J Neurophysiol 121: 2001–2012, 2019.
First published April 3, 2019; doi:10.1152/jn.00763.2018.

INNOVATIVE METHODOLOGY | *Control of Movement*

A supervised machine learning approach to characterize spinal network function

A. N. Dalrymple,[1] S. A. Sharples,[2,3] N. Osachoff,[4] A. P. Lognon,[2,3] and P. J. Whelan[2,4]
[1]*Neuroscience and Mental Health Institute, University of Alberta, Edmonton, Alberta, Canada;* [2]*Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada;* [3]*Graduate Program in Neuroscience, University of Calgary, Calgary, Alberta, Canada; and* [4]*Department of Comparative Biology and Experimental Medicine, University of Calgary, Calgary, Alberta, Canada*

Submitted 12 November 2018; accepted in final form 20 March 2019

Dalrymple AN, Sharples SA, Osachoff N, Lognon AP, Whelan PJ. A supervised machine learning approach to characterize spinal network function. *J Neurophysiol* 121: 2001–2012, 2019. First published April 3, 2019; doi:10.1152/jn.00763.2018.—Spontaneous activity is a common feature of immature neuronal networks throughout the central nervous system and plays an important role in network development and consolidation. In postnatal rodents, spontaneous activity in the spinal cord exhibits complex, stochastic patterns that have historically proven challenging to characterize. We developed a software tool for quickly and automatically characterizing and classifying episodes of spontaneous activity generated from developing spinal networks. We recorded spontaneous activity from in vitro lumbar ventral roots of 16 neonatal [postnatal day (P)0–P3] mice. Recordings were DC coupled and detrended, and episodes were separated for analysis. Amplitude-, duration-, and frequency-related features were extracted from each episode and organized into five classes. Paired classes and features were used to train and test supervised machine learning algorithms. Multilayer perceptrons were used to classify episodes as rhythmic or multiburst. We increased network excitability with potassium chloride and tested the utility of the tool to detect changes in features and episode class. We also demonstrate usability by having a novel experimenter use the program to classify episodes collected at a later time point (P5). Supervised machine learning-based classification of episodes accounted for changes that traditional approaches cannot detect. Our tool, named SpontaneousClassification, advances the detail in which we can study not only developing spinal networks, but also spontaneous networks in other areas of the nervous system.

NEW & NOTEWORTHY Spontaneous activity is important for nervous system network development and consolidation. Our software uses machine learning to automatically and quickly characterize and classify episodes of spontaneous activity in the spinal cord of newborn mice. It detected changes in network activity following KCl-enhanced excitation. Using our software to classify spontaneous activity throughout development, in pathological models, or with neuromodulation, may offer insight into the development and organization of spinal circuits.

machine learning; neural recording; spinal cord; spontaneous activity

INTRODUCTION

For decades, isolated embryonic or neonatal spinal cord preparations have been used to study sensorimotor integration (Clemens and Hochman 2004; Kudo and Yamada 1987; Mentis et al. 2011; O'Donovan and Landmesser 1987; Smith and Feldman 1987) and to examine patterns of movement-related network activity (Cowley and Schmidt 1997; Smith and Feldman 1987; Whelan et al. 2000). Neonatal mice generate stepping movements when exposed to olfactory or cutaneous stimulation (Fady et al. 1998; Jamon and Clarac 1998), and patterns of locomotor-like activity can be generated in vitro by inducing a high excitability state in the network through electrical, optical, or neurochemical stimulation (Gordon et al. 2008; Gordon and Whelan 2006; Hägglund et al. 2010; Kiehn 2016; Sharples and Whelan 2017). However, the vast majority of hindlimb movements produced by neonatal rodents are sporadic ataxic episodes of uncoordinated rhythmic action (Jamon and Clarac 1998). A fictive correlate of these movements is recognizable in vitro as spontaneous activity, which represents the motor output of the developing spinal networks in a more natural activity state. Spontaneous activity plays a critical role in the development of motor networks, in synaptic and receptor plasticity, and in guiding the innervation of descending motor and ascending sensory afferents, both of which are sparse at birth.

The qualitative complexity and stochastic nature of spontaneous activity patterns make them challenging to study. Previous work investigating spontaneous network behavior has primarily focused on quantifying AC-coupled or high-pass filtered activity (for examples, see Fellippa-Marques et al. 2000; Hanson and Landmesser 2003), which can detect activity peaks but may miss the subthreshold components. In contrast, DC-coupled recordings detect extraordinarily rich and detailed subthreshold information (O'Donovan 1987; Whelan et al. 2000; Fig. 1), which provides important clues about network dynamics that could be used as biomarkers for developmental disorders, such as cerebral palsy.

To enable the examination of such activity including the subthreshold components, we created a software tool that uses machine learning to classify episodes of spontaneous patterns of motor activity recorded from isolated mouse neonatal spinal cords in vitro. Using MATLAB (MathWorks, Natick, MA), we

Address for reprint requests and other correspondence: P. Whelan, HMRB 168, 3330 Hospital Dr. NW, Calgary, Alberta, T2N 4N1, Canada (e-mail: whelan@ucalgary.ca).

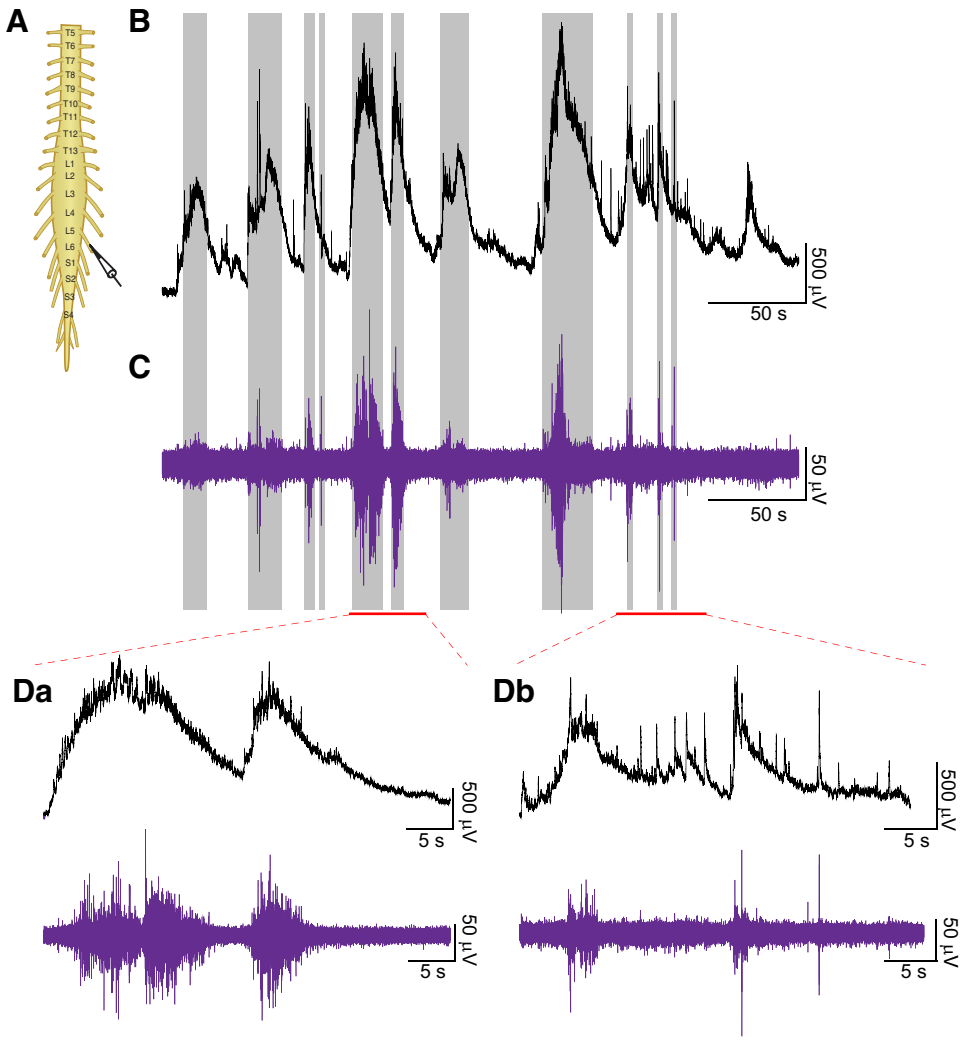

Fig. 1. Spontaneous network activity recorded from the lumbar spinal cord. *A*: experimental setup: suction electrodes recorded activity from the L2 and L5 ventral roots of neonatal mice (P0–P3). *B*: DC-coupled neurogram; gray bars highlight episodes detected in AC-coupled neurogram. *C*: high-pass filtered (100 Hz) neurogram; gray bars highlight episodes detected in AC-coupled neurogram. *Da* and *Db*: enlarged regions. Top traces depict DC-coupled neurograms; *bottom traces* depict AC-coupled neurograms.

developed the software in a graphical user interface (GUI) that allows users to detect, visualize, and characterize episodes of spontaneous motor activity. This tool, called SpontaneousClassification, quickly and automatically extracted features and classified episodes using supervised machine learning. This approach is substantially faster and easier than manual classification. In addition, it detected changes in stochastic spontaneous network activity after an induced increase in spinal network excitability.

