THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Enhancing Stock Price Prediction Using Support Vector Machine Approach

Nbaakee, Lebari Goodday
Student, Department of Computer Science,
Ignatius Ajuru University of Education, Port Harcourt, Nigeria
Kabari Giok Ledisi
Lecturer, Department Computer Science,
Kenule Polytechnic Bori, Rivers State, Nigeria

Abstract:
The unpredictability of the capital market has throughout the year's involved extraordinary worry to scientists and financial specialists in the market who are wary on the following move as components past simple financial matters of scale can cause an unexpected upsurge or sharp drop in the cost of stock either making an open door for an increase or misfortune. The hazard factor is eccentric in this way the requirement for the usage of AI way to deal with stock value forecast. This paper sends bolster vector machines an AI way to deal with the expectation of stock value that has been less executed in time arrangement investigation and forecast with stunning outcomes.

Keywords: *Prediction, Stock Market and SVM*

1. Introduction

Due to the unconstrained and non-direct nature of money related time arrangement information in the capital market, precise forecast of costs of budgetary instruments which is fundamental for each player in the market to have the option to settle on better venture choices with least hazard has become a delusion. These complexities result from the immense volume of exchange and information handled day by day and various different components extending from financial and political likewise pulls a string available making an increasingly unpredictable and insecure situation. Speculators in the capital market are cautious regarding which organization stock to purchase or sell at normal interims as the unpredictability increments on a regular schedule. The ceaseless development of this exceptionally fluctuating and sporadic information has advanced the basic requirement for the emanant of a mechanized methodology for the proficient investigation of such gigantic budgetary information to separate significant insights from them.

Stock value expectation and anticipating has in a late time accepted the front burner as the development and solidness of the financial exchange assumes a significant job in monetary advancement and development of a country. Thusly, it has gotten a matter of genuine worries to speculators and investors in the market to guarantee exact and legitimate gauges of stock costs to have the option to pull in high rewards and gainful preferences as mistaken and inconsistent forecasts can have hopeless results. In this manner, it is basic to give a precise and effective model for stock expectations [5]. Stock value gauging approaches and investigation are partitioned into three classes that incorporate basic examination, specialized examination and innovative investigation separately [7].
The paper presents a stock prediction system using support vector machine

2.1. Aim and Objectives

This paper intends to build up an AI way to deal with the securities exchange forecasts utilizing Support Vector Machine. The objectives are:

1. Develop an SVM model to break down and anticipate the stock cost
2. Training the model utilizing dataset from the Nigeria Stock Exchange
3. Test the Model precision
4. Predict future exchange the financial exchange

2. Literature Review

The securities exchange forecast has become an inexorably significant issue right now. One of the techniques utilized is a specialized investigation; however, such strategies don't generally yield precise outcomes. So it is imperative to create strategies for an increasingly exact expectation. For the most part, speculations are made utilizing forecasts that are gotten from the stock cost in the wake of considering every one of the variables that may influence it. The procedure that was utilized in this example was a relapse. Since money related stock imprints produce tremendous measures of information at some random time an extraordinary volume of information needs to experience investigation before an expectation can be made. Every one of the systems recorded under relapse has its own focal points and impediments over its different partners. One of the critical strategies that were referenced was a straight relapse. The manner in which direct relapse models work is that they are regularly fitted utilizing the least-squares approach, however, they may, on the other hand, be additionally be fitted in different manners, for example, by reducing the "absence of fit" in some other standard, or by lessening an impaired adaptation of the least-squares misfortune work. Then again, the least-squares approach can be used to fit the nonlinear model [6].

2.1. Machine Learning for Stock Prediction

AI (ML) "learns" a model from past information so as to anticipate future information [1]. The key procedure is realizing which one of the computerized brains is. AI includes diverse measurable, probabilistic, prescient and advancement systems that can be executed as the learning strategies, for example, the calculated relapse, counterfeit neural systems (ANN), K-closest neighbour (KNN), choice trees (DT) and Naive Bayes. There are significant learning calculations in ML that are managed learning and solo learning. Regulated learning manufactures a model by gaining from known classes (named preparing information). Interestingly, solo learning techniques take in the normal highlights from obscure class information (unlabelled preparing information) [2]. Utilization of AI strategies in securities exchanges expectation and examination has demonstrated to be of incredible supplement and improvement to the conventional techniques with an upgrade in the forecast of real-time exchanging and high hazard decrease making improving speculators trust in the market.

