Efficient Feature Extraction and Classification Methods in Neural Interfaces

Mahsa Shoaran, Benyamin A. Haghi, Masoud Farivar, and Azita Emami

California Institute of Technology

Brain disorders such as dementia, epilepsy, migraine and autism are a major concern of the modern world. Despite the heavy economic burden they impose on the society, many of these disorders remain largely undertreated. Within this context, neural devices are increasingly being utilized to interface with brain, monitor and detect the neurological abnormalities, and trigger an appropriate type of therapy such as neuromodulation to restore normal function. A key challenge to enable such new treatments is to integrate state-of-the-art signal acquisition techniques, as well as efficient biomarker extraction and classification methods to accurately identify symptoms of various disorders, using low-cost, highly integrated, wireless and miniaturized devices.

OVERVIEW: THERAPEUTIC NEURAL DEVICES

The general block diagram of a closed-loop neural interface system is shown in Figure 1. The neural signals recorded by an array of electrodes (e.g. intracranial, scalp, or other types) are initially amplified, filtered and digitized. A feature extraction processor is subsequently activated to extract the disease-associated biomarkers. Upon abnormality detection, a programmable neural stimulator is triggered to suppress the symptoms of disease (e.g. a seizure or migraine attack, Parkinson's tremor, memory dysfunction, etc.) through periodic charge delivery to the tissue. In addition to high sensitivity (*true positive rate*), sufficient specificity (*true negative rate*) and low latency, the abnormality detector device has to satisfy the safety, portability and biocompatibility requirements of the human body.

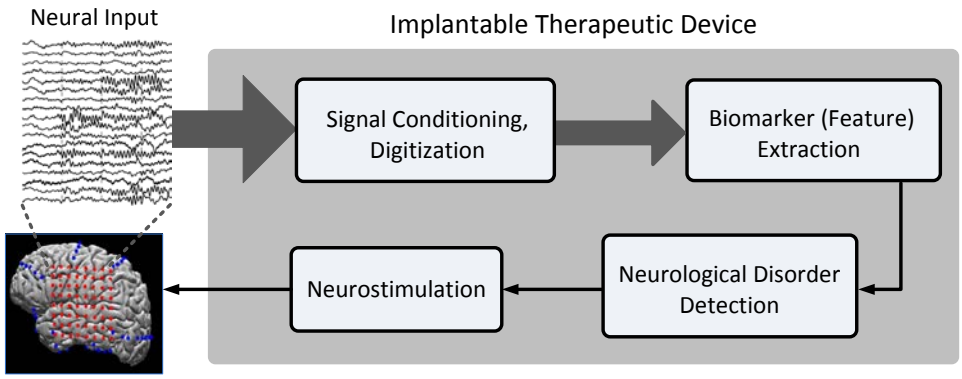

FIGURE 1 General block diagram of a closed-loop therapeutic system for detection and suppression of disabling neurological symptoms.

Current and Future Directions

The emerging field of neuroengineering uses engineering technologies to investigate and treat neurological diseases. Epilepsy has been one of the primary targets, along with movement disorders, stroke, chronic pain, affective disorders and paralysis (Stacey and Litt 2008).

Approximately one third of epileptic patients exhibit seizures that are not controlled by medications. Neuromodulation offers a new avenue of treatment for intractable epilepsy. Over decades, research on epilepsy has led to fundamental understandings of brain function, with strong implications for other neurological disorders. In addition, due to severity of refractory epilepsy and the need for surgery, human tissue and epileptic EEG datasets are largely available. Therefore, the majority of therapeutic neural interfaces in literature have focused on extracting epileptic biomarkers for automated seizure detection (Shoeb et al. 2004; Verma et al. 2010; Shoaran et al. 2015).

Spectral energy of neural channels in multiple frequency bands, as well as various time and frequency domain features have been used as potential seizure biomarkers. To improve the power and area efficiency in multichannel systems, a spatial filtering technique was proposed to precede the seizure detection unit (Shoaran et al. 2016). However, in the majority of existing devices, the classification of neural features is performed either remotely, or by means of moderately accurate thresholding techniques. A patient-specific Support Vector Machine (SVM) classifier was

implemented by Yoo et al. (2013), where the classification processor contributes to a significant portion of chip area and power. To improve the accuracy of detection, resource-efficient on-chip learning is becoming an essential element of next generation implantable and wearable diagnostic devices.

