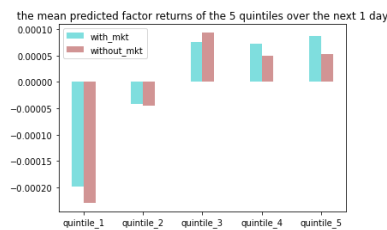
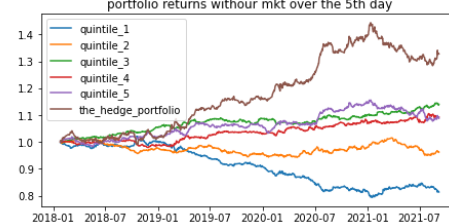
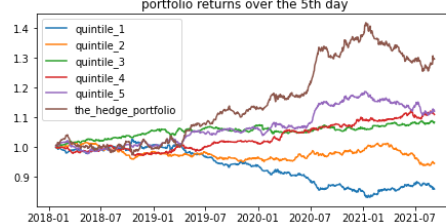
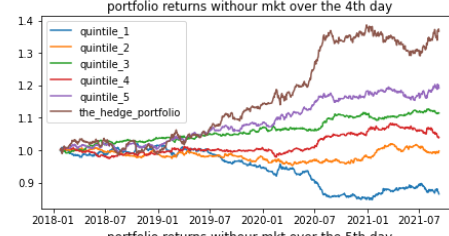
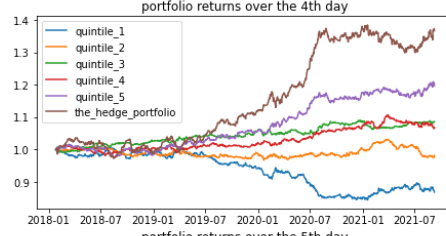
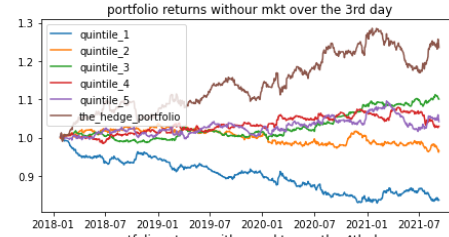
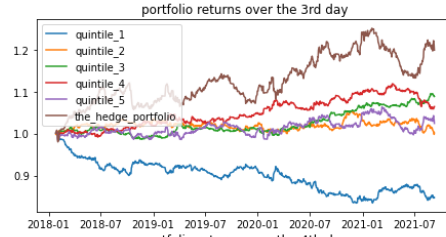
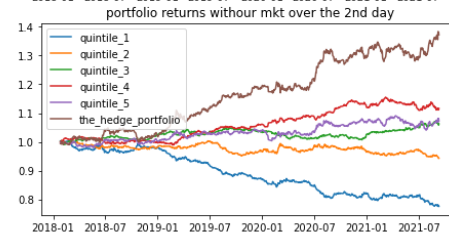
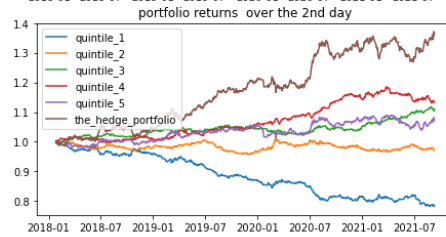
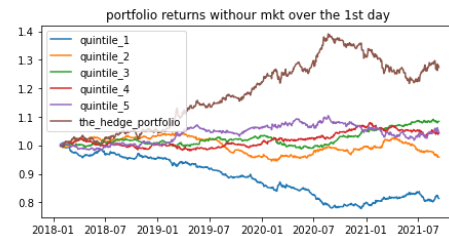
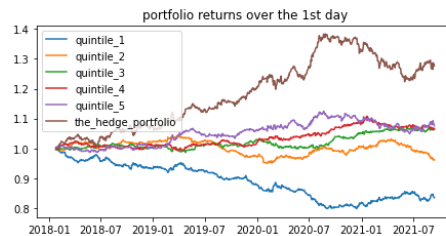


The Model is predicting the **next day** return with **next day return as label** in Random Forest
The input is the **all** factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]
The model is updated **after 125 days** before next day prediction without normalized
Due to sklearn does have low CPU occupation, the computation could be every slow.
The parameters every half of year for grid search are set like

"n_estimators": [20,80,160],
"max_depth": [2,4,8],
"min_samples_leaf": [4,10,20,40,80]

