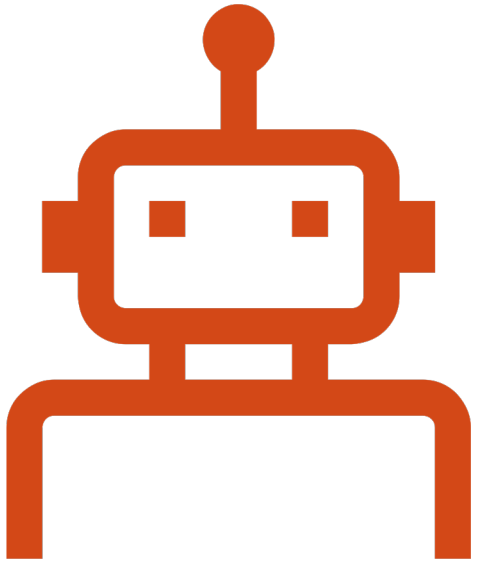
ITC05F MACHINE LEARNING PROJECT 2

Group 2	**1**	**Hoang Tuan Linh**	**m5232108**	**(presenter)**
	2	Tran Thi Thoa	s1242006	

CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - Autoencoder-based classifier
  - Boltzmann machine-based classifier
  - Ensemble learning of decision trees
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - Results on CIFAR-10 dataset
- **Conclusions**



DATASETS

Fashion-MNIST

T-shirt/top : 0
Trouser : 1
Pullover : 2
Dress : 3
Coat : 4
Sandal : 5
Shirt : 6
Sneaker : 7
Bag : 8
Ankle boot : 9

70,000 28x28 grayscale images
(60,000 for training; 10,000 for testing)

10 classes

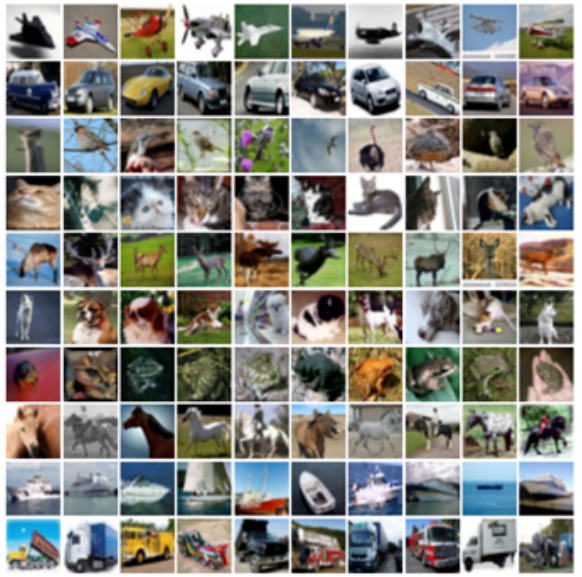
8-bit **grayscale** images
(256 steps from white "0" to black "255")

Size: 150.2 MB

Source: https://www.kaggle.com/zalando-research/fashionmnist

CIFAR-10

airplane : 0
automobile : 1
bird : 2
cat : 3
deer : 4
dog : 5
frog : 6
horse : 7
ship : 8
truck : 9



60,000 32x32 colour images
(50,000 for training; 10,000 for testing)

10 classes (6000 images/class)

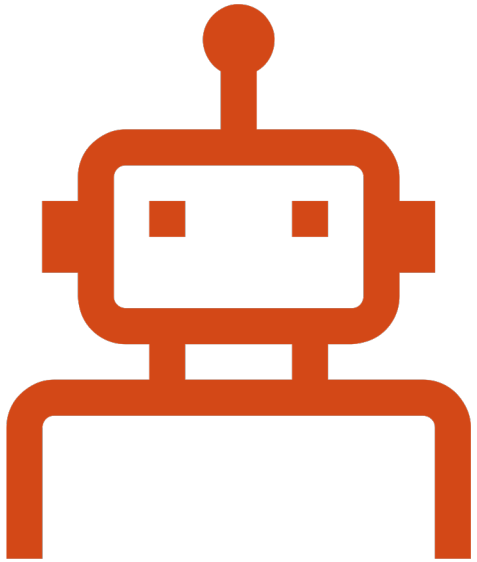
24-bit RGB full-color image (combination of 3 colors of red/green/blue, 256 levels from "0" to "255")

Size: 132.4 MB

Source: https://www.cs.toronto.edu/~kriz/cifar.html

CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - **VGG16-based classifier with transfer learning**
  - Autoencoder-based classifier
  - Boltzmann machine-based classifier
  - Ensemble learning of decision trees
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - Results on CIFAR-10 dataset
- **Conclusions**



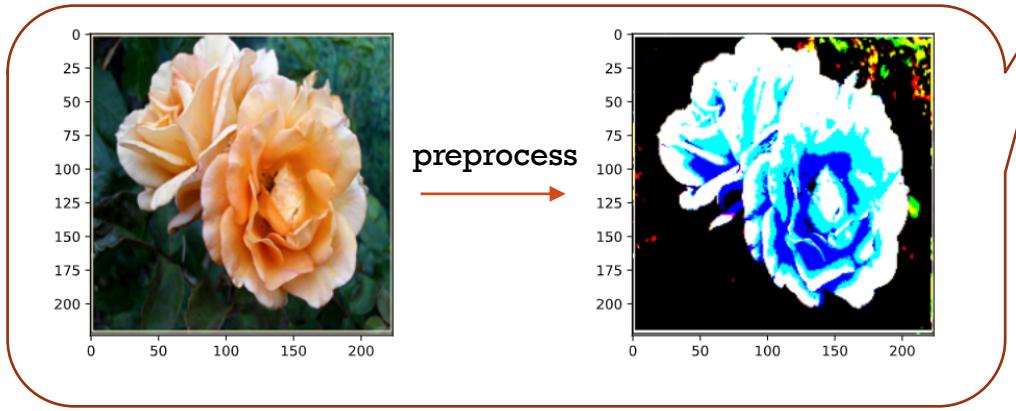
VGG16-BASED CLASSIFIER (TRANSFER LEARNING)

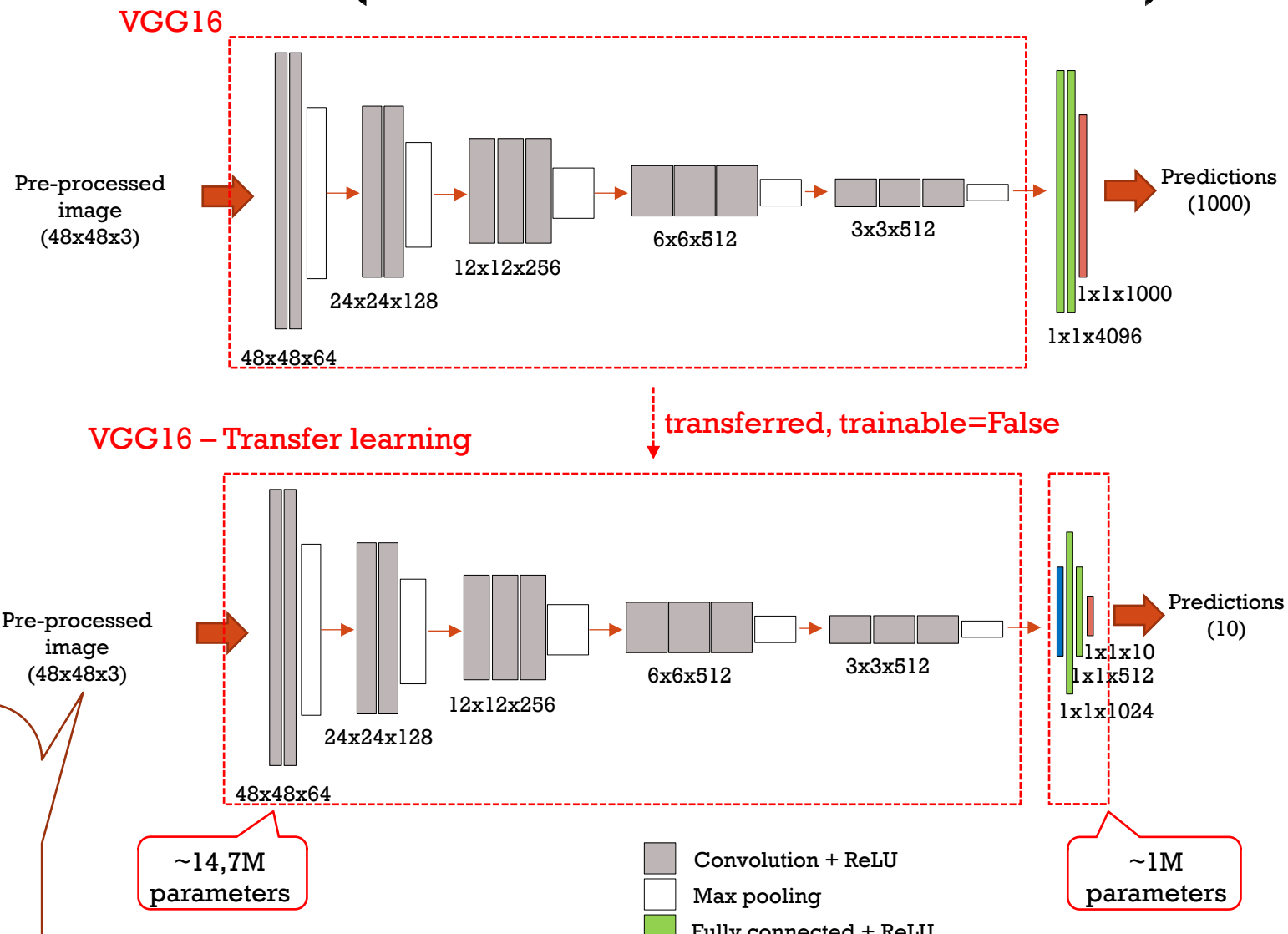
VGG16:

- a CNN model proposed from the University of Oxford in 2014.
- achieves 92.7% top-5 test accuracy on the ImageNet dataset of over **14 million images** belonging to **1000 classes.**

Thanks to transfer learning, we might save a lot of time for training 15.7M parameters.