2.2. Support Vector Machine Learning for Stock Prediction

Bolster Vector Machine is an amazing AI strategy for building a classifier planned for making a choice edge between two classes that encourages the forecast of marks from at least one component vectors [3]. This choice edge otherwise called the hyperplane is situated so that it is beyond what many would consider possible from the nearest information focuses from every one of the classes called bolster vectors.

Given a labeled training dataset:

$(x_1, y_1), ..., (x_n, y_n), x_i \in R^d$ and $y_i \in (-1, +1)$

where x_i is a feature vector representation and y_i the class label (negative or positive) of a training compound i.

The optimal hyperplane can then be defined as:

$wx^T + b=0$

where w is the weight vector, x is the input feature vector, and b is the bias.

The w and b would satisfy the following inequalities for all elements of the training set:

$wx_i^T + b \geq +1$ if $y_i=1$

$wx_i^T + b \leq -1$ if $y_i=-1$

The objective of training an SVM model is to find the w and b so that the hyperplane separates the data and maximizes the margin $1 / || w ||^2$.

Vectors x_i for which $|y_i|$ $(wx_i^T + b)= 1$ will be termed support vector

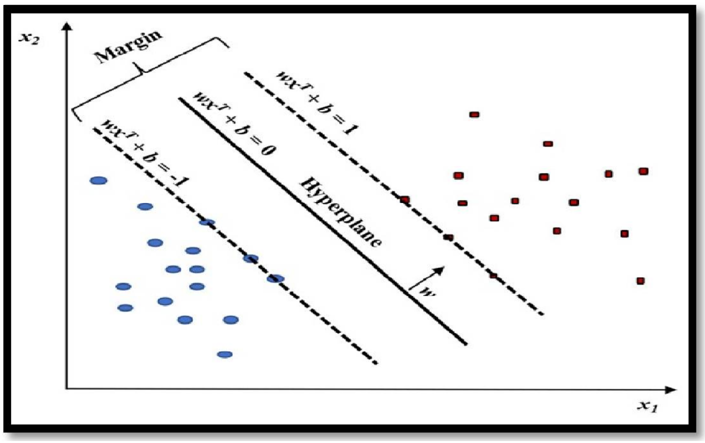


Figure 1: Linear SVM Model. Two Classes (Red Versus Blue) Were Classified [2]

[4] proposed different AI strategies for stock forecast with accentuation to stock longevity in "Securities exchange Prediction Using Machine Learning Algorithms" and found that systems like irregular timberland RF and support vector machine SVM have not been completely abused by analyst for financial exchange expectation contrasted with different time series for stock forecast. In implementing the same concerning different techniques it was found that SVM shows a significant level of exactness in regard to stock longevity and speculators certainty working with an approval precision of 89.43 and a blunder margin 12.29 individually. They opined that the effective forecast of the stock utilizing RF and SVM will be an incredible resource for the financial exchange foundations and will give genuine answers for the issues that stock speculators face.

3. Materials and Method

The AI strategy conveyed in this paper is the help vector machine. SVM is a discriminative classifier that uses administered learning with marked information. The yield is hyperplanes that gatherings the new dataset. The tuning parameters of the SVM classifier are piece parameter, gamma parameter, and regularization parameter.

- Kernels can be arranged as direct and polynomial portions ascertains the forecast line. In direct portions expectation for another information is determined by the dab item between the information and the help vector.
- C parameter is known as the regularization parameter; it decides if the exactness of the model is increments or diminishes. The default estimation of c=10. Lower regularization esteem prompts misclassification.
- Gamma parameter quantifies the impact of a solitary preparation on the model. Low qualities mean a long way from the conceivable edge and high qualities connotes closeness from the conceivable edge. This are shown in figure 2.

