

0 Ergänzung zu der Aufgabestellung

Die Beispieldatenbank „bank.mdb“ (Access → Datenmodell, SQL) können Sie in folgendem Verzeichnis finden: Wwz-ivory:\Wi\Lehre\Eus\Beispiele\BeispieleAufgaben\Bank

1 Kreditkartenbetrug

a) Folgende Attribute eignen sich zur Klassifikation oder Vorhersage betrügerischer Transaktionen:

- Transaktionstyp (Kreditkarte)
- Transaktionszahl (Abweichung vom Durchschnitt)
- Durchschn. Transaktionszahl
- Karte neu (Missbrauch v.a. bei neuer Karte)
- Kartenverlust
- frühere Betrugsfälle

b) Weitere interessante Attribute wären:

- Ungültige Transaktionen (Ist beispielsweise versucht worden das Kartenlimit bei einer Transaktion zu überschreiten?)
- Kartenstatus (Ist die Karte gesperrt?)
- Betrugszeitpunkt (An welchem Wochentag, Monat, zu welcher Stunde?)

c) SQL-Code (MS Access):

```
SELECT DISTINCT
    TRANSAKTION.[Konto-ID], Kartenverlust,
    Transaktionszahl, Transaktionsdurchschnitt,
    Betrugsfälle, [Karte neu]
FROM
    TRANSAKTION, KONTO, KONTOKORRENT
WHERE
    Transaktionstyp = "Kreditkarte"
    AND Datum >= #6/1/97# AND Datum < #7/1/97#
    AND Transaktionszahl - Transaktionsdurchschnitt>5
    AND TRANSAKTION.[Konto-ID] = KONTO.[Konto-ID]
    AND KONTO.[Konto-ID] = KONTOKORRENT.[Konto-ID]
ORDER BY
    TRANSAKTION.[Konto-ID];
```

Erläuterungen:

i) Transaktionstyp = "Kreditkarte":

Nur Kreditkartentransaktionen sollen verwendet werden

ii) AND Datum >= #6/1/97# AND Datum < #7/1/97#:

Betrachtung einer bestimmten Periode (hier Juni 1997)

Achtung: Je nach Sprachversion müssen die Datumsangaben angepasst werden!

iii) AND Transaktionszahl - Transaktionsdurchschnitt>5:

Nur bei bedeutender Erhöhung (hier min. 5 gewählt) der betreffenden Transaktionsanzahl von der Ø Transaktionsanzahl soll der Datensatz angezeigt werden.

- d) Geeignete Data Mining-Methoden:
Regelinduktion, Clusteranalyse, Neuronale Netze, Regressionsanalyse

2 Retention Management

- a) Geeignete Vorhersageattribute:
- Kontenzahl (je mehr desto loyaler)
 - Konto bei Konkurrenz (wenn ja → eher Wechsel der Bank)
 - Alter des HH-Vorstands (→ Lebenszyklus)
 - Transaktionszahl (Rückgang?)
 - Nächste Filiale (→ Entfernung, Öffnungszeiten etc.)
- b) Weitere interessante Attribute wären:
- Brancheninformation
 - Anzahl Bankenwechsel
 - Anzahl Kontenänderungen
 - Bankwechsel (z.B. Bank diesen Monat verlassen?)
- c) Geeignete Data Mining-Methoden:
- Clusteranalyse: Gibt es Gruppen, die eher als andere zu einem Bankenwechsel tendieren?
 - Statistische Verfahren: Trend Analysen → Spezielle Betrachtung der Fakten in Abhängigkeit der Zeitdimension etc.
- d) SQL-Code (MS Access):

```
PARAMETERS
    [Heutiges Datum] DateTime;
TRANSFORM
    Sum(KONTOKORRENT.Transaktionszahl) AS Transaktionszahl
SELECT
    KONTOKORRENT.[Konto-ID]
FROM
    KONTOKORRENT
WHERE
    (KONTOKORRENT.Datum>=([Heutiges Datum]-181))
GROUP BY
    KONTOKORRENT.[Konto-ID]
PIVOT
    DatePart("m",Datum);
```

Erläuterungen:

- i) PARAMETERS [Heutiges Datum] DateTime;
Variablendeklaration
- ii) TRANSFORM Sum(KONTOKORRENT.Transaktionszahl)
Transform enthält die Aggregatfunktion der Pivottabelle. (Sum wäre hier eigentlich gar nicht notwendig, da das Attribut *Transaktionszahl* schon aufsummiert ist und nur einmal pro Monat vorkommt. Eine Aggregatsfunktion wird hier aber verlangt.)
- iii) (KONTOKORRENT.Datum>=([Heutiges Datum]-181))
Daten der letzten 6 Monate (181 Tage)
- iv) PIVOT DatePart("m",Datum);
Die Zeitdimension (Datum) wird nach der Monatszahl aufgelistet.