



Dynamic Routing between Capsules - Hinton et al. (2017)

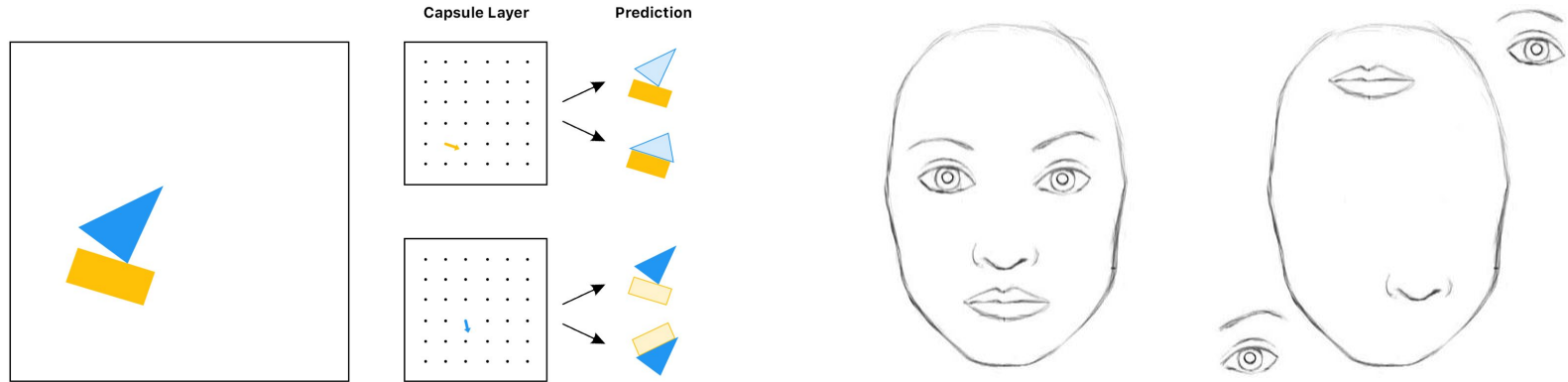
Maxime Allard - Selim Amrouni - Thibault Duplay - Phillipe Mizrahi



Planning

- 1) What are Capsules and why should we use them ?
- 2) Capsnet on MNIST
- 3) Capsnet applied to financial time-series classification

What are Capsules and Why Should We Use them?

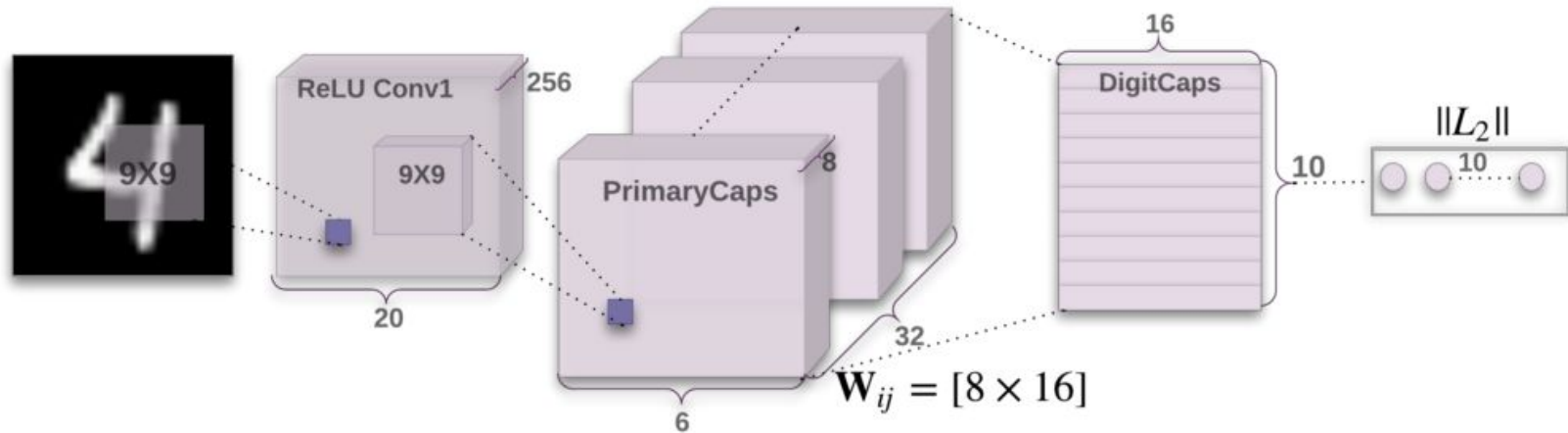


Source:

