

# **Benutzerhandbuch „Entscheidungsbaum“**

Verfasser: Arne Krawielitzki, Alexander Landmann, Engin Yilmaz, Jan Zimmer

- 1. Ziel des Programms**
- 2. Systemanforderungen**
- 3. Ansichten**
  - 3.1. Tabellenansicht**
    - 3.1.1. Zeile hinzufügen**
    - 3.1.2. Zeile löschen**
    - 3.1.3. Spalte hinzufügen**
    - 3.1.4. Spalte löschen**
    - 3.1.5. Zielattribut setzen**
  - 3.2. Baum interaktiv**
  - 3.3. Baum automatisch**
- 4. Funktionen**
  - 4.1. Datei öffnen**
  - 4.2. Datei speichern**

## **1. Ziel des Programms**

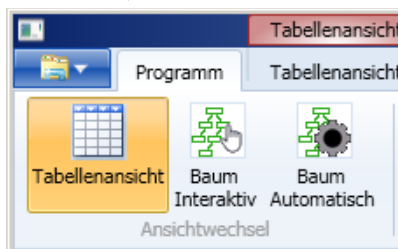
Dieses Programm dient der datenanalytischen Darstellung von Entscheidungsbäumen auf Grundlage von Informationen in Tabellenform. Durch die Erstellung von Regeln können Entscheidungsbäume Interaktiv oder Automatisch erstellt werden um so große Datenmengen zu analysieren.

## **2. Systemanforderungen**

Damit Sie das Programm "Entscheidungsbaum" nutzen können, wird ein Microsoft Windows System mit XP oder höher vorausgesetzt. Zudem muss das Microsoft .NET Framework 4.0 auf dem System installiert sein. Sollte dies nicht der Fall sein, können Sie dieses über die Microsoft-Website tun.

### 3. Ansichten

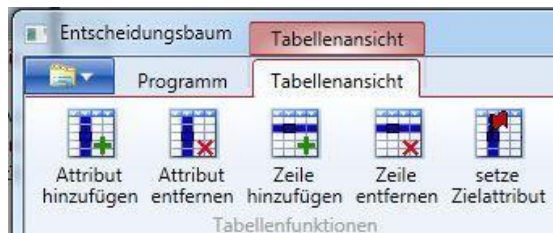
In dem Programm stehen Ihnen 3 Ansichten zur Verfügung: Tabellenansicht, Baum Interaktiv, Baum Automatisch.



Sie können unter dem "Programm"-Tab zwischen den verschiedenen Ansichten wechseln, indem Sie auf die entsprechende Schaltfläche klicken.

Der Wechsel auf die Ansichten "Baum Interaktiv" und "Baum Automatisch" ist erst möglich, sobald Sie in der Tabellenansicht ein Zielattribut festgelegt haben.

#### 3.1 Tabellenansicht



In der Tabellenansicht, können Sie die Daten auf denen gearbeitet werden soll, bearbeiten. Dazu können Sie auf den "Tabellenansicht"-Tab wechseln um die Funktionen der Tabellenansicht angezeigt zu bekommen. Auf die Funktionen dieser Ansicht wird im Folgenden genauer eingegangen.

##### 3.1.1 Attribut hinzufügen



Durch Klicken dieses Icons kann ein Attribut der Tabelle hinzugefügt werden. Über ein Dialogfeld wird man unmittelbar aufgefordert den Namen des Attributs anzugeben.

Attribute können via Drag&Drop an andere Positionen verschoben werden, um die Reihenfolge der Attribute zu ändern. Des Weiteren kann durch Anklicken des Attributs die Tabelle nach Werten steigend bzw. fallend sortiert werden.

**Hinweis:** Der Attributnamen sind hinterher nicht mehr änderbar! Nur durch löschen (inkl. aller Werte) und neu anlegen. Ein Maximum von 16 Attributen ist möglich.

### 3.1.2 Attribut entfernen



Um ein Attribut zu entfernen, klicken Sie in der Tabelle auf einen Eintrag in der Spalte mit dem Attribut welches Sie löschen möchten. Anschließend klicken Sie in dem “Tabellenansicht”-Tab auf die Schaltfläche “Attribut entfernen”.

### 3.1.3 Zeile hinzufügen



Um einen neuen Eintrag in die Tabelle einzufügen, klicken Sie im “Tabellenansicht”-Tab auf die Schaltfläche “Zeile hinzufügen”. Anschließend können Sie durch selektieren einer einzelnen Zelle den Wert ändern.

**Hinweis:** Es können maximal 1000 Einträge in die Tabelle aufgenommen werden.

### 3.1.4 Zeile entfernen



Um einen Eintrag aus der Tabelle zu entfernen, klicken Sie in der Tabelle auf den Eintrag den Sie entfernen möchten. Anschließend klicken Sie in dem “Tabellenansicht”-Tab auf die Schaltfläche “Zeile entfernen”.

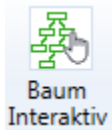
### 3.1.5 Zielattribut setzen



Wenn Sie ein Zielattribut setzen möchten, wählen Sie dieses aus und wechseln Sie durch klicken auf den “Tabellenansicht”-Tab zu den Tabellenwerkzeugen. Hier klicken Sie auf das Icon mit der Beschriftung “setze Zielattribut”. Das Attribut des ausgewählten Feldes ist nun als Zielattribut gesetzt und wird für die Baumdarstellungen verwendet.