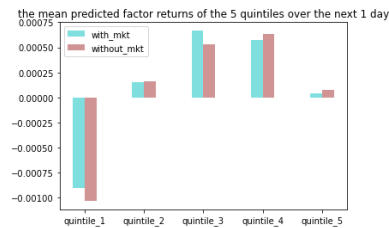


	annualized_return	annualized_volatility	sharpe_ratio
1st_the_hedge_portfolio	0.069666	0.066224	1.087611
2nd_the_hedge_portfolio	0.088708	0.067251	1.34265
3rd_the_hedge_portfolio	0.051481	0.067564	0.803608
4th_the_hedge_portfolio	0.090259	0.066719	1.375933
5th_the_hedge_portfolio	0.074432	0.062915	1.21469
1st_the_hedge_portfolio_without_mkt	0.068905	0.065054	1.094653
2nd_the_hedge_portfolio_without_mkt	0.091067	0.065381	1.413453
3rd_the_hedge_portfolio_without_mkt	0.060275	0.064416	0.973632
4th_the_hedge_portfolio_without_mkt	0.090334	0.063668	1.439876
5th_the_hedge_portfolio_without_mkt	0.081524	0.06237	1.334165

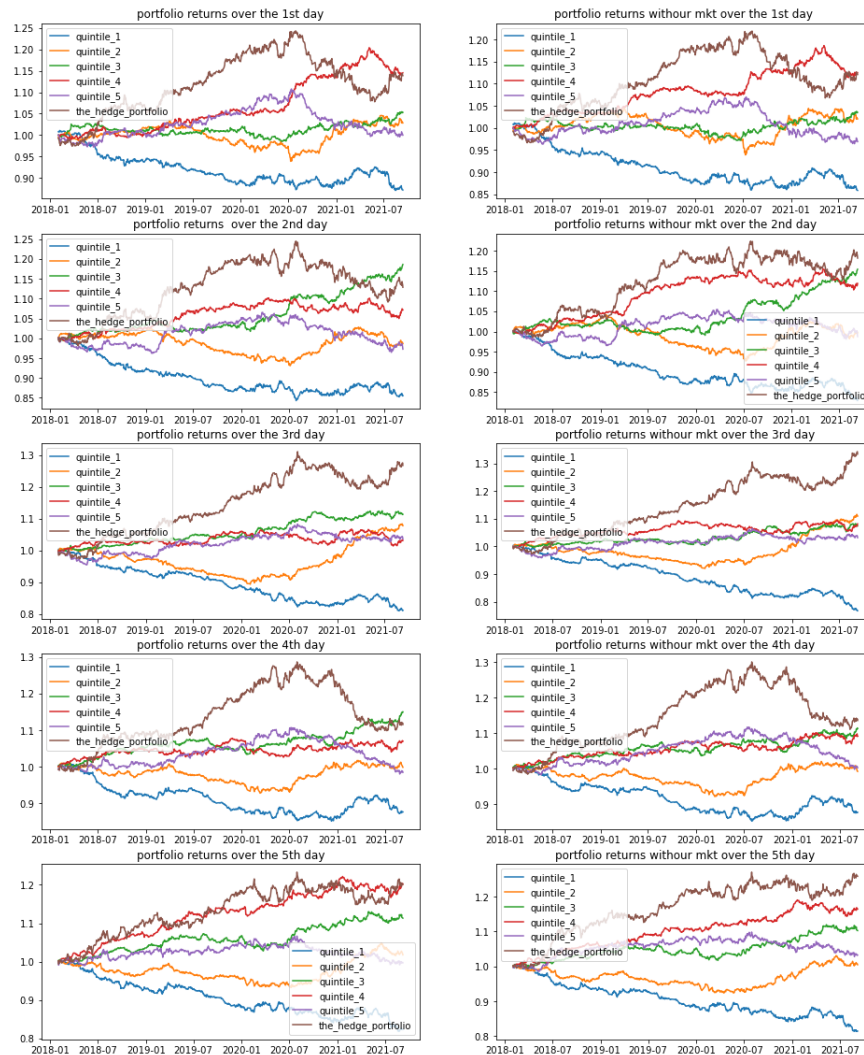


The Model is predicting the next day return with next week return as label in Random Forest
The input is the all factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]
The model is updated after 125 days before next day prediction without normalized
The parameters are set like

