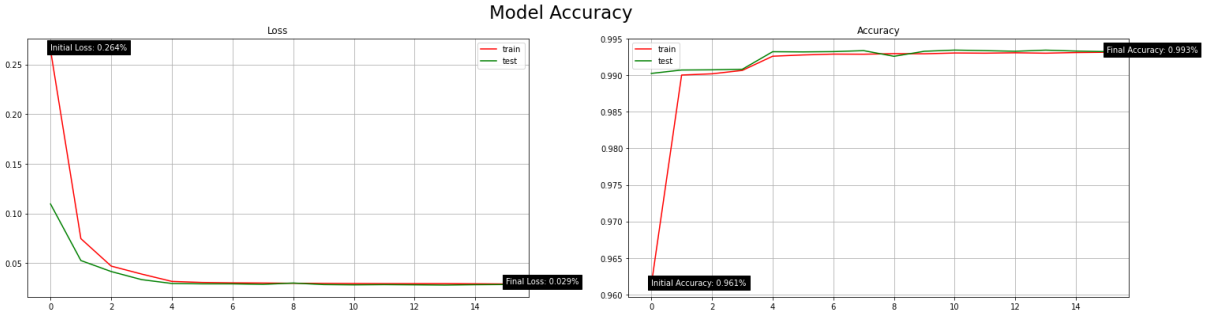
Model Accuracy



Model Accuracy



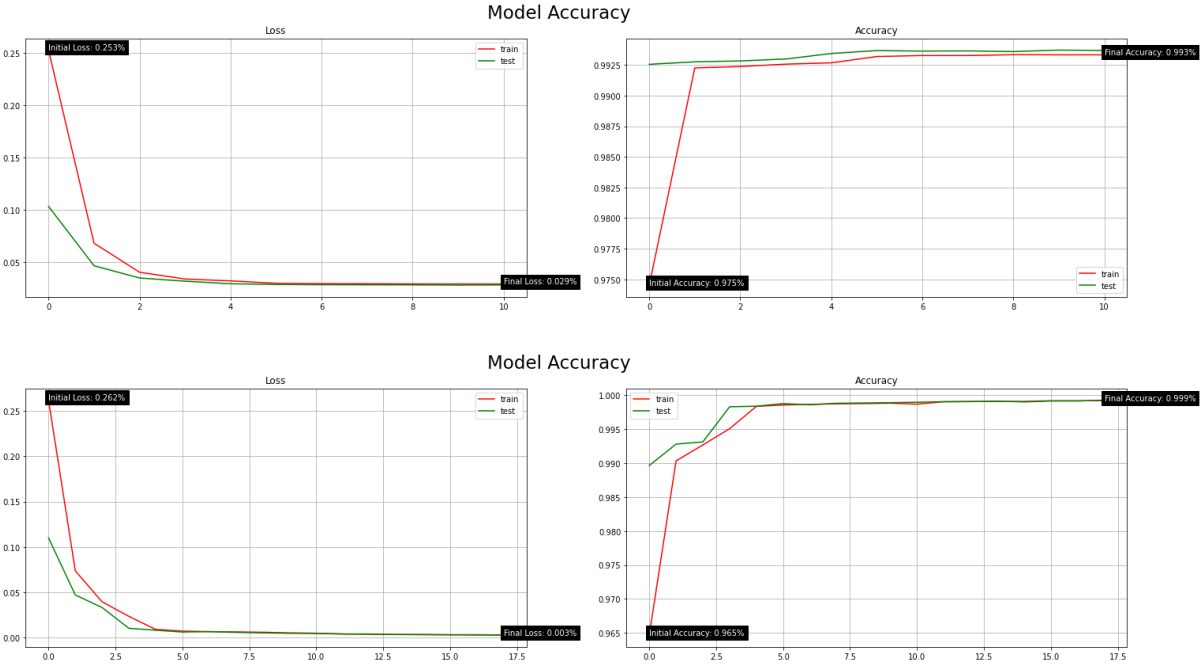
Model Accuracy



Principal Component Analysis (PCA)

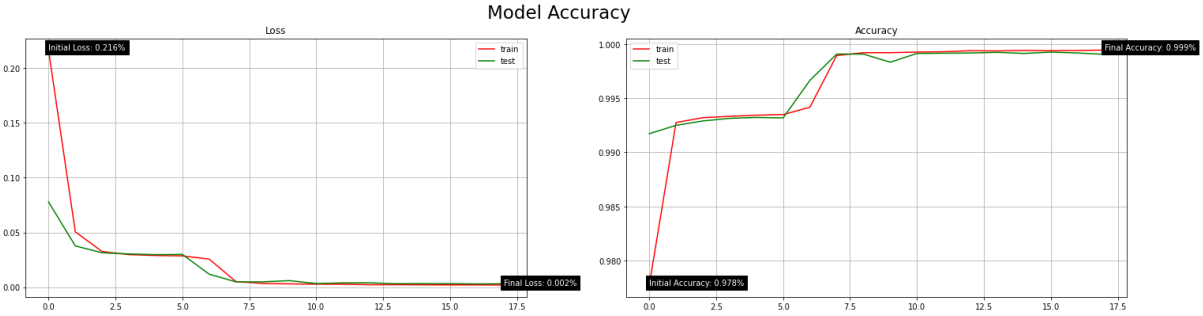
Dimensionality Reduction	Model Loss (Error)	Model Accuracy (%)
30	0.030	99.32%
25	0.031	99.28%
20	0.028	99.37%
15	0.031	99.29%
10	0.031	99.29%
5	0.028	99.34%

- Principal Component Analysis (PCA) activeness (when correlation features not used) alteration

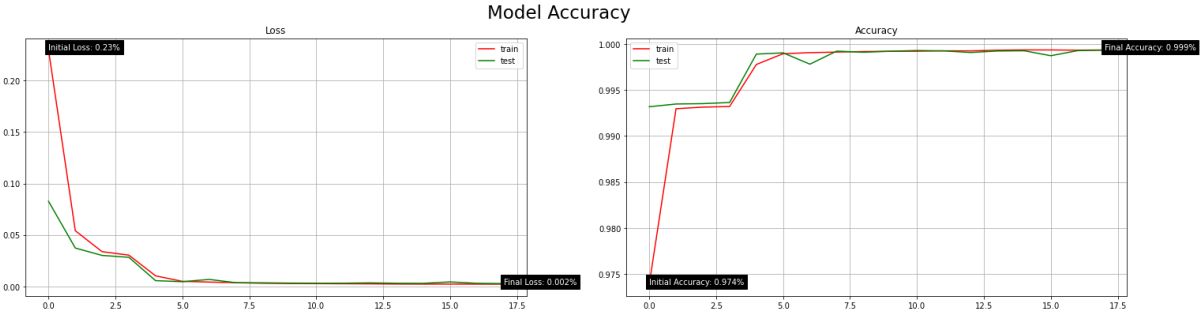


Principal Component Analysis (PCA)		
Active?	Model Loss (Error)	Model Accuracy (%)
True	0.028	99.37%
False	0.003	99.92%

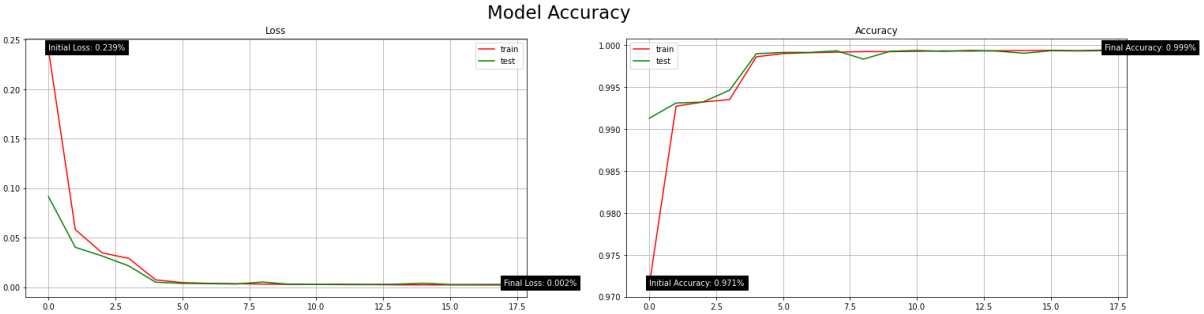
- Training and testing data split alteration