MACHINE LEARNING IN NEURAL DEVICES

Conventional classification techniques such as SVMs and k-Nearest Neighbors (KNNs) are hardware intensive and require high processing power and large memory units to perform complex computations on chip (Figure 2).

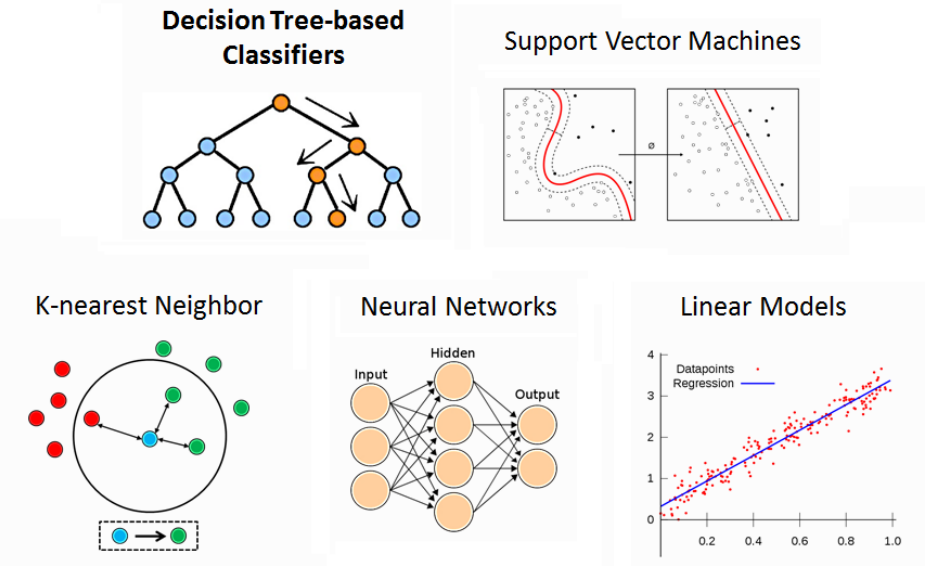

FIGURE 2 Schematic of common learning models as potential candidates for hardware implementation.

Scalability Challenges

Numerous studies show that a large number of acquisition channels are required to obtain an accurate representation of brain activity, and the therapeutic potential of neural devices is limited at low spatiotemporal resolution. Thus, it is expected that future interfaces integrate thousands of channels at relatively high sampling rates, making it crucial to operate at extremely low power. A small size of device is also critical to minimize implantation issues.

Despite a substantial literature on machine learning, hardware-friendly implementation of such techniques is not sufficiently addressed. Indeed, even the simple arithmetic operations performed in conventional classification methods can become very costly with increasing number of channels. In addition, filter banks and in general, feature extraction units can be hardware intensive, particularly at higher frequencies associated with intracranial EEG. Extensive system-level design improvement is needed to meet the requirements of implantable device, while preserving the high-resolution recording capability.

DECISION TREE-BASED CLASSIFIERS

We present and evaluate a seizure detection algorithm using an ensemble of decision tree classifiers. The general schematic of a single decision tree (DT) is shown in Figure 2. With only simple comparators as their core building blocks, such classifiers are a preferable solution to reduce hardware design complexity. Employing a gradient-boosted ensemble of decision trees, we achieve a reasonable trade-off between detection accuracy and implementation cost.

Gradient-boosting (Friedman 2001) is one of the most successful machine learning techniques that exploits Gradient-based optimization and boosting, by adaptively combining many simple models to get an improved predictive performance. Binary split decision trees are commonly used as the "weak" learners. Boosted trees are at the core of the state-of-the-art solutions in a variety of learning domains, given their excellent accuracy, fast computation and operation.

Combined with an efficient feature extraction model, we show that these classifiers quickly become competitive with more complex learning models, with only a small number of low-depth "shallow" trees (Shoaran et al. 2016). These ensembles of axis-parallel DT classifiers are excellent candidates for on-chip integration, by eliminating the multiplication operation and offering significant power and area savings.

Performance Evaluation and Hardware Design

As a benchmark, we consider a boosted ensemble of 8 trees with depth of 3 and compare it to linear SVM, cubic SVM and KNN-3 models proposed for on-chip classification, using the following set of features: line-length, time-domain variance, and multiple band powers. The proposed approach is tested on a large dataset of over 140 days of intracranial EEG data from 23 epileptic patients. Figure 3 (left) shows the average F1-measure of classifiers. This benchmark is already competitive with its peers, and it can outperform with larger ensemble sizes. It achieves an average seizure detection sensitivity of 98.3%.

