



FIND YOUR PASSION AT NVIDIA.

We're **hiring interns** and **new college grads**. Come join the industry leader in virtual reality, artificial intelligence, self-driving cars, and gaming.

We're hiring in hardware, architecture, software, research, deep learning, and more!

Learn more at: www.nvidia.com/university





DEEP
LEARNING
INSTITUTE

Algorithmic Trading using Statistical Arbitrage with Deep Autoencoders

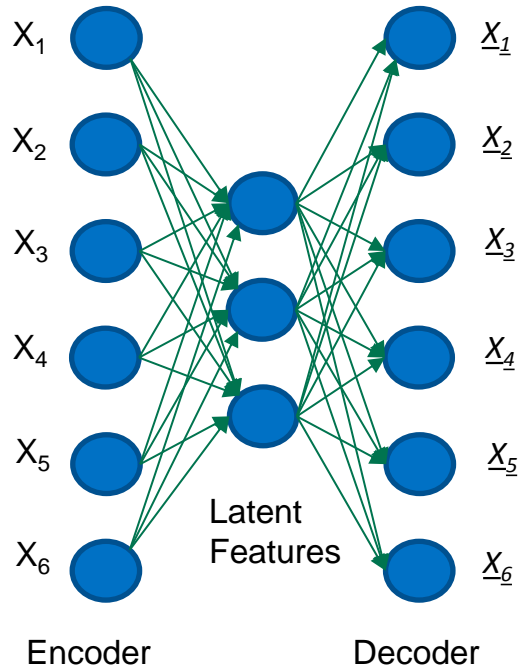
Onur Yilmaz, Ph.D.

Deep Learning Solution Architect, NVIDIA Corporation
Certified Instructor, NVIDIA Deep Learning Institute

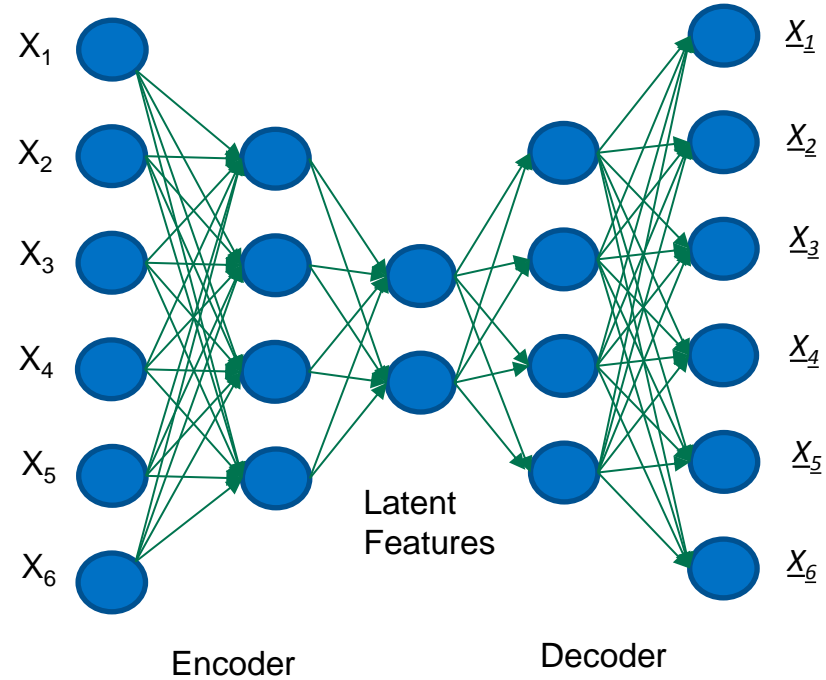
AGENDA

- Autoencoders
- Anomaly Detection
- Algorithmic Trading using Statistical Arbitrage with Deep Autoencoders

Autoencoder



(a)



(b)

Autoencoder

- Which of the following number sequences do you find the easiest to memorize?
- 40, 27, 25, 36, 81, 57, 10, 73, 19, 69
- 50, 25, 76, 38, 19, 58, 29, 88, 44, 22, 11, 34, 17, 52, 26, 13, 40, 20

