



**SIAT**  
QuanTech 2019  
CONFERENCE  
MILANO





# Applicazione in Finanza di Modelli Quantitativi E Machine Learning

**Luigi Piva**

**[luigi.piva@quantlab.co.uk](mailto:luigi.piva@quantlab.co.uk)**





### PROFILO

- Laurea in Scienze Statistiche presso l'Università di Bologna
- Master Econometria all'Università di Siena
- CQF (Certificate Quantitative Finance) - 2010- London - Math Premiere

### Attività

- 2007-2012 Trader Quantitativo per il gruppo Equity Line Solutions –Seven Mills Ltd Londra - UK
- Dal 2012 Trader Quantitativo per Quantlab Limited – Londra -UK
- Dal 2004 Socio SIAT

### Pubblicazioni

- “Analisi Tecnica Efficace Applicata ai Trading Systems” Experta, 2004
- “Top Trader Readers Digest 2005-2006” Experta 2006
- “Uncertain volatility with static hedge by Finite Difference: pricing vanilla and binary options ” Fitch 7City Londra 2011
- “Multivariate Financial Time Series Analysis: energy futures” Fitch 7City Londra 2011

### Campionati

- Campionati di Trading: +61% TopTraderCup 2011, +42% TopTraderCup 2010, +38% TopTraderCup 2009





**Possiamo prevedere il rendimento di  
attività finanziarie usando Machine  
Learning?  
E' più o meno efficace di altri  
Modelli previsionali?**







## CASO DI STUDIO



- **Usiamo:** dati storici monovariati EurUsd indicatori tecnici
- **Creiamo:** un modello previsionale regressione/machine learning back-test





