technische universität
dortmund

Fakultät Statistik

Stützvektormaschinen für die Überlebenszeitanalyse mit R

Support vector machines for survival analysis with R

Césaire Fouodo

Computergestützte Statistik
Fakultät Statistik

Inhalt

Klassische Überlebenszeitanalyse

Überlebenszeit: die Zeit bis zum Auftreten eines Ereignis

Beispiele:

- Untersuchung von zwei Therapien (in der Medizin):

  vollständige Remission von einer Erkrankung oder Tod
- Überlebenszeiten von elektronischen Geräten (in der Elektrotechnik)
- Dauer der ersten Ehe (in der Soziologie)

Klassische Überlebenszeitanalyse

Überlebensfunktion

Sei $T \geq 0$ eine Zufallsvariable

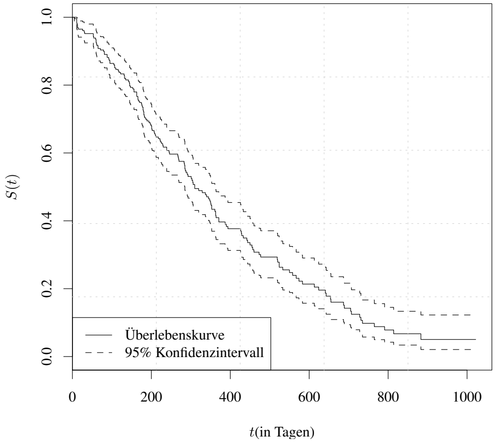
$S(t) = Pr(T > t)$ ist die **Überlebensfunktion**

Es gilt:

$$S(t) = \begin{cases} 1, & \text{falls } t = 0 \\ 0, & \text{falls } t \longrightarrow \infty \end{cases}$$

Klassische Überlebenszeitanalyse

Überlebensfunktion

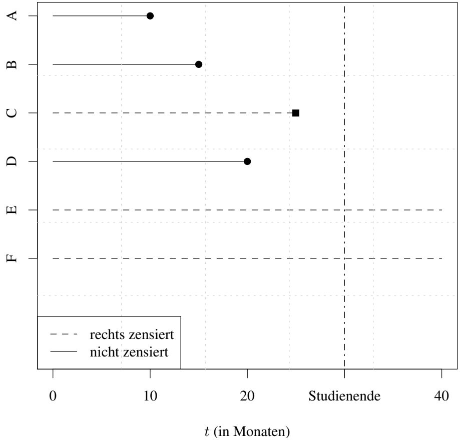


Klassische Stützvektormaschinen

Zensierung

Individuen	δ	T
A	1	10
B	1	15
C	0	25
D	1	20
E	0	30
F	0	30

δ: Status



Klassische Überlebenszeitanalyse

Das proportionale Hazard (Cox) Modell[1]:

- $h(t|X) = h_0(t)\exp(\beta' X)$,

  wobei $\beta' = (\beta_1, \beta_2, ..., \beta_d) \in \mathbb{R}^d$ der zu schätzende Parametervektor ist

- $h_0(t)$ heißt ***Basis-Hazard-Funktion***

- Schätzung der Parameter anhand der **partiellen Likelihood**

- Proportionale Hazards zwischen zwei Individuen i und j

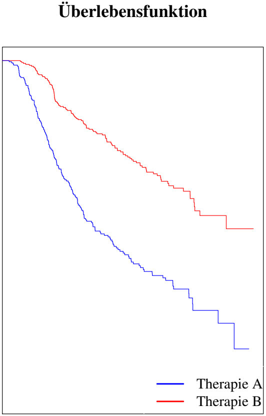
$$\frac{h(t|X_i)}{h(t|X_j)} = \frac{\exp(\beta' X_i)}{\exp(\beta' X_j)} = \exp(\beta'(X_i - X_j))$$

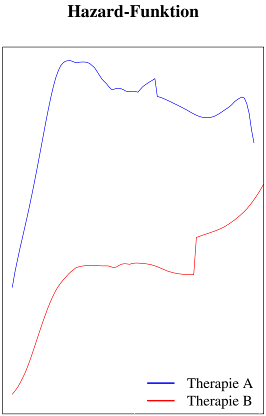
[1] D. R. Cox. 1972. 34. S. 187–220.