Autoencoder

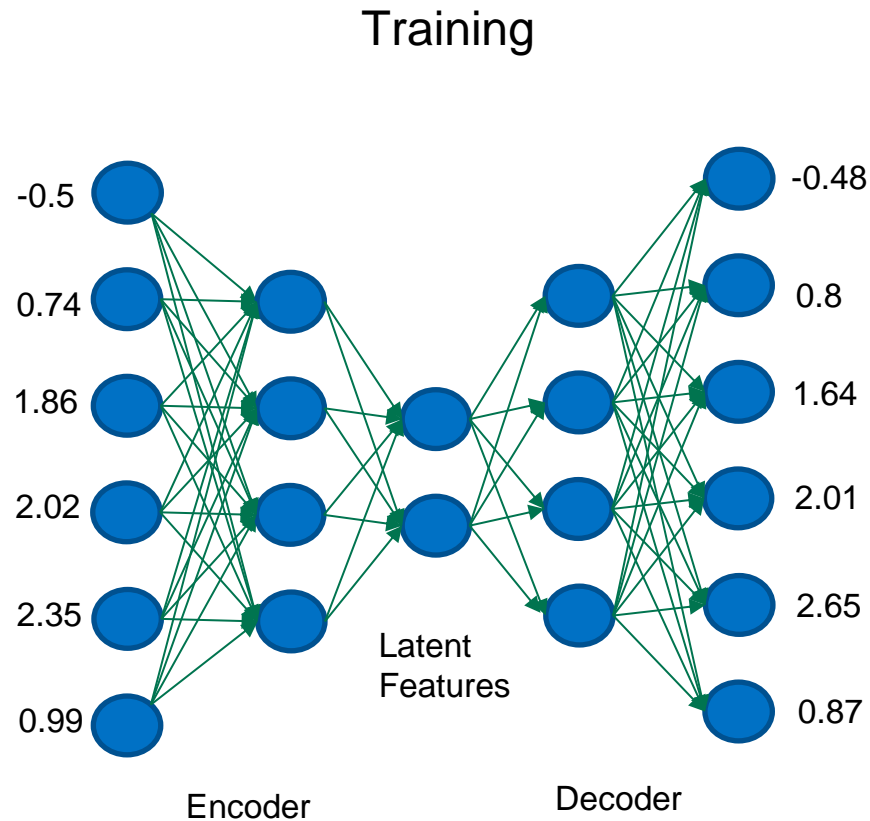
- Even though the second sequence is longer, it is easier to memorize the second one
- It follows a simple rule: even numbers are followed by their half, and odd numbers are followed by triple plus one (*hailstone sequence*)
- Ref: Aurelien Geron, Hands-On Machine Learning with Scikit-Learn and Tensorflow

Autoencoder - Use Cases

- Image clustering for labeling
- Image generation
- Anomaly Detection

Deep Autoencoder - Anomaly Detection

Time	Signal
1	-0.5
2	0.74
3	1.86
4	2.02
5	2.35
6	0.99
7	4.26
...	
200	-0.5

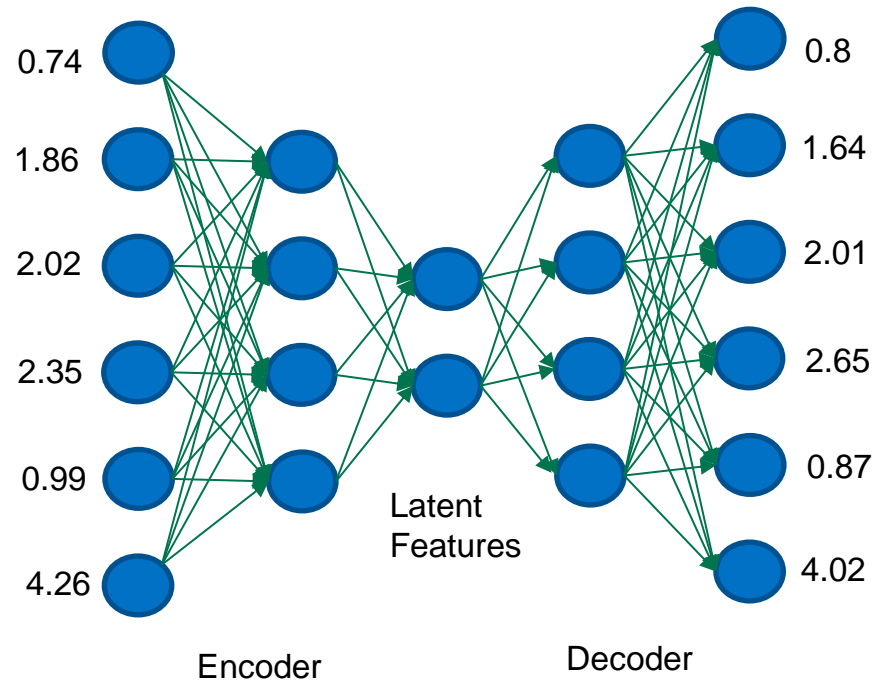


Iteration: 1

Deep Autoencoder - Anomaly Detection

Time	Signal
1	-0.5
2	0.74
3	1.86
4	2.02
5	2.35
6	0.99
7	4.26
...	
200	-0.5

Training

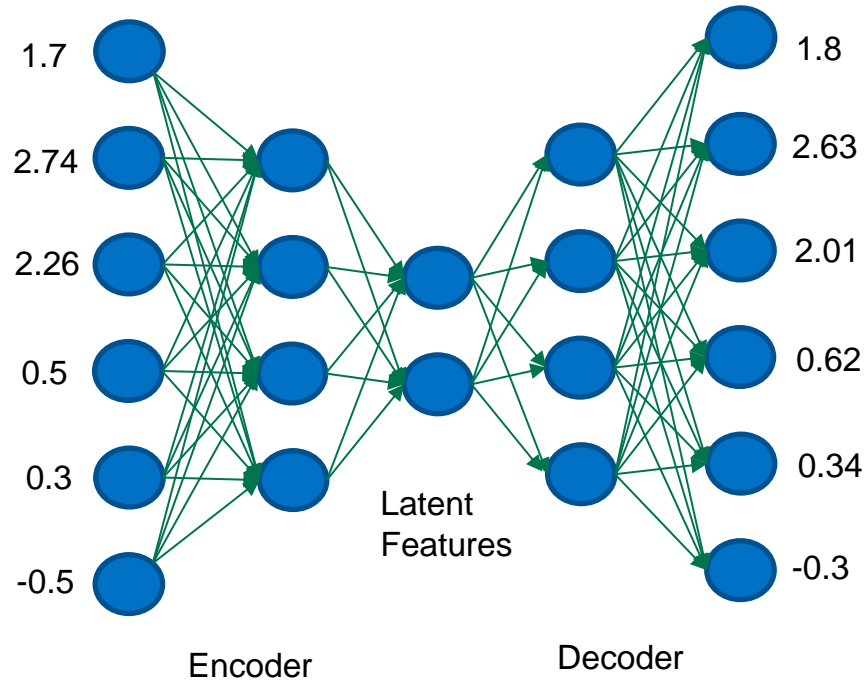


Iteration: 2

Deep Autoencoder - Anomaly Detection

Inference

Time	Signal
201	1.7
202	2.74
203	2.26
204	0.5
205	0.3
206	-0.5
207	-0.9
...	
300	-0.2

