A Novel Seismic Wave Analysis and Prediction Model using SVM Kernel and Time Series Prediction

Yongyang Zheng
Nanchang University, Nanchang, China

Abstract — Earthquakes and Volcanos are very important and often fatal phenomena taking place around the globe on an unpredictable time scale which makes it more dangerous. Earthquakes occur due to collision of tectonic plates deep underground which send out vibrations of great magnitude through the layers of the earth. These vibrations are termed as seismic waves which are an important class of signals for analysis of their magnitude, frequency which also consequently help in prediction of occurrence and study of deep seismic activity. A novel SVM based kernel model has been proposed in this research paper for predicting seismic activity on a time scale. This method is computationally efficient and a frequency domain method has been used for analysis of the frequency content. The experimentations have been done on sample seismic activity observation inputs and results justify the efficiency of the proposed approach.

Keywords — Seismic waves, prediction models, time series prediction, SVM kernel

I. INTRODUCTION

Earthquakes have been prevalent over the decades bringing catastrophic devastation throughout the globe. The level of devastation is directly influence by the magnitude of the earthquake that has occurred at the epicenter. Earthquakes are mainly due to change in position of tectonic plates deep beneath the earth’s crust. Off late, there has been seismic activity on a continual basis at tectonic points identified throughout the globe. Histories of seismic activity have become more prevalent and of greater magnitudes right from 1890s to till date. A recent earthquake that hit Haiti had a magnitude of nearly 7.0 on the Richter scale costing over 200 thousand human lives and 400 thousand displaced people. This has provided the necessary motivation and need to predict the occurrence of earthquake in advance so as to prevent mass scale of destruction. An advance determination of intense seismic activity could prove to be very essential so as to mobilize the evacuationary procedures so as to save precious human lives.

Numerous research has been carried out to develop and implement intelligent algorithms to predict the occurrence of seismic activities based on seismic recordings obtained from vulnerable zones throughout the globe. Geographers are continuously stationed at permanent points on the globe identified to be seismic prone activities based on history and intensity of occurrences to monitor the movement and nature of the tectonic plates underneath. A sample seismic wave is shown in figure 1 indicating the amplitude versus the time plot.

The proposed research paper deals with analysis and hence prediction of seismic activity based on the inputs obtained by analyzing the seismic wave. A seismic wave is strongly correlated with three key factors namely intensity, spectrum and duration. Two categories of analysis are defined for extraction and study of seismic waves in the form of static and dynamic activities. A static system studies the amplitude and frequency of the input signal to predict the damage that could be caused due to the seismic activity while dynamic analysis takes into consideration the entire three dimensional quantities which includes intensity of spectrum, duration and spectral information. There are several methodologies that could be adopted to study the frequency domain content most of which are in the frequency domain. Most common techniques include the well known as Fourier transform, Sine and Cosine transforms, Wavelet family transforms etc., The temporal frequency response is described as:

\[ S_d = \left( a, a, t \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t) \sin(\omega t) e^{-j\omega t} \, d\omega \]

(1)

\[ S_r = \left( a, a, t \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t) \cos(\omega t) e^{-j\omega t} \, d\omega \]

(2)
The above three equations denote the displacement spectrum, velocity and acceleration spectrum from the given seismic signal.

\[
S_2 = (\eta, \omega, \omega) = \frac{1}{\alpha} \int_0^\infty \gamma(t) \sin [\omega (t - \omega)] d\omega
\]

Figure 2 depicts a typical acceleration spectrum from the recorded seismic signal wave. Earlier methods found in the literature make use of temporal frequency response spectrum analysis (TRFS) scheme to effectively obtain the depth details of the given seismic signal under study.

A prediction model utilizing a feed forward neural network [5] has been implemented in the literature where a multilayer perceptron was trained using back propagation algorithm. The activation function for the neural network is a sigmoid type. Mean squared error and standard deviation have been taken as the output parameters and compared with Markov based models and 81% accuracy has been reported. In Findings of the literature also depict models using a hybrid model [9] comprising of the ANN and ARIMA (Auto regressive integrated moving average) [16] [18] which has been implemented and the output of ARIMA have been used to train the neural feed. The hybrid model is observed to perform better than the linear autoregressive and random walk models and produced similar results to that of the ANN. A new fuzzy time series model [7] [13] to improve forecasting have been presented where membership functions of varying degrees are used to relate the fuzzy outputs to the crisp outputs. These fuzzy relationships are then used to forecast the stock index in Taiwan. This study performs out of sample for recasting and the results are compared with those of previous studies to demonstrate the performance of the proposed model.

Standard Back propagation (SBP), Scaled Conjugate Gradient (SCG) and Back propagation with Bayesian Regularization (BPR) to predict probabilities [14] [7] of occurrences have been seen in the literature for application of forex time series prediction against Australian dollar and their evaluation results have been elaborated in [11]. A feed forward multilayer perceptron (MLP) [6] model by using the historic records regarding the stock share has been implemented in order to predict a company’s stock value and compared it with Elman Recurrent Network and Regression Model. Results indicate that MLP has lower MSE, MAPE, and MAE values in comparison with Elman [4] [6] [9] and linear regression [10] [8] whereas the Elman recurrent neural network outperforms the multilayer perceptron in predicting the direction of changes. Time series forecasting has been carried out using two ANN techniques [14], multilayer perceptron (MLP) and Volterra. On the other hand, the findings [1] report that hidden Markov Models are unstable for trading tool on foreign exchange data as the results are dependent on too many factors. Although the multilayer perceptron (MLP) neural network is used widely in forecasting systems, it has the drawback of being time consuming and not being able to restore the memory of past events. To improve the past forecasting models, a hybrid forecasting model [18] that refines past models and optimizes the Elman Recurrent Neural Network (Elman NN) for predicting the Taiwan stock price trends has been implemented. The proposed model outperforms the other listed models due to nonlinear prediction capabilities, faster convergence, and accurate mapping ability. A new neural network based hybrid model [2] is developed in which a predefined number of simulated data series is generated and they are tuned by an efficient particle swarm optimization algorithm. Using two datasets of real Euro/ Dollar rates, how the proposed hybrid model could reasonably enhance the results of GARCH-type models [5] and the traditional neural network in terms of different performance measures is demonstrated. Illustrations on how the respective simulated data series as the input variable to the network could contribute to improve the accuracy of volatility forecasting are presented in the paper.

