International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:9, No:5, 2015

A Study of Behavioral Phenomena Using ANN

Yudhajit Datta

Abstract—Behavioral aspects of experience such as will power are rarely subjected to quantitative study owing to the numerous complexities involved. Will is a phenomenon that has puzzled humanity for a long time. It is a belief that will power of an individual affects the success achieved by them in life. It is also thought that a person endowed with great will power can overcome even the most crippling setbacks in life while a person with a weak will cannot make the most of life even the greatest assets. This study is an attempt to subject the phenomena of will to the test of an artificial neural network through a computational model. The claim being tested is that will power of an individual largely determines success achieved in life. It is proposed that data pertaining to success of individuals be obtained from an experiment and the phenomenon of will be incorporated into the model, through data generated recursively using a relation between will and success characteristic to the model. An artificial neural network trained using part of the data, could subsequently be used to make predictions regarding data points in the rest of the model. The procedure would be tried for different models and the model where the networks predictions are found to be in greatest agreement with the data would be selected; and used for studying the relation between success and will.

Keywords—Will Power, Success, ANN, Time Series Prediction, Sliding Window, Computational Model, Behavioral Phenomena

I. Introduction

THE current understanding of psychology tells us that the two major factors governing success are motivation and will [6]. As per neuro-biological studies, the origin of motivation is partially attributed to the neuro-transmitter dopamine, which is thought to be responsible for the instinct of target setting and achieving in humans [1]. Any amount of success achieved causes the release of dopamine in the body, causing the drive for further success [2]. It has been scientifically established that low levels of dopamine make people and other animals less likely to work for things [7]. The second factor, i.e. will power is less understood. The model in this study is based on the assumption that will power and the dopamine response of an individual together mould the success story of an individual. The success of an individual in turn affects an individuals will. In other words, success is a function of will power and will power is a function of success. The model is characterized by the function through which will is related to success. Success on the other hand is assumed to be linearly dependent on four primary factors, namely: the individuals will power; the detachment level of an individual towards success, accounting for the predominance of the effect dopamine has on them; the individuals recent history of success, reflecting the amount of dopamine currently present in their system; and lastly random factors. The study proposes an experiment on one hundred candidates to obtain data pertaining to the model. The story of success and failure of an individual, dubbed the *Life Story* is taken to be a sequence of events. The data points corresponding to life events in the model are determined after performing multiple linear transformations on the data set obtained from the experiment. Once the model is arrived at on the basis of the available data, the final objective is to get an artificial neural network to solve a Time Series Prediction problem. The expected outcome from this step is that the network will succeed in making predictions that agree with the data. The experiment will be repeated several times for different models. After the best-fit model has been selected and verified to be correct, the plots obtained from the model will be used to study the relation between success and will for each of the candidates.

Y. Datta is a student at the Department of Electrical and Electronics Engineering, National Institute of Technology, Calicut, Kerala, 673601 India e-mail: yudhajit.datta@gmail.com.

II. Theory

A. Artificial Neural Network

In the feed-forward architecture of the Neural Model each neuron k is characterized by a transfer function f_k, such that:

$$Output_k = f_k(w_{ik}x_k) \tag{1}$$

From a mathematical point of view, a Neural Network is a function $f : R^N \rightarrow R^M$ Where the function f is defined as the composition of some other function g_i :

$$f = (g_n) * (g_{n1}) * ... * (g_1) \tag{2}$$

Therefore a neural network defines a function f_w where w is the vector of weights. The idea is to find the best approximator of a function in the space defined by :

$$C = \{f_{w1,w2,...,wn}\}_{w\epsilon(R+)^n} \tag{3}$$

Where n is the total number of weights [3].The fundamental steps involved in creating a neural net are:

1) Variable Selection
2) Data collection
3) Data processing
4) Partitioning Data set into Training, Testing and Validation set
5) Setting Neural network parameters:
   a) Number of hidden layers
   b) Number of hidden neurons
   c) Number of output neurons
   d) Transfer functions
6) Selecting Evaluation Criteria
7) Setting Neural Network training parameters:
   a) Number of training iterations
   b) Learning rate and momentum
8) Implementation

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:9, No:5, 2015

Fig. 1. Ratio in which Data Set is partitioned into Training Set, Testing Set and Validation Set

