אוניברסיטת בן-גוריון בנגב

הצעת תוכנית מחקר ללימודי דוקטורט

טרם נקבע שם

TBD

מנחם סדיגורסקי

אפריל 2018

שם המנחה: ________________________________חתימת המנחה: _________________________

חתימת יו״ר ועדת מוסמכים מחלקתי: _________________________

BEN-GURION UNIVERSITY OF THE NEGEV
THE FACULTY OF NATURAL SCIENCES
DEPARTMENT OF COMPUTER SCIENCE

TBD

This dissertation proposal is submitted in partial fulfillment of the requirements for Doctor in Philosophy degree

in

Computer Science

by

Menachem Sadigurschi

April 2018

תקציר

תרגום של התקציר לעברית יופיע כאן בסופו של דבר.

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Table of Contents

Chapter 1

Introduction

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.1 Definitions and notation

We will write $[k] := \{1, \ldots, k\}$. An *instance space* is an abstract set $\mathcal{X}$. For a concept class $\mathcal{C} \subset \{0,1\}^{\mathcal{X}}$, if say that $\mathcal{C}$ *shatters* a set $\{x_1, \ldots, x_k\} \subset \mathcal{X}$ if

$$\mathcal{C}(S) = \{(f(x_1), f(x_2), \ldots, f(x_k)) : f \in \mathcal{C}\} = \{0,1\}^k .$$

The VC-dimension $d = d_{\mathcal{C}}$ of $\mathcal{C}$ is the size of the largest shattered set (or ∞ if $\mathcal{C}$ shatters sets of arbitrary size) [Vapnik and Červonenkis, 1971]. When the roles of $\mathcal{X}$ and $\mathcal{C}$ are exchanged — that is, an $x \in \mathcal{X}$ acts on $f \in \mathcal{C}$ via $x(f) = f(x)$, — we refer to $\mathcal{X} = \mathcal{C}^*$ as the *dual* class of $\mathcal{C}$. Its VC-dimension is then $d^* = d^*_{\mathcal{C}} := d_{\mathcal{C}^*}$, and referred to as the *dual VC dimension.* Assouad [1983] showed that $d^* \leq 2^{d+1}$.

For $\mathcal{F} \subset \mathbb{R}^{\mathcal{X}}$ and $t > 0$, we say that $\mathcal{F}$ t-shatters a set $\{x_1, \ldots, x_k\} \subset \mathcal{X}$ if

$$\mathcal{F}(S) = \{(f(x_1), f(x_2), \ldots, f(x_k)) : f \in \mathcal{F}\} \subseteq \mathbb{R}^k$$

contains the translated cube $\{-t, t\}^k + r$ for some $r \in \mathbb{R}^k$. The t-fat-shattering dimension $d(t) = d_{\mathcal{F}}(t)$ is the size of the largest t-shattered set (possibly ∞) [Alon et al., 1997]. Again, the roles of $\mathcal{X}$ and $\mathcal{F}$ may be switched, in which case $\mathcal{X} = \mathcal{F}^*$ becomes the dual class of $\mathcal{F}$. Its t-fat-shattering dimension is then $d^*(t)$, and Assouad's argument shows that $d^*(t) \leq 2^{d(t)+1}$.

A *sample compression scheme* (κ, ρ) for a hypothesis class $\mathcal{F} \subset \mathcal{Y}^{\mathcal{X}}$ is defined as follows. A k-*compression* function κ maps sequences $((x_1, y_1), \ldots, (x_m, y_m)) \in \bigcup_{\ell \geq 1} (\mathcal{X} \times \mathcal{Y})^\ell$ to elements in $\mathcal{K} = \bigcup_{\ell \leq k'} (\mathcal{X} \times \mathcal{Y})^\ell \times \bigcup_{\ell \leq k''} \{0,1\}^\ell$, where $k' + k'' \leq k$. A *reconstruction* is a function $\rho : \mathcal{K} \to \mathcal{Y}^{\mathcal{X}}$. We say that (κ, ρ) is a k-size sample compression scheme for $\mathcal{F}$ if κ is a k-compression and for all $h^* \in \mathcal{F}$ and all $S = ((x_1, h^*(x_1)), \ldots, (x_m, h^*(y_m)))$, we have $\hat{h} := \rho(\kappa(S))$ satisfies $\hat{h}(x_i) = h^*(x_i)$ for all $i \in [m]$.