"n_estimators": [20,80,160],
"max_depth": [2,4,8],
"min_samples_leaf": [4,10,20,40,80]



	annualized_return	annualized_volatility	sharpe_ratio
1st_the_hedge_portfolio	0.036571	0.061459	0.636696
2nd_the_hedge_portfolio	0.034094	0.064967	0.566934
3rd_the_hedge_portfolio	0.068049	0.063395	1.107687
4th_the_hedge_portfolio	0.030785	0.06258	0.533458
5th_the_hedge_portfolio	0.05232	0.064447	0.852737
1st_the_hedge_portfolio_without_mkt	0.032002	0.060634	0.56899
2nd_the_hedge_portfolio_without_mkt	0.047338	0.063661	0.784292
3rd_the_hedge_portfolio_without_mkt	0.083315	0.063785	1.331871
4th_the_hedge_portfolio_without_mkt	0.035746	0.06337	0.606134
5th_the_hedge_portfolio_without_mkt	0.066009	0.064289	1.063175



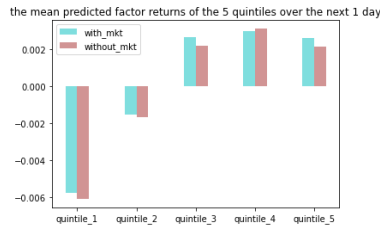
The Model is predicting the **next day** return with **next month return as label** in Random Forest
The input is the **all** factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]
The model is updated **after 125 days** before next day prediction without normalized

The parameters are set like

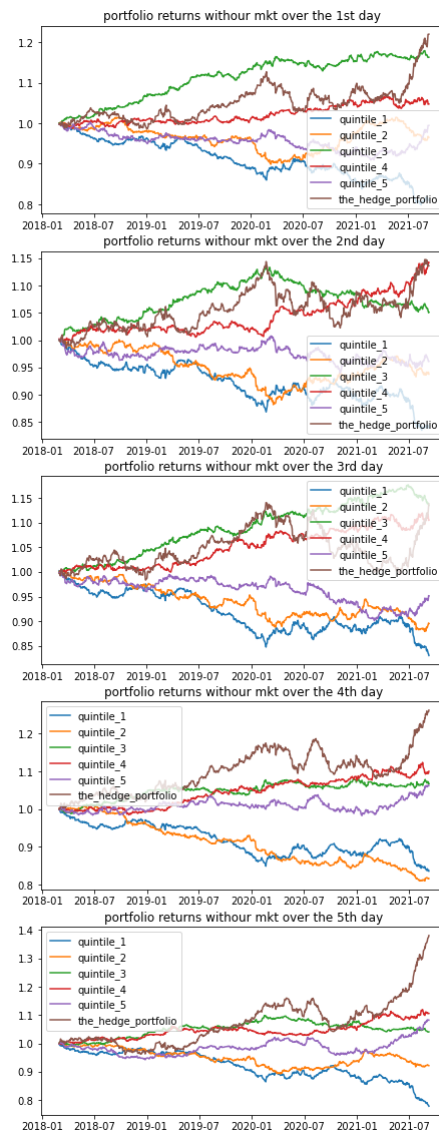
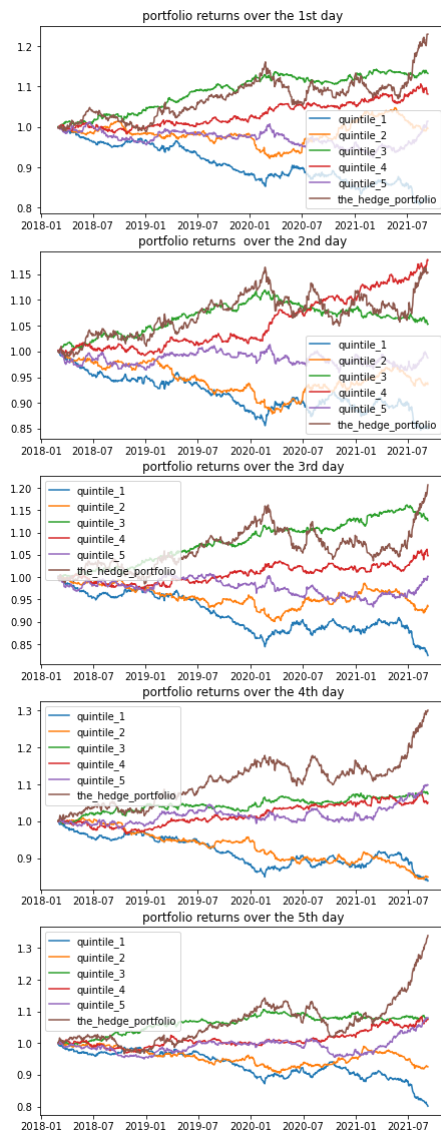
"n_estimators": [20,80,160],