The testing set size ranges from 10 % to 30 % of the training set. To prevent risk of over-fitting, the size of the training set must be at least five times the number of weights. The hidden layers provide the network with its ability to generalize. In theory one layer is enough to approximate any continuous function. Increasing the number of hidden layers, increases the risk of over-fitting and increases computation time [3].For a three-layer network, it has been suggested that the hidden layer should have hidden neurons approximately equal to:

$$n_{hidden} = (n_{input} * n_{output})^{1/2} \tag{4}$$

B. Time Series Prediction Problem

A time series is a sequence of vectors $x(t), t = 0, 1, ...$ where t represents elapsed time. Neural Networks have been widely used as time series forecasters. Most often these are feed-forward networks which employ a sliding window over the input sequence. Typical examples of this approach are market predictions, meteorological and network traffic forecasting. An important issue that must be addressed in such systems is that of the number of data points which should be used in the input representation. It is not always the case that the model with the highest resolution has the best predictive power, so that superior results may be obtained by employing only every n^{th}point in the series. Work in neural networks has concentrated on forecasting future developments of the time series from values of x up to the current time. Formally this can be stated as: find a function $f : R^N \rightarrow R$ such as to obtain an estimate of x at time $t + d$, from the N time steps back from time t, so that:

$$x(t + d) = f(x(t), x(t - 1), ..., x(t - N + 1)) \tag{5}$$

$$x(t + d) = f(y(t)) \tag{6}$$

Where $y(t)$ is the $N - ary$ vector of lagged values. Normally d will be one, so that f will be forecasting the next value of x. The standard neural network method of performing Time Series Prediction is to induce the function f using any feed-forward function approximating Neural Network architecture, such as a standard Multi Layer Perceptron or a Radial Bias Function architecture, using a set of $N - tuples$ as inputs and a single output as the target value of the network. This method is often called the sliding window technique as the $N - tuple$ input slides over the full training set [4].

Time series are generally sequences of measurements of one or more visible variables of an underlying dynamic system,

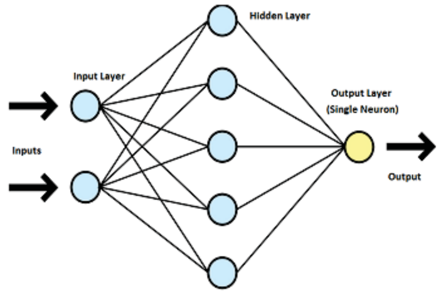


Fig. 2. Architecture of an Feed Forward Multi Layer Neural Network

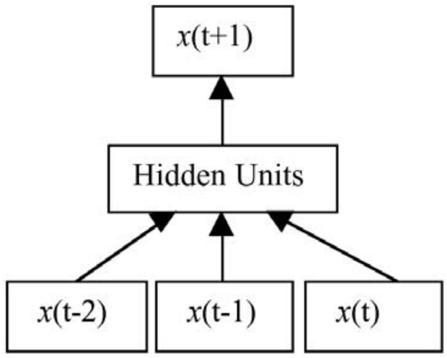


Fig. 3. ANN Architecture for implementing Sliding Window Technique

whose state changes with time as a function of its current state vector:

$$d(u(t))/dt = G(u(t)) \tag{7}$$

For the discrete case, the next value of the state is a function of the current state:

$$u(t + 1) = u(F(t)) \tag{8}$$

Such dynamic systems may evolve over time to an attracting set of points that is regular and of simple shape; any time series derived from such a system would also have a smooth and regular appearance. However the other result that is possible is that the system may evolve to a chaotic attractor. Here, the path of the state vector through the attractor is non-periodic and because of this any time series derived from it will have a complex appearance and behavior. In a real-world system such as a stock market, the nature and structure of the state space is obscure; so that the actual variables that contribute to the state vector are unknown or debatable. The task for a time series predictor can therefore be rephrased as; given measurements of one component of the state vector of a dynamic system, reconstruct the (possibly) chaotic dynamics of the phase space and thereby predict the evolution of the measured variable. The embedding theorem of Mane and Takens shows that the space of time lagged vectors y with sufficiently large dimension will capture the structure of the original phase space [5].