For real-valued functions, we say it is a *uniformly* ε-*approximate* compression scheme if

$$\max_{1 \leq i \leq m} |\hat{h}(x_i) - h^*(x_i)| \leq \varepsilon.$$

1.2 Main results

Throughout the paper, we implicitly assume that all hypothesis classes are *admissible* in the sense of satisfying mild measure-theoretic conditions, such as those specified in Dudley [1984, Section 10.3.1] or Pollard [1984, Appendix C]. We begin with an algorithmically efficient version of the learner-to-compression scheme conversion in Moran and Yehudayoff [2016]:

Theorem 1.2.1 (Efficient compression for classification). *Let $\mathcal{C}$ be a concept class over some instance space $\mathcal{X}$ with VC-dimension d, dual VC-dimension d^*, and suppose that $\mathcal{A}$ is a (proper, consistent) PAC-learner for $\mathcal{C}$: For all $0 < \varepsilon, \delta < 1/2$, all $f^* \in \mathcal{C}$, and all distributions D over $\mathcal{X}$, if $\mathcal{A}$ receives $m \geq m_{\mathcal{C}}(\varepsilon, \delta)$ points $S = \{x_i\}$ drawn iid from D and labeled with $y_i = f^*(x_i)$, then $\mathcal{A}$ outputs an $\hat{f} \in \mathcal{C}$ such that*

$$\mathbb{P}_{S \sim D^m}\left(\mathbb{P}_{X \sim D}\left(\hat{f}(X) \neq f^*(X) \,|\, S\right) > \varepsilon\right) < \delta.$$

For every such $\mathcal{A}$, there is a randomized sample compression scheme for $\mathcal{C}$ of *size $O(k \log k)$, where $k = O(dd^*)$. Furthermore, on a sample of any size m, the compression set may be computed in expected time*

$$O\left((m + T_{\mathcal{A}}(cd)) \log m + mT_{\mathcal{E}}(cd)(d^* + \log m)\right),$$

where $T_{\mathcal{A}}(\ell)$ is the runtime of $\mathcal{A}$ to compute $\hat{f}$ on a sample of size ℓ, $T_{\mathcal{E}}(\ell)$ is the runtime required to evaluate $\hat{f}$ on a single $x \in \mathcal{X}$, and c is a universal constant.

Although for our purposes the existence of a distribution-free sample complexity $m_{\mathcal{C}}$ is more important than its concrete form, we may take $m_{\mathcal{C}}(\varepsilon, \delta) = O(\frac{d}{\varepsilon} \log \frac{1}{\varepsilon} + \frac{1}{\varepsilon} \log \frac{1}{\delta})$ [Vapnik and Chervonenkis, 1974, Blumer et al., 1989], known to bound the sample complexity of empirical risk minimization; indeed, this loses no generality, as there is a well-known efficient reduction from empirical risk minimization to any proper learner having a polynomial sample complexity [Pitt and Valiant, 1988, Haussler et al., 1991]. We allow the evaluation time of $\hat{f}$ to depend on the size of the training sample in order to account for non-parametric learners, such as nearest-neighbor classifiers. A naive implementation of the Moran and Yehudayoff [2016] existence proof yields a runtime of order $m^{cd} T_{\mathcal{A}}(c'd) + m^{cd^*}$ (for some universal constants c, c'), which can be doubly exponential when $d^* = 2^d$; this is without taking into account the cost of computing the minimax distribution on the $m^{cd} \times m$ game matrix.