Klassische Überlebenszeitanalyse

Proportionalitätsannahme erfüllt

Beispiel



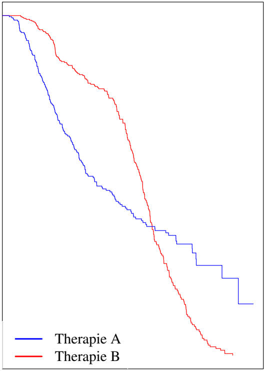


Klassische Überlebenszeitanalyse

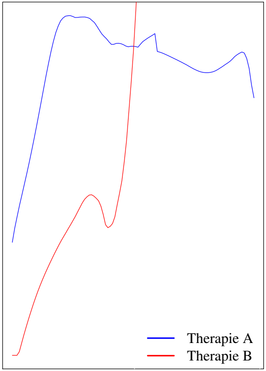
Proportionalitätsannahme nicht erfüllt

Beispiel

Überlebensfunktion



Hazard-Funktion



Klassische Überlebenszeitanalyse

Referenzmethoden dieser Studie:

- Das proportionale Cox (oder PH) Modell
- Zufallswälder für die Überlebenszeitanalyse[2]
- Das Gradient-Boosting für die Überlebenszeitanalyse[3]

[2] H. Ishwaran und U. Kogalur. 2007. 2. S. 25–31.

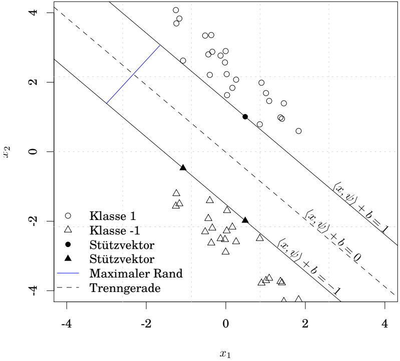
[3] J. H. Friedman. 2001. 29. S. 1189–1232.

Inhalt

Klassische Stützvektormaschinen[4]

streng linear trennbare Trainingsgruppen

Seien $X \in \mathbb{R}^d$ und $Y \in \{-1, 1\}$



SVM-Problem

Minimiere $\tau(\psi) = \frac{1}{2}\|\psi\|^2$

u.d.N. $y_i(\langle x_i, \psi \rangle + b) \geq 1$

[4] V. N. Vapnik. 1995. ISBN: 978-0-387-94559-0.

Klassische Stützvektormaschinen

Lagrangefunktion

$$L(\psi, b, \lambda) = \frac{1}{2}\|\psi\|^2 - \sum_{i=1}^{n} \lambda_i (y_i(\langle x_i, \psi \rangle + b) - 1)$$

Es folgt mit den Karush-Kuhn-Tucker (KKT)-Bedingungen[5]

$$-(y_i(\langle x_i, \psi \rangle + b) - 1) \leq 0 \tag{1a}$$

$$\lambda_i \geq 0 \tag{1b}$$

$$\lambda_i (y_i(\langle x_i, \psi \rangle + b) - 1) = 0 \tag{1c}$$

[5] H. W. Kuhn und A. W. Tucker. 1951.

Klassische Stützvektormaschinen[6]

überlappende Trainingsgruppen

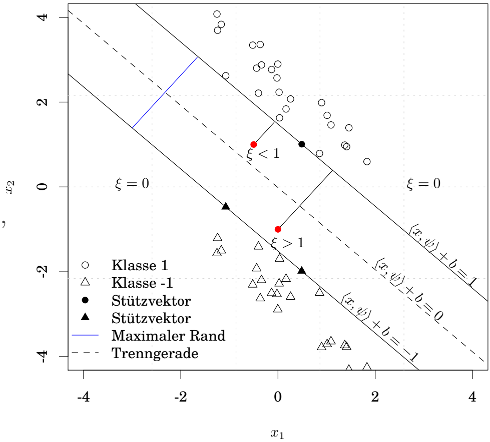
Das SVM-Problem

$$\min_{\psi,b,\xi} \quad \tau(\psi) = \frac{1}{2}\|\psi\|^2 + C\sum_{i=1}^{n}\xi_i$$

u.d.N. $\quad -\left(y_i(\langle x_i,\psi\rangle + b) + \xi_i - 1\right) \leq 0,$

$\quad \xi_i \geq 0,$

wobei C ein Tuningparameter ist

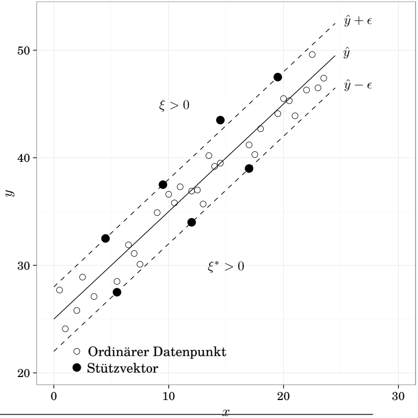


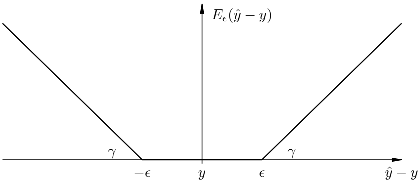
[6] V. N. Vapnik. 1995. ISBN: 978-0-387-94559-0.