SpontaneousClassification may facilitate investigations of spinal network function in a different manner than currently available techniques, expanding the toolbox of methods used to investigate spinal neural networks. It may also enable the analysis of patterns in developing retinal networks (Galli and Maffei 1988; Sernagor and Grzywacz 1995; Torborg and Feller 2005; Wong et al. 1995) and in structures such as the hippocampus (Garaschuk et al. 1998; Menendez de la Prida et al. 1996; Sipilä et al. 2006), cerebellum (Watt et al. 2009), and thalamus (Mooney et al. 1996; Pangratz-Fuehrer et al. 2016) in greater detail than currently available techniques. The custom MATLAB code and an executable file, along with the raw files used to generate the data illustrated in this report, are available in a data repository (Dalrymple et al. 2019). Portions of these data were previously presented in abstract form (Sharples SA et al. 2014).

METHODS

Ethical Approval

All procedures used were approved by the University of Calgary Health Sciences Animal Care Committee in accordance with the Canadian Council on Animal Care guidelines (Protocol AC16-0182). Pregnant mice were sourced from Charles River Laboratories, Wilmington, MA and were provided ad libitum access to food and water.

Tissue Preparation

Experiments were performed in neonatal C57BL/6 mice [postnatal days (P)0–P5; $n = 32$] obtained from timed-pregnant females housed in the animal care facility at the University of Calgary until they gave birth. Neonates were anesthetized by cooling, then decapitated and eviscerated to expose the vertebral column. The remaining tissue was placed ventral side up in a dissection chamber filled with room-temperature carbogenated (95% O_2-5% CO_2) artificial cerebrospinal fluid (aCSF) (128 mM NaCl, 4 mM KCl, 1.5 mM $CaCl_2$, 1 mM $MgSO_4$, 0.5 mM Na_2HPO_4, 21 mM $NaHCO_3$, 30 mM D-glucose). A ventral laminectomy exposed the spinal cord. We cut the dorsal and ventral roots, removed and transferred the spinal cord to a recording chamber, ventral side up, with recirculating carbogenated aCSF at a flow rate of 20 ml/min, and gradually heated them from room temperature to 27°C. This temperature is closer to the physiological neonatal core temperature (32°C; Goodrich 1977), is above room

temperature, which tends to fluctuate, and is known to produce more reliable activity in our experience.

Electrophysiological Recordings

Using Clampex software (Molecular Devices, Sunnyvale, CA), we acquired DC-coupled neurograms from the isolated spinal cords (i.e., spontaneous motor activity) by drawing the ventral roots from the bilateral second and fifth lumbar segments (L2, L5) into tight-fitting suction electrodes filled with aCSF (Fig. 1*A*). Neurograms were amplified 1,000 times, digitized using a 2.5-kHz sampling frequency (Digidata 1440, Molecular Devices), and analyzed with our custom-designed MATLAB program SpontaneousClassification on a laboratory computer. The recordings from each root were treated independently for analysis. In certain experiments we manipulated the excitability of the network by increasing the concentration of KCl in the aCSF from 4 to 8 mM. In these experiments spontaneous activity was recorded and analyzed during 20 min of baseline at 4 mM. Spontaneous activity was recorded for 30 min after increasing KCl to 8 mM, from which the final 20 min were analyzed. Most experiments were completed by one experimenter; however, in a final set of experiments using physiological aCSF, we further tested the algorithm's reliably to detect activity from older preparations (P5) completed by a different experimenter.

Classification Software

SpontaneousClassification has three modules: visualization, data preparation and feature extraction, and classification. The classification module includes two supervised machine learning classifiers trained on spontaneous motor activity episodes recorded from the isolated spinal cords. A user manual and video describing SpontaneousClassification are available as supplemental files (Dalrymple et al. 2019).

Visualization module. This module displays a trace from the loaded neurogram file in the GUI where users can access multiple zooming, shifting, and sliding options that operate along the *x*- and *y*-axes. Traces can also be split into multiple files for further analysis. See the user manual in the supplemental documents for further details.

Data preparation and feature extraction module. This module prepares the data for further stages of analysis and extracts features of interest from the episodes.

EPISODE DETECTION. Episode detection first requires the trace to be detrended to remove the DC drift acquired during the recording. After detrending the trace, thresholds must be selected for each trace to separate the episodes. Episode threshold was determined by first sorting all data points from each trial into bins of ascending amplitude. Figure 2*B* depicts the highest amplitude data points, which likely represent bursting activity, on the right side of the *x*-axis, and a flat region to their left, with lower amplitude, nonbursting activity. The flat region's limits can be determined by viewing the bins and will likely need to be tuned for each different trace, based on variation in recording quality (i.e., signal to noise) and amplitude range, which depends on the amount of spontaneous activity. To determine the episode threshold, we added a multiple of the standard deviation of nonbursting baseline data points to the mean of the data points within a selected region of the nonbursting baseline. We found that a multiple of four times the standard deviation optimized results. We then used the episode threshold to determine temporal criteria for episode onset and offset boundaries (time required above threshold and time required below episode threshold, respectively). In our experiments, we defined episode onset as the recording spending 0.25 s above threshold and episode offset as the recording spending 0.25 s below threshold. For some traces, the automatic separation of episodes did not correctly capture all the episodes; therefore, some episodes were separated manually. SpontaneousClassification allows for both automatic and manual episode separation. Displaying the detected episodes in the SpontaneousClassification GUI can be helpful when tuning detection parameters (Fig. 2*C*).

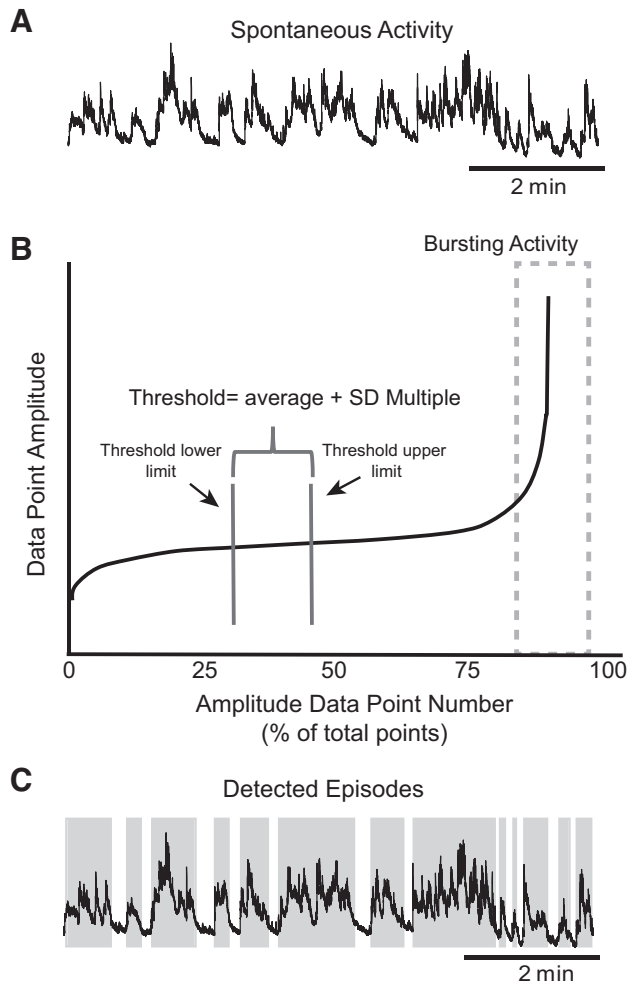

Fig. 2. Detecting episodes of spontaneous activity. *A*: DC-coupled neurogram. *B*: amplitude of data points from *A* sorted in ascending order. Threshold reflects user-defined upper and lower limit values of the sorted amplitudes within the flat region. The average amplitude of the data points plus a user-defined standard deviation of the data points within this region is then set as the episode detection threshold. *C*: boundaries of detected episodes are highlighted in gray.

FEATURE EXTRACTION. SpontaneousClassification identified 902 episodes of spontaneous activity in 40 traces recorded from the 17 spinal cords. Twelve episodes from one animal were excluded from analysis due to aberrant multispike activity that was extremely low frequency and long in duration. The remaining 890 episodes were categorized into one of five classes.

From each episode, SpontaneousClassification extracted four time-, two amplitude-, and three frequency-related features (Fig. 3*Aa*). Time-related features for a given episode *e* include *1*) the start and end times for the episode, *2*) the time from the previous episode's offset to *e*'s onset (time from previous), *3*) the time from the previous episode's onset to *e*'s onset (start-start time), and *4*) the duration of *e*. Amplitude-related features included the maximum and the average amplitude, expressed both as raw voltage and as a percentage of the absolute maximum in a trace. The frequency-related features (Fig. 3*Ab*), obtained using a discrete Fourier transform, include the largest frequency component (peak frequency), the bandwidth, and the power at the largest frequency component (peak power). To focus on features related to activities of interest, such as locomotion or scratching, we limited the

frequency components to those below 10 Hz. We defined bandwidth as the highest frequency component (limited to 10 Hz) minus the lowest (nonzero) frequency component. Table 1 lists the values for each feature. On average, episodes were 18.2 s (SD = 17.6) long, had the largest frequency component near 1 Hz (0.9 Hz, SD = 1.4), and had a wide range of amplitudes, averaging 654.1 μV (SD = 414.3). The remarkable diversity with respect to the qualitative nature of the recorded episodes was well suited to inform supervised machine learning approaches to classify the different types of episodes we observed.

Classification module. After automatically extracting features from each episode of spontaneous motor activity, the classification module uses those features to classify episodes according to generalizations obtained from a trained supervised machine learning algorithm programmed into SpontaneousClassification. SpontaneousClassification comes with pretrained machine learning classifiers. The following sections describe how we trained and tested them.

SUPERVISED MACHINE LEARNING. Supervised machine learning entails using prelabeled data to derive generalizations about the relationship between inputs and outputs. Features that describe or quantify the data serve as inputs, and the outputs may be classification or numerical values (also referred to as supervisory signals; Witten et al. 2017). The inferred function from training with prelabeled examples is then presented with new data and tested for how well it can generalize on novel data. To determine the accuracy of the generalization on a testing data set, predicted outputs are compared with actual outputs. Ideally, the generalizations are broad enough to properly classify new data without overfitting the inference function to the training data.