**Hinweis:** Dieser Schritt ist immer nötig bevor Sie zur Baumansicht wechseln.

### 3.2 Baum interaktiv



Nachdem Daten in der Tabellenansicht eingelesen oder geschrieben wurden und ein Zielattribut festgelegt wurde, können Sie mit einem Klick auf die Schaltfläche "Baum Interaktiv" im "Programm"-Tab auf die interaktive Baum Ansicht wechseln.

Hier können Sie eigene Regeln definieren und so eigenständig einen Entscheidungsbaum aufbauen. Dies geschieht durch Doppelklick auf den ersten Knoten. Daraufhin erscheint ein Fenster mit den Einträgen die durch diesen Knoten repräsentiert werden. Innerhalb des Identifikationsfensters können Sie das Attribut festlegen welches als Regel für den zu identifizierenden Knoten durch auswählen der "Attribut auswählen" Schaltfläche bestimmen.

Nach dem Schließen des Fensters aktualisiert sich der interaktive Baum automatisch und zeigt nun die neuen Knoten an. Diesen Prozess können Sie mit allen Knoten wiederholen.

In der Grafik wird jeder Knoten wie folgt dargestellt.

keine Regel festgelegt	
Objekte:	11
Objekte j:	5
Objekte n:	6
Entropie:	0.994030211476956
Gew. Entropie:	0

Sie können das aktuell als Regel eingestellte Attribut des Knotens sehen, sowie die Anzahl an Objekten die durch diesen Knoten repräsentiert werden und wieviele dieser Knoten den Zielattributwert "j" oder "n" besitzen. Damit Sie den besten Entscheidungsbaum erstellen können, wird ihnen die Entropie des Knotens angezeigt, sowie die Gewichtete Entropie. Je niedriger der Wert ist, desto besser der Knoten.

### 3.3 Baum automatisch



Nachdem Daten in der Tabellenansicht eingelesen oder geschrieben wurden und ein Zielattribut festgelegt wurde, können Sie mit einem Klick auf die Schaltfläche “Baum Automatisch” im “Programm”-Tab auf die automatische Baum Ansicht wechseln. Auf dieser Ansicht wird automatisch der Optimale Entscheidungsbaum ermittelt. Dabei wird die gewichtete Entropie für jeden Knoten berechnet und jeweils das Attribut für einen Knoten ausgewählt welches die niedrigste gewichtete Entropie besitzt. Wie auch im Interaktiven Baum können Sie einen Knoten durch doppeltes Klicken auf einen Knoten identifizieren. Dazu wird ein Identifikationsfenster geöffnet welches alle Einträge auflistet, die durch diesen Knoten repräsentiert werden.

Um Einfluss auf die Generierung zu nehmen, können Sie den Parameter “Mindestanzahl Objekte” im “Baumautomatisch”-Tab ändern, um festzulegen ab welcher Anzahl an Objekten je Knoten nicht weiter aufgeteilt werden soll. Damit die Änderung übernommen wird, müssen Sie den Baum neu berechnen indem Sie auf die Schaltfläche “Baum neu berechnen” klicken.

In der Grafik wird jeder Knoten wie folgt dargestellt.

keine Regel festgelegt	
Objekte:	11
Objekte j:	5
Objekte n:	6
Entropie:	0.994030211476956
Gew. Entropie:	0

Sie können das aktuell als Regel eingestellte Attribut des Knotens sehen, sowie die Anzahl an Objekten, die durch diesen Knoten repräsentiert werden und wieviele dieser Knoten den Zielattributwert “j” oder “n” besitzen. Damit Sie den besten Entscheidungsbaum erstellen können, wird ihnen die Entropie des Knotens angezeigt, sowie die Gewichtete Entropie. Je niedriger der Wert ist, desto besser der Knoten.

## 4. Funktionen

Folgend werden die einzelnen Funktionalitäten des Programms beschrieben inkl. wichtiger Spezifikationen.

- **Datei öffnen**



Über das Optionsmenü kann die Funktion zum Öffnen von Dateien gefunden werden. Die Schaltfläche "Datei öffnen" anklicken und im darauf folgenden File Selector Fenster die zu öffnende Datei auswählen.

**Hinweis:** Hierbei ist zu beachten, dass das Format der Datei .csv sein muss. Als Trennzeichen ist ausschließlich das Semikolon ( ; ) akzeptiert. Des weiteren führt es zu einer Fehlermeldung wenn nicht mit der ersten Zeile/Spalte die Tabelle beginnt. In die erste Reihe kommt immer die Reihenbezeichnung.

- **Datei speichern**



Über das Optionsmenü kann die Funktion zum Datei speichern aufgerufen werden. Die Schaltfläche "Datei speichern" klicken und im darauf folgenden File Selector Fenster den Zielort der zu speichernden Datei auswählen.

Als Format muss zwingend .csv als Endung angegeben werden, da es sonst nicht wieder eingelesen werden kann. Als Trennzeichen ist ausschließlich ein Semikolon ( ; ) erlaubt.