Klassische Stützvektormaschinen[7]

Stützvektorregression

ϵ-sensitive Fehlerfunktion





[7] C. Bishop. 2007. ISBN: 978-0-387-31073-2.

Klassische Stützvektormaschinen

Stützvektorregression

Das SVM-Problem[8]

$$\min_{\psi,b,\xi,\xi*} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

u.d.N. $$y_i - \langle \psi, F(x_i) \rangle - b \leq \epsilon + \xi_i,\ \xi_i \geq 0,\ i = 1, ..., n$$

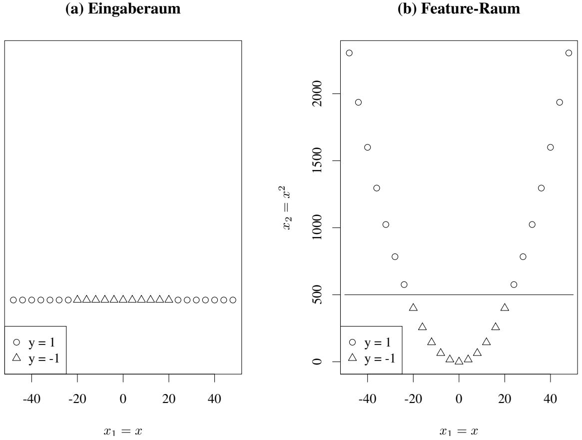
$$\langle \psi, F(x_i) \rangle + b - y_i \leq \epsilon + \xi_i^*,\ \xi_i^* \geq 0,\ i = 1, ..., n$$

[8] C. Bishop. 2007. ISBN: 978-0-387-31073-2.

Klassische Stützvektormaschinen

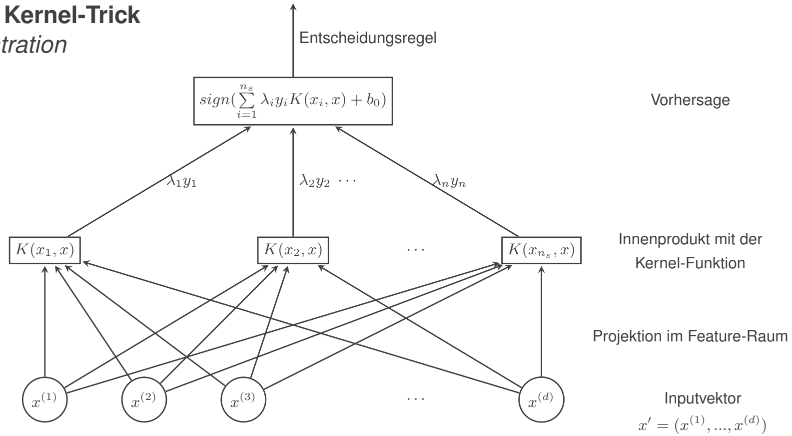
Der Kernel-Trick

Beispiel



Klassische Stützvektormaschinen

Der Kernel-Trick
Illustration



Klassische Stützvektormaschinen

Der Kernel-Trick

Verwendete Kernel

- linearer Kernel: $K(x_i, x_j) = \sum_{p=1}^{d} x_{ip} x_{jp}$
- additiver Kernel[9]: $K(x_i, x_j) = \sum_{p=1}^{d} K_p(x_i^p, x_j^p)$, wobei
  - $K_p(x_i^p, x_j^p) = \frac{c - |x_i^p - x_j^p|}{c}$

    für stetige und ordinale Variablen, $c = \max_p - \min_p$
  - $K_p(x_i^p, x_j^p) = \begin{cases} 1, & \text{falls } x_i^p = x_j^p, \\ 0 & \text{sonst} \end{cases}$

    für binäre und kategorielle Variablen

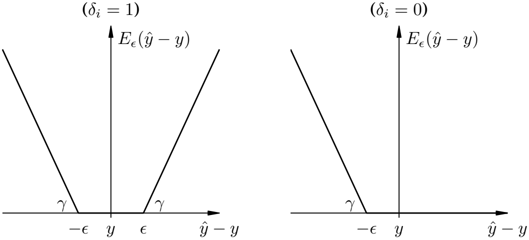
[9] A. Daemen und B. De Moor. 2009. S. 5913–5917.

Inhalt

Stützvektormaschinen für die Überlebenszeitanalyse

Regressionsansatz[10] [11]

Regression



[10] P. K. Shivaswamy, W. Chu und M. Jansche. 2007. S. 655–660.