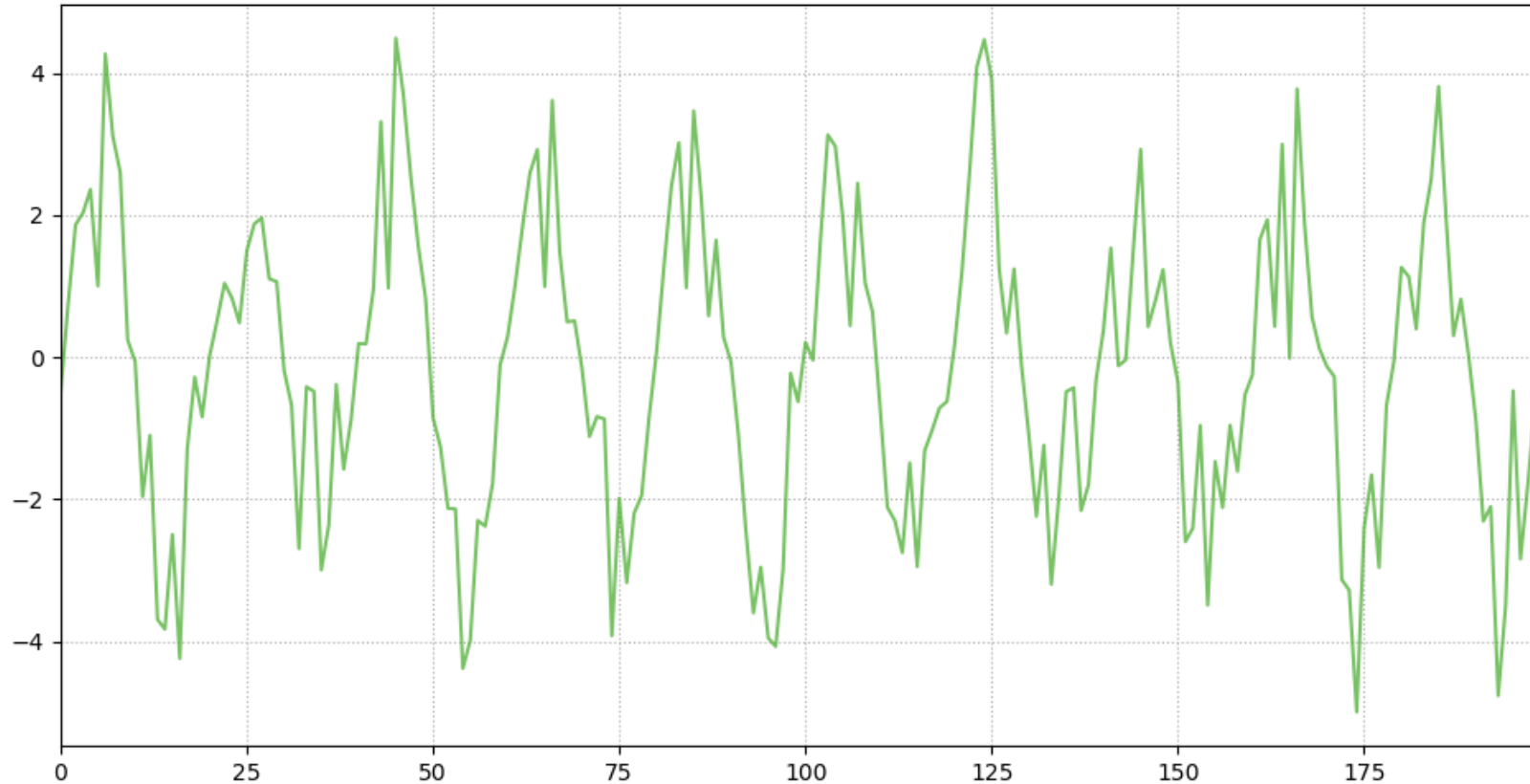


Reconstruction Error

Time	Signal
201	-0.1
202	0.11
203	0.25
204	-0.12
205	0.3
206	-0.04
207	-0.2
...	
300	