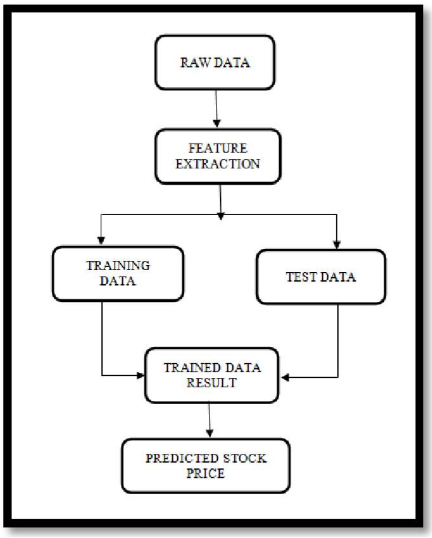


Figure 2: SVM Stock Prediction Architecture Using

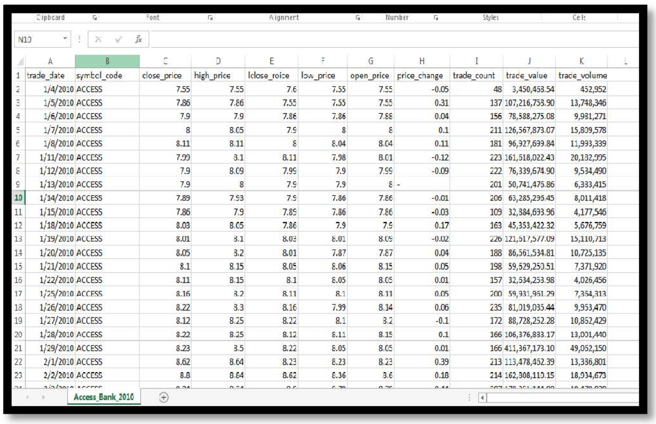
In the model engineering in Figure 2, the model gets crude information and concentrate highlight from them. This information is additionally isolated into preparing test information which is utilized for making the last expectation of stock. The dataset utilized was sourced from the Nigerian Stock trade information demonstrating the exchange on from 2000 to 2011 of all organizations recorded. The dataset was pre-prepared and the supply of Access banks was picked for this investigation.

M18

	A	B	C	D	E	F	G	H	I	J	K	L
1	200	trade_date	symbol_code	close_pri	high_pric	lclose_rc	low_pric	open_pri	price_cha	trade_co	trade_value	trade_volume
2	1	2/1/2000	ASHAKACEM	22.5	8.8	22.09	8.8	8.8	0.41	1	352,000.00	40,000
3	2	2/1/2000	AVONCROWN	0.68	1.1	0.68	1.1	1.1	-	2	2,475.00	2,250
4	3	2/1/2000	BERGER	2.75	3.1	2.75	3.1	3.1	-	3	20,770.00	6,700
5	4	2/1/2000	CADBURY	26.6	16.27	26.1	16	16.27	0.5	15	1,642,871.99	101,150
6	5	2/1/2000	CAP	2.7	2	2.7	2	2	-	1	600	300
7	6	2/1/2000	CCNN	5.98	1.7	5.98	1.7	1.7	-	2	313,607.50	184,475
8	7	2/1/2000	CORNERST	2.01	1.25	1.96	1.25	1.25	0.05	1	12,750.00	10,200
9	8	2/1/2000	CRUSADER	2.7	1.05	2.7	1.05	1.05	-	1	673.05	641
10	9	2/1/2000	CUTIX	2.51	1.59	2.51	1.59	1.59	-	2	4,771.59	3,001
11	10	2/1/2000	DUNLOP	3.25	4.21	3.27	4.2	4.21	-0.02	9	213,980.00	50,900
12	11	2/1/2000	ETERNAOIL	1.1	0.76	1.1	0.76	0.76	-	5	50,160.00	66,000
13	12	2/1/2000	FIRSTBANK	23.5	13.99	23.47	13.55	13.55	0.03	31	1,203,425.30	87,848
14	13	2/1/2000	FLOURMILL	14.61	8.91	14.61	8.91	8.91	-	1	14,416.38	1,618
15	14	2/1/2000	GUARANTY	6.2	2.21	6.2	2.2	2.2	-	10	786,820.00	357,600
16	15	2/1/2000	GUINNESS	36	19.05	35.96	19.05	19.05	0.04	4	274,129.50	14,390
17	16	2/1/2000	IPWA	1.2	1.2	1.2	1.2	1.2	-	1	600	500
18	17	2/1/2000	JBERGER	44	17.84	44	17.84	17.84	-	1	35,680.00	2,000
19	18	2/1/2000	JOSBREW	0.52	0.53	0.53	0.52	0.53	-0.01	2	106,650.00	205,000
20	19	2/1/2000	LAWUNION	1.38	1.56	1.38	1.56	1.56	-	1	1,560.00	1,000
21	20	2/1/2000	MOBIL	70	52.87	70	52.87	52.87	-	3	176,057.10	3,330
22	21	2/1/2000	MORISON	3.27	3.22	3.27	3.22	3.22	-	1	6,440.00	2,000
23	22	2/1/2000	NB	30.2	19.83	30.18	19.08	19.81	0.02	15	1,357,382.62	69,073
24	23	2/1/2000	NBC	22	12	22	11.8	12	-	15	480,750.40	40,292
25	24	2/1/2000	NCR	2.2	1.95	2.2	1.95	1.95	-	1	1,950.00	1,000
26	25	2/1/2000	NEM	0.52	0.5	0.52	0.5	0.5	-	2	5,429.50	10,859
27	26	2/1/2000	NESTLE	60	26.79	60	26.79	26.79	-	10	179,707.32	6,708
28	27	2/1/2000	NIG-GERMAN	2.6	5.07	2.6	5.07	5.07	-	1	2,535.00	500