SUPERVISED CLASSIFICATION PROCEDURE. The episodes in the training data set were recorded from spinal cords from animals aged P0–P3 and were classified by the experts in two batches; *batch one* contained 576 episodes and *batch two* contained 253 episodes. Three individuals classified each episode of activity as either rhythmic, not rhythmic, multiburst and rhythmic, or multiburst and not rhythmic. The final classes also indicated whether an episode's amplitude was larger or smaller than 50% of the maximum amplitude of a single trace. Therefore, five classes were derived based on the following descriptors: small amplitude; large amplitude, not rhythmic; large amplitude, rhythmic; multiburst, not rhythmic; and multiburst, rhythmic (Fig. 3, *Ba–c* and *Ca–b*). Further segregation of small amplitude classes was not necessary as they were nonrhythmic and not multiburst.

For the first batch, each expert classified the episodes independently. Any episodes that did not have 100% agreement were reclassified by all three experts as a group. We added the second batch of episodes to increase the representation in each class; those episodes were classified by the same individuals as a group. After classification, the episodes were divided into a training set, consisting of 90% of the episodes (n = 735), and a testing set, consisting of the remaining 10% of the episodes (n = 82). The episodes in the testing set were selected proportionally according to class to reflect the true proportion of the episodes among the data sets. We used a random number generator to select the episodes within each class to the training or testing set. A third batch (73 episodes) was used as an additional independent test set. The classes assigned by the experts were considered to be true and used to train and evaluate supervised machine learning algorithms.

TRAINING PROCEDURE. We paired the eight automatically extracted features related to time, amplitude, and frequency (inputs) to the class labeled by the experts (output) for each episode in the training data set (735 episodes). These paired inputs and outputs were used to train the machine learning algorithms. Using the data mining platform Weka 3.8.0 (Frank et al. 2016) and the training data set from the first batch of episodes we swept through numerous supervised machine learning algorithms, using varied parameters. The second batch of episodes were added after this initial sweep through the algorithms. Each run of the training data set through an algorithm and parameter setting was 10-fold cross-validated. We tested 13 classifying algorithms: Bayesian network, naive Bayes, multilayer perceptron (MLP), simple logistic model, k-nearest neighbor, sequential minimal optimization; for training a support vector classifier), instance-based learning (K*), J48 (a decision tree), reduced error pruning tree, decision stump, Hoeffding tree, logistic model tree, and random tree. We selected the best algorithms, and parameters within algorithms, for further testing, based on the cross-validated classification accuracy results.

Across all classification methods tested, we executed a total of 2,947 sweeps through algorithms and parameters. Initially, classification to one of the five classes was performed by a single algorithm; the best training accuracy achieved was 56.5% using an MLP. An MLP is a type of feed-forward artificial neural network modeled after biological neural networks (Minsky and Papert 1969). The neurons, or nodes, in an MLP are connected by weights, forming a network capable of representing complex expressions. Additional hidden layers, such as those in deep neural networks, can further increase the complexity of the representations (see Fig. 4). This technique has been applied to facial recognition (Parkhi et al. 2015; Sun et al. 2015), speech recognition (Amodei et al. 2016; Hinton et al. 2012), and medical diagnostics (Burt et al. 2018; Lu et al. 2018; Song et al. 2018).

Since the episodes could be separated by their descriptors, two binary classifiers were found to have the best overall classification performance. One binary classifier determined whether episodes were rhythmic or not, and the other determined whether episodes were multiburst or not. The normalized maximum amplitude determined the size of each episode. According to classification performance, the top contenders for binary classification during training for rhythmicity were an MLP and a Bayesian network; an MLP was the top choice for determining whether an episode was multiburst or not.

TESTING PROCEDURE. After determining the final classifying algorithms, we used the testing data sets to evaluate performance of the trained algorithms on new data. The testing data set included the 10% hold-out set from the initial two batches of episodes (n = 82 episodes), as well as episodes from a third, independent batch of

Table 1. *Number and proportion of episodes within each class and overall for each feature*

Class	Total, no., %	Training, no.	Testing, no.	Time from Prev, s	St-st Time, s	Duration, s	Peak Freq, Hz	Bandwidth, Hz	Peak Power, V	Max Amp, μV	Avg Amp, μV
S	367(44.9%)	330	48	15.4 (19.2)	34.9 (27.5)	7.5 (7.2)	1.6 (1.7)	8.7 (1.3)	23.6 (30.9)	371.9 (180.2)	156.2 (76.5)
LnR	124 (15.2%)	112	24	20.2 (26.2)	36.6 (30.9)	16.5 (13.8)	0.8 (1.0)	9.3 (0.7)	79.1 (169.1)	807.5 (428.2)	282.5 (154.6)
LR	42 (5.1%)	38	19	14.3 (15.8)	28.1 (23.9)	20.2 (12.8)	0.8 (0.9)	9.6 (0.3)	78.3 (142.5)	876.4 (367.9)	312.1 (129.4)
MnR	77 (9.4%)	69	27	15.9 (24.3)	30.1 (29.5)	33.6 (18.1)	0.2 (0.4)	9.8 (0.2)	90.1 (82.2)	944.6 (416.4)	324.0 (142.5)
MR	207 (25.3%)	186	37	15.8 (25.0)	33.1 (31.5)	30.3 (20.3)	0.3 (0.5)	9.7 (0.4)	87.3 (82.4)	856.9 (414.5)	304.3 (142.5)
Overall	890	735	155	16.2 (22.3)	33.8 (29.1)	18.2 (17.6)	0.9 (1.4)	9.2 (1.0)	58.7 (96.8)	654.1 (414.3)	240.7 (140.7)

Values are average (SD). Episodes were classified as small (S), large nonrhythmic (LnR), large rhythmic (LR), multiburst not rhythmic (MnR), or multiburst rhythmic (MR). Amp, amplitude; Avg, average; Freq, frequency; Max, maximum; Prev, previous; St-st, start to start.

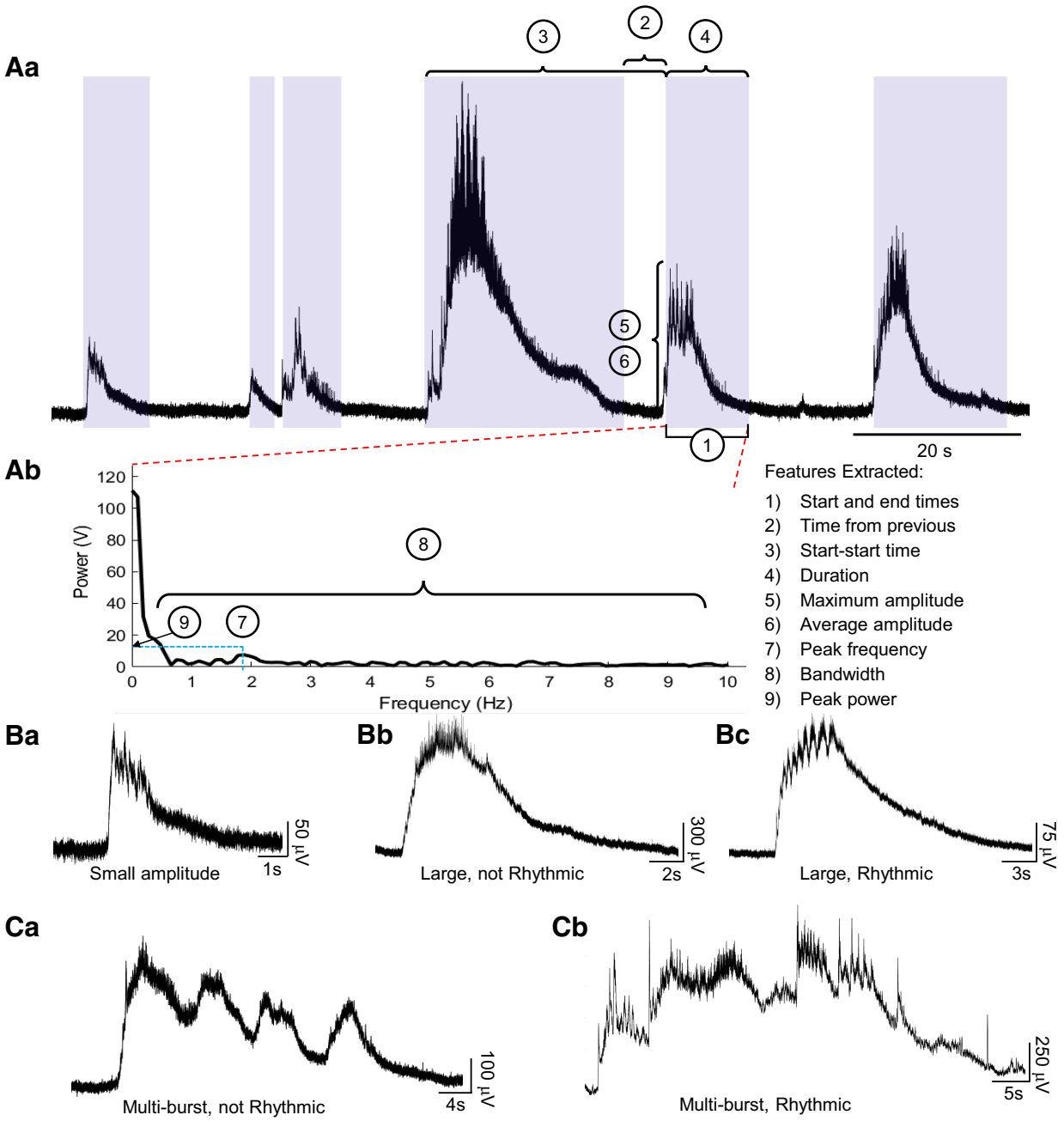

Fig. 3. Features extracted from each episode of spontaneous activity. *Aa*: time- and amplitude-based features extracted from each episode: (1) time from the previous episode, (2) start-start time, (3) duration, (4) maximum amplitude, and (5) average amplitude; both amplitudes expressed as raw voltage and percentage of largest amplitude in the trace. *Ab*: Fourier transform of the episode to extract frequency-related features; (6) largest nonzero frequency component; (7) bandwidth of peak components, excluding zero and frequency components larger than 10 Hz; and (8) power at most prominent nonzero frequency component. *B* and *C*: examples of classes of spontaneous activity: *Ba*, small; *Bb*, large, not rhythmic; *Bc*, large, rhythmic; *Ca*, multiburst, not rhythmic; *Cb*, multiburst, rhythmic.

episodes ($n = 73$ episodes). Performance metrics from training and testing were used to compare algorithms for the binary classifications. These metrics give a well-rounded picture of how well the algorithms generalize to the data.