# Uso di linguaggio matematico per fare previsioni sul futuro

Input/  
Predictors  
→

Predictive  
model

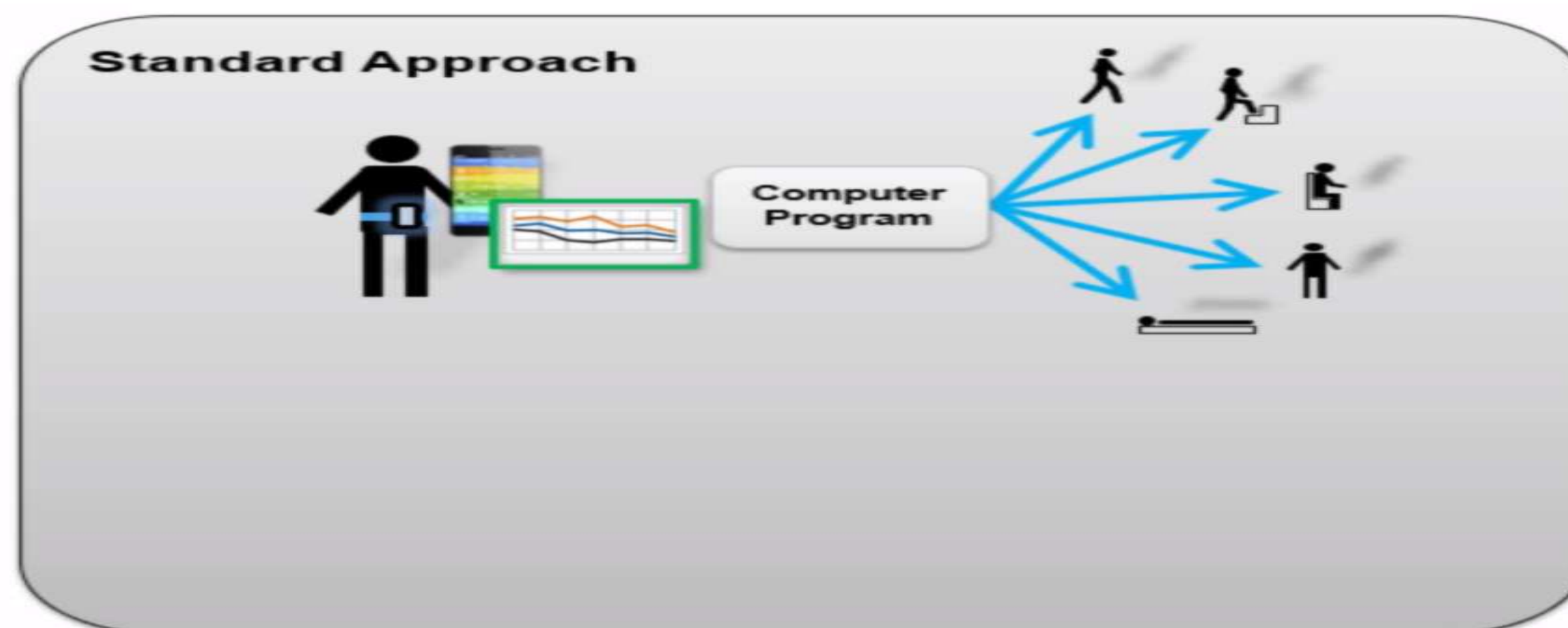
Output/  
Response  
→

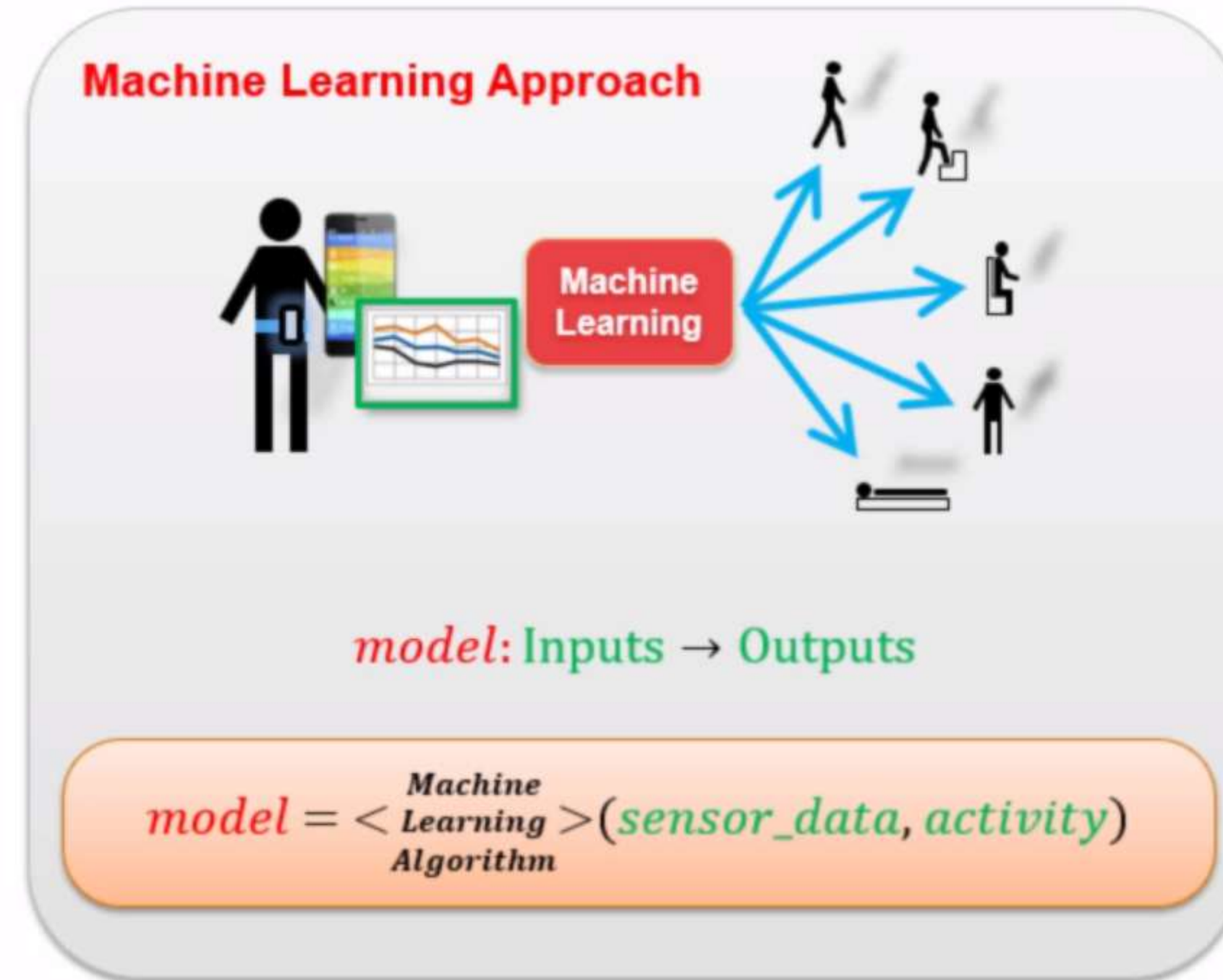
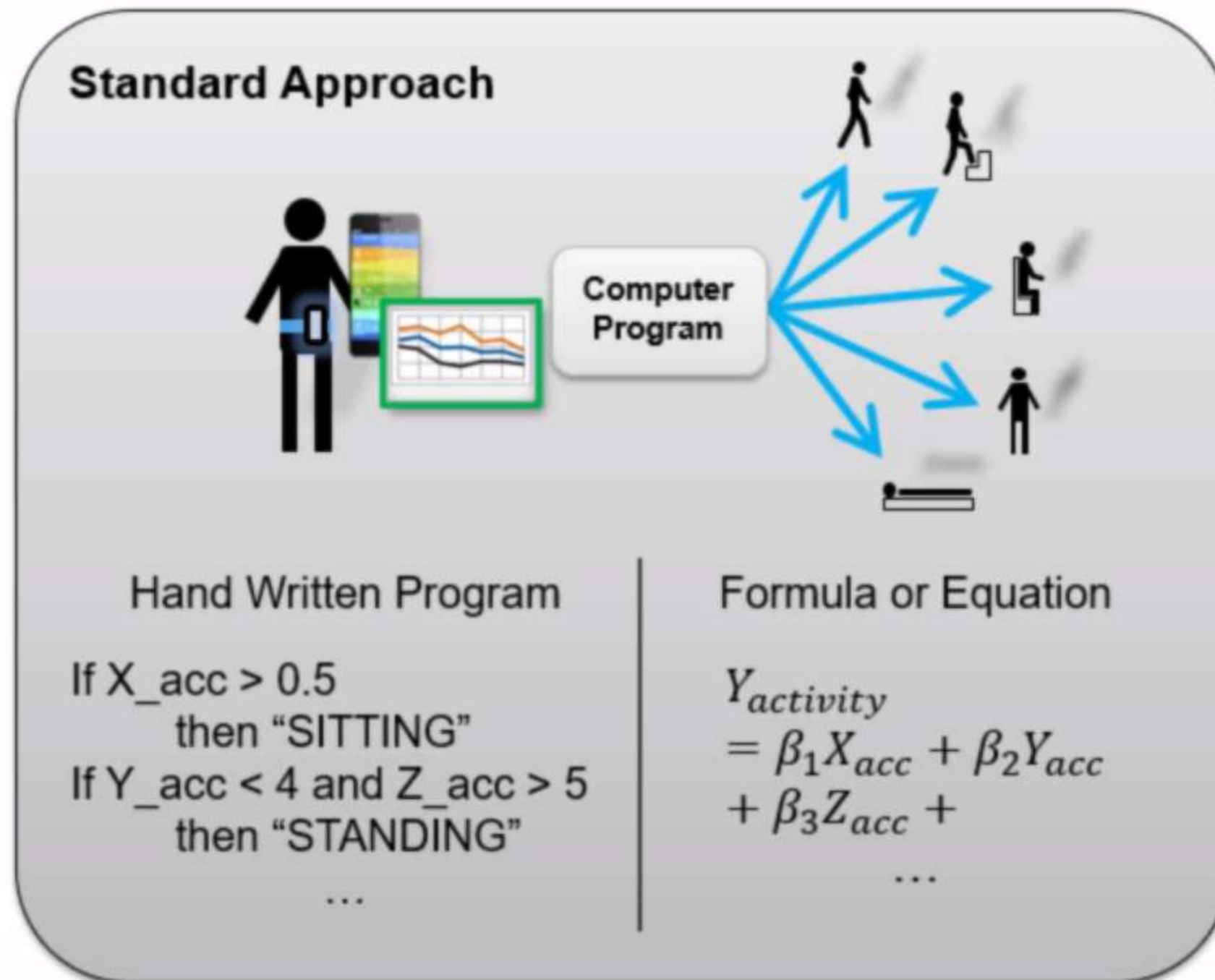




## Cos'è il Machine Learning?

Il Machine Learning usa i dati per generare un programma che svolge la funzione desiderata