Next, we extend the result in Theorem 1.2.1 from classification to regression:

Theorem 1.2.2 (Efficient compression for regression)**.** *Let $\mathcal{F} \subset [0,1]^{\mathcal{X}}$ be a function class with t-fat-shattering dimension $d(t)$, dual t-fat-shattering dimension $d^*(t)$, and suppose that $\mathcal{A}$ is an ERM (i.e., proper, consistent) learner for $\mathcal{F}$: For all $f^* \in \mathcal{C}$, and all distributions D over $\mathcal{X}$, if $\mathcal{A}$ receives m points $S = \{x_i\}$ drawn iid from D and labeled with $y_i = f^*(x_i)$, then $\mathcal{A}$ outputs an $\hat{f} \in \mathcal{F}$ such that $\max_{i \in [m]} |\hat{f}(x_i) - f^*(x_i)| = 0$. For every such $\mathcal{A}$, there is a randomized uniformly ε-approximate sample compression scheme for $\mathcal{F}$ of size $O(k\tilde{m} \log(k\tilde{m}))$, where $\tilde{m} = O(d(c\varepsilon) \log(1/\varepsilon))$ and $k = O(d^*(c\varepsilon) \log(d^*(c\varepsilon)/\varepsilon))$. Furthermore, on a sample of any size m, the compression set may be computed in expected time*

$$O(mT_{\mathcal{E}}(\tilde{m})(k + \log m) + T_{\mathcal{A}}(\tilde{m}) \log(m)),$$

where $T_{\mathcal{A}}(\ell)$ is the runtime of $\mathcal{A}$ to compute $\hat{f}$ on a sample of size ℓ, $T_{\mathcal{E}}(\ell)$ is the runtime required to evaluate $\hat{f}$ on a single $x \in \mathcal{X}$, and c is a universal constant.

A key component in the above result is our construction of a generic (η, γ)-weak learner.

Definition 1.2.1. For $\eta \in [0,1]$ and $\gamma \in [0,1/2]$, we say that $f : \mathcal{X} \to \mathbb{R}$ is an an (η, γ)-*weak hypothesis* (with respect to distribution D and target $f^* \in \mathcal{F}$) if

$$\mathbb{P}_{X \sim D}(|f(X) - f^*(X)| > \eta) \leq \frac{1}{2} - \gamma.$$

Theorem 1.2.3 (Generic weak learner)**.** *Let $\mathcal{F} \subset [0,1]^{\mathcal{X}}$ be a function class with t-fat-shattering dimension $d(t)$. For some universal numerical constants $c_1, c_2, c_3 \in (0, \infty)$, for any $\eta, \delta \in (0,1)$ and $\gamma \in (0, 1/4)$, any $f^* \in \mathcal{F}$, and any distribution D, letting $X_1, \ldots, X_m$ be drawn iid from D, where*

$$m = \left\lceil c_1 \left(d(c_2 \eta) \ln\left(\frac{c_3}{\eta}\right) + \ln\left(\frac{1}{\delta}\right)\right)\right\rceil,$$

with probability at least $1 - \delta$, every $f \in \mathcal{F}$ with $\max_{i \in [m]} |f(X_i) - f^*(X_i)| = 0$ is an (η, γ)-weak hypothesis with respect to D and f^*.

In fact, our results would also allow us to use any hypothesis $f \in \mathcal{F}$ with $\max_{i \in [m]} |f(X_i) - f^*(X_i)|$ bounded below η: for instance, bounded by $\eta/2$. This can then also be plugged into the construction of the compression scheme and this criterion can be used in place of consistency in Theorem 1.2.2.

In the supplementary material, we give applications to sample compression for nearest-neighbor and bounded-variation regression.