Accuracy describes the number of episodes for which class was correctly predicted by the algorithm

$$Accuracy = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) \times 100$$

where TP is the number of true positive classifications, TN is the number of true negative classifications, FP is the number of false positive classifications, and FN is the number of false negative classifications. For example, for the rhythmicity classifier, positive refers to a rhythmic episode, while negative refers to a nonrhythmic episode.

Specificity indicates the number of correctly predicted negative episodes, calculated as follows:

$$Specificity = \left(\frac{TN}{TN + FP}\right) \times 100$$

Sensitivity measures the number of positive class episodes correctly identified:

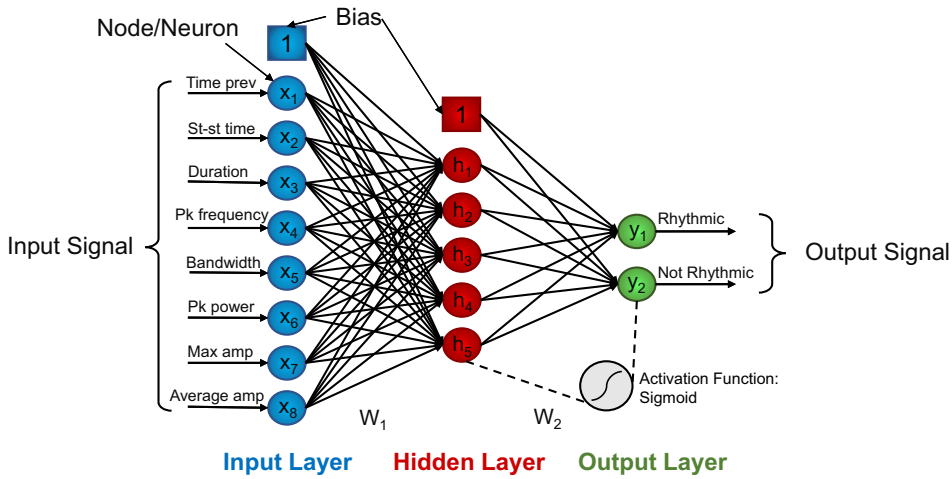

Fig. 4. Structure of multilayer perceptron (MLP) used for classifying episodes as rhythmic or not. The input layer (x, blue) receives normalized values of the features extracted from each episode. The weights between the input layer and the hidden layer (h, red), and the hidden layer and the output layer (y, green) were determined during training. The output layer consists of two binary nodes: 0 = not rhythmic, 1 = rhythmic. The hidden layer and output layer nodes are activated using the sigmoid function. The MLP for classifying episodes as multiburst or not has a similar structure, with a different number of nodes in the hidden layer and different learning parameters (learning rate and momentum) used during training to obtain the weight values.

$$Sensitivity = \left(\frac{TP}{TP + FN}\right) \times 100$$

Finally, precision measures the number of positive episodes correctly classified:

$$Precision = \left(\frac{TP}{TP + FP}\right) \times 100$$

We selected the supervised machine learning algorithms implemented in SpontaneousClassification based on a balance between training performance and testing performance. The same performance metrics were used to assess agreement between the experts who initially classified the episodes.

During testing of the final algorithms, the MLP had a higher accuracy, specificity, and precision than the Bayesian network for training and testing rhythmicity (see Table 2) and thus we implemented the MLP in SpontaneousClassification. The MLP during testing had a relatively smaller sensitivity to rhythmicity due to higher false negative classifications but had a moderate precision and very high accuracy and specificity due to a large number of true negative and satisfactory true positive identifications.

MULTILAYER PERCEPTRONS. As mentioned earlier, an MLP, such as those we chose to implement in SpontaneousClassification, is a feed-forward artificial neural network with an input layer, a hidden layer, and an output layer, each with nodes that represent neurons. Figure 4 depicts a schematic of an MLP for rhythmic binary classification.

For our MLPs, the number of nodes in the input and output layers was defined by the number of features and classes, respectively. Therefore, each MLP has eight input nodes, one for each normalized feature, and two output nodes, 0 (false, or nonrhythmic) and 1 (true, or rhythmic). The number of nodes in the hidden layer was based on performance during training. The MLP trained for classifying rhythmic and nonrhythmic episodes has five nodes in the hidden layer (Fig. 4). The MLP for classifying whether or not episodes are multiburst has 10 nodes in the hidden layer. The input and hidden layers also contain a bias unit of 1 that connects to the nodes in the following layer.

The weights connecting the nodes of the input layer to the hidden layer, and the hidden layer to the output layer, were learned through back-propagation during training in Weka (ver. 3.8.0; Frank et al. 2016). Back-propagation uses gradient descent to update the weights by minimizing errors (Witten et al. 2017). The step size for gradient descent and learning momentum were determined based on training performance. We obtained the best results using a learning rate equal to 0.7 and a momentum equal to 0.5 for the rhythmic or not MLP, and a learning rate of 0.1 and momentum of 0.4 for the multiburst or not MLP. The hidden and output layers are activated using a sigmoid function. The binary classification is finalized at the output layer, where the node with the greatest activation determines an episode's class.

GUI capability summary. Figure 5 depicts the final classification procedure implemented in the GUI. It includes the two MLPs (one to classify episodes for rhythmicity and the other for whether it is multiburst or not) that use weights and parameters obtained through training, features extracted from the data preparation and feature extraction module, the output of the MLPs, and the normalized maximum amplitude to automatically classify each episode into one of the five classes (cf. Fig. 3*B*). The entire process of episode detection, feature extraction, and classification takes a matter of

Table 2. *Classification performance for each expert who labeled episodes for training and for the final supervised machine learning algorithms chosen for testing*

	Expert A	*Expert B*	*Expert C*	MLP Train	MLP Test	Bayes Train	Bayes Test
Rhythmic							
Accuracy, %	94	87	83	84	78	78	71
Specificity, %	91	77	97	91	91	80	76
Sensitivity, %	98	98	69	39	44	70	42
Precision, %	91	79	95	43	66	37	23
Multiburst							
Accuracy, %	97	93	90	86	83		
Specificity, %	97	100	89	92	89		
Sensitivity, %	98	78	92	74	73		
Precision, %	92	98	78	83	82		

Classification included labeling episodes as rhythmic or not and multiburst or not. MLP, multilayer perceptron; Test, testing; Train, training.

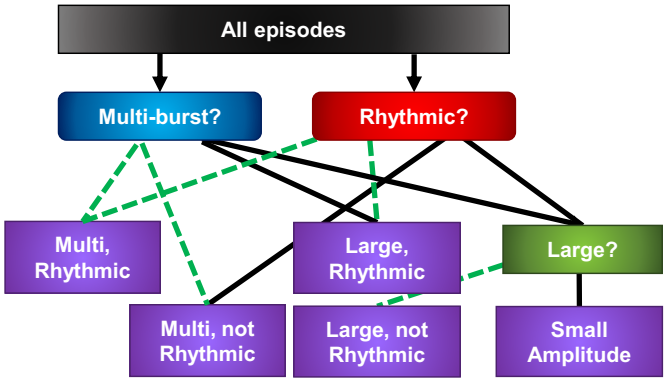

Fig. 5. Final classification procedure. Two multilayer perceptrons were used to classify episodes, one as rhythmic or not, and the other as multiburst or not. The results from these two binary classifiers combine to label episodes into classes: large rhythmic, multiburst not rhythmic, multiburst rhythmic. Amplitude information serves to further classify episodes as large nonrhythmic and small amplitude.

minutes to perform. The automated classification process is completed in seconds or less, significantly reducing the analysis process from manual classification, and conveniently compiling everything into a spreadsheet.

Statistics

We used χ^2 tests to compare performance measures from three expert (human) classifiers, with cross tabulations generated for all pairwise combinations. We compared episode features, before and after inducing excitation with KCl, via paired *t*-tests. We used Wilcoxon (*W*) signed-rank tests whenever data failed tests of normality or equal variance. The proportion of episodes within each class between the two excitability states were compared using the χ^2 test, with cross-tabulations generated all pairwise combinations. $P \leq 0.05$ was used to indicate significance for all tests.