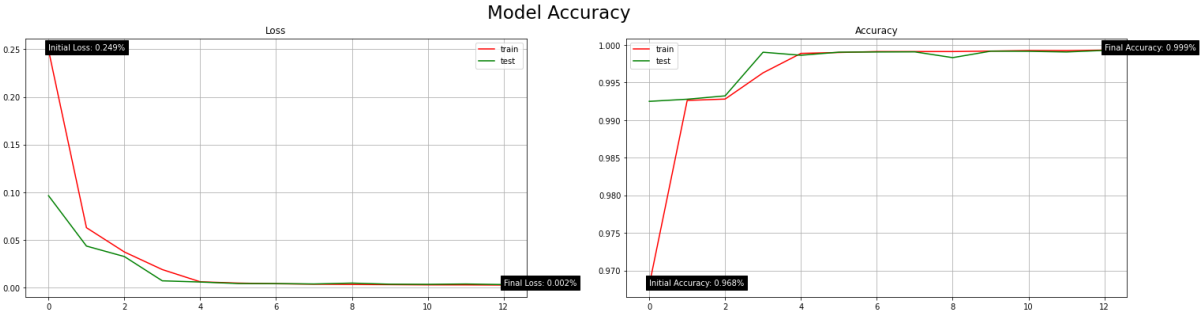
Model Accuracy



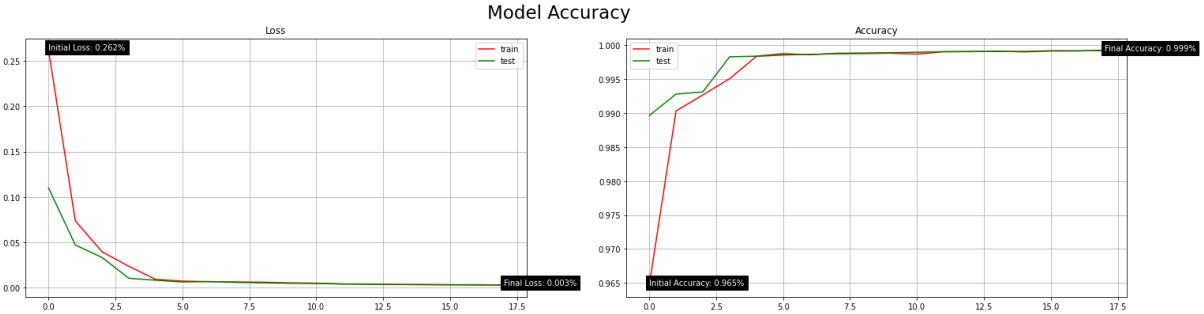
Model Accuracy



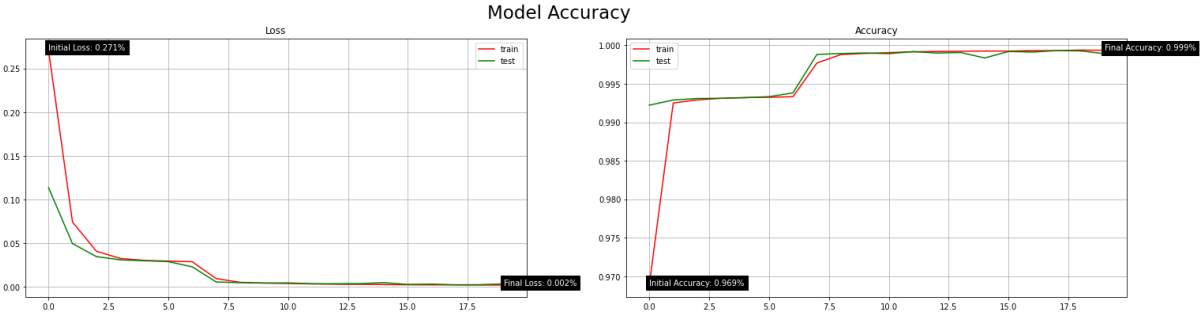
Model Accuracy

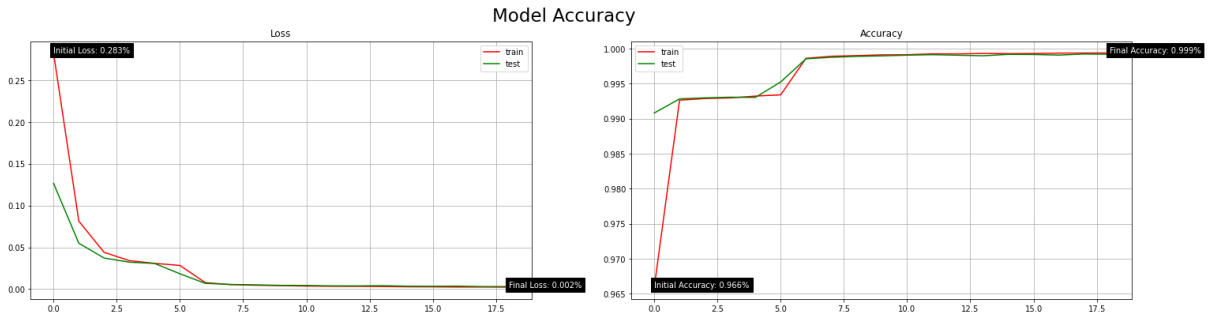


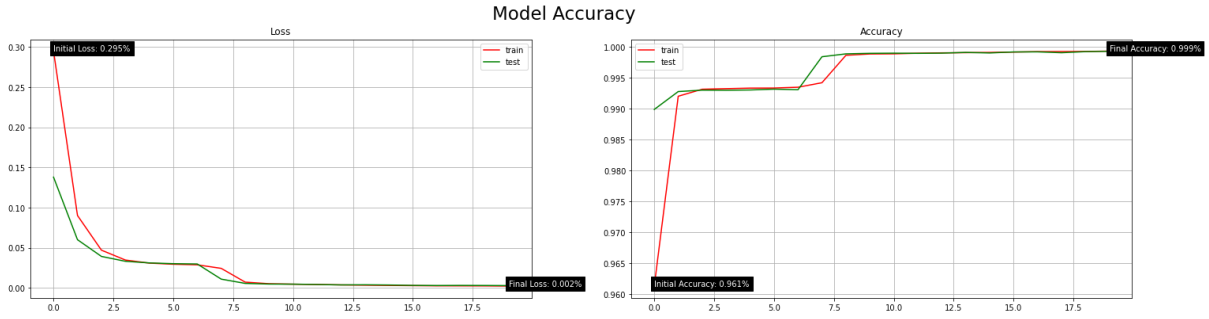
Model Accuracy



Model Accuracy



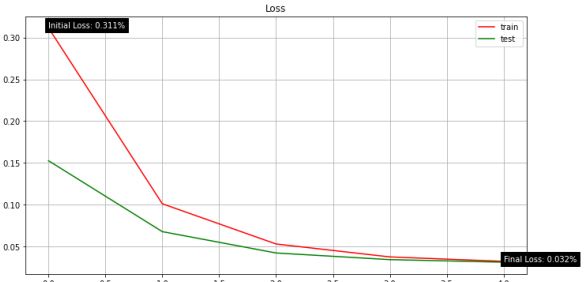


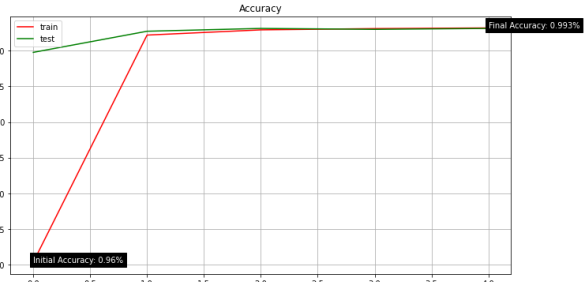


Training and Testing Dataset Partition		
Dataset Partition (%)	Model Loss (Error)	Model Accuracy (%)
0.9 | 0.1	0.004	99.92%
0.85 | 0.15	0.003	99.91%
0.8 | 0.2	0.003	99.94%
0.75 | 0.25	0.004	99.91%
0.7 | 0.3	0.003	99.92%
0.65 | 0.35	0.005	99.84%
0.6 | 0.4	0.004	99.90%
0.55 | 0.45	0.004	99.91%
0.5 | 0.5	0.004	99.92%

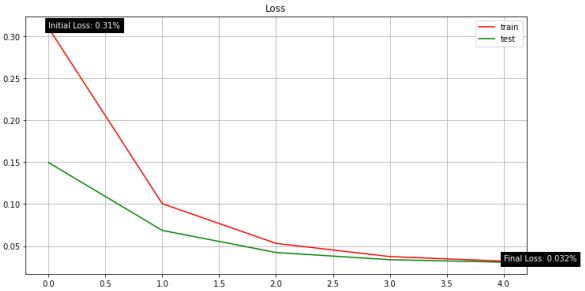
- K-fold Cross Validation (CV) dataset partitioning alteration