1.3 Related work

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^\infty a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^n a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this

text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1.4 Our contribution

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2}\mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^\infty e^{-\alpha x^2}\mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2}\mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Chapter 2

Boosting Real-Valued Functions

As mentioned above, the notion of a *weak learner* for learning real-valued functions must be formulated carefully. The naïve thought that we could take any learner guaranteeing, say, absolute loss at most $\frac{1}{2} - \gamma$ is known to not be strong enough to enable boosting to ε loss. However, if we make the requirement too strong, such as in Freund and Schapire [1997] for `AdaBoost.R`, then the sample complexity of weak learning will be so high that weak learners cannot be expected to exist for large classes of functions. However, our Definition 1.2.1, which has been proposed independently by Simon [1997] and Kégl [2003], appears to yield the appropriate notion of *weak learner* for boosting real-valued functions.

2.1 The MedBoost Algorithm

In the context of boosting for real-valued functions, the notion of an (η, γ)-weak hypothesis plays a role analogous to the usual notion of a weak hypothesis in boosting for classification. Specifically, the following boosting algorithm was proposed by Kégl [2003].

2.1.1 The algorithm

As it will be convenient for our later results, we express its output as a sequence of functions and weights; the boosting guarantee from Kégl [2003] applies to the weighted quantiles (and in particular, the weighted median) of these function values.

Here we define the weighted median as

$$\mathrm{Median}(y_1, \ldots, y_T; \alpha_1, \ldots, \alpha_T) = \min\left\{ y_j : \frac{\sum_{t=1}^{T} \alpha_t \mathbb{I}[y_j < y_t]}{\sum_{t=1}^{T} \alpha_t} < \frac{1}{2} \right\}.$$

Also define the weighted *quantiles*, for $\gamma \in [0, 1/2]$, as

$$Q_\gamma^+(y_1, \ldots, y_T; \alpha_1, \ldots, \alpha_T) = \min\left\{ y_j : \frac{\sum_{t=1}^{T} \alpha_t \mathbb{I}[y_j < y_t]}{\sum_{t=1}^{T} \alpha_t} < \frac{1}{2} - \gamma \right\}$$
$$Q_\gamma^-(y_1, \ldots, y_T; \alpha_1, \ldots, \alpha_T) = \max\left\{ y_j : \frac{\sum_{t=1}^{T} \alpha_t \mathbb{I}[y_j > y_t]}{\sum_{t=1}^{T} \alpha_t} < \frac{1}{2} - \gamma \right\},$$

and abbreviate $Q_\gamma^+(x) = Q_\gamma^+(h_1(x), \ldots, h_T(x); \alpha_1, \ldots, \alpha_T)$ and $Q_\gamma^-(x) = Q_\gamma^-(h_1(x), \ldots, h_T(x); \alpha_1, \ldots, \alpha_T)$ for $h_1, \ldots, h_T$ and $\alpha_1, \ldots, \alpha_T$ the values returned by `MedBoost`.

2.1.2 Analysis

Lemma 5 Then Kégl [2003] proves the following result.

Lemma 2.1.1. *(Kégl [2003]) For a training set $Z = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ of size m, the return values of `MedBoost` satisfy*

$$\frac{1}{m} \sum_{i=1}^{m} \mathbb{I}\left[\max\left\{ \left|Q_{\gamma/2}^+(x_i) - y_i\right|, \left|Q_{\gamma/2}^-(x_i) - y_i\right| \right\} > \eta/2\right] \leq \prod_{t=1}^{T} e^{\gamma \alpha_t} \sum_{i=1}^{m} P_t(i) e^{-\alpha_t \theta_i^{(t)}}.$$

We note that, in the special case of binary classification, `MedBoost` is closely related to the well-known AdaBoost algorithm [Freund and Schapire, 1997], and the above results correspond to a standard margin-based analysis of Schapire et al. [1998].

Corollary 6 For our purposes, we will need the following immediate corollary of this, which follows from plugging in the values of α_t and using the weak learning assumption, which implies $\sum_{i=1}^{m} P_t(i)\mathbb{I}[\theta_i^{(t)} = 1] \geq \frac{1}{2} + \gamma$ for all t.

