Printed Machine Learning Classifiers

Muhammad Husnain Mubarik[*$] ,Dennis D. Weller[*†], Nathaniel Bleier[$], Matthew Tomei[$]
Jasmin Aghassi-Hagmann[§], Mehdi B. Tahoori[†] and Rakesh Kumar[$]

[$]University of Illinois Urbana-Champaign, [†]Karlsruhe Institute of Technology, [§]University of Applied Sciences Offenburg

Abstract—A large number of application domains have requirements on cost, conformity, and non-toxicity that silicon-based computing systems cannot meet, but that may be met by printed electronics. For several of these domains, a typical computational task to be performed is classification. In this work, we explore the hardware cost of inference engines for popular classification algorithms (Multi-Layer Perceptrons, Support Vector Machines (SVMs), Logistic Regression, Random Forests and Binary Decision Trees) in EGT and CNT-TFT printed technologies and determine that Decision Trees and SVMs provide a good balance between accuracy and cost. We evaluate conventional Decision Tree and SVM architectures in these technologies and conclude that their area and power overhead must be reduced. We explore, through SPICE and gate-level hardware simulations and multiple working prototypes, several classifier architectures that exploit the unique cost and implementation tradeoffs in printed technologies - a) *Bespoke printed classifiers* that are customized to a model generated for a given application using specific training datasets, b) *Lookup-based printed classifiers* where key hardware computations are replaced by lookup tables, and c) *Analog printed classifiers* where some classifier components are replaced by their analog equivalents. Our evaluations show that bespoke implementation of EGT printed Decision Trees has $48.9\times$ lower area (average) and $75.6\times$ lower power (average) than their conventional equivalents; corresponding benefits for bespoke SVMs are $12.8\times$ and $12.7\times$ respectively. Lookup-based Decision Trees outperform their non-lookup bespoke equivalents by 38% and 70%; lookup-based SVMs are better by 8% and 0.6%. Analog printed Decision Trees provide $437\times$ area and $27\times$ power benefits over digital bespoke counterparts; analog SVMs yield $490\times$ area and $12\times$ power improvements. Our results and prototypes demonstrate feasibility of fabricating and deploying battery and self-powered printed classifiers in the application domains of interest.

Index Terms—printed electronics, machine learning

I. INTRODUCTION

While the impact of computing appears to be ubiquitous in today's society and economy, a large number of important domains are still minimally touched. Consider the over 10-trillion dollar fast-moving consumer goods (FMCG) market [47], for example. Disposables such as packaged foods, beverages, toiletries, over-the-counter drugs, and other consumables are sold largely without any embedded computing devices that could help with identification and tracking [69] (*is this today's pill?*), quality monitoring [14] (*is this milk bad?*), brand authentication [56] (*is this apple Golden Delicious?*), or interactivity [72] (*is my beer at the temperature I like?*).

The primary reason why FMCG domains (as well as domains such as low-end healthcare (e.g., bandages and wound dressings), agriculture [13], and environment [3]) have not seen much penetration of computing is the cost limitation of today's silicon-based computing systems. Silicon-based systems continue to cost much more than the cost requirements of these domains. For example, item-level tagging of several FMCG products - consider apples, milk, soda bottles, and bandages - must have sub-cent costs [64] (i.e., equivalent to the cost of a barcode [64]) that silicon-based systems cannot meet due to the high manufacturing, testing, and assembly costs of such systems [46]; even the cheapest microcontrollers and RFIDs cost several cents [64]. Many of the above domains also have stretchability, porosity, non-toxicity, and flexibility requirements that silicon-based systems cannot meet [41].

Low voltage printed electronics [27] has emerged as promising technology to target such application domains. Printing technologies often rely on maskless [76], portable [44], and additive [27] manufacturing methods which can greatly reduce costs and production timelines [28]. Such technologies also lead to devices that are conformable [36] and non-toxic [49]. Furthermore, recently developed printed technologies (e.g., EGT [26]) are low-voltage, allowing them to be battery-powered or potentially self-powered [5] when used in context of the above applications.

In this paper, we focus on printed machine learning (ML) classifiers. A large number of printed applications may need to make classification decisions in the field. For example, a printed smart wound dressing [48] may be used to determine if a wound has healed. A printed in-situ sensor [72] may determine if a packaged food item has expired. A printed pulse oximeter [45] may determine when the oxygenation or pulse rate levels are abnormal. Prior work has not explored the design space of printed machine learning classifier architectures for any learning algorithm. This is not surprising since such exploration relies on design tools which require process design kits (PDKs) for printed technologies. Such PDKs have just started becoming available [66], [80] as the technologies have begun to mature, making developing such PDKs worthwhile.

In this research, we perform an exploration of low-cost classifier architectures for printed technologies (Fig. 1) using recently developed EGT and CNT-TFT PDKs [10]. Our exploration yields several interesting observations. First, since printed technologies have orders of magnitude larger feature sizes than state-of-the-art CMOS, the circuits designed and fabricated in printed technologies have significantly worse area and power characteristics than silicon counterparts (Table I). Therefore, simple classification algorithms and models that can be implemented at low gate count are strongly favored (e.g., Decision Trees and Support Vectors Machines (SVMs)). Second, since both non-recurring engineering (NRE) costs and per unit-area fabrication costs in printed technology are low

* Equal contribution

[31], highly custom bespoke classifiers - classifier architectures that are customized to a model generated for a given application using specific training datasets (Section IV) are enabled; such degree of customization is mostly infeasible in silicon due to high NRE and fabrication costs. Third, unique implementation tradeoffs in printed technologies encourage new optimizations. For example, low ROM overhead in EGT printed technology compared to logic enables lookup-based implementation of classifiers (Section V). Similarly, the high absolute area and power overhead of digital printed architectures and the lower relative overhead of analog elements (e.g., resistors, diodes, and capacitors) compared to digital elements in printed technologies versus silicon encourage analog implementation of classifiers (Section VI). This paper makes the following contributions:

- We perform the first exploration of different classification algorithms in terms of accuracy and potential cost for two printed technologies (EGT [26] and CNT-TFT [65]). Our results show that simple classification algorithms such as Decision Trees and SVMs provide a good balance in terms of accuracy and potential hardware costs compared to more complex algorithms such as multi-layer perceptrons and logistic regression.
- We evaluate conventional Decision Tree and SVM architectures in printed technologies for a set of simple sensor-based machine learning applications. Our evaluations show that conventional classifier architectures have high area and power overheads which must be reduced.
- We develop and evaluate bespoke printed classifier architectures. We show that EGT-bespoke Decision Tree implementations have $4\times$ lower delay, $75\times$ lower power, and $48\times$ lower area (on average) than their conventional general-purpose counterparts. Corresponding benefits for bespoke SVM implementations are $1.4\times$, $12.7\times$, and $12.8\times$ respectively. To the best of our knowledge, this is the first quantification of the benefits of bespoke classifiers in printed technology. We also fabricated a working EGT prototype of a bespoke Decision Tree. This is the first prototype of a Decision Tree in a printed technology.
- We develop and evaluate lookup-based printed classifier architectures, where certain logic functions (e.g., comparators in Decision Trees and MAC units in SVMs) are replaced by lookup tables. Our results show that lookup-based EGT Decision Trees improve the area of bespoke counterparts by $1.93\times$ and power by $1.65\times$, with 50% delay overhead (on average). Lookup-based SVMs, see in the best case, a 40% reduction in delay with 8% and 1% improvement in area and power respectively. To the best of our knowledge, these are the first lookup table-based implementations of machine learning classifiers for printed technology. We also fabricated a working prototype of an EGT ROM that can be used to implement lookup-based printed classifiers.
- We develop and evaluate analog printed classifiers[1] where

---

[1] Some may prefer to call these classifiers mixed-signal or semi-analog since digital elements are still present
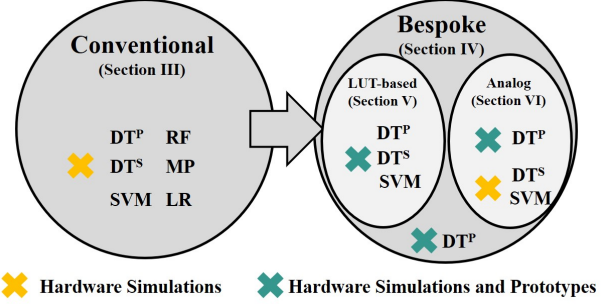

Fig. 1: An overview of this research: we explore the space of conventional and printing-specific classifier architectures through hardware simulations and prototypes. DT^s: Serial Decision Tree, DT^p: parallel Decision Tree, *SVM*: Support Vector Machine, *MLP*: Multilayer Perceptron, *LR*: Logistic Regression, *RF*: Random Forests

TABLE I: PPA analysis of common ML operations in PPDK, CNT-TFT and TSMC40nm. D,A,P: Delay, Area, Power

Components	EGT			CNT-TFT			TSMC-40nm		
	D msec	A cm2	P mW	D usec	A mm2	P mW	D nsec	A um2	P mW
Comparator	11.2	0.15	0.61	9.5	0.21	8.32	0.23	94	0.14
MAC	27	1.12	4.12	16.14	1.4	57	0.57	255	0.51
Relu	2.54	0.03	0.14	1.44	0.35	10	0.1	67	0.46

data representation and computation (comparisons for Decision Trees and MACs for SVMs) are implemented using analog logic. Our results show that EGT analog Decision Tree classifiers outperform their digital bespoke counterparts by $437\times$ and $27\times$ for area and power respectively and are $1.63\times$ slower. Corresponding area and power benefits for analog SVM architectures are $490\times$ and $1212\times$ respectively; analog SVMs are $1.36\times$ slower than bespoke counterparts. These are the first printed implementations of analog Decision Trees and SVMs and quantification of their benefits. We also fabricate and evaluate a prototype analog signal Decision Tree - this is the first prototype of an analog Decision Tree in a printed technology.

II. BACKGROUND AND RELATED WORK

Printed electronics is an emerging technology which holds promise to enable flexible [29], large-area [30], and ultra-low-cost computing systems [31] through the use of printing-based fabrication techniques such as screen [5], [37], roll-to-roll [21], [39], and inkjet [21], [25], [27], [73] printing. Inkjet printing, for example, has attracted a lot of attention, as it allows for contact-less printing on a wide range of carrier materials such as flexible substrates due to its mask-less fabrication process, where jetting of droplets is controlled by a CAD software, thus enabling digital printing [20].

Some printing methods rely purely on additive manufacturing steps, while others are based on both additive and subtractive processes. Comparable to subtractive silicon-based processes, subtractive printing processes include fabrication steps which involve the development of photoresists and subsequent etching. Due to this, the subtractive processes are relatively expensive compared to additive processes, as they demand expensive equipment and infrastructure [16]. In contrast, a fully additive

TABLE II: Accuracy and computation requirements of different classification algorithms; models generated by scikit-learn. DT-1/2/4/8: Decision Tree Classifiers with depth 1/2/4/8, RF-2/4/8: Random Forest classifiers with 4/8/16 estimators with max depth of each tree is up to 8, MLP-1: Multi-Layer Perceptron with 1 hidden layer and up to 5 hidden nodes, MLP-3: Multi-Layer Perceptron with 3 hidden layers up to 5 nodes per hidden layer, A: Accuracy on test data, #C: Number of comparisons, #M: Number of MAC operations.