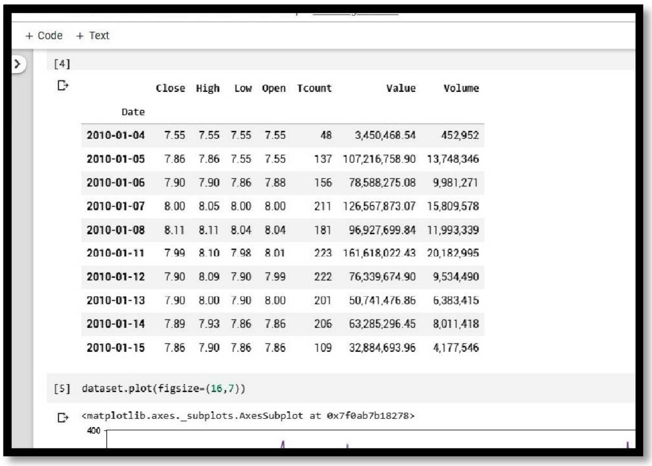
NSE_Stock_Data_2000_2011

Figure 3: NSE Stock Prices from 2000 to 2011
(Source: NSE, 2011)



	trade_date	symbol_code	close_price	high_price	lclose_price	low_price	open_price	price_change	trade_count	trade_value	trade_volume
2	1/4/2010	ACCESS	7.55	7.55	7.6	7.55	7.55	-0.05	48	3,450,458.54	452,952
3	1/5/2010	ACCESS	7.86	7.86	7.55	7.55	7.55	0.31	137	107,216,758.90	13,748,346
4	1/6/2010	ACCESS	7.9	7.9	7.86	7.86	7.88	0.04	156	78,588,275.08	9,981,271
5	1/7/2010	ACCESS	8	8.05	7.9	8	8	0.1	211	126,567,873.07	15,809,578
6	1/8/2010	ACCESS	8.11	8.11	8	8.04	8.04	0.11	181	96,927,699.84	11,993,339
7	1/11/2010	ACCESS	7.99	8.1	8.11	7.98	8.01	-0.12	223	161,518,022.43	20,182,995
8	1/12/2010	ACCESS	7.9	8.09	7.99	7.9	7.99	-0.09	222	76,339,674.90	9,534,490
9	1/13/2010	ACCESS	7.9	8	7.9	7.9	8	-	201	50,741,475.86	6,383,415
10	1/14/2010	ACCESS	7.89	7.93	7.9	7.86	7.86	-0.01	206	63,285,295.45	8,011,418
11	1/15/2010	ACCESS	7.86	7.9	7.89	7.86	7.86	-0.03	109	32,884,693.96	4,177,546
12	1/18/2010	ACCESS	8.03	8.05	7.86	7.9	7.9	0.17	163	45,353,422.32	5,676,759
13	1/19/2010	ACCESS	8.01	8.1	8.03	8.01	8.09	-0.02	226	121,617,577.09	15,110,713
14	1/20/2010	ACCESS	8.05	8.2	8.01	7.87	7.87	0.04	188	86,561,534.81	10,725,135
15	1/21/2010	ACCESS	8.1	8.15	8.05	8.06	8.15	0.05	198	59,629,250.51	7,371,920
16	1/22/2010	ACCESS	8.11	8.15	8.1	8.05	8.05	0.01	157	32,634,253.98	4,026,456
17	1/25/2010	ACCESS	8.16	8.2	8.11	8.1	8.11	0.05	200	59,931,951.29	7,364,313
18	1/26/2010	ACCESS	8.22	8.3	8.16	7.99	8.14	0.06	235	81,019,035.44	9,953,470
19	1/27/2010	ACCESS	8.12	8.25	8.22	8.1	8.2	-0.1	172	88,728,252.28	10,852,429
20	1/28/2010	ACCESS	8.22	8.25	8.12	8.11	8.15	0.1	166	106,376,883.17	13,001,440
21	1/29/2010	ACCESS	8.23	8.5	8.22	8.05	8.05	0.01	166	411,367,173.10	49,052,150
22	2/1/2010	ACCESS	8.62	8.64	8.23	8.23	8.23	0.39	213	113,478,452.39	13,336,801
23	2/2/2010	ACCESS	8.8	8.84	8.62	8.36	8.6	0.18	214	162,308,110.15	18,934,673

