

# Weitere Anwendungen von WEKA / Machine Learning

Stefan Warwas

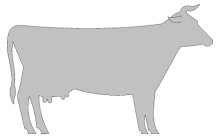
Universität des Saarlandes

30.11.2006

# Übersicht

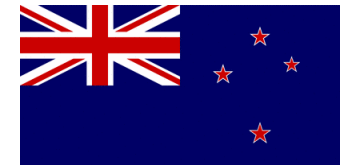
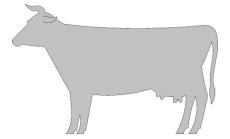
---

- Landwirtschaft
  - Einsatz von Machine Learning für die Auswahl von Milchkühen
  - allgemeines Prozessmodell
- Music Information Retrieval
  - automatische Klassifizierung von Musik
  - Vergleich verschiedener Algorithmen und Feature Sets

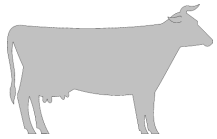


# Landwirtschaft Übersicht

---

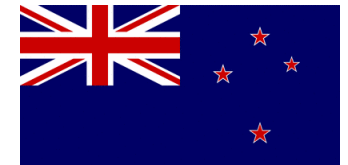
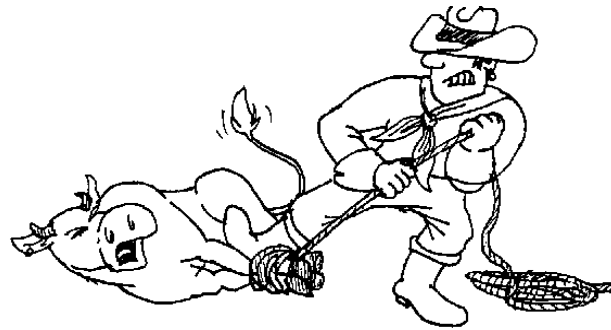
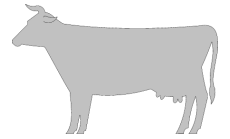


[1] Stephen R. Garner et al. Applying a Machine Learning Workbench: Experience with Agricultural Databases. 1995

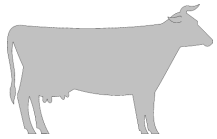


# Landwirtschaft Übersicht

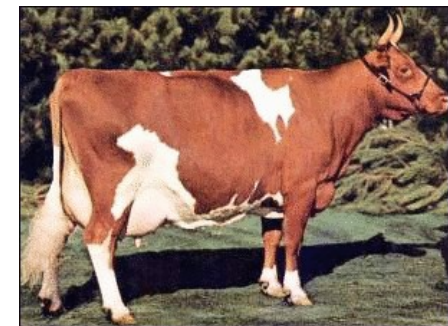
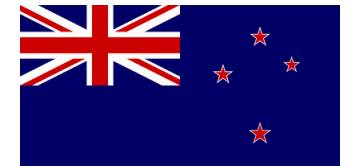
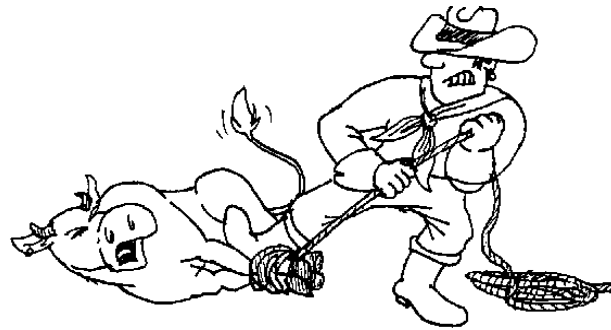
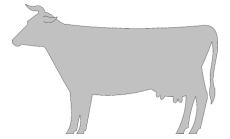
---