	DT-1		DT-2		DT-4		DT-8		RF-2		RF-4		RF-8		MLP-1		MLP-3		LR			SVM-C			SVM-R		
	A	#C	A	#C	A	#C	A	#C	A	#C	A	#C	A	#C	A	#M	A	#M	A	#M	#C	A	#M	#C	A	#M	#C
Arrhythmia	0.56	1	0.58	3	0.62	12	0.60	46	0.52	98	0.61	166	0.65	351	0.62	1380	0.58	1430	0.63	2893	11	0.64	14k	55	0.25	263	12
Cardio	0.79	1	0.79	3	0.87	13	0.94	47	0.91	104	0.92	225	0.94	479	0.90	110	0.91	160	0.91	57	3	0.90	57	3	0.84	19	4
GasID	0.37	1	0.49	3	0.67	14	0.90	97	0.94	239	0.95	471	0.96	899	0.98	665	0.98	715	0.91	762	6	0.99	1.9k	15	0.99	127	7
HAR	0.57	1	0.82	2	0.90	4	0.99	4	0.99	146	0.99	334	0.99	812	0.99	85	0.99	69	0.94	60	5	0.99	120	10	0.90	12	6
Pendigits	0.19	1	0.32	3	0.66	15	0.92	122	0.94	266	0.96	570	0.97	1147	0.94	130	0.91	180	0.92	160	10	0.98	720	45	0.19	16	11
RedWine	0.44	1	0.47	3	0.53	14	0.51	112	0.57	169	0.56	361	0.61	745	0.57	85	0.60	19	0.56	66	6	0.55	165	15	0.57	11	7
WhiteWine	0.45	1	0.48	3	0.51	15	0.54	160	0.55	277	0.56	569	0.59	1122	0.56	90	0.55	19	0.54	66	6	0.51	165	15	0.51	11	7

printing environment requires only a simple deposition machine, where functional inks are printed layer-wise on substrates to realize transistors, passive components (such as resistors and capacitors), and interconnects. This simplification of the manufacturing chain is, however, associated with a reduction of performance (inkjet-printed EGTs versus subtractively printed CNT-TFTs, for example [10]), mainly due to larger features sizes attributed to low resolution and, hence, low-cost printing.

Irrespective of the deployed fabrication process, performance and transistor-densities of silicon-based electronics can never be reached, due to the large feature sizes in printed electronics, which are in the range of μms, compared to nms in silicon. As a consequence, printed electronics is considered a complementary technology to silicon-based circuits, with the potential to support application domains, the so-called "large-area electronics", with stringent requirements in terms of cost, conformity, or non-toxicity (Section I). Please see [28] for a much more detailed overview of printed electronics and its benefits and applications.

Work on supporting machine learning in printed technologies has been limited. Weller *et al* [77] develop an EGT neuron implementation that can be used to support both multiply-accumulate (MAC) operation as well as non-linear activation functions. Douthwaite *et al* [23] propose a time domain encoding-based implementation of a MAC engine for flexible substrates. Ling *et al* [43] discuss the possibility of implementing spiking neural networks using EGTs. Ozer *et al* [55] suggest that benefits may be possible from bespoke classifiers and describe an automated methodology to generate such classifiers; no quantification was performed. In contrast to previous work, we explore and compare several conventional and printing-specific classifier architectures in multiple low-voltage printed technologies using hardware simulations and prototypes.

Our exploration critically requires the use of printed PDKs. Several PDKs have been developed for printed technologies [66], [80]. We choose to perform our exploration using EGT and CNT-TFT technologies since a) these technologies can operate at low voltage and can, therefore, be potentially used to build battery or self-powered classification systems, and b) the corresponding PDKs and standard cell libraries are open source [10] making it easier for us to perform our exploration.

III. Choosing Classification

Algorithm for Printed Applications

The first question we attempt to answer is - *what classification algorithms can be feasibly supported in printed technologies*? The choice of the classification algorithm to support in printed hardware for a given application depends both on the characteristics of the application (which determine the accuracy for a given algorithm) and the costs of the implementation of the algorithm in hardware (e.g., power, area, and latency). We studied five different classification algorithms - Decision Trees [59] (DTs), Random Forests [12] (RFs), Multi-Layer Perceptrons [62] (MLPs), Logistic Regression (LR), and Support Vector Machines (SVMs) [11] - and evaluated them in terms of accuracy and potential hardware cost using scikit-learn over seven datasets belonging to simple machine learning applications that consume at least one sensor input and have low precision, duty cycle, and sample rate requirements. Printed [8], [19], flexible [15], wearable [52], [63], [67], [78], or RFID [6], [17], [53], [68], [74] sensors already exist for these applications making them suitable for our study. Six datasets (Arrhythmia [34], cardiotocography [9], pendigits [4], GasID [32], and RedWine/WhiteWine [18]) were chosen from the UCI Machine learning repository [24]; HAR (Human activity recognition) dataset was taken from [7]. Arrhythmia [34] classifies heart rhythms based on ECG sensor outputs. Cardiotocography (cardio) [9] classifies cardiotocograms based on several sensed values and their histograms. Pendigits [4] classifies hand-written numbers based on pen-tip pressure and location. GasID [32] uses chemical sensors to classify gasses. Red and White Wine [18] classify wines by their quality using pH and metal trace sensors. The HAR (Human Activity Recognition) dataset was taken from [7] and classifies a person's activity (walking, standing, sitting, etc) based on accelorometer outputs.

We performed pre-processing for each application dataset to remove all the non-sensor (categorical) features. Each dataset is then divided (70/30) into a training dataset and a test dataset. In the training set, all input features are normalized to have zero mean and unit variance. During training, hyperparameters are selected using scikit-learn's built-in hyperparameter search functionality (`RandomizedSearchCV`) with 5-fold cross validation and 100 iterations. MLPs, SVMs, and logistic regression are trained until convergence with the default tolerance. Once training is over, we generate classification accuracy (in Table II) on the test dataset.

Table II, in conjunction with Table I, allows us to understand the accuracy-cost tradeoffs between different classification
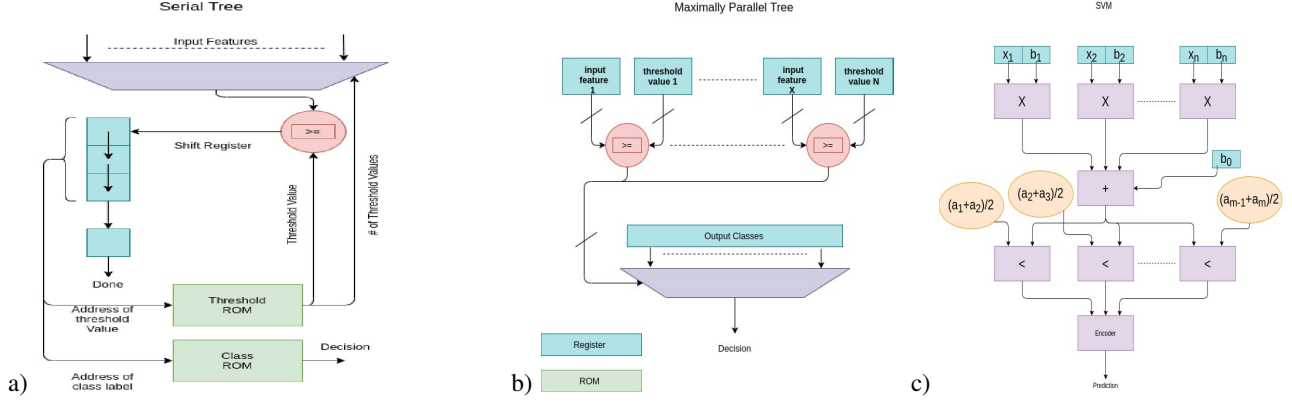
Fig. 2: **a**): Conventional architectures for Serial Trees, **b**) Maximally Parallel Trees, and **c**) Regression-based SVMs. $a_1, a_2, a_3, ..., a_n$ are class-labels.

algorithms in printed and silicon technologies. We present results for Decision Trees with depth 1, 2, 4 and 8, Random Forests with 2, 4, and 8 trees, and MLPs with 1 and 3 hidden layers and 5 nodes per hidden layer. We evaluated both classification and regression versions of SVMs (SVM-C and SVM-R respectively). SVM-C is using the one-vs-one multi-class classification strategy, so there is one binary classifier for every pair of class labels. For SVM-R, the class labels (integers) are treated as real values used to train a single SVM regressor. During inference, the output (which is a real value) is mapped to the nearest class label.

To estimate the potential hardware cost of each classifier in printed and silicon technologies, we observe that the computation in each classifier (during classification) is dominated by two operations: comparisons and two-input multiply-accumulates (MACs). The potential hardware cost of each classifier, therefore, depends on the number of these operations in the trained model and their cost of hardware implementation (Table II). We count the number of each operation in the trained model of each classifier generated by scikit-learn. To calculate the implementation cost of an operation, we implemented each operation natively in RTL and synthesized the designs with Synopsys Design Compiler using EGT, CNT-TFT, and 40nm TSMC libraries. We observe (Table I) that an EGT MAC requires $7.5\times$ more area, $6.8\times$ more power, and has $2.4\times$ higher latency relative to a comparison. Therefore, conventional classifiers with large number of MACs (e.g., MLPs, SVMs, and logistic regression) will have high overall hardware overhead. We also observe that hardware overheads in printed technologies are orders of magnitude higher than the silicon implementation. This means that a classification algorithm whose hardware implementation exceeds a certain number of MAC units and comparators may be infeasible for printed applications, even as its silicon implementation may have acceptable overheads.

Results show that the Decision Trees are clearly the lowest cost classifiers. MLPs, LR and SVM-C (SVM classification) algorithms have high hardware cost (19 to almost 2000 MAC units) making the corresponding area and power overheads (21 to $2250\,\text{cm}^2$ and 0.078 to $8.2\,\text{W}$ in EGT) likely prohibitive for most printed applications (even as the corresponding area

and power overheads in silicon - 0.004 to $0.51\,\text{mm}^2$ and 0.009 to $1\,\text{W}$ - are most likely acceptable for most applications). RFs also have high hardware cost for high accuracy goals (52 to $168\,\text{cm}^2$ and 0.21 to $0.68\,\text{W}$ in EGT for RF-8) due to a large number of comparators used. However, the cost can be scaled down at the expense of accuracy by reducing the depth of each tree in the forest. SVM-Rs have higher hardware cost than most Decision Trees, but still have much lower cost than other classifiers.

Overall, our results show that simple classification algorithms such as Decision Trees (for all applications) and SVM-Rs (for some applications) provide a good balance in terms of accuracy and estimated costs. In fact, for HAR, Decision Trees have the highest accuracy tied with more complex classifiers. SVM-Rs performs better than Decision Trees for applications such as wine quality where there are simple linear relationships between input features and class labels. More complex algorithms such as MLP, and LR have higher accuracy on average, but have high overhead in printed technologies. Random Forests may allow tunable accuracy-cost tradeoffs. However, Decision Trees are the kernel of a Random Forest ensemble; any optimization for Decision Trees is a natural optimization for Random Forests. As a result, we restrict our detailed evaluations in subsequent sections to Decision Trees and SVM-Rs.

A. Conventional Classifier Architectures in Printed Technologies

Decision Tree and SVM-R classifiers have low cost while still providing reasonable accuracy. For this reason, we choose to study these classifiers in greater depth. We implemented conventional decision tree and SVM-R classifiers in EGT and CNT-TFT to understand how well these architectures may meet the requirements of printed applications and where the cost bottlenecks and opportunities for improvement are.