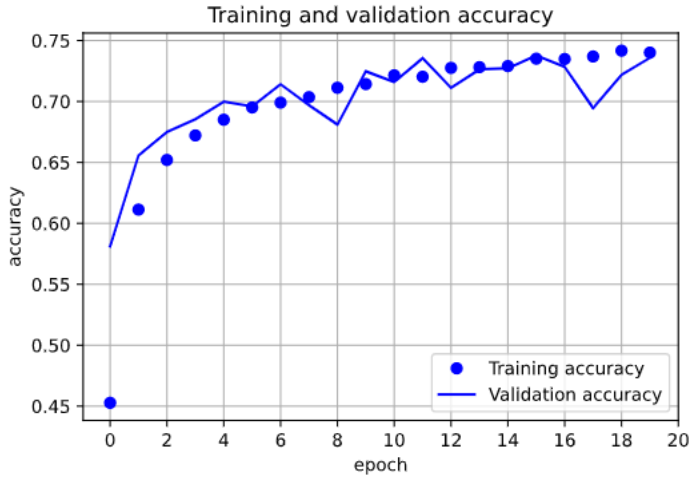
Input image	48x48x3
# of classes	10
Total # of params	15,767,365
Trainable params	1,052,677
Non-trainable params	14,714,688

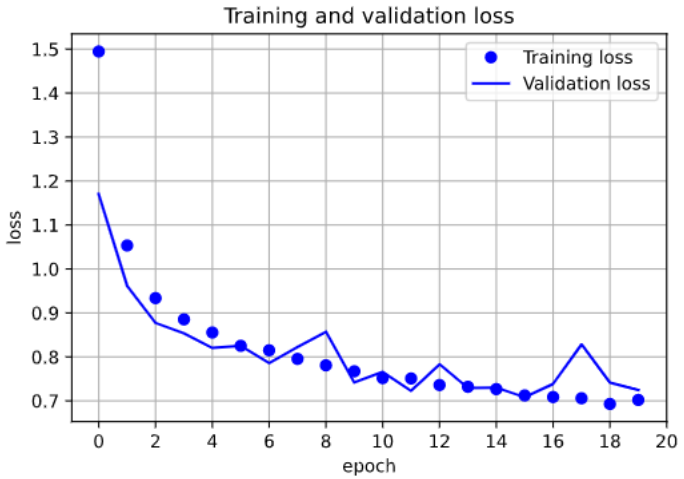




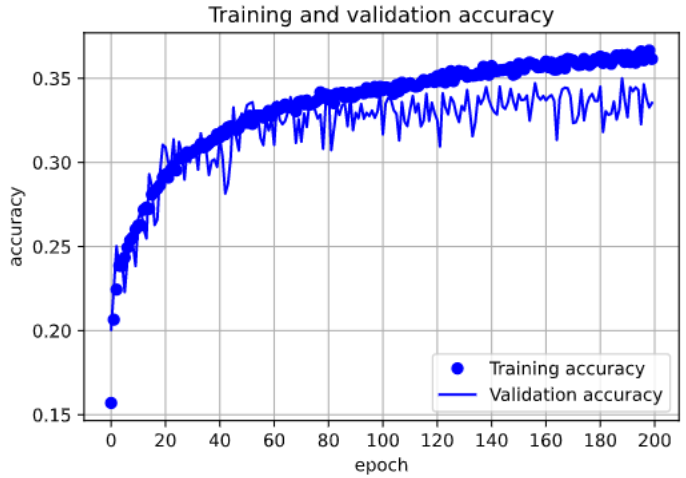
TRAINING OF THE VGG16-BASED CLASSIFIER

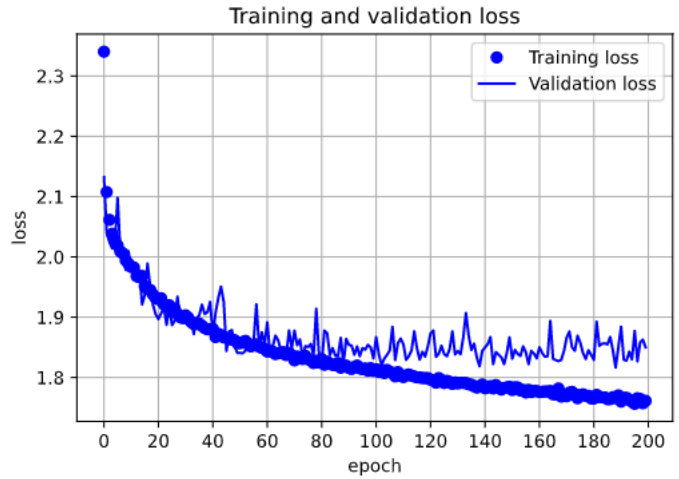
On the Fashion-MNIST dataset





On the CIFAR-10 dataset





Batch size = 64

Validation size = 0.2

Optimizer = Adam

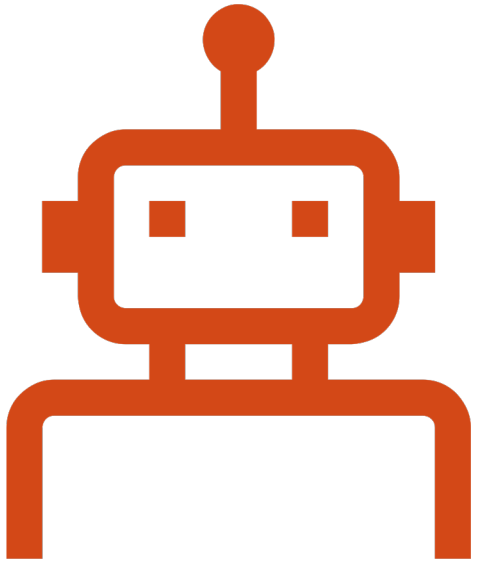
On Fashion-MNIST dataset: the accuracy is quite good with 20 epochs.

However, the VGG16-based classifier is not suitable for CIFAR-10 dataset.

Even with 200 epochs (~15 hrs of training), the accuracy is still not good.

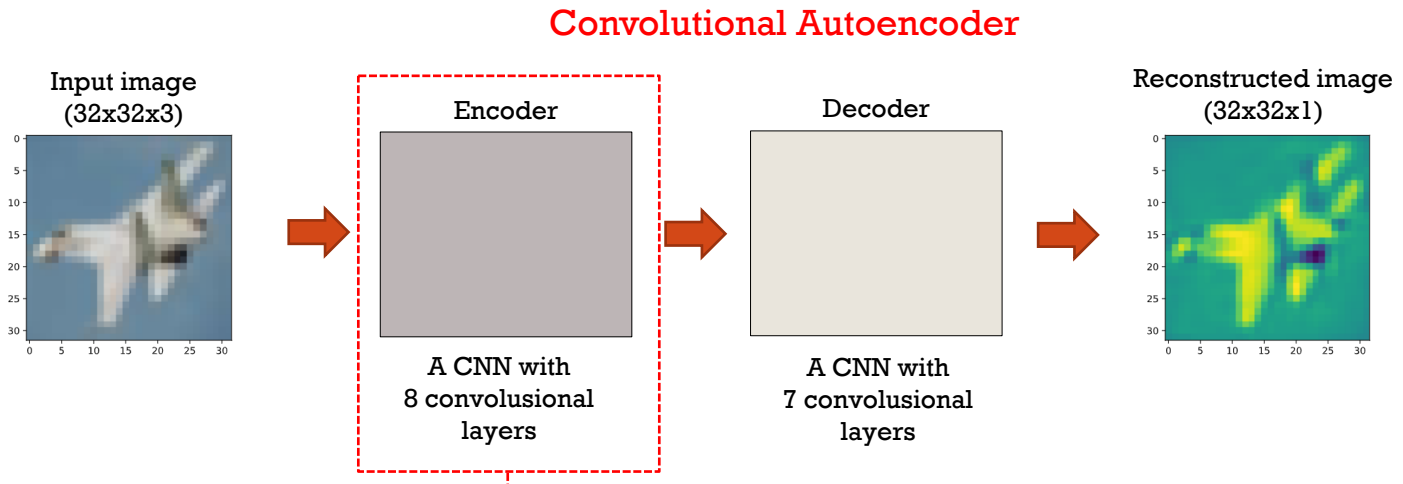
CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - **Autoencoder-based classifier**
  - Boltzmann machine-based classifier
  - Ensemble learning of decision trees
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - Results on CIFAR-10 dataset
- **Conclusions**

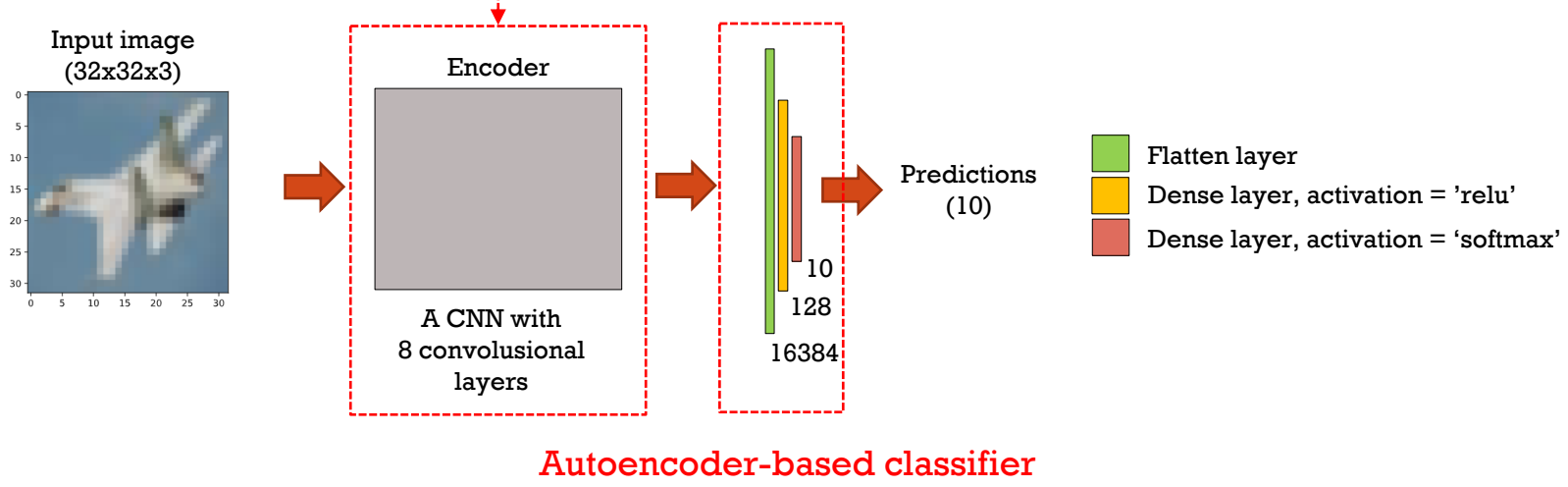


AUTOENCODER-BASED CLASSIFIER

Total # of params	1,759,233
Trainable params	1,756,417
Non-trainable params	2,816
Optimizer	RMSprop
Loss	MSE

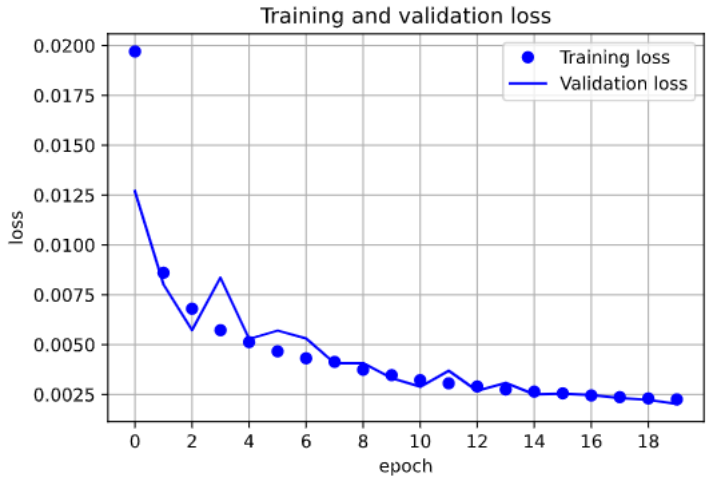


# of classes	10
Total # of params	3,274,666
Trainable params	2,098,570
Non-trainable params	1,176,096
Optimizer	Adam
Loss	Categorial crossentropy



