

**Predicting Short Term
Stock Trends using
Gaussian Regression**

Summary:

Market participants have successfully used technical indicators (TI) to execute profitable trades. TI are metrics derived from price history of a financial instrument. It has been observed in the literature that profitability in 'technical stock trading' has moved from daily to intraday trades, over the past couple of decades. In this study, we will consider the problem of using Gaussian Process Regression (GPR), a non-parametric supervised learning method, to forecast financial time series as a function of one or more technical indicators, for high frequency data. We hope this will be of use in intraday technical trading.

1. Introduction:

A financial instruments time series history has the potential to influence price evolution³. This is the bases of technical trading. In this study, we aim to use Gaussian Process Regression (GPR) to forecast returns on stocks using widely used technical indicators as inputs. The basic methodology is adopted form a study by Goshal et.al.⁶ who have used, amongst other variables, technical indicators to predict future daily returns on S&P 500 index.

It has been argued recently that the profitability of technical trading has moved to intraday trades². Our study according looks at predicting intraday returns using TI. We consider commonly used TIs that include Moving Average Oscillator (MAO) and Relative Strength Index (RSI), Support and Resistance (SR) and Moving Average Convergence Divergence (MACD) for high frequency data. We use GPR on our training data using these technical indicators and assess the relevance of each one to our forecast. The 'best' indicators are then used to predict short term stock price movements.

2. Technical Indicators:

Technical Analysis aims to exploit price trends in the asset price dynamics. The underlying idea here being that the pattern of asset price dynamics is a sequence of trends, interrupted by 'whipsaws', that repeats itself across different timescales.

Commonly used technical indicators include Moving Average models, Momentum models, and Relative Strength Index models. We describe these, using terminology in Schulmeister², below.

2.1 Moving Average Oscillators: These are constructed from a short term moving average (M_s) and a long term moving average (M_L) of past prices. The length of M_s usually varies between 1 and 10 periods, and that of M_L between 10 and 30 periods.

The trading rule to be followed is: Go long when M_s crosses M_L from below and go short when the converse occurs or equivalently go long when $(M_s - M_L)$ becomes positive and go short if the converse occurs.

Based on this we define the moving average oscillator $MAO_{j,k}$ as

$$MAO_{j,k}(t) = \frac{(M_{s,j}(t) - M_{L,k}(t))}{M_{L,k}(t)} \times 100$$

2.2 Momentum Oscillators: These are constructed from the relative price difference between the current price and that from k periods in the past.

$$MO_i(t) = \frac{(P(t) - P(t - i))}{P(t - i)} \times 100$$

The basic trading rule here is: Go long when the momentum turns from negative to positive and go short if the converse occurs.

The idea behind these rules is: following trends. This is so because $M_S(P(t))$ exceeds (or falls below) $M_L(P(t - i))$, only if an upward (downward) price movement has persisted over some periods.

2.3 Trading Rules for MAO and MO

The basic trading rule is given in the leftmost graph in figure 1. As mentioned above the rule is that when the momentum turns from negative to positive, one takes a long position and when the momentum goes from negative to positive take a short position.

