



Gated recurrent unit(GRU)

구현 및 실험 결과 보고

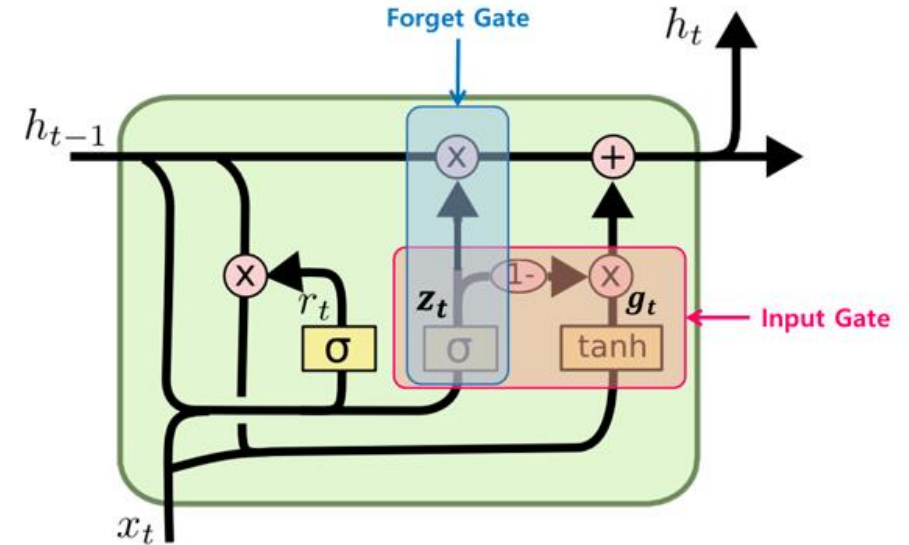
Tae-Won Lee,
leetee95@gmail.com
2020. 02. 09.

Introduction

- LSTM 모델을 이용한 자산가격의 변동성 예측이 가능해짐.
- KOSPI, NASDAQ과 같은 금융자산에서부터 , Gold, Oil과 같은 실물자산의 변동성 예측을 목표로 함.
- 전통적인 시계열 예측에서 부족했던 부분을 보완해줄 것을 기대함.

Related works

- Gated Recurrent Unit.
 - lstm은 cell state가 존재하지만, GRU에는 Cell state가 존재하지 않음.
 - Gate 가 2개로 lstm보다 연산량이 적음.



Proposed method



Experimental setting

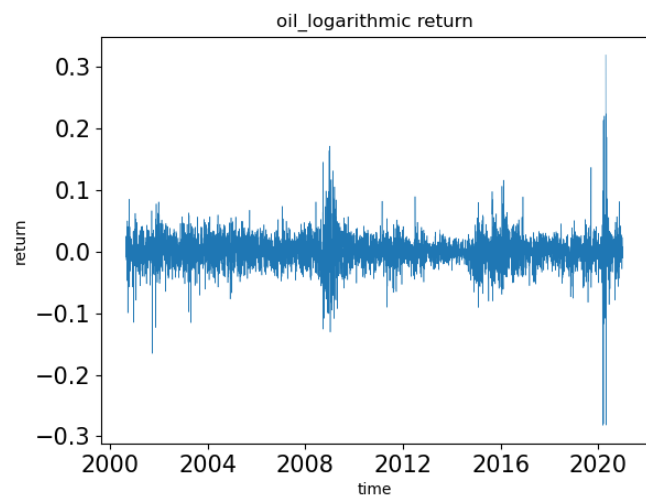
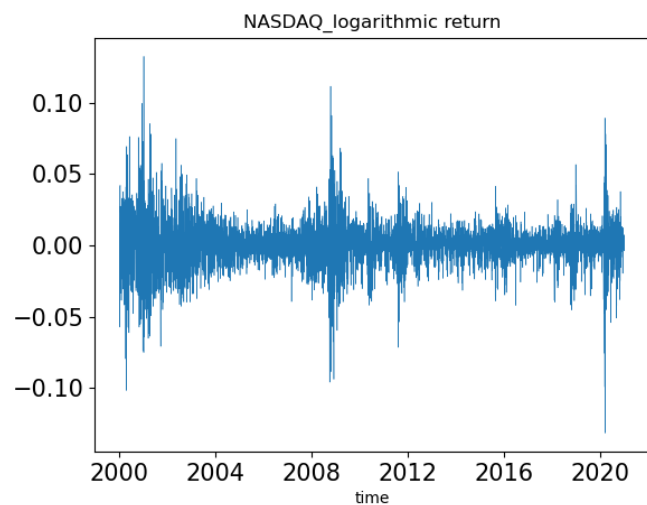
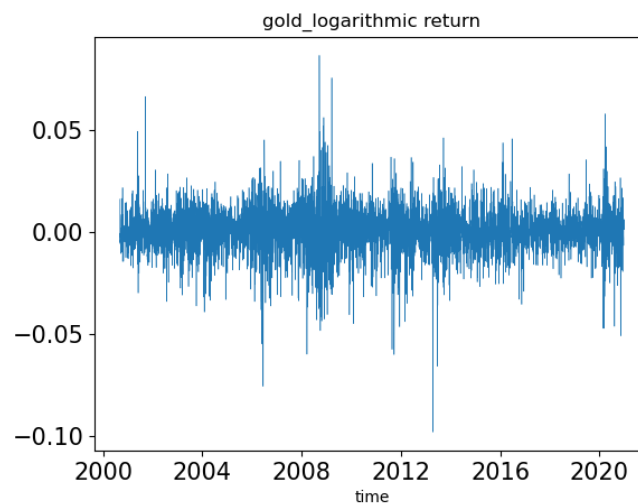
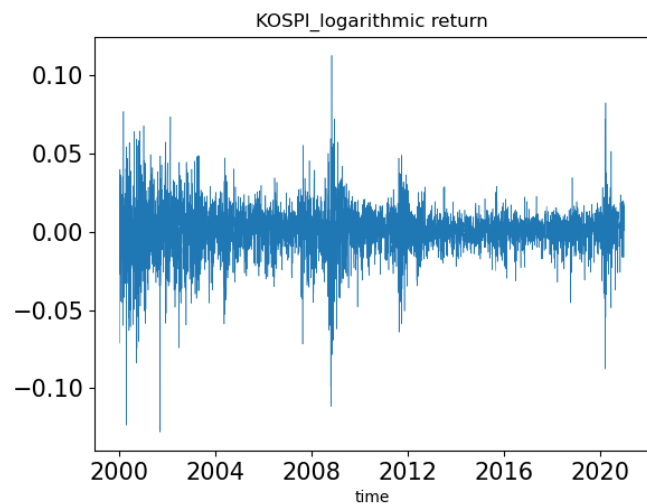
- Dataset

