

Deterministic Finite Automata in the Detection of EEG Spikes and Seizures

Rory A. Lewis^{1,2}, Doron Shmueli¹, and Andrew White¹

¹ Departments of Pediatrics & Neurology, University of Colorado Denver, Anschutz Medical Campus, Denver, CO 80262

² Department of Computer Science, University of Colorado at Colorado Springs, Colorado Springs, CO, 80933

Abstract. This Paper presents a platform to mine epileptiform activity from Electroencephalograms (EEG) by combining the methodologies of Deterministic Finite Automata (DFA) and Knowledge Discovery in Data Mining (KDD) TV-Tree. Mining EEG patterns in human brain dynamics is complex yet necessary for identifying and predicting the transient events that occur before and during epileptic seizures. We believe that an intelligent data analysis of mining EEG Epileptic Spikes can be combined with statistical analysis, signal analysis or KDD to create systems that intelligently choose when to invoke one or more of the aforementioned arts and correctly predict when a person will have a seizure. Herein, we present a correlation platform for using DFA and Action Rules in predicting which interictal spikes within noise are predictors of the clinical onset of a seizure.

1 Introduction

Epilepsy is a neurological disorder that makes people susceptible to recurrent unprovoked seizures due to electrical disturbances in the brain. Unfortunately, 30% of patients that suffer from epilepsy are not well controlled on medication. Only a small fraction of these can be helped by seizure surgery [5]. Therefore, it would be life changing to a large number of individuals if a system could be developed that would predict a seizure hours, minutes, or even seconds before its clinical onset. The challenge in this problem is that the dimensionality is huge; in the human brain there are approximately 100 billion neurons, each with about 1000 connections (synapses)[28]. Even in the rat brain it is estimated that there are approximately 200 million neurons [4], [1]. The connections are wired such that the problem is highly chaotic. In a certain class of seizures it would be helpful if they could be detected even a few seconds prior to the start of a seizure. The dimensionality of the problem can be significantly reduced, with only a small loss of information by recording electrical potentials at multiple points on the surface of the skull or, using depth electrodes, in the hippocampus (EEG). EEGs are accepted as one of the best means of evaluating neurocognitive functions [15]. EEG spike/seizure detection and prediction is made more complicated by the

following: (1) For a single individual, no two seizures or even their EEG correlates are exactly alike, (2) seizures from different individuals vary significantly, (3) there is no single metric that consistently changes during all seizures, (4) correlation among channels can change significantly from one seizure to the next, and (5) even experts disagree as to what constitutes a seizure [27]. Occasionally, the reduction in dimensionality does result in an indeterminate mapping from EEG record to animal state (i.e. it is not surjective or onto). For the reasons listed above, rigid seizure detection rules do not produce good results [7], [26]. Interictal spikes are brief (20 - 70 ms) sharp spikes of electrical phenomena that stand out when compared to background EEG rhythms and may be indicative of an underlying epileptic process. Because they are considered as an indicator of the presence of epileptic seizures, and may actually precede a seizure (sentinel spike), the detection of these interictal, transient spikes which may be confused with artifact or noise is indeed a crucial element in the prediction of epileptic conditions.

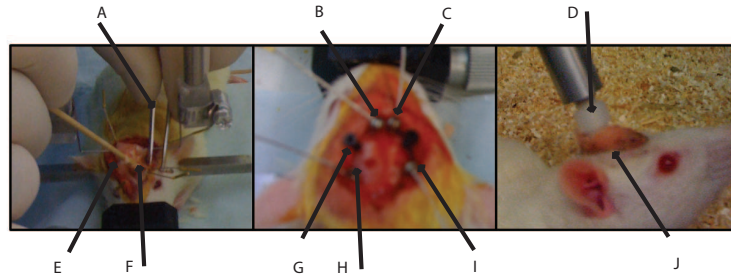


Fig. 1. *Implantable Tethered System Devices:* (A-D) Stereotaxic placement of cortical electrodes. (E) Dental cement polymer applied to hold the electrodes in place. Note dental cement on q-tip. (F) The tethered pre-amplifier connects to the implanted electrodes and sends the signals to Epilepsy Monitoring Unit)