1) Decision Trees: First, we consider decision trees. Complexity of the decision tree classification algorithm can be scaled by changing the number of levels in the tree and the number of comparisons at each level. For deeper trees, there exists a meaningful parallelism versus work-efficiency tradeoff. If comparisons are evaluated in serial, only the comparisons leading to the correct classification are required. However, we might want to evaluate comparisons at multiple levels in parallel

TABLE III: Conventional Serial Trees, L: Logic, M: Memory

Trees	EGT							CNT-TFT							TSMC40nm						
	Delay ms		Area mm²		Power mW		Gate Count	Delay µs		Area mm²		Power mW		Gate Count	Delay ns		Area µm²		Power mW		Gate Count
	L	M	L	M	L	M		L	M	L	M	L	M		L	M	L	M	L	M	
DT-1	18.1	2.04	19.2	0.85	1.57	0.08	59	10.2	604	0.24	0.9	27	0.16	65	0.29	100	62	0.15	0.125	62	54
DT-2	44.4	4.08	38	2	4.92	0.17	115	20	1208	0.42	2.1	94	0.33	115	0.7	150	87	0.47	0.26	124	74
DT-4	86.8	8.16	182	9.6	28	0.35	626	52	2416	2.2	9.7	472	0.66	742	1.64	250	472	2.1	1.04	249	350
DT-8	181	16.3	196	204	71	0.71	638	130	4832	2.49	207.4	968	1.34	765	3.4	450	495	34	2.48	528	353

TABLE IV: Conventional Maximally Parallel Trees, D,A,P: Delay, Area, Power

Trees	EGT				CNT-TFT				TSMC40nm			
	D ms	A cm²	P mW	Gate #	D µs	A mm²	P W	Gate #	D ns	A µm²	P mW	Gate #
DT-1	29	1.55	13.1	257	16	2.2	183	288	0.47	386	0.4	129
DT-2	37	4.2	30.6	790	43	5.9	0.18	864	0.61	1071	0.75	492
DT-4	50	20	118	4.1k	52	28.9	0.7	5.1k	0.92	4036	1.99	1.8k
DT-8	59	213	1027	49k	89	300	3.25	62k	1.25	39035	12	19k

to reduce latency. The more levels we evaluate in parallel, the more wasteful comparisons we will have to do.

We evaluate two implementations of decision trees in hardware, which correspond to the two ends of the parallelism versus wasted-work tradeoff space. The first implementation performs comparisons fully serially, so we call this the serial decision tree. The corresponding architecture diagram is shown in Fig. 2. In the serial implementation, there is a single comparator. There are two ROM memories, one for thresholds that the input features are compared to, and the other for classifications. During inference, the working node in the tree is stored in a shift register. The shift register is initialized with the value 1, and, in each subsequent cycle, the result of the current comparison is stored in the least significant bit of the shift register. Since the shift register works for any tree with a given depth, it is capable of indexing nodes that do not exist for a given (possibly unbalanced) tree. Therefore, we must either transform the value of the shift register before indexing into threshold memory or size the threshold memory assuming a full tree. We chose the latter. Once the initial set bit in the shift register reaches the most significant bit, we know inference is complete and use the stored index to look up the classification in the classification ROM.

For each decision tree depth, we generated RTL for the corresponding decision tree performing comparisons of 8-bit values. For the number of input features, we used the number of nodes in the tree or the maximum number of input features in any of our applications, whichever is smaller. For our modeling, we use standard cell libraries and ROM models from [10]. ROM models for TSMC-40nm are derived from [79]. Table III shows the latency, area, and power of logic and memory (ROM) in the decision trees for different technologies.

The results show that inference latency (logic delay + ROM delay) is high, especially for deep trees (almost 200ms for depth-8 trees!). Area and power overheads are also excessive. For example, the power requirement of EGT DT-8 exceeds the peak power produced by several printed and hybrid harvesters [40], [42] (Fig. 3). Similarly, EGT DT-4 and DT-8 cannot be powered by Blue Spark 10 mAh (2 mA peak

current) and 30 mAh (2 mA peak current) printed batteries [71]; Molex [2] 90 mAh (20 mA peak current) printed batteries have three times bigger area footprint. Analogously, high area of the serial trees has direct impact on yield, bill of materials (BOM), and fabrication throughput.

Note that DT-4 and DT-8 have similar gate count. For trees in Table III, the number of unique features is equal to $\min\{2^d - 1, 14\}$ where d is the depth of the tree and 14 is the average of the number of unique features in the datasets used. For both DT-4 and DT-8, the number of unique features is 14, which fixes the size of input multiplexers. The small difference in the gate counts of DT-4 and DT-8 is due to the increase in the width of the shift register, which is linear with the depth of the tree.

Our second implementation performs comparisons fully in parallel, so we call it a maximally parallel decision tree. The architecture of maximally parallel tree is shown in Fig. 2. For every node in the decision tree, there is a comparator and two registers. One register holds a threshold and the other holds an input feature. The result of all the comparisons is then used to select the classification using a multiplexer.

The delay, area, and power of the maximally parallel trees are shown in Table IV. These trees are, on average, 1.32× faster than the serial counterparts. However, area and power overheads are excessive when compared to serial implementations (20× bigger and consume 8.07× more power for EGT implementation). In fact, only a depth-1 EGT parallel tree can be powered either by Blue Spark or Molex printed batteries (Fig. 3); the deeper trees consume too much power to be powered by any printed battery or a printed energy harvester.

2) Support Vector Machines: We also implemented conventional inference architecture for SVM-Rs. The implementation is fully parallel. I.e., every MAC operation is assigned to its own MAC unit in hardware. The input features and coefficients (chosen equal in number to the max input features in all the datasets - 263, for arrhythmia) are stored in registers. The hardware multipliers, equal in number to the number of input features, multiply the input features with the corresponding trained coefficient. All the multiplication results are then added and mapped to the nearest class using comparators and class encoder as shown in Fig. 2.

Table V shows the delay, area and power analysis of 4, 8, 12 and 16 bit (width of input features and coefficient) SVMs implemented in EGT, CNT-TFT and TSMC40nm. Our evaluation shows that even the smallest SVM (4-bit) has an area of 85cm² and a power of 288mW. In fact, no conventional SVM can be powered by a printed battery or energy harvester (Fig. 3).
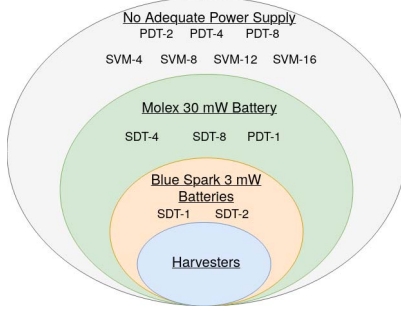
Fig. 3: Conventional Parallel (PDT) and Serial (SDT) EGT Decision Trees of depths 1, 2, 4, and 8 placed into sets based on which sources can power them: Molex [2], and Blue Spark [70], [71] printed batteries. Printed and hybrid harvesters [40], [42] cannot power any conventional EGT classifier architecture.

TABLE V: Conventional SVMs: 4,8,12 and 16 bits.

	EGT				CNT-TFT				TSMC40nm			
	D msec	A cm2	P mW	Gates #	D msec	A cm2	P W	Gates #	D nsec	A mm2	P mW	Gates #
SVM-4	85.6	85.3	288	32k	0.042	1.02	2.48	38k	1.47	0.019	6.1	9.98k
SVM-8	125	439	1424	170k	0.07	5.39	15	202k	2.07	0.09	25	46k
SVM-12	142	860	2632	252k	0.06	10	27	309k	2.4	0.19	46	95k
SVM-16	151	1445	4294	403k	0.07	17	45	624k	2.6	0.32	76	153k

Overall, our evaluations show that conventional classifier architectures have high delay, area, and power overheads when implemented in printed technologies. Subsequent sections present printing-specific architectures that are capable of reducing these overheads for several applications by multiple orders of magnitude.

IV. BESPOKE PRINTED CLASSIFIERS

Since both NRE costs and per unit-area fabrication costs in printed technology are low [31], even sub-cent [54] [2], especially for additive and mask-less technologies such as inkjet printing that may even allow portable and on-demand printing, this enables highly custom bespoke classifiers - classifier architectures that are customized to a model generated for a given application using specific training datasets - even at low to moderate volumes. Such degree of customization is mostly infeasible in lithography-based silicon technologies, especially at low to moderate volumes, due to high NRE costs (lithography equipment, material processing equipment, etc.) as well as high fabrication costs (maskset costs, etc.). This degree of customization enables reduced area and gate count designs, which further reduces marginal costs[3]. In this section, we present the first quantification of benefits of bespoke classifiers in printed technologies.

A. Bespoke Decision Trees

To generate bespoke serial decision trees (Fig. 4), we explore trees with width 4, 8, 12 and 16 bits for a given application

[2]As reference, Fujifilm Dimatix 2850 Materials inkjet printer [57] that we use to print electronics costs 50000USD and achieves sub-cent marginal cost per printed circuit when accounting for the cost of cartridges, ink, and other materials; in contrast, even older silicon foundries may cost hundreds of millions of dollars [35]

[3]Marginal costs may get even lower for higher degree of commercialization (since cartridges/inks/materials may become cheaper). Fixed costs (e.g., printer) may also decrease at high volumes.
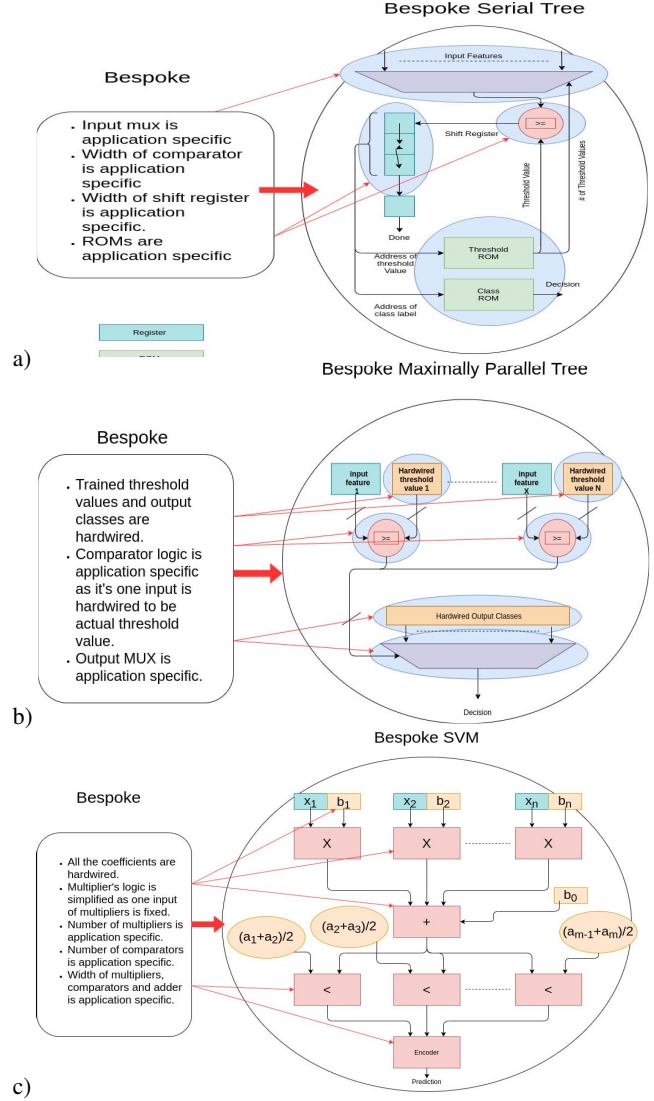

Fig. 4: Architecture of **a)**: bespoke serial trees, **b)** bespoke maximally parallel trees, and **c)** bespoke regression-based SVMs.