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:9, No:5, 2015

	Test1	Test2	Test3	Test4	Test5	...	Test100	MAX MARKS BY INDIVIDUAL
Individual 1	9	8	9	9	10	...	8	10
Individual 2	6	5	6	7	8	...	8	8
Individual 3	7	5	6	3	5	...	8	9
Individual 4	3	4	6	5	5	...	7	7
Individual 5	2	8	9	6	8	...	9	9
Individual 6	3	2	1	1	0	...	3	5
Individual 7	9	9	9	9	8	...	10	10
Individual 8	10	10	8	8	9	...	9	10
Individual 9	7	6	6	8	9	...	10	10
Individual 10	8	7	7	9	10	...	8	10
Individual 11	6	7	8	9	10	...	10	10
Individual 12	9	7	10	10	8	...	9	10
...	...	...	...	...	...	...	...	...
Individual 100	7	6	6	7	7	...	7	8
MAX MARKS OF ANY INDIVIDUAL	10	10	10	10	10	...	10	

Fig. 4. Data Set obtained from Experiment on 100 Individuals

	Test1	Test2	Test3	Test4	Test5	...	Test100
Individual 1	0.9	0.8	0.9	0.9	1	...	0.8
Individual 2	0.75	0.625	0.75	0.875	1	...	1
Individual 3	0.7778	0.5556	0.6667	0.3333	0.5556	...	0.88889
Individual 4	0.4286	0.5714	0.8571	0.7143	0.7143	...	1
Individual 5	0.2222	0.8889	1	0.6667	0.8889	...	1
Individual 6	0.6	0.4	0.2	0.2	0	...	0.6
Individual 7	0.9	0.9	0.9	0.9	0.8	...	1
Individual 8	1	1	0.8	0.8	0.9	...	0.9
Individual 9	0.7	0.6	0.6	0.8	0.9	...	1
Individual 10	0.8	0.7	0.7	0.9	1	...	0.8
Individual 11	0.6	0.7	0.8	0.9	1	...	1
Individual 12	0.9	0.7	1	1	0.8	...	0.9
...	...	...	...	...	...	...	...
Individual 100	0.875	0.75	0.75	0.875	0.875	...	0.875

Fig. 5. Data Set Normalized with respect to best performance of individual (N_1)

	Test1	Test2	Test3	Test4	Test5	...	Test100
Individual 1	0.9	0.8	0.9	0.9	1	...	0.8
Individual 2	0.6	0.5	0.6	0.7	0.8	...	0.8
Individual 3	0.7	0.5	0.6	0.3	0.5	...	0.8
Individual 4	0.3	0.4	0.6	0.5	0.5	...	0.7
Individual 5	0.2	0.8	0.9	0.6	0.8	...	0.9
Individual 6	0.3	0.2	0.1	0.1	0	...	0.3
Individual 7	0.9	0.9	0.9	0.9	0.8	...	1
Individual 8	1	1	0.8	0.8	0.9	...	0.9
Individual 9	0.7	0.6	0.6	0.8	0.9	...	1
Individual 10	0.8	0.7	0.7	0.9	1	...	0.8
Individual 11	0.6	0.7	0.8	0.9	1	...	1
Individual 12	0.9	0.7	1	1	0.8	...	0.9
...	...	...	...	...	...	...	...
Individual 100	0.7	0.6	0.6	0.7	0.7	...	0.7

Fig. 6. Data Set Normalized with respect to best individuals performance (N_2)

	Test1	Test2	Test3	Test4	Test5	...	Test100
Individual 1	0.9	0.8	0.9	0.9	1	...	0.8
Individual 2	0.6791	0.5659	0.6791	0.7923	0.9055	...	0.9055
Individual 3	0.7399	0.5285	0.6342	0.3171	0.5285	...	0.8456
Individual 4	0.3699	0.4932	0.7398	0.6165	0.6165	...	0.8631
Individual 5	0.2114	0.8456	0.9513	0.6342	0.8456	...	0.9513
Individual 6	0.4743	0.3162	0.1581	0.1581	0	...	0.4743
Individual 7	0.9	0.9	0.9	0.9	0.8	...	1
Individual 8	1	1	0.8	0.8	0.9	...	0.9
Individual 9	0.7	0.6	0.6	0.8	0.9	...	1
Individual 10	0.8	0.7	0.7	0.9	1	...	0.8
Individual 11	0.6	0.7	0.8	0.9	1	...	1
Individual 12	0.9	0.7	1	1	0.8	...	0.9
...	...	...	...	...	...	...	...
Individual 100	0.7923	0.6791	0.6791	0.7923	0.7923	...	0.7923

Fig. 7. Net Normalized Data Set (N_{net})

