On the approximability of the Largest Sphere Rule Ensemble Classification problem[☆]

Vahan Mkrtchyan^a, K. Subramani^{a,1,*}

^aLane Department of Computer Science and Electrical Engineering, West Virginia University, United States of America

Abstract

In this paper, we discuss the approximability of the Largest Sphere Rule Ensemble Classification problem (**LSRES**). This problem is closely related to the Weighted Rule Ensemble Classification Reversal problem (**WRECR**) [2].

Key Words: Weighted Rule Ensemble Classification Reversal problem; feasibility; **NP-completeness**; approximation; inapproximability.

2010 MSC: 68Q17; 03B70

1. Introduction

In this paper, we discuss the approximability of the Largest Sphere Rule Ensemble Classification problem (**LSRES**). This problem is closely related to the Weighted Rule Ensemble Classification Reversal problem (**WRECR**) [2].

WRECR is a well-known problem, widely referenced in the literature. This problem has applications in domains such as biology, medicine and computer security [2, 3, 5, 6]. In this paper, we introduce a related problem called **LSRES**, and motivate it through an application in computer security (see Section 4).

The main contributions of this paper are as follows:

- (i) The design of a $\frac{C}{n}$ -approximation algorithm for **LSRES**, with running time $O(n^{C+1})$ for each fixed $C > 0$.
- (ii) The proof of absence of a polynomial-time $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm for **LSRES**, unless **P=NP**. This result holds for each $\epsilon \in (0, 1)$.
- (iii) The proof of **NP-completeness** of the feasibility version of **WRECR**.

The rest of the paper is organized as follows: In Section 2, we discuss the preliminaries and notations used throughout the paper. Section 3 presents the formulations of the main problems discussed in the paper. Section 4 discusses a problem that motivates the study of **LSRES** and presents related work in the literature. Section 5 discusses the main results obtained in this paper. In Section 6, we summarize our contributions and outline directions for future research.

2. Preliminaries

In this paper, we use the standard convention that 0 corresponds to **false**, and 1 corresponds to **true**. Let $B = \{0, 1\}$. The operations defined on B are algebraic or Boolean depending on the context.

A Boolean function f is a mapping $f : B^n \rightarrow B$, where $B^n = B \times B \times \dots \times B$.

Definition 1. Let $f(x_1, \dots, x_n)$ be a Boolean function. Its dual Boolean function [7] is defined as:

$$f^*(x_1, \dots, x_n) = \bar{f}(\bar{x}_1, \dots, \bar{x}_n).$$

From the definition, it follows that for any Boolean function f , we have,

$$(f^*)^* = f.$$

[☆]This research has been supported in part by the Air Force of Scientific Research through Award FA9550-12-1-0199.

*Corresponding author

Email addresses: vahan.mkrtchyan@mail.wvu.edu (Vahan Mkrtchyan), ksmmani@csee.wvu.edu (K. Subramani)

¹The author is supported by the National Science Foundation through Award CCF-1305054

For example, if

$$f_1(x) = x, f_2(x) = \bar{x}, \text{ and } f_3(x, y) = x \wedge y,$$

then

$$f_1^*(x) = x, f_2^*(x) = \bar{x}, \text{ and } f_3^*(x, y) = x \vee y.$$

The first two equalities show that the identity and complement functions are self-dual (i.e., the function coincides with its dual). The last equality can be generalized to show that the dual of a conjunction is a disjunction and vice versa.

Let $\mathbf{x} = (x_1, \dots, x_n)$ be an n -dimensional vector over $\{0, 1\}$. A decision rule r is a conjunction of literals of the variables $x_1, \dots, x_n$. Assume, we are given a set of decision rules $R = \{r_1, \dots, r_t\}$, and a weight-function $w : R \rightarrow \mathbb{Z}$. Consider the expression $\Delta(R, w, \mathbf{x})$ defined as follows:

$$\Delta(R, w, \mathbf{x}) = \sum_{i=1}^t w(r_i) \cdot r_i(\mathbf{x}).$$

In the expression above we interpret the values of rules algebraically.