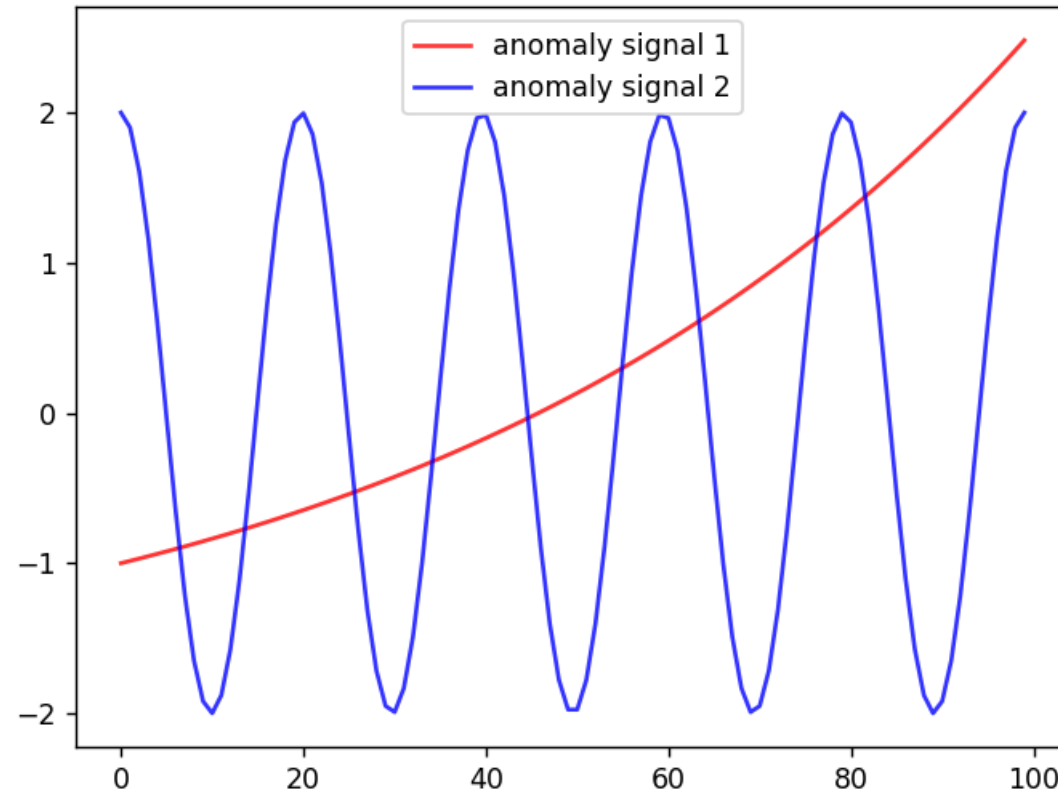
Deep Autoencoder - Anomaly Detection

- Original Signal



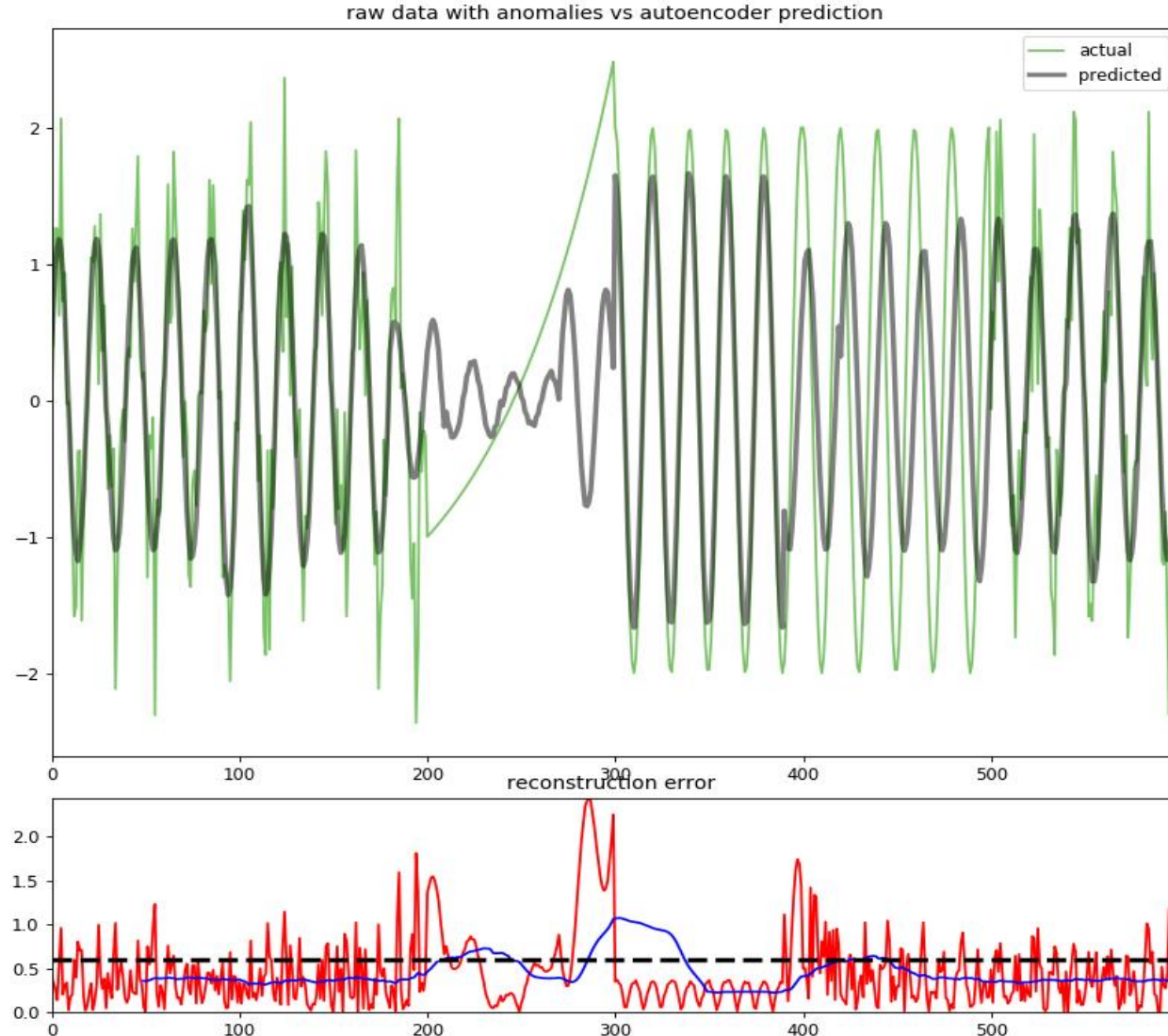
Deep Autoencoder - Anomaly Detection

- Anomaly Signals



Deep Autoencoder - Anomaly Detection

- Original Signal and Anomaly Signals



Statistical Arbitrage with Deep Autoencoders

TOPICS

- Lab Overview
- Financial Terminology
- A Simple Mean-Reversion Based Algorithmic Trading Example
- Lab
 - Deep Autoencoder
 - Stat-Arb
 - Backtesting
 - Performance Evaluations