2 Recording Epileptogenesis

Until 1992 most EEG analysis was based on analysis of brain slices [14] or anesthetized animals [3]. Kainic acid, a chemoconvulsant extracted from seaweed, was introduced to induce seizures in animals. This provided a major breakthrough particularly with the advent of monitoring the animals on video, but the equally significant subclinical seizures were impossible to detect with video monitoring alone. The field was further advanced through the development of a tethered recording system [2] in which multi-channel cortical and sub-cortical recordings could be obtained. The quality of recordings were further improved by incorporating a small pre-amplifier close to the skull, allowing for a significant increase in the signal to noise (S/N) ratio. As shown in Figure 1, electrodes were placed

stereotaxically in the hippocampus and secured in the skull [25], [24] Additional electrodes were placed directly on the dura. Dental cement was applied to hold the electrode pins together in a plastic cap that was later connected to the pre-amplifier. The pre-amplified signal was sent to an amplifier and from there to a computer for storage.

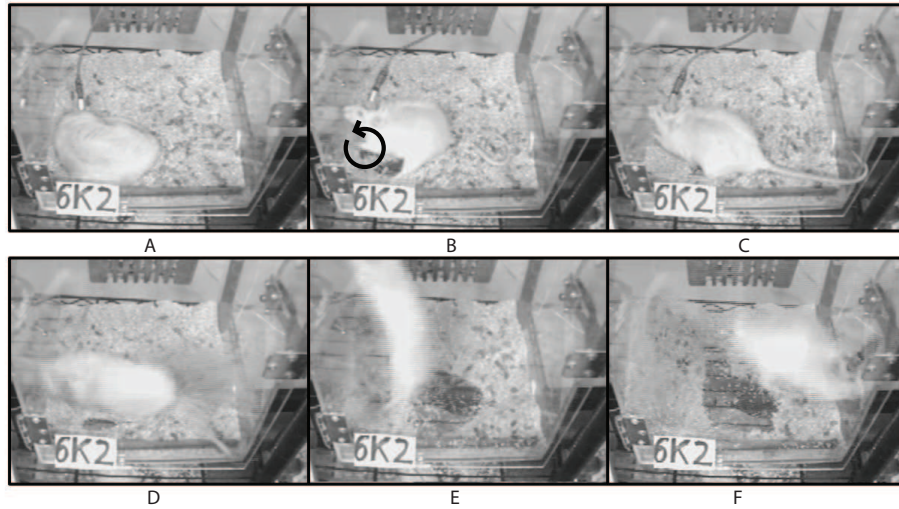


Fig. 2. *Rat 6K2*: Progression of a Clinical convulsive seizure: (A) Video capture shows the Rat to be in a normal sleeping stage. (B-C) Rat exits sleeping stage and starts having a p3 seizure (racine scale). (D-F) The seizure magnitude escalates and leads to a violent uncontrollable seizure - p5 (racine scale).

Our facility has the capability of continuously monitoring up to 64 tethered or untethered rats. Untethered rats underwent video monitoring and the tethered rats underwent both video and EEG monitoring. This paper will discuss an algorithm for analyzing the EEG of a rats experiencing an event that evolves into a P5 Seizure [18] This event occurs somewhat infrequently, and we have collected EEG data set with three events that includes both the seizures and several minutes surrounding the events. As seen in Figure 2, a rat experienced a kainite-induced seizure that evolved from stage P3 to P5. In frame A, the rat was sleeping. In Frame B, 58 seconds later, the rat experienced a P3 seizure evidenced by the circular clawing (forelimb clonus). In Frame C, 28 seconds later, convulsive activity stopped and there was no epileptic activity on the EEG. Frame D was taken 1 minute later and at this point the rat began to experience a P5 seizure that lasted several seconds. Frames E and F were taken subsequently and demonstrate the intensity of the seizure. Not seen in this figure is that this rat was eating calmly shortly after the end of the seizure. The availability of EEG data for this seizure and 2 other similar seizures, along with video,

allowed us to test the hypothesis that a novel deterministic finite automata (DFA) methodology will be able to differentiate the different aspects (sleeping, P3 seizure, between seizures, and P5 seizure) of the EEG record.