Let $F(\mathbf{x})$ be a function defined as:

$$F(\mathbf{x}) = \text{sgn}(\Delta(R, w, \mathbf{x})), \quad (1)$$

where

$$\text{sgn}(z) = \begin{cases} 1, & \text{if } z > 0, \\ -1, & \text{if } z < 0, \\ 0, & \text{otherwise.} \end{cases}$$

For a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$ and a set $S \subseteq \{1, \dots, n\}$, let $\mathbf{x}^S$ be the vector obtained from $\mathbf{x}$ by complementing the elements with indices from S .

The Weighted Rule Ensemble Classification Reversal problem (**WRECR**) is defined as follows:

WRECR: Given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$, a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$ and a positive integer $k \leq n$, is there a set $S \subseteq \{1, \dots, n\}$ with $|S| \leq k$, such that $F(\mathbf{x}^S) \neq F(\mathbf{x})$?

If we partition the set of n -dimensional Boolean vectors into 3 groups based on the values of F (see Figure 1), then the **WRECR** problem can be interpreted as the problem of checking whether there is a set $S \subseteq \{1, \dots, n\}$ with $|S| \leq k$, such that the vector $\mathbf{x}^S$ belongs to a different group (see Figure 1).

Figure 1: Partitioning n -dimensional Boolean vectors according to the values of F .

In this paper, we derive some results that are concerned with the approximation properties of a problem that is related to **WRECR**.

3. Formulation of the Main Problems

We first consider the feasibility version of **WRECR**:

WRECRF: Given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$, and a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$, is there a set $S \subseteq \{1, \dots, n\}$, such that $F(\mathbf{x}^S) \neq F(\mathbf{x})$?

Observe that **WRECR** is the decision problem corresponding to the problem of constructing a smallest cardinality feasible set S if one exists.

Now, we consider the **LSRES** problem, which is defined as follows:

LSRES: Given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$ and a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$, the goal is to find the maximum $d \geq 0$, such that for every set $S \subseteq \{1, \dots, n\}$ with $|S| \leq d$, $F(\mathbf{x}^S) = F(\mathbf{x})$ (see (1)).

Observation 1. *LSRES is always feasible. In other words, in this problem the parameter d is well-defined.*

One can always take $d = 0$, in order to get a feasible solution.

Lemma 1. **LSRES** is as hard as **WRECR** (**WRECR** $\prec$ **LSRES**).

Proof. We Turing-reduce **WRECR** to **LSRES**. Let $P = \langle R, w, \mathbf{x}, k \rangle$ be an instance of **WRECR**. We construct the instance $Q = \langle R, w, \mathbf{x} \rangle$ of **LSRES**. Assume that there is an algorithm which determines the parameter d in Q . Then it suffices to check whether $(d + 1) \leq k$ in order to solve P . $\square$

Finally, consider the decision version of **LSRES**:

LSRES_D: Given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$, a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$ and a number l , the goal is to check whether for every set $S \subseteq \{1, \dots, n\}$ with $|S| \leq l$, $F(\mathbf{x}^S) = F(\mathbf{x})$.

Lemma 2. **LSRES_D** is in coNP.

Proof. Assume that we are given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$, a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$, a number l and a set $S \subseteq \{1, \dots, n\}$ with $|S| \leq l$ and $F(\mathbf{x}^S) \neq F(\mathbf{x})$. We can verify that S satisfies $|S| \leq l$ and $F(\mathbf{x}^S) \neq F(\mathbf{x})$ in polynomial time. Thus, negative instances of **LSRES_D** have a certificate of polynomial length, which can be checked in polynomial time. $\square$

4. Motivation and Related Work

Assume that we are trying to test some patients for cancer and we have a large data set of observations. Based on the data, we build decision trees that contain gene expression levels of the patients. Each such tree corresponds to a decision rule. Each test checks whether the expression level of a gene is below a threshold [6]. A gene that has a significant impact on the classification decision is considered as an indicator, i.e., a characteristic showing the severity of the condition of a patient. Coming up with sets of genes that act as indicators can lead to new methods for overcoming the disease [3].