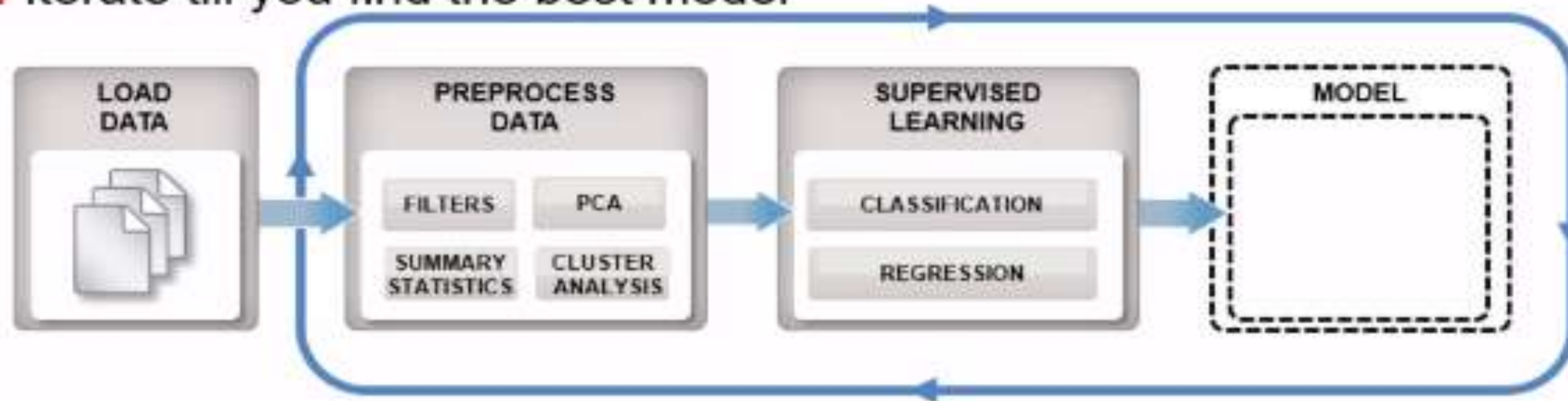








**Train:** Iterate till you find the best model



**Predict:** Integrate trained models into applications





## Come procediamo?

- Analizziamo i **Dati**
- Costruiamo i **Fattori**
- Eseguiamo il **Training** del modello
- Facciamo il **Back-Test**





## DATI

- Serie storica **EURUSD**
- Dati ad un minuto del **Mid-Price Bid/Ask**
- Salvati in un oggetto **Timetable**

10x3 timetable

Time	Mid	Bid	Ask
01-Jan-2007 00:00:00	1.31916	1.31908	1.31924
01-Jan-2007 00:01:00	1.31929	1.31921	1.31937
01-Jan-2007 00:02:00	1.31954	1.31946	1.31962
01-Jan-2007 00:03:00	1.31963	1.31958	1.31968
01-Jan-2007 00:04:00	1.31952	1.31945	1.31959
01-Jan-2007 00:05:00	1.3195	1.31942	1.31958
01-Jan-2007 00:06:00	1.31945	1.3194	1.3195
01-Jan-2007 00:07:00	1.31965	1.31962	1.31968
01-Jan-2007 00:08:00	1.31958	1.31953	1.31963
01-Jan-2007 00:09:00	1.319525	1.31945	1.3196





## Fattori- Previsori

- Indicatori di Analisi Tecnica  
**Rsi - Macd**
- **Rendimenti** a n-Minuti (5,10,15,20,25,30 e 60 )







## Training

Usiamo i Fattori definiti e proviamo a prevedere le future variazioni dei prezzi, tradando in base a queste previsioni

## Modello

Performiamo una Regressione Lineare Stepwise previsioni e utilizziamo Classification Trees

## Back-Testing

Su 10 anni di dati usando Bid/Ask per considerare i costi.  
Variamo le finestre In-Sample e Out-of-Sample





## In ambiente Matlab

\_carichiamo i dati a 1 minuto

\_definiamo una frequenza di trading (60 minuti) e raggruppiamo i dati

\_otteniamo una prima tabella dei Fattori (Previsori) e un vettore  
*response*

24487x18 timetable

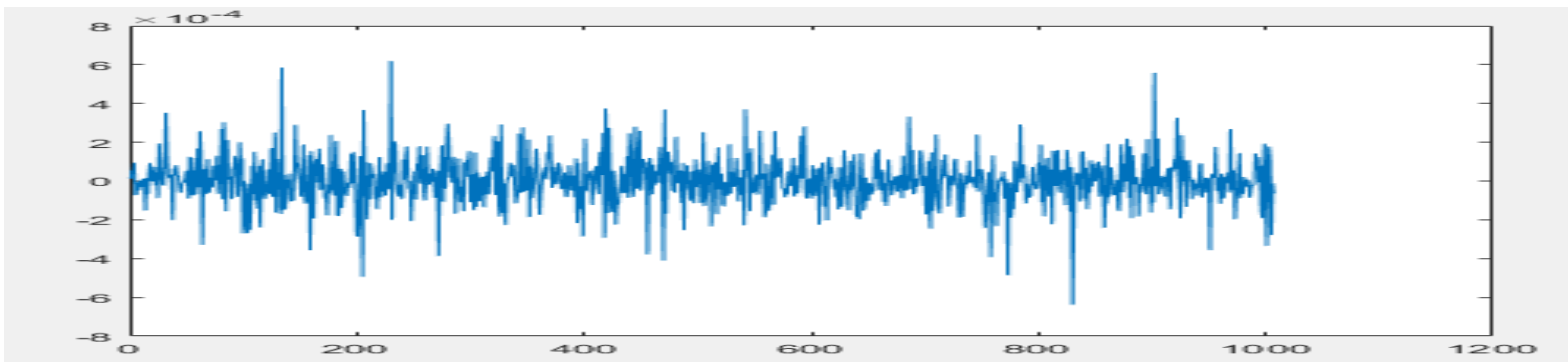
	Time	1 Mid	2 Bid	3 Ask	4 Return_5	5 Return_10	6 Return_15	7 Return_20	8 Return_25	9 Return_30	10 Return_60	11 macd	12 rsindex_5
1	01-Jan-2007 01:00:00	1.3193	1.3193	1.3194	4.5481e-05	-3.0318e-05	-6.0635e-05	6.0642e-05	1.0613e-04	1.8950e-05	1.0613e-04	7.8319e-06	68.75
2	01-Jan-2007 02:00:00	1.3190	1.3189	1.3191	3.7908e-05	-7.5808e-05	0	2.2745e-05	-6.0647e-05	-1.2129e-04	-2.1223e-04	3.8748e-06	55.31
3	01-Jan-2007 03:00:00	1.3196	1.3196	1.3197	8.3364e-05	3.7891e-05	9.8522e-05	0	2.2734e-05	7.5779e-06	4.6246e-04	1.3548e-05	60.00
4	01-Jan-2007 04:00:00	1.3194	1.3194	1.3195	-9.8518e-05	-8.3363e-05	-7.5785e-05	-1.8944e-04	-1.5158e-05	-4.5472e-05	-1.5914e-04	-1.5871e-06	30.30
5	01-Jan-2007 05:00:00	1.3195	1.3194	1.3195	1.4402e-04	9.8534e-05	1.8951e-04	2.1225e-04	1.8192e-04	1.2128e-04	3.7895e-05	2.0963e-05	70.21
6	01-Jan-2007 06:00:00	1.3195	1.3194	1.3195	8.7165e-05	6.0635e-05	1.3644e-04	1.1749e-04	7.9585e-05	9.4745e-05	-1.1368e-05	1.3824e-05	68.25
7	01-Jan-2007 07:00:00	1.3192	1.3192	1.3192	-1.2885e-04	-1.5161e-05	-1.4401e-04	-7.5798e-05	-1.2127e-04	-6.8219e-05	-2.0084e-04	-1.9716e-05	32.29
8	01-Jan-2007 08:00:00	1.3192	1.3192	1.3193	-3.7901e-06	-1.1370e-05	-5.3058e-05	7.2017e-05	-1.1370e-05	-9.8532e-05	3.4112e-05	-2.4004e-06	49.20
9	01-Jan-2007 09:00:00	1.3169	1.3169	1.3169	1.5187e-05	-5.3152e-05	-1.2148e-04	0	-4.5559e-05	-1.1389e-04	-0.0018	-4.7755e-05	52.08
10	01-Jan-2007 10:00:00	1.3170	1.3170	1.3171	0	6.8341e-05	-2.2778e-05	2.2779e-05	9.8717e-05	0	7.5935e-05	6.9980e-06	
11	01-Jan-2007 11:00:00	1.3169	1.3169	1.3170	-7.5933e-06	-5.6947e-05	-6.0743e-05	2.2781e-05	0	1.5187e-05	-6.0743e-05	-5.2364e-06	48.78
12	01-Jan-2007 12:00:00	1.3170	1.3170	1.3171	6.8341e-05	-4.1759e-05	6.0747e-05	-3.7963e-05	0	6.8341e-05	6.0747e-05	5.7777e-06	58.82
13	01-Jan-2007 13:00:00	1.3168	1.3167	1.3169	-1.8222e-04	-1.0630e-04	-9.8711e-05	-1.5188e-05	-1.3667e-04	-6.0748e-05	-1.3667e-04	-4.8640e-06	21.42
14	01-Jan-2007 14:00:00	1.3169	1.3169	1.3170	-6.0745e-05	-4.5559e-05	-3.7967e-05	6.0752e-05	2.2781e-05	-1.5187e-05	4.5564e-05	-5.4604e-06	21.42
15	01-Jan-2007 15:00:00	1.3194	1.3193	1.3194	1.2887e-04	4.5478e-05	4.5478e-05	9.8541e-05	1.3645e-04	8.3379e-05	0.0019	1.2015e-04	82.69
16	01-Jan-2007 16:00:00	1.3193	1.3192	1.3194	6.8225e-05	-1.5160e-05	-4.5478e-05	4.5482e-05	-7.5799e-06	7.5806e-05	-9.0952e-05	-7.1001e-06	67.30
17	01-Jan-2007 17:00:00	1.3190	1.3189	1.3191	-9.8548e-05	-1.0613e-04	-2.1981e-04	-1.0992e-04	-1.3266e-04	-1.6676e-04	-1.8192e-04	-3.2340e-05	36.73
18	01-Jan-2007 18:00:00	1.3194	1.3194	1.3195	5.3055e-05	-0.0011	-2.5004e-04	1.1370e-04	1.1749e-04	1.8951e-04	3.2600e-04	3.3926e-06	55.07
19	01-Jan-2007 19:00:00	1.3195	1.3194	1.3195	4.0942e-04	1.5539e-04	1.2886e-04	1.8192e-04	1.5918e-04	1.0611e-04	1.5158e-05	-4.9894e-05	86.48
20	01-Jan-2007 20:00:00	1.3194	1.3194	1.3195	-6.0629e-05	-5.3051e-05	5.1564e-04	5.8393e-04	-6.0629e-05	3.0317e-05	-3.7894e-05	6.3481e-05	33.33
21	01-Jan-2007 21:00:00	1.3192	1.3188	1.3195	-1.1369e-04	-1.6674e-04	-1.3643e-04	2.2742e-05	-1.5159e-04	-3.7901e-05	-1.8190e-04	-1.2191e-05	42.26
22	01-Jan-2007 22:00:00	1.3197	1.3196	1.3198	-9.8500e-05	1.5156e-05	2.9562e-04	2.4254e-04	6.8204e-05	1.8190e-04	3.7144e-04	1.1579e-04	44.67