Figure 4: Filtered 2010 Access Bank Stock Dataset



Date	Close	High	Low	Open	Tcount	Value	Volume
2010-01-04	7.55	7.55	7.55	7.55	48	3,450,468.54	452,952
2010-01-05	7.86	7.86	7.55	7.55	137	107,216,758.90	13,748,346
2010-01-06	7.90	7.90	7.86	7.88	156	78,588,275.08	9,981,271
2010-01-07	8.00	8.05	8.00	8.00	211	126,567,873.07	15,809,578
2010-01-08	8.11	8.11	8.04	8.04	181	96,927,699.84	11,993,339
2010-01-11	7.99	8.10	7.98	8.01	223	161,618,022.43	20,182,995
2010-01-12	7.90	8.09	7.90	7.99	222	76,339,674.90	9,534,490
2010-01-13	7.90	8.00	7.90	8.00	201	50,741,476.86	6,383,415
2010-01-14	7.89	7.93	7.86	7.86	206	63,285,296.45	8,011,418
2010-01-15	7.86	7.90	7.86	7.86	109	32,884,693.96	4,177,546

```
[5] dataset.plot(figsize=(16,7))
<matplotlib.axes._subplots.AxesSubplot at 0x7f0ab7b18278>
```

Figure 5; Pre-processed Dataset

4. Results and Discussion

The traits of the dataset sourced from the CVS record from the Nigerian Stock Exchange utilized for this examination and incorporates the "HIGH" which portrays the most elevated worth the stock had in earlier year, "LOW" is a remarkable in opposition to "HIGH" and looks like the least worth the stock had in earlier year, "OPEN" the estimation of the stock at the absolute starting point of the exchanging day, and "CLOSE" represents the cost at which the stock is esteemed before the exchanging day closes. There are different qualities such as TCOUNT, VOLUME, and VALUE; however, the previous assumed a significant job in our discoveries.

The stock value forecast was re-enacted with the help vector machine in python with high preparation and approval exactness. The precision after 50 ages was 89.55 and the preparation misfortune 23.32 individually, in this manner the execution of the stock value expectation with the help vector machine could be viewed as another AI way to deal with a stock forecast like LSTM and the RNN. Bolster vector machine has been a magnificent classifier and can likewise work in a prescient situation productively in this manner filling in as prescient instrument that gives the premonition required by speculators and key players in the financial exchange to anticipate and make an okay venture in spite of the multifaceted nature.

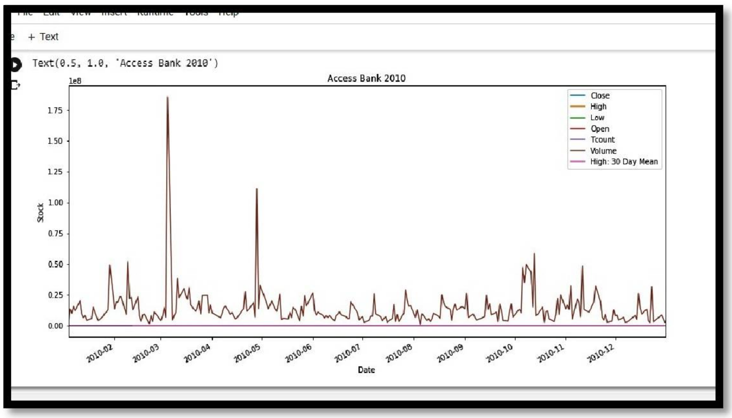


Figure 5: Access Bank Stock Price Signals

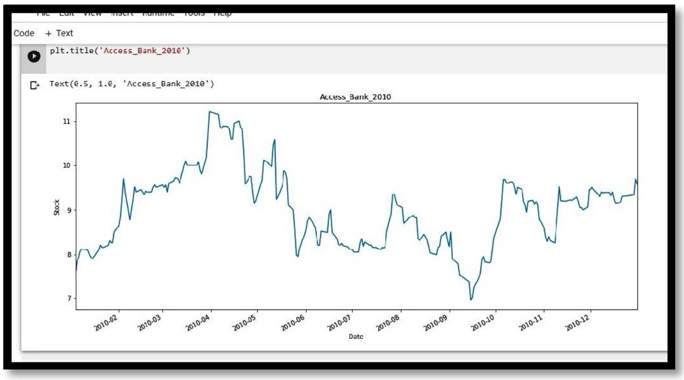


Figure 6: Access Bank Stock Showing Low

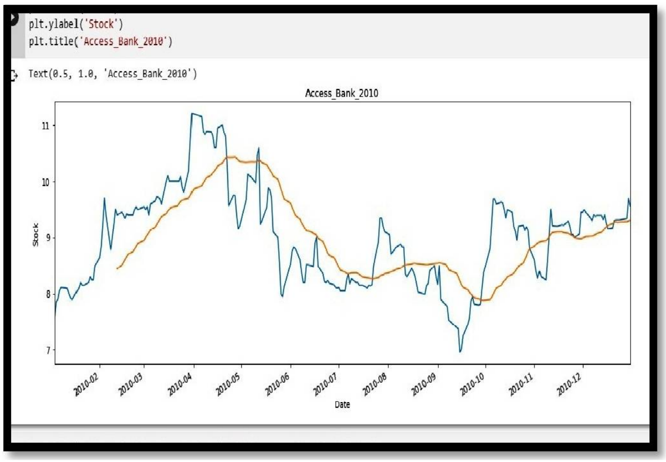


Figure 7: Access Bank Stock Showing Predicted price

5. Conclusion

By estimating the exactness of the model, we found that the help vector machine is one of the most reasonable methodologies for foreseeing stock cost. SVM will be an incredible resource for dealers and financial specialists to diminish hazards and settle on educated venture choice as the unpredictable idea of the market can be followed and predicated utilizing this AI calculation.

6. References

i. Cruz, JA and Wishart, DS. (2006). Applications of machine learning in cancer prediction and prognosis. Cancer Inform, 2, 59–77.
ii. Huang, S; Cai, N; Pacheco, P.P; Narades, S; Wang, Y and Xu, W. (2018). Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics, 15(1), 41–51.
iii. Noble, W. (2006). What is a support vector machine. Nat biotechnol, 24(12), 1557-1565.
iv. Sadia,kH; Sharma,A; Paul, A; SarmisthaPadh, and Sanyal, S. (2019). Stock Market Prediction Using Machine Learning Algorithms. International Journal of Engineering and Advanced Technology (IJEAT), 8(4), 25-31.
v. Sedighi, M ;Jahangirnia, H;Gharakhani, M and Fard, S.F. (2019). A Novel Hybrid Model for Stock Price Forecasting Based on Metaheuristics and Support Vector Machine. MDPI Journal, 4(75), 2-28.
vi. Sharma, A; Bhuriya, D and Singh, U. (2017). Survey of Stock Market Prediction Using Machine Learning Approach. ICECA.
vii. Tsang, P.M.; Kwok, P.; Choy, S.O.; Kwan, R.; Ng, S.C.; Mak, J.; Tsang, J.; Koong, K. and Wong, T.L. (2007). Design and implementation of NN5 for Hong Kong stock price forecasting. Journal of Engineering and Applied Artificial Intelligence, 20, 453-461.