<https://medium.freecodecamp.org/understanding-capsule-networks-ais-alluring-new-architecture-bdb228173ddc>

<https://medium.com/ai%C2%B3-theory-practice-business/understanding-hintons-capsule-networks-part-i-intuition-b4b559d1159b>

CapsNet Structure



Source:

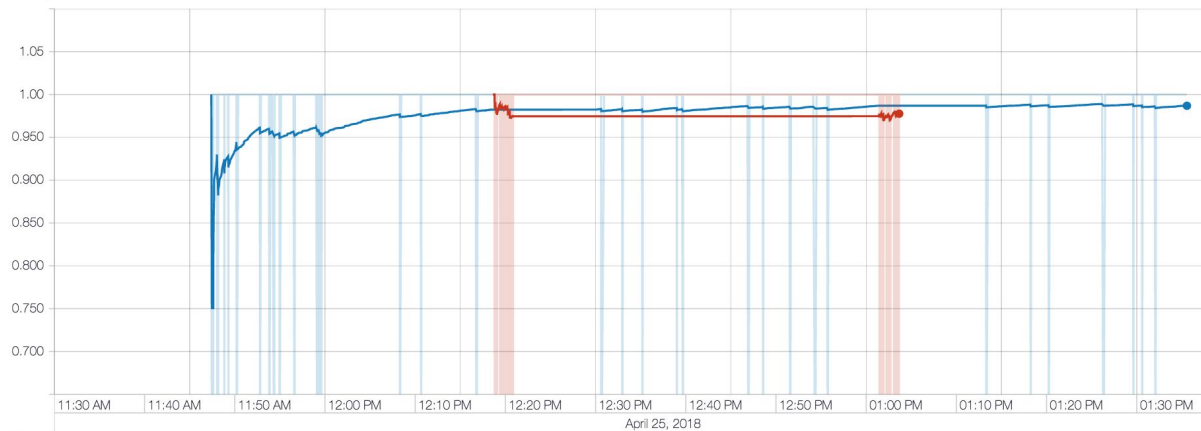
<https://medium.com/ai/C2%B3-theory-practice-business/understanding-hintons-capsule-networks-part-i-intuition-b4b559d1159b>



Reproduction of the Results From the Paper

MNIST - Accuracy

accuracy_1



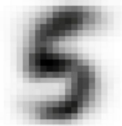
Training and
testing accuracy

MNIST - Reconstruction From the Decoder

Input Label :5



Output Label :[3]



Input Label :3



Output Label :[2]



Input Label :9



Output Label :[0]



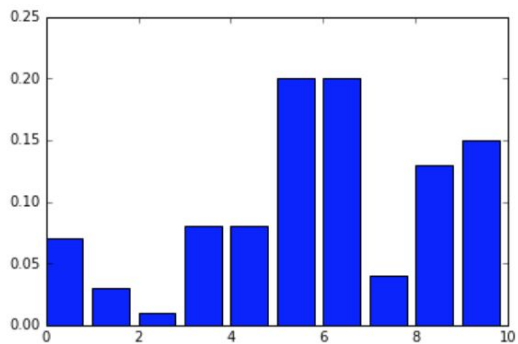
Input Label :3



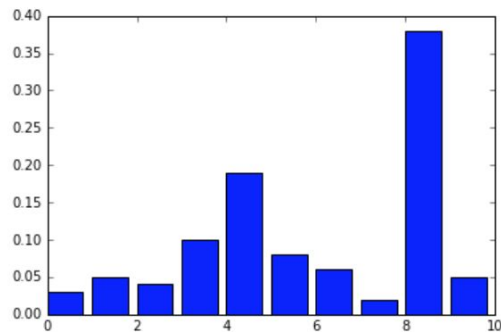
Output Label :[3]