RESULTS

The expert individuals who classified the first batch of episodes reached majority consensus for 95.3% of the episodes [55.4% (319/576) unanimous agreement, 39.9% (230/576) two of three agreed; 4.7% (27/576) no agreement]. Table 2 presents a detailed assessment of the experts' agreement and human error, evaluated using the metrics described above. Each individual was compared with the consensus of the other two experts to demonstrate the consistency among the experts as they were considered to be the gold standard for the algorithms. Overall, all experts had high accuracy. *Expert A* performed well at labeling episodes as rhythmic and multiburst. *Expert B* had high specificity for labeling episodes as multiburst or not ($P \leq 0.001$; χ^2), but not rhythmic or not ($P < 0.0001$; χ^2). *Expert C* was not very sensitive to labeling episodes as rhythmic or not ($P < 0.0001$; χ^2) but was for multiburst or not ($P < 0.0001$; χ^2). *Expert B* was less precise at labeling episodes as rhythmic or not ($P \leq 0.001$; χ^2), while *expert C* was less precise at labeling episodes as multiburst or not ($P < 0.0001$; χ^2). Since the second and third batches were classified as a group or pair, these metrics could not be calculated for classifying those episodes.

We first determined the accuracy of episode classification by the algorithms (for an overview of classes see Fig. 3, *B* and *C*), as that information was required to form the training set for the machine learning aspect of the software and to reveal the value of classifying spontaneous activity. As can be seen in Table 1, the data are rich with detail. We found that a majority of the episodes recorded from spinal cords of animals age P0–P3 ($n = 17$ preparations; 890 episodes) were small in amplitude and duration (44.9%). While large episodes were most often not rhythmic (15.2%), multiburst episodes were most often rhythmic (25.3%). Multiburst, nonrhythmic episodes had the largest maximum and average amplitudes, suggesting that larger amplitudes may not correlate with the probability of rhythmic activity. The duration of multiburst nonrhythmic episodes was longer than all other classes. Notably, frequency-related features extracted from each episode and analyzed using traditional approaches, such as peak frequency, bandwidth, and peak power, were similar for rhythmic and nonrhythmic episodes. This implies that rhythmicity within an episode cannot be detected simply using extracted features and thresholds. Therefore, the data were used to train and test two multilayer perceptrons to classify episodes as rhythmic or not, and multiburst or not. SpontaneousClassification also used the normalized amplitude to classify episodes into one of five classes (see METHODS and Figs. 3 and 5). The ability to discriminate between rhythmic episodes using a combination of all of the features demonstrates the power of machine learning methods to make inferences and predictions using data beyond simple time- and frequency-domain based analyses.

We next tested the ability of SpontaneousClassification to detect changes in spontaneous network activity following an increase in excitability. This allowed us to compare normal network features with features of a perturbed network. Increasing the extracellular concentration of KCl from 4 to 8 mM (for $n = 8$ spinal cord preparations; age: P2) resulted in a robust depolarization of the DC potential in the ventral roots of the lumbar spinal cord [ΔVoltage = 2,100 (SD = 2,183) μV]. An overall increase in the number of episodes [baseline: 271 episodes, KCl: 387 episodes; $t_{(7)} = 4.1$; $P = 0.005$] accompanied the depolarization, but the increase was not uniform across classes; only small-amplitude [baseline: 137 episodes, KCl: 203 episodes; $t_{(7)} = 2.3$; $P = 0.05$] and multiburst, rhythmic [baseline: 89 episodes, KCl: 132 episodes; $t_{(7)} = 3.1$; $P = 0.02$] episodes increased significantly in number (Fig. 6*Ca*). Of note, while the number of episodes changed, the relative proportion of each episode class when normalized to the total number of episodes remained unchanged compared with baseline ($P = 0.429$, χ^2).

In some cases, global episode features changed, due to changes in their assigned episode classes. For example, globally, episodes were shorter [duration, baseline: 24.0 s (SD = 3.5), KCl: 19.0 s (SD = 3.4); $t_{(7)} = 3.40$; $P = 0.01$; Fig. 6*Da*] and had smaller mean and maximum amplitudes [mean amplitude, baseline: 252.2 μV (SD = 156.7), KCl: 112.3 μV (SD = 79.7); $t_{(7)} = 6.7$; $P = 0.0003$; maximum amplitude, baseline: 819.4 μV (SD = 546.3), KCl: 547.3 μV (SD = 334.4); $t_{(7)} = 5.8$; $P = 0.00065$; see Fig. 6, *Db* and *Dc*]. A reduction in duration was only significant in the small amplitude episode class [baseline: 9.9 s (SD = 3.4), KCl: 6.6 s (1.8); $t_{(7)} = 2.9$; $P = 0.02$] and a reduction in amplitude only in the small amplitude [maximum amplitude, baseline: 451.3 μV (SD = 112.8), KCl: 340.7 μV (SD = 86.1); $t_{(7)} = 4.4$, $P = 0.003$; average amplitude, baseline: 150.3 μV (SD = 38.7), KCl: 67.6 μV(SD = 22.5); $t_{(7)} = 7.3$, $P = 0.0002$] and multiburst rhythmic classes [maximum amplitude, baseline: 1,167 μV (SD = 422), KCl: 736 μV (SD = 268); $W = -36$, $T^+ = 0.00$; $T^- = -36$,

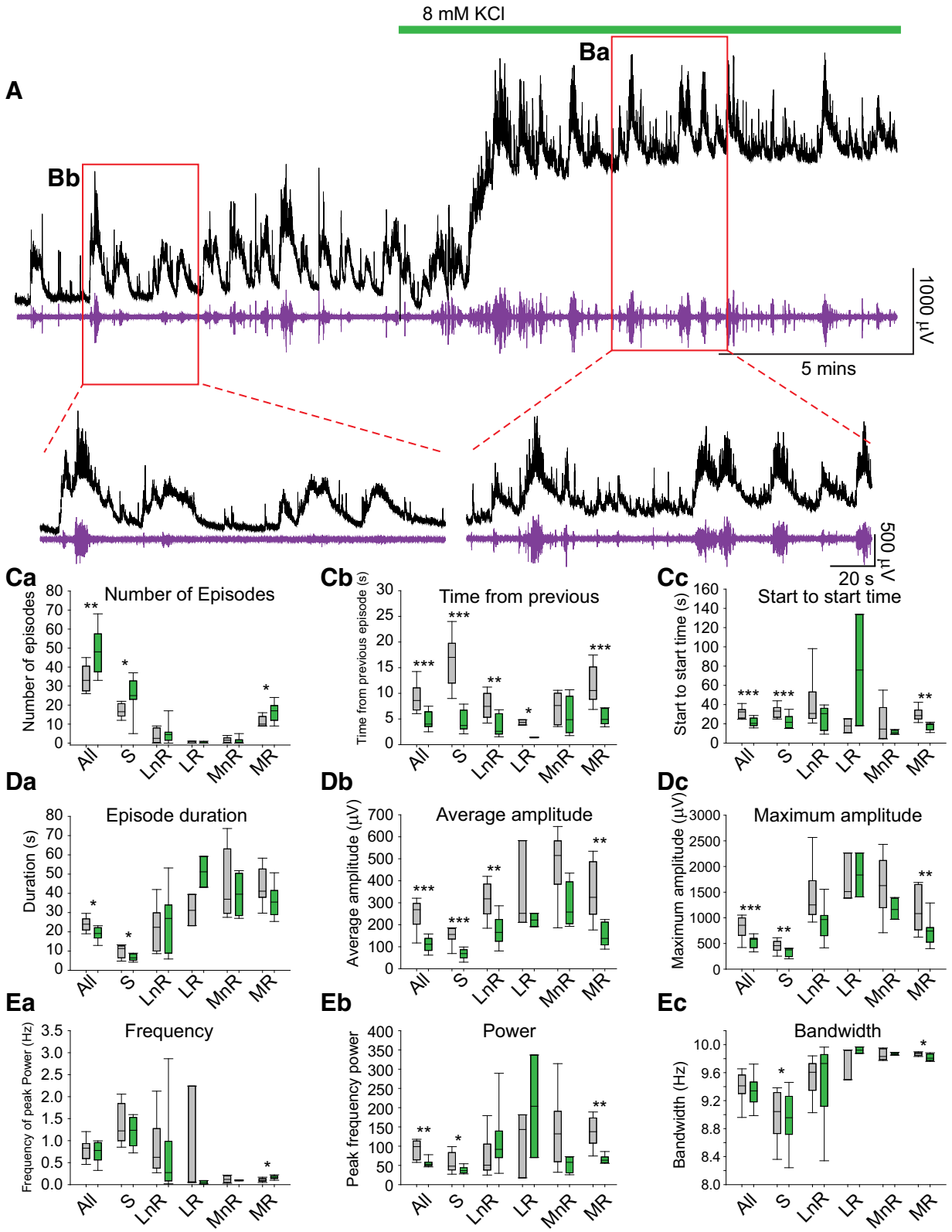

Fig. 6. Alterations in spontaneous episode class activity under conditions of enhanced excitability by increased KCl concentration. *A*: DC-coupled neurogram recorded from lumbar ventral roots at baseline and after increasing the concentration of KCl from 4 to 8 mM (green bar). The traditionally analyzed high-pass filtered (100 Hz) neurogram is below the DC-coupled recording (purple). *B*: enlargement of regions at baseline (*Ba*) and at the higher concentration of KCl (*Bb*). The addition of KCl increased the excitation of network activity, depolarizing the DC potential, and altered several features extracted from each episode (*C–E*). Episode features were compared overall and for each of the five episode classes [small (S), large nonrhythmic (LnR), large rhythmic (LR), multiburst not rhythmic (MnR) or multiburst rhythmic (MR)]. Box-and-whisker plots display interquartile range (boxes), median (horizontal black lines), maximum and minimum values in data range (whiskers). Means were compared using paired *t*-tests. Asterisks denote significant differences before and after increasing KCl: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

$P = 0.008$; average amplitude, baseline: 353 μV (SD = 126), KCl: 155 μV (SD = 53); $t_{(7)} = 4.7$, $P = 0.003$].