"max_depth": [2,4,8],

"min_samples_leaf": [4,10,20,40,80]



	annualized_return	annualized_volatility	sharpe_ratio
1st_the_hedge_portfolio	0.060209	0.062231	1.001558
2nd_the_hedge_portfolio	0.041292	0.063581	0.688344
3rd_the_hedge_portfolio	0.054705	0.062704	0.908042
4th_the_hedge_portfolio	0.077514	0.061516	1.283877
5th_the_hedge_portfolio	0.086545	0.062325	1.406867
1st_the_hedge_portfolio_without_mkt	0.057742	0.060842	0.983466
2nd_the_hedge_portfolio_without_mkt	0.036821	0.063028	0.623393
3rd_the_hedge_portfolio_without_mkt	0.037243	0.061112	0.648105
4th_the_hedge_portfolio_without_mkt	0.068246	0.060774	1.152036
5th_the_hedge_portfolio_without_mkt	0.095923	0.061812	1.561664



Conclusion: the all factors were taken into account, but the week and the month return as label has a poorer performance than the next day return as label.

For one thing, it could be that we use the week and month return as label but in back test we just rebalance the portfolio each day. The target we need in reality is not the same as the train label in definition, for example tarin label is next day return but the target is next 5th day return, the train label is next month return and the target is next day return.

For another, it could be the search grid is too large and even the optimal parameters are not the global optimal. The performance of back test is not so good.

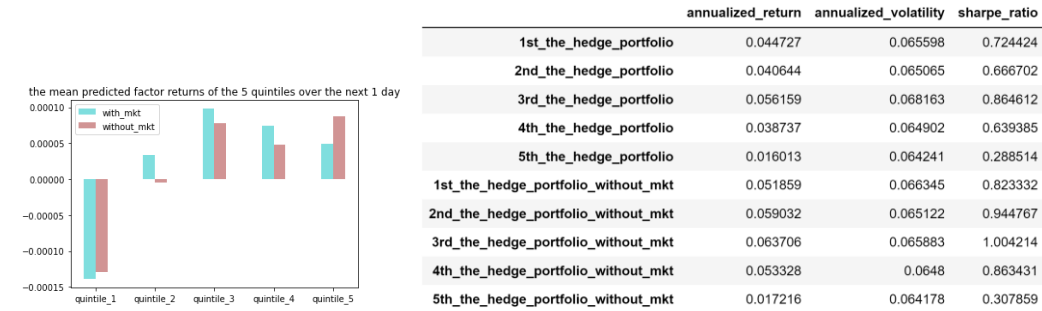
Finds: as I have trained multiple models, after 2020-07, there is a reversal effect nearly in all model for factor we generate.