In order to solve the problem of finding indicators, we solve the following problem: Assume that we have

an observation $\mathbf{x} = (x_1, \dots, x_n)$, which has features x_i , $1 \leq i \leq n$. For simplicity of the presentation, we assume that $x_i \in \{0, 1\}$. Furthermore, assume that we have some decision rules $\{r_1, \dots, r_t\}$ defined on the set of observations, which take values 0, 1 or -1 . These values represent the impact of a rule on $\mathbf{x}$. Let $F(\mathbf{x})$ be a function indicating the sign of the total impact (sum of all impacts) of rules on the observation $\mathbf{x}$. $F(\mathbf{x})$ is 1, if the impact is positive, -1 , if the impact is negative and 0 otherwise. **LSRES** is concerned with the determination of the largest value of k , for which the modification of any $l \leq k$ features results in an observation $\mathbf{x}'$, such that $F(\mathbf{x}') = F(\mathbf{x})$.

The **LSRES** problem also has an application in security. Suppose we have an object which can be in one of several states. We represent these states as Boolean vectors $\mathbf{x} = (x_1, \dots, x_n)$. Assume that the object can transition from one state to another. Moreover, we assume that there are certain rules $R = \{r_1, \dots, r_t\}$ having positive and negative weights that apply in each state (rules apply in a state, and not during the transition). These rules are represented as conjunctions $r_i(x_1, \dots, x_n) = C_1^{(i)} \wedge \dots \wedge C_{s_i}^{(i)}$, $1 \leq i \leq t$ of positive or negative weights. If a conjunction is 0 on a state, we say that the corresponding rule is passive, otherwise it is active. $\Delta(R, w, \mathbf{x})$ as defined in Section 2, represents the total impact of all the rules. Depending on the sign of $\Delta(R, w, \mathbf{x})$, we partition the set of states into three disjoint groups. Assume that there is a certain penalty for an object, if it moves from a state from one group to that of another group. Then, the **LSRES** problem can be interpreted as the problem of finding the largest radius, within which the transition for the object is secure. In other words, for such transitions, the object is not penalized.

WRECR is considered in [2], where it is shown to be **NP-complete**. A problem that is closely related to **WRECR** is considered in [1]. However, in contrast with [2] (and this paper), in [1], the vectors are considered as elements of R^n - n -dimensional vector space. Moreover, in [2] it is argued that the two problems are independent, in a sense that neither of them is a particular case of the other one.

5. Main Results

In [2], it is shown that **WRECR** is **NP-complete**.

In this paper, we establish that even the feasibility version of **WRECR**, viz., **WRECRF** is **NP-complete**. It is important to note that our **NP-completeness** result subsumes the result in [2]. Furthermore, our proof is significantly simpler.

Theorem 1. ***WRECRF** is **NP-complete**.*

Proof. We first show that **WRECRF** is in **NP**. Assume that we are given a set of rules $R = \{r_1, \dots, r_t\}$, a weight function $w : R \rightarrow \mathbb{Z}$, a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$, and a set $S \subseteq \{1, \dots, n\}$, such that $F(\mathbf{x}^S) \neq F(\mathbf{x})$. We can verify that $F(\mathbf{x}^S) \neq F(\mathbf{x})$ in polynomial time.

In order to show that **WRECRF** is complete for **NP**, we present a polynomial time reduction from **SAT** [4].

Recall that in an instance of **SAT**, we are given a set $X = \{x_1, \dots, x_n\}$ of Boolean variables, a CNF formula $K(x_1, \dots, x_n) = C_1 \wedge \dots \wedge C_s$, where C_j is a clause over the literals of $\{x_1, \dots, x_n\}$, $1 \leq j \leq s$. The goal is to check whether there exists an assignment $\alpha = (\alpha_1, \dots, \alpha_n)$ to the variables $\{x_1, \dots, x_n\}$, such that $K(\alpha) = 1$.

Now we describe the reduction. Consider an arbitrary instance I of **SAT**. Without loss of generality, we can assume that $K(1, 1, \dots, 1) = 0$. Add a new variable z to X to obtain $X' = X \cup \{z\}$. Define the set of rules as $R = \{C_1^*, \dots, C_s^*, r_{s+1}, r_{s+2}\}$. Here, for $j = 1, \dots, s$, C_j^* is the dual of C_j , $r_{s+1} = z$, $r_{s+2} = \bar{z}$. Observe that each C_j^* is a conjunction of literals. The weight function on the set of rules is defined as follows: $w(C_j^*) = 1$, and $w(r_{s+1}) = w(r_{s+2}) = -\epsilon$ for some $\epsilon \in (0, 1)$. The initial $(n+1)$ -dimensional Boolean vector is defined as $\mathbf{0} = (0, \dots, 0, 0)$. Let I' be the resulting instance of **WRECRF**.