[11] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Regressionsansatz[12] [13]

Regression

Formulierung des SVM-Problems

$$\min_{\psi,b,\xi,\xi*} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

$$\text{u.d.N.} \quad y_i - \langle \psi, F(x_i) \rangle - b \leq \xi_i,$$

$$\langle \psi, F(x_i) \rangle + b - y_i \leq \xi_i^*,$$

$$\xi_i, \xi_i^* \geq 0,$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{y} = \sum_{i=1}^{n} (\beta_i - \delta_i \beta_i^*) K(x_i, x^*) + b_0$$

[12] P. K. Shivaswamy, W. Chu und M. Jansche. 2007. S. 655–660.

[13] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Regressionsansatz[14] [15]

Regression

Formulierung des SVM-Problems

$$\min_{\psi,b,\xi,\xi*} \quad \tfrac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$\text{u.d.N.} \quad y_i - \langle\psi, F(x_i)\rangle - b \leq \xi_i,$$

$$\delta_i(\langle\psi, F(x_i)\rangle + b - y_i) \leq \xi_i^*,$$

$$\xi_i, \xi_i^* \geq 0,$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{y} = \sum_{i=1}^{n}(\beta_i - \delta_i\beta_i^*)K(x_i, x^*) + b_0$$

[14] P. K. Shivaswamy, W. Chu und M. Jansche. 2007. S. 655–660.

[15] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Ranking-Ansatz[16] [17]

- Datenpunkte werden in steigender Reihenfolge nach den Überlebenszeiten geordnet
- Zwei Datenpunkte i und j ($i < j$) vergleichbar, wenn $\delta_i = 1 \Rightarrow t_i < t_j$
- i und j sind konkordant, wenn gilt: $u(x_i) < u(x_j)$, wobei u eine Vorhersagefunktion ist

$$\text{Sei } v_{ij} = \begin{cases} 1, & \text{falls } i \text{ und } j \text{ vergeichbar sind,} \\ 0 & \text{sonst} \end{cases}$$

[16] L. Evers und C.-M. Messow. 2008. 24. S. 1632–1638.

[17] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Ranking-Ansatz[18]

*vanbelle*1

$$\min_{\psi,\varepsilon} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{\substack{j<i \\ \delta_i=1}} v_{ij}\varepsilon_{ij}$$

$$\text{u.d.N.} \quad \langle \psi, F(x_i) \rangle - \langle \psi, F(x_j) \rangle \geq 1 - \varepsilon_{ij},$$

$$\varepsilon_{ij} \geq 0, \; i, j = 1, ..., n,$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{u}(x^*) = \sum_{j<i} \alpha_{ij} \left(K(x_i, x^*) - K(x_j, x^*) \right)$$

[18] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Ranking-Ansatz[19]
*vanbelle*2

k-nächste Nachbarn

$$\mathcal{C}_i = \{(i, j) : t_j \text{ ist } k\text{-nächste bei } t_i\}, i = 1, ..., n,$$

- $\mathcal{C}_i$ bestimmt die k-nächsten Nachbarn des Datenpunktes x_i
- x_i wird mit seinem nächsten Nachbar verglichen
- $\bar{j}(i)$ bezeichnet den nächsten Nachbar von i, wenn er existiert
- Voraussetzung: $\delta_{(1)} = 1$

[19] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Ranking-Ansatz[20]
*vanbelle*2

Formulierung des SVM-Problems

$$\min_{\psi,\varepsilon} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n} \varepsilon_i$$

$$\text{u.d.N.} \quad \langle \psi, F(x_i) \rangle - \langle \psi, F(x_{\bar{j}(i)}) \rangle \geq y_i - y_{\bar{j}(i)} - \varepsilon_i,$$

$$\varepsilon_i \geq 0,$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{u}(x^*) = \sum_{i=1}^{n} \alpha_i \left(K(x_i, x^*) - K(x_{\bar{j}(i)}, x^*) \right)$$

[20] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Der hybride Ansatz[21] (**Ranking** + **Regression**)

hybrid

Das SVM-Problem

$$\min_{\psi,\varepsilon} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n} \varepsilon_i$$

$$\text{u.d.N.} \quad \langle \psi, F(x_i) \rangle - \langle \psi, F(x_{\bar{j}(i)}) \rangle \geq y_i - y_{\bar{j}(i)} - \varepsilon_i,$$

$$\varepsilon_i \geq 0$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{u}(x^*) = \sum_{i=1}^{n} \langle \alpha_i \left(F(x_i) - F(x_{\bar{j}(i)}) \right) \quad , F(x^*) \rangle$$

[21] Van Belle et al. 2011. 53. S. 107–118.