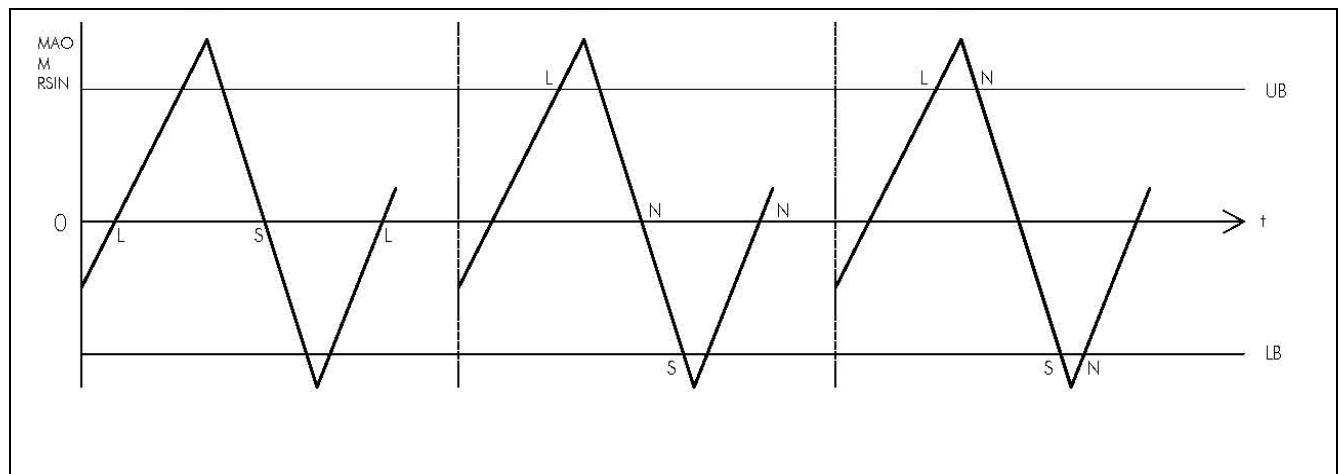


Figure 1: Trading Rules based on MAO and MO

There are variations on this rule, for example the center graph in figure 1, here one defines upper and lower bounds and the trading rule is, one opens a long (short) position whenever the oscillator crosses the upper (lower) bound from below (above). When a long (short) position is held and the oscillator crosses the zero line from above (below) the model switches to a neutral position.

2.4 Relative Strength Index:

The Relative Strength Index (RSI) is used for identifying overbought or oversold conditions. The n day RSI is defined as

$$RSI_n(t) = 100 - \left[\frac{100}{1 + \frac{up_n(t)}{dn_n(t)}} \right]$$

Where $D_i = P_{t-i+1} - P_{t-i}$ is the daily price change, and $up_n(t) = \frac{\sum D_i}{n}$ for $D_i > 0$ and $dn_n(t) = \frac{\sum D_i}{n}$ for $D_i < 0$

The size of the RSI oscillator depends not only on the overall price change but also on the degree of monotonicity of this change. If the RSI falls (rises) again below (above) a certain level (the upper/lower bound of the RSI oscillator) the situation is considered overbought (oversold).

The original RSI fluctuates between 0 and 100, to make it comparable to the moving average and momentum oscillators we consider a 'normalized' RSI which fluctuates around zero.

$$RSIN_n(t) = \frac{RSI_n - 50}{50}$$

2.5 Trading Rules for RSI:

Based on the description above, one can use the rules described graphically in figure 2. For example in the leftmost graph, the trader switches from a long (short) to a short (long) position, if the oscillator crosses the upper (lower) bound from above (below). Some variations on this basic idea are also given in figure 2. These rules can be applied to the MAO and MO as well as the RSI.

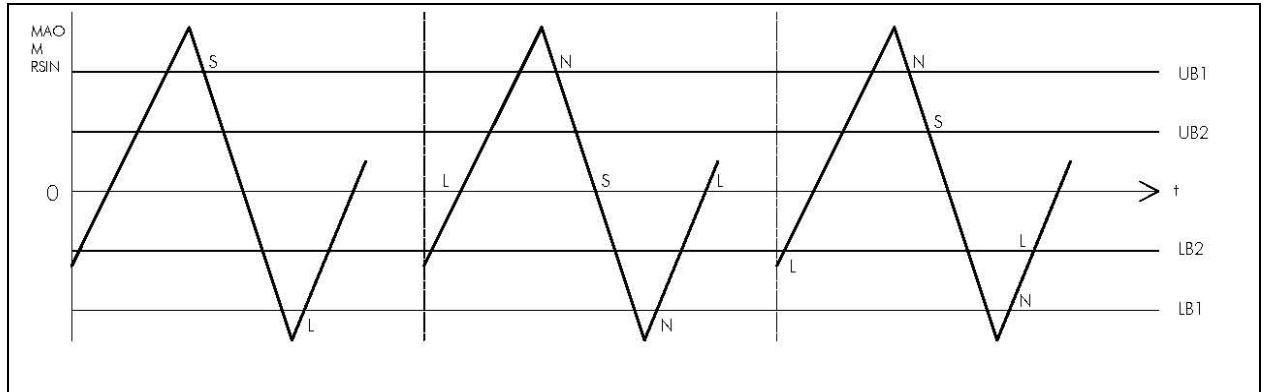


Figure 2: Trading Rules for MAO, MO and RSI

Our interest of course in the project is not on the actual trades, but in the predictive power of these indicators.