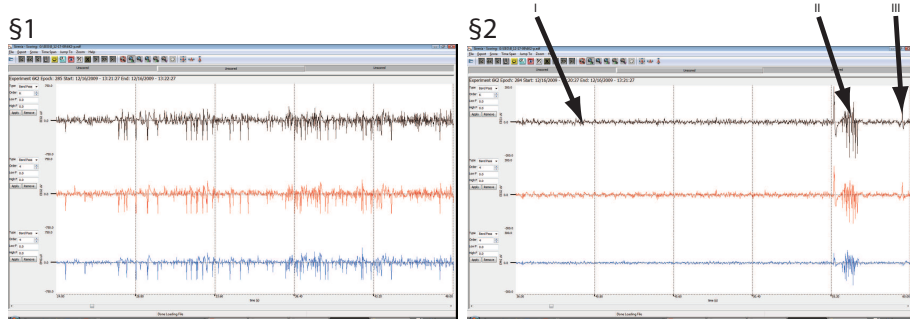


Fig. 3. Rat p_3 seizure (racine scale): §1: 6K2's EEG state correlating to Figure 2's Stage B denoted by circular clawing §2: I. Normal EEG wave-form while Rat is a sleep. II. Appearance of Sentinel Spike prior to P3 (Racine Scale) seizure. III. Possible Interictal Spike often misinterpreted as Artifact and vice-versa..

2.1 EEG Analysis

For our analysis EEG potentials were sampled at 800Hz. EEG electrodes were placed bilaterally in the hippocampi (referenced to a common dural screw) and a separate channel recorded from the dura. Each EEG contains a approximately 100,000 time points. As such, its interpretation is non-trivial and attempts at automating the analysis have met with only limited success. In this paper, we seek to demonstrate the efficacy of the DFA algorithm to distinguish all 4 states in each seizure event and distinguish artifact from interictal spikes and other noise. An author (RL) has begun to integrate statistical analysis with Action Rules [9], [19], [21] in Signals with a system influenced Dr. Zdzislaw Pawlak [17], [20], [23]. Fourier Action Rules Trees of signal distortion [10] and 3) Machine Learning with Signal Noise, Genetic algorithms [12] and FS-trees, Rough Sets, LERS, PNC2, J45, CART, & Orange. [11]

Figure 3§1 illustrates one rat's normal EEG wave-form while asleep. Point II in Figure 3§1 illustrates a Sentinel Spike, the hallmark indicator that a seizure is imminent. Point III in Figure 3§1 shows a region in which the EEG record deviates from baseline. An important task is to determine whether this deviation is simply artifact or if it represents epileptiform activity (an interictal spike) Figure 3§2 provides details of the P3 seizure. Figure 4§3 provides a zoomed-out overall view of all four stages and Figure 4§4 provides details of the P5 (Racine Scale) stage.

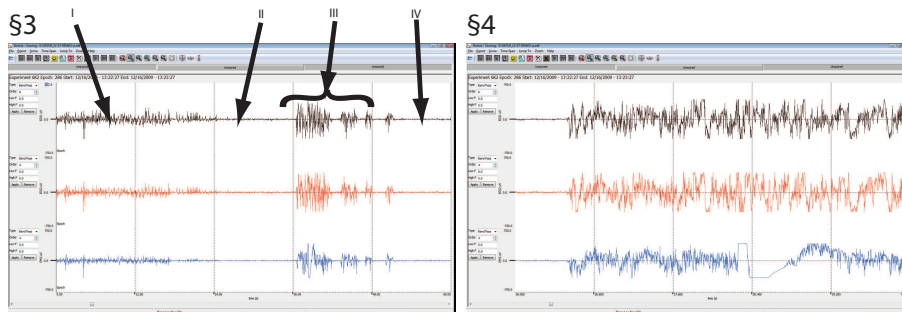


Fig. 4. *Rat P3's Seizure (Racine Scale): §3: I.* A P3 (Racine Scale) seizure in progress. *II.* Period of no Electrographic seizing activity. *III.* Electrographic outburst indicating the violent seizure time. *IV.* End of Electroencephalographic seizure event. **§4:** Covers periods D, E and F in Figure 2 showing the EEG analysis while 6K2 experiences the P5 (Racine Scale) seizure *Figure 2.*