Stützvektormaschinen für die Überlebenszeitanalyse

Der hybride Ansatz[22] **(Ranking + Regression)**

hybrid

Das SVM-Problem

$$\min_{\psi,\varepsilon,b,\xi,\xi^*} \quad \frac{1}{2}\|\psi\|^2 + \gamma \sum_{i=1}^{n} \varepsilon_i + \mu \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

$$\text{u.d.N.} \quad \langle \psi, F(x_i)\rangle - \langle \psi, F(x_{\bar{j}(i)})\rangle \geq y_i - y_{\bar{j}(i)} - \varepsilon_i,$$

$$y_i - \langle \psi, F(x_i)\rangle - b \leq \xi_i,$$

$$\delta_i(\langle \psi, F(x_i)\rangle + b - y_i) \leq \xi_i^*,$$

$$\varepsilon_i \geq 0, \xi_i \geq 0, \xi_i^* \geq 0$$

Vorhersage gegeben einen Datenpunkt x^*

$$\hat{u}(x^*) = \sum_{i=1}^{n} \langle \alpha_i \left(F(x_i) - F(x_{\bar{j}(i)})\right) + (\beta_i - \delta_i \beta_i^*) F(x_i), F(x^*)\rangle + b_0$$

[22] Van Belle et al. 2011. 53. S. 107–118.

Inhalt

1. Klassische Überlebenszeitanalyse
2. Klassische Stützvektormaschinen
3. Stützvektormaschinen für die Überlebenszeitanalyse
4. **Umsetzung in R und Vergleichsstudie**
5. Zusammenfassung
6. Diskussion und Ausblick

Umsetzung in R und Vergleichsstudie

Verwendete R-Pakete

1. `survival`[23]
2. `randomForestSRC`[24]
3. `mboost`[25]
4. **`survivalsvm`**
5. `mlr`[26]

[23] T. M. Therneau. R package version 2.38. 2015.
[24] H. Ishwaran und U. Kogalur. R package version 2.3.0. manual, 2016.
[25] H. Binder. R package version 1.4. 2013.
[26] B. Bischl u. a. R package version 2.9. 2016.

Umsetzung in R und Vergleichsstudie

Das R-Paket survivalsvm

Verwendungsbeispiel

```
library(survivalsvm)
library(survival)  # 'survival' ist für den Befehl 'Surv' benötigt.
data(veteran, package = "survival")
```

Umsetzung in R und Vergleichsstudie

Das R-Paket survivalsvm

Verwendungsbeispiel

```
library(survivalsvm)
library(survival)  # 'survival' ist für den Befehl 'Surv' benötigt.
data(veteran, package = "survival")
```

```
# Bestimmung der Trainings- und Testdatensätze.
set.seed(123)
n <- nrow(veteran)
train.index <- sample(1:n, 120, replace = FALSE)
test.index <- setdiff(1:n, train.index)
```

Umsetzung in R und Vergleichsstudie

Das R-Paket survivalsvm

Verwendungsbeispiel

```
library(survivalsvm)
library(survival)  # 'survival' ist für den Befehl 'Surv' benötigt.
data(veteran, package = "survival")
```

```
# Bestimmung der Trainings- und Testdatensätze.
set.seed(123)
n <- nrow(veteran)
train.index <- sample(1:n, 120, replace = FALSE)
test.index <- setdiff(1:n, train.index)
```

```
# Anpassung des auf dem Regressionsansatz basierenden SVM-Modells anhand der
# Trainingsdaten.
survsvm.reg <- survivalsvm(Surv(time, status) ~ ., subset = train.index, data = veteran,
    type = "regression", gamma.mu = 1, opt.meth = "quadprog", kernel = "add_kernel")
```

Umsetzung in R und Vergleichsstudie

Das R-Paket survivalsvm

Verwendungsbeispiel

```
print(survsvm.reg)
```

```
##
## survivalsvm result
##
## Call:
##
##  survivalsvm(Surv(time, status) ~ ., subset = train.index, data = veteran, type = "regression", gamm
##
## Survival svm approach              : regression
## Type of Kernel                     : add_kernel
## Optimization solver used           : quadprog
## Number of support vectors retained : 104
```

Umsetzung in R und Vergleichsstudie

Das R-Paket `survivalsvm`

Verwendungsbeispiel

```r
# Vorhersage für die Testindex.
predict(object = survsvm.reg, newdata = veteran, subset = test.index)
```

```
##
##
## survivalsvm prediction
##
## Type of survivalsvm                       : regression
## Type of kernel                            : add_kernel
## Optimization solver used in model         : quadprog
## predictions                               : 154.9441 138.9379 129.4329 138.882 126.5404 ...
```