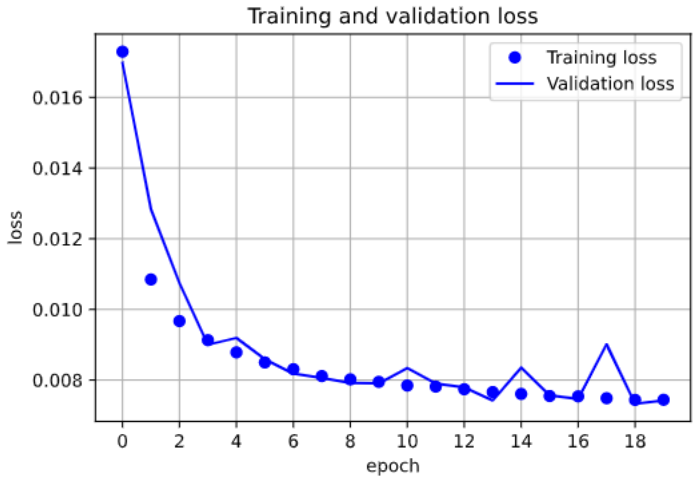
TRAINING OF THE CONVOLUTIONAL AUTOENCODER

Fashion-MNIST



Total params	1,758,657
Trainable params	1,755,841
Non-trainable params	2,816
Avg. training time for 1 epoch	374.6 sec
Total training time (# of epochs)	2.08 hrs (20 epochs)

CIFAR-10



Total params	1,759,233
Trainable params	1,756,417
Non-trainable params	2,816
Avg. training time for 1 epoch	380.9 sec
Total training time (# of epochs)	2.11 hrs (20 epochs)

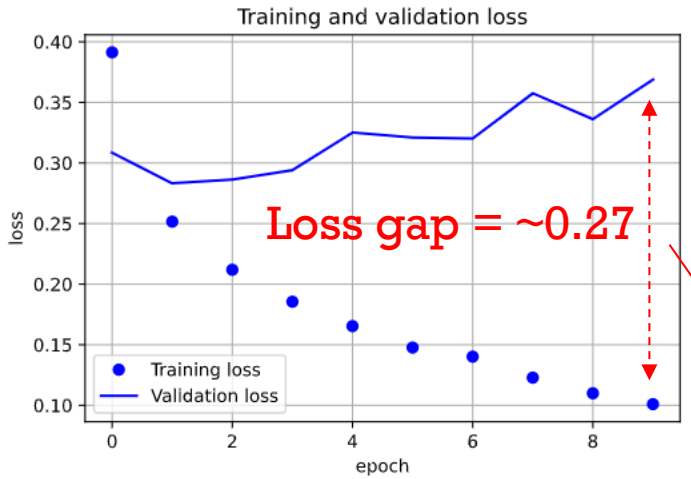
Loss = MSE

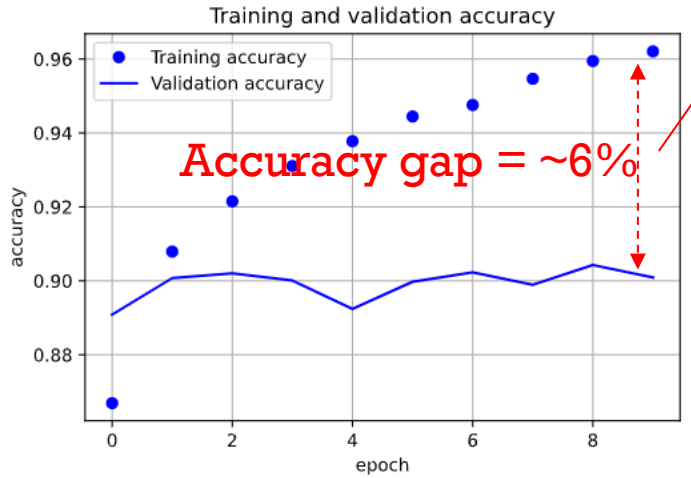
The training loss match well with the validation loss.

-> the autoencoder generalize very well on unseen images.

TRAINING OF THE AUTOENCODER-BASED CLASSIFIER / FASHION-MNIST

Without regularization

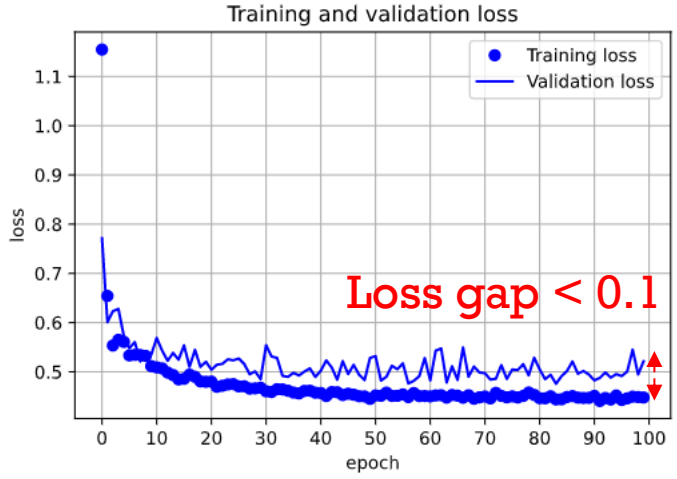


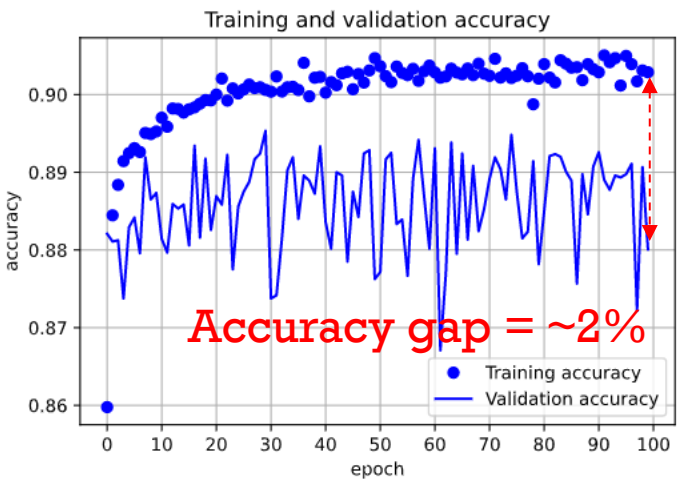


overfitting!

On the test set:
Accuracy : 90.51%
Loss : 0.3635

L1- & L2-norm regularization (on Dense layers)





On the test set:
Accuracy : 88.16%
Loss : 0.5048

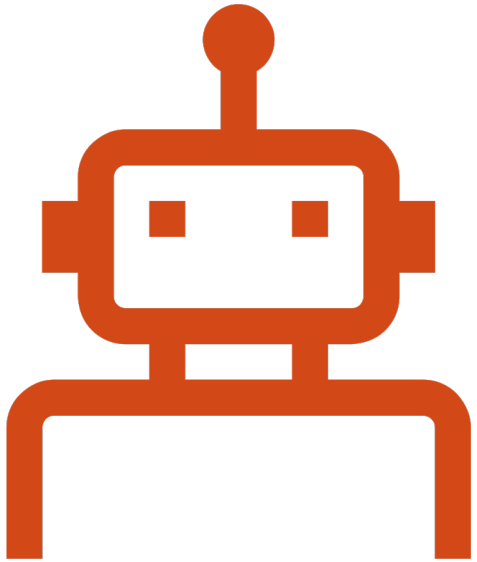
L1- & L2-norm regularization helps reduce the gap between training and validation results.

However, regularized classifiers take much longer time for training to obtain good accuracy.

Note:
lambda (L1) = 1e-4
lambda (L2) = 1e-3

CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - Autoencoder-based classifier
  - **Boltzmann machine-based classifier**
  - Ensemble learning of decision trees
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - Results on CIFAR-10 dataset
- **Conclusions**



BOLTZMANN MACHINE-BASED CLASSIFIER

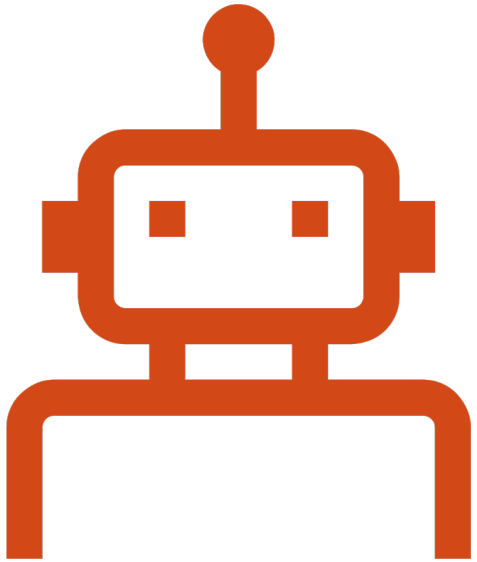
- The restricted Boltzmann machine (RBM), a two-layered network of stochastic units with undirected connections between pairs of units in the two layers, called visible and hidden nodes, has been implemented.
- Then implement MLP model to classify the data.

