MIAMI HERBERT BUSINESS SCHOOL


# SENet Application in Portfolio Construction

## Big Data and Factor Modeling

Ivy Yang

May 11, 2022

# Overview

1. Topic and Contribution
2. Algorithm
3. Evaluation
4. Plan

# Topic and Contribution

# Main topic

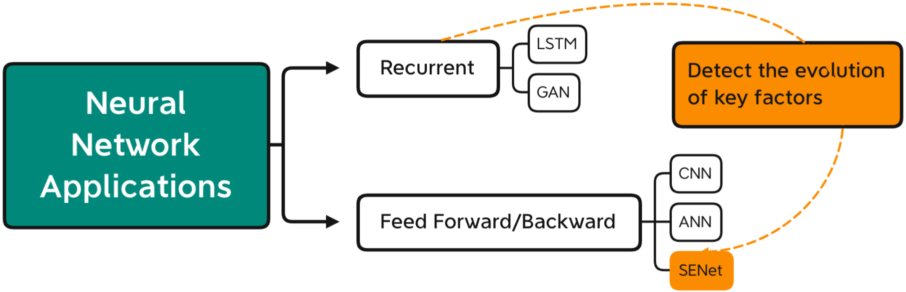
**The aim of this paper is to find the stochastic discount factor (SDF) with a new machine learning tool - Squeeze and Excitation Network (SENet).**

**Why it is important?**

- Neural network shows only the result, not the reason.
- Tree algorithms cannot generate a result as good as neural networks
- GAN and LSTM are using recurrent neural network but not showing the channels. We may want to update the recurrent method that shows the channel of factors.

# Literature

**Recall the neural network algorithm that has been applied to portfolio management:**



# Literature

**Literature in Asset Pricing Using Machine Learning Methods:**

- Chen, Pelger and Zhu (2019)
  This paper utilized the recurrent Long-Short-Term-Memory network to capture a small set of hidden economic state processes. They also use a feedforward network to capture the non-linear effect. Finally, they utilized the Generative-Adversarial network to enforce a no-arbitrage condition to their portfolio.
- Guijarro-Ordonez, Pelger and Zanotti (2021)
  This paper utilized the convolutional transformer technique. Their factors include the conditional latent asset pricing factors.

# Literature

**Literature in Asset Pricing Using Machine Learning Methods:**

- Chernozhukov et.al (2018)[1]
  This paper tried to solve the high-dimension problem. It uses the general double machine learning (DML) framework to mitigate bias and restore valid inference on a low-dimensional parameter of interest in the presence of high-dimensional parameters.
- Feng, Polson and Xu (2021) - working paper
  This paper summarized the current machine learning methods in factor selecting. They

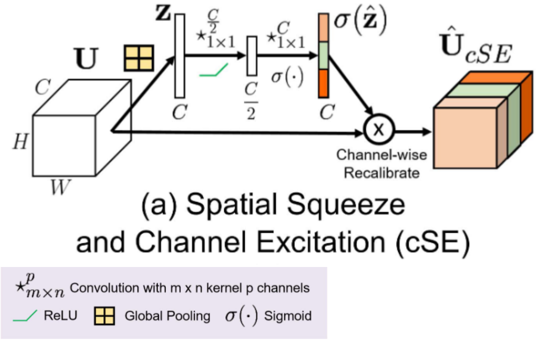
# Literature

**Literature in Artificial In:**

- Hu et.al (2018) “Competitive Inner-Imaging Squeeze and Excitation for Residual Network”
  This paper applied the SENet in image recognition.
- Li et.al (2018) “ Recurrent Squeeze-and-Excitation Context Aggregation Net for Single Image Deraining”
  This paper proposed the recurrent learning process of SENet.

# Algorithm

# SENet - cSE (Spatial Squeeze and Channel Excitation)




(a) Spatial Squeeze and Channel Excitation (cSE)

- **Initial transformation:**
  $\mathbf{U} = [\mathbf{u_1}, \mathbf{u_2}, \cdots, \mathbf{u_C}]$ with $\mathbf{u}_i \in \mathbb{R}^{H\times W}$
- **Squeeze:**

$$z_k = \mathbf{F}_{sq}(\mathbf{U}_c) = \frac{1}{F_1 \times F_2} \sum_{i=1}^{F_1} \sum_{j=1}^{F_2} \mathbf{u}_k(i,j)$$