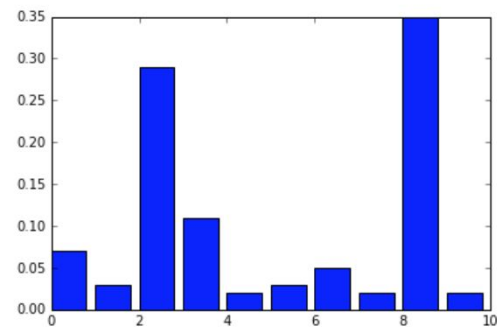
MNIST - Overlapping of Digits



5 & 0

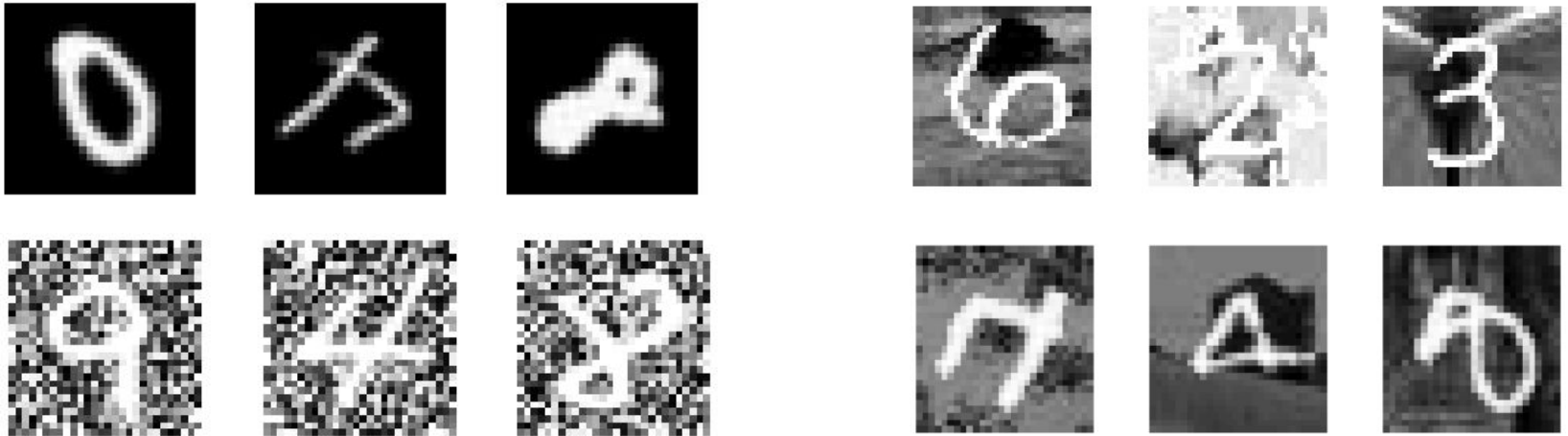


7 & 8



2 & 8

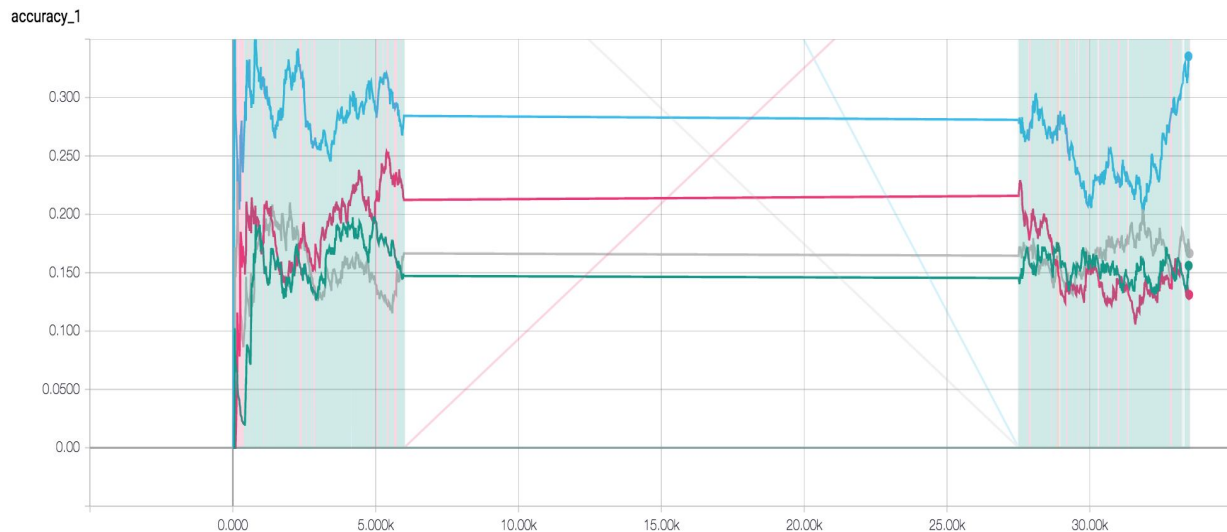
MNIST Variations - LISA Lab - Accuracies



Source:

<http://www.iro.umontreal.ca/~lisa/twiki/bin/view.cgi/Public/MnistVariations>

MNIST Variations - LISA Lab - Accuracies



- summaries/test_rotated
- summaries/test_back_rand
- summaries/test_back_img
- summaries/test_rot_back_img