 - Next Steps

WHAT THIS LAB IS

- An introduction to:
 - Financial Terminology
 - Financial Time Series Data
 - Keras with TensorFlow backend
 - Algorithmic Trading
- Hands-on exercises using Keras with TensorFlow backend for algorithmic trading
- Complete trading strategy that generates profit and loss curve (P&L)

ASSUMPTIONS

- You are familiar with Autoencoders
- Helpful to have:
 - Keras experience
 - Python experience

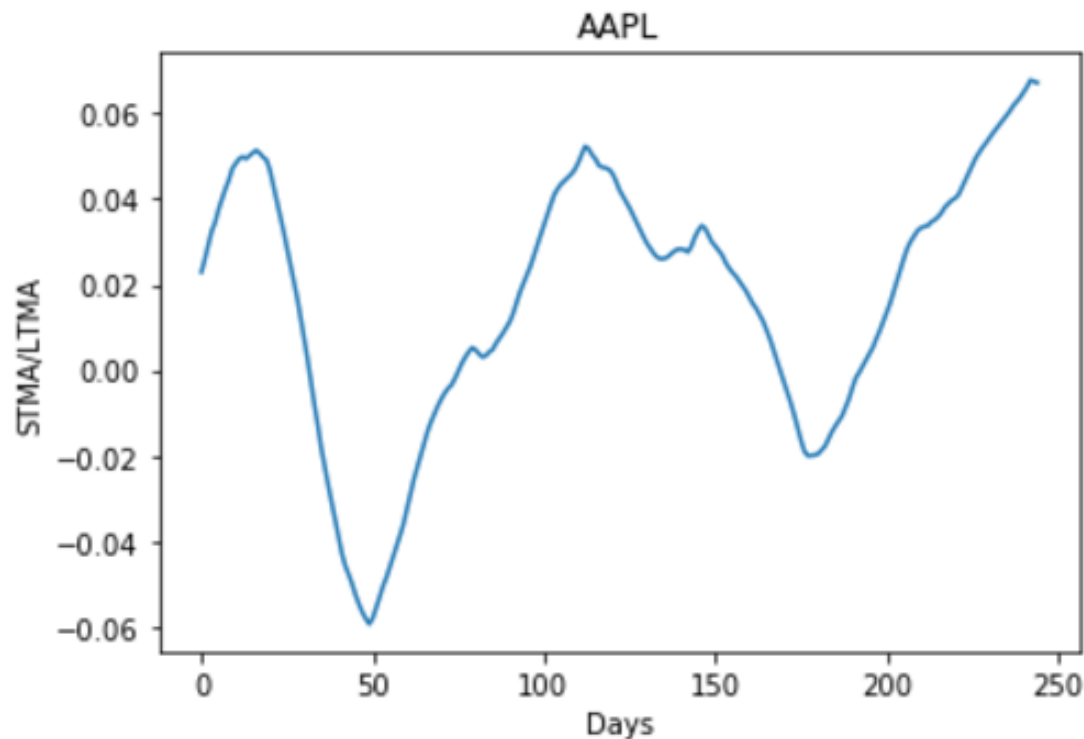
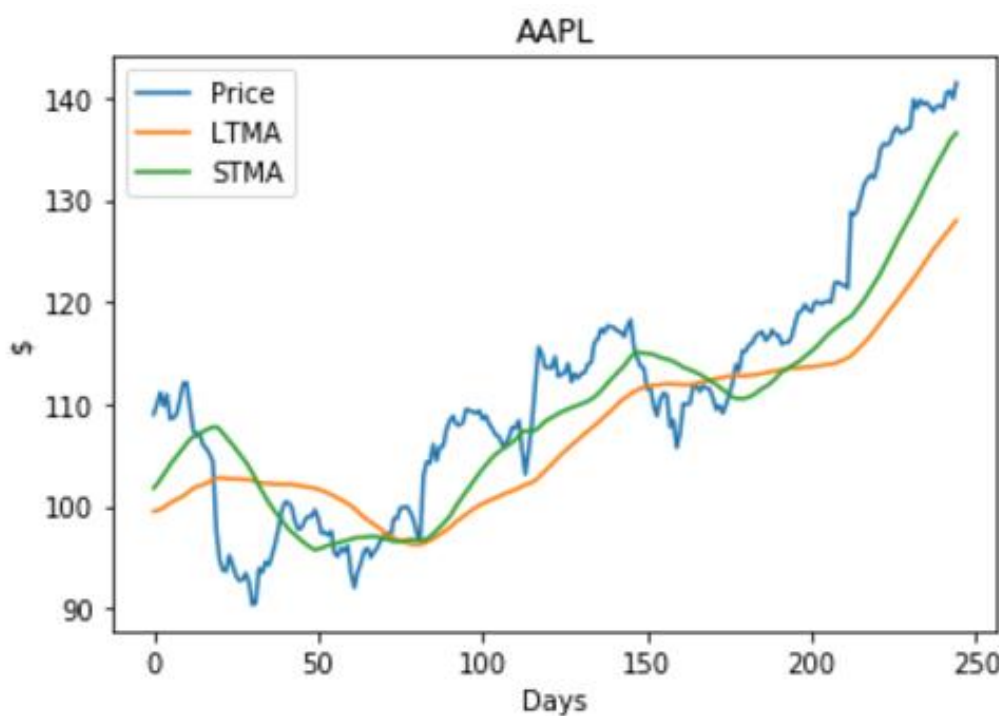
TAKE AWAYS