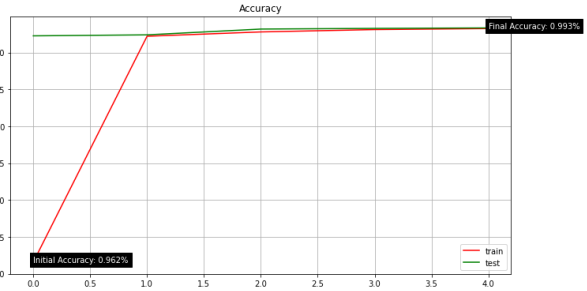
Dataset Fold 1



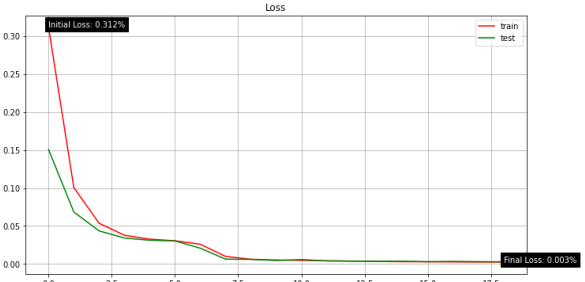


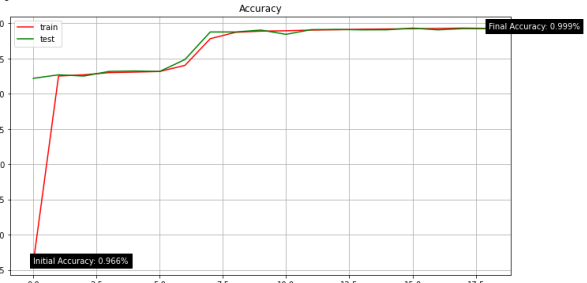
Dataset Fold 2



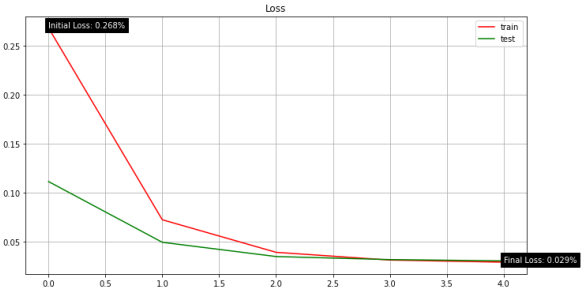


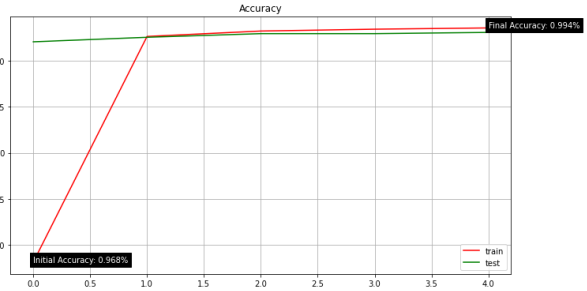
Model Accuracy



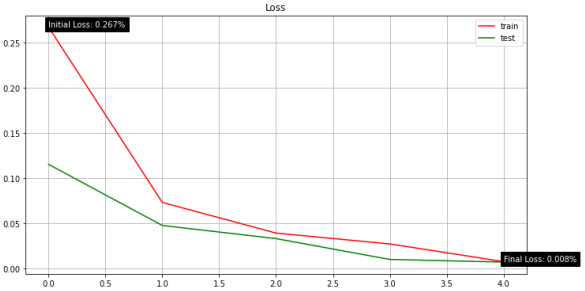


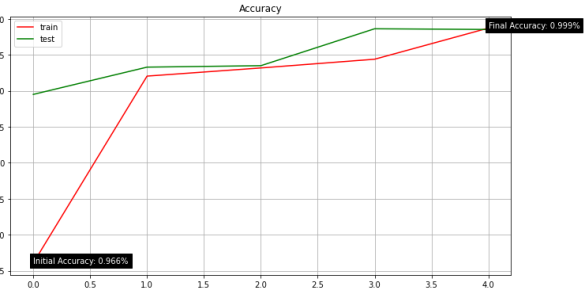
Dataset Fold 1



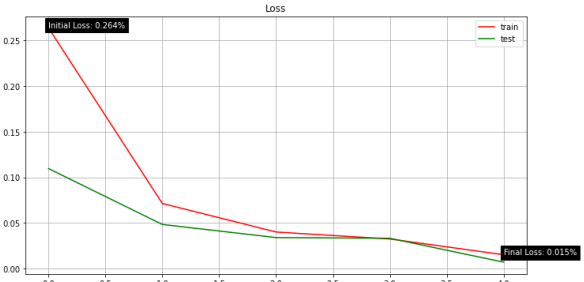


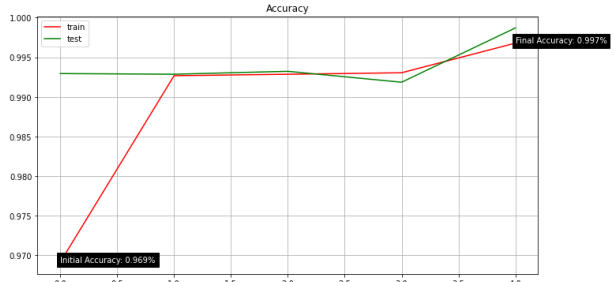
Dataset Fold 2



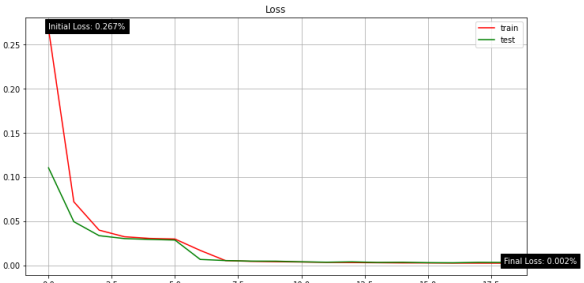


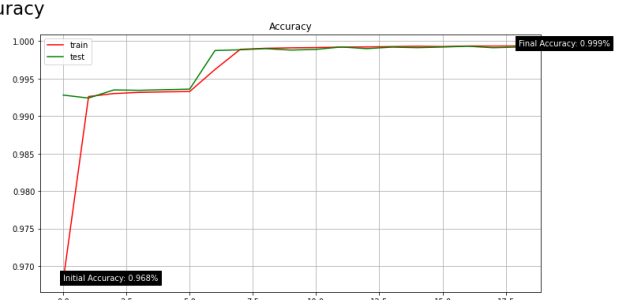
Dataset Fold 3



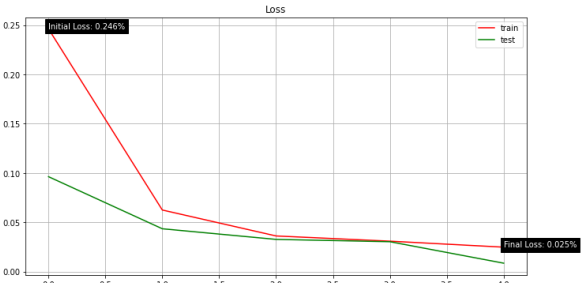


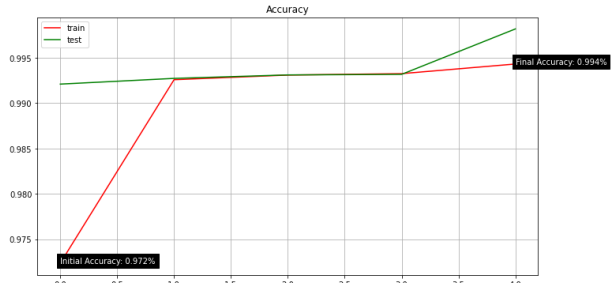
Model Accuracy



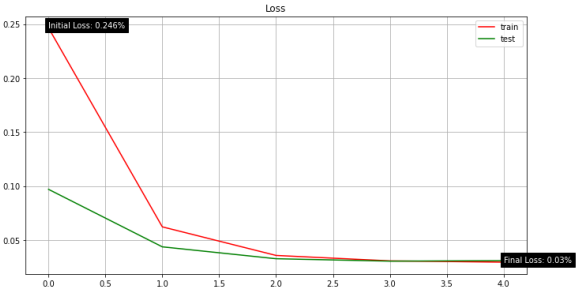


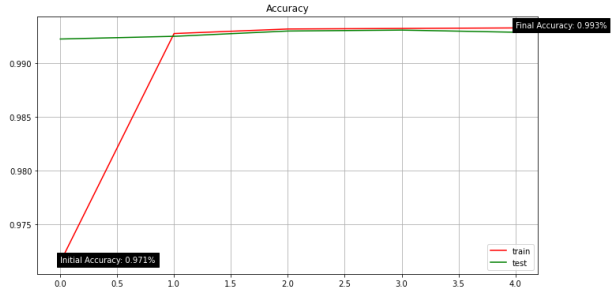
Dataset Fold 1



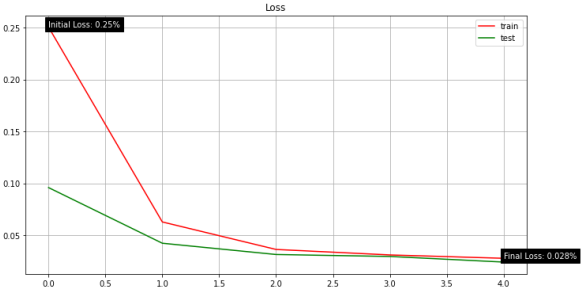


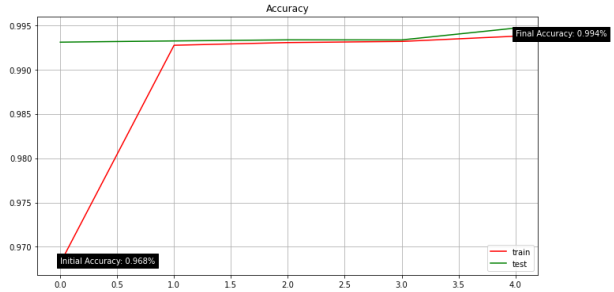
Dataset Fold 2



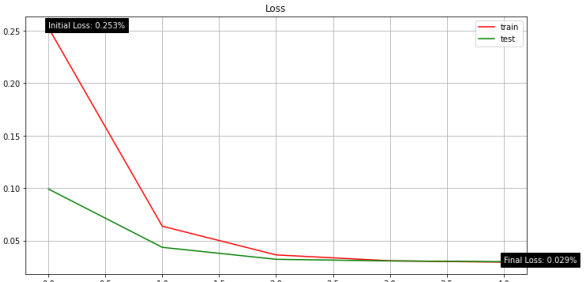


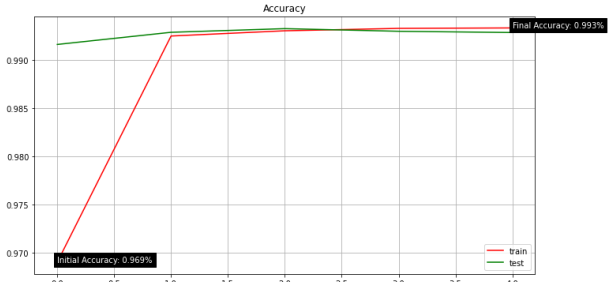
Dataset Fold 3



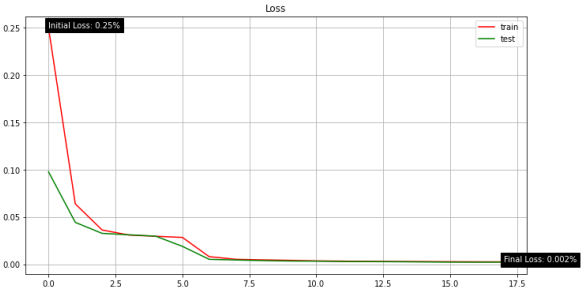


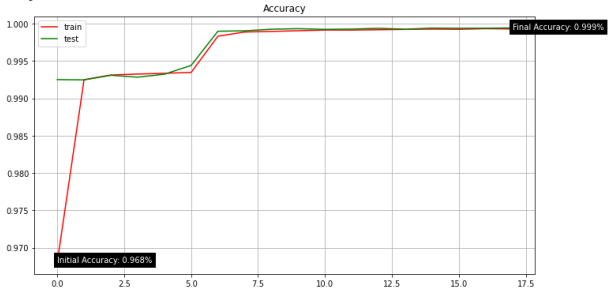