01-Jan-2007 01:00:00	-2.1223e-04
01-Jan-2007 02:00:00	4.6246e-04
01-Jan-2007 03:00:00	-1.5914e-04
01-Jan-2007 04:00:00	3.7895e-05
01-Jan-2007 05:00:00	-1.1368e-05
01-Jan-2007 06:00:00	-2.0084e-04
01-Jan-2007 07:00:00	3.4112e-05
01-Jan-2007 08:00:00	-0.0018
01-Jan-2007 09:00:00	7.5935e-05
01-Jan-2007 10:00:00	-6.0743e-05
01-Jan-2007 11:00:00	6.0747e-05
01-Jan-2007 12:00:00	-1.3667e-04
01-Jan-2007 13:00:00	4.5564e-05
01-Jan-2007 14:00:00	0.0019
01-Jan-2007 15:00:00	-9.0952e-05
01-Jan-2007 16:00:00	-1.8192e-04
01-Jan-2007 17:00:00	3.2600e-04
01-Jan-2007 18:00:00	1.5158e-05
01-Jan-2007 19:00:00	-3.7894e-05
01-Jan-2007 20:00:00	-1.8190e-04
01-Jan-2007 21:00:00	3.7144e-04
01-Jan-2007 22:00:00	5.7590e-04





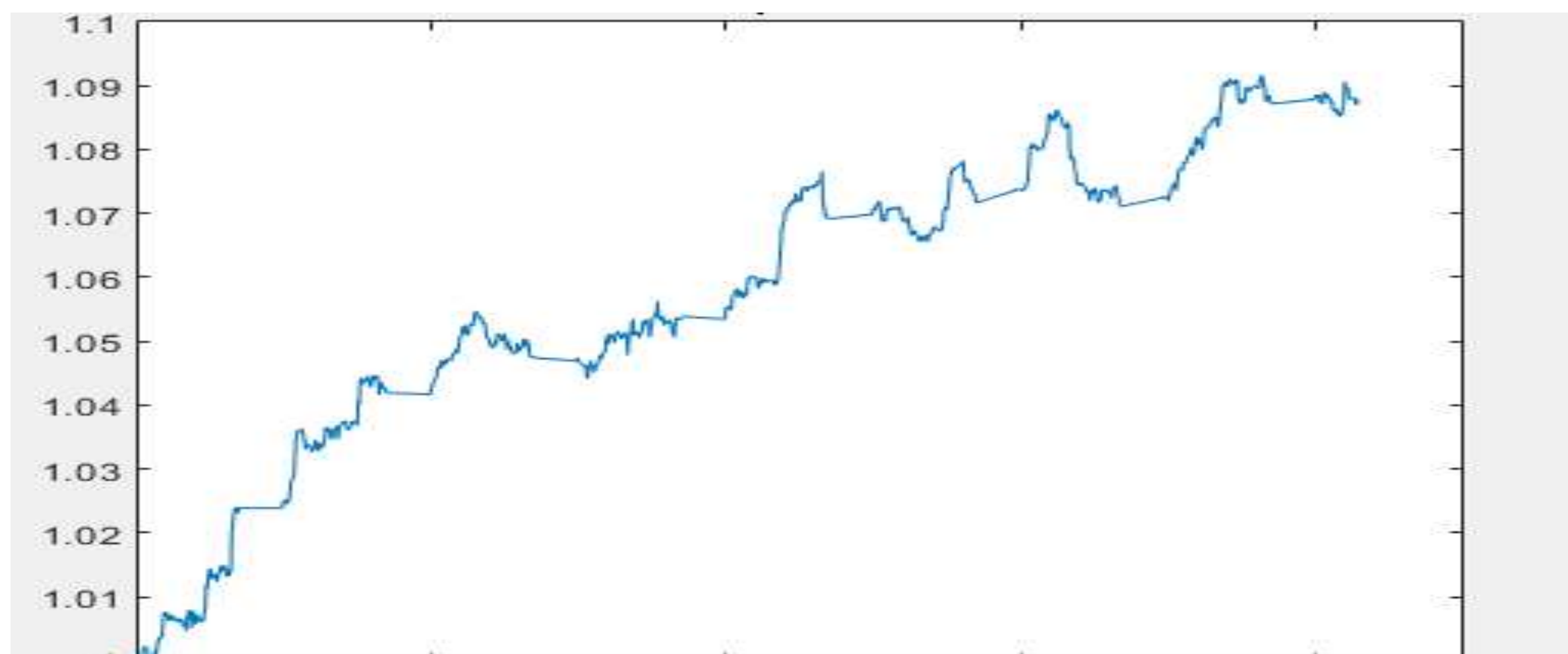
- Creiamo un **In-Sample** di dati utilizzando una parte, i 4/5 del database
- Dato il sottoinsieme di dati In-Sample facciamo Training del **Modello Lineare**  
Usiamo *fitlm* che ha come output la colonna response che è ciò che vogliamo modellare, dati tutti i nostri Fattori (o previsori)
- Avendo l'oggetto **Model Train** richiamiamo il metodo *predict* sull'In-Sample per ottenere la previsione dei rendimenti