3. Gaussian Processes:

A Gaussian process is any collection of random variables such that an arbitrary subset of these has a joint Gaussian distribution¹. The process is completely specified by the mean $m(x)$ and covariance $k(x, x')$ functions i.e. $m(x) = E[f(x)]$ and $k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))]$

We can represent a Gaussian Process in the form

$$f(x) \sim GP(m(x), k(x, x'))$$

wolog, we assume mean to be zero, all results can be easily extended to non zero mean data.

For a given training data set $X = \{x_1, x_2, \dots, x_n\}$ with corresponding output variables $y = \{y_1, y_2, \dots, y_n\}$ and Gaussian Process f , the distribution of $f = \{f(x_1), f(x_2), \dots, f(x_n)\}$ will be multivariate Gaussian $f \sim N(0, K)$ where $K_{ij} = k(x_i, x_j)$. Conditional on f we have a Gaussian observation model given by $y_i | f \sim N(0, \sigma^2)$. We can marginalize out f to get $y_i \sim N(0, K + \sigma^2 I)$.

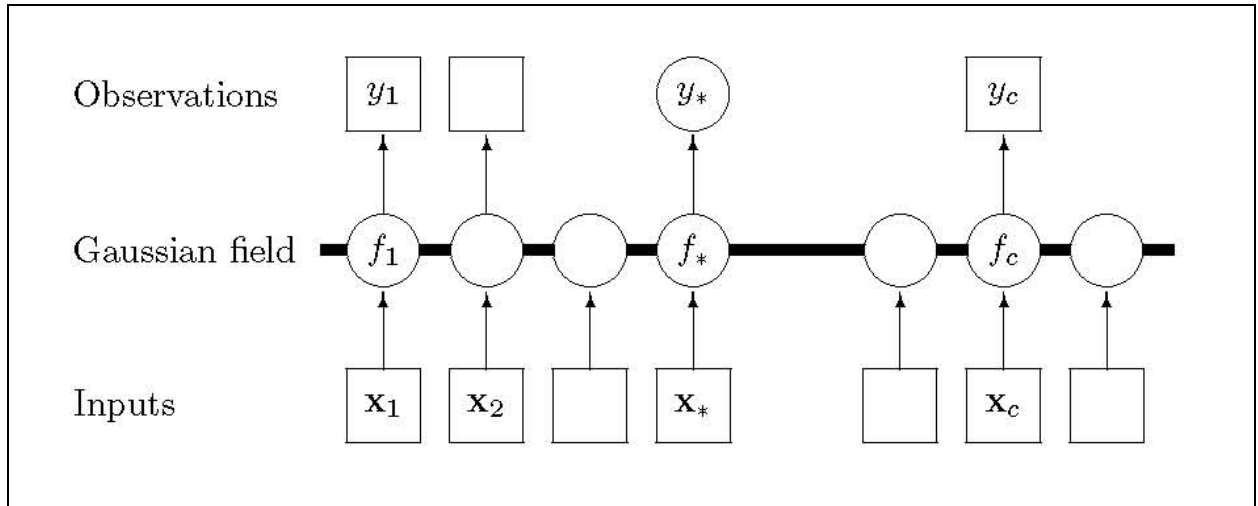


Figure 3: Gaussian Process Regression

Conditioning on the training data gives the following predictive distribution y for a new data point x^*

$$y^* | x^*, X, y \sim N(k^*(K + \sigma^2 I)^{-1} y, k^{**} - k^*(K + \sigma^2 I)^{-1} k^{*T})$$

Here $K_{ij} = k(x_i, x_j)$, $k^* = [k(x_1, x^*), k(x_2, x^*), \dots, k(x_n, x^*)]$ and $k^{**} = k(x^*, x^*)$

What this does in essence is to take information from a 'prior' encoded in the covariance matrix $k(x, x')$, combine it with the observation data to produce a posterior distribution for forecasting.