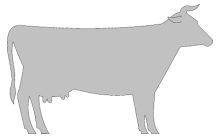


pro Jahr werden ca. 20% aller Kühe gekeult

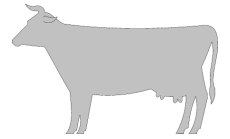


# Landwirtschaft Übersicht

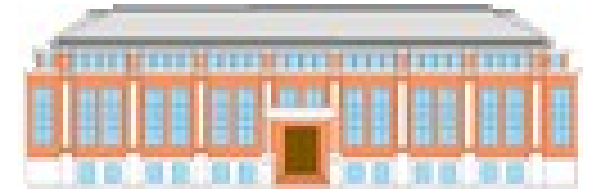
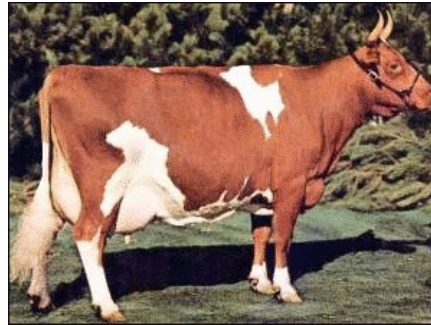




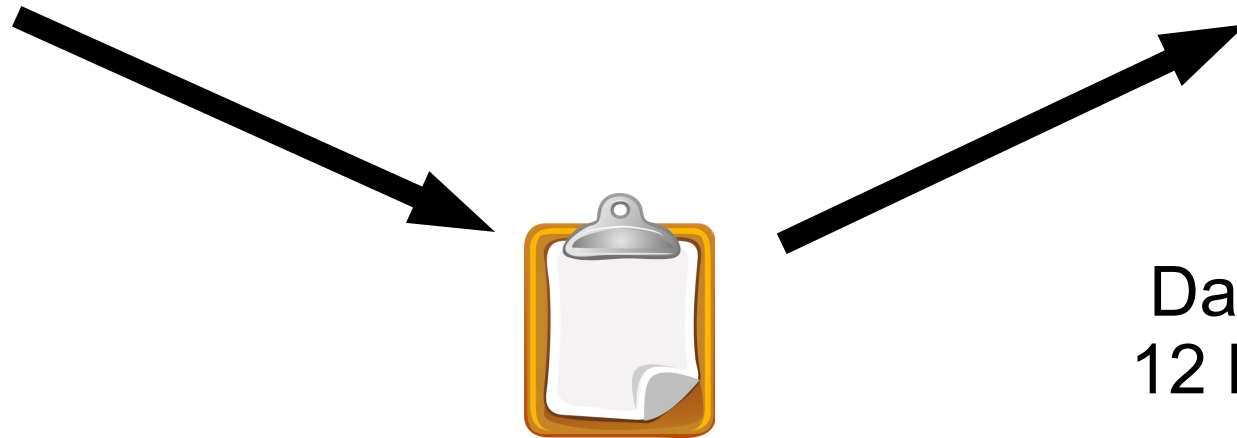
# Landwirtschaft Übersicht



Bauer

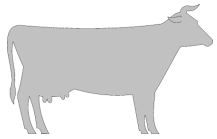


Livestock  
Corporation

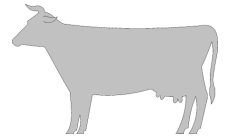


Daten über  
12 Mio. Kühe

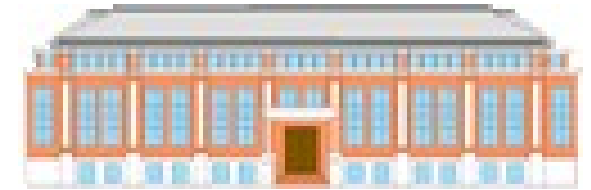
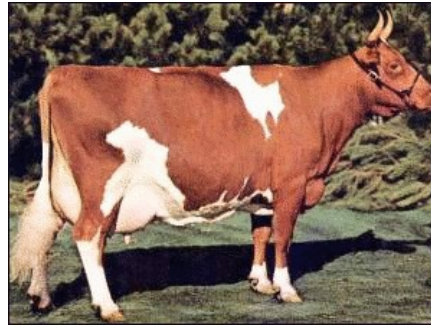
4-12 mal im Jahr