Umsetzung in R und Vergleichsstudie

Verschachtelte Kreuzvalidierung

Iteration	**Teildatensätze**				**Fehler**
$j = 1$	D_1	D_2	$\cdots$	D_L	θ_1
	Testfehler	Tuning mittels der K-fachen CV			
$j = 2$	D_1	D_2	$\cdots$	D_L	θ_2
		Testfehler	Tuning mittels K-facher CV		
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$j = L$	D_1	D_2	$\cdots$	D_L	θ_L
	Tuning mittels der K-fachen CV			Testfehler	

Geschätzter Vorhersagefehler $\hat{\theta}$

Umsetzung in R und Vergleichsstudie

Datensätze

Datensatz	**Größe (Beob. $\times$ Var.)**	**Studie**
Veteran	137×8	*Veteran's administration lung cancer Trial*[27]
LVR	129×10	*Interpretation of a Leukemia Trial Stopped Early*[28]
LT	129×10	*Interpretation of a Leukemia Trial Stopped Early*
GBCSG2	686×10	*German Breast Cancer Study Group*[29]
MLC	228×10	*Mayo clinic lung cancer*[30]

[27] J. D. Kalbfleisch und R. L. Prentice. 2002. ISBN: 978-0-471-36357-6.

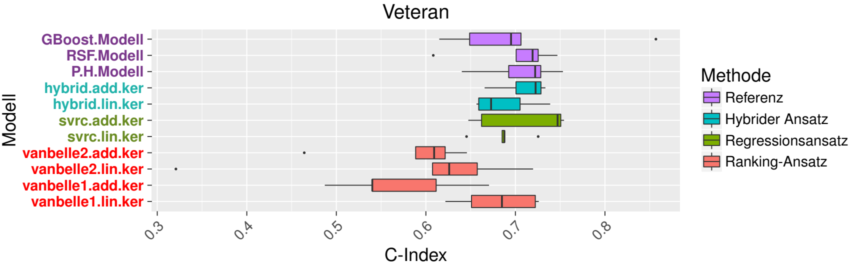
[28] S. Ermerson und P. Banks. 1994.

[29] M. Schumacher u. a. 1994. 12. S. 2086 –2093.

[30] C. L. Loprinzi u. a. 1994. 12. S. 601–607.