and use scikit-learn to calculate the accuracy corresponding to those widths. We pick the tree that gives the best accuracy for the application (up to three significant digits) with minimum hardware cost (e.g. for Arrhythmia DT-1, accuracy remains the same when we increase the classifier width from 4 to 16, hence we pick DT-1 with 4-bit comparator width). In addition, we customize the mux to have a number of inputs equal to the number of input features. We also customize the width of the shift register to be equal to the depth of the tree. Size of each entry in the Threshold ROM is customized to the width of threshold values. The number of entries in the class ROM size is customized to match the number of classes. Fig. 6 shows the delay, area, power analysis of EGT bespoke serial trees relative to conventional serial trees. EGT bespoke serial trees have 1.2%, 37%, and 22% improvements in latency, area, and power (on average). Corresponding benefits for CNT-TFT bespoke serial trees (not shown) are 1.02%, 33%, and 26% respectively.
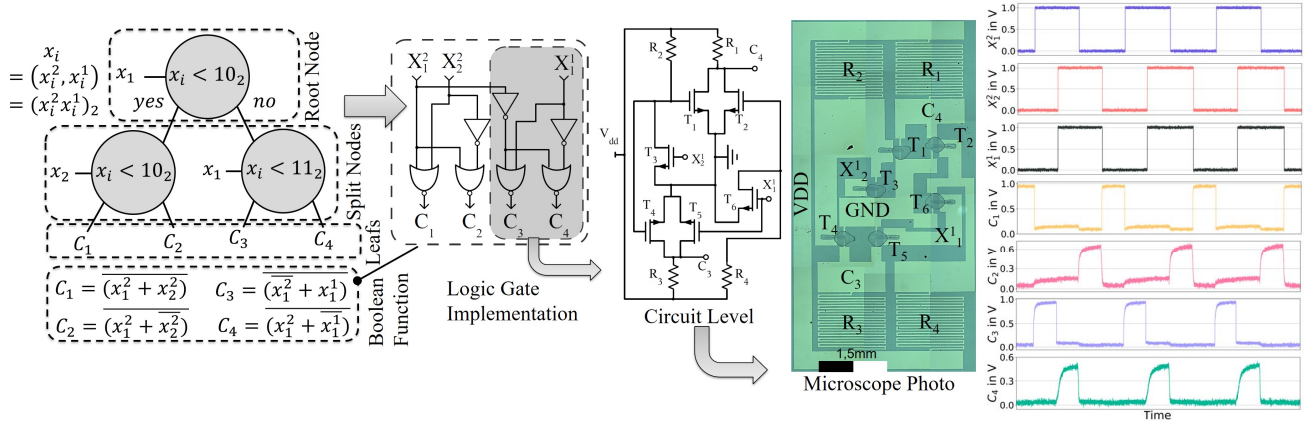
Fig. 5: **Left/Middle**: Design flow of the proposed hardware prototype of a 2-bit balanced and depth-of-2 bespoke digital Decision Tree. Top-level design of the hardwired Decision Tree is converted into logic level representation and then into transistor level circuit description. Next layout-extraction and microscope photo of the fabricated Decision Tree extracted. **Right**: Transient Measurements
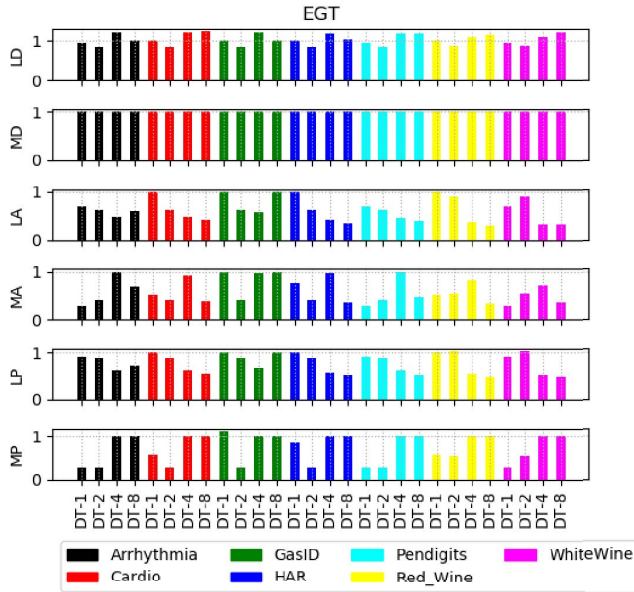

Fig. 6: Bespoke Serial Trees normalized against the conventional Serial Trees.

Fig. 7: Bespoke Maximally Parallel Trees normalized against conventional Maximally Parallel Trees.

To generate a bespoke maximally parallel tree (Fig. 4), the registers that are used for holding thresholds and inputs in the conventional architecture are removed. We train all the trees using our scikit-learn framework to get the trained threshold values which we then hardwire in the RTL, replacing the threshold registers in the conventional maximally parallel implementations. The input feature registers are replaced with connections directly to the input feature port they will use. The synthesis tool can then optimize away unnecessary gates in the design. For example, now that the actual trained threshold values are hardwired, the comparators have only one variable input which greatly simplifies overall design. We use the same methodology as bespoke serial trees to finalize the bitwidth of the tree. Results for EGT bespoke parallel trees are shown in Fig. 7. Our results show that the bespoke maximally parallel decision trees perform much better in all measures compared to their conventional counterparts. Latency, area, and
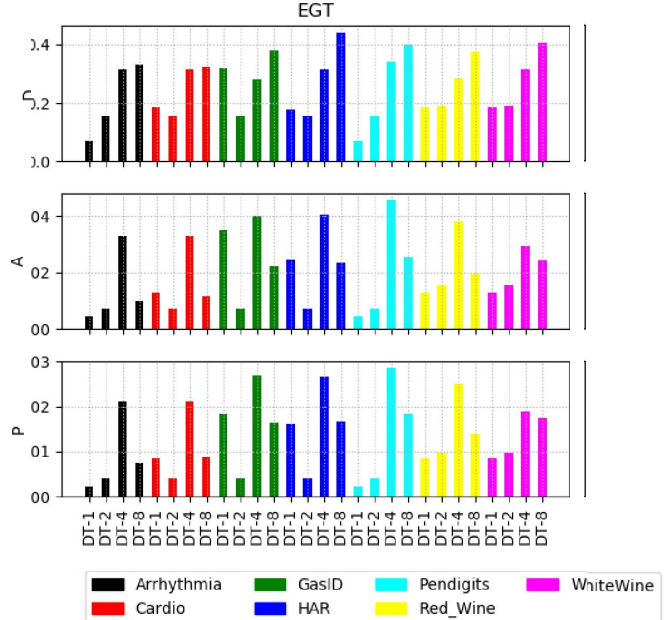
power improve by $3.9\times$, $48.9\times$, and $75.6\times$ respectively (on average). CNT-TFT bespoke maximally parallel trees yield similar benefits - latency, area, and power benefits (not shown) are $6.6\times$, $62.6\times$, and $27.3\times$ respectively (on average).

EGT parallel bespoke trees have $10.3\times$ lower area, $28.8\times$ lower power, and $9.51\times$ lower latency, on average, compared to their bespoke serial equivalents. In fact, unlike conventional counterparts, parallel bespoke trees are strictly better than serial bespoke trees.

B. Bespoke SVMs

We also developed bespoke SVM classifiers (Fig. 4). The number of input features, coefficients, multipliers, and comparators, as well the width of the registers, multipliers, adder, and comparators were fixed to the corresponding application specific values which we get after training the SVM regression models on the corresponding datasets using our scikit-learn

framework. In addition, since registers are expensive in printed technologies (a DFF is $1.41\,mm^2$, $0.018\,mm^2$, and $3.99\,\mu m^2$ in EGT, CNT-TFT, and TSMC 40nm respectively; corresponding power values are $121\,\mu W$, $77\,\mu W$, and $4.7\,\mu W$), we replace the registers with hardwired trained coefficient values. Now that the coefficients are hardwired, our multipliers have only one variable input which further optimizes the logic of hardware multipliers. Fig. 11 shows that EGT bespoke SVMs have $1.4 \times$ lower delay, $12.8 \times$ lower area, and $12.7 \times$ lower power (on average) compared to conventional SVM implementations. Corresponding benefits for CNT-TFT bespoke SVMs (not shown) are $1.7\times$, $16\times$, and $8.96\times$ respectively.

C. A Bespoke Decision Tree Prototype

Finally, to demonstrate the feasibility of printing bespoke classifiers, we designed and fabricated an EGT-based 2-bit encoded bespoke balanced and binary Decision Tree of depth 2 with threshold 10_2 (Fig. 5). Such a tree is easily transformed into a simple logic gate representation (Fig. 5). We extracted the corresponding circuit-level layout and used the Fujifilm Dimatix 2850 Materials inkjet printer for EGT-based fabrication on a ITO-sputtered glass substrate, which was structured by laser-ablation to obtain the passive conductive tracks. A microscope photo of the fabricated circuit after printing the EGTs is shown in Fig. 5.

To test operation of the prototype, all output class label pins C_1, C_2, C_3, C_4 were measured against all possible input signals of the relevant bit positions in x_1 and x_2, which are: x_1^2, x_2^2, x_1^1. As can be seen from Fig. 5, only one class label C_i is activated at the same time, in accordance to the functional description in Fig. 5. Thus the fabricated circuit is fully functional. Also, as the outputs of root and split nodes in this design have a high input impedance and low output impedance, the presented depth-2 tree can be used as a building block for building arbitrary larger trees.

To the best of knowledge, this is the first quantification of the benefits of bespoke classifiers in printed technology. Also, this is the first prototype of a digital Decision Tree in a printed technology.

V. LOOKUP-BASED PRINTED CLASSIFIERS

It is well known [22] that a computational function typically implemented using digital or analog logic can often also be implemented as a lookup table (LUT). The practicality of this approach depends on a judicious selection of the computation to replace with a LUT as well as the overheads of the technology the LUT is implemented in. In several printed technologies, including EGT, ROMs have low area and power overhead (e.g., 1-bit EGT ROM has an area of $0.05 mm^2$, while one-input inverter has an area of $0.22\ mm^2$ [10], corresponding power values are $3.13\ uW$ and $9.6\ uW$ respectively) since ROMs can be built simply as a crossbar architecture where the cross-points are shorted by printing a conductive material (such as PEDOT:PSS) to represent a bit-value [10]. Unlike silicon where ROM cells can have high delay (e.g., $900\times$ slower than the inverter cell [79]), delay of an EGT crossbar-based ROM cell

is also low (within $1.5\times$ of inverter cell [10])[4]. This opens up the possibility of implementing classifiers using lookup tables (LUTs) where computation logic (e.g., comparators in Decision Trees and multipliers in SVMs) are replaced by LUTs to reduce area and power overhead.

Two issues arise when replacing logic with ROMs. The first issue is: how much computation should be replaced with each LUT? If the amount of computation we replace with a ROM is too large, the number of ROM entries in the replacement is also large, and we may not see benefits. If the amount of computation we replace is too small, the surrounding logic (e.g., muxes and decoders) needed to access ROM can become too expensive. An exploration is needed to find the right amount of computation to replace with ROM. The second related issue is: how much reuse occurs for the surrounding logic (e.g., decoders)? For the ROM sizes that we are interested in, the decoder, for example, is expensive enough that a ROM-based comparison is always more expensive than its logic-based counterpart. The effective overhead of the decoder per lookup can be decreased, however, if the same decoder is used multiple times (i.e., if it can be shared across multiple computations). Fortunately, this is often the case with classifiers since classifiers compute on the same input feature multiple times.

A. Lookup Replacements in Decision Trees and SVMs

In serial Decision Trees, there is only one comparator to replace and none of the inputs are used simultaneously, so ROMs are not a good fit for logic replacement (the area and power of individual lookup-based comparator was $6.7\times$ and $7.3\times$ respectively compared to the non-lookup based comparator due to decoder overhead). In parallel Decision Trees, however, there are often many comparisons using the same input feature leading to signficant decoder reuse. Fig. 8 shows the architecture of lookup-based implementation of maximally parallel trees. In Fig. 9, we show the latency, area and power benefits we obtained from replacing all comparators in the EGT parallel tree implementations with lookup based equivalents (results are normalized against bespoke maximally parallel trees). In many cases, especially with shallow trees, there is not enough input feature reuse for lookup tables to be useful. But, in the best case, we see 13%, 38%, and 70% improvements in delay, area, and power. For CNT-TFT, the ROM area is typically greater than logic area ($0.05\ mm^2$ 1-bit ROM vs $0.002\ mm^2$ one-input inverter [10]), while ROM power is lower than logic power ($2.77\ \mu W$ 1-bit ROM vs $8.08\ \mu W$ one-input inverter [10]). As a result, a lookup-based parallel tree implementation provides 76.2% power benefit, on average, at the cost of increasing the area $69\times$.