3 Methods

3.1 Deterministic Finite Automata (DFA)

Deterministic finite automata can be used in many applications. We used this methodology to track the current state of a finite-state EEG system. As time moved forward the particular system state would change dependent on such quantities as amplitude, slope, second derivative, Short-Term Fourier Transform STFT (average frequency) as well as other signal features. For the current analysis, programming was done using Visual Basic subroutines and data was stored using the European Data Format (EDF). At time zero we begin at state zero. The state at the next time step is assumed to be dependent only on the current system state and conditions (input state) during the current time step (e.g. slope). A consequence of this is that the current state is independent of the order in which the input states occurred (in this sense it is similar to a Markov chain).

Illustrative Example of our DFA Methodology To motivate reader understanding we illustrate the concept of our usage of DFA using a simplified transition table given in Figure 5. The table in Figure 5 has a total of ten columns. The first column is the current system state. As one moves from one time point of the EEG to the next, the state of the system changes. The state to which the system changes is based upon the transition matrix; each of the columns in the transition matrix represents a current parametric set (e.g. slope within a particular range). More generally, the columns may represent a condition in which current or past parameters have specified values. It should comprise a collectively exhaustive and mutually exclusive set such that there are no events that either fall into multiple columns or do not fall into any of the columns. One proceeds from one system state at a given time point to the next system state

at the next time point until a terminal event occurs. Terminal events can be the identification of spikes, seizures or artifact.

As mentioned previously, the first column of the transition table corresponds to the current system state. The particular input state at the current time (columns 5 - 10) is ascertained by investigating parameters at that time or at prior times. For this example, we have six mutually exclusive, collectively exhaustive input states. These are presented in Table 1 where α , β , and γ are limits selected by the authors using expert knowledge of what parameters would be characteristic of spikes. We note that states 4, 5 and 6 are the same as 1, 2, and 3 except that the absolute value of the second derivative ($f''(x)$ or $\frac{d^2y}{dt^2}$) is less than γ . The purpose of the use of slope ($m = \frac{y_2 - y_1}{t_2 - t_1}$) is to differentiate between the normal state, the possibility of a spike, and likely artifact (artifact, such as that noted when the animal is chewing, is often distinguished from spike because the slope is much greater).

State	# of 1/4	# of 2/5	# of 3/6	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6
0	0	0	0	0	2	0	0	2	0
1	0	0	1	0	3	0	0	3	0
2	0	1	0	10	4	3	10	4	3
3	0	1	1	11	5	0	11	5	0
4	0	2	0	12	6	5	12	6	5
5	0	2	1	13	7	0	13	7	0
6	0	3	0	14	6	7	14	16	7
7	0	3	1	15	7	0	15	16	0
8	1	0	0	0	10	0	0	10	0
9	1	0	1	0	11	0	0	11	0
10	1	1	0	0	12	11	0	12	11
11	1	1	1	0	13	0	0	13	0
12	1	2	0	0	14	13	0	14	13
13	1	2	1	0	15	0	0	15	0
14	1	3	0	0	14	15	0	16	15
15	1	3	1	0	15	0	0	16	0
16	0	0	0	24	18	17	24	18	17
17	0	0	1	25	19	0	25	19	0
18	0	1	0	26	20	19	26	20	19
19	0	1	1	27	21	0	27	21	0
20	0	2	0	28	22	21	28	22	21
21	0	2	1	29	23	0	29	23	0
22	0	3	0	30	22	23	30	32	23
23	0	3	1	31	23	0	31	32	0
					⋮				
$n-1$	1	3	0	0	30	31	0	32	31
n	1	3	1	0	31	0	0	32	0

Fig. 5. *Sample Transition Table:* where the number of possible states is 31, the number of states with slope too high required for rejection is 2, number of states required for slopes in the range for acceptance is 4, and the number of states with slope too low requiring rejection is 2.