- Understanding the methods for algorithmic trading
- Ability to setup backtesting and train a autoencoder network
- Enough info to start using Keras with TensorFlow backend to learn from your own data

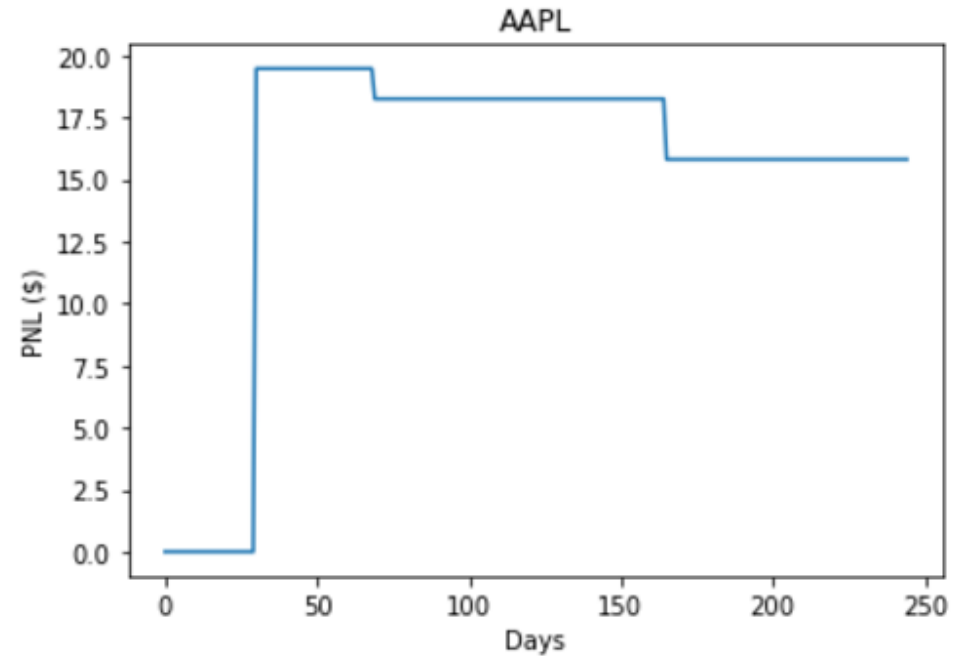
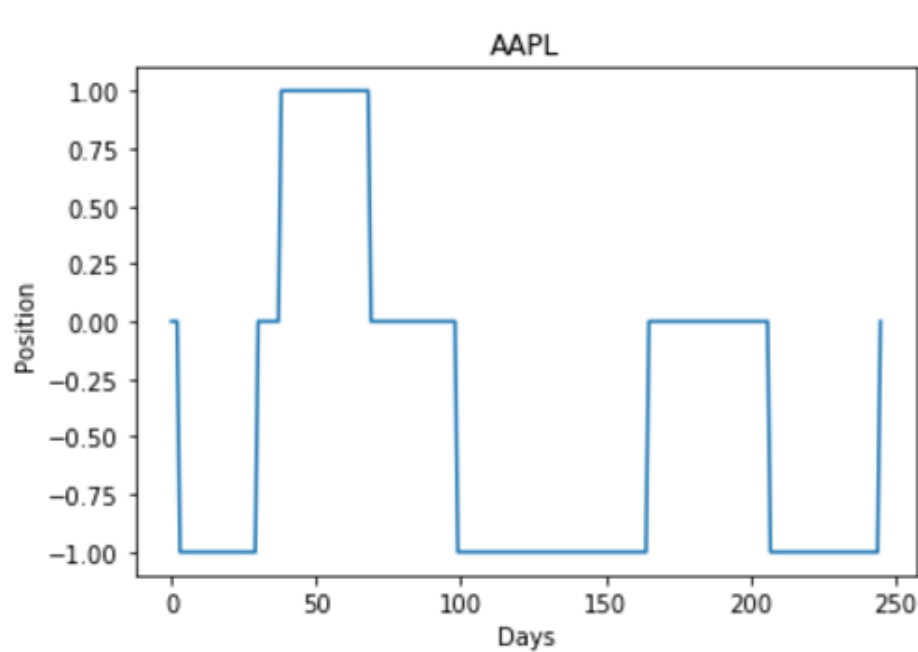
Statistical Arbitrage

- Arbitrage: simultaneous purchase and sale of an asset to profit from a difference in the price
- Let's assume that we have a group of stock from a same sector in our portfolio
- Typically, analysis of residuals calculated from a regression analysis of each stock in portfolio with respect to sector's ETF(s) will suggest that some stocks are cheap or over-priced. The residuals are mean-reverting (Ref: M. Avellaneda, Statistical Arbitrage in the U.S. Equities Market, 2009)
- Sector's ETF(s) or PCA are used to create the independent variables in regression analysis
- What happens if we use deep autoencoders?