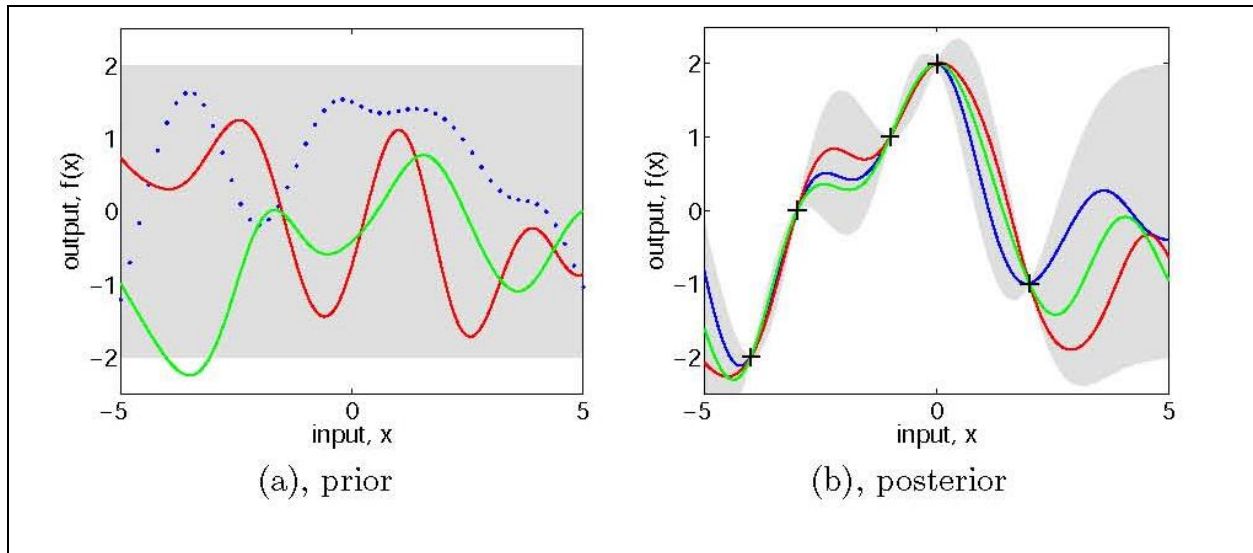


Figure 4: Learning the distribution via GPR

3.1 Covariance Kernels:

There are several choices for the appropriate covariance kernel^{1,8}, we will try three different kernels to see which gives the best fit of the training data.

The first is the squared exponential covariance $K(x, x') = \sigma_f^2 e^{-\left(\frac{(x-x')^2}{2l^2}\right)}$, where σ_f and l are the hyperparameters, do be determined from the data.

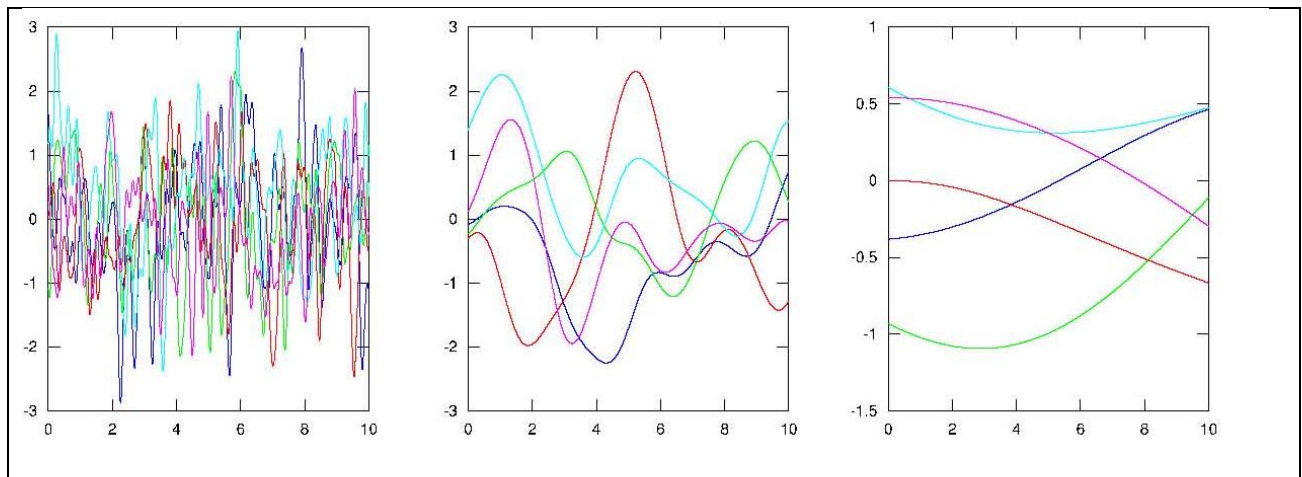


Figure 5: Varying 'length parameter' l for a SE Kernel with $\sigma_f = 1$

The second is a Matern Class covariance $K(x, x') = \left(1 + \frac{\sqrt{3}(x-x')}{l}\right) e^{-\frac{\sqrt{3}(x-x')}{l}}$