III. Model

The model is characterized by a *Characteristic Function* F. The aim is to study the outputs of the model correlating the effect of will power on an individuals success. The model uses data from the experiment to represent the sequence of events in the lives of individuals. The features of the model are:

1) There are 100 individuals in the model.
2) Each individual is characterized by a set of attributes.
3) The 5 attributes associated with each individual are *Life story*, *Will*, *Success*, *Apathy Factor*, and *Random Factor*.
4) The attributes, *Will* and *Success* uniquely characterize each individual, and are the subjects of interest in the study.
5) *Success* is a function of *Success* (past success), *Will*, *Apathy Factor* and *Random Factor*.
6) As each individual goes through events in his/her *Life Story*, his/her *Will Power* and *Success* change, while *Apathy Factor* remains unaltered. *Random Factor* varies with each life event.

	Test1	Test2	Test3	Test4	Test5	...	Test100
Individual 1	0.8	0.6	0.8	0.8	1	...	0.6
Individual 2	0.3582	0.1318	0.3582	0.5846	0.811	...	0.811
Individual 3	0.4798	0.057	0.2684	-0.3658	0.057	...	0.6912
Individual 4	-0.2602	-0.0136	0.4796	0.233	0.233	...	0.7262
Individual 5	-0.5772	0.6912	0.9026	0.2684	0.6912	...	0.9026
Individual 6	-0.0514	-0.3676	-0.6838	-0.6838	-1	...	-0.0514
Individual 7	0.8	0.8	0.8	0.8	0.6	...	1
Individual 8	1	1	0.6	0.6	0.8	...	0.8
Individual 9	0.4	0.2	0.2	0.6	0.8	...	1
Individual 10	0.6	0.4	0.4	0.8	1	...	0.6
Individual 11	0.2	0.4	0.6	0.8	1	...	1
Individual 12	0.8	0.4	1	1	0.6	...	0.8
...						...	
Individual 100	0.5846	0.6791	0.6791	0.5846	0.5846	...	0.5846

Fig. 8. Final Net Normalized Data Set(N'_{net})

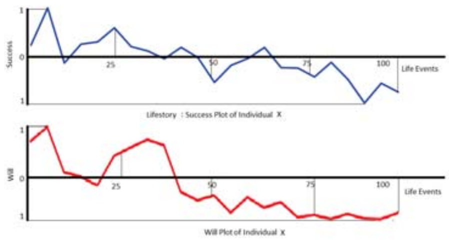


Fig. 9. Success Plot(Life Story) and Will Plot of an Individual

A. Life Story

It is possible to obtain the data pertaining to success by the means of a study conducted in a controlled environment on a 100 individuals. In the study each individual would be subjected to a 100 small tests over a period of time. The data presented in Fig.4 has been generated randomly to illustrate the procedure that must follow data collection. Since success is a subjective experience for every individual, absolute success needs to be normalized so as to make the stories of different individuals more easily comparable. Normalization is once performed with respect to each individuals best performance, and once with respect to the best performance among all the individuals. We thus get two orthogonal normalizations of the dataset. The net normalization accounting for the two is calculated using a Euclidean distance formula:

$$(N_{net})^2 = ((N_1)^2 + (N_2)^2)/2 \tag{9}$$

Where N_1 and N_2 represent the two orthogonal normalizations and N_{net} represents the net normalized dataset. At last, the Net Normalized set is made to undergo a transformation mapping the range of the dataset from (0,1) to (-1,1).This is done using the linear map:

$$x = ax + b \tag{10}$$

Where substituting the end-points of the ranges (0,1) and (-1,1) we get:

$$-1 = a(0) + b \tag{11}$$

$$1 = a(1) + b \tag{12}$$

Solving the equations; we obtain a and b as:

$$a = 2; b = -1 \tag{13}$$

Applying the transformation, we get the final Net Normalized data. In the final Net Normalized dataset, each horizontal column corresponds to a life story of an individual. Formally, the *Life Story* is a one-dimensional array of 100 real numbers in the range of -1 and 1. These numbers represent outcomes of 100 events in the *Life Story* of an individual. A positive outcome in life event is represented by a positive number, while a negative event is represented by a negative number. The magnitude of the outcome indicates the degree of importance of the outcome to the individual.