Corollary 2.1.1.1. *For $T = \Theta\left(\frac{1}{\gamma^2} \ln(m)\right)$, every $i \in \{1, \ldots, m\}$ has*

$$\max\left\{ \left|Q_{\gamma/2}^+(x_i) - y_i\right|, \left|Q_{\gamma/2}^-(x_i) - y_i\right| \right\} \leq \eta/2.$$

noend 1 `MedBoost`($\{(x_i, y_i)\}_{i\in[m]}$,T,γ,η)

Define P_0 as the uniform distribution over $\{1, \ldots, n\}$
for do$t = 0, \ldots, T$
  Call weak learner to get h_t and $(\eta/2, \gamma)$-weak hypothesis wrt $(x_i, y_i) : i \sim P_t$ (repeat until it succeeds)
  for do$i = 1, \ldots, m$
    $\theta_i^{(t)} \leftarrow 1 - 2\mathbb{I}[|h_t(x_i) - y_i| > \eta/2]$
  end for
  $\alpha_t \leftarrow \frac{1}{2}\ln\left(\frac{(1-\gamma)\sum_{i=1}^m P_t(i)\mathbb{I}[\theta_i^{(t)}=1]}{(1+\gamma)\sum_{i=1}^m P_t(i)\mathbb{I}[\theta_i^{(t)}=-1]}\right)$
  if then$\alpha_t = \infty$
    Return T copies of h_t, and $(1, \ldots, 1)$
  end if
  for $i = 1, \ldots, m$ **do**
    $P_{t+1}(i) \leftarrow P_t(i)\frac{\exp\{-\alpha_t\theta_i^{(t)}\}}{\sum_{j=1}^m P_t(j)\exp\{-\alpha_t\theta_j^{(t)}\}}$
  end for
end for
Return $(h_1, \ldots, h_T)$ and $(\alpha_1, \ldots, \alpha_T)$

for do$t = 0, \ldots, T$
  Call weak learner to get h_t and $(\eta/2, \gamma)$-weak hypothesis wrt $(x_i, y_i) : i \sim P_t$ (repeat until it succeeds)
  for do$i = 1, \ldots, m$
    $\theta_i^{(t)} \leftarrow 1 - 2\mathbb{I}[|h_t(x_i) - y_i| > \eta/2]$
  end for
  $\alpha_t \leftarrow \frac{1}{2}\ln\left(\frac{(1-\gamma)\sum_{i=1}^m P_t(i)\mathbb{I}[\theta_i^{(t)}=1]}{(1+\gamma)\sum_{i=1}^m P_t(i)\mathbb{I}[\theta_i^{(t)}=-1]}\right)$
  if then$\alpha_t = \infty$
    Return T copies of h_t, and $(1, \ldots, 1)$
  end if
  for $i = 1, \ldots, m$ **do**
    $P_{t+1}(i) \leftarrow P_t(i)\frac{\exp\{-\alpha_t\theta_i^{(t)}\}}{\sum_{j=1}^m P_t(j)\exp\{-\alpha_t\theta_j^{(t)}\}}$
  end for
end for
Return $(h_1, \ldots, h_T)$ and $(\alpha_1, \ldots, \alpha_T)$

2.2 The Sample Complexity of Weak Learning

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^{\infty} a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^{n} a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.2.1 The Notion of "Weak Learning"

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of

the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2}\mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^{\infty} e^{-\alpha x^2}\mathrm{d}x \int_{-\infty}^{\infty} e^{-\alpha y^2}\mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$\sum_{k=0}^{\infty} a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^{n} a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{-p \pm \sqrt{p^2-4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.2.2 Upper Bound on The Sample Complexity of Weak Learning

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a \sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.2.3 Tightness of The Upper Bound

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2} \sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this

text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^{\infty} a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^{n} a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{-p \pm \sqrt{p^2-4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of

the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Simon's lower bound Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^\infty a_0 q^k = \lim_{n\to\infty}\sum_{k=0}^n a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Simon's open question Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Chapter 3