Params for RBM	
Visible size	784/3072
Hidden size	500
Batch size	200
# of epochs	50
Learning rate	0.01

Params for MLP	
Layer sizes	[500 20 10]
Batch size	100
# of epochs	50
Learning rate	0.05

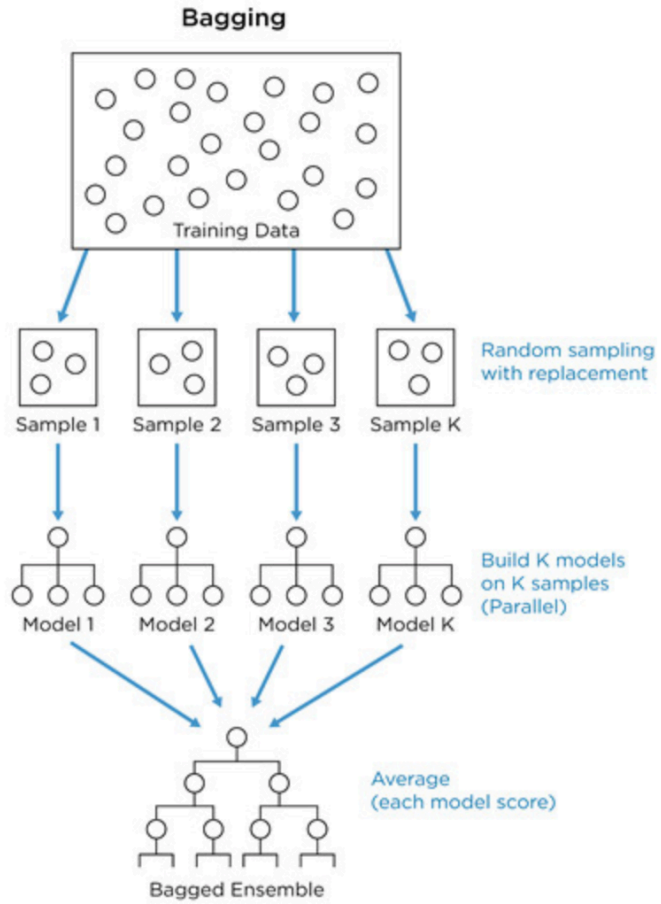
CONTENTS

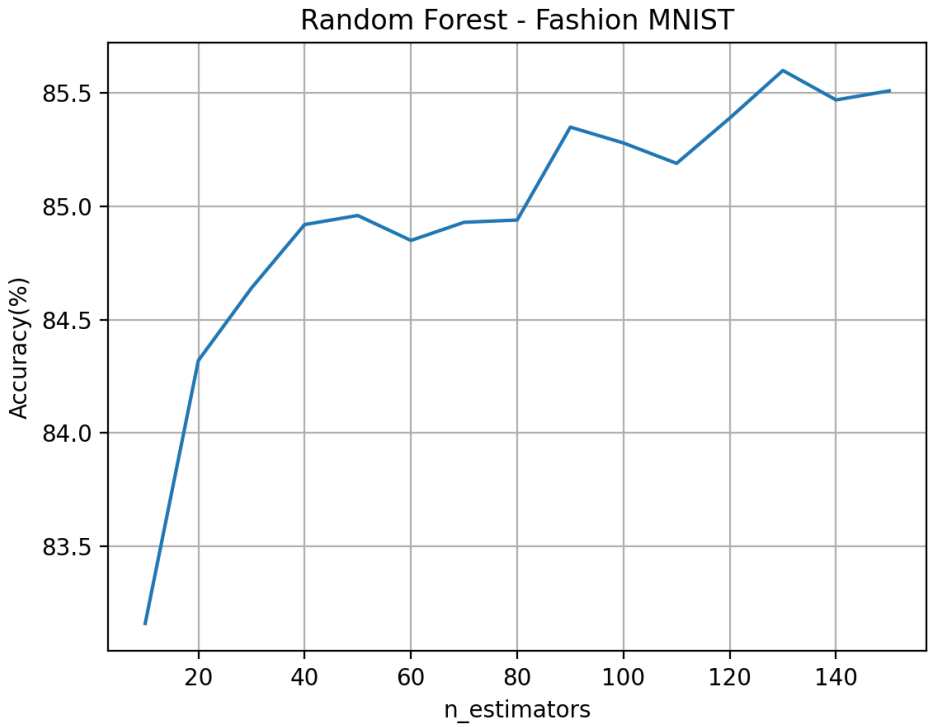
- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - Autoencoder-based classifier
  - Boltzmann machine-based classifier
  - **Ensemble learning of decision trees**
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - Results on CIFAR-10 dataset
- **Conclusions**



ENSEMBLE OF DECISION TREES (1/3)

BAGGING = **B**ootstrap **agg**regat**ing**

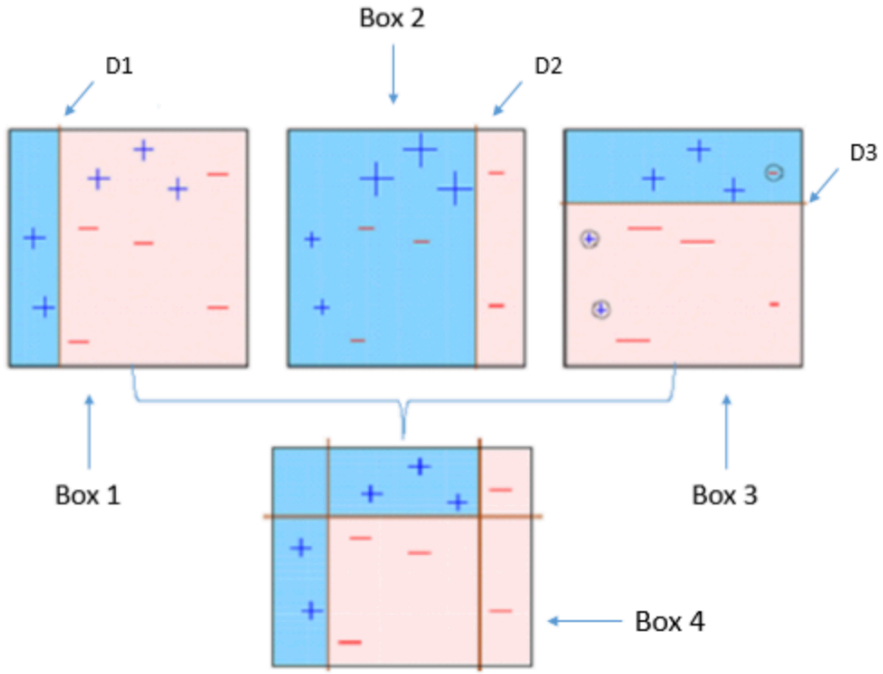


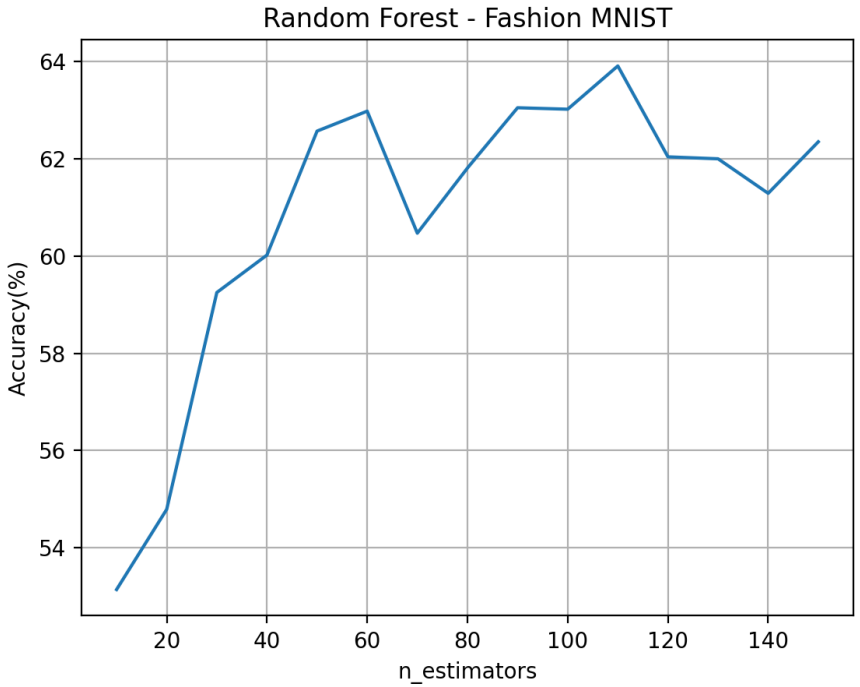


Increasing # of estimators (classifiers) improves the accuracy of the ensemble model.

ENSEMBLE OF DECISION TREES (2/3)

AdaBoost:

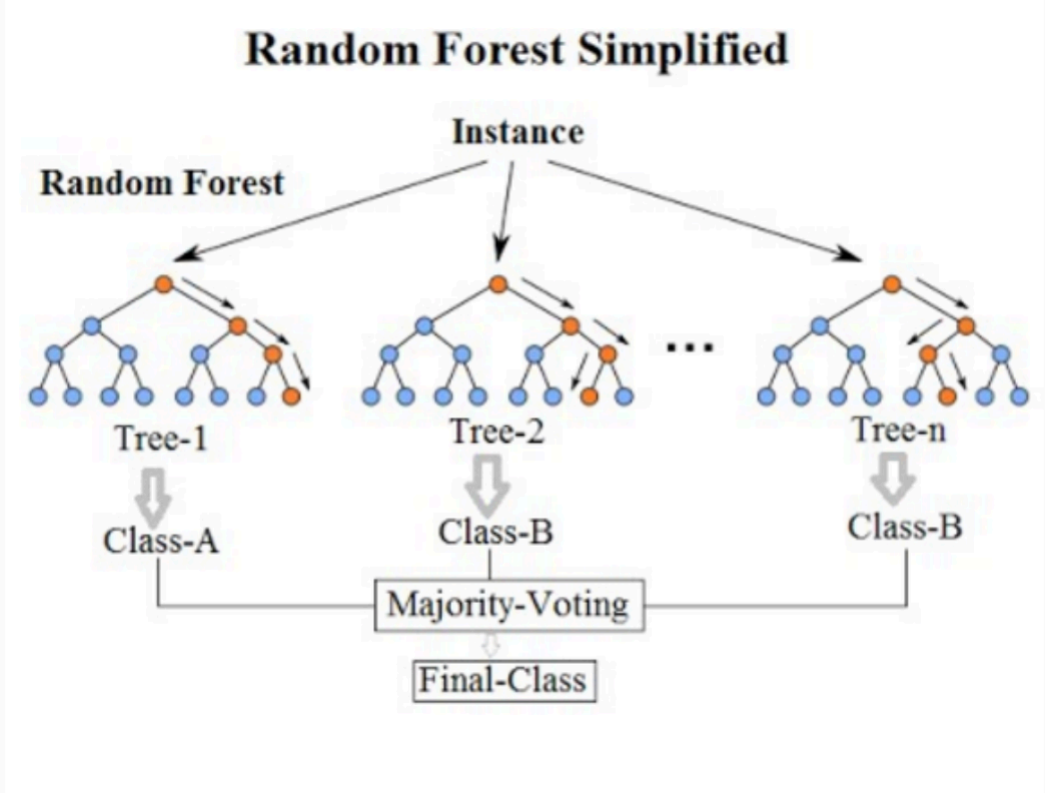


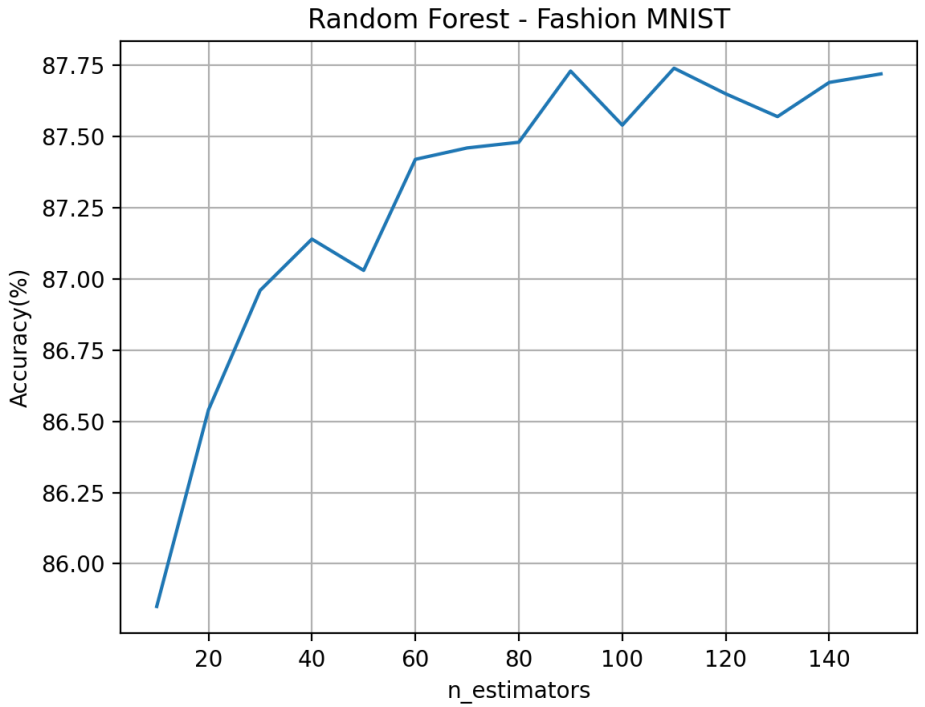


Increasing # of estimators (classifiers) improves the accuracy of the ensemble model.