Observe that I' can be constructed from I in polynomial time. We now show that for each n -dimensional Boolean vector $(\alpha_1, \dots, \alpha_n)$, we have,

- (a) $K(\alpha_1, \dots, \alpha_n) = 0$, if and only if $F(\bar{\alpha}_1, \dots, \bar{\alpha}_n, \beta) = 1$ for any $\beta \in \{0, 1\}$,
- (b) $F(\alpha_1, \dots, \alpha_n, \beta) \neq 0$ for any $\beta \in \{0, 1\}$.

For the proof of observation (a), note that if $K(\alpha_1, \dots, \alpha_n) = 0$, then there is j ($1 \leq j \leq s$), such that $C_j(\alpha_1, \dots, \alpha_n) = 0$. This implies that $C_j^*(\bar{\alpha}_1, \dots, \bar{\alpha}_n) = 1$. Since $w(r_{s+1}) = w(r_{s+2}) = -\epsilon$, we have, $F(\bar{\alpha}_1, \dots, \bar{\alpha}_n, \beta) = 1$ for any $\beta \in \{0, 1\}$.

For the proof of converse statement, assume that $F(\bar{\alpha}_1, \dots, \bar{\alpha}_n, \beta) = 1$ for any $\beta \in \{0, 1\}$. Since the weights of $C_1^*, \dots, C_s^*$ are 1, the weights of r_{s+1}, r_{s+2} are $-\epsilon$, and $\epsilon \in (0, 1)$, there is j ($1 \leq j \leq s$), such that $C_j^*(\bar{\alpha}_1, \dots, \bar{\alpha}_n) = 1$. This implies that $C_j(\alpha_1, \dots, \alpha_n) = 0$. Hence $K(\alpha_1, \dots, \alpha_n) = 0$.

The proof of observation (b) follows from the fact that the weights of $C_1^*, \dots, C_s^*$ are 1, the weights of r_{s+1}, r_{s+2} are $-\epsilon$, and $\epsilon \in (0, 1)$.

We now show that I is a “yes” instance if and only if I' is a “yes” instance.

Assume that I is a “yes” instance. Since $K(1, 1, \dots, 1) = 0$, we have, $F(0, \dots, 0, 0) = 1$ (see observation (a)). As I is satisfiable, there exists a set $S \subseteq \{1, \dots, n\}$, such that modifying $(1, \dots, 1)$ on S leads to a satisfying assignment. Using that assignment and observation (a), one can show that the instance of **WRECRF** is a “yes” instance.

Now, assume that I' is a “yes” instance. We claim that $F(0, \dots, 0, 0) = 1$. Suppose that $F(0, \dots, 0, 0) \neq 1$. It follows from observation (b) that $F(0, \dots, 0, 0) = -1$. Observation (a) implies that $K(1, 1, \dots, 1) = 1$ contradicting our initial assumption that $K(1, 1, \dots, 1) = 0$. Thus $F(0, \dots, 0, 0) = 1$. Since I' is a “yes” instance, there is a set S such that $F((0, \dots, 0, 0)^S) \neq F(0, \dots, 0, 0)$, hence $F((0, \dots, 0, 0)^S) = -1$ (see observation (b)). Observation (a) implies that the instance I of **SAT** is satisfiable.

The proof of the theorem is completed. $\square$

We now show that **LSRES** does not admit a $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm (for any $\epsilon \in (0, 1)$), unless **P=NP**.

Theorem 2. *There is no a polynomial time $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm for **LSRES** for any $\epsilon \in (0, 1)$, unless **P=NP**.*

Proof. Let $\epsilon \in (0, 1)$. We prove this theorem by the following strategy: We assume that there exists a $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm $\mathcal{A}$ for **LSRES**, and

using this algorithm, we design a polynomial time algorithm for a **coNP-complete** problem. This proves the theorem.