In other instances, we observed no change in global episode features but did find changes for some features within a specific episode class. For example, global peak frequency [Fig. 6*Ea*; $t_{(7)} = 0.5$, $P = 0.6$] and bandwidth [Fig. 6*Ec*; $W = -10$, $T^{+} = 13$, $T^{-} = -23$; $P = 0.5$] were comparable before and after exciting the network; however, for the multiburst rhythmic episode class, the peak frequency increased [$t_{(7)} = 3.3$, $P = 0.01$] and its mean bandwidth decreased [$t_{(7)} = 2.7$, $P = 0.03$]. Our approach, therefore, affords the capacity to account for changes in global episode features in specific episode classes and to tease out more subtle changes in episode features within a class that might be missed when considering the population of all episodes.

In a final experiment, we tested the ability of SpontaneousClassification to detect changes in spontaneous activity from episodes recorded under novel conditions from which the algorithms were trained. Here, a different experimenter collected data and analyzed episodes of spontaneous activity from older preparations (P5; $n = 7$) outside of the age range from which the algorithms were trained (P0–P3). This experimenter was able to easily use the software to extract features and classes for episodes of activity at this time point. Visual inspection of the episodes reveal similar underlying characteristics of the activity, consistent with the descriptors that were used to create the five classes (Fig. 7*A*). Similar to our other data sets, the vast majority of detected episodes were of the small [46.5% (SD = 17.0)] and multiburst nonrhythmic classes [36.7% (SD = 7.40); Fig. 7*B*]. There was a small proportion of the rhythmic multiburst class [1.93% (SD = 2.17)]. The proportion of large nonrhythmic episodes was modest [14.0% (SD = 13.4)], and the large rhythmic class comprised the smallest proportion of total episodes [0.77% (SD = 0.97)].

DISCUSSION

In vitro preparations of neonatal rodent spinal cords are a powerful tool for understanding how spinal networks generate the rhythmic activity that enables mammals to walk (Gordon et al. 2008; Gordon and Whelan 2006; Hägglund et al. 2010). In the past, investigations involved evoking in vitro patterns of fictive locomotor-like rhythmicity consistent with walking. Although newborn rodents can air-step at birth (Jamon and Clarac 1998), the vast majority of movements observed in neonates are sporadic and ataxic. In vitro recordings of neuronal activity underlying those sporadic, ataxic movements reveal spontaneous activity, which is critical for maturation of spinal network function (Cang et al. 2005; Hanson and Landmesser 2004; Yu et al. 2004). However, the qualitatively

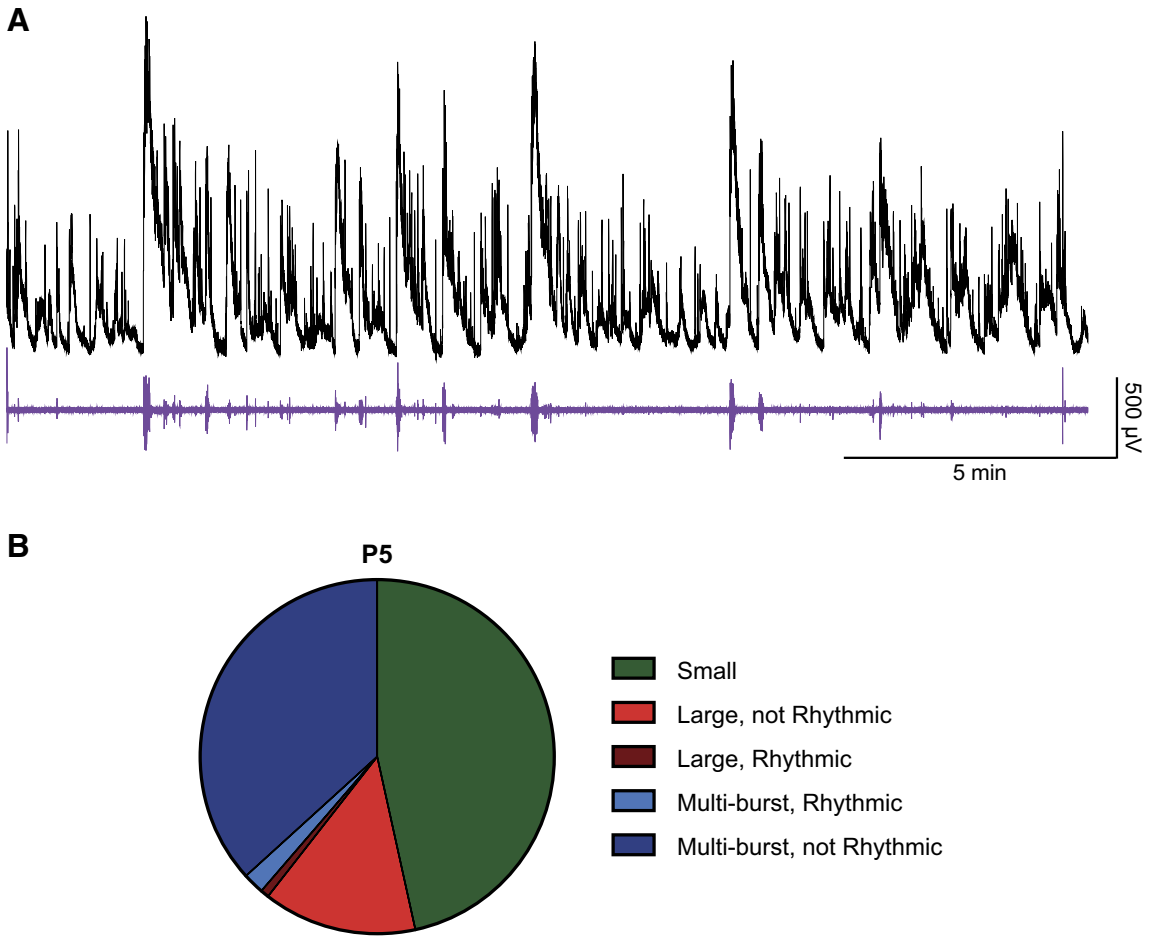

Fig. 7. SpontaneousClassification was tested on episodes recorded from the L5 ventral root of P5 spinal cords collected by a novel experimenter. *A*: representative 20 min neurograms depicting spontaneous activity from P5 L5 ventral roots. High-pass filtered (100 Hz) neurograms are presented in purple under respective raw DC neurograms. *B*: proportions of episode classes detected from spinal cords at P5. Percentages are indicated in text.

rich and diverse nature of the spontaneous activity makes it challenging to study. We developed SpontaneousClassification, a software tool to detect episodes of spontaneous activity and use supervised machine learning to identify five distinct episode classes based on quantitative and qualitative features of neuronal recordings. SpontaneousClassification enables evaluation of subtle components of DC-coupled traces by detrending without the use of high-pass filters. Furthermore, SpontaneousClassification automatically characterizes episodes of spontaneous activity based on their features and then groups them into one of five classes of episodes. This automated classification process is extremely fast, enabling large volumes of data to be analyzed quickly. Analyzing DC-coupled neurograms through machine learning classification enables researchers to investigate the nature and development of spontaneous activity in finer detail and with a new outlook than previously available techniques.

When Should This Tool Be Used?

SpontaneousClassification was designed to discriminate subtleties in stochastic, DC-coupled neurograms recorded from neonatal mouse spinal cords. Because its classifiers were trained on episodes of this specific type of recording, they are only appropriate for similarly acquired and structured data. We demonstrated its utility of analyzing episodes outside of the training conditions by classifying episodes of activity collected under varying excitation states (increased KCl) and at a later stage in development (P5). SpontaneousClassification should not be used on AC-coupled or high-pass filtered neurograms, nor on very-high-frequency episodes (i.e., $>$10 Hz). However, the feature extraction methods employed in SpontaneousClassification may be of interest for analyzing AC-coupled data. After the features are extracted, they can help determine whether the classifiers are appropriate for the data and whether the classification is working correctly. All investigators should have intimate knowledge of their data. If the class distributions seem odd considering the appearance of the data, researchers should recheck the feature extraction and consider other methods.

Classification of Other Biological Signals

A similar machine learning approach could be applied to other types of activity. Supervised machine learning has been used to detect abnormalities in the electrocardiogram characteristic waveforms to automatically diagnose arrhythmias (Gao et al. 2004; Khazaee and Ebrahimzadeh 2010; Li et al. 2014), ischemic events (Choi et al. 2017; Papadimitriou et al. 2001; Papaloukas et al. 2003), and acute coronary syndromes (Harrison and Kennedy 2005; Myers et al. 2017). It can also predict brain disorders such as epilepsy (Diambra et al. 1998; Diambra and Malta 1999; Subasi and Ismail Gursoy 2010) and schizophrenia (Sabeti et al. 2007) from electroencephalography (EEG) signals. MLPs are often used to classify electrical activity recorded from muscles (electromyography; EMG) for the control of myoelectric prostheses (Hargrove et al. 2007; Hiraiwa et al. 1992; Karlık 2014; Kelly et al. 1990).

These methods differ from spike sorting (Lewicki 1998), which uses clustering methods to separate, or sort, waveforms. Clustering in this manner is unsupervised learning; features of the waveform are compared with values defining individual clusters, sorting the waveforms into one of the clusters. Clustering methods have been used to classify AC-coupled potentials recorded from the dorsum of the lumbosacral spinal cord (Martin et al. 2015); episodes were detected based on their peak and selected a set time surrounding the peak, but this selection method is useful only when episode duration varies little. This method replaced template matching, which was previously used by the same group (Chávez et al. 2012), but they noted that it was time consuming and required constant supervision.

