



Klausur, Fragen

Business Analytics (IN2028) (Technische Universität München)

Rekonstruktion der Business Analytics Klausur aus dem WS2015/2016 (Gedächtnisprotokoll; Klausur war eigentlich auf Englisch)

Area 1 - Regression

Coeffizient interpretieren usw

Jahreseinkommen wurde untersucht anhand von diversen Variablen

Income

Age {20, ..., 70}

Male: {0,1}

a) i) Welche sind abhängig, welche unabhängig?

Cross sectional, time-series oder panel data?

	coefficient	p value	Significance level
Intercept			
Age			
Male			

Welche Variablen sind signifikant?

Wie wirkt sich Age und Male auf das Einkommen aus?

Die Leute werden alle noch einmal ein Jahr später befragt. Was hat sich geändert?

Cross sectional, time-series oder panel data?

was sagt der p-value in Bezug auf die Signifikanz?

Area 2 - Clustering

3 Instanzen mit (Anzahl Personen, Anzahl Autos)

a) 2 Means Clustering durchführen (2 Centroids gegeben)

i) Instance 1: (2,2)

ii) Instance 2: (4,4)

iii) Instance 3: (7,4)

iv) Cluster A starting center (2,1)

v) Cluster B starting Center (6,4)?

b) Hierarchical Clustering (Bottom-up, minimum/single linkage clustering)

c) Ergebnis aus a) und b) vergleichen und beurteilen

Area 3 - Gain Curve & paired t-test

Folgende Werte sind gegeben

ID		Classifier 1	Classifier 2
1	+	0,98	0,97
2	-	0,86	0,84
3	+	0,79	0,77
4	-	0,67	0,65

Table 4

a) Zeichne 2 Gain kurven mit 25% Schritten in ein gegebenes Koordinatensystem und beschrifte die Achsen. Gebe gesondert den Wert für X und Y an nach jedem Schritt.

ID	State	X	Y
1	+		
2	+		
3	-		
4	-		

Table 5

ID	State	X	Y
1	+		
2	-		
3	+		
4	-		

Table 6

b) Die Classifier haben keinen signifikanten unterschied bei einem Signifikanzniveau von 5%. Formel für paired t-test und d sowie eine t-tabelle waren gegeben.

Area 4 - Supp und Conf

Lord of the Rings (L)	Forest Gump (F)	Noch ein Film (N)	The Hobbit (H)
+	-	-	+
+	-	+	+
+	+	+	+
-	+	+	-
-	+	-	+

- Apriori Algorithm, um alle Werte mit $\text{supp} \geq 0.6$ zu finden.
- Finde alle Paare mit $\text{conf} = 1$
- Finde alle Regeln mit $\text{supp} \geq 0.4$ und $\text{conf} \geq 0.6$
- Eine Instanz wird hinzugefügt und nochmals Regeln bestimmen
+ | - | - | -