For SVMs, we replaced the multipliers with ROM-based implementations (Fig. 8). However, the area and power of individual lookup-based multiplier was $1.30\times$ and $1.17\times$ respectively compared to the non-lookup based multiplier due to decoder overhead. Fig. 12 shows the latency, area, and power

[4]ROM cells also have high power in silicon; $\sim 1200\times$ the power of an inverter cell [79]
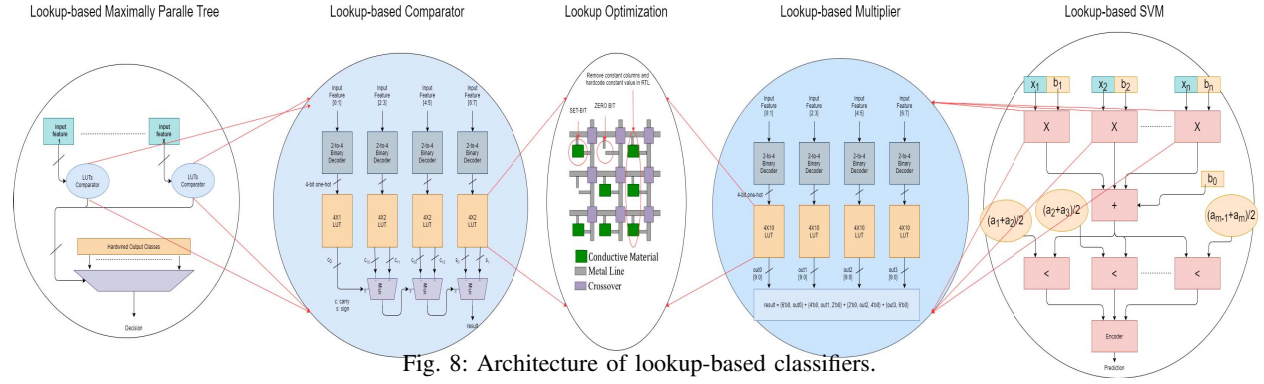
Fig. 8: Architecture of lookup-based classifiers.
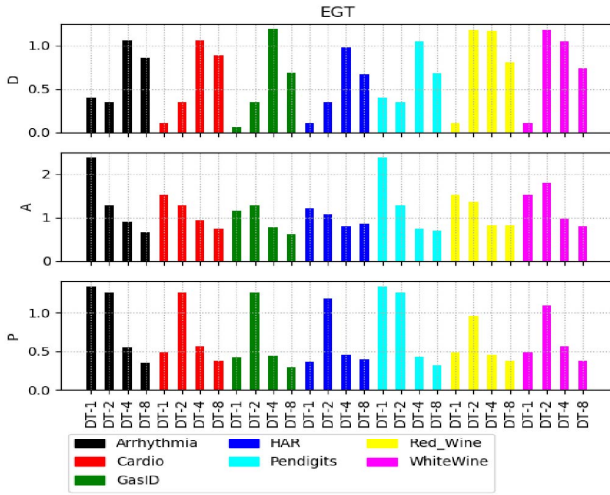

Fig. 9: Lookup-based implementation of Bespoke Maximally Parallel Trees. Normalized against Bespoke Maximally Parallel Trees.
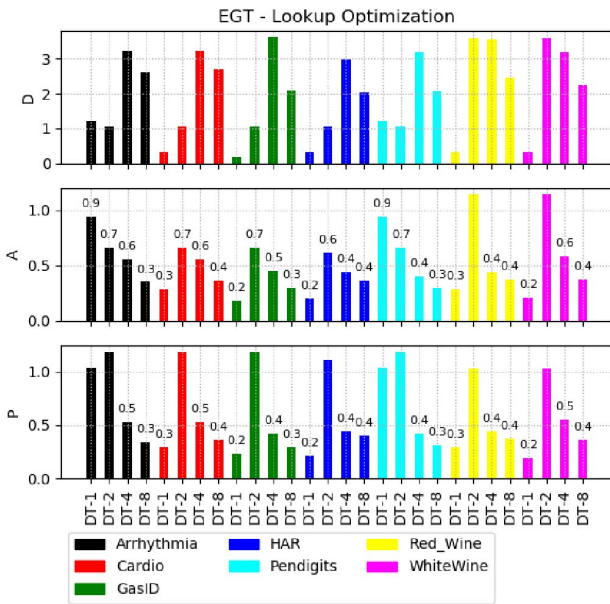

Fig. 10: Impact of additional optimizations for Lookup-based Maximally Parallel Trees. Normalized against Bespoke Maximally Parallel Trees.
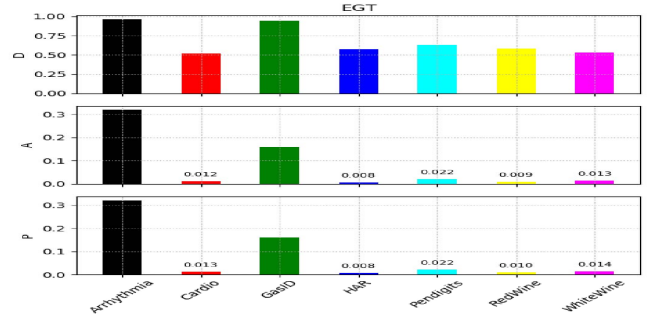

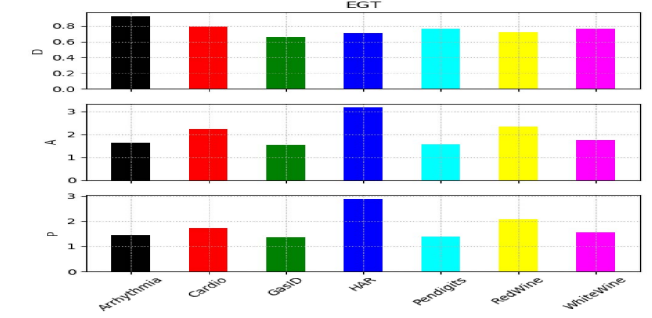Fig. 11: Bespoke SVMs normalized against conventional SVMs.

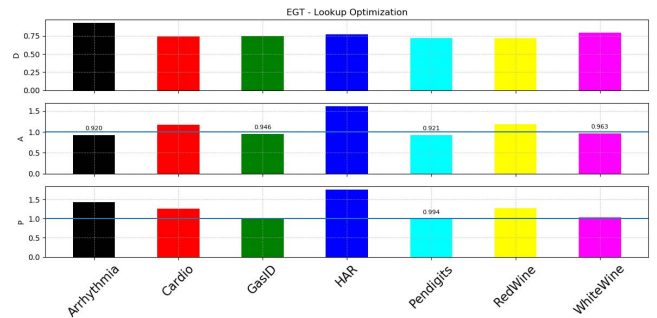Fig. 12: Lookup-based implementation of Bespoke SVMs. Normalized against Bespoke SVMs.

Fig. 13: Impact of additional optimizations for Lookup-based SVMs. Normalized against Bespoke SVMs.

of lookup-based SVMs normalized with respect to bespoke SVMs in EGT. Since every input feature is used only once, there is no decoder sharing like there was for parallel Decision Trees. As such, we do not see any benefits.

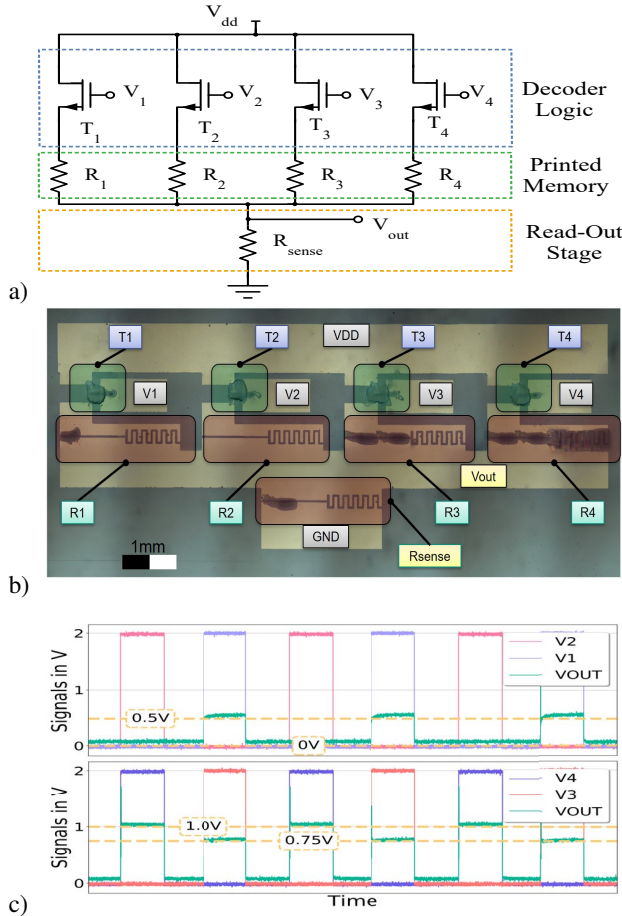Fortunately, lookup-based classifier architecture introduces

Fig. 14: **a)/b)**: Schematic/Microscope Photo of a 4x1 ROM. **c)**: Transient Measurements of the ROM.

additional printing-specific optimization opportunities. First, if certain output bits from lookup tables stay constant (i.e., all values within a column of the table are identical) - a frequent occurrence for the small lookup tables we use, the corresponding pieces of the table (e.g., the columns with identical bits) can be removed. Also, the constant is hard coded in the RTL to allow the logic synthesis tool to optimize logic downstream. Second, we can use a bespoke design for the ROM. When a bit stored in the ROM is set, we print a dot resistor at the crossover point of column and row. When a bit is clear, no dot is printed resulting in infinite resistance (Fig. 14). Since the dot resistors are much larger than a wire crossover point in the ROM, clear bits consume practically zero area, while the decoder remains unchanged.

Fig. 10 and Fig. 13 show the impact of the two optimizations. For EGT maximally parallel Decision Trees area and power benefits are increased by 52%, and 12% respectively compared to non-optimized lookup-based implementations. Perhaps more interestingly, lookup-based SVMs now see area and delay benefits for multiple applications. Note that these optimization would be mostly infeasible in silicon due to the high degree of customization.

B. Lookup Prototyping

Finally, to demonstrate feasibility of implementing lookup-based classifier architectures, we fabricated a 4x1 printed one-time programmable ROM element. The four rows of the ROM in Fig. 14 are accessed by a decoder logic block consisting of pass transistors $T_1 - T_4$, while data is stored in a resistive crossbar architecture with printed resistors $R_1 - R_4$ at the crossbar interconnects. The read signal of a read operation is obtained from the output voltage V_{out} across the sensing resistor R_{sense}. The printed ROM basically implements a voltage divider structure, with the fixed R_{sense} in the pull-down network and the variable printed resistors R_i in the pull-up network. By tuning the geometry of the printed resistors, different resistance states can be encoded which represent multiple bits of information. The chosen resistances of the printed prototype were: $R_1 = 2R_{sense}$, $R_2 = \infty$ (not printed), $R_3 = R_{sense}/2$ and $R_4 \sim 0\Omega$ (maximum resistor area).

Thus 2-bit of information could be encoded per ROM element, and thus 8-bit information for the whole 4x1 ROM. Transient measurements in Fig. 14 show data being read out successfully based on the decoded address.

The delay of the prototyped ROM element was about 10ms with an average power consumption of 39μW. The area requirement was 38mm^2. The printed prototyped ROM can be easily scaled to larger memory sizes by adding additional rows or columns and can, therefore, serve as the building block for lookup-based printed classifiers.

VI. ANALOG PRINTED CLASSIFIERS

One effective method for significantly reducing transistor count, and, therefore, overall area and power overhead, in printed classifiers could be to judiciously substitute complex logic (e.g., multi-bit comparators in Decision Trees) by small analog circuits with only few transistors [23]. For some applications, an analog classifier architecture may also allow sensor outputs to be connected directly to the classifier, avoiding costly analog-to-digital converter (and reverse).

In silicon-based (SI) classifiers, such substitutions are particularly challenging since a) it introduces additional verification and test challenges [58] [33], and b) noise and mismatch constraints force the analog devices to be large and, therefore, any area and power benefits may be lost [33]. In printed technologies, low fabrication costs allow iterative refinement to fix/reduce noise/mismatch issues. Therefore, such analog substitutions may be more feasible.