In the present study, episodes of activity recorded from the ventral root were too variable in shape and duration to apply template matching. Furthermore, we attempted several clustering methods and they failed to produce the classes previously defined (results not shown). Therefore, visual identification of classes and training with supervised machine learning prevailed as the most viable option for classifying episodes of spontaneous activity.

Limitations

Detrending each trace is unique to the DC drift in the recording and must be customized by the user. Selecting thresholds for episode detection is also a subjective process. After choosing threshold values to detect episodes, users can view them to ensure they were detected correctly. There may be some variability in settings between traces, especially if background excitability or noise show large fluctuations. Future work may explore a more automated detection process.

Supervised machine learning is guided by the experts classifying the data. It allows experts to relay their knowledge to an algorithm to perform tasks more quickly and automatically. Therefore, the algorithm is limited by the class labels it trains on. To reduce subjectivity as much as possible, we had three expert individuals classify episodes of activity and used those data to train the algorithm. Initial classification of the first batch of episodes resulted in 95.3% agreement between at least two of three experts. Despite this high percentage, initial training accuracy was low. This prompted the experts to refine their criteria for rhythmicity. Revising the visual criteria for rhythmicity increased the experts' confidence in labeling the episodes as rhythmic or not, as well as their classification performance. Although this introduced some subjectivity to the class labels, it was a necessary step that clarified what was considered to be rhythmic. The classification accuracy achieved with the revised criteria confirmed that the individuals and the program classified episodes similarly. This was further corroborated by the high accuracy obtained during classification of novel episodes during the testing phase, which presented the algorithms with fresh data. This demonstrated that the individuals were able to relay their expertise to the program for fast, automatic, and accurate classification.

Increasing the number of episodes in the training data set would increase the representation of each class and may help the algorithm generalize more effectively. However, the process of manual classification is time-consuming and there is no universal criterion for the number of instances of data that are needed. A balance between data availability and training accuracy must be attained.

Conclusions

This work provides and demonstrates a software tool (SpontaneousClassification) for characterizing and classifying episodes of spontaneous activity from the lumbar spinal cord. Supervised machine learning classification of episodes allowed for the analysis of DC-coupled neurograms. We demonstrated the utility of SpontaneousClassification by comparing the features and class distributions between episodes of two different activity states. The data reconfirmed that the addition of KCl nonspecifically excites spinal motor networks, resulting in a similar proportion of episodes in each class, even with an increase in overall number. This is an interesting result since it points to the fact that although the activity may appear stochastic, the overall composition is quite robust and resistant to perturbations. While it is difficult to speculate, this may be accomplished by a decrease in calcium-activated potassium conductances or other voltage-gated potassium conductances. SpontaneousClassification is easy to use and enables fast and automatic characterization of spontaneous activity. Furthermore, no knowledge of machine learning is needed to use SpontaneousClassification. Future applications for SpontaneousClassification may include characterization of spontaneous activity from other regions of the spinal cord, such as the cervical and thoracic regions, throughout different time points in development, or with the addition of neuromodulators and antagonists, as we can now investigate and analyze features of spontaneous activity quickly and automatically. This analysis may be used to further reveal the cellular components and network structure of the spinal cord during development, complementing other methods of studying spinal motor activity.

ACKNOWLEDGMENTS

The authors thank Vivian Mushahwar for valuable comments on the manuscript. We acknowledge technical support from Claude Veillette and Michelle Tran.

GRANTS

Studentships were provided from the Queen Elizabeth II Graduate Scholarship (A. N. Dalrymple), Natural Sciences and Engineering Research Council (NSERC) of Canada (S. A. Sharples), Alberta Innovates Health Solutions (S. A. Sharples), and the Hotchkiss Brain Institute (S. A. Sharples, A. P. Lognon), and the Cumming School of Medicine (A. P. Lognon). This work was supported by funding from Wings for Life (P. J. Whelan) and an NSERC Discovery Grant (210833842 to P. J. Whelan). This collaboration was facilitated by funds derived from Campus Alberta Neuroscience (A. N. Dalrymple).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

A.N.D., S.A.S., and P.J.W. conceived and designed research; S.A.A. and A.P.L. performed experiments; A.N.D., S.A.S., N.O., A.P.L., and P.J.W. analyzed data; A.N.D., S.A.S., N.O., and P.J.W. interpreted results of experiments; A.N.D., S.A.S., A.P.L., and P.J.W. prepared figures; A.N.D., S.A.S., and P.J.W. drafted manuscript; A.N.D., S.A.S., N.O., A.P.L., and P.J.W. edited and revised manuscript; A.N.D., S.A.S., N.O., A.P.L., and P.J.W. approved final version of manuscript.

REFERENCES

Amodei D, Ananthanarayanan S, Anubhai R, Bai J, Battenberg E, Case C, Casper J, Catanzaro B, Cheng Q, Chen G, Chen J, Chen J, Chen Z, Chrzanowski M, Coates A, Diamos G, Ding K, Du N, Elsen E, Engel J, Fang W, Fan L, Fougner C, Gao L, Gong C, Hannun A, Han T, Johannes L, Jiang B, Ju C, Jun B, LeGresley P, Lin L, Liu J, Liu Y, Li W, Li X, Ma D, Narang S, Ng A, Ozair S, Peng Y, Prenger R, Qian S, Quan Z, Raiman J, Rao V, Satheesh S, Seetapun D, Sengupta S, Srinet K, Sriram A, Tang H, Tang L, Wang C, Wang J, Wang K, Wang Y, Wang Z, Wang Z, Wu S, Wei L, Xiao B, Xie W, Xie Y, Yogatama D, Yuan B, Zhan J, Zhu Z. Deep Speech 2: end-to-end speech recognition in English and Mandarin. *Proc Machine Learning Res* 48: 173–182, 2016.

Burt JR, Torosdagli N, Khosravan N, RaviPrakash H, Mortazi A, Tissavirasingham F, Hussein S, Bagci U. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *Br J Radiol* 91: 20170545, 2018. doi:10.1259/bjr.20170545.

Cang J, Rentería RC, Kaneko M, Liu X, Copenhagen DR, Stryker MP. Development of precise maps in visual cortex requires patterned spontaneous activity in the retina. *Neuron* 48: 797–809, 2005. doi:10.1016/j.neuron.2005.09.015.

Chávez D, Rodríguez E, Jiménez I, Rudomin P. Changes in correlation between spontaneous activity of dorsal horn neurones lead to differential recruitment of inhibitory pathways in the cat spinal cord. *J Physiol* 590: 1563–1584, 2012. doi:10.1113/jphysiol.2011.223271.

Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 24: 361–370, 2017.

Clemens S, Hochman S. Conversion of the modulatory actions of dopamine on spinal reflexes from depression to facilitation in D3 receptor knock-out mice. *J Neurosci* 24: 11337–11345, 2004. doi:10.1523/JNEUROSCI.3698-04.2004.

Cowley KC, Schmidt BJ. Regional distribution of the locomotor pattern-generating network in the neonatal rat spinal cord. *J Neurophysiol* 77: 247–259, 1997. doi:10.1152/jn.1997.77.1.247.

Dalrymple AN, Sharples SA, Osachoff N, Lognon AP, Whelan PJ. *A Supervised Machine Learning Approach to Characterize Spinal Network Function* (Online). OSF [28 February 2019]. doi:10.17605/OSF.IO/R7SG6.

Diambra L, Capurro A, Plastino A. Neural networks that learn how to detect epileptic spikes. *Phys Lett A* 241: 61–66, 1998. doi:10.1016/S0375-9601(98)00089-9.

Diambra L, Malta CP. Nonlinear models for detecting epileptic spikes. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 59: 929–937, 1999.

Fady JC, Jamon M, Clarac F. Early olfactory-induced rhythmic limb activity in the newborn rat. *Brain Res Dev Brain Res* 108: 111–123, 1998. doi:10.1016/S0165-3806(98)00040-6.

Fellippa-Marques S, Vinay L, Clarac F. Spontaneous and locomotor-related GABAergic input onto primary afferents in the neonatal rat. *Eur J Neurosci* 12: 155–164, 2000. doi:10.1046/j.1460-9568.2000.00895.x.

Frank E, Hall MA, Witten IH. The WEKA workbench. Online appendix to *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Burlington, MA: Morgan Kaufmann, 2016. https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf.

Galli L, Maffei L. Spontaneous impulse activity of rat retinal ganglion cells in prenatal life. *Science* 242: 90–91, 1988. doi:10.1126/science.3175637.

Gao D, Madden M, Schukat M, Chambers D, Lyons G. Arrhythmia identification from ECG signals with a neural network classifier based on a Bayesian framework [Online]. In: Proceedings of the 24th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence. Citeseer, 2004. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.75.2054&rep=rep1&type=pdf.

Garaschuk O, Hanse E, Konnerth A. Developmental profile and synaptic origin of early network oscillations in the CA1 region of rat neonatal hippocampus. *J Physiol* 507: 219–236, 1998. doi:10.1111/j.1469-7793.1998.219bu.x.

Goodrich CA. Measurement of body temperature in neonatal mice. *J Appl Physiol* 43: 1102–1105, 1977. doi:10.1152/jappl.1977.43.6.1102.

Gordon IT, Dunbar MJ, Vanneste KJ, Whelan PJ. Interaction between developing spinal locomotor networks in the neonatal mouse. *J Neurophysiol* 100: 117–128, 2008. doi:10.1152/jn.00829.2007.