ENSEMBLE OF DECISION TREES (3/3)

RANDOM FOREST

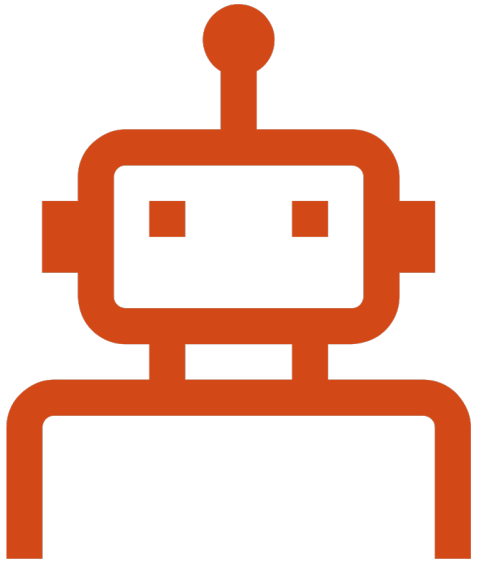




Increasing # of estimators (classifiers) improves the accuracy of the ensemble model.

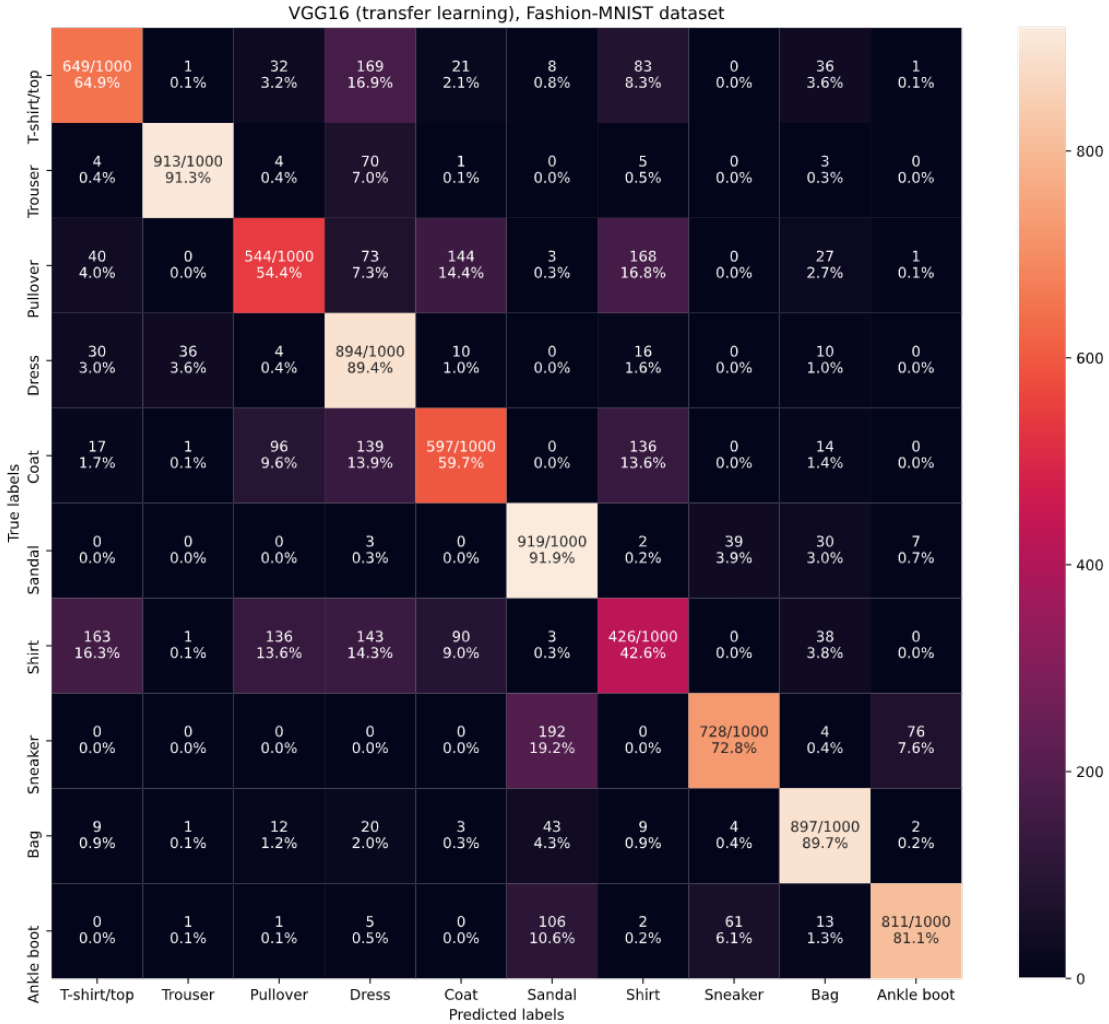
CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - Autoencoder-based classifier
  - Boltzmann machine-based classifier
  - Ensemble learning of decision trees
- **Numerical results:**
  - **Results on Fashion-MNIST dataset**
  - Results on CIFAR-10 dataset
- **Conclusions**

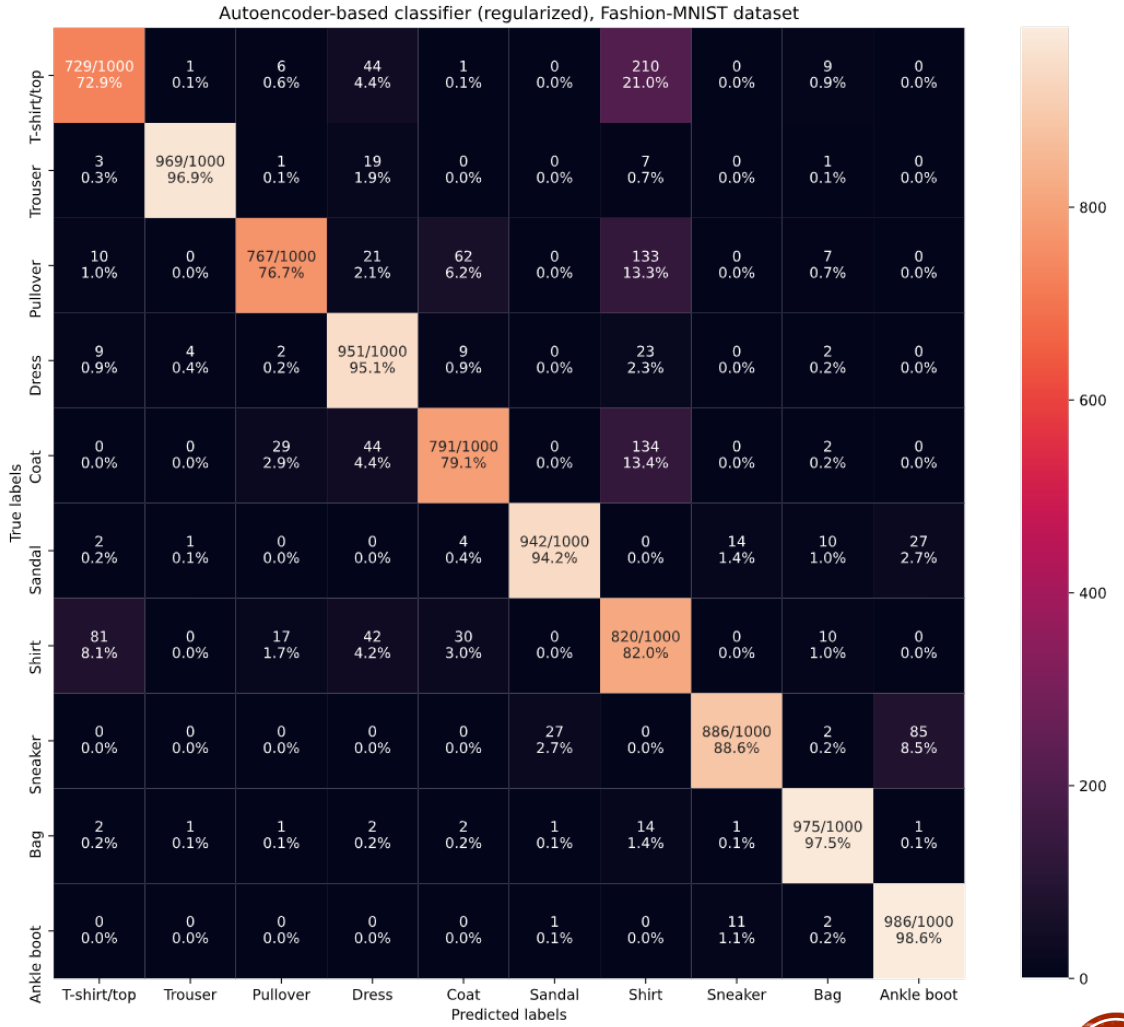


RESULTS / FASHION-MNIST DATASET (1/3)