A. Analog Replacements in Decision Trees and SVMs

To build an analog Decision Tree architecture, we observe that, at each node of the Decision Tree, the binary decision can be formulated as an if-else-statement of the form: $x_k \leq \tau_j$, where τ_j is a pre-defined threshold determined by the learning phase. For an analog implementation, this binary comparison can be realized by a back-to-back inverter, which has a printed resistor in the pull-up network of one of the inverters, and a transistor in the pull-up network of the opposite inverter (see root node in Fig. 15). Moreover, the features x_k are encoded as voltage signals, and are normalized to the interval $[0V, 1V]$.
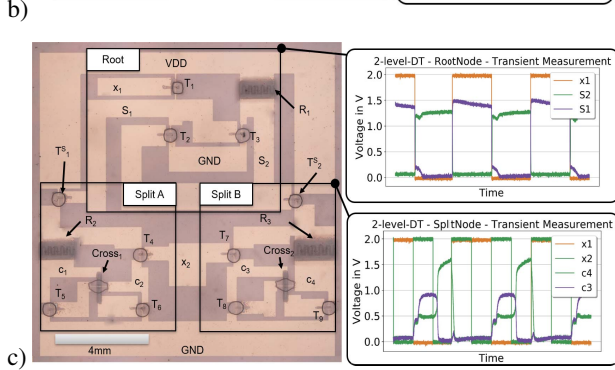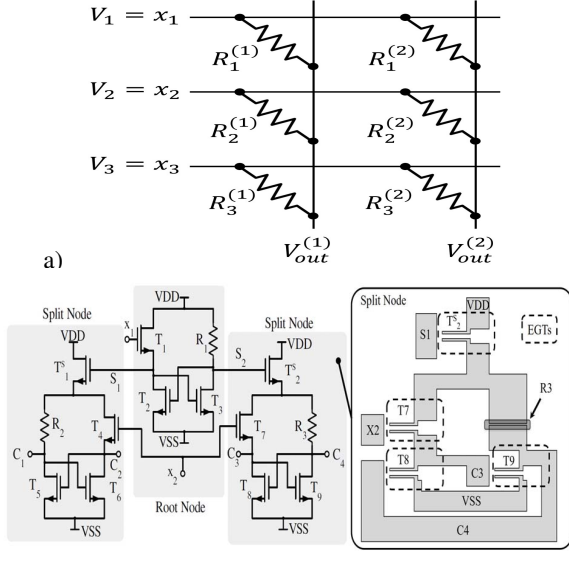
a)

b)

c)

Fig. 15: **a)**: Circuit of a MAC engine for analog SVM implementation. **b)**: Schematic and layout of a 2-level analog Decision Tree with one root and two split nodes **c)**: Digital post-processed image from multiple microscope photos of the printed 2-level analog Decision Tree with corresponding transient measurements.
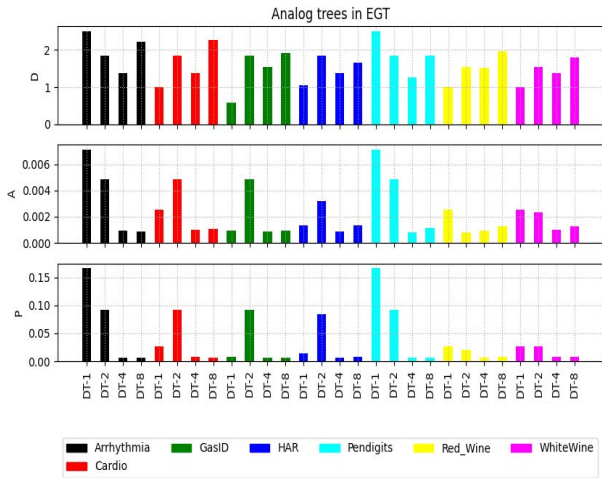

Fig. 16: Analog Trees in EGT. Tree results normalized against Bespoke Maximally Parallel Trees.

Subsequently, the input is then applied to the gate of a transistor in the pull-up network, and converted into a resistance value.
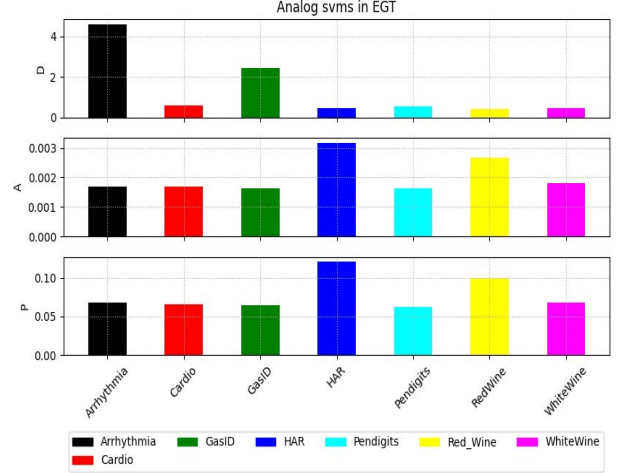

Fig. 17: Analog SVMs in EGT. SVM results normalized against Bespoke SVMs.

Dependent on the input voltage level, the equivalent transistor resistance changes its value in the range of certain On- and Off-resistances ($[R_{on}, R_{off}]$), dependent on the transistor characteristics. Next, the threshold τ_j is encoded as a resistor, by using the following mapping function:

$R_j = \frac{\tau_j - \tau_j^{min}}{\tau_j^{max} - \tau_j^{min}} \cdot (R_{max} - R_{min}) + R_{min}$, where R_{min}, R_{max} are

the technology-dependent and feasible (printable) resistor values, and $\tau_j^{min}, \tau_j^{max}$ are determined by the trained decision tree model.

Based on the difference of resistances of the transistor and resistor, one output node (S_1 or S_2) is pulled up stronger to VDD than the other, and the bi-stable back-to-back inverter converges to a state, where the output nodes are complementary ('1'/'0' or '0'/'1'). These output signals are then passed to the child nodes, where only one is enabled at a time, based on the analog binary comparison. This structure guarantees that at any level of the tree, only one child is selected, and hence, at the leaf level, one and only one leaf is selected. An interesting side-effect of this is that switching activity is limited to the depth of the tree. In effect, there is implicit logic which gates off unused portions of the circuit from consuming dynamic power. For any Decision Tree design, the process of adding split nodes to the last layers is repeated until the desired Decision Tree architecture is reached, and the class labels are read out from the leaves of the last split nodes in each branch. Due to the insertion of selector transistors in the split nodes, the resulting voltage levels deteriorate from the root node down to the split nodes in the last layer. This signal attenuation across a cascade of split nodes can be compensated for by using additional inverters (buffers) before the input of the selector transistors, to improve the signal levels.

Fig. 16 shows the latency, area and power of analog implementation of bespoke maximally parallel trees in EGFET. Our calculations show that analog trees have $437\times$ less area and $27\times$ less power (on average) with slight increase in latency ($1.6\times$) compared to digital bespoke maximally parallel trees.

Authorized licensed use limited to: University of Illinois. Downloaded on August 02,2021 at 15:32:35 UTC from IEEE Xplore. Restrictions apply.

We similarly developed an analog SVM implementation where we replaced the MAC operation by a one-time programmed resistive crossbar architecture, depicted in Fig. 15. The crossbar architecture is programmed by printing resistors with different geometries at the crossbar interconnects. The inputs to the crossbar are voltage signals, and the output current is sensed per each column separately.

The MAC operation is performed by applying Kirchoff's rule to the resistor network. The output voltage $V_{out}^{(c)}$ of a virtually grounded column c is computed by:

$$V_{out}^{(c)} = \sum_{i=1}^{P} \frac{V_i}{R_i^{(c)}} \left(\sum_{i=1}^{P} \frac{1}{R_i^{(c)}} \right)^{-1} = \sum_{i=1}^{P} V_i \, w_i^{(c)} \tag{1}$$

with

$$w_i^{(c)} = \frac{1}{R_i^{(c)}} \left(\sum_{i=1}^{P} \frac{1}{R_i^{(c)}} \right)^{-1} \tag{2}$$

where P is the number of rows per column.

Thus, the MAC operation can be directly derived from (1), where the values w_i can be determined by printing appropriate resistance values R_i, which solve (2). As the voltages are set as the MAC inputs $V_i = x_i$, the mathematical form of a multi-input MAC operation is obtained: $y = \sum_{i=1}^{P} w_i x_i$.

Fig. 17 shows the latency, area and power of analog implementation of bespoke SVMs. Our calculations show that, in analog implementations, on average, area and power improve by 490× and 12× with slight increase in latency 1.3× respectively. To the best of our knowledge, these are the first printed implementations of analog Decision Trees and SVMs and quantification of their benefits.

B. Analog Decision Tree Prototype

Finally, to demonstrate the feasibility of analog printed classifiers, we fabricated an analog 2-level Decision Tree based on EGT-technology. The fabricated 2-level Decision Tree consists of one root node and two split nodes (Fig. 15) resulting in 11 EGTs and 3 printed resistors. In this layout, PEDOT resistor is printed into the gap between the source of T_S and c_4. The overall fabrication process was similar to one described in Section IV-C. In addition to EGTs and resistors, crossovers were also inkjet printed to make the connections, by using an isolation layer (Dimethylsulfoxide (DMSO) and Polycarbonate (PC)) in combination with a conductive layer (PEDOT:PSS). A microscope photo of the fabricated 2-level Decision Tree is provided in Fig. 15.

The transient measurements for the root node are depicted in Fig. 15. As expected, when the input x_1 is at logical '1', S_1/S_2 are in state '1'/'0'. When x_1 is '0', the state changes to '0'/'1'. The transient response of the right split node to all 4 input combinations is also shown in Fig. 15. In the case the split node is unselected (x_1 is high), the output voltages C_3/C_4 are pulled down to 0V. On the other hand, if the split node is selected (x_1 is low), C_3/C_4 are pulled up or down, according to the input signal x_2. The worst case output signals of the split node are clearly distinguishable (405mV), and hence the printed analog Decision Tree was functioning correctly. This is the first prototype of an analog Decision Tree in a printed technology.
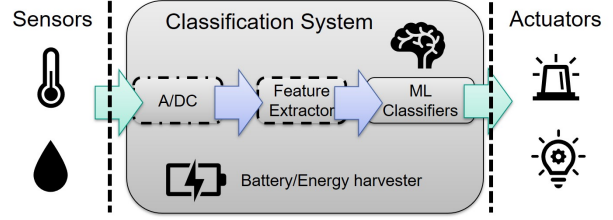

Fig. 18: Block Diagram of a canonical printed classifier system. The power source may be a printed battery or a harvester. Components outlined using dashes may be omitted in certain contexts.
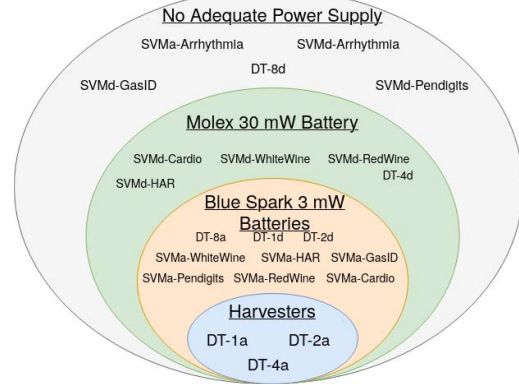

Fig. 19: Optimized analog and digital classifiers placed into sets based on which sources can power them (similar to Fig. 3). For each decision tree depth, we pick the set for the power source that can power all bespoke benchmark variants. E.g., Molex printed batteries can power depth-4 digital decision trees for each benchmark. Harvesters [42] are now capable of powering several decision trees. The 'a' and 'd' suffixes on 'SVM' and 'DT' are for 'analog' and 'digital' respectively.