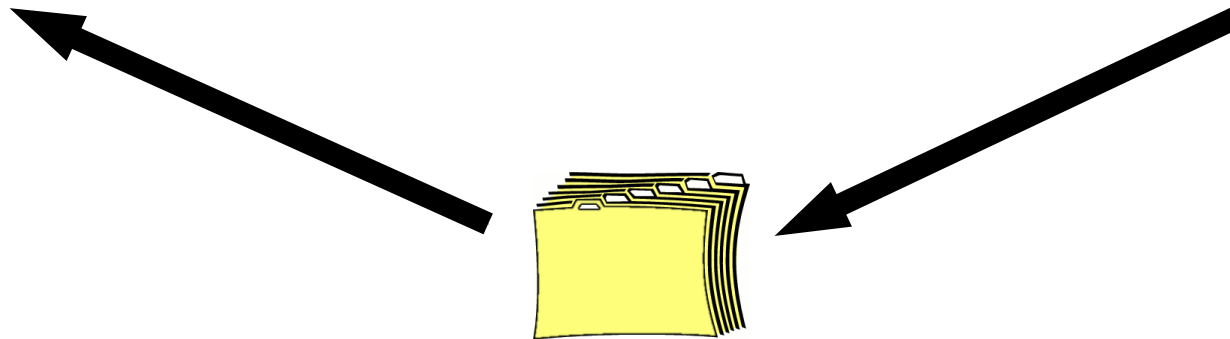
# Landwirtschaft Übersicht



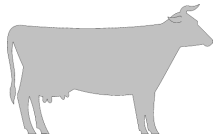
Bauer



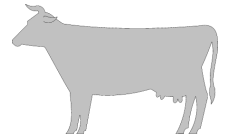
Livestock  
Corporation



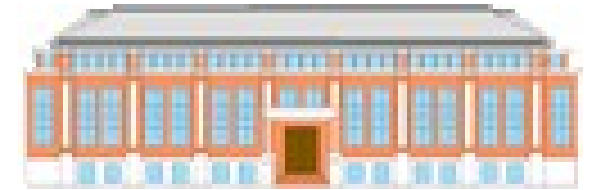
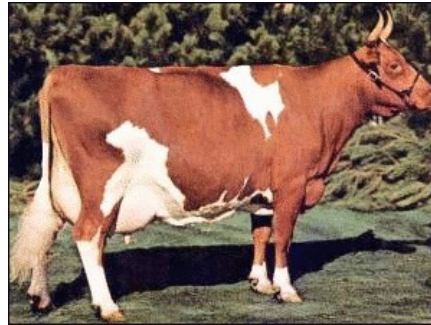
Auswertung der Daten



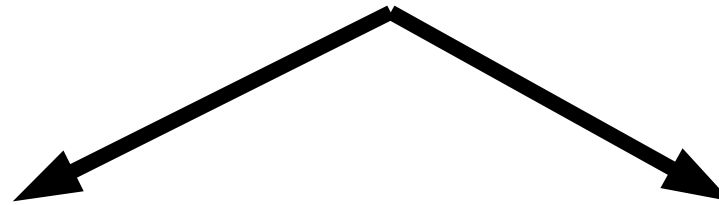
# Landwirtschaft Übersicht



Bauer



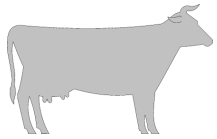
Livestock  
Corporation



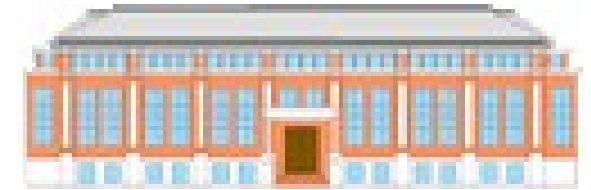
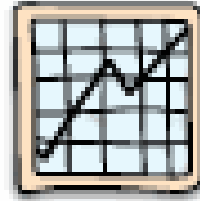
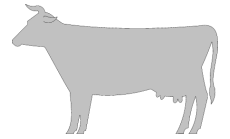
production index



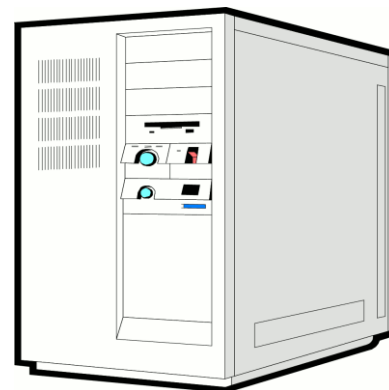
breeding index

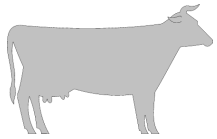


# Landwirtschaft Übersicht