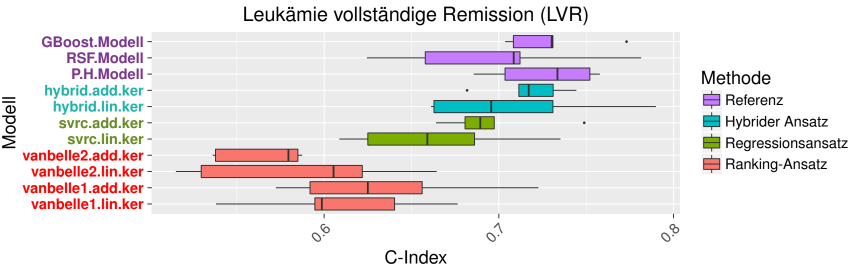
Umsetzung in R und Vergleichsstudie

Boxplot der C-Indizes



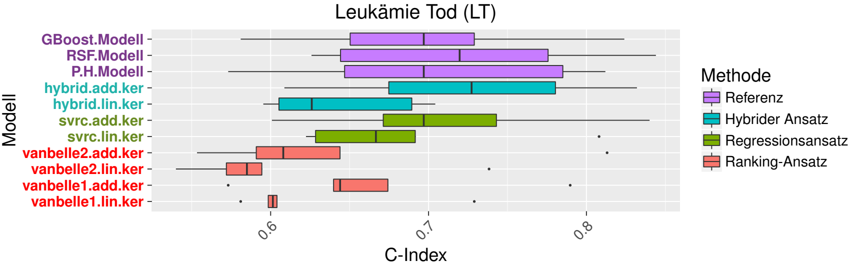
Umsetzung in R und Vergleichsstudie

Boxplot der C-Indizes



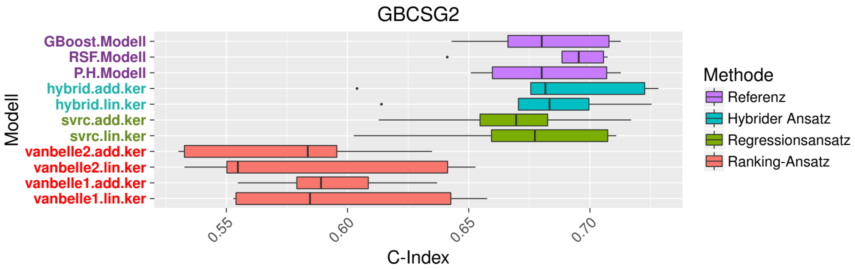
Umsetzung in R und Vergleichsstudie

Boxplot der C-Indizes



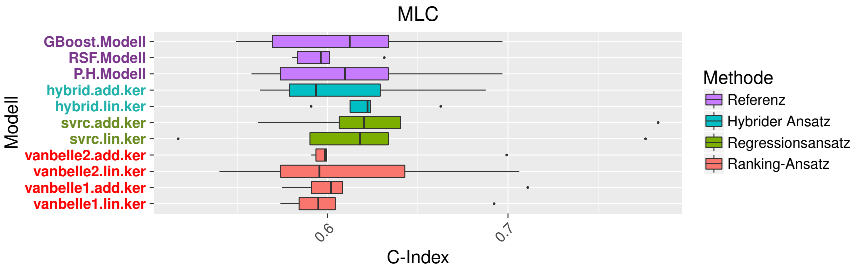
Umsetzung in R und Vergleichsstudie

Boxplot der C-Indizes



Umsetzung in R und Vergleichsstudie

Boxplot der C-Indizes



Inhalt

Zusammenfassung

- **Referenzmodelle**
  - Cox Modell
  - Zufallswälder für die Überlebenszeitanalyse
  - Grandient-Boosting für die Überlebenszeitanalyse

Zusammenfassung

- **Referenzmodelle**
  - Cox Modell
  - Zufallswälder für die Überlebenszeitanalyse
  - Grandient-Boosting für die Überlebenszeitanalyse
- **SVM-Modelle**
  - Regressionsansatz
  - Ranking-Ansatz
  - hybrider Ansatz (Ranking + Regression) $\Rightarrow$ bessere Ergebnisse

Inhalt

Diskussion und Ausblick

- **Diskussion**
  - *hybrid*-Modell: zwei Tuningparameter
  - Andere SVM-Modelle: ein Tuningparameter

    $\Rightarrow$ längere Laufzeit bei hybrid-Modell, dafür besser
  - Die Gütemaße der SVM-Modelle sind von den Kernelfunktionen beeinflusst

Diskussion und Ausblick

- **Diskussion**
  - *hybrid*-Modell: zwei Tuningparameter
  - Andere SVM-Modelle: ein Tuningparameter

    ⇒ längere Laufzeit bei hybrid-Modell, dafür besser
  - Die Gütemaße der SVM-Modelle sind von den Kernelfunktionen beeinflusst
- **Ausblick:**
  - Das R-Paket und den Methodenvergleich veröffentlichen
  - Die SVM für die Überlebenszeitanalyse mit noch mehr Verfahren vergleichen

Danke für die Aufmerksamkeit!

Umsetzung in R und Vergleichsstudie

Datensatz	Methode	Kernel	C-Index (SD_{CI})	Log-Rank (SD_{LR})	Hazard Rate (SD_{HR})
Veteran	vanbelle1	lineare	0, 68 (0, 05)	6, 72 (4, 72)	0, 12 (0, 07)
		additive	0, 57 (0, 07)	2, 90 (4, 30)	0, 46 (0, 34)
	vanbelle2	lineare	0, 59 (0, 15)	3, 70 (3, 72)	1, 39 (2, 73)
		additive	0, 59 (0, 07)	1, 99 (1, 40)	0, 67 (0, 79)
	svrc	lineare	0.69 (0, 03)	4, 32 (5, 82)	0, 11 (0, 07)
		additive	0, 71 (0, 05)	5, 22 (2, 88)	0, 15 (0, 13)
	hybrid	lineare	0.69 (0, 04)	4, 31 (4, 62)	0, 13 (0, 07)
		additive	0, 71(0, 02)	5, 18 (2.35)	0, 10 (0.04)
	P.H. Modell		0, 71 (0, 04)	7, 77 (7, 95)	0, 09 (0, 06)
	RSF		0, 70 (0, 05)	6, 31 (3, 69)	0, 08 (0, 06)
	GBoost		0, 70 (0, 09)	8, 79 (6, 98)	0, 10 (0, 08)
LVR	vanbelle1	lineare	0, 61 (0, 05)	1, 92 (2, 08)	0, 20 (0, 20)
		additive	0, 63 (0, 06)	5, 76 (4, 58)	0, 18 (0, 20)
	vanbelle2	lineare	0, 59 (0.06)	2, 04 (2, 77)	0, 37 (0, 23)
		additive	0, 57 (0.03)	1, 20 (0, 78)	0, 43 (0, 26)
	svrc	lineare	0, 66 (0, 05)	3, 33 (1, 60)	0, 08 (0, 07)
		additive	0, 70 (0, 03)	3, 68 (3, 69)	0, 11 (0, 06)
	hybrid	lineare	0, 71 (0, 05)	5, 29 (4, 37)	0, 05 (0, 03)
		additive	0, 72 (0, 02)	5, 31 (2, 44)	0, 06 (0, 04)
	P.H. Modell		0, 73 (0, 03)	4, 42 (3, 49)	0, 05 (0, 02)
	RSF		0, 70 (0, 06)	5, 66 (4, 39)	0, 10 (0, 10)
	GBoost		0, 73 (0, 03)	3, 82 (1, 91)	0, 05 (0, 03)