VII. DISCUSSION

A printed ML classifier is only a component of a complete classification system (Fig. 18). Some classification applications (e.g. HAR, Pendigits, Red-Wine and White-Wine) require no feature extraction, as the accelerator performs inference directly on the sensed signals. For other applications, feature extraction can be performed either with a custom designed fixed-function unit, or with software running on a printed microprocessor [10]. Sensors are either integrated directly into computational units, bypassing ADCs [60], [61] or use printed ADCs [1]. Since the role of any printed system is known at print-time, interfaces can be custom. Additionally, due to the low cost of wires and high cost of flip-flops [10], parallel ports may be preferred to serial ports. Since these components are integrated to the same substrate at print time and are unpackaged, there are no concerns over pin counts. In general, the ideal applications for printed classifiers are those with non-categorical data (signals have to be measured in the field) with minimal feature extraction requirements, and relaxed latency requirements.

Fig. 19 shows that, unlike conventional classifier architectures (Fig. 3), most printing-specific classifier architectures can be powered by a printed battery or an energy harvester. However, we still see that several classifiers are currently infeasible due to power limitations. We also see that the power supply requirements of bespoke classifiers is dataset dependent (all bespoke analog SVMs can be powered by Blue Spark

batteries, except for Arrythmia). These results motivate the need for further research in printed batteries and harvesters. Additionally, it suggests that certain applications will need to sacrifice classification accuracy in order to meet power budget requirements.

Conventional printed classifiers are expensive even when a full system is considered. For example, we estimate EGT-printed 2-bit / 4-bit ADCs to cost $3.76\,mm^2$ / $25.4\,mm^2$ in area and $60\,\mu W$ / $360\,\mu W$ power [10]. Conventional EGT-printed classifiers (Tables III, IV, V) are often much bigger ($\sim$20 to $1445\,cm^2$) and consume orders of magnitude more power (1.6 to $4200\,mW$). Similarly, a printed microprocessor based feature extraction (FE) may cost $\sim$2 to $3\,cm^2$ [10]; printed sensors can be as small as $\sim$0.5 mm^2 and consume $<$2 mW [38]. Again, conventional printed classifiers are much more expensive (Tables III, IV, V). A classifier's system-level overhead will be even higher when ADCs and FE engines are optimized for area or power, or not used entirely (e.g, in case of direct interfacing [60]. The techniques proposed in this paper, therefore, would provide significant system-level benefits. This would be true even when power supply is considered; commercial printed batteries occupy 20 to $50\,cm^2$, while printed backscatter mechanisms (e.g., for RFIDs) are at least an order of magnitude smaller [75].

Design of printed classifiers is highly automatable. After training a decision tree on a dataset, scikit-learn's Decision-TreeClassifier class exposes the internal structure of the trained tree, allowing us to traverse the tree and recursively transform it into RTL for a bespoke decision tree classifier (we perform a similar transformation for SVM). Ozer *et al* propose a similar design flow. For our bespoke parallel trees, the generated RTL module body consists of a single assignment statement for each node in the tree, as well as a 'casex' statement to choose the correct output class, a relatively straightforward transformation. For the serial trees, the tree is transformed into hexfiles representing the class and threshold values which are stored in ROM. For lookup-based classifiers, we replace comparisons in decision trees and MACs in SVMs by generating in RTL lookup tables in place of those operations. An automated process for finding what to replace with lookup tables and how to replace it has the potential to lead to better results. We know that one indicator of a beneficial replacement is MISD parallelism (since it allows decoder reuse). Finding MISD parallelism in a design tool can be done by finding high fanout nets. Then, for each output connected to a given net, the tool finds maximum amount of logic that only depends on that net and replaces that logic with a ROM. The correct methodology for automatically determining how to break up a given computation into multiple lookup tables and logic that combines their results is not obvious and requires further study. Generation of analog classifiers is understandably more difficult to automate (e.g., we do not know of any printed analog libraries) - we consider it a subject of future work. Our digital SVM and DT generator is published under an open source license at https://github.com/PrintedComputing.

Compared to CMOS baselines (Table IV), EGT-based analog / digital trees are 10$\times$ / 1000$\times$ larger, and consume $10^5\times$ / $10^6\times$ more energy per inference. Similarly large area and energy differences exist for SVMs. Thus it is unlikely that there exist system design points such that an EGT-based system outperforms a silicon CMOS system in terms of power, performance, or area (PPA) – argument for printing must be in non-PPA terms (cost, conformality, time-to-market, non-toxicity, etc.)

Finally, envisioned applications for printed classifiers have resilience requirements (against bending, dirt, humidity, wear and tear, etc.). EGTs can be bent reliably to a radius of 10mm over 100 times with ¡10% change in electrical characteristics [43]; resilience against dirt, humidity can be provided easily by a printed passivation layer [50]. This level of resilience is adequate for the short-shelf-life applications we target. For higher mechanical and temperature resilience, Kapton films [51] can be used.

VIII. Conclusion

A large number applications domains have not seen much penetration of computing due to the cost, conformity, and toxicity limitations of silicon-based computer systems. Recent low-voltage printed computer systems have the potential to address these limitations. A common computational task in these application domains is anticipated to be machine learning classification. In this work, we explored the hardware cost of inference engines for popular classification algorithms in EGT and CNT-TFT printed technologies. We found that Decision Trees and SVMs provide a good balance between accuracy and cost. Subsequently, we evaluated conventional Decision Tree and SVM architectures in these technologies and concluded that their area and power overhead must be reduced for them to be feasible. Then, we explored, through SPICE and gate-level simulations and multiple working prototypes, several printing-specific classifier architectures that exploit the unique cost and implementation tradeoffs in printed technologies. Our evaluations showed that bespoke EGT printed Decision Trees have 48.9$\times$ lower area (average) and 75.6$\times$ lower power (average) than their conventional equivalents; corresponding benefits for bespoke SVMs are 12.8$\times$ and 12.7$\times$ respectively. Lookup-based Decision Trees outperformed their non-lookup bespoke equivalents by 38% and 70%; lookup-based SVMs were better by 8% and 0.6%. Analog printed Decision Trees provided 437$\times$ and 27$\times$ benefits over digital bespoke counterparts; analog SVMs yielded 490$\times$ and 12$\times$ improvements. Our results and prototypes demonstrate feasibility of fabricating and deploying battery and self-powered printed classifiers in the application domains of interest.

IX. Acknowledgments

REFERENCES

[1] "Flex-adc, american semiconductor." [Online]. Available: https://www.americansemi.com/flex-adc-development-kit.html

[2] "Thin-film battery, molex." [Online]. Available: https://www.molex.com/molex/products/family/thinfilm_battery

[3] K. Ahmed and M. Gregory, "Integrating wireless sensor networks with cloud computing," in *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks*, 2011, pp. 364–366.

[4] F. Alimoglu and E. Alpaydin, "Methods of combining multiple classifiers based on different representations for pen-based handwritten digit recognition," in *Proceedings of the Fifth Turkish Artificial Intelligence and Artificial Neural Networks Symposium (TAINN 96.* Citeseer, 1996.

[5] A. Almusallam, R. Torah, D. Zhu, M. Tudor, and S. Beeby, "Screen-printed piezoelectric shoe-insole energy harvester using an improved flexible pzt-polymer composites," in *Journal of Physics: Conference Series*, vol. 476, no. 1. IOP Publishing, 2013, p. 012108.

[6] K. Andrew and P. Gudimetla, "Internal positioning system for cardiotocograph (ctg) transducers at mackay base hospital," *Procedia Manufacturing*, vol. 30, pp. 365–372, 2019.

[7] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones." in *Esann*, 2013.

[8] A. Avramescu, T. Noguer, M. Avramescu, and J.-L. Marty, "Screen-printed biosensors for the control of wine quality based on lactate and acetaldehyde determination," *Analytica Chimica Acta*, vol. 458, no. 1, pp. 203–213, 2002.

[9] D. Ayres-de Campos, J. Bernardes, A. Garrido, J. Marques-de Sa, and L. Pereira-Leite, "Sisporto 2.0: a program for automated analysis of cardiotocograms," *Journal of Maternal-Fetal Medicine*, vol. 9, no. 5, pp. 311–318, 2000.

[10] N. Bleier, M. Mubarik, F. Rasheed, J. Aghassi-Hagmann, M. Tahoori, and R. Kumar, "Printed microprocessors," in *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, (accepted) 2020.

[11] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152.

[12] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[13] J. Burrell, T. Brooke, and R. Beckwith, "Vineyard computing: sensor networks in agricultural production," *IEEE Pervasive Computing*, vol. 3, no. 1, pp. 38–45, 2004.

[14] J. A. Cavallo, M. C. Strumia, and C. G. Gomez, "Preparation of a milk spoilage indicator adsorbed to a modified polypropylene film as an attempt to build a smart packaging," *Journal of Food Engineering*, vol. 136, pp. 48 – 55, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0260877414001344

[15] Y. Cha, K. Nam, and D. Kim, "Patient posture monitoring system based on flexible sensors," *Sensors*, vol. 17, no. 3, p. 584, 2017.

[16] J. Chang, X. Zhang, T. Ge, and J. Zhou, "Fully printed electronics on flexible substrates: High gain amplifiers and dac," *Organic Electronics*, vol. 15, no. 3, pp. 701–710, 2014.

[17] B. S. Cook, R. Vyas, S. Kim, T. Thai, T. Le, A. Traille, H. Aubert, and M. M. Tentzeris, "Rfid-based sensors for zero-power autonomous wireless sensor networks," *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2419–2431, 2014.

[18] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Modeling wine preferences by data mining from physicochemical properties," *Decision Support Systems*, vol. 47, no. 4, pp. 547–553, 2009.

[19] J. Courbat, D. Briand, L. Yue, S. Raible, and N. De Rooij, "Drop-coated metal-oxide gas sensor on polyimide foil with reduced power consumption for wireless applications," *Sensors and Actuators B: Chemical*, vol. 161, no. 1, pp. 862–868, 2012.

[20] Z. Cui, *Printed electronics: materials, technologies and applications*. John Wiley & Sons, 2016.

[21] A. de la Fuente Vornbrock, D. Sung, H. Kang, R. Kitsomboonloha, and V. Subramanian, "Fully gravure and ink-jet printed high speed pbttt organic thin film transistors," *Organic Electronics*, vol. 11, no. 12, pp. 2037–2044, 2010.

[22] H. de Lassus Saint-Geniès, D. Defour, and G. Revy, "Exact lookup tables for the evaluation of trigonometric and hyperbolic functions," *IEEE Transactions on Computers*, vol. 66, no. 12, pp. 2058–2071, 2017.

[23] M. Douthwaite, F. García-Redondo, P. Georgiou, and S. Das, "A time-domain current-mode mac engine for analogue neural networks in flexible electronics," in *2019 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2019, pp. 1–4.

[24] D. Dua and C. Graff, "UCI machine learning repository," 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[25] A. T. E. et al., "Inkjet-printed egfet-based physical unclonable function—design, evaluation, and fabrication," *IEEE TVLSI*, vol. 26, no. 12, pp. 2935–2946, 2018.

[26] C. M. et al., "Progress report on "from printed electrolyte-gated metal-oxide devices to circuits"," *Advanced Materials*, p. 1806483, 2019.

[27] D. W. et al., "An inkjet-printed low-voltage latch based on inorganic electrolyte-gated transistors," *IEEE EDL*, vol. 39, no. 6, pp. 831–834, 2018.

[28] J. S. C. et al., "A circuits and systems perspective of organic/printed electronics: review, challenges, and contemporary and emerging design approaches," *IEEE JESTCS*, vol. 7, no. 1, pp. 7–26, 2017.

[29] M. G. M. et al., "All-printed flexible and stretchable electronics," *Advanced Materials*, vol. 29, no. 19, p. 1604965, 2017.

[30] P.-Y. C. et al., "30.1: Invited paper: 65-inch inkjet printed organic light-emitting display panel with high degree of pixel uniformity," in *SID Symposium Digest of Technical Papers*, vol. 45, no. 1. Wiley Online Library, 2014, pp. 396–398.