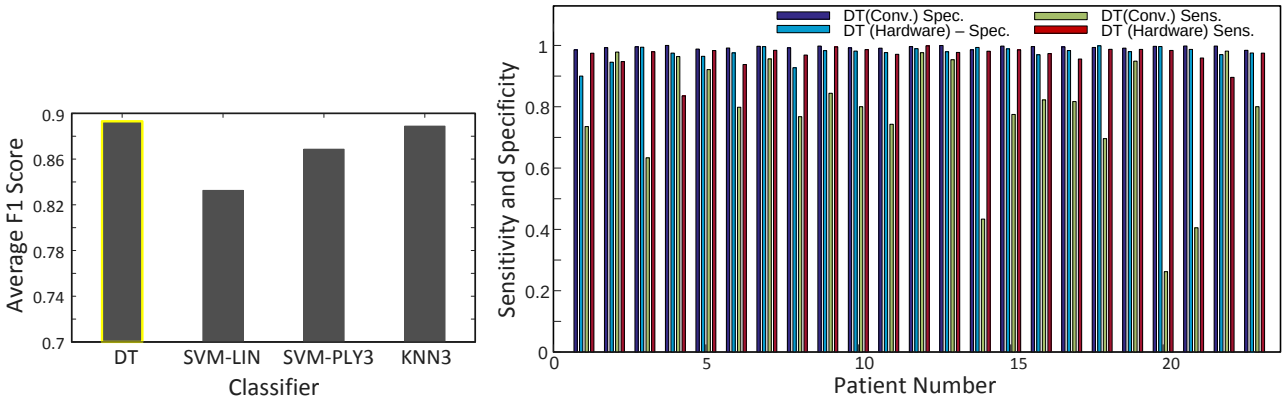

FIGURE 3 Comparison of predictive ability of different classification methods with an ensemble of 8 trees of depth 3 (left), and the classification performance of the asynchronous hardware model compared to conventional DT (right).

Decision trees are very efficient, but also susceptible to overfitting in problems with high feature space dimensionality. To address this, we limit the number of nodes in each tree, i.e., design shallow trees using small number of features. Shorter trees are also more efficient in hardware and equally important, incur less detection delay. In our simulations, the detection accuracy is not significantly improved (<0.5%) with depth values of 4 and higher.

We propose the architecture shown in Figure 4 to implement ensembles of decision trees. At each comparison step, only the features appearing in the active nodes of trees are needed, thus the rest of recording array can be switched off to save power. Since the final decision is made upon completing

decisions in prior levels, one single feature extraction unit can be sequentially used per tree. This results in a significant hardware saving, in contrast to SVM that requires all features from the entire array. For example, the memory required to classify 32-channel neural data with 8 trees, a maximum depth of 3 and threshold resolution of 8 bits, is as low as 100 bytes, while SVM and KNN-based arrays would need over 500kB of memory to classify this data. Upon training for a specific patient and depending on the difficulty of detection task, additional “supportive” trees can be used to further boost the classification accuracy.