Gordon IT, Whelan PJ. Monoaminergic control of cauda-equina-evoked locomotion in the neonatal mouse spinal cord. *J Neurophysiol* 96: 3122–3129, 2006. doi:10.1152/jn.00606.2006.

Hägglund M, Borgius L, Dougherty KJ, Kiehn O. Activation of groups of excitatory neurons in the mammalian spinal cord or hindbrain evokes locomotion. *Nat Neurosci* 13: 246–252, 2010. doi:10.1038/nn.2482.

Hanson MG, Landmesser LT. Characterization of the circuits that generate spontaneous episodes of activity in the early embryonic mouse spinal

cord. *J Neurosci* 23: 587–600, 2003. doi:10.1523/JNEUROSCI.23-02-00587.2003.

Hanson MG, Landmesser LT. Normal patterns of spontaneous activity are required for correct motor axon guidance and the expression of specific guidance molecules. *Neuron* 43: 687–701, 2004. doi:10.1016/j.neuron.2004.08.018.

Hargrove LJ, Englehart K, Hudgins B. A comparison of surface and intramuscular myoelectric signal classification. *IEEE Trans Biomed Eng* 54: 847–853, 2007. doi:10.1109/TBME.2006.889192.

Harrison RF, Kennedy RL. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Ann Emerg Med* 46: 431–439, 2005. doi:10.1016/j.annemergmed.2004.09.012.

Hinton G, Deng L, Yu D, Dahl GE, Mohamed AR, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T, Kingsbury B. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag* 29: 82–97, 2012. doi:10.1109/MSP.2012.2205597.

Hiraiwa A, Uchida N, Shimohara K. EMG pattern recognition by neural networks for prosthetic fingers control. *Annu Rev Autom Program* 17: 73–79, 1992. doi:10.1016/S0066-4138(09)91014-X.

Jamon M, Clarac F. Early walking in the neonatal rat: a kinematic study. *Behav Neurosci* 112: 1218–1228, 1998. doi:10.1037/0735-7044.112.5.1218.

Karlık B. Machine learning algorithms for characterization of EMG signals. *Int J Inform Electron Eng* 4: 189–194, 2014. doi:10.7763/IJIEE.2014.V4.433.

Kelly MF, Parker PA, Scott RN. The application of neural networks to myoelectric signal analysis: a preliminary study. *IEEE Trans Biomed Eng* 37: 221–230, 1990. doi:10.1109/10.52324.

Khazaee A, Ebrahimzadeh A. Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. *Biomed Signal Process Control* 5: 252–263, 2010. doi:10.1016/j.bspc.2010.07.006.

Kiehn O. Decoding the organization of spinal circuits that control locomotion. *Nat Rev Neurosci* 17: 224–238, 2016. doi:10.1038/nrn.2016.9.

Kudo N, Yamada T. Morphological and physiological studies of development of the monosynaptic reflex pathway in the rat lumbar spinal cord. *J Physiol* 389: 441–459, 1987. doi:10.1113/jphysiol.1987.sp016665.

Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network* 9: R53–R78, 1998. doi:10.1088/0954-898X_9_4_001.

Li Q, Rajagopalan C, Clifford GD. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Trans Biomed Eng* 61: 1607–1613, 2014. doi:10.1109/TBME.2013.2275000.

Lu D, Popuri K, Ding GW, Balachandar R, Beg MF; Alzheimer's Disease Neuroimaging Initiative. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Sci Rep* 8: 5697, 2018. doi:10.1038/s41598-018-22871-z.

Martin M, Contreras-Hernández E, Béjar J, Esposito G, Chávez D, Glusman S, Cortés U, Rudomin P. A machine learning methodology for the selection and classification of spontaneous spinal cord dorsum potentials allows disclosure of structured (non-random) changes in neuronal connectivity induced by nociceptive stimulation. *Front Neuroinform* 9: 21, 2015. doi:10.3389/fninf.2015.00021.

Menendez de la Prida L, Bolea S, Sanchez-Andres JV. Analytical characterization of spontaneous activity evolution during hippocampal development in the rabbit. *Neurosci Lett* 218: 185–187, 1996. doi:10.1016/S0304-3940(96)13095-0.

Mentis GZ, Blivis D, Liu W, Drobac E, Crowder ME, Kong L, Alvarez FJ, Sumner CJ, O'Donovan MJ. Early functional impairment of sensory-motor connectivity in a mouse model of spinal muscular atrophy. *Neuron* 69: 453–467, 2011. doi:10.1016/j.neuron.2010.12.032.

Minsky M, Papert S. *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press, 1969.

Mooney R, Penn AA, Gallego R, Shatz CJ. Thalamic relay of spontaneous retinal activity prior to vision. *Neuron* 17: 863–874, 1996. doi:10.1016/S0896-6273(00)80218-4.

Myers PD, Scirica BM, Stultz CM. Machine learning improves risk stratification after acute coronary syndrome. *Sci Rep* 7: 12692, 2017. doi:10.1038/s41598-017-12951-x.

O'Donovan MJ. In vitro methods for the analysis of motor function in the developing spinal cord of the chick embryo. *Med Sci Sports Exerc* 19, *Suppl*: S130–S133, 1987.

O'Donovan MJ, Landmesser L. The development of hindlimb motor activity studied in the isolated spinal cord of the chick embryo. *J Neurosci* 7: 3256–3264, 1987. doi:10.1523/JNEUROSCI.07-10-03256.1987.

Pangratz-Fuehrer S, Sieghart W, Rudolph U, Parada I, Huguenard JR. Early postnatal switch in $GABA_A$ receptor α-subunits in the reticular thalamic nucleus. *J Neurophysiol* 115: 1183–1195, 2016. doi:10.1152/jn.00905.2015.

Papadimitriou S, Mavroudi S, Vladutu L, Bezerianos A. Ischemia detection with a self-organizing map supplemented by supervised learning. *IEEE Trans Neural Netw* 12: 503–515, 2001. doi:10.1109/72.925554.

Papaloukas C, Fotiadis DI, Likas A, Michalis LK. Automated methods for ischemia detection in long duration ECGs. *Cardiovasc Rev Rep* 24: 313–319, 2003.

Parkhi OM, Vedaldi A, Zisserman A. Deep face recognition. In: *Proceedings of the British Machine Vision Conference (BMVC)*, edited by Xie X, Jones MW, Tam GKL. Durham, UK: BMVA Press, 2015, p. -41.1–41.12. 10.5244/C.29.41.

Sabeti M, Sadreddini MH, Nezhad JT. EEG signal classification using an association rule-based classifier. 2007 IEEE International Conference on Signal Processing and Communications. Dubai, United Arab Emirates, November 24–27, 2007, p. 620–623. doi:10.1109/ICSPC.2007.4728395.

Sernagor E, Grzywacz NM. Emergence of complex receptive field properties of ganglion cells in the developing turtle retina. *J Neurophysiol* 73: 1355–1364, 1995. doi:10.1152/jn.1995.73.4.1355.

Sharples SA, Osachoff N, Whelan PJ. A new method for detection of spontaneous rhythmic activity from ventral roots of the neonatal mouse isolated spinal cord (Abstract). Neuroscience 2014. Washington, DC, November 15–19, 2014.

Sharples SA, Whelan PJ. Modulation of rhythmic activity in mammalian spinal networks is dependent on excitability state. *eNeuro* 4: ENEURO.0368-16.2017, 2017. doi:10.1523/ENEURO.0368-16.2017.

Sipilä ST, Schuchmann S, Voipio J, Yamada J, Kaila K. The cation-chloride cotransporter NKCC1 promotes sharp waves in the neonatal rat hippocampus. *J Physiol* 573: 765–773, 2006. doi:10.1113/jphysiol.2006.107086.

Smith JC, Feldman JL. In vitro brainstem-spinal cord preparations for study of motor systems for mammalian respiration and locomotion. *J Neurosci Methods* 21: 321–333, 1987. doi:10.1016/0165-0270(87)90126-9.

Song Y, Zhang Y-D, Yan X, Liu H, Zhou M, Hu B, Yang G. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. *J Magn Reson Imaging* 48: 1570–1577, 2018. doi:10.1002/jmri.26047.

Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst Appl* 37: 8659–8666, 2010. doi:10.1016/j.eswa.2010.06.065.

Sun Y, Liang D, Wang X, Tang X. *DeepID3: face recognition with very deep neural networks* (Preprint). *arXiv* 1502.00873, 2015.

Torborg CL, Feller MB. Spontaneous patterned retinal activity and the refinement of retinal projections. *Prog Neurobiol* 76: 213–235, 2005. doi:10.1016/j.pneurobio.2005.09.002.

Watt AJ, Cuntz H, Mori M, Nusser Z, Sjöström PJ, Häusser M. Traveling waves in developing cerebellar cortex mediated by asymmetrical Purkinje cell connectivity. *Nat Neurosci* 12: 463–473, 2009. doi:10.1038/nn.2285..

Whelan P, Bonnot A, O'Donovan MJ. Properties of rhythmic activity generated by the isolated spinal cord of the neonatal mouse. *J Neurophysiol* 84: 2821–2833, 2000. doi:10.1152/jn.2000.84.6.2821.

Witten IH, Frank E, Hall MA, Pal CJ. *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Burlington, MA: Morgan Kaufmann, 2017.

Wong RO, Chernjavsky A, Smith SJ, Shatz CJ. Early functional neural networks in the developing retina. *Nature* 374: 716–718, 1995. doi:10.1038/374716a0.

Yu CR, Power J, Barnea G, O'Donnell S, Brown HEV, Osborne J, Axel R, Gogos JA. Spontaneous neural activity is required for the establishment and maintenance of the olfactory sensory map. *Neuron* 42: 553–566, 2004. doi:10.1016/S0896-6273(04)00224-7.