Using Capsules on Financial Time-Series



Problem Statement - Architecture

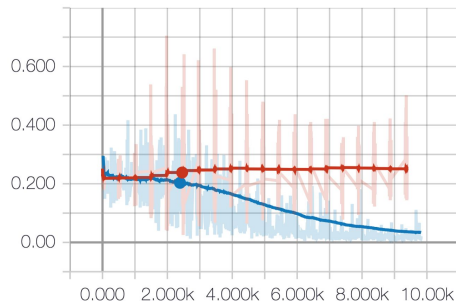
Pipe:

- Transformed Financial Time-Series of the S&P 500 into frames of 30 days (2013 - 2018)
- 5 normalized features: Open, Close, High, Low, Volume
- => Gives us a matrix of shape (5,30)
- Predict sign of the return at day+1 (2 labels)
- Trained for each stock (1 Network per stock)
- Only 32 filters (vs. 256 for MNIST) to reduce overfitting

Overfitting and Dropout (AAPL: 20 Epochs)

Training without dropout

loss_1

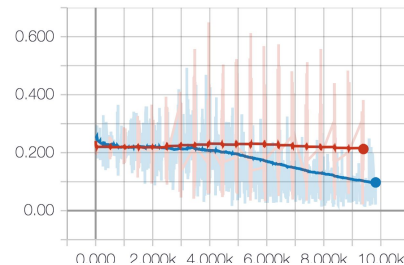


ticker: AAPL max_train_accuracy: 99.9 %
ticker: AAPL max_val_accuracy: 52.42 %
ticker: AAPL Mean Train Accuracy over 20 epochs: 85.16 %
ticker: AAPL Mean Train Loss over 20 epochs: 0.15148963
ticker: AAPL Mean Validation Accuracy over 20 epochs: 49.27 %
ticker: AAPL Mean Validation Loss over 20 epochs: 0.24998944

Dropout Conv1: 0.75
Same Dataset

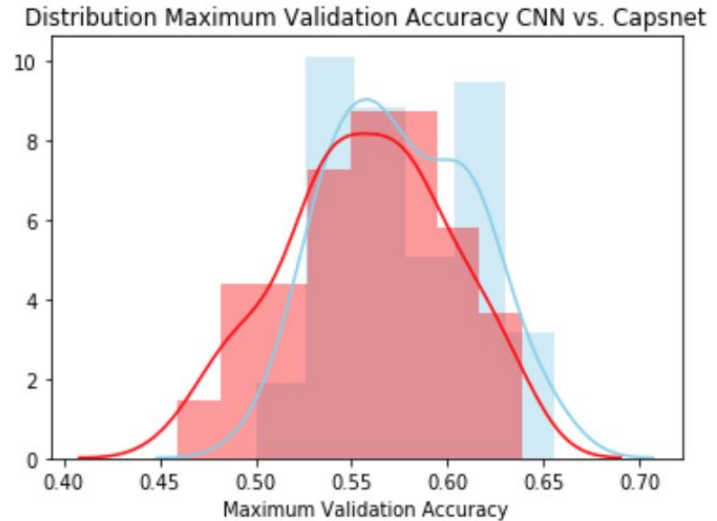
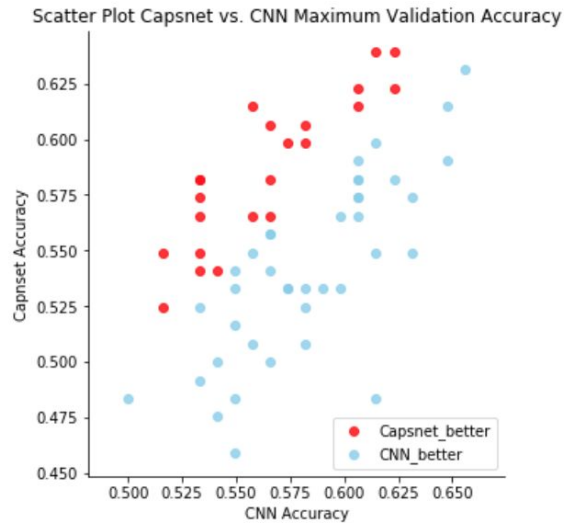
Training with dropout

loss_1



ticker: AAPL max_train_accuracy: 96.64 %
ticker: AAPL max_val_accuracy: 66.13 %
ticker: AAPL Mean Train Accuracy over 20 epochs: 74.91 %
ticker: AAPL Mean Train Loss over 20 epochs: 0.19669774
ticker: AAPL Mean Validation Accuracy over 20 epochs: 56.98 %
ticker: AAPL Mean Validation Loss over 20 epochs: 0.22132969

Performance Comparison CNN vs. Capsnet





Work in Progress - Multiple Regression

- Block Matrix as input
- 5 blocks corresponding to the 5 features
- Each block is of shape (nb of tickers, 30)
- Change of Architecture