VGG16-based classifier

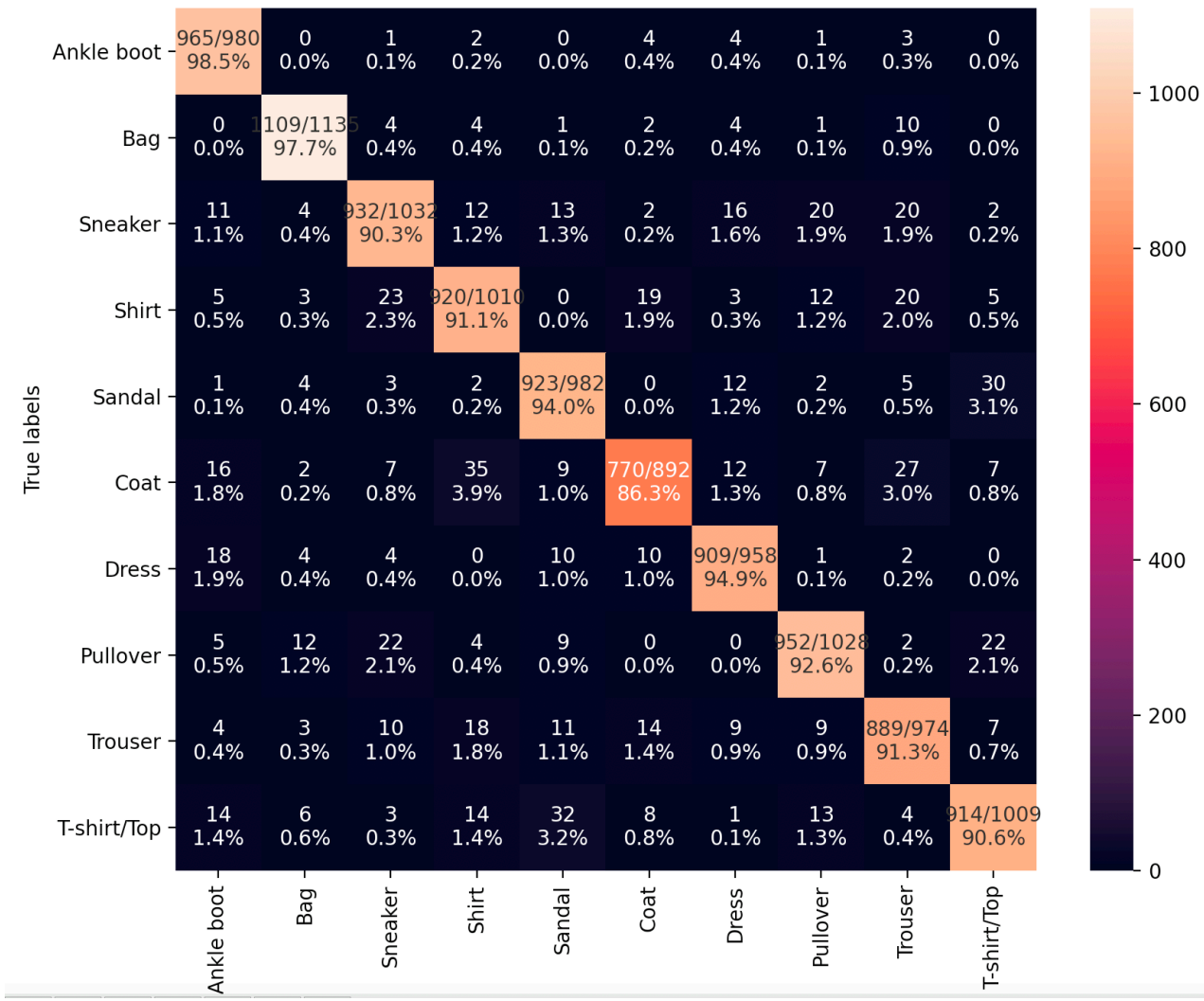


Autoencoder-based classifier

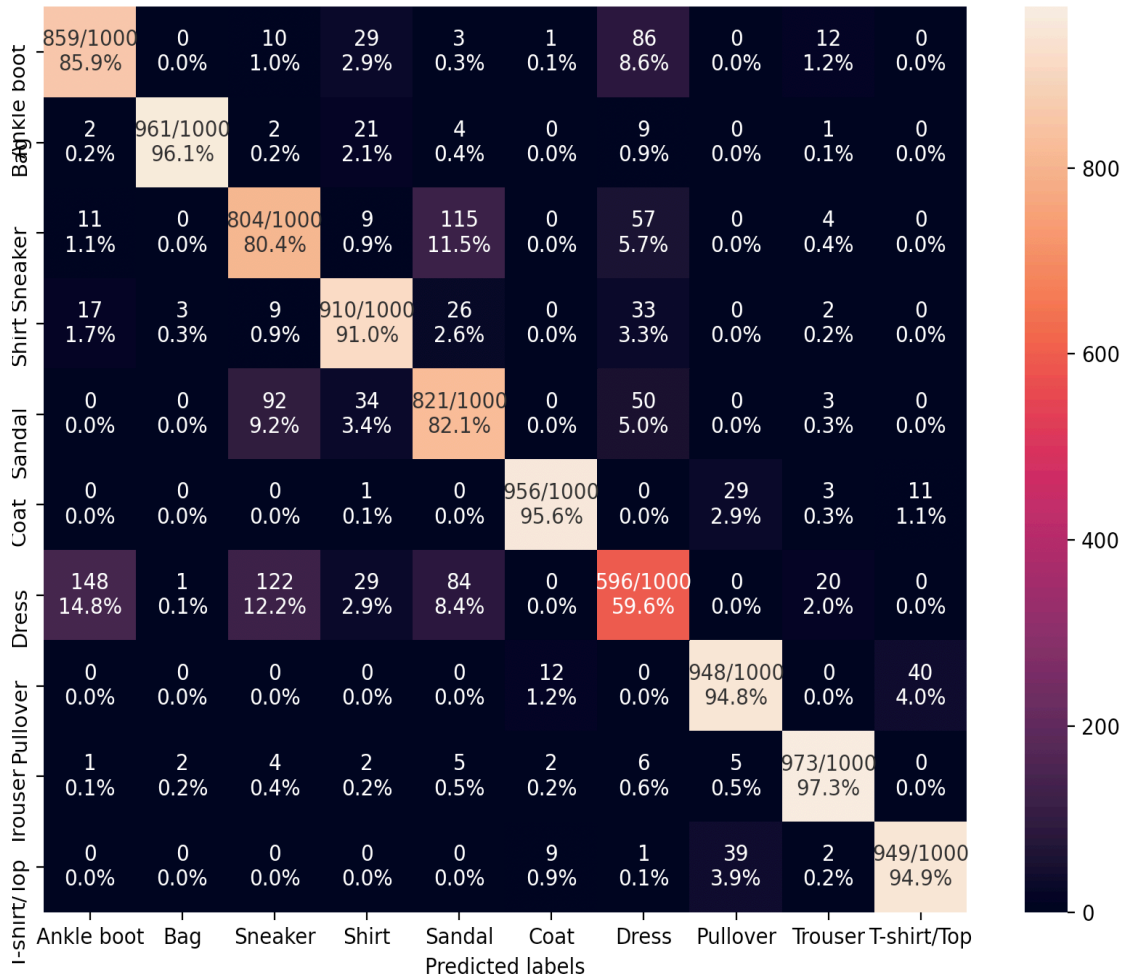


RESULTS / FASHION-MNIST DATASET (2/3)

Restricted BM-based classifier



Ensemble classifier : Random forest



RESULTS / FASHION-MNIST DATASET (3/3)

		VGG16-based classifier	Autoencoder-based classifier	RBM-based classifier	Ensemble of decision trees		
					Random forest	Bagging	Adaboost
1	Total params	15,769,930	2,782,570	--	--	--	--
2	Trainable params	1,055,242	1,607,050	--	--	--	--
3	Non-trainable params	14,714,688	1,175,520	--	--	--	--
4	Training time for 1 epoch	323.75 sec	58.25 sec	5.22sec	0.3 msec	0.3 msec	0.5 msec
5	Total training time (# of epochs)	6,475 sec (20 epochs)	5,825 sec (100 epochs)	261.49 sec (50 epoch)	0.3 msec	0.3 msec	0.5 msec
6	Loss (categorical cross-entropy)	0.7122	0.5084	0.0144	0.5046	0.5784	1.1103
7	Accuracy	73.78%	88.16%	92.63%	89.02%	85.44%	64.35%
8	Precision score	0.7443	0.8930	0.9257	0.8691	0.8531	0.6445
9	Recall score	0.7378	0.8816	0.9253	0.8676	0.8521	0.5523
10	F1 score (average='macro')	0.7343	0.8833	0.9252	0.8687	0.8526	0.5953
11	Test time	54.63 sec	10.89 sec	12.3 msec	16.55 sec	351.19 sec	322.29 sec

VGG16-based classifier is the most complex model and thus requires long time for training (even with transfer learning).

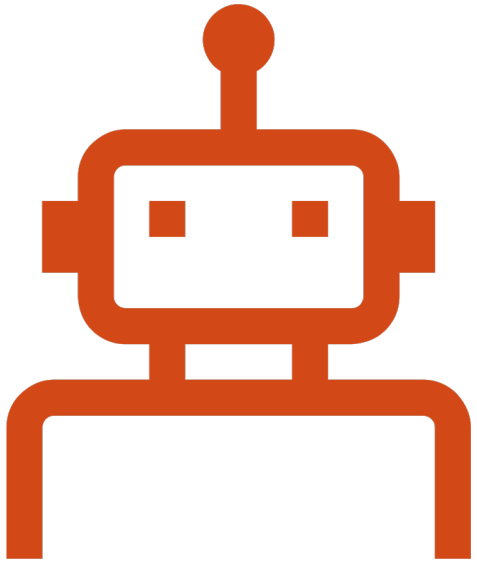
Autoencoder-based classifier: requires the least time for testing and produces quite good accuracy.

RBM-based classifier: produces the best accuracy.

Ensemble of DTs: requires the least time for training.