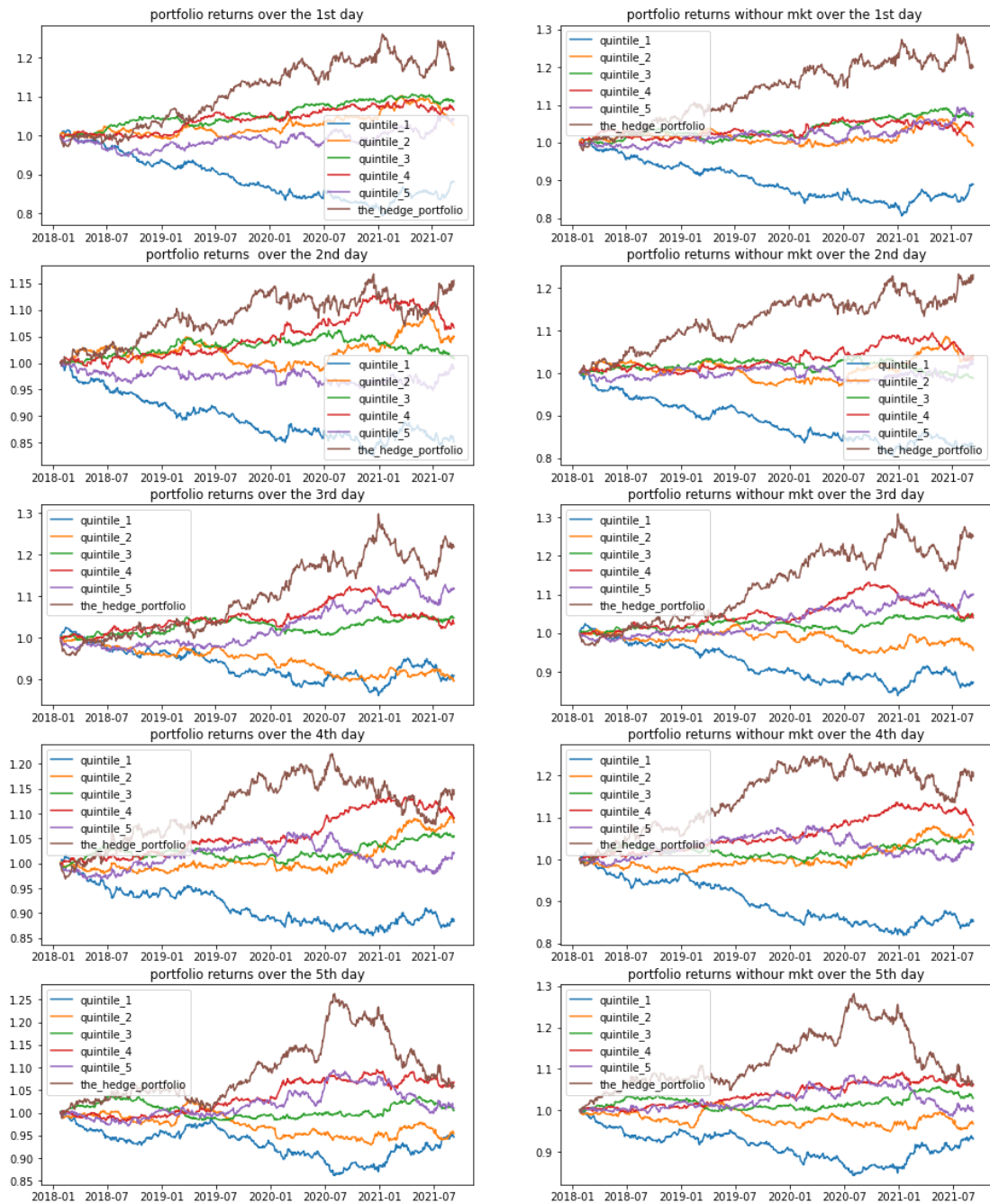
The Model is predicting the next day return with next day return as label in Random Forest
The input is the single kind factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]

The model is updated after 125 days before next day prediction with grid search.

The parameters are set like

"n_estimators": [20,40,80,120,160,200],
"max_depth": [2,3,5,8,12,16],
"min_samples_leaf": [2, 5,10,20,40,80]





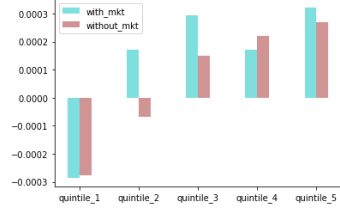
The Model is predicting the **next day** return with **next week return as label** in Random Forest
The input is the **single** kind factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]

The model is updated **after 125 days** before next day prediction.

