



Technische
Universität
Braunschweig



Chair for
Chip Design for
Embedded Computing



Seminar Technische Informatik

Top 10 algorithms in data mining

Stephan Mielke, 22.01.2015

Inhalt

Data Mining

Klassifikation

Clustering

Assoziation

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Unterkapitel 1

Ende

Einleitung - Der Weltraum enendliche Weiten



Abbildung 1: Hubble Ultra Deep Field[1]

Einleitung - Einsatz von DM in der Astronomie

- Teleskope erfassen pro Bild ca 10.000 Objekte
- Manuelle Klassifizierung unmöglich [2]
- Benutzung von Klassifizierungsalgorithmen aus DM
- Je Objekt 9 Attribute (8 Isophotenformen, Leuchtkraft)
- Ausgabewert „stellar“
 - 0.0 – 0.1 Galaxie
 - 0.9 – 1.0 Stern

Einleitung - Einsatz von DM in der Astronomie

- Teleskope erfassen pro Bild ca 10.000 Objekte
- Manuelle Klassifizierung unmöglich [2]
- Benutzung von Klassifizierungsalgorithmen aus DM

Name	Erkennung
Random Forest	82, 89%
Decision Tree	80, 68%
Artificial Neural Network	75.82%
Support Vector Machines	37, 82%

Tabelle 1: Erkennungsraten der Algorithmen Stern / Galaxie[3]

Überblick

Data Mining

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Ende

Data Mining - Einleitung

Data Mining - Klassifikation

Data Mining

Klassifikation

Clustering

Assoziation

Data Mining - Clustering

Data Mining

Klassifikation

Clustering

Assoziation

Data Mining - Assoziation

Data Mining

Klassifikation

Clustering

Assoziation

Überblick

Data Mining

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Ende

Überblick

Data Mining

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Ende



Hier steht der Titel der Folie

Wir beginnen mit einer Aufzählung

- Aufzählzeichen werden als Quadrate dargestellt
 - Unterpunkte ebenfalls
 - Allerdings etwas kleiner

Überblick

Data Mining

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Ende

Itemize-Test

- Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam
- At vero eos et accusam et justo duo dolores et ea rebum.
 - Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet!
 - Nam eget dui.
 - Maecenas tempus, tellus eget condimentum rhoncus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum.
 - Duis leo
- Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus.

Kapitel 2 - Unterkapitel 1

Kapitel 2

Unterkapitel 1

Mathe-Test

Gaußsche Summenformel:

$$1 + 2 + 3 + 4 + \dots + n = \sum_{k=1}^n k = \frac{n(n+1)}{2}$$

Faltung:

$$(f * g)(\xi) := \int_{\mathbb{R}^n} f(y)g(\xi - y)dy$$

Überblick

Data Mining

Top 10 algorithms in data mining

Neue Algorithmen

Kapitel 2

Ende

Farbtest

Dies ist ein Text in tuRed.

Dies ist ein Text in tuSecondaryDark80.

Dies ist ein Text in tuSecondaryLight.

Verwendung von Spalten

Dies ist die erste Spalte. Die Angabe der Option `[onlytextwidth]` sorgt dafür, dass die Spaltenbreite korrekt eingehalten wird.

Dies ist die zweite Spalte mit weiteren Informationen.

Blöcke

Diest ist ein Block

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam

Diest ist ein Example-Block

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam

Diest ist ein Alert-Block

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam

Quellcode

Quellcode-Frames müssen immer die Option [fragile] tragen.

```
#include <iostream>

using namespace std;





// main function
int main() {
    cout << "Hello_World!";
}
```

Wichtig

Diese Folie ist wichtig!

Literatur

bla[4]

-  S. B. S. NASA, ESA and the HUDF Team. (2004) Hubble ultra deep field. [Online]. Available:
<http://imgsrc.hubblesite.org/hu/db/images/hs-2004-07-a-pdf.pdf>
-  M. Ester and J. Sander, *Knowledge discovery in databases: Techniken und Anwendungen*. Springer Heidelberg, 2000, vol. 2, no. 4.
-  P. J. O’Keefe, M. G. Gowanlock, S. M. McConnell, and D. R. Patton, “Star-galaxy classification using data mining techniques with considerations for unbalanced datasets,” in *Astronomical Data Analysis Software and Systems XVIII*, vol. 411, 2009, p. 318.
-  X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, S. Y. Philip *et al.*, “Top 10 algorithms in data mining,” *Knowledge and Information Systems*, no. 1, pp. 1–37, 2008.