The purpose of the use of $f''(x)$ is to ensure that there is actually a peak and not just a baseline shift. Columns 2 - 4 indicate the number of times that input states 1 or 4, 2 or 5, or 3 or 6 respectively have occurred. For example, looking at system state 12 one notes that there has been a single event in which the input state 1 or 4 existed, two events in which input state 2 or 5 existed and no events in which the state 3 or 6 existed. To register a spike, there must be two time points in which ($m = \frac{y_2 - y_1}{t_2 - t_1}$) falls in the range expected for a spike (input states 2 or 5), followed by one time point in which $f''(x)$ is high (input state 5) which corresponds to a peak, followed by two time points in which the slope again falls in the correct range (input states 2 or 5). This sequence must occur before one obtains two slopes greater than the range or two slopes less than the range. In this example a spike is indicated by system state 32. A heuristic definition can then be used to establish the seizure state by requiring a certain number of spikes in a particular time interval (e.g. 20 detected spikes in 10 seconds).

Input State Conditions	
1	$ m = \frac{y_2 - y_1}{t_2 - t_1} > \alpha \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} < \gamma$
2	$\alpha > m = \frac{y_2 - y_1}{t_2 - t_1} > \beta \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} < \gamma$
3	$ m = \frac{y_2 - y_1}{t_2 - t_1} < \beta \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} < \gamma$
4	$ m = \frac{y_2 - y_1}{t_2 - t_1} > \alpha \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} > \gamma$
5	$\alpha > m = \frac{y_2 - y_1}{t_2 - t_1} > \beta \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} > \gamma$
6	$ m = \frac{y_2 - y_1}{t_2 - t_1} < \beta \wedge f''(x) \text{ or } \frac{d^2 y}{dt^2} > \gamma$

Table 1. Six mutually exclusive, collectively exhaustive input states. Where α , β , and γ are user selected constants. States 4, 5 and 6 are the same as 1, 2, and 3 except that the second derivative is less than a given value

We now consider the sample path through the transition matrix illustrated by the chain of circles noted. For this sample the sequence of input states are assumed to be: 2, 2, 3, 1, 2, 5, 5, 5, 5. We initially start with state 0, time interval 0. At this time, the slope was calculated to be appropriate for a spike, i.e. $\alpha < |slope| < \beta$, with the second derivative γ (input state = 2). As a result the transition matrix indicated a change to system state 2. For the second time interval the input state was calculated to be the same as that in the first time interval (input state = 2), and the transition matrix (row 3, column 6) indicated a change to system state 4. In the next time interval input state 3 was calculated and the system state of 5 (row 5, column 7) was determined. Subsequent input states could then be coupled with the current system states to draw a time path through the transition matrix. In this case, a spike is registered at the end of the

path because state 32 is obtained at the end of the chain. Had there been too many slopes that did not meet criteria, the system state would return to zero (see for example system state 9, input state 1).

3.2 Results

For our analysis of the data we used the same six input states given in the example above. These used only the slope and standard deviation to determine the input state. The transition table was significantly bigger, having 336 entries. This required 7 slopes in the correct range, but allowed 4 slopes to be too great and 6 slopes to be too small to terminate a spike search sequence. A screen capture of the code correctly identifying a spike is given in Figure 6. Similarly, a screen capture of the code correctly rejecting artifact is given in Figure 7. The algorithm was quite successful at determining the presence of spikes and using the heuristic definitions of seizures ($> 20spikes$ in $10seconds$), was able to detect all seizures without difficulty. This determination was made within 6 seconds of the onset of the seizure. Figure 8 gives the EEG recording on which the detected spikes are indicated. Unfortunately, the code was unable, using only slope and standard deviation, to differentiate between either sleep and interseizure period or between the P3 and P5 portions of the seizure.