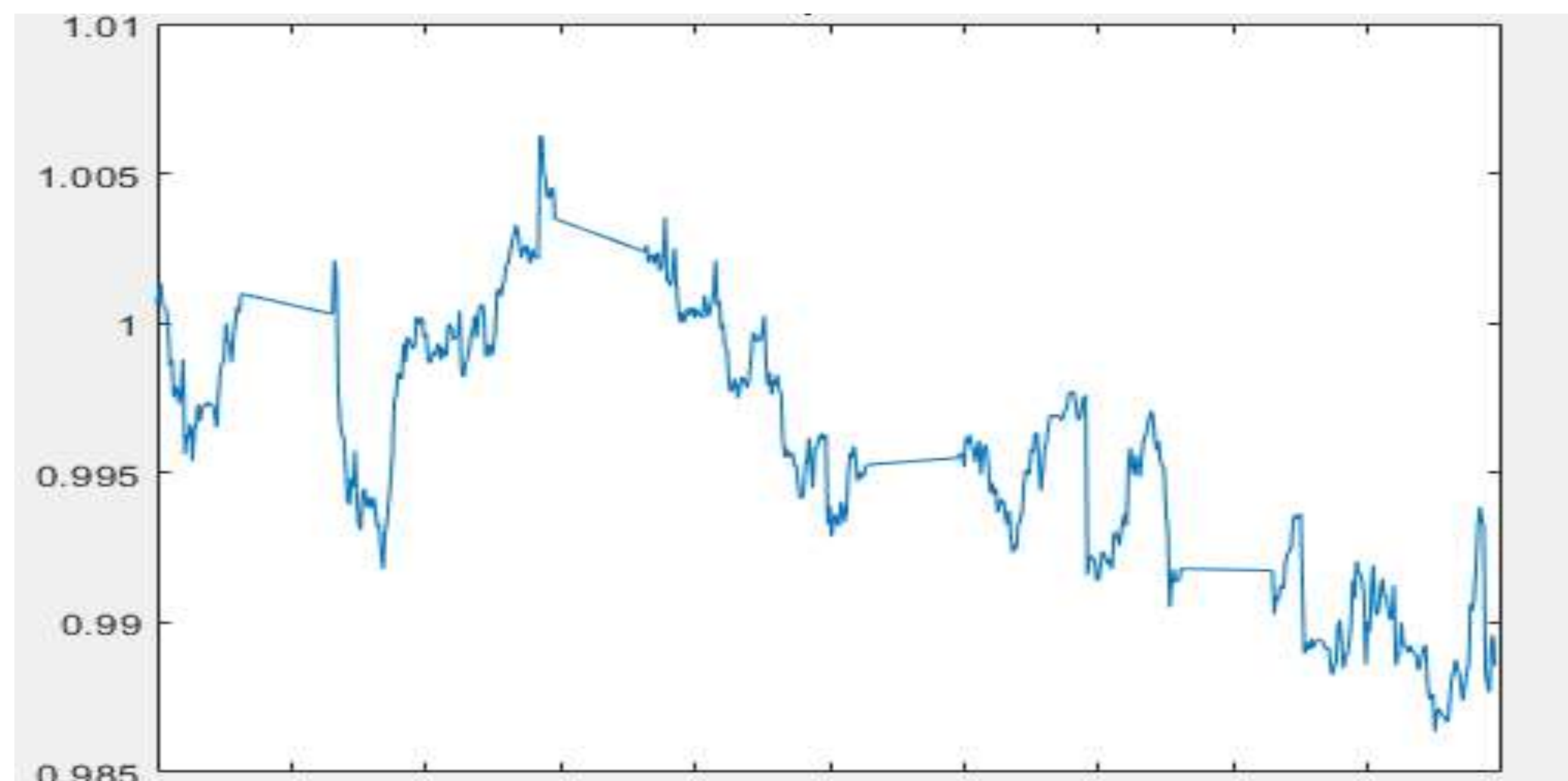
- Per vedere i risultati i termini di strategia, associamo un valore positivo (1) se i rendimenti previsti sono in rialzo e un valore negativo (-1) se sono in ribasso. Queste due colonne, moltiplicate per la variazione effettiva ci permettono di avere un'idea del P&L del modello.
- Vediamo i risultati sull'In-Sample







- Sono, ovviamente, abbastanza buoni, ma dobbiamo vedere quello che importa veramente, cioè l'Out-pf-Sample



- Non è un buon risultato, il modello lineare non è abbastanza valido, proviamo uno ***Stepwise Model***