Dataset Fold 4



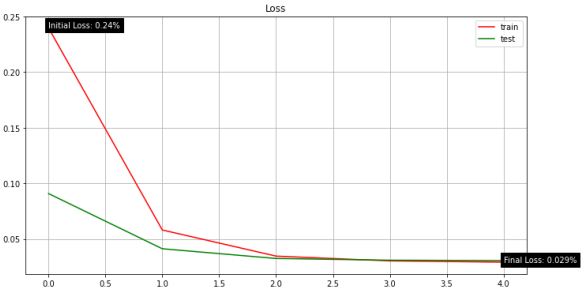


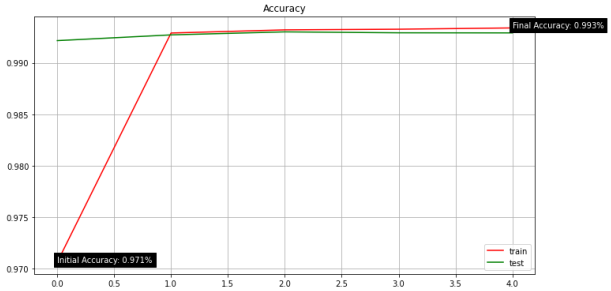
Model Accuracy



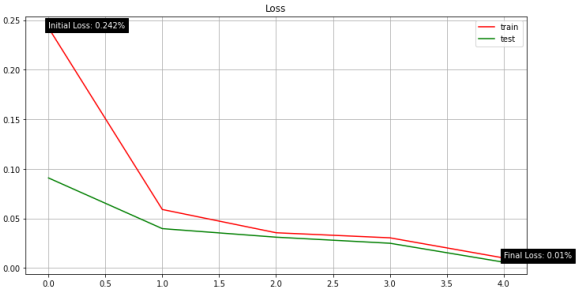


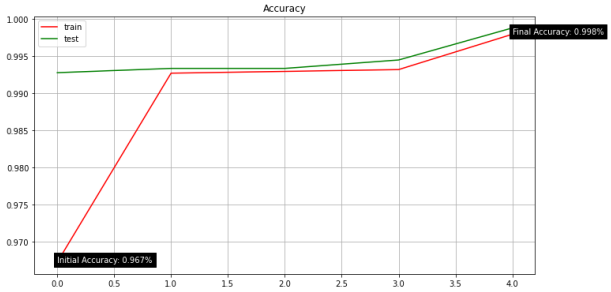
Dataset Fold 1



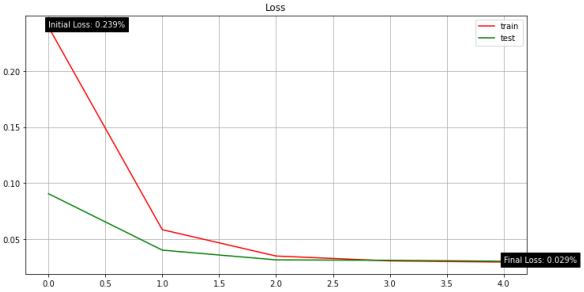


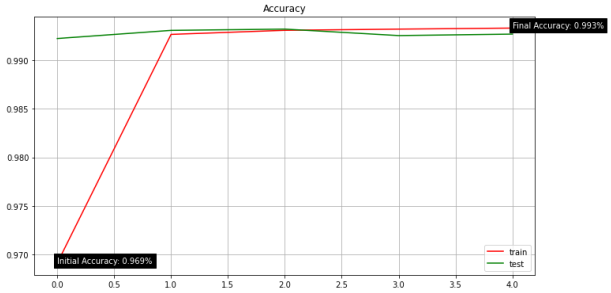
Dataset Fold 2



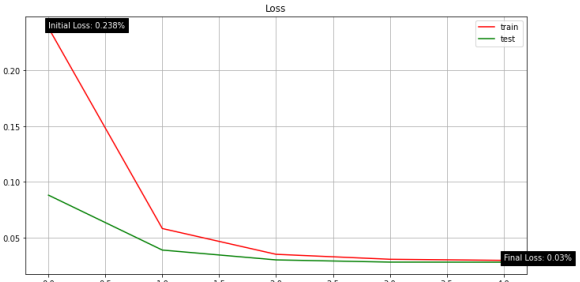


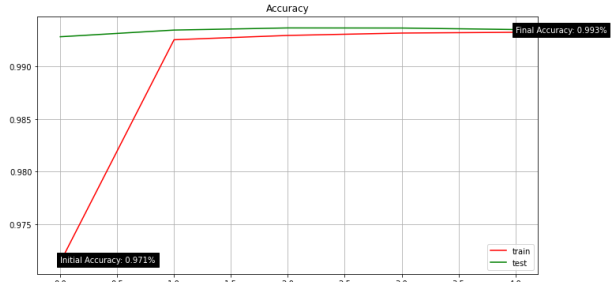
Dataset Fold 3



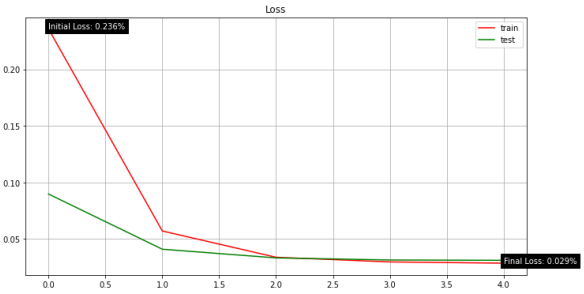


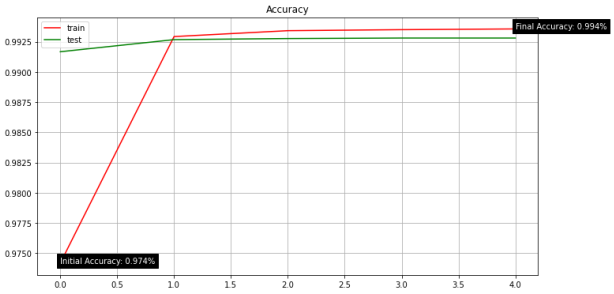
Dataset Fold 4



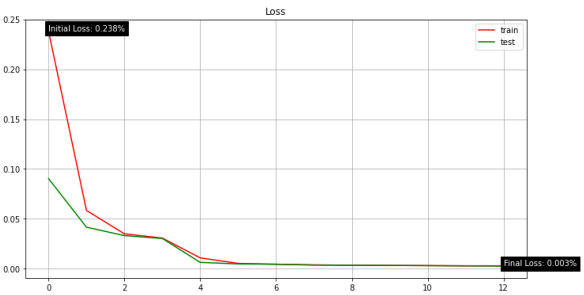


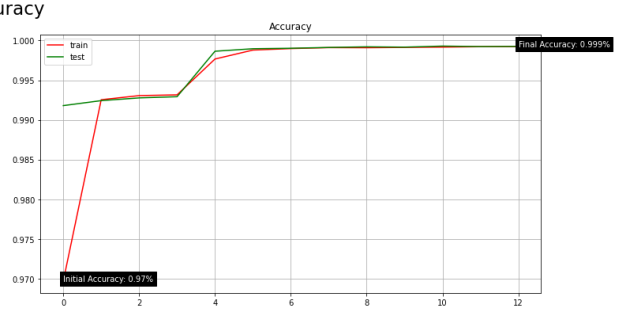
Dataset Fold 5



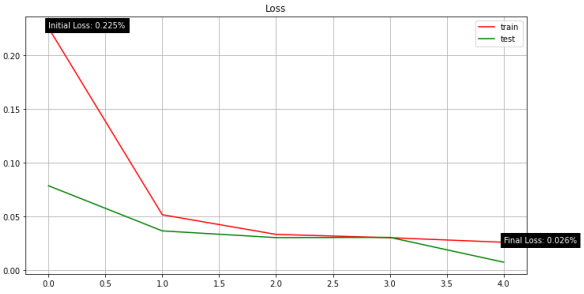


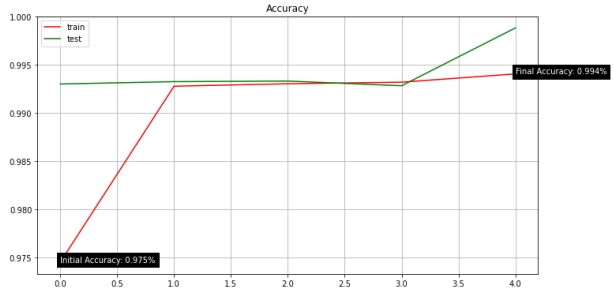
Model Accuracy



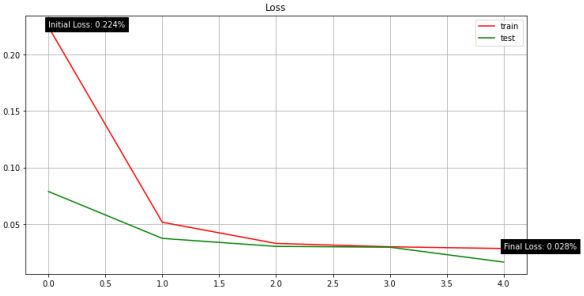


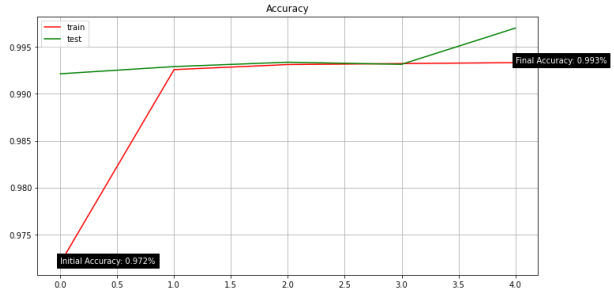
Dataset Fold 1



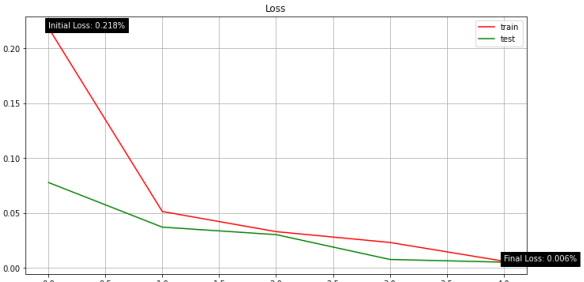


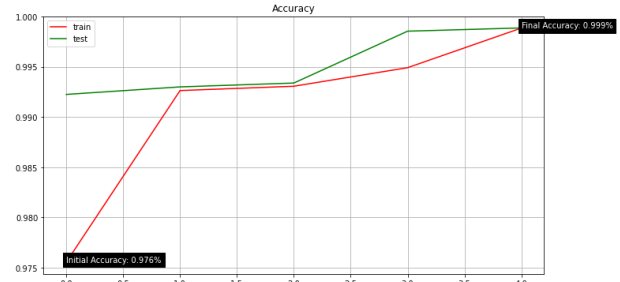
Dataset Fold 2



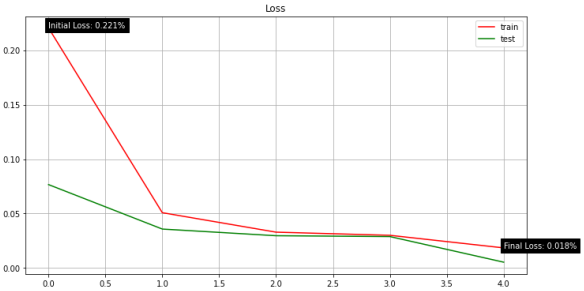


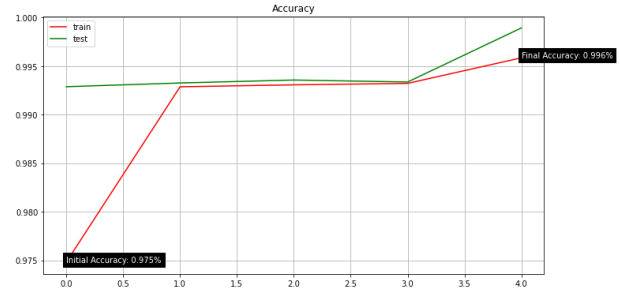
Dataset Fold 3