Umsetzung in R und Vergleichsstudie

Datensatz	Methode	Kernel	C-Index (σ_{CI})	Log-Rank (σ_{LR})	Hazard Rate (σ_{HR})
LT	vanbelle1	linearer	$0,62\,(0,06)$	$1,90\,(2,18)$	$0,31\,(0,20)$
		additive	$0,66\,(0,08)$	$4,31\,(5,41)$	$0,25\,(0,25)$
	vanbelle2	lineare	$0,61\,(0,08)$	$1,76\,(2.89)$	$0,37\,(0,26)$
		additive	$0,64\,(0,10)$	$1,87\,(1,99)$	$0,28\,(0,25)$
	svrc	lineare	$0,68\,(0,08)$	$4,38\,(4,89)$	$0,12\,(0,08)$
		additive	$0,71\,(0,09)$	$7,10\,(6,05)$	$0,12\,(0,09)$
	hybrid	lineare	$0,68\,(0,05)$	$4,50\,(4.14)$	$0,14\,(0,10)$
		additive	$0,72\,(0,09)$	$6,89\,(6,45)$	$0,09\,(0,09)$
	P.H. Modell		$0,70\,(0,10)$	$7,38\,(6,06)$	$0,08\,(0,08)$
	RSF		$0,72\,(0.09)$	$5,66\,(4,39)$	$0,10\,(0,11)$
	GBoost		$0,69\,(0,09)$	$8,51\,(6,72)$	$0,09\,(0,07)$
GBCSG2	vanbelle1	lineare	$0,60\,(0,05)$	$7,57\,(6,63)$	$0,16\,(0,10)$
		additive	$0,59\,(0,03)$	$4,59\,(1,59)$	$0,16\,(0,09)$
	vanbelle2	lineare	$0,59\,(0,06)$	$4,58\,(4,36)$	$0,23\,(0,23)$
		additive	$0,58\,(0,04)$	$2,92\,(1,78)$	$0,25\,(0,13)$
	svrc	lineare	$0,67\,(0,04)$	$14,81\,(8,79)$	$0,05\,(0,05)$
		additive	$0,67\,(0,04)$	$10,93\,(7,94)$	$0,10\,(0,08)$
	hybrid	lineare	$0,68\,(0,04)$	$13,26\,(6,29)$	$0,03\,(0,03)$
		additive	$0,68\,(0,05)$	$16,13\,(8,40)$	$0,07\,(0,08)$
	P.H. Modell		$0,68\,(0,03)$	$10,66\,(1,70)$	$0,02\,(0,02)$
	RSF		$0,69\,(0,03)$	$16,18\,(8,38)$	$0,10\,(0,06)$
	GBoost		$0,68\,(0,03)$	$11,69\,(2,48)$	$0,02\,(0,02)$

Umsetzung in R und Vergleichsstudie

Datensatz	Methode	Kernel	C-Index (σ_{CI})	Log-Rank (σ_{LR})	Hazard Rate (σ_{HR})
MLC	vanbelle1	lineare	$0,61\ (0,05)$	$2,03\ (2,07)$	$0,46\ (0,28)$
		additive	$0,62\ (0,05)$	$1,83\ (2,28)$	$0,49\ (0,33)$
	vanbelle2	lineare	$0,61\ (0,06)$	$1,42\ (1,66)$	$0,38\ (0,27)$
		additive	$0,62\ (0,05)$	$1,73\ (2,26)$	$0,48\ (0,35)$
	svrc	lineare	$0,63\ (0,09)$	$2,59\ (3,11)$	$0,44\ (0,36)$
		additive	$0,64\ (0,08)$	$3,37\ (5,81)$	$0,45\ (0,31)$
	hybrid	lineare	$0,62\ (0,03)$	$2,36\ (2,52)$	$0,28\ (0,15)$
		additive	$0,61\ (0,05)$	$1,18\ (1,43)$	$0,45\ (0,34)$
	P.H. Modell		$0,61\ (0.05)$	$1,39\ (1.18)$	$0,30\ (0.22)$
	RSF		$0,60\ (0.02)$	$0,34\ (0,40)$	$0,39\ (0,11)$
	GBoost		$0,61\ (0,06)$	$1,83\ (1,62)$	$0,32\ (0,22)$