The **coNP-complete** problem that we consider is a restriction of **Tautology** [4]:

Restriction of Tautology. Given a set $X = \{x_1, \dots, x_n\}$ of Boolean variables, a DNF formula $f(x_1, \dots, x_n) = K_1 \vee \dots \vee K_t$ such that $f(1, 1, \dots, 1) = 1$, the goal is to check whether f is identically one. Here K_i is a conjunction of literals of $\{x_1, \dots, x_n\}$, for $i = 1, \dots, t$.

Clearly, this problem is **coNP-complete**. Now we describe a polynomial time algorithm for this problem using the $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm $\mathcal{A}$ for **LSRES**.

Let I be an instance of **Restriction of Tautology**. Consider an instance I' of **LSRES** obtained as follows: Add new variables $z_0, z_1, \dots, z_q$ to X to obtain $X' = (X \cup \{z_0, z_1, \dots, z_q\})$, where q is the smallest integer satisfying $q \geq (n^{\frac{1}{\epsilon}} - n - 1)$. The set of rules is defined as $R = \{K_1, \dots, K_t, r_0, r'_0, \dots, r_q, r'_q\}$, where $r_0 = z_0$, $r'_0 = \bar{z}_0, \dots, r_q = z_q$, $r'_q = \bar{z}_q$.

The weight function on the set of rules is defined as follows: $w(K_j) = 1$, for $j = 1, \dots, t$ and $w(r_i) = w(r'_i) = -\delta$ for $i = 0, 1, \dots, q$, where δ is chosen to satisfy the inequality $(q+1) \cdot \delta < 1$ ($\delta = \frac{1}{q+2}$ will be fine for our purpose). The initial $(n+q+1)$ -dimensional Boolean vector is defined as $\mathbf{1} = (1, \dots, 1, 1, \dots, 1)$. Note that I' can be constructed in polynomial time.

We claim that I is a “yes” instance, if and only if $\mathcal{A}(I') \geq n$.

First, suppose that I is a “yes” instance, that is, f is identically one. This means that for each Boolean vector $\gamma = (\gamma_1, \dots, \gamma_n)$ there is a j , $1 \leq j \leq t$, such that $K_j(\gamma_1, \dots, \gamma_n) = 1$. Hence, by the choice of the weights, we have,

$$\Delta(R, w, \gamma) \geq 1 - (q+1) \cdot \delta. \quad (2)$$

Since $(q+1) \cdot \delta < 1$, we have for any $S \subseteq \{1, \dots, (n+q)\}$, $F(\mathbf{x}^S) = F(\mathbf{x})$. Hence

$$OPT = n + q + 1,$$

where OPT is the largest d from **LSRES**. Now, observe that since $\mathcal{A}$ is a $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm, we have,

$$\begin{aligned} \mathcal{A}(I) &\geq \frac{OPT}{(n+q+1)^{1-\epsilon}} \\ &= \frac{n+q+1}{(n+q+1)^{1-\epsilon}} \\ &= (n+q+1)^\epsilon \\ &\geq n \end{aligned}$$

by the choice of q .

For the proof of the converse statement, assume that the output of $\mathcal{A}(I) \geq n$. Since $f(1, 1, \dots, 1) = 1$, we have $F(1, \dots, 1, 1, \dots, 1) = 1$ (see (2)), and the same equality is true if we change the values of any n elements of the Boolean vector $(1, \dots, 1, 1, \dots, 1)$. Since $(q+1) \cdot \delta < 1$, (2) implies that for each Boolean vector $\theta = (\theta_1, \dots, \theta_n)$ there is a j , $1 \leq j \leq t$, such that $K_j(\theta_1, \dots, \theta_n) = 1$. This means that f is one on all 2^n n -dimensional Boolean vectors. Hence f is identically one.

The proved claim implies that $\mathcal{A}$ can be used to decide whether f is identically one in polynomial time. This means that **P=coNP**. Since **P=coP**, we have **P=NP**.

The proof of the theorem is completed. $\square$

Despite the previous negative result, it turns out that the problem admits a $\frac{C}{n}$ -approximation algorithm for any fixed $C > 0$.