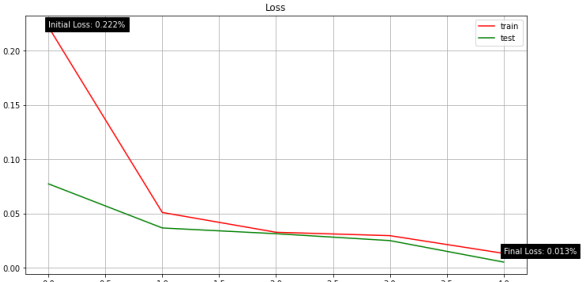


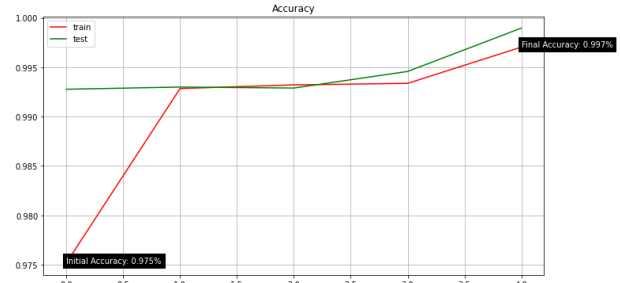
Dataset Fold 4



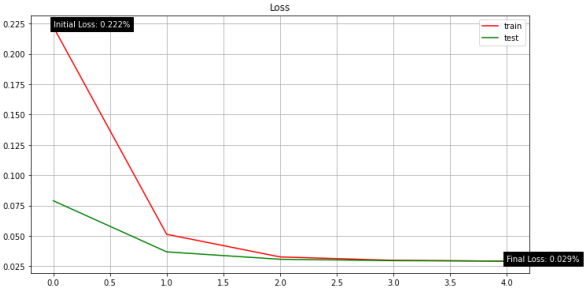


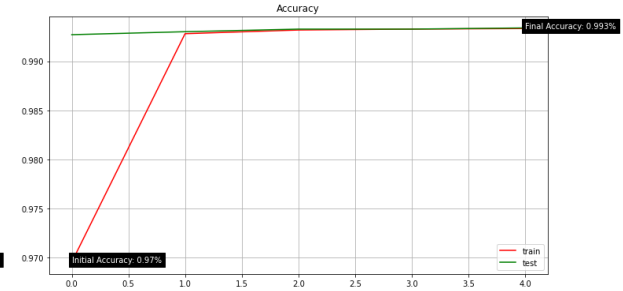
Dataset Fold 5



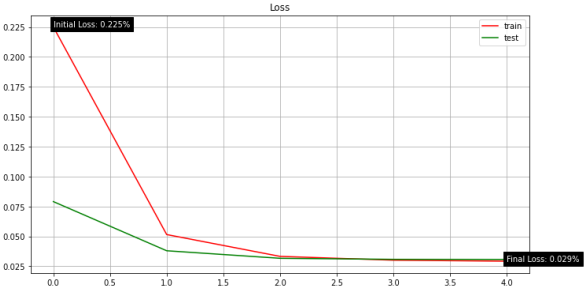


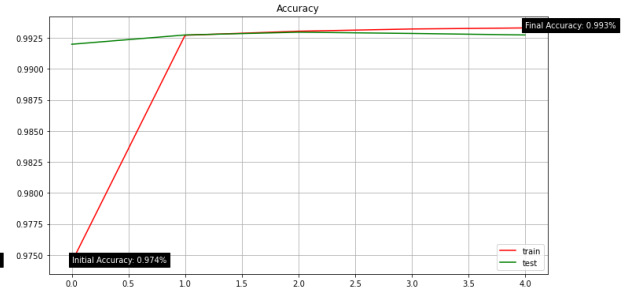
Dataset Fold 6



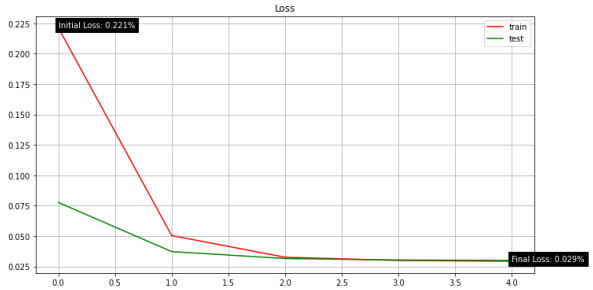


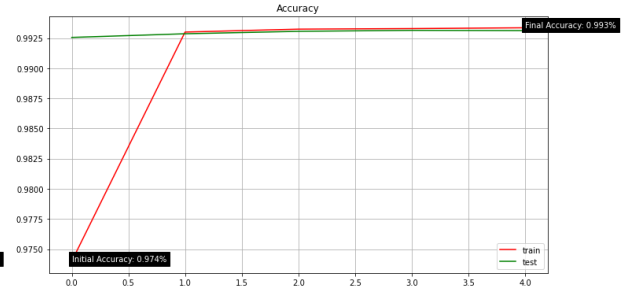
Dataset Fold 7



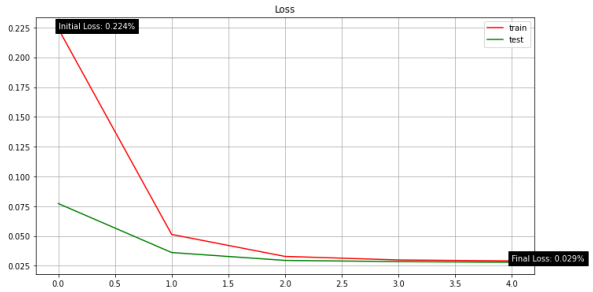


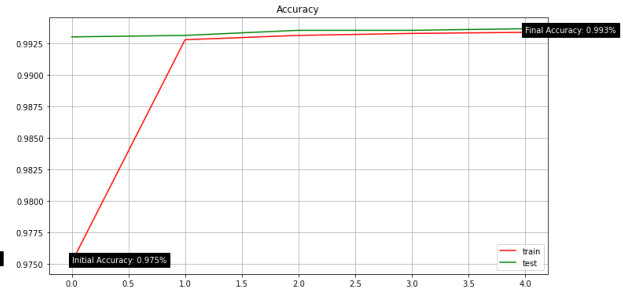
Dataset Fold 8



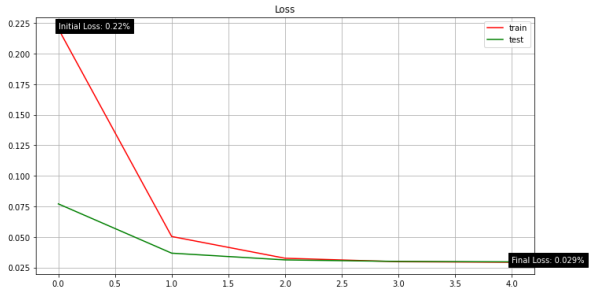


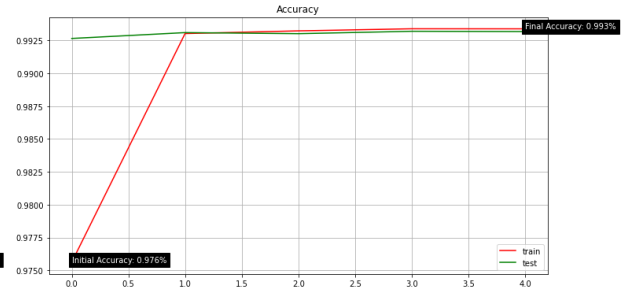
Dataset Fold 9



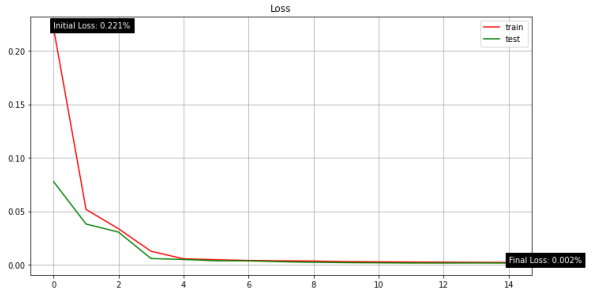


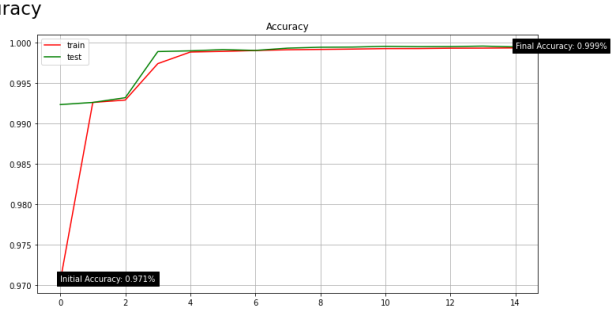
Dataset Fold 10





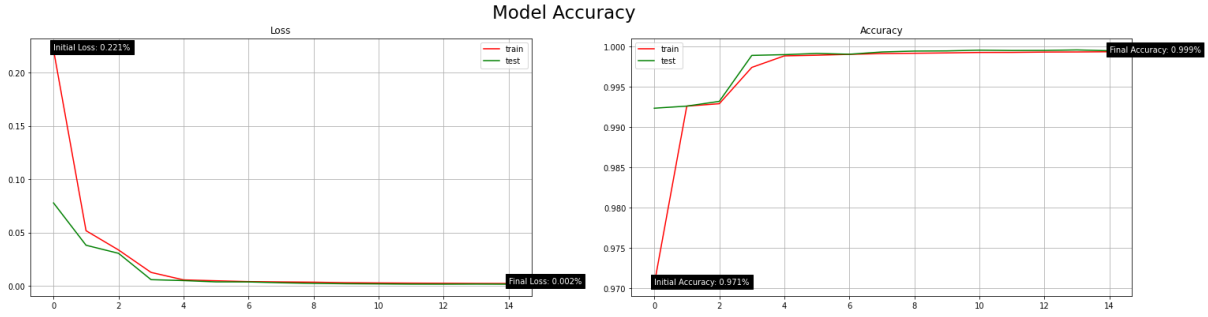
Model Accuracy

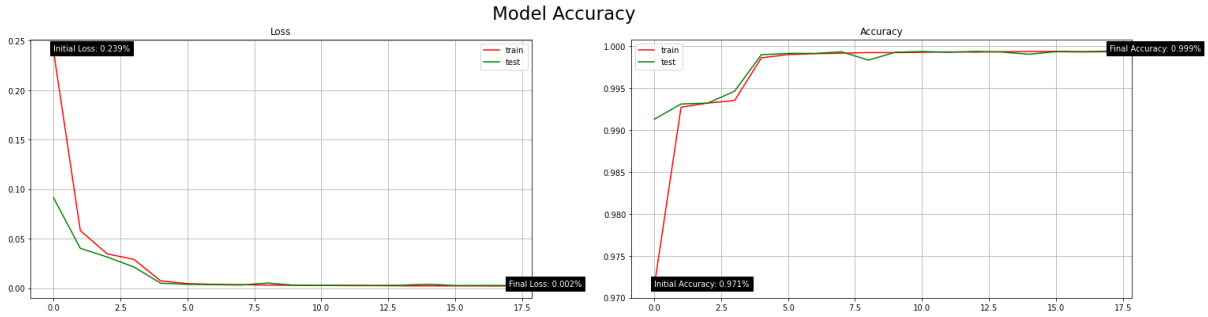




K-fold Cross Validation (CV)		
K-folds	Model Loss (Error)	Model Accuracy (%)
2	0.004	99.91%
3	0.003	99.92%
4	0.003	99.94%
5	0.004	99.91%
10	0.002	99.95%

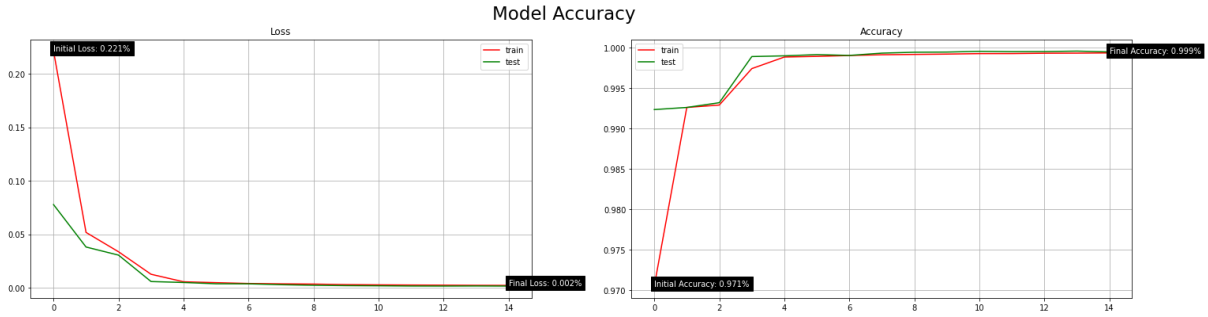
- K-fold Cross-validation (CV) activeness alteration