Another hybrid neural network architecture [3] of Particle Swarm Optimization and Adaptive Radial Basis Function (ARBF–PSO), has been implemented for financial forecasting purposes by benchmarking the ARBF–PSO results with those of three different neural networks architectures, a Nearest Neighbors algorithm (k-NN), an autoregressive moving average model (ARMA), a moving average convergence/divergence model (MACD) plus a naive strategy. More specifically, the trading and statistical performance of all models is investigated in a forecast simulation of the varying currencies in the globe, rainfall prediction, weather prediction etc.

II. PROPOSED WORK

Support vector machines are a very important and quite essential statistical tool for deriving a set of functions from the given class with an objective function of optimizing and minimizing the risk function. The objective function is defined as:

\[
\theta(f) = \frac{1}{2} \int M(\varphi \cdot f(x))dQ(\varphi)\] (3)

The classical feature of SVM kernel is that it minimizes the risk function by reducing the normalized and regularized risk function. The regularized risk is nothing but the weighted sum of empirical risk with respect to the data. SVMs are an important class of predictors since they
determine a linear decision $y = f(x)$ thus producing a generalized performance. The minimization of kernel function is described as:

$$\min (\alpha, \sigma, \epsilon^*) = \frac{1}{2} \sum \alpha^i \alpha^j y^i y^j K(x^i, x^j) + C \sum \epsilon^i$$

(4)

The proposed technique utilizes kernels which extend the solution of decision to non-linear problems. This linearization is achieved by mapping the input $A$ into the feature space as defined by the general mapping:

$$A \mapsto \Phi(A)$$

(5)

could be defined in feature space such that $S^f(x) = x$ where $S$ is the contraction constant in the feature space to maintaining the controllability of the system. If $S$ is a closed subset of feature space $Y$ and $A \mapsto \Phi$ defined as a contraction in the subspace then a unique fixed point function could be defined in the Banach space as

$$|S^f(x) - x| \leq \frac{2}{1 - k} |S(x) - x|$$

(6)

Here we define for a continuous system that

$$\|S(y + v) - S(y)\| = \|S(S(y + v), y + v) - S(S(y), y + v)\|$$

$$= \|S(y + v) - S(y)\| + S(S(y + v), y + v) - S(S(y), y + v)$$

(7)

Simplifying the above inequality it could be concluded that

$$|S(y + v) - S(y)\| \leq \frac{2}{1 - k} |S(y + v) - S(y)\|$$

(8)

Extending the above equation for a continuous solution we can prove that

$$S(y + v) - S(y) = \gamma(x)$$

(9)

If $P, Q$ are spaces defined in Banach domain and $X, Y$ are subsets of $P$ and $Q$ respectively then

$$P \subseteq A_X(X \times Y, \gamma)$$

(10)

Utilizing the shift operation, $A \mapsto A - A(\gamma, q)$ where $A(\gamma, q)$ is taken to be zero and hence the solution converges to the equation

$$P(x, y) = y - \zeta S(\xi, y)^{-1}$$

(11)

The Fouries analysis of the predicted time series is depicted in figure 3.

III. RESULTS AND DISCUSSION

The experimentation has been done in MATLAB for a SISO system modelling a Wiener filter to remove the noise from the input signal. The processor is an Intel 3 2.5 GHz processor with a 4GB RAM capacity. Figure 2 illustrates the nonlinear system modelled in SIMULINK environment with known input parameters to each of the blocks. The output signal is viewed across the scope. Separability of SVM kernel in the proposed work has been depicted in figure 4 shown below. A kernel basically is a mathematical operator which accepts two inputs and outputs the correlation between the two inputs in time domain. The frequency spectra filtered through the Wiener filter in the proposed work has been given to the SVM kernel which establishes the relationship between the data sets provided to the network.

SVM is a classical tool that helps to solve classification and prediction models and establishes a single minima in time series or polynomial time. It has an important property of convex optimization as well as duality.
Mean absolute error is given as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |x_i - \hat{x}_i|$$  \hspace{1cm} (3)

The MSE on the other hand specifies the difference between the estimator and what is estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. Figure 5 depicts the observed mean absolute error values against the epoch number. 500 epochs have been used for convergence of the error. The difference occurs because of randomness or because the estimator doesn’t account for information that could produce a more accurate estimate.

And MSE is given as

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$  \hspace{1cm} (4)

Figure 5 illustrates the convergence of error where the point of stability is achieved. A near to convergence is achieved at about 20 – 25s which is a satisfactory convergence rate with respect to non-linear systems.

IV. CONCLUSION

This paper presents a novel SVM kernel based prediction in time series for a seismic activity detection model. Spectral signal obtained from Fourier analysis of the input signal is used as the source of input to the SVM kernel to get the time prediction series for accurate prediction of intense seismic activity. An important advantage of using SVM kernel is that it maximizes the minimum distance between the data sets to the corresponding separators. It also aids in efficient feature expansion thus reducing the convergence time and computational complexity of the proposed technique when compared to conventional techniques. The advantage of SVM kernels lie in the fact that they are robust to noise and provide high classification and prediction accuracy.

REFERENCES