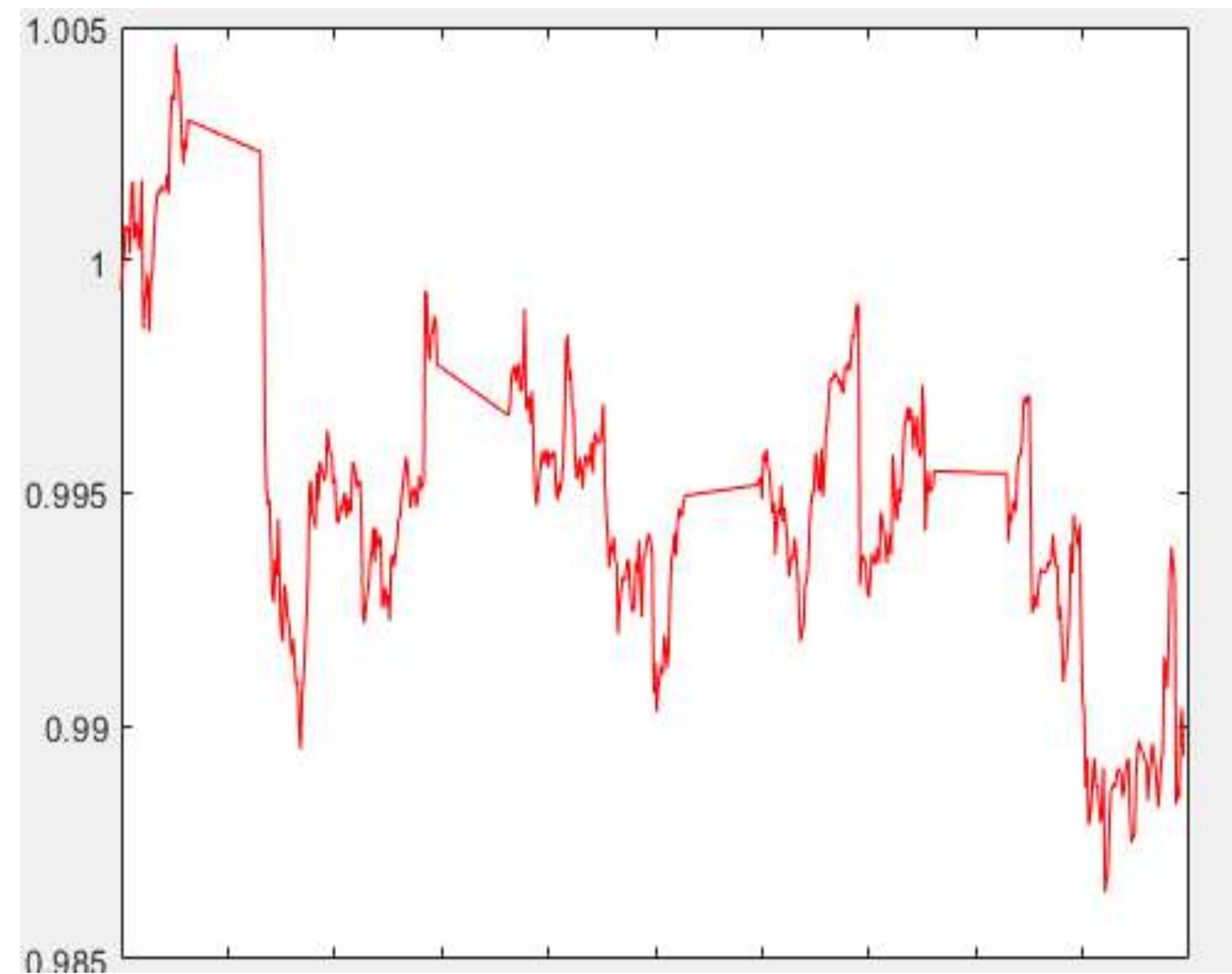
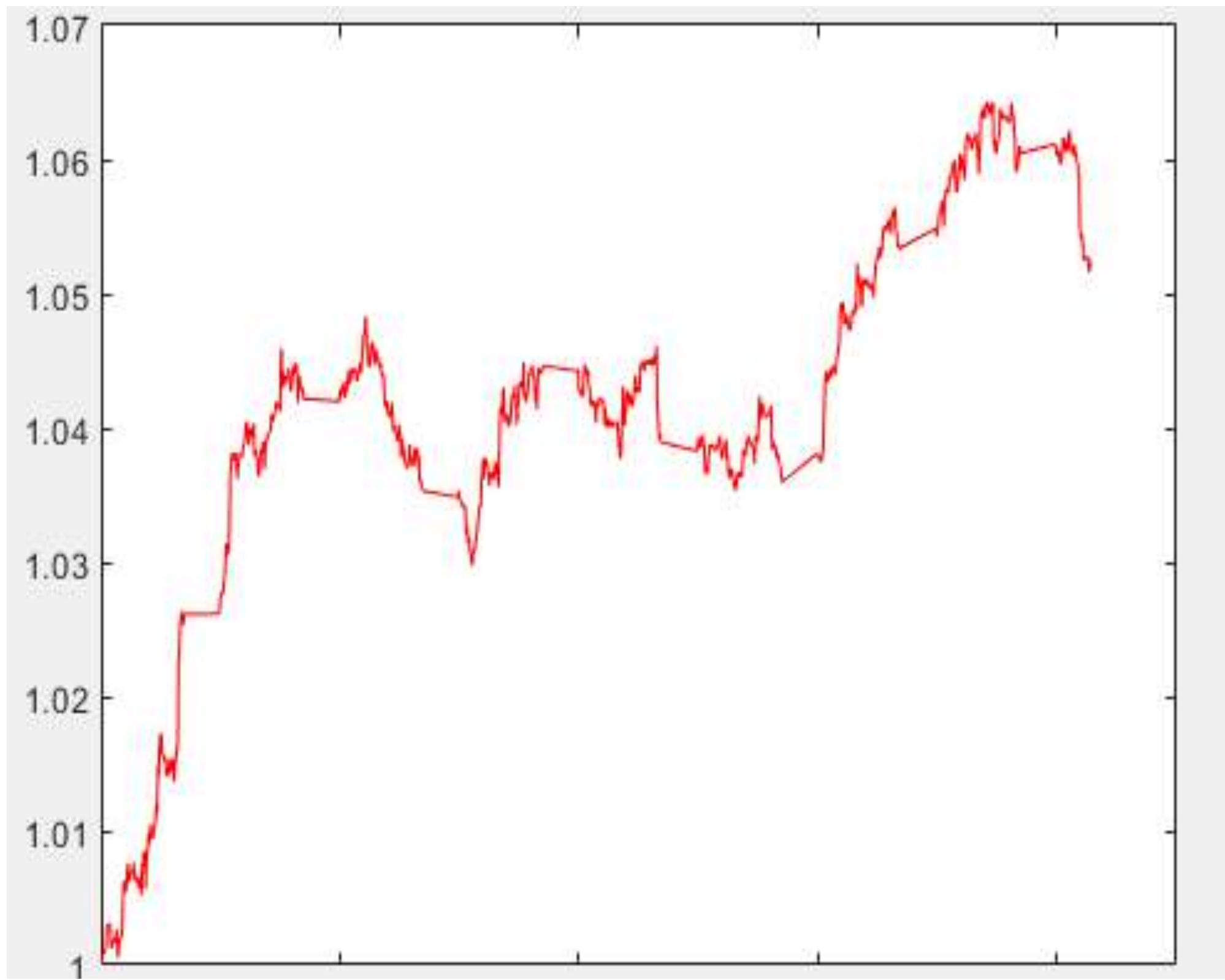


- Lo Stepwise Model rimuove i nostri Fattori (Previsori) se, al test di quanto sono efficaci, non presentano buoni risultati. Creiamo un oggetto **StepwiseModel** e , ancora, richiamiamo il metodo **predict**. Calcoliamo i risultati, come in precedenza su In-Sample e Out-of-Sample