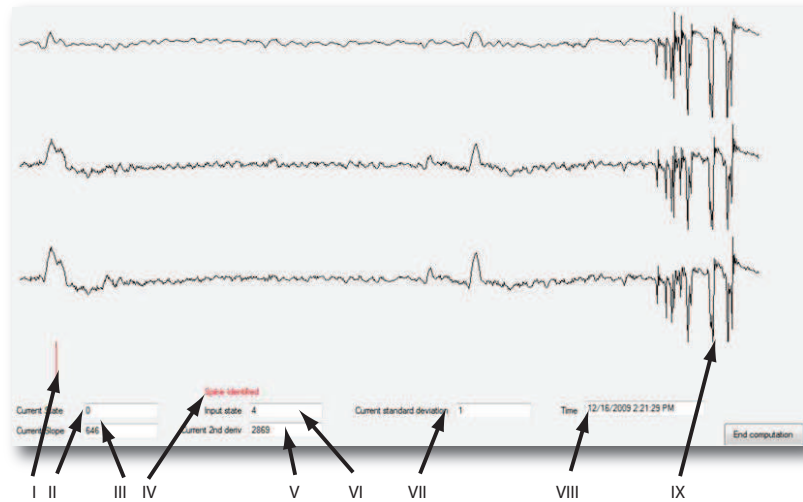


Fig. 6. *Correctly Detecting Spike:* Direct Screen Capture from Visual Studio Platform where **I.** Indicates red line programmed to identify the existence of pre-seizure spike. **II.** Current State. **III.** Current Slope. **IV.** Red "pop-Up" programmed to alert a pre-seizure spike is detected. **V.** Current 2^{nd} Derivative. **VI.** Current Standard Deviation. **VII.** Input State. **VIII.** Time of Day, and **IX.** Beginning of correctly predicted Seizure.

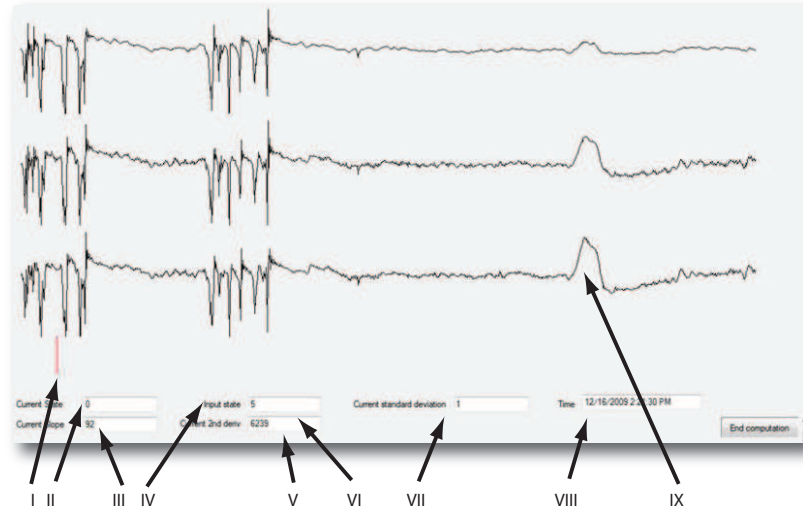


Fig. 7. *Correctly Detecting Artifact:* Direct Screen Capture from Visual Studio Platform where *I.* Indicates red line programmed to identify the existence of Artifact. *II.* Current State. *III.* Current Slope. *IV.* Red "pop-Up" programmed to alert a pre-seizure spike is detected. *V.* Current 2nd Derivative. *VI.* Current Standard Deviation. *VII.* Input State. *VIII.* Time of Day, and *IX.* Possible spike located visually. Program soon detects it as a positive.

3.3 Conclusions

DFA is an extremely flexible platform that can be used to identify spikes and seizures and to sort various events that occur during seizures. The flexibility allows it to mimic and also use other techniques in its determination. In the present case, we have only scratched the surface of the true capability of the methodology. It is possible to greatly extend our analysis through the use of much more sophisticated input states. These could use algorithms such as Fourier series or wavelet analysis to determine the best path through the transition matrix. It is not restricted to linear analysis such as Neural Networks, Random Forest and Machine Learning's J45 to define strong classifiers for items such as Sentinel Spikes; it is also possible to use sequential non-linear analysis to establish whether or not spikes have occurred. By generalizing the input states to include past parameters, it is even possible to force the current state to be dependent on the path taken to get to the current state. The transition matrix can also be modified in such a way that multiple final deterministic states are possible (i.e. multiple end points could be identified). It is our plan to investigate further methods in which the DFA algorithm can be successfully employed. This includes the process of integrating KDD with the DFA methods and also considering the use of time domain analysis of EEG signal by statistical analysis and characteristics computation [13] with different frequencies [22], non-linear