Moving Average - Mean Reversion



Moving Average - Mean Reversion



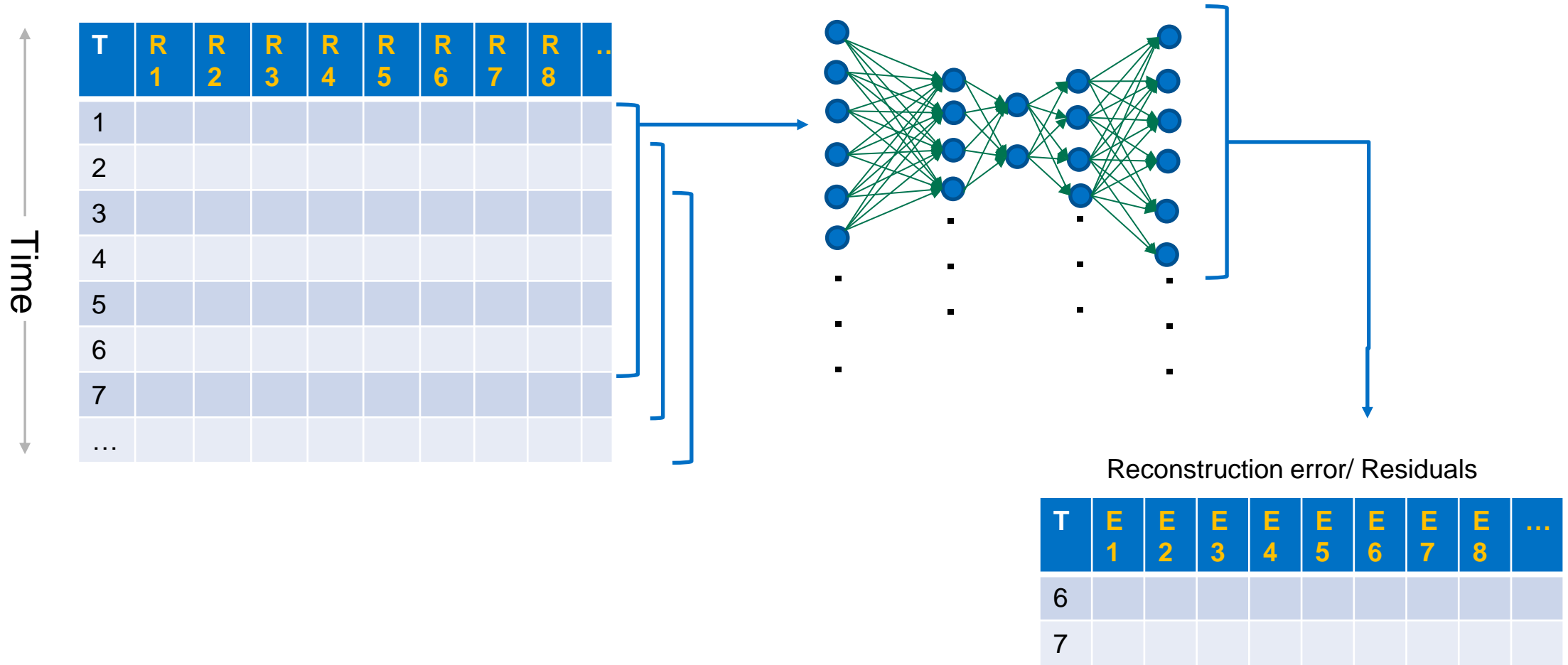
Deep Autoencoder based Stat-Arb

- Autoencoder
- Backtesting
- Data
- Performance Evaluation
- Hyperparameters

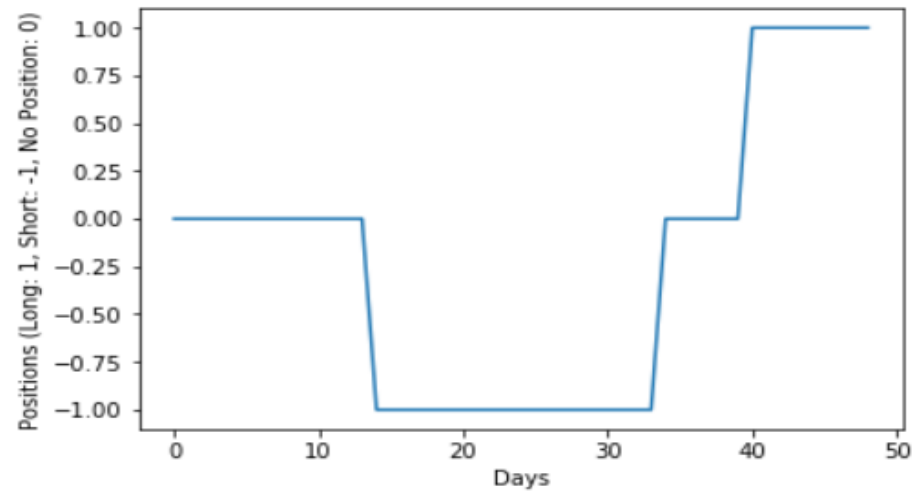
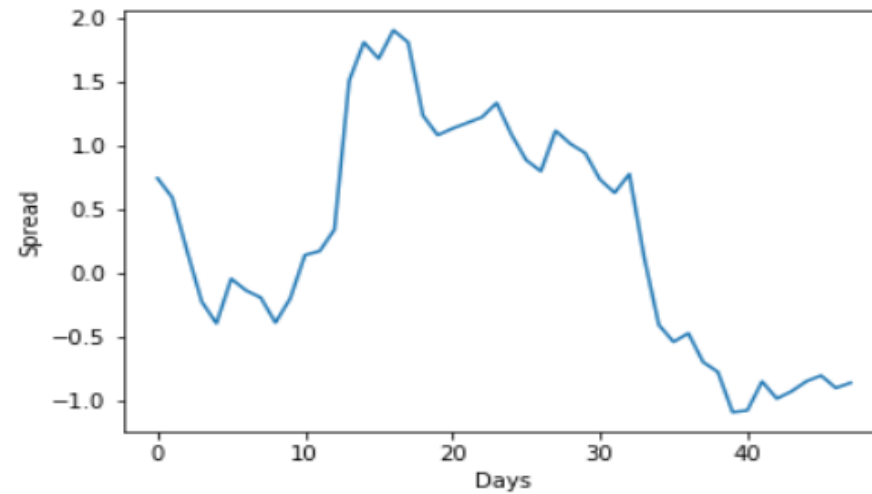
Dataset

	T	R1	R2	R3	R4	R5	R6	R7	R8	..
1										
2										
3										
4										
5										
6										
7										
...										

Backtesting



Residuals

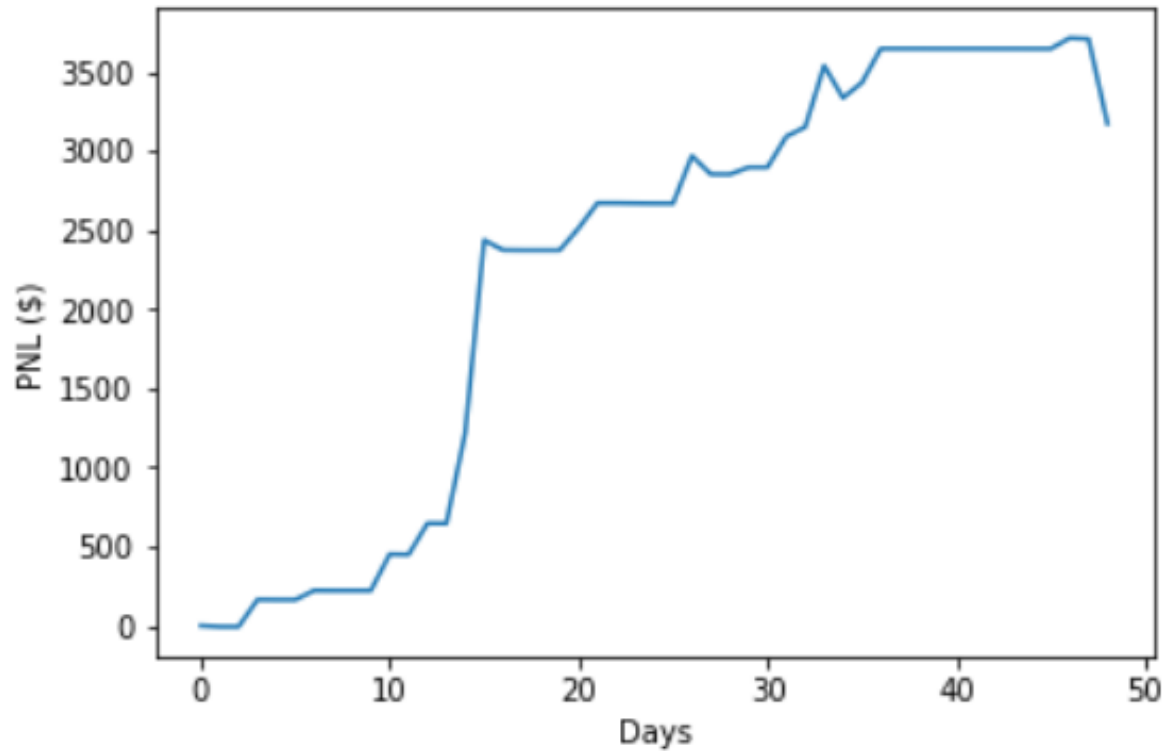


Hyperparameters

- T1: threshold to enter a position
- T2: threshold to exit a position
- Number of layers in deep autoencoder
- Historical data window
- Window for residual analysis

P&L

- Annual Sharpe Ratio: 4.58



Next Steps

After the lab, we recommend you doing the following options;

- Build a HPC system with multiple GPUs for overnight or real-time hyperparameter search.
- Develop the multi-GPU version of the code.
- Perform hyperparameter search periodically and use the best performing hyperparameters for a period (a day, a week, etc)
- Try deeper autoencoders and see the performance.

Instructor: Onur Yilmaz, Ph.D.



DEEP
LEARNING
INSTITUTE

www.nvidia.com/dli