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	0.00096483	0.00037517	2.5717	0.010263
Return_20	0.70075	0.21252	3.2973	0.0010105
Return_25	-0.3597	0.10408	-3.4559	0.00057144
rsindex_20	-1.9225e-05	7.4912e-06	-2.5664	0.010421







- Lo Stepwise Model non migliora i risultati, ci rivolgiamo quindi al **Machine Learning** , mantenendo la stessa struttura (con Fattori previsionali e Response).
- Utilizziamo Supervised Learning dove , cioè, etichettiamo i nostri dati, Prendiamo i nostri rendimenti futuri previsti e li etichettiamo come buy, sell o hold.
- Carichiamo i dati
- Creiamo una tavola con tutti i nostri Fattori (Previsori)





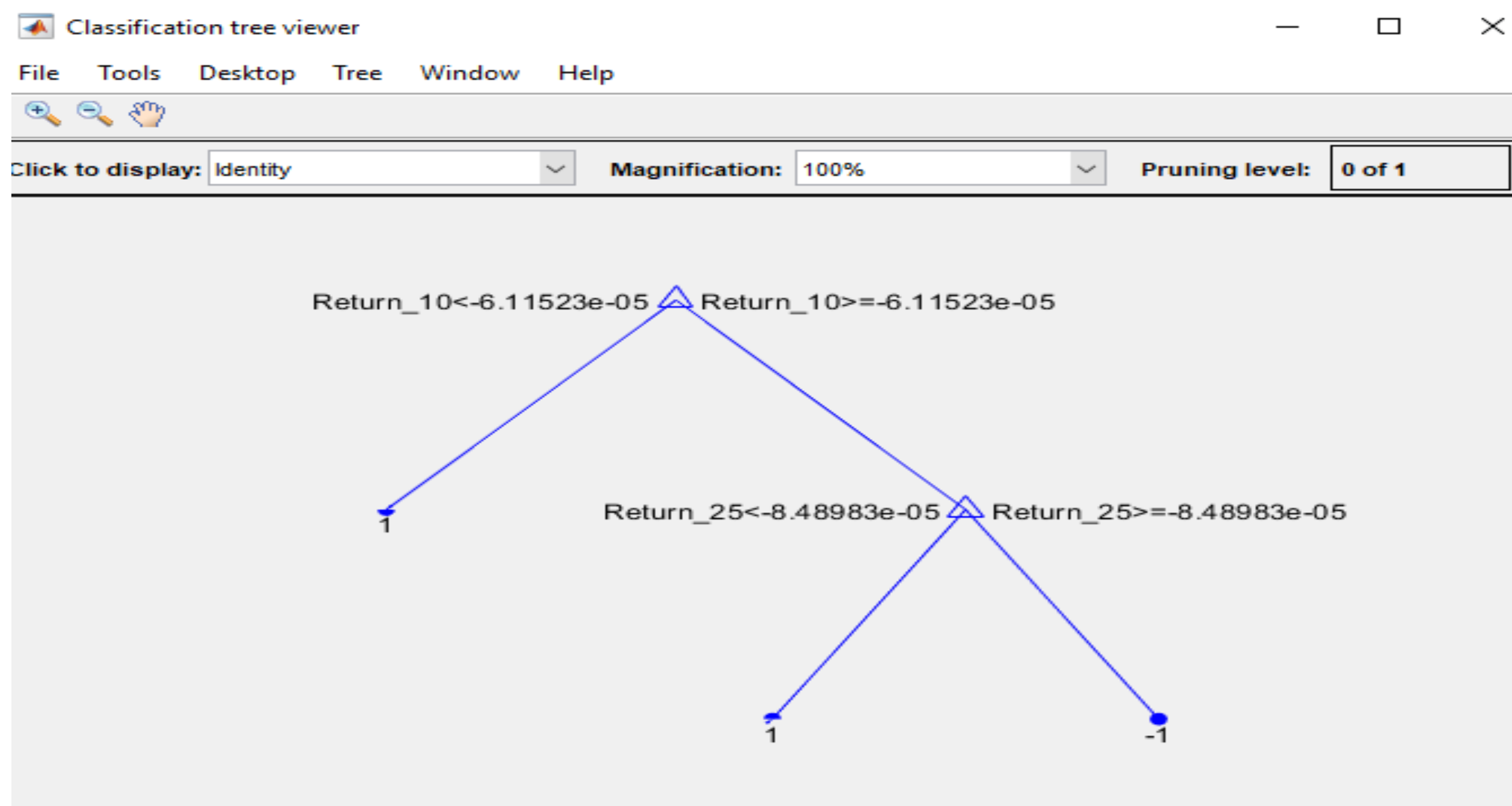


- Lo Stepwise Model non migliora i risultati, ci rivolgiamo quindi al **Machine Learning** , mantenendo la stessa struttura (con Fattori previsionali e Response).
- Utilizziamo **Supervised Learning** dove , cioè, etichettiamo i nostri dati, Prendiamo i nostri rendimenti futuri previsti e li etichettiamo come buy, sell o hold.
- Carichiamo i dati
- Creiamo una tabella con tutti i nostri Fattori (Previsori)
- Aggiungiamo una colonna Response alla tabella