stocks	N	Train	val	test
KOSPI INDEX	5156	3210	979	967
NASDAQ	5265	3263	1002	1000
Gold price	5050	3067	994	989
Oil price	5060	3076	994	990

- (2000.01.01 – 2020.12.31) 기간 내의 데이터.
- 각 데이터는 시초가(Open), 고가(High), 저가(Low), 종가(Close), 수정주가(Adj Close), 거래량(Volume) 6개의 변수가 존재. 종가(close)를 사용하여 예측.
- pandas_datareader 모듈에서 data 수집.

Experimental setting

- logarithmic return



- logarithmic return : $\ln((t+1)/t)$ 로 나타냄
t : 시점이 t일때 종가(close)

Experimental setting

- Statistics

stocks	Mean	Standard deviation	Skewness	Kurtosis	Jerque Bera	ADF
KOSPI INDEX	1582.518	587.343	-0.421	1.861	432.61***	-0.438834
NASDAQ	3791.769	2367.556	1.388	4.395	2125.9183***	3.346301
Gold price	997.637	479.258	-0.105	1.796	315.6375***	-0.459757
Oil price	61.785	25.912	0.400	2.381	216.52333***	-1.997886
KOSPI INDEX log difference	0.000193	0.0149	-0.572	10.03	10943.0983***	-33.339841***

- Experiment Settings

Hyperparameters	Number
hidden state, cell state dimension	10
input sequence length	7
output sequence length	1
epoch	15
Batch size	128

Comparison results

- Results of out-of-sample forecasts with various loss function

Model	data	MAE	MSE
RNN	KOSPI INDEX	0.000222346732958353	0.21690651135785238
	NASDAQ	0.001294612086244992	0.10981855647904533
	Gold price	0.0012410988232919148	0.19466256030968257
	Oil price		
LSTM	KOSPI INDEX	9.395793832871796e-06	0.03744102056537356
	NASDAQ	0.00020127013392214264	0.022208844178489277
	Gold price	0.0012053362359958036	0.0440918505191803
	Oil price		
GRU	KOSPI INDEX	0.7917789646557399	0.15417489835194179
	NASDAQ	0.007145870210868972	0.0060817864058273175
	Gold price	0.016360039689711163	0.08163328043052129
	Oil price		

Performance Measurements

- Stock Dataset
 - pandas_datareader 모듈에서 data 수집.
 - 각 반복마다 7개의 input과 1개의 target 값이 tuple 형태로 묶여 반환.
(7개 input, 1개 output)
- Train, validation
 - loss function 으로 MSE, MAE Loss를 사용.
 - metric에서 사용하는 함수에 따라 손실함수 변환.
- Test
 - loss function 으로 MSE, MAE Loss를 사용.
 - metric을 적용하여 모형 최종 평가.

Reference

- 1 Ha Young Kim, & Chang Hyun Won (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems With Applications, Vol.103. 25-37.
<https://doi.org/10.1016/j.eswa.2018.03.002>

Appendix

Appendix

- Jarque-bera test
 - 왜도와 첨도가 정규분포와 일치하는지 판단하는 적합도 검정
- ADF test
 - 시계열 데이터가 정상성(stationary)을 만족하는지에 대한 검정
- LM-ARCH 검정
 - ARCH, GARCH와 같은 모형을 적용할 수 있는지에 대한 검정.