The third covariance kernel we consider is the rational quadratic kernel, $K(x, x') = \left(1 + \frac{(x-x')^2}{2\alpha l^2}\right)^{-\alpha}$

3.2 Hyperparameters:

Given a covariance kernel, we would need to obtain the hyperparameters from the data. This is referred to as 'training' of the model. To accomplish this we will maximize the marginal likelihood

$$\frac{\partial}{\partial \theta_j} \log(p(\underline{y}|\underline{X}, \underline{\theta})) = \frac{1}{2} \text{tr} \left((\underline{\alpha} \underline{\alpha} - K^{-1}) \frac{\partial K}{\partial \theta_j} \right)$$

Here $\underline{\alpha} = K^{-1} \underline{y}$.

3.3 Multiple Inputs and Outputs:

The formalism described above can easily be extended to the general case multiple inputs and multiple outputs⁸. The simplest approach is to take the covariance function to be the product of one dimensional covariance functions over each input. Let x^e be the e th element in the input space and x_i^e be the value of the e th element at the index point i . The covariance is then given by

$$k(x_i, x_j) = \prod_e k^e(x_i^{(e)}, x_j^{(e)})$$

Where k^e is the covariance function over the e th input. We also define the distance function

$$d(x_i, x_j, \Sigma) = \sqrt{(x_i - x_j)^T \Sigma^{-1} (x_i - x_j)}$$

Here Σ represents hyperparameters of the model. For example a diagonal form of Σ simply provides an individual length scale $l^e = \sqrt{\Sigma(e, e)}$ in the distance metric. By allowing off diagonal entries, one allows for correlations amongst the input dimensions.

As mentioned earlier the length scale is a measure of how distant points in input space need to be for them to be substantially uncorrelated. Such a covariance function implements automatic relevance determination (ARD), as the inverse of the length determines the importance of a particular input.

Previous Work:

Technicals driven GPR has been used in time series forecasting for a wide variety of asset classes including stock prices, stock volatility and commodity spreads^{6,7}. There has also been work on financial prediction using text data⁵. Most relevant to our proposed project are recent papers by Farrell et.al.⁷ and Goshal et.al.⁶

Farrell et.al use GPR with two different inputs (based on daily stock returns) to predict next day returns. Their first input considers a simple metric based on stock price trend (whether the price goes up or

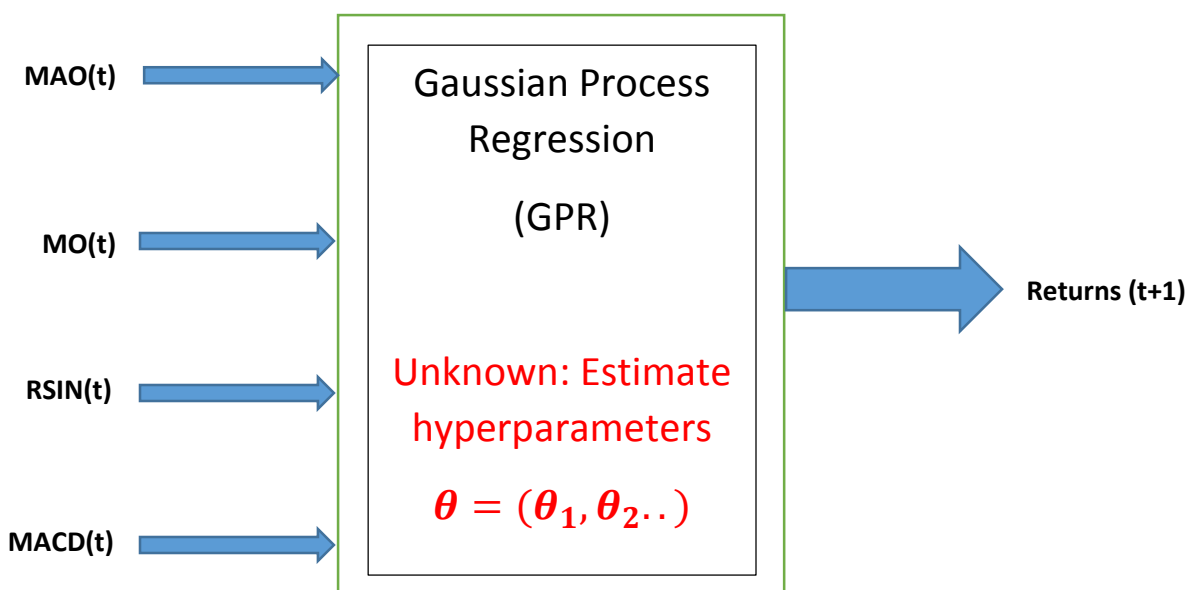
down). The second input is simply the value of stock at close on the previous day. Their results show that quality of prediction is heavily dependent on the size of the training set.