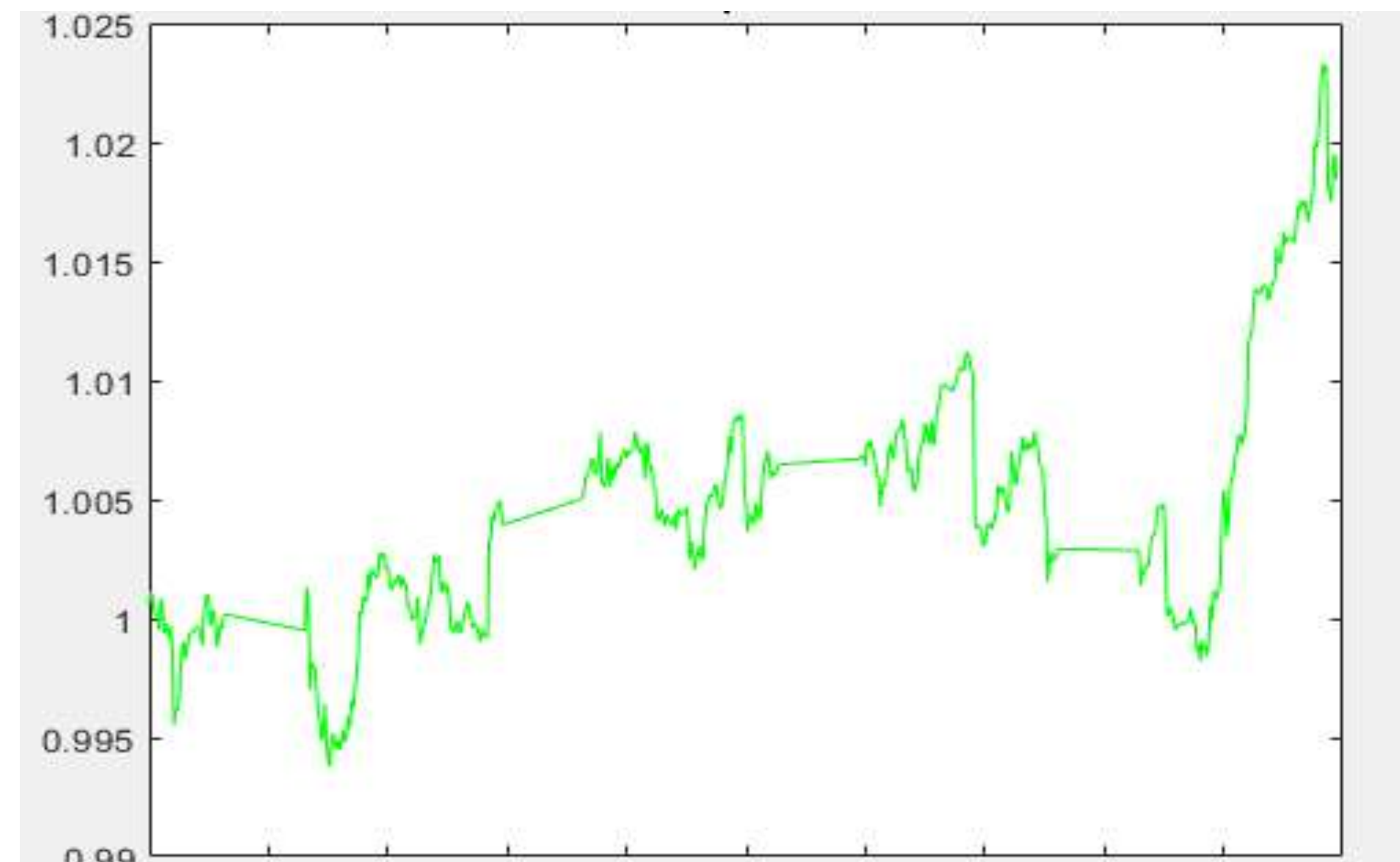
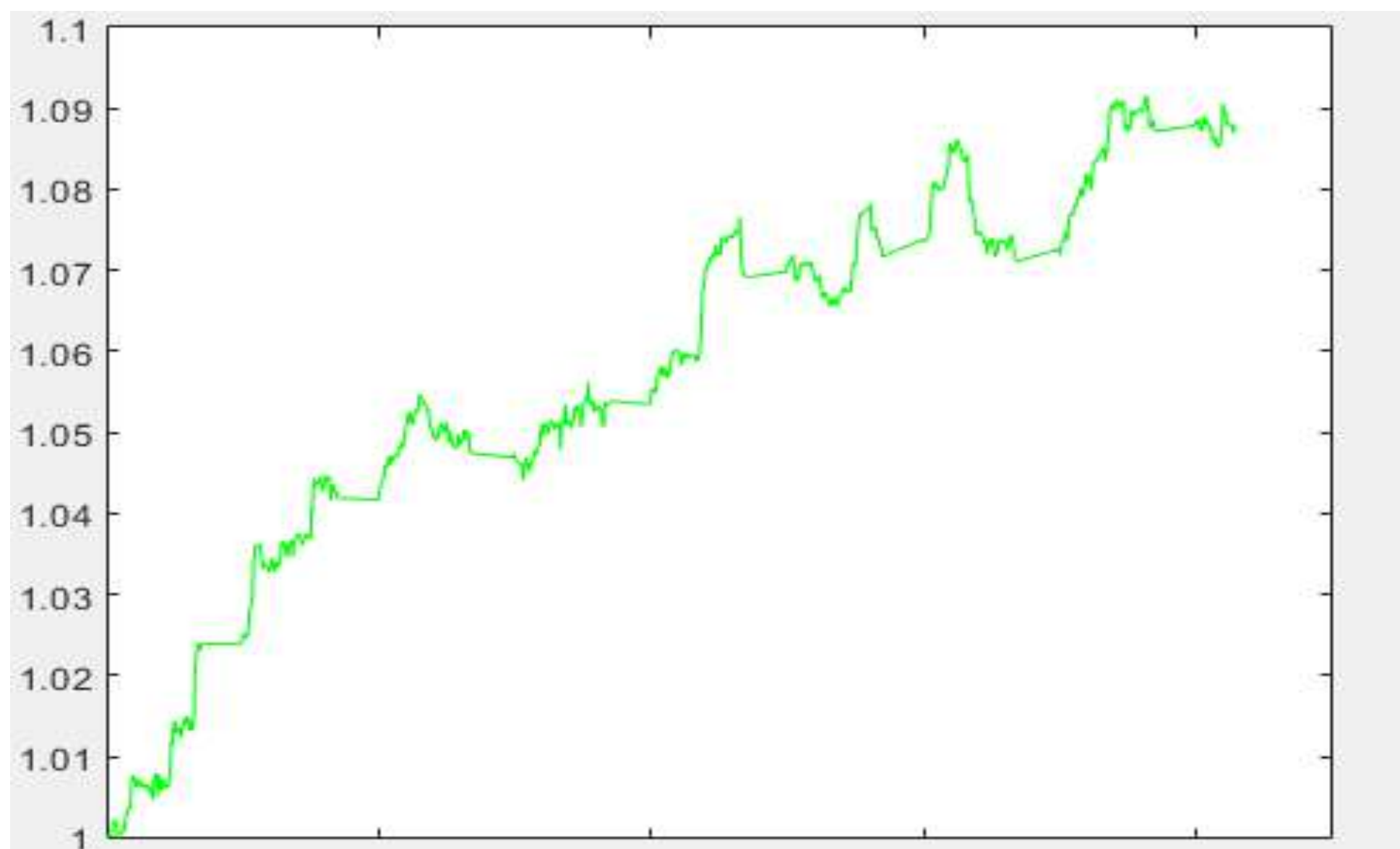


- In Matlab è possibile utilizzare l'App Classification Learner per scegliere, in maniera guidata, il miglior modello. Il miglior candidato risulta essere un modello ad albero, molto semplice, quello che vediamo qui sotto:





- Quindi *fittiamo* il modello richiamando il metodo `fictree`
- Applichiamo il metodo utilizzato in precedenza ai rendimenti previsti futuri e valutiamo In-Sample e Out-of-Sample





## Conclusioni

- Sia un modello di Regressione Lineare semplice che un modello Stepwise non producono risultati accettabili in Out-of-Sample, mentre un modello Tree di Supervised Machine Learning mostra risultati promettenti...

## Sviluppi

- Quindi a Quantlab Limited vogliamo fare questa metodologia più «nostra» e affiancare, potenzialmente, questi modelli previsionali ad altri, che già utilizziamo, come modelli Econometrici e Reti Neurali

**Grazie e a presto...**

**luigi.piva@quantlab.co.uk**