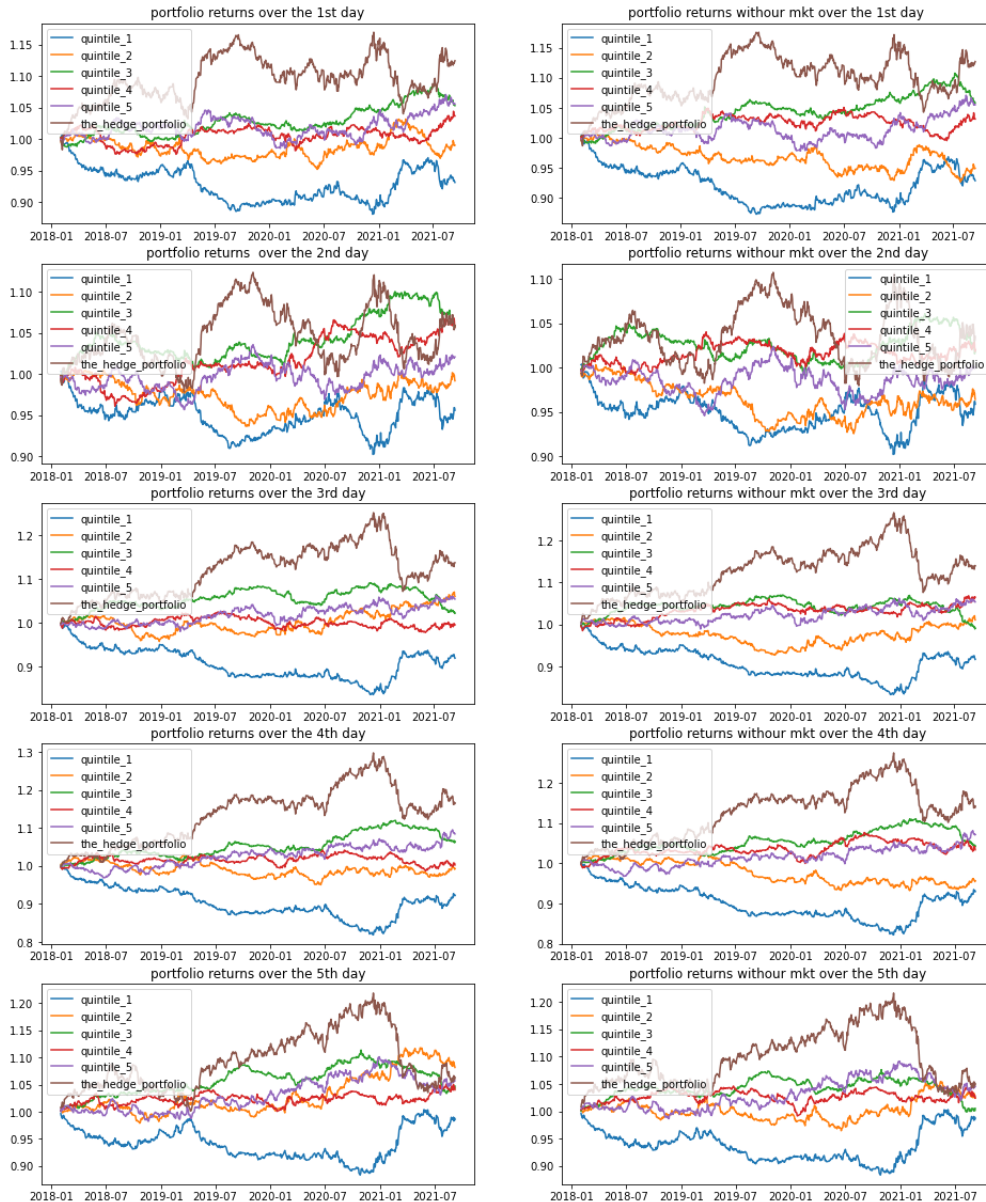
The parameters are set like

```
"n_estimators": [20,40,80,120,160],
"max_depth": [2,3,5,8,12,16],
"min_samples_leaf": [ 2, 5,10,20,40]
```


the mean predicted factor returns of the 5 quintiles over the next 1 day



	annualized_return	annualized_volatility	sharpe_ratio
1st_the_hedge_portfolio	0.032981	0.06722	0.534149
2nd_the_hedge_portfolio	0.015979	0.066822	0.279127
3rd_the_hedge_portfolio	0.03642	0.066212	0.592888
4th_the_hedge_portfolio	0.043425	0.068725	0.675454
5th_the_hedge_portfolio	0.017307	0.067571	0.29707
1st_the_hedge_portfolio_without_mkt	0.033699	0.066493	0.550092
2nd_the_hedge_portfolio_without_mkt	0.009544	0.066384	0.181389
3rd_the_hedge_portfolio_without_mkt	0.037091	0.066272	0.602545
4th_the_hedge_portfolio_without_mkt	0.037575	0.069017	0.588451
5th_the_hedge_portfolio_without_mkt	0.013831	0.068428	0.242341



The Model is predicting the next day return with next month return as label in Random Forest
The input is the single kind factors return over day -1, -2, -3, [-4,-5],[-6,-21],[-22,-64],[-65,-126],[-127,-252]

The model is updated after 125 days before next day prediction

The parameters are set like

"n_estimators": [20,40,80,120,160],
 "max_depth": [2,3,5,8,12,16],
 "min_samples_leaf": [2, 5,10,20,40]