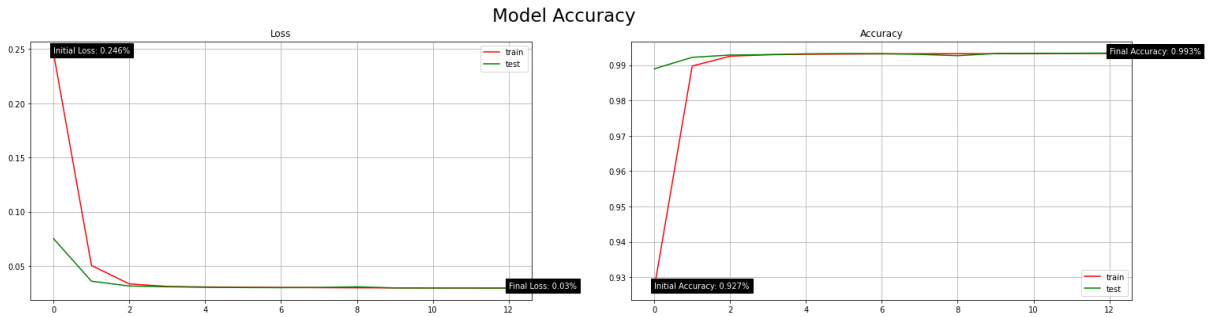


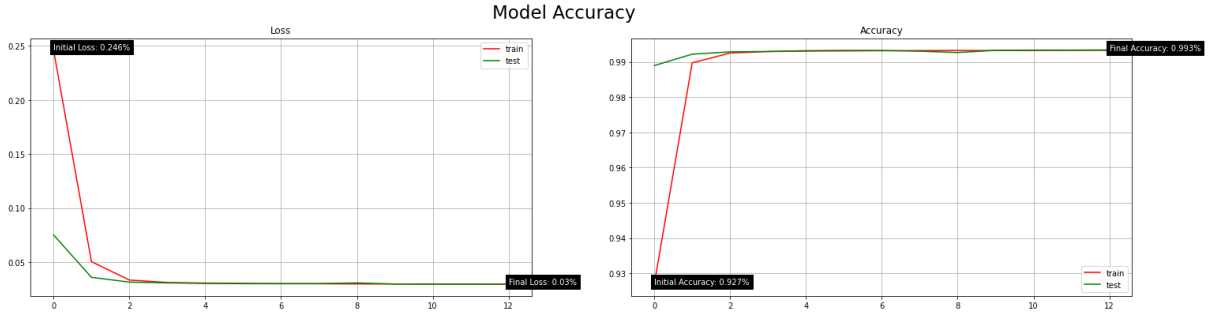


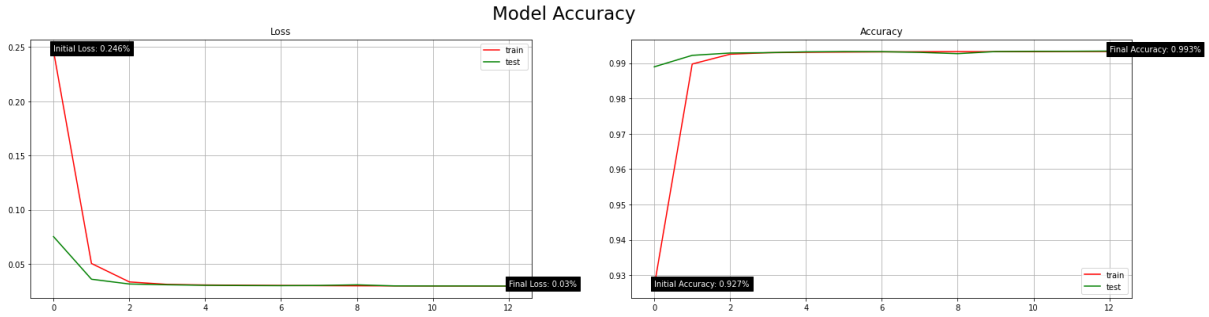
K-fold Cross Validation (CV)		
Active?	Model Loss (Error)	Model Accuracy (%)
True	0.002	99.95%
False	0.003	99.94%

- Active model optimiser algorithm alteration



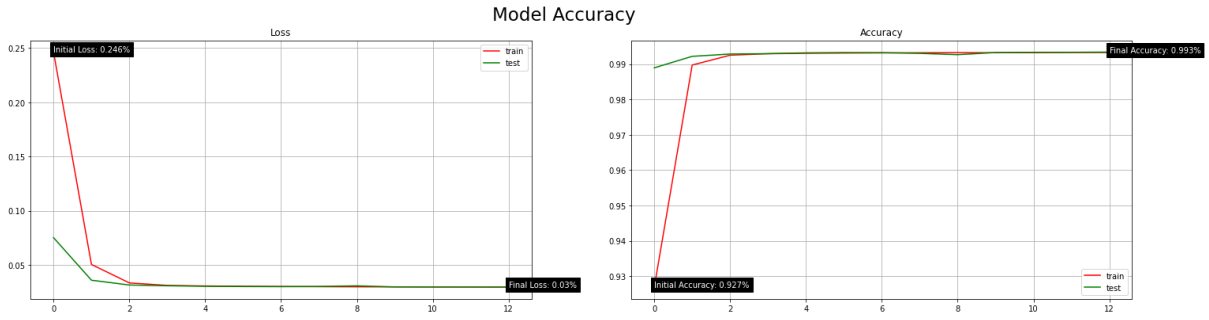


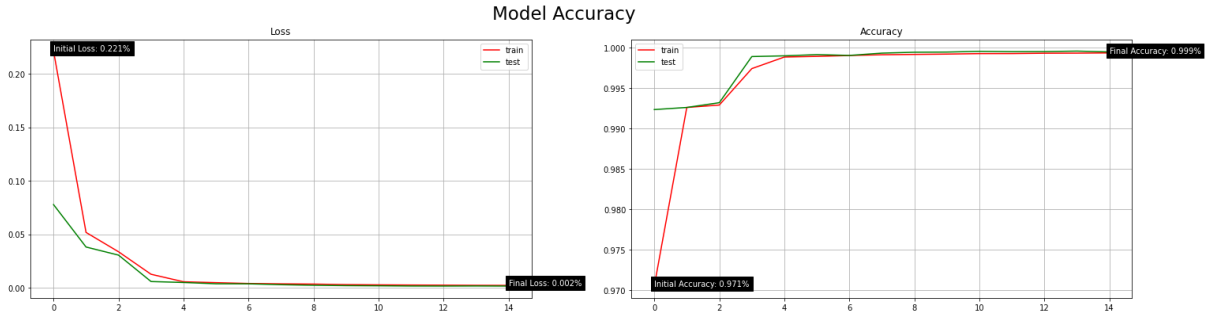


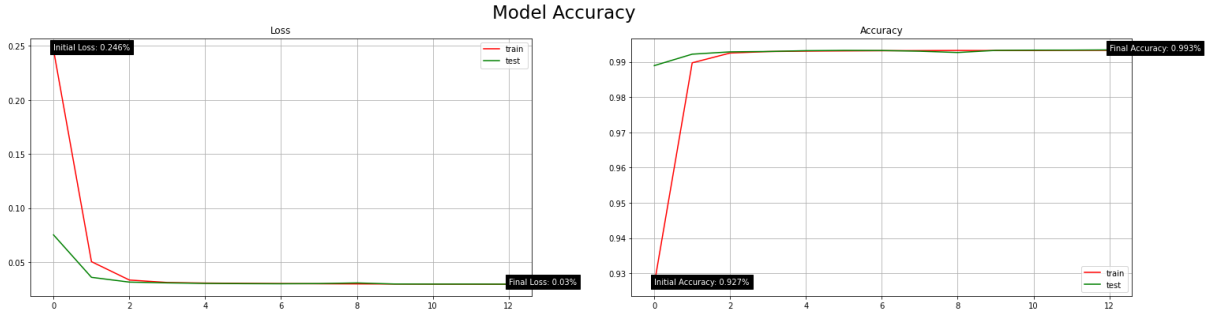


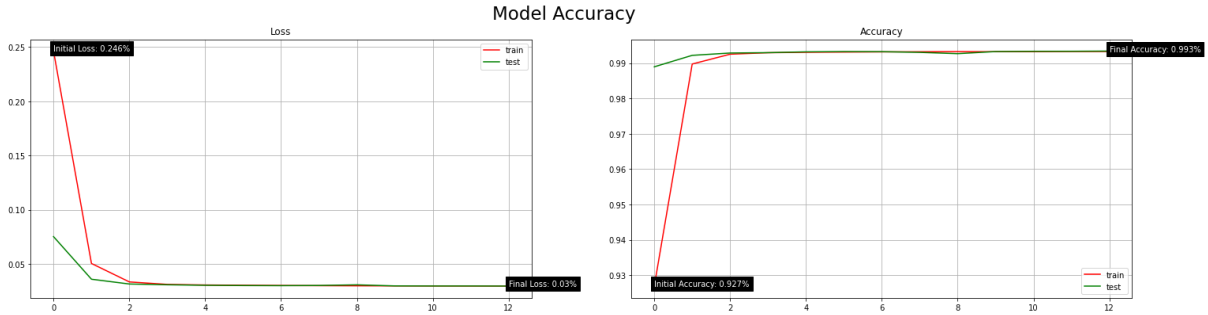
Model Optimiser Algorithm		
Optimiser Algorithm	Model Loss (Error)	Model Accuracy (%)
Adam	0.002	99.95%
Adamax	0.030	99.30%
Nadam	0.007	99.89%
RMSprop	0.493	80.56%