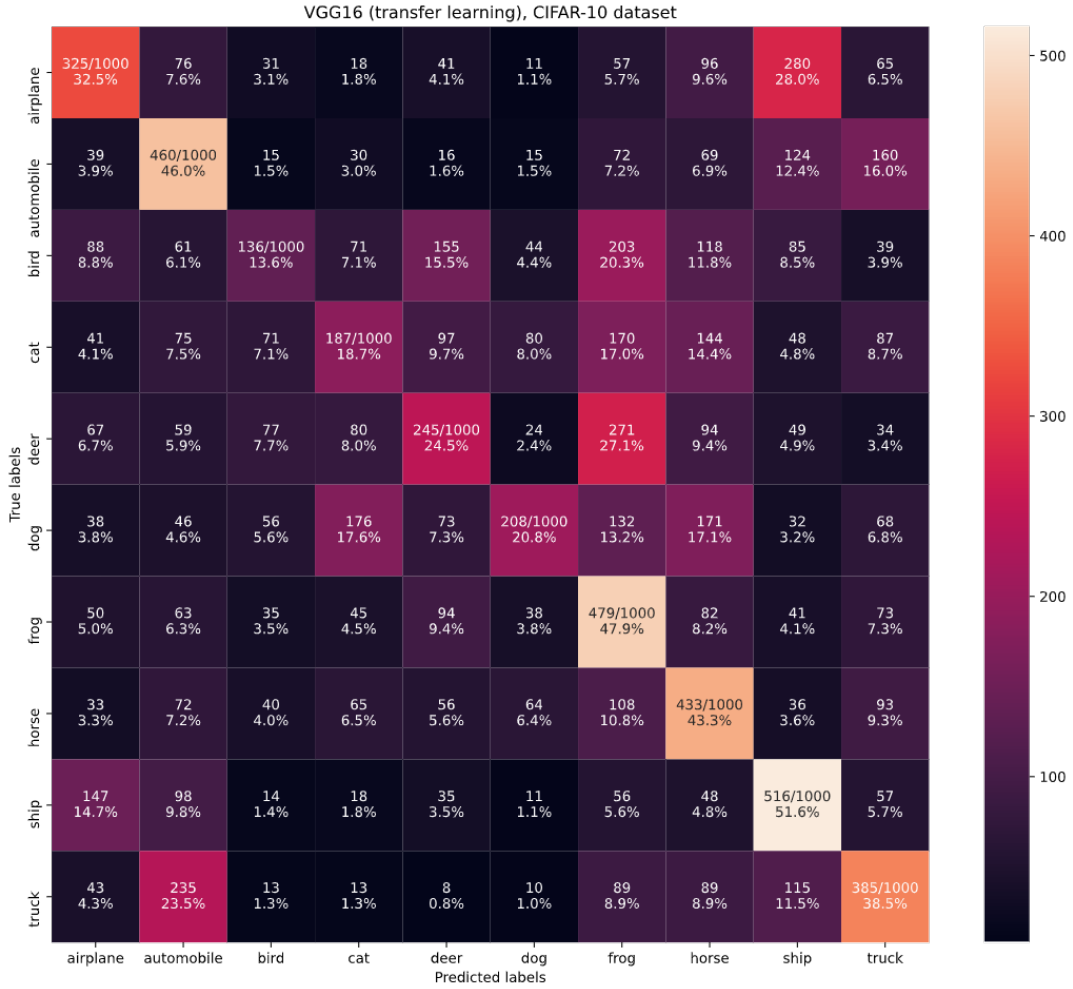
CONTENTS

- **Datasets:**
  - Fashion-MNIST
  - CIFAR-10
- **Methods:**
  - VGG16-based classifier with transfer learning
  - Autoencoder-based classifier
  - Boltzmann machine-based classifier
  - Ensemble learning of decision trees
- **Numerical results:**
  - Results on Fashion-MNIST dataset
  - **Results on CIFAR-10 dataset**
- **Conclusions**

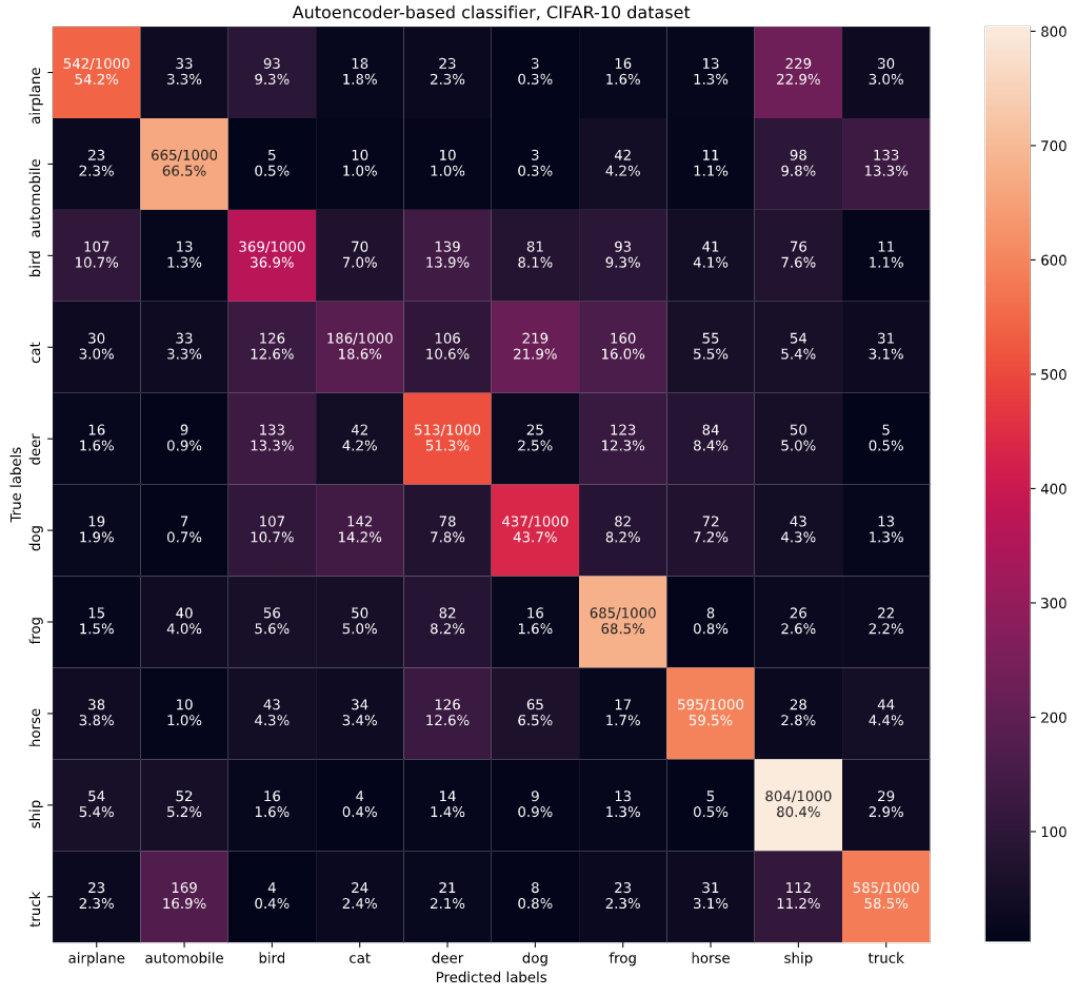


RESULTS / CIFAR-10 DATASET (1/3)