Goshal et.al provide a systematic framework for predicting next day returns using GPR. They consider different classes of inputs, including Technical Indicators, metrics derived from Market Sentiment, Options based modelling of Price Space and metrics derived from Broker Recommendations. They ascertain the importance of the different metrics by first considering the correlation of each of these with the next day returns. The metrics showing significant correlations in each of the above described 'groups' are then used in a GPR to determine feature relevance within each group. Finally, the most relevant inputs are then used for predictive purposes and the quality of predictions tested.

Schulmeister in his study has shown that using commonly used technical indicators, one can execute profitable trades. However, this profitability has steadily declined over the past few decades. In fact he shows that profitable technical trading has shifted from models based on daily data to those based on higher frequency data. He considers MAO, MO and RSI in his analysis.

Proposed Framework:

Our proposed work is based on the framework established by Goshal et.al. We will first test various available TI for significant correlation with returns, for the training data. This will be done for four classes of TI: MAO, MA, RSIN and MACD. Once the most important Tis have been determined based on this analysis, a GPR will be run on the test data and the TIs having most 'relevance' as defined below ascertained. Finally, GPR will be used, with these as inputs for predictive purposes. This will be done in two steps, each of the selected TI will be used individually, and then various combinations of the TIs will be used and the results evaluated.



We now describe this process in detail.

Step 1: Choosing the ‘best’ TIs

- We consider 4 different classes of oscillators. The MAO, given by $MAO_{j,k}(t) = \frac{(M_{s,j}(t) - M_{L,k}(t))}{M_{L,k}(t)} \times 100$, the MO, given by, $MO_i(t) = \frac{(P(t) - P(t-i))}{P(t-i)} \times 100$, the RSI given by, $RSI_n(t) = 100 - \left[\frac{100}{1 + \frac{up_n(t)}{dn_n(t)}} \right]$ and the MACD. Note there is a wide variety of these oscillators based on the different periods.
- For the MAO and MACD we will consider M_s between 1 and 10 periods and M_L between 10 and 30 periods, with the restriction that the lengths of M_s and M_L differ by at least 5 periods.
- For the RSIN and MO, we consider periods between 3 and 30 steps.
- An upper bound of 0.3 and a lower bound of -0.3 is chosen for the RSIN signal.
- Correlation of each one of these are calculated with the next step return, for data obtained at 1 sec, 5 sec and 30 sec intervals.
- Significant correlation will be used to determine the TI to be used, one from each class.