- Training dataset batch size alteration



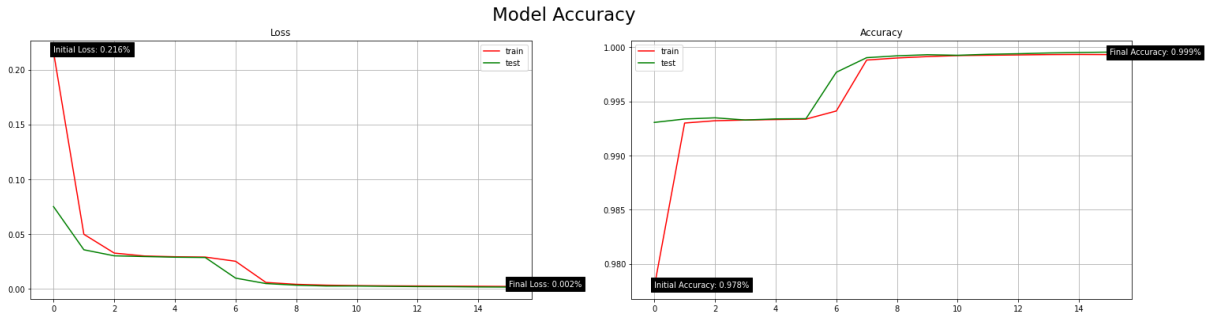


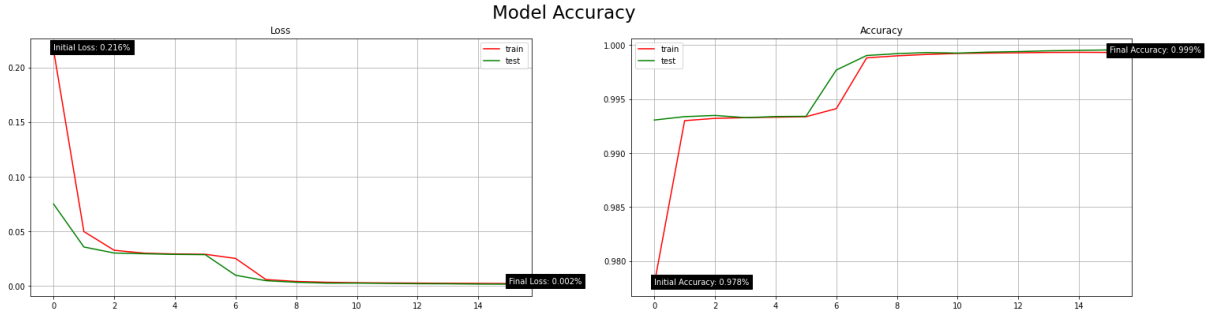


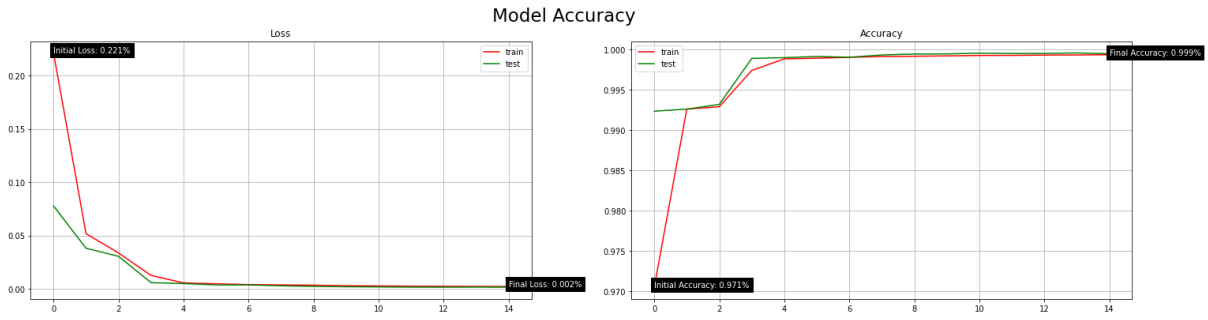


Training Dataset Sample (Batch) Size		
Batch Sample Size	Model Loss (Error)	Model Accuracy (%)
256	0.003	99.94%
128	0.002	99.95%
64	0.003	99.91%
32	0.003	99.92%

- Fully-connected layer (FCL) count in the model (when model topology is [32, 32, 32, 5, 5]) alteration

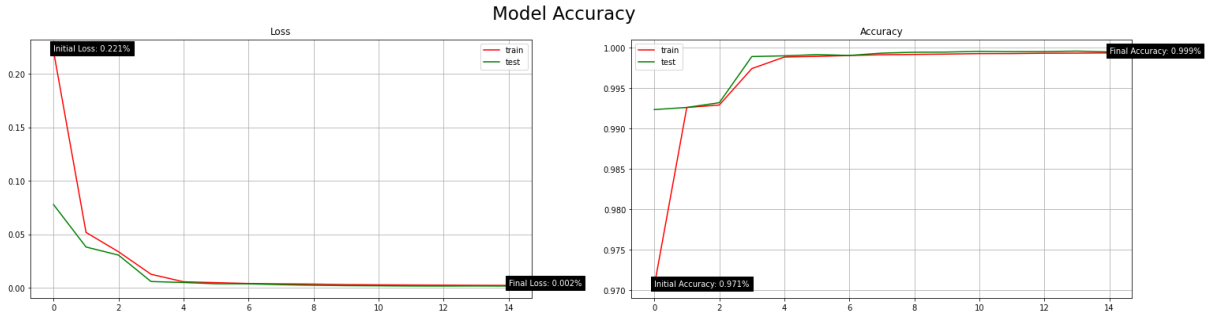


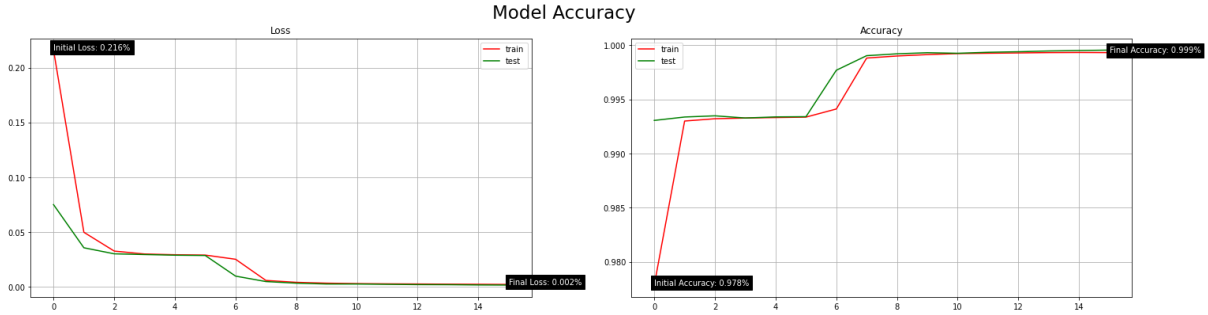


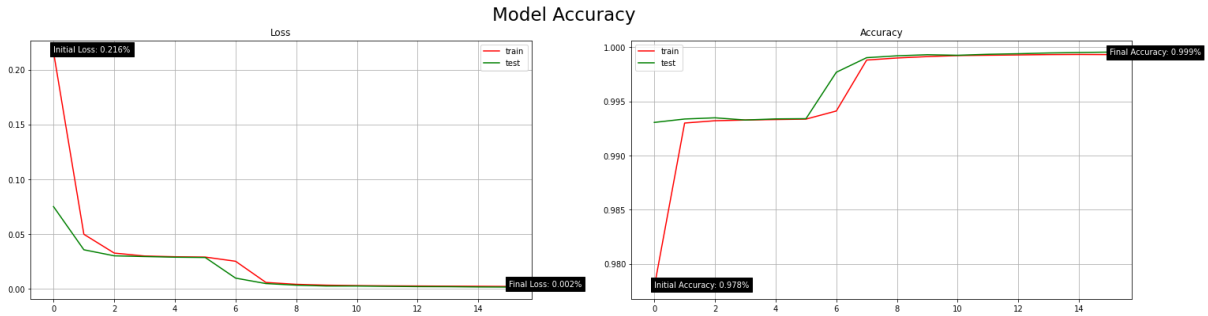


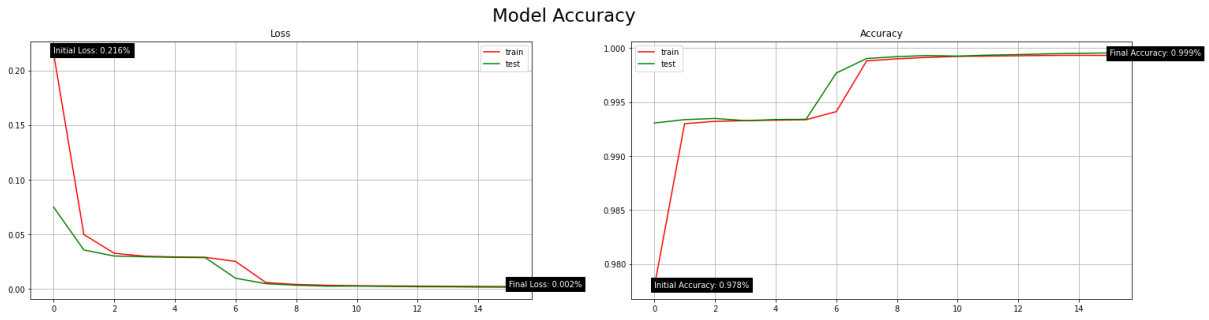
Model Fully-Connected Layer (FCL) Count		
FCL Count	Model Loss (Error)	Model Accuracy (%)
3	0.003	99.92%
4	0.003	99.92%
5	0.002	99.95%

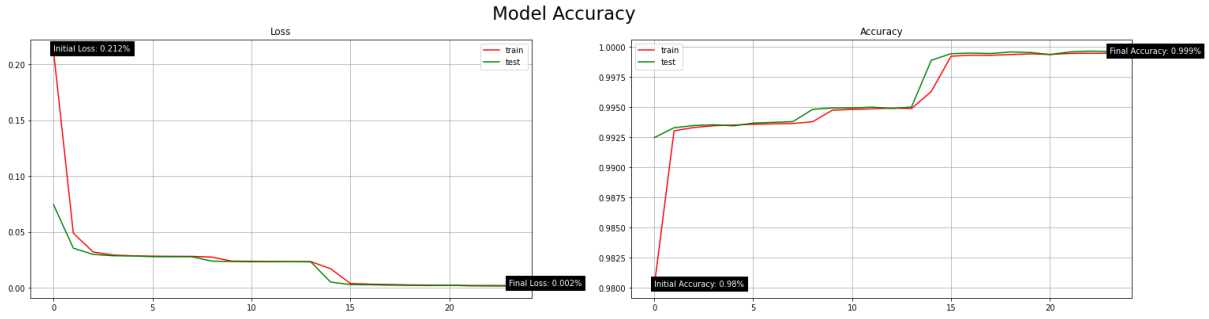
- Model topology (neuron counts) alteration