B. Will Power

Will Power of an individual is a variable initialized to a Random value at the start of a *Life Story*. As an individual goes through the outcome of events in their *Life Story*, *Will* is updated as:

$$\Delta Will = \begin{cases} F(Success) & F(Success) \leq 1 \\ 1 & Otherwise \end{cases} \tag{14}$$

$$Will = \begin{cases} Will + \Delta Will & Will + \Delta Will \leq 1 \\ 1 & Otherwise \end{cases} \tag{15}$$

Where the function F which characterizes the model is fixed for that model.

C. Success

The recent history of success of the individual is represented as *Success* and is the previous value of success or an average of the past few values of success. *Success* is updated after each life event as follows:

$$Success = (1 - A) * (Will) * (Success)R \tag{16}$$

A is a constant gain called the *Apathy Factor* that determines how much the individual is affected by the outcome of life events. R is a *Random Factor*. R is a *Gaussian* white noise to set in to account for factors that affect *Success* independent of *Will*. It assumes a random value for every life event.

TABLE I
SYMBOLS - PARAMETERS OF MODEL

Symbol	Parameter	Type	Range
F	Characteristic Function	Function	–
$\Delta Will$	Incremental change in Will Power	float	(0,1)
$Will$	Will Power	float	(0,1)
$Will$	Will Power	1-D Array of float	(0,1)
$Success$	Past Success	float	(-1,1)
$Success$	Success	float	(-1,1)
$Success$	Success	1-D Array of float	(-1,1)
A	Apathy Factor	float	(0,1)
R	Random Factor	float	(0,1)
N_1	Data Set Normalized with respect to best performance of individual	2-D Array of float	(0,1)
N_2	Data Set Normalized with respect to best individuals performance	2-D Array of float	(0,1)
N_{net}	Net Normalized Data Set	2-D Array of float	(0,1)
N'_{net}	Final Net Normalized Data Set	2-D Array of float	(-1,1)

D. Apathy Factor

Apathy Factor is the amount of indifference of an Individual towards success and failure. It is constant for an Individual. It is a factor which denotes the individuals detachment towards outcomes, accounting for the predominance of the effect that dopamine has on them.

E. Random Factor

Random Factor accounts for the randomness in life. Since this factor is random, it is different for every event in the *Life Story* of an individual, i.e. it is a randomly generated number. After each life event, R gets updated to a new random number between 0 and 1.

IV. Algorithm

For every individual:

1) Initialize (float) $Will \rightarrow$ Gaussian random number
2) Store current value of (float) $Will$ in (Array) $Will$
3) Initialize (float) $A \rightarrow$ Gaussian random number
4) Set (float) $R \rightarrow$ Gaussian random number
5) Set (float) $Success$ = 1
6) Calculate (float) $Success$ based on initial will
7) Store current value of (float) $Success$ in (array) $Success$
8) Repeat till every event in $LifeStory$ is exhausted:
    a) Calculate (float) $\Delta Will$ based on latest value of $Success$
    b) Calculate (float) $Will$
    c) Store current value of (float) $Will$ in (Array) $Will$
    d) Set (float) $R \rightarrow$ Gaussian random number
    e) Determine (float) $Success$
    f) Calculate (float) $Success$
    g) Store current value of (float) $Success$ in (array) $Success$
9) Plot (Array) $Success$ versus event number
10) Plot (Array) $Will$ versus event number

V. Procedure

1) For every odd integer m in the range of [-9 , 9), Calculate p such that $p = m + 0.5$
2) Let the union of the sets containing all possible values of p and m be Set S
3) For every number n of Set S, do:
   a) Set the Function F to be a Polynomial function of power n, such that $F(x) = x^n$
   b) Generate the model corresponding to F
   c) Obtain the Plots of $Success$ and $Will$. Tabulate the data
   d) Train a feed forward neural network using first 80 % of the data from the table
   e) Use the network in prediction mode on the rest 20 % of the data from the corresponding $Success$ plot, and tabulate the Outputs corresponding to $Will$. This step makes the network predict a data point of $Will$ corresponding to every data point of $Success$
   f) Check if the Plot of $Will$ generated by the network matches the Plot generated from the model
   g) Re-run the entire procedure to verify the results predicted by the network remain the same
   h) Retain the models that give a greater than 80 % match between predicted and obtained data
   i) Select the best-fit model out of the retained models

VI. Observations

At the end of the experiment when the best-fit model has been selected, Plots of *Will* and Success with time will be generated for each individual in the model. It should be interesting to:

1) Compare individuals with similar *Life Stories* and Initial *Will* to see if they have drastically different *Will* Plots. This juxtaposes the *Will Power* and *Success* plots of individuals with similar overall fortunes.
2) Compare individuals with similar *Life Stories* but drastically different Initial *Wills* to see if they have notably different *Will* plots.
3) Compare individuals whose *Will* increases overall and see if they all have *Apathy Factors* in a common range. This observation will help determine if individuals who are resilient in life are in fact individuals who are more detached towards the outcomes of their actions.