  This vector is transformed to
  $\hat{\mathbf{z}} = \mathbf{W}_1 (\delta (\mathbf{W}_2 \mathbf{z}))$, with
  $\mathbf{W}_1 \in \mathbb{R}^{C\times \frac{C}{r}}, \mathbf{W}_2 \in \mathbb{R}^{\frac{C}{r}\times C}$.
- **Excitation:**
  $\hat{\mathbf{U}}_{cSE} =$
  $[\sigma(\hat{z}_1)\mathbf{u}_1, \sigma(\hat{z}_2)\mathbf{u}_2, \cdots, \sigma(\hat{z}_C)\mathbf{u}_C]$

# SENet - Understand the "Weight"

- Weight in model. - filter: ReLu and Tanh.
  `https://cs231n.github.io/convolutional-networks/`
- We care about: Tanh
  Is the Tanh weight comparable if they come from different dimensions? - Seems yes, because they are proportions.

**The target for our research is to obtain the Tanh whlie maintain a relatively high return of the portfolio (should be comparable to Chen et.al (2019)**

# SENet - Main difficulties

**Squeeze by what:**

- FF-3 factors (monthly data)
- FF-5 factors (monthly data)
- Factors used in Chen et.al (2019) - 49 factors in 5 categories.
- All factors in Open Source Asset Pricing

Since we need to average each channel, so the channels have to be factors. Otherwise it makes no sense to squeeze. Hence, we will have more channels than any of the existing literature.

# SENet - Main difficulties

**Is there any mathematical issues?**

- Since we utilize the sigmoid or Tanh in the SE algorithm, with the elevation of dimension, do we expect a much severer vanishing gradient issue?

  We can only have one middle layer to avoid this problem.

- When the inputs grow extremely small or extremely large, the sigmoid function saturates at 0 and 1 while the tanh function saturates at -1 and 1.

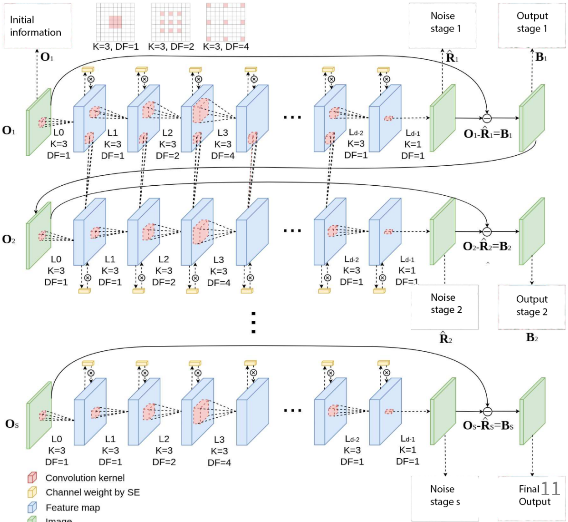
  To avoid this issue, I will first try to eliminate stocks with extreme factor values - using SP100 as a base, delete stocks with missing values (85 stocks)

# SENet - Recurrent

We have two ways to go:

1. Utilize the Recurrent SENet.

2. ⋆ Combine the SENet with LSTM.
   (The LSTM has some issues in practical, need more research)



# Evaluation

# Way to evaluate

**Until now, the performance of GAN PCA IPCA RNN are compared in Chen et.al (working paper)[2]**

- Sharpe ratio
- Max 1 month loss & max draw-down
- Performance with respect to the market, 3-factor FF model and 5-factor FF model

# Plan

# Future work

**Current Work**

- Replicate Chen et. al (2019)
  Learn how to organize factors into data layer. Understand the LSTM algorithm and its weakness. Provide a theoretical base for SENet research.

**Next Step**

- Find a better way to solve the vanishing gradient problem for our activation function "tanh".
- After get the result for FF-5 factors, I will compare it to the result from FF-3 factors. Then I will investigate if higher dimension is applicable to this algorithm.
- If the learning rate and result of higher dimension seems unreal, I will tackle with high dimension problems.

# Code Source

Squeeze and Excitation Network:
`https://amaarora.github.io/2020/07/24/SeNet.html`

`https://blog.paperspace.com/channel-attention-squeeze-and-excitation-networks/amp/`

Recurrent Squeeze and Excitation Network:
`https://xialipku.github.io/RESCAN/`

Details:

- The kernel setting: `https://towardsdatascience.com/understanding-2d-dilated-convolution-operation-with-examples-in-numpy-an`

# The End

# References I

V. Chernozhukov.
Double/debiased machine learning for treatment and structural parameters.
*The Econometics Journal*, 2018.

M. P. Luyang Chen and J. Zhu.
Deep learning in asset pricing.
*SSRN ELECTRONIC JOURNA*, 2019.