[31] V. S. et al., "Printed electronics for low-cost electronic systems: Technology status and application development," in *2008-34th ESSCIRC*. IEEE, 2008, pp. 17–24.

[32] S. Feng, F. Farha, Q. Li, Y. Wan, Y. Xu, T. Zhang, and H. Ning, "Review on smart gas sensing technology," *Sensors*, vol. 19, no. 17, p. 3760, 2019.

[33] G. Gielen and W. Dehaene, "Analog and digital circuit design in 65 nm cmos: End of the road?" in *Design, Automation and Test in Europe*. IEEE, 2005, pp. 37–42.

[34] H. A. Guvenir, B. Acar, G. Demiroz, and A. Cekin, "A supervised machine learning algorithm for arrhythmia analysis," in *Computers in Cardiology 1997*. IEEE, 1997, pp. 433–436.

[35] J. Handy, "Why are computer chips so expensive?" 2014.

[36] Y. Huang, H. Wu, L. Xiao, Y. Duan, H. Zhu, J. Bian, D. Ye, and Z. Yin, "Assembly and applications of 3d conformal electronics on curvilinear surfaces," *Materials Horizons*, vol. 6, no. 4, pp. 642–683, 2019.

[37] W. J. Hyun, E. B. Secor, M. C. Hersam, C. D. Frisbie, and L. F. Francis, "High-resolution patterning of graphene by screen printing with a silicon stencil for highly flexible printed electronics," *Advanced Materials*, vol. 27, no. 1, pp. 109–115, 2015.

[38] W. Jia, A. J. Bandodkar, G. Valdes-Ramirez, J. R. Windmiller, Z. Yang, J. Ramírez, G. Chan, and J. Wang, "Electrochemical tattoo biosensors for real-time noninvasive lactate monitoring in human perspiration," *Analytical chemistry*, vol. 85, no. 14, pp. 6553–6560, 2013.

[39] M. Jung, J. Kim, J. Noh, N. Lim, C. Lim, G. Lee, J. Kim, H. Kang, K. Jung, A. D. Leonard *et al.*, "All-printed and roll-to-roll-printable 13.56-mhz-operated 1-bit rf tag on plastic foils," *IEEE Transactions on Electron Devices*, vol. 57, no. 3, pp. 571–580, 2010.

[40] S. Kim, M. M. Tentzeris, and A. Georgiadis, "Hybrid printed energy harvesting technology for self-sustainable autonomous sensor application," *Sensors*, vol. 19, no. 3, p. 728, 2019.

[41] Y. J. Kim, S.-E. Chun, J. Whitacre, and C. J. Bettinger, "Self-deployable current sources fabricated from edible materials," *Journal of Materials Chemistry B*, vol. 1, no. 31, pp. 3781–3788, 2013.

[42] J. Lim, H. Jung, C. Baek, G.-T. Hwang, J. Ryu, D. Yoon, J. Yoo, K.-I. Park, and J. H. Kim, "All-inkjet-printed flexible piezoelectric generator made of solvent evaporation assisted batio3 hybrid material," *Nano Energy*, vol. 41, pp. 337–343, 2017.

[43] H. Ling, D. A. Koutsouras, S. Kazemzadeh, Y. van de Burgt, F. Yan, and P. Gkoupidenis, "Electrolyte-gated transistors for synaptic electronics, neuromorphic computing, and adaptable biointerfacing," *Applied Physics Reviews*, vol. 7, no. 1, p. 011307, 2020.

[44] L. Liu, Y. Feng, and W. Wu, "Recent progress in printed flexible solid-state supercapacitors for portable and wearable energy storage," *Journal of Power Sources*, vol. 410, pp. 69–77, 2019.

[45] C. M. Lochner, Y. Khan, A. Pierre, and A. C. Arias, "All-organic optoelectronic sensor for pulse oximetry," *Nature Communications*, vol. 5, no. 1, 2014.

[46] W. Maly, "Cost of silicon viewed from vlsi design perspective," in *Proceedings of the 31st annual Design Automation Conference*, 1994, pp. 135–142.

[47] R. a. Markets, "Global fmcg market opportunity analysis, 2018-2019 2025 - a billion new consumers in emerging markets," Sep 2019. [Online]. Available: https://www.prnewswire.com/news-releases/global-fmcg-market-opportunity-analysis-2018-2019--2025---a-billion-new-consumers-in-emerging-markets-300913934.html

[48] P. Mostafalu, W. Lenk, M. R. Dokmeci, B. Ziaie, A. Khademhosseini, and S. R. Sonkusale, "Wireless flexible smart bandage for continuous monitoring of wound oxygenation," *IEEE Transactions on biomedical circuits and systems*, vol. 9, no. 5, pp. 670–677, 2015.

[49] S. Mühl and B. Beyer, "Bio-organic electronics—overview and prospects for the future," *Electronics*, vol. 3, no. 3, pp. 444–461, 2014.

[50] S. Nam, H. Jeon, S. H. Kim, J. Jang, C. Yang, and C. E. Park, "An inkjet-printed passivation layer based on a photocrosslinkable polymer for long-term stable pentacene field-effect transistors," *Organic Electronics*, vol. 10, no. 1, pp. 67–72, 2009.

[51] C. Neff, "Analysis of printed electronic adhesion, electrical, mechanical, and thermal performance for resilient hybrid electronics," 2018.

[52] K. Nguyen, E. Bamgbose, B. P. Cox, S. P. Huang, A. Mierzwa, S. Hutchins, B. Caso, M. Culjat, C. Connelly, D. Y. Lacoursiere *et al.*, "Wearable fetal monitoring solution for improved mobility during labor & delivery," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 4397–4400.

[53] G. A. Oguntala, R. A. Abd-Alhameed, N. T. Ali, Y.-F. Hu, J. M. Noras, N. N. Eya, I. Elfergani, and J. Rodriguez, "Smartwall: Novel rfid-enabled ambient human activity recognition using machine learning for unobtrusive health monitoring," *IEEE Access*, vol. 7, pp. 68 022–68 033, 2019.

[54] E. Ozer, "Towards low-cost bespoke computing with flexible electronics," Mar 2019.

[55] E. Ozer, J. Kufel, J. Biggs, G. Brown, J. Myers, A. Rana, A. Sou, and C. Ramsdale, "Bespoke machine learning processor development framework on flexible substrates," in *2019 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)*. IEEE, 2019, pp. 1–3.

[56] M. Peris and L. Escuder-Gilabert, "Electronic noses and tongues to assess food authenticity and adulteration," *Trends in Food Science & Technology*, vol. 58, pp. 40–54, 2016.

[57] M. Pietsch, T. Rödlmeier, S. Schlisske, J. Zimmermann, C. Romero-Nieto, and G. Hernandez-Sosa, "Inkjet-printed polymer-based electrochromic and electrofluorochromic dual-mode displays," *Journal of Materials Chemistry C*, vol. 7, no. 23, pp. 7121–7127, 2019.

[58] F. Poehl, F. Demmerle, J. Alt, and H. Obermeir, "Production test challenges for highly integrated mobile phone socs—a case study," in *2010 15th IEEE European Test Symposium*. IEEE, 2010, pp. 17–22.

[59] J. R. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1, no. 1, pp. 81–106, 1986.

[60] F. Reverter, "The art of directly interfacing sensors to microcontrollers," *Journal of Low Power Electronics and Applications*, vol. 2, no. 4, p. 265–281, 2012.

[61] F. Reverter, "Interfacing sensors to microcontrollers," *Smart Sensors and MEMs*, p. 23–55, 2018.

[62] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," California Univ San Diego La Jolla Inst for Cognitive Science, Tech. Rep., 1985.

[63] M. Sadrawi, C.-H. Lin, Y.-T. Lin, Y. Hsieh, C.-C. Kuo, J. C. Chien, K. Haraikawa, M. F. Abbod, and J.-S. Shieh, "Arrhythmia evaluation in wearable ecg devices," *Sensors*, vol. 17, no. 11, p. 2445, 2017.

[68] H. Song, J. Choi, C.-W. Park, D.-B. Shin, S.-S. Kang, S. H. Oh, and K. Hwang, "Study of quality control of traditional wine using it sensing

[64] B. Shao, "Fully printed chipless rfid tags towards item-level tracking applications," Ph.D. dissertation, KTH Royal Institute of Technology, 2014.

[65] L. Shao, T.-C. Huang, T. Lei, Z. Bao, R. Beausoleil, and K.-T. Cheng, "Compact modeling of carbon nanotube thin film transistors for flexible circuit design," in *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2018, pp. 491–496.

[66] L. Shao, T.-C. Huang, T. Lei, Z. Bao, R. Beausoleil, and K.-T. Cheng, "Process design kit for flexible hybrid electronics," in *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*. IEEE, 2018, pp. 651–657.

[67] M. G. Signorini, G. Lanzola, E. Torti, A. Fanelli, and G. Magenes, "Antepartum fetal monitoring through a wearable system and a mobile application," *Technologies*, vol. 6, no. 2, p. 44, 2018.
technology," *Journal of the Korean Society of Food Science and Nutrition*, vol. 44, no. 6, pp. 904–911, 2015.

[69] J. A. Spaeder, "Radio frequency identification pharmaceutical tracking system and method," 2005.

[70] B. Spark. Blue spark 10mah battery. [Online]. Available: https://www.bluesparktechnologies.com/index.php/products-and-services/battery-products/ultra-thin-series

[71] B. Spark. Blue spark 30mah battery. [Online]. Available: https://www.bluesparktechnologies.com/index.php/products-and-services/battery-products/101-blue-spark-st-standard-series2

[72] M. Vanderroost, P. Ragaert, F. Devlieghere, and B. De Meulenaer, "Intelligent food packaging: The next generation," *Trends in food science & technology*, vol. 39, no. 1, pp. 47–62, 2014.

[73] W. Voit, W. Zapka, P. Dyreklev, O.-J. Hagel, A. Hägerström, and P. Sandström, "Inkjet printing of non-volatile rewritable memory arrays," in *NIP & Digital Fabrication Conference*, vol. 2006, no. 3. Society for Imaging Science and Technology, 2006, pp. 34–37.

[74] S. Vora, K. Dandekar, and T. Kurzweg, "Passive rfid tag based heart rate monitoring from an ecg signal," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 4403–4406.

[75] A. Walla, B. Hassan, J. Yong, Y. Liang, Y. Yu, B. Nasr, A. Nirmalathas, and E. Skafidas, "A fully printed backscatter radio transceiver," in *2019 IEEE MTT-S International Microwave Symposium (IMS)*. IEEE, 2019, pp. 1442–1445.

[76] Y. Wang, J. Bokor, and A. Lee, "Maskless lithography using drop-on-demand inkjet printing method," in *Emerging Lithographic Technologies VIII*, vol. 5374. International Society for Optics and Photonics, 2004, pp. 628–636.

[77] D. D. Weller, M. Hefenbrock, M. B. Tahoori, J. Aghassi-Hagmann, and M. Beigl, "Programmable neuromorphic circuit based on printed electrolyte-gated transistors," in *Proceedings of the Asia South Pacific design automation conference (ASP-DAC)*, 2020.

[78] T. Yamada, Y. Hayamizu, Y. Yamamoto, Y. Yomogida, A. Izadi-Najafabadi, D. N. Futaba, and K. Hata, "A stretchable carbon nanotube strain sensor for human-motion detection," *Nature nanotechnology*, vol. 6, no. 5, p. 296, 2011.

[79] Y. Ye, Y. Kang, C. Zhang, Y. Chan, H. Wu, S. Lee, Z. Song, and B. Chen, "A 40-nm 16-mb contact-programming mask rom using dual trench isolation diode bitcell," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 4, pp. 1333–1341, 2015.

[80] J. Zhou, T. Ge, and J. S. Chang, "Fully-additive printed electronics: Process development kit," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2016, pp. 862–865.