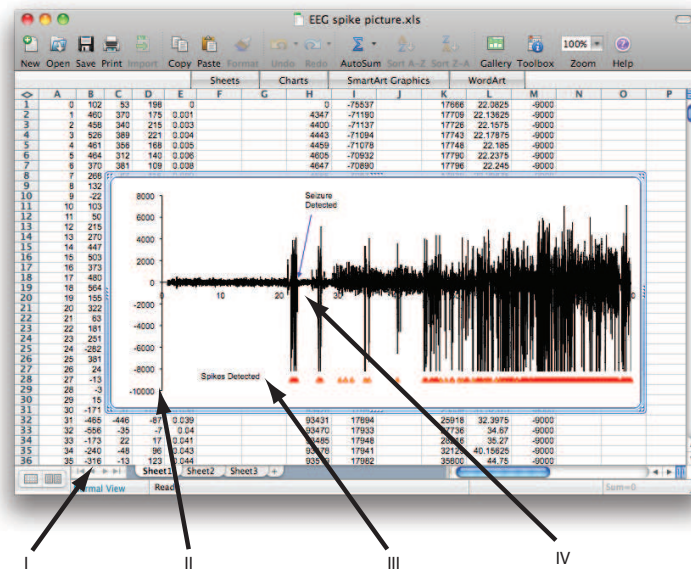


Fig. 8. Results: Seizure Correctly Predicted: Code correctly identifies a seizure 6 seconds before its onset. **I.** Excel Spreadsheet with CSV output. **II.** Excel built in graph. **III.** The Spikes Detected references and **IV.** Seizure Detected at location 6 seconds before onset.

dynamics and chaos theory [8], and intelligent systems such as artificial neural network and other artificial-intelligence structures [6], [16].

References

1. F. Bandeira, R. Lent, and S. Herculano-Houzel. Changing number of neuronal and non-neuronal cells underly post natal brain growth in the rat. *Proceedings of The National Academy of Sciences of the USA*, 106 (33):14108 – 13, August 18th 2009.
2. E. H. Bertram, J. M. Wiliamson, J. F. Cornett, S. Spradlin, and Z. F. Chen. Design and construction of a long-term continuous video-eeeg monitoring unit for simultaneous recording of multiple small animals. *Brain Res Brain Res Protoc*, 2:8597, 1997.
3. P. S. Buckmaster and F. E. Dudek. Neuron loss, granule cell axon reorganization, and functional changes in the dentate gyrus of epileptic kainate-treated rats. *The Journal of Comparative Neurology*, 385(3):385 – 404, 1998.
4. L. K. et al. An efficient method for estimating the total number of neurons in rat brain cortex. *Journal of Neuroscience Methods*, 31 (2) February:93–100, 1990.
5. H. Firpi, E. Goodman, and J. Echazu. Genetic programming artificial features with applications to epileptic seizure prediction. *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference, Shanghai*, 17-18 Jan::4510 – 4513, 2005.