Theorem 3. *For each fixed $C > 0$, **LSRES** admits a $\frac{C}{n}$ -approximation algorithm, whose running time is $O(n^{C+1})$.*

Proof. We show that Algorithm 1 meets the conditions of the theorem. Observe that its running time is $O(n^{C+1})$. This follows from the fact that the number of all Boolean vectors whose Hamming distance from $\mathbf{x} = (x_1, \dots, x_n)$ is at most $O(n^{C+1})$. Let us show that it has performance ratio $\frac{C}{n}$. First observe that for OPT , which is the largest value of d , we have the following upper bound

$$OPT \leq n.$$

Algorithm 1 A $\frac{C}{n}$ -approximation algorithm for **LSRES**

- 1: Input: a set of rules $R = \{r_1, \dots, r_t\}$, weight function $w : R \rightarrow Z$ and a Boolean vector $\mathbf{x} = (x_1, \dots, x_n)$.
 - 2: Output: Maximum $d \geq 0$, such that for any $S \subseteq \{1, \dots, n\}$ with $|S| \leq d$, we have $F(\mathbf{x}^S) = F(\mathbf{x})$.
 - 3: Consider all Boolean vectors whose Hamming distance from $\mathbf{x} = (x_1, \dots, x_n)$ is at most C .
 - 4: **if** (There is an $S \subseteq \{1, \dots, n\}$ such that $F(\mathbf{x}^S) \neq F(\mathbf{x})$) **then**
 - 5: The maximum value for d can be found in polynomial time, which means the problem can be solved exactly.
 - 6: **else**
 - 7: return C ;
 - 8: **end if**
-

Now observe that ALG , which is the output of the algorithm, is C . Hence

$$ALG = C = \frac{C}{n} \cdot n \geq \frac{C}{n} \cdot OPT,$$

or

$$OPT \leq \frac{n}{C} \cdot ALG.$$

The proof of the theorem is completed. $\square$

6. Conclusion

In the paper we considered the **LSRES** and **WRECR** problems. We have shown that it is **NP-hard** to check the feasibility of **WRECR** (**WRECRF**) which strengthens a result presented in [2]. Moreover, we have shown that **LSRES** does not admit a $\frac{1}{n^{1-\epsilon}}$ -approximation algorithm ($\epsilon \in (0, 1)$), unless **P=NP**, and it admits a $\frac{C}{n}$ -approximation algorithm for each fixed $C > 0$.

It would be interesting to investigate the case when the performance ratio is between $O(\frac{1}{n})$ and $O(\frac{1}{n^{1-\epsilon}})$. This case is already interesting because the reduction presented in the inapproximability result becomes non-polynomial for example for the case $\frac{1}{\frac{n}{\log n}} = \frac{\log n}{n}$ -performance ratio.

Though **WRECR** and **LSRES** are closely related from the perspective of exact algorithms, the two problems seem to be different from the perspective of approximation algorithms.

Acknowledgment. We are indebted to two anonymous referees of this paper for their useful comments, which helped us to improve the presentation of the paper.

References

- [1] D. Axelrod, T. Bonates, P. Hammer, I. Lozina, From Diagnosis to Therapy via LAD. Invited Lecture at INFORMS Annual Meeting, Denver, CO. 2004.
- [2] T. O. Bonates, J. W. V. Morais Neto, Reversing the Classification of Weighted Rule Ensembles, Proc. ISAIM 2014.
- [3] L. Chin, W. C. Hahn, G. Getz, M. Meyerson, Making Sense of Cancer Genomic Data. Genes and Development 25(6): 534-555, 2011.
- [4] M. R. Garey, D. S. Johnson, Computers and Intractability: A Guide to the Theory of **NP-completeness**. W. H. Freeman, 1979.
- [5] T. K. Ho, The Random Subspace Method for Constructing Decision Forests. Pattern Analysis and Machine Intelligence, IEEE Transactions on 20(8): 832-844, 1998.
- [6] A. Statnikov, C. F. Aliferis, Are random forests better than support vector machines for microarray-based cancer classification? In AMIA Annual Symposium Proceedings, 686-690. American Medical Informatics Association, 2007.
- [7] I. Wegener, The Complexity of Boolean Functions. John Wiley and Sons Ltd, and B. G. Teubner, Stuttgart. 1987.