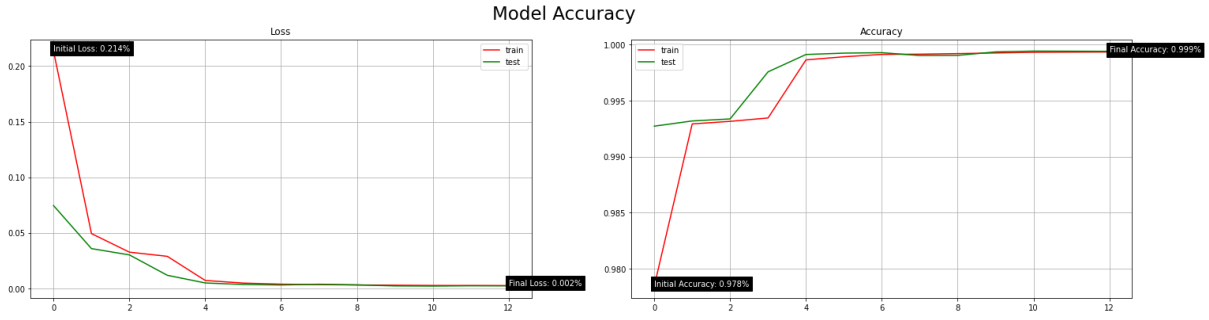


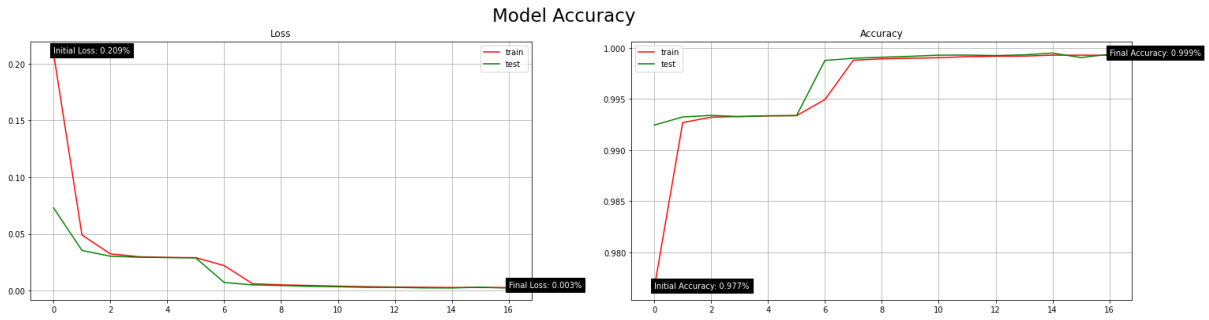








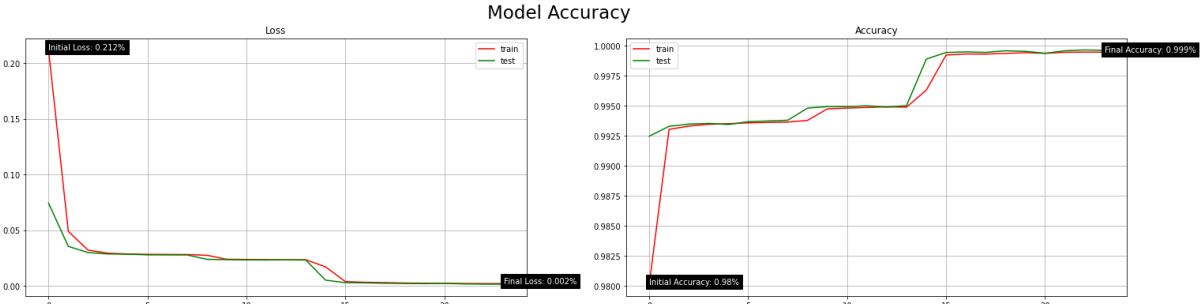




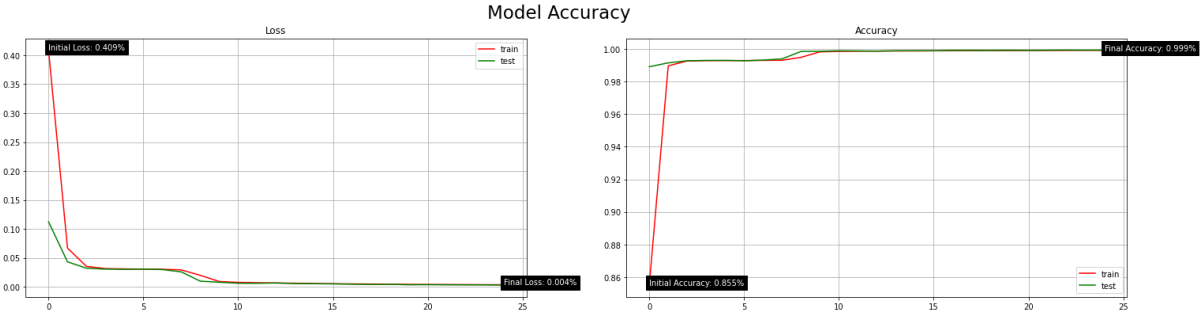
Model Topology		
Neurons Per Layer	Model Loss (Error)	Model Accuracy (%)
32, 32, 32, 5, 5	0.002	99.95%
80, 80, 80, 80, 5	0.002	99.95%
70, 70, 70, 70, 5	0.003	99.93%
60, 60, 60, 60, 5	0.002	99.95%
50, 50, 50, 50, 5	0.002	99.96%
40, 40, 40, 40, 5	0.004	99.90%
30, 30, 30, 30, 5	0.003	99.93%

- Model layer transfer function alteration

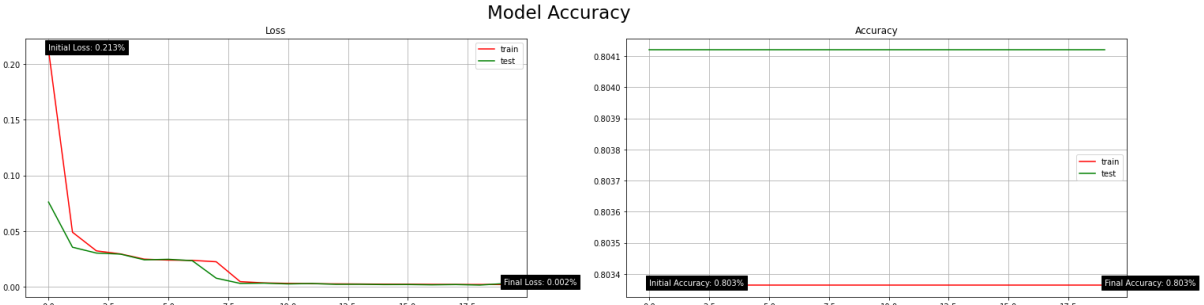
Model Accuracy



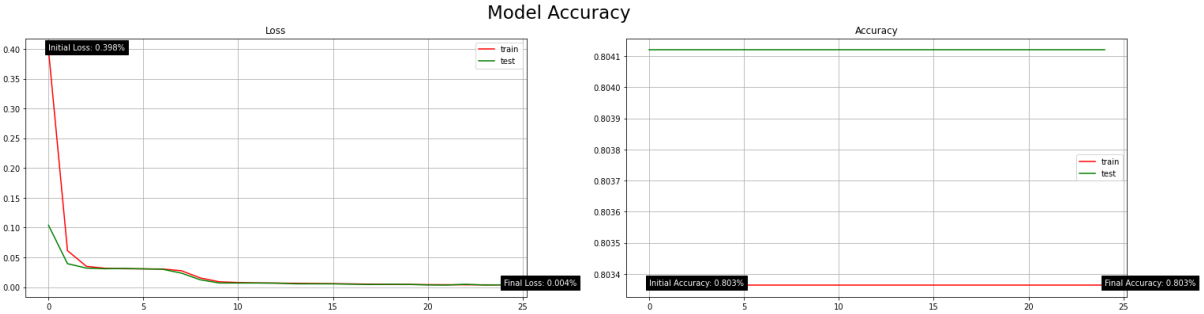
Model Accuracy



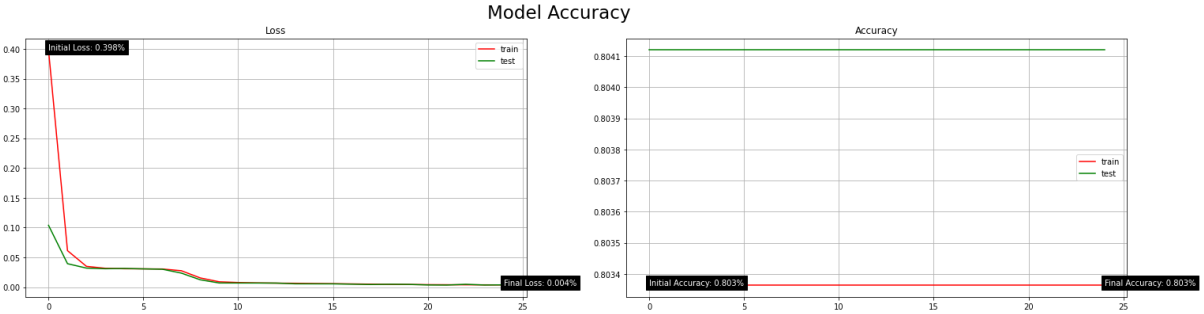
Model Accuracy

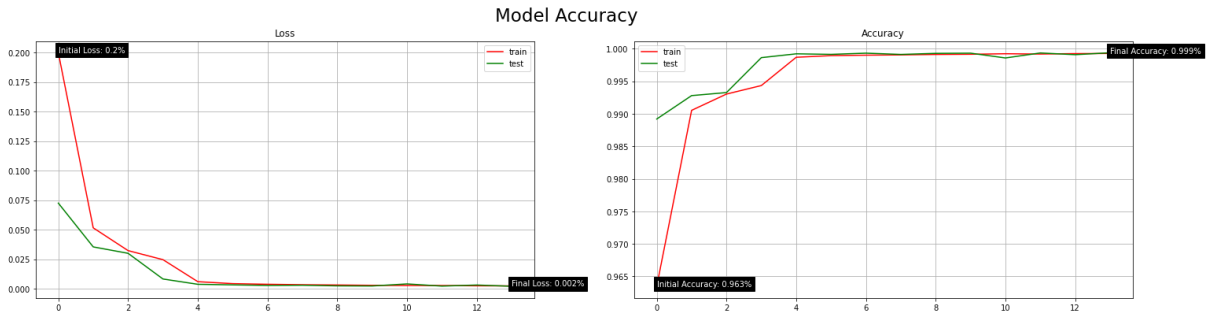


Model Accuracy



Model Accuracy

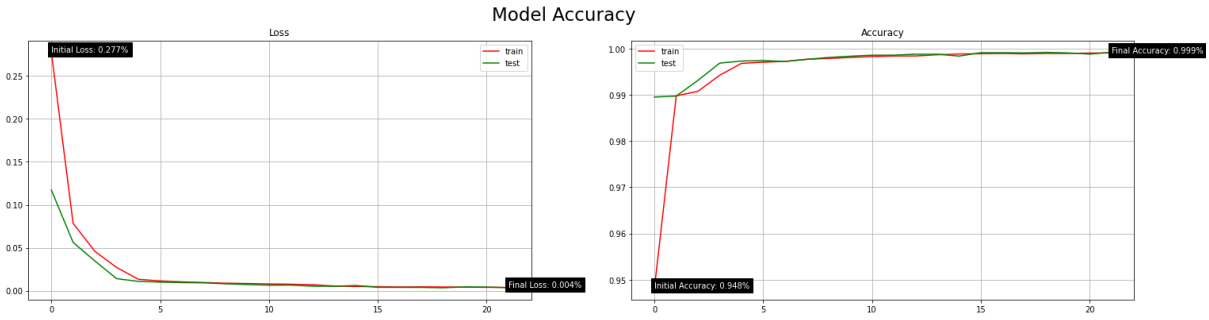




Model Layer Transfer (Activation) Functions		
Transfer Function Per layer	Model Loss (Error)	Model Accuracy (%)
relu, relu, relu, softmax, sigmoid	0.002	99.96%
sigmoid, sigmoid, sigmoid, softmax, sigmoid	0.004	99.92%
relu, relu, relu, softmax, softmax	0.002	80.41%
sigmoid, sigmoid, sigmoid, softmax, softmax	0.005	80.41%
relu, relu, relu, relu, softmax	0.005	80.41%
relu, softmax, relu, softmax, sigmoid	0.003	99.93%

Appendix C

Multiclass Classification Model Performance



Binary Classification Model Performance