6. A. B. Geva and D. H. Kerem. Forecasting generalized epileptic seizures from the eeg signal by wavelet analysis and dynamic unsupervised fuzzy clustering. *IEEE Transaction of Biomed Engineering*, 45:1205-1216, 1998.
7. J. L. Helliera, A. M. White, P. A. Williams, F. E. Dudek, and K. J. Staley. Nmda receptor-mediated long-term alterations in epileptiform activity in experimental chronic epilepsy. *Neuropharmacology*, 56 (2):414 – 421, 2009.
8. K. Lehnertz and C. Elger. Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss. *Electroencephalography And Clinical Neurophysiology*, 95:1081-117, 1995.
9. R. Lewis and Z. Raś. Rules for processing and manipulating scalar music theory. *Proceedings of the International Conference on Multimedia and Ubiquitous Engineering*, pages 819–824, 2007.
10. R. Lewis and A. Wiczorkowska. Categorization of musical instrument sounds based on numerical parameters. *June 28-30 in Warsaw Poland*, 4585/2007:784–792, 2007.
11. R. Lewis, X. Zhang, and Z. Raś. Mirai: Multi-hierarchical fs-tree based music information retrieval system. *June in Warsaw Poland*.
12. R. Lewis, A. Cohen, W. Jiang, and Z. Raś. Hierarchical tree for dissemination of polyphonic noise. In *RSCTC '08: Proceedings of the 6th International Conference on Rough Sets and Current Trends in Computing*, pages 448–456, Berlin, Heidelberg, 2008. Springer-Verlag.
13. B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, and M. Dichter. Epileptic seizures may begin hours in advance of clinical onset. a report of five patients. *Neuron*, 30(1):51 – 64, 2001.
14. P. Molnar and J. V. Nadler. Mossy fiber-granule cell synapses in the normal and epileptic rat dentate gyrus studied with minimal laser photostimulation. *The Journal of Neurophysiology*, 82(4):1883–1894, 1999.
15. E. Niedermeyer and F. L. da Silva. Electroencephalography basic principles, clinical applications & related fields. *Eds. Niedermeyer E, Lopes da Silva F,*, Williams and Wilkins: Baltimore (MD), 1999.
16. Y. Pan, S. S. Ge, F. Tang, and A. A. Mamun. Detection of epileptic spike-wave discharges using svm, in: *Proceedings of the 2007 IEEE International Conference on Control Applications*. *Suntec City*, Singapore:467 – 472, 2007.
17. Z. Pawlak. Information systems - theoretical foundations. *Information Systems Journal*, 6:205–218, 1991.
18. R. Racine. Modification of seizure activity by electrical stimulation: II. motor seizure. *Electroencephalogr Clin Neurophysiol*, 32:281–294, 1972.
19. Z. Raś and A. Dardzińska. Intelligent query answering. *Encyclopedia of Data Warehousing and Mining*, pages 639–643, 2005.
20. Z. Raś, A. Tzacheva, and L. Tsay. Action rules. *Encyclopedia of Data Warehousing and Mining*, pages 1–5, 2005.
21. Z. Raś and A. Wiczorkowska. Action rules: how to increase profit of a company. *Principles of Data Mining and Knowledge Discovery*, No. 1910:87–592, 2000.
22. Y. Salant, I. Gath, and O. Henriksen. Prediction of epileptic seizures from two-channel eeg. *Medical and Biological Engineering and Computing*, 36:549 – 556, 1998.
23. L. Tsay and Z. Raś. Action rules discovery system dear method and experiments. *Journal of Experimental and Theoretical Artificial Intelligence*, Vol. 17 No. 1-2:119–128, 2005.

24. A. M. White, P. A. Williams, D. J. Ferraro, S. Clark, S. D. Kadam, F. E. Dudek, and K. J. Staley. Efficient unsupervised algorithms for the detection of seizures in continuous eeg recordings from rats after brain injury. *Journal of Neuroscience Methods*, 152(1-2):255 – 266, 2006.
25. P. Williams, A. White, D. Ferraro, S. Clark, K. Staley, and F. E. Dudek. The use of radiotelemetry to evaluate electrographic seizures in rats with kainate-induced epilepsy. *Journal of Neuroscience Methods*, 155(1):39 – 48, 2006.
26. P. A. Williams, J. L. Hellier, A. M. White, S. Clark, D. J. Ferraro, W. Swiercz, K. J. Staley, and F. E. Dudek. Development of spontaneous recurrent seizures after kainate-induced status epilepticus. *The Journal of Neuroscience, February 18, 2009*, 29(7):21032112, 2009.
27. P. A. Williams, J. L. Hellier, A. M. White, K. J. Staley, and F. E. Dudek. Development of spontaneous seizures after experimental status epilepticus: Implications for understanding epileptogenesis. *Epilepsia (Series 4)*, 48:157 – 163, 2007.
28. R. W. Williams and K. Herrup. The control of neuron number. *The Annual Review of Neuroscience*, 11:423–453, 1988.