VGG16-based classifier

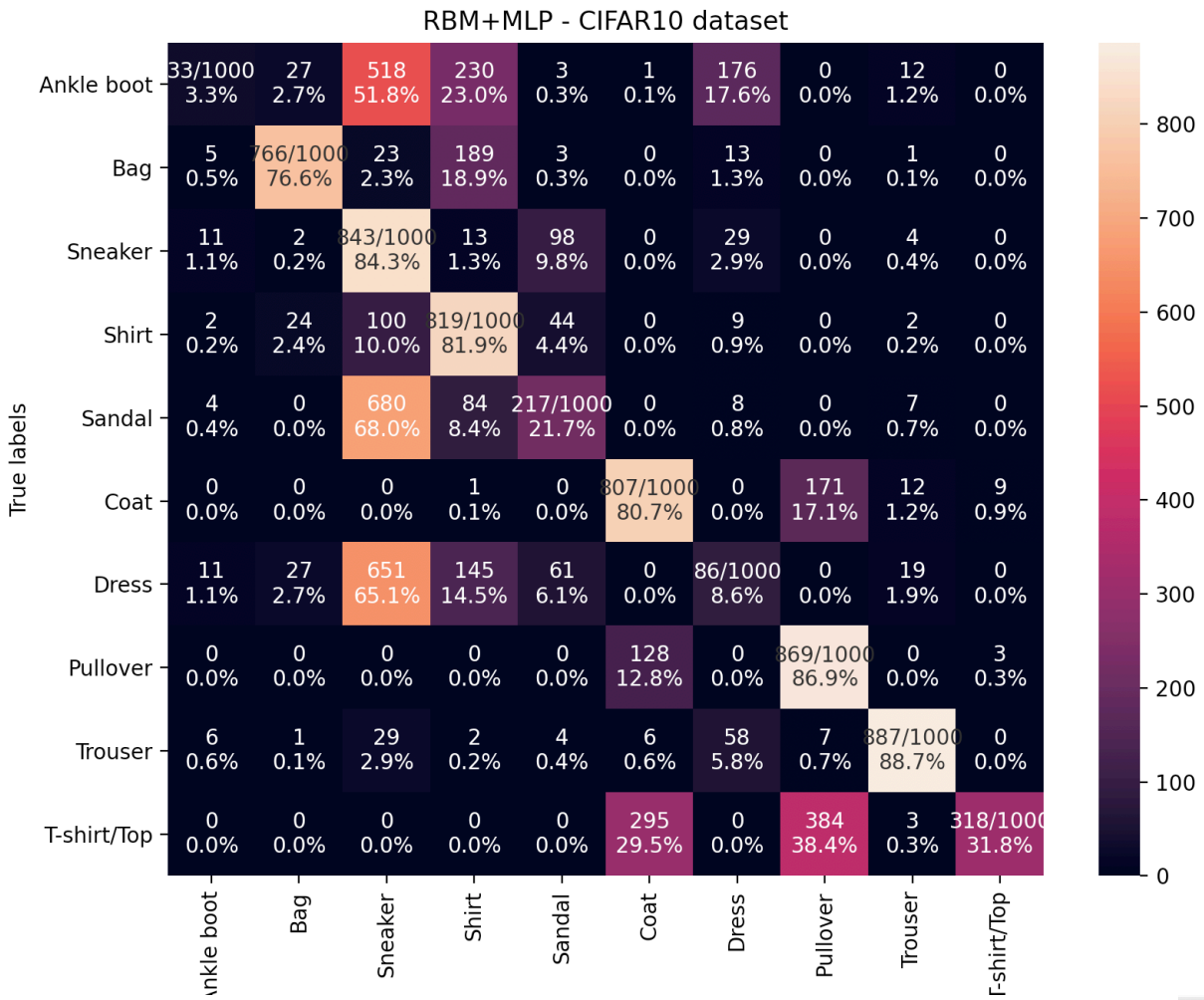


Autoencoder-based classifier

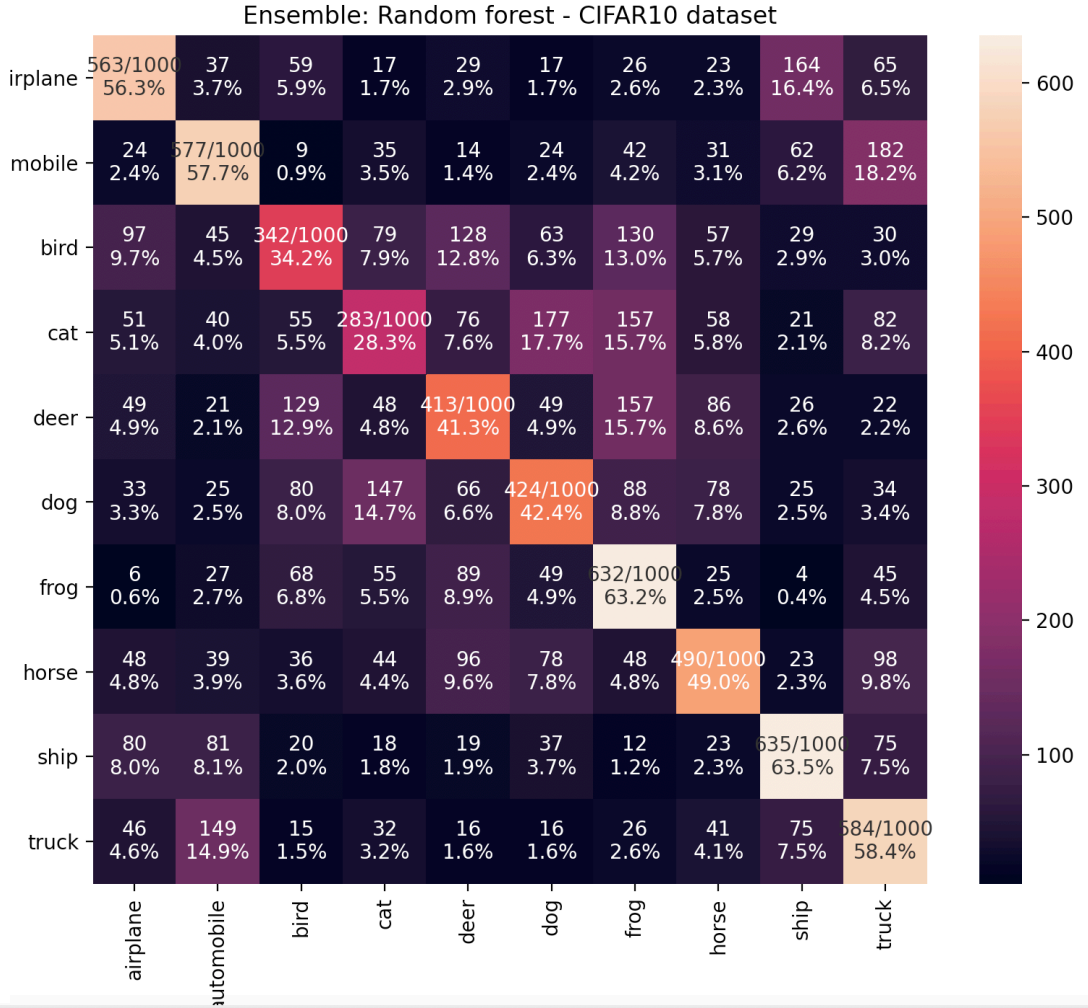


RESULTS / CIFAR-10 DATASET (2/3)

RBM-based classifier



Ensemble classifier : Random forest



RESULTS / CIFAR-10 DATASET (3/3)

		VGG16-based classifier	Autoencoder-based classifier	RBM-based classifier	Ensemble of decision trees		
					random forest	bagging	Adaboost
1	Total params	15,769,930	3,274,666	--	--	--	--
2	Trainable params	1,055,242	2,098,570	--	--	--	--
3	Non-trainable params	14,714,688 (obtained from VGG16)	1,176,096 (the encoder part)	--	--	--	--
4	Training time for 1 epoch	**268.8 sec**	61.9 sec	**225.6 sec**	**0.7 msec**	**0.8 msec**	**0.8 msec**
5	Total training time (# of epochs)	**14.93 hrs (200 epochs)**	619 sec (10 epochs)	4,512 sec (20 epoch)	0.7 msec	0.8 msec	0.8 msec
6	Loss (categorical cross-entropy)	1.8691	1.3718	1.3875	1.4865	1.4230	1.4125
7	Accuracy	**33.74%**	**53.81%**	**52.25%**	**48.01%**	**49.26%**	**49.43%**
8	Precision score	0.3365	0.5318	0.6102	0.4760	0.4888	0.4897
9	Recall score	0.3374	0.5381	0.5612	0.4759	0.4883	0.4893
10	F1 score (average='macro')	0.3255	0.5285	0.5243	0.4763	0.4880	0.4891
11	Test time	54.76 sec	22.89 sec	**12.16sec**	**374.33 sec**	**1,754 sec**	**2,517 sec**

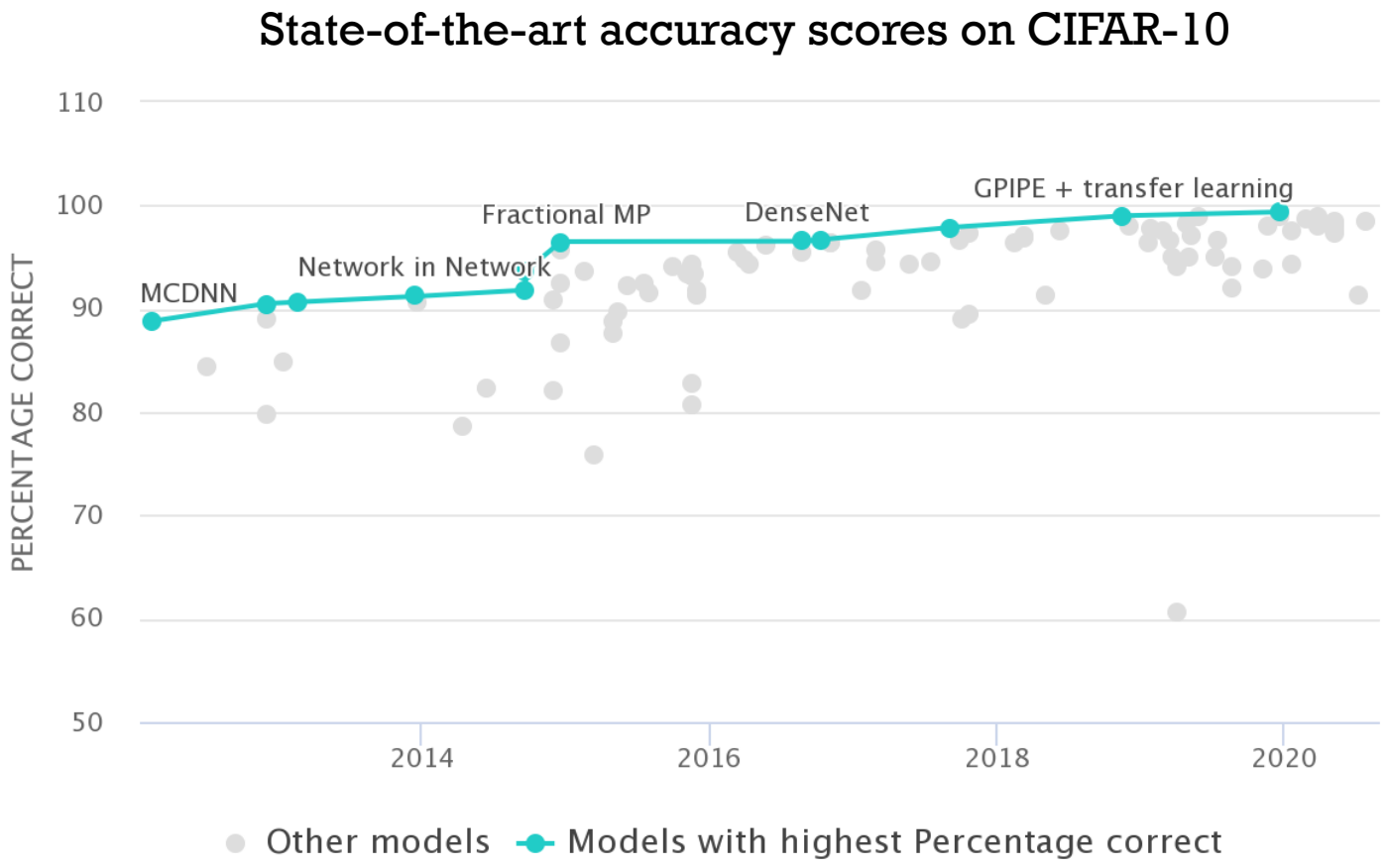
In general, all methods produce not good accuracy on CIFAR-10 dataset.

VGG16-based classifier requires long time for training.

Autoencoder-based classifier: produces the best accuracy (53.81%).

RBM-based classifier: requires the least time for testing.

Ensemble of DTs: requires the least time for training but very long time for testing.



Source: https://paperswithcode.com/sota/image-classification-on-cifar-10

COMMENTS / CIFAR-10 DATASET

- There are many models that achieve very good accuracy on CIFAR-10 dataset.
- However, all four models of interest are not suitable for CIFAR-10 (the accuracy is not good, ~50%)
- To obtain better accuracy, model tuning and longer training should be conducted.

CONCLUSIONS

- VGG16-based model produces quite good accuracy on Fashion-MNIST dataset (acc=73.78%) but is not suitable for CIFAR-10 dataset (acc=33.74%).
  - Possible reasons: CIFAR-10 has images of too low resolution (32x32) with various types of background (in comparison: Fashion-MNIST's images has only black background)
    -> it's harder to filter out the background while preserving the main object of a picture
  - Transfer learning save a lot of time for training.
- Autoencoder-based classifier produces better results and takes less time than VGG16-based model on both datasets.
  - To mitigate the overfitting problem, L1- and L2-norm regularization is very effective but requires more time for training.
- Restricted Boltzmann machine-based classifier:
  - Produces very good accuracy (92.63%) on Fashion-MNIST dataset.
  - Requires small time (in comparison with 3 other counterparts) to make predictions.
- Ensemble classifer:
  - Increasing the # of estimators (classifiers) improves the accuracy of the ensemble model.
  - Requires longer time to make predictions in comparison with 3 counterparts

REFERENCES

1. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Apr. 2015, Accessed: Aug. 26, 2020. [Online]. Available: http://arxiv.org/abs/1409.1556.
2. K. Team, "Keras documentation: Transfer learning & fine-tuning." https://keras.io/guides/transfer_learning/ (accessed Aug. 22, 2020).
3. "Autoencoder as a Classifier," *DataCamp Community*, Jul. 20, 2018. https://www.datacamp.com/community/tutorials/autoencoder-classifier-python (accessed Aug. 23, 2020).
4. A. Nagpal, "L1 and L2 Regularization Methods," *Medium*, Oct. 14, 2017. https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c (accessed Aug. 28, 2020).
5. M. Tanaka and M. Okutomi, "A Novel Inference of a Restricted Boltzmann Machine," in *2014 22nd International Conference on Pattern Recognition*, Stockholm, Sweden, Aug. 2014, pp. 1526–1531, doi: 10.1109/ICPR.2014.271.
6. "Restricted Boltzmann Machines (RBM)" http://deeplearning.net/tutorial/rbm.html.
7. "Restricetd Bolzmann machine", Learning Machines course Taught by Patrick Hebron at NYU ITP.
8. Lujing Chen, "Basic Ensemble Learning (Random Forest, AdaBoost, Gradient Boosting)- Step by Step Explained", toward data science, Jan. 3, 2019.

ROLES OF MEMBERS

1. **Hoang Tuan Linh (m5232108)** implements:
   - CNN + Transfer learning from VGG16 model
   - Autoencoder-based classifier with & without regularization
   - On datasets: Fashion-MNIST + CIFAR-10
   - Platform: Python3 / Tensorflow
   - Hardware: Intel Core i5 3.4GHz / 8GB RAM
2. **Tran Thi Thoa (s1242006)** implements:
   - Boltzmann machine-based classifier
   - Ensemble of decision trees (Bagging, Adaboost, Random forest)
   - On datasets: Fashion-MNIST + CIFAR-10
   - Platform: Python3 / sklearn
   - Hardware: Intel Core i5 3.4GHz / 8GB RAM

- Thank you for your attention
- Q&A