From Boosting to Compression

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

3.1 Binary Classification

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^\infty a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^n a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

3.2 Real-Valued Functions

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this

text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2}\sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Chapter 4

Examples

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

4.1 Sample compression for BV functions

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^{\infty} a_0 q^k = \lim_{n\to\infty} \sum_{k=0}^{n} a_0 q^k = \lim_{n\to\infty} a_0 \frac{1-q^{n+1}}{1-q} = \frac{a_0}{1-q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

4.2 Sample compression for nearest-neighbor regression

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$a\sqrt[n]{b} = \sqrt[n]{a^n b}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of

the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\int_0^\infty e^{-\alpha x^2} \mathrm{d}x = \frac{1}{2} \sqrt{\int_{-\infty}^\infty e^{-\alpha x^2} \mathrm{d}x \int_{-\infty}^\infty e^{-\alpha y^2} \mathrm{d}y} = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

$$\sum_{k=0}^\infty a_0 q^k = \lim_{n \to \infty} \sum_{k=0}^n a_0 q^k = \lim_{n \to \infty} a_0 \frac{1 - q^{n+1}}{1 - q} = \frac{a_0}{1 - q}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no

need for special content, but the length of words should match the language.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Chapter 5

Future

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Bibliography

Noga Alon, Shai Ben-David, Nicolò Cesa-Bianchi, and David Haussler. Scale-sensitive dimensions, uniform convergence, and learnability. *Journal of the ACM*, 44(4):615–631, 1997. URL `citeseer.ist.psu.edu/alon97scalesensitive.html`.

Patrice Assouad. Densité et dimension. *Ann. Inst. Fourier (Grenoble)*, 33(3): 233–282, 1983. ISSN 0373-0956. URL `http://www.numdam.org/item?id=AIF_1983__33_3_233_0`.

Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. Assoc. Comput. Mach.*, 36(4):929–965, 1989. ISSN 0004-5411.

Richard M. Dudley. A course on empirical processes. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 1–142. Springer, Berlin, 1984.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 55(1):119–139, 1997. ISSN 0022-0000. doi: http://dx.doi.org/10.1006/jcss.1997.1504.

David Haussler, Michael Kearns, Nick Littlestone, and Manfred K Warmuth. Equivalence of models for polynomial learnability. *Information and Computation*, 95(2):129–161, 1991.

Balázs Kégl. Robust regression by boosting the median. In *Learning Theory and Kernel Machines*, pages 258–272. Springer, 2003.

Shay Moran and Amir Yehudayoff. Sample compression schemes for VC classes. *J. ACM*, 63(3):21:1–21:10, 2016. doi: 10.1145/2890490. URL `http://doi.acm.org/10.1145/2890490`.

Leonard Pitt and Leslie G Valiant. Computational limitations on learning from examples. *Journal of the ACM (JACM)*, 35(4):965–984, 1988.

David Pollard. *Convergence of Stochastic Processes.* Springer-Verlag, 1984.

Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. Boosting the margin: a new explanation for the effectiveness of voting methods. *Ann. Statist.*, 26(5):1651–1686, 1998. ISSN 0090-5364. doi: 10.1214/aos/1024691352. URL `http://dx.doi.org/10.1214/aos/1024691352`.

Hans Ulrich Simon. Bounds on the number of examples needed for learning functions. *SIAM J. Comput.*, 26(3):751–763, 1997. doi: 10.1137/S0097539793259185. URL `https://doi.org/10.1137/S0097539793259185`.

V. N. Vapnik and A. Ja. Červonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Teor. Verojatnost. i Primenen.*, 16:264–279, 1971. ISSN 0040-361x.

V. N. Vapnik and A. Ya. Chervonenkis. *Teoriya raspoznavaniya obrazov. Statisticheskie problemy obucheniya.* Izdat. "Nauka", Moscow, 1974.