Step 2: Determining Feature Relevance

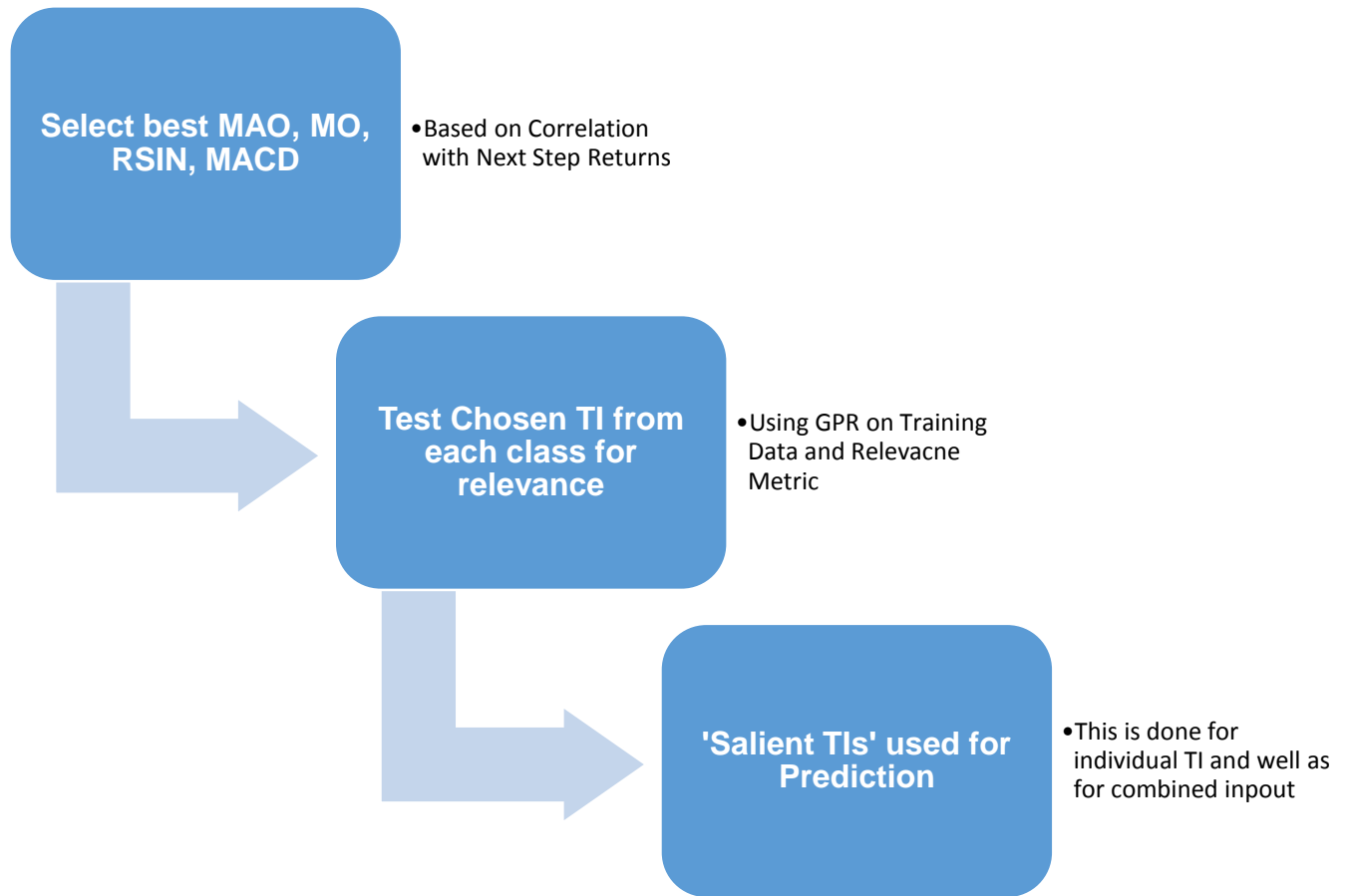
- Feature relevance for each TI chosen as above will be determined using an ARD GPR.
- The relevance will be determined based on relevance ratio

$$\text{Relevance Ratio}_i = \frac{\text{Relevance Score}_i}{\text{Relevance Score}_{\text{noise}}}$$

- Here, $\text{Relevance Score}_i = l_i^{-1}$. This has been discussed in the hyperparameters section above.

Step 3: Evaluating Model Performance

- TI from each class will be used to predict the next step returns
- This will be tested by finding the correlation against the actual value, Mean Absolute Deviation (MAD) and Normalized Root Mean Square Error (NRMSE) of the two.
- The TIs from different classes will be used as a combined input for GPR and the predictions tested again as above



References:

1. **Rasmussen C.E. and Williams K.I.**, Gaussian Processes for Machine Learning, MIT Press (2005)
2. **Schulmeister S.**, The Profitability of Technical Stock Trading has Moved from Daily to Intraday Data, WIFO Working Paper No. 289 (2010)
3. **Park C.O. and Irwin S.H.**, What do we know about the Profitability of Technical Analysis?, J. Econ Surveys, 21: 786-826 (2007)
4. **Scholtus M.L. and van Dijk D.**, High Frequency Technical Trading: The Importance of Speed, Tinbergen Institute Discussion Paper 12-018/4 (2012)
5. **Wuthrich B. et.al.** Daily Stock Market Forecast from Textual Web Data, IEEE Intl. Conf. on Systems, Man and Cybernetics (1998)
6. **Goshal S. and Roberts S.**, Extracting Predictive Information from Heterogeneous Data Streams using Gaussian Processes, Alg. Finance, vol 5, no. 1-2, pp. 21-30 (2016)
7. **Ferrell T. and Correa A.**, Gaussian Process Regression Models for Predicting Stock Trends, MIT Univ. Technical Report (2007)
8. **Roberts S. et.al.**, Gaussian Processes for Timeseries Modelling, Phil Trans. R Soc. A, (2012)