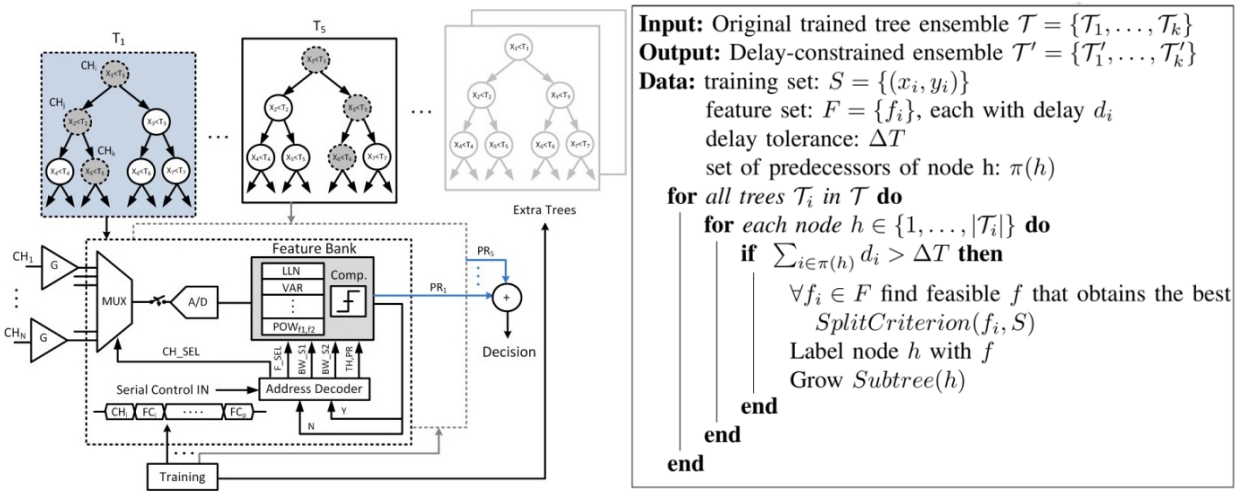

FIGURE 4 Hardware-level architecture for an ensemble of decision tree classifier with primary and supportive trees (left) and a greedy training algorithm to meet the delay constraints (right).

The proposed architecture faces a practical challenge of designing decision trees under application-specific delay constraints, that is, given any ensemble $\tau = \{\tau_1, \cdots, \tau_k\}$ of decision trees obtained from our original method, we need to ensure that each tree τ_i satisfies the delay constraint: $\sum_{i \in \pi(h)} d_i \leq \Delta T$, where d_i is the time required to compute feature f_i, ΔT is the maximum tolerable detection delay, and $\pi(h)$ is the set of all predecessors of node h. We propose a ``greedy'' algorithm to solve this practical constraint, by building trees that satisfy the delay requirement (Figure 4). However, this algorithm may result in a sub-optimal solution. We further investigate a novel asynchronous model to learn from neural data streams, the results of which are shown in

Figure 3 (right). Here, the trees are built with features that maximize the accuracy, regardless of their computational delay. Based on averaged results of completed trees and previous results of incomplete trees, decisions are frequently updated over 0.5-sec intervals to avoid long latencies and maximize the sensitivity. Once completed, longer trees contribute to a decision at future time steps.

CONCLUSIONS

Based on a simple, yet sufficiently accurate decision tree model, we introduce efficient hardware architectures and related training algorithms to predict the abnormal neurological states in various disorders, such as epilepsy, Parkinson's disease and migraine. Such classifiers would potentially allow the full integration of processing circuitry with the sensor array in various resource-constrained biomedical applications.

REFERENCES

Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics. pp.1189−1232.

Shoaran M, Shahshahani M, Farivar M, Almajano J, Shahshahani A, Schmid A, Bragin A, Leblebici Y, Emami A. 2016. A 16-Channel 1.1mm^2 Implantable Seizure Control SoC with Sub-μW/Channel Consumption and Closed-Loop Stimulation in 0.18μm CMOS. IEEE Symposium on VLSI Circuits (VLSIC), HI.

Shoaran M, Farivar M, Emami A. 2016. Hardware-Friendly Seizure Detection with a Boosted Ensemble of Shallow Decision Trees. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1826−1829, Orlando.

Shoaran M, Pollo C, Schindler K, Schmid A. 2015. A Fully-Integrated IC with 0.85μW/Channel Consumption for Epileptic iEEG Detection. IEEE Transactions on Circuits and Systems II: Express Briefs (TCAS-II). 62(2): 114−118.

Shoeb A, Edwards H, Connolly J, Bourgeois B, Treves ST, Guttag J. 2004. Patient-specific Seizure Onset Detection. Epilepsy & Behavior. 5(4): 483−498.

Stacey WC and Litt B. 2008. Technology Insight: Neuroengineering and Epilepsy—Designing Devices for Seizure Control. Nature Clinical Practice Neurology. 4(4): 190−201.

Verma N, Shoeb A, Bohorquez J, Dawson J, Guttag J, Chandrakasan AP. 2010. A Micro-Power EEG Acquisition SoC with Integrated Feature Extraction Processor for a Chronic Seizure Detection System. IEEE J. Solid-State Circuits. 45: 804−16.

Yoo J, Yan L, El-Damak D, Altaf MAB, Shoeb A, Chandrakasan AP. 2013. An 8-Channel Scalable EEG Acquisition SoC with Patient-Specific Seizure Classification and Recording Processor. IEEE J. Solid-State Circuits. 48: 214−28.