Maschinelles Lernen
Übungsblatt 3

Dr. Steffen Rendle, Christoph Freudenthaler
Wirtschaftsinformatik und Maschinelles Lernen (ISMLL)
Universität Hildesheim

9. November 2010
Abgabe bis 19. November

Aufgabe 1: Multiple Lineare Regression

Auf einer Website, die DVD-Bewertungen sammelt und mit den gesammelten Bewertungen ihren Benutzern neue DVDs empfiehlt, haben unter anderem zwei Benutzer die folgenden Bewertungen (1 Stern ist die schlechteste, 5 Sterne die beste) abgegeben:

Index	**Benutzer**	**Film**	**Bewertung**
1	A	*The Big Lebowski*	4 Sterne
2	A	*Brazil*	2 Sterne
3	A	*Titanic*	5 Sterne
4	B	*Brazil*	3 Sterne
5	B	*The Godfather*	4 Sterne
6	B	*Toy Story*	4 Sterne

Drei verschiedene Vorhersageverfahren würden anhand der restlichen gesammelten Bewertungen folgende Vorhersagen treffen:

Index	$\hat{r}_s$	$\hat{r}_r$	$\hat{r}_k$
1	3.7	3.8	3.9
2	2.4	2.5	2.3
3	2.2	3.0	4.1
4	3.2	3.1	2.9
5	4.7	4.4	4.2
6	4.1	3.9	4.2

a)

Berechnen Sie für jedes Verfahren den durchschnittlichen absoluten und den durchschnittlichen quadratischen Fehler im Vergleich zu den tatsächlichen Bewertungen.

b)

Ein Modell für die Kombination der ersten beiden Verfahren ist

$$r(x) = \beta_0 + \beta_1 \cdot \hat{r}_s(x) + \beta_2 \cdot \hat{r}_r(x) + \epsilon$$

Berechnen Sie die Schätzungen $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ mit Hilfe der in der Vorlesung vorgestellten Methode. Verwenden Sie zur Lösung des auftretenden linearen Gleichungssystems das Gaußsche Eliminationsverfahren. Geben Sie die Gleichungen des zu lösenden Systems explizit an. Geben Sie die Zwischenschritte (gerundet auf zwei Nachkommastellen) in Matrixschreibweise an.

c)

Berechnen Sie für das kombinierte Verfahren die Residuenquadratsumme, den durchschnittlichen absoluten und den durchschnittlichen quadratischen Fehler zu den bekannten Daten. Wie aussagekräftig sind die so berechneten Fehlermaße? Begründen Sie!

d)

Berechnen Sie die kombinierte Vorhersage für $\hat{r}_s(x) = 3.0$ und $\hat{r}_r(x) = 4.6$. Welcher negative Umstand fällt Ihnen dabei auf? Was passt nicht mit den Parametern $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$? Wie könnte man solche Ergebnisse verhindern?

Aufgabe 2: Nicht-lineare Regression als lineare Regression in höherdimensionalen Räumen

Index	x	y
1	-2	0.135
2	-1	0.368
3	0	1
4	1	2.718
5	2	7.389
6	3	20.08

a)

Gegeben sei obiger Datensatz. Berechnen Sie die Modellparameter des linearen Modells $M1: \ \hat{y} = x\beta_1$ und erstellen Sie ein Streudiagramm der Variablen x und y und zeichnen Sie die Vorhersagefunktion des Modells ein.

b)

Berechnen Sie als nächstes die Modelle $M2: \ \hat{y} = x\beta_1 + x^2\beta_2$ und $M3: \ \hat{y} = x\beta_1 + x^2\beta_2 + x^3\beta_3$ und zeichnen Sie die Vorhersagefunktion ebenfalls in das (gleiche) Streudiagramm. Was fällt ihnen auf?

c)

Berechnen Sie die Taylorpolynomdarstellung der Exponentialfunktion $y = e^x$. Vergleichen Sie die Koeffizienten des Taylorpolynoms mit denen aus Model $M3$. Warum entsprechen die geschätzten Koeffizienten nicht den Koeffizienten des Taylorpolynoms?

Aufgabe 3: Ridge Regression

a)

Zeigen Sie, dass die Formel zur Parameterschätzung des linearen Modells mit Ridge Regression wiefolgt aussieht:

$$\hat{\beta} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

b)

Berechnen Sie für das lineare Modell aus Aufgabe 1b) die Parameterschätzer per Ridge Regression mit $\lambda = 0.1$.