VII. Criteria of Evaluation

The criteria of evaluation for the correctness of the model is that the results predicted by the network must remain same each time the best-fit model is re-run. The relationship between *Will* and *Success* stated in the model should be considered to be a fair approximation of reality if the criteria of evaluation is satisfied. The model predicts that for a constant value of *Will*, *Success* increases in an inverse proportional with *Apathy Factor*. Also, *Will* is related to *Success* through a power function F. Should the index of the function be positive *Will* would be inversely proportional to *Apathy Factor* , and should it be negative, *Will* would be proportional to it. For a model with a *Characteristic Function* with a positive index, we would need to check if individuals with smaller values of *Apathy Factor* indeed experience increased fluctuations in their *Will* plots as compared to individuals with greater *Apathy Factor*. If either this or the opposite case was found to be true, it would confirm the correctness of the model. At last a study of the plots of *Will* and *Success* as generated from the final model data would give a hint to the extent up to which will power affects the success of each of the individuals subjected to the experiment.

VIII. Conclusion

This study is a proposal for a quantitative analysis of behavioral phenomena. The results may validate or challenge a long standing belief regarding the same. The model, though based on simple premises has the potential to address the consistency of fundamental beliefs. The model could prove to be a step toward bridging the gap between the subjectivity of experience and objectivity of quantitative sciences. The approach of using a computational model and a neural network in collaboration to verify behavioral models might pave the way for a new technique of conducting psychological studies based on objective measurements. If the positive dependence of the two parameters is confirmed and the model is able to generate the *Will* and *Success* plots of each individual in the study, it might result in a fundamental shift in the way we conduct capability analysis of potential candidates in eligibility detection and selection scenarios. One can imagine a situation where one might be interested to select candidates on the basis a specific criteria that would maximize the chances of success of that individual at a particular task. Additional constraints may be built into the model to accommodate for the additional specificity required for such purposes. For instance, a model created for short-listing potential candidates for a heavy combat training program would have to have constraints built in, that would cause it to ascribe increased weight to success in activities involving physical endurance over other kinds of activities.

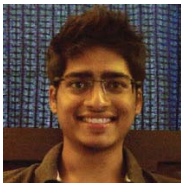
Acknowledgment

The author would like to thank Dr. V.K. Govindan, Professor and Ms Lijiya A, Assistant Professor at CSE Dept., National Institute of Technology Calicut for their continued guidance, encouragement and support through the period of conceptualization of the study.

References

[1] Seth Herd ,1, Brian Mingus and Randall OReilly, Dopamine and self-directed learning, Biologically Inspired Cognitive Architectures : DBLP, 2010, pp. 58-63.
[2] Kent C. Berridge , Terry E. Robinson , What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?, Brain Research Reviews : Elsevier, 1998, pp.309369.
[3] Marilyn M Nelson, W T Illingworth, A Practical Guide to Neural Networks, Prentice Hall PTR, 1991
[4] R.J.Frank, N.Davey, S.P.Hunt,Time Series Prediction and Neural Networks, Journal of Intelligent and Robotic Systems,May 2001, Volume 31, Issue 1-3, pp 91-103

[5] Ababarnel H., D., I., Brown R., Sidorowich J., L. and Tsimring L., S., 1993, The analysisof observed chaotic data in physical systems.Reviews of Modern Physics, Vol. 65, No. 4,pp1331-1392.
[6] www.apa.org/helpcenter/willpower.aspx (as of 16-11-13).
[7] today.uconn.edu/blog/2012/11/uconn-researcher(as of 16-11-13).



Yudhajit Datta Yudhajit Datta was born in Kolkata, India, in 1993. He is currently pursuing the final year of his B.Tech degree in Electrical and Electronics Engineering from National Institute of Technology Calicut, India. He has a keen interest in the applications of computer engineering. His research interests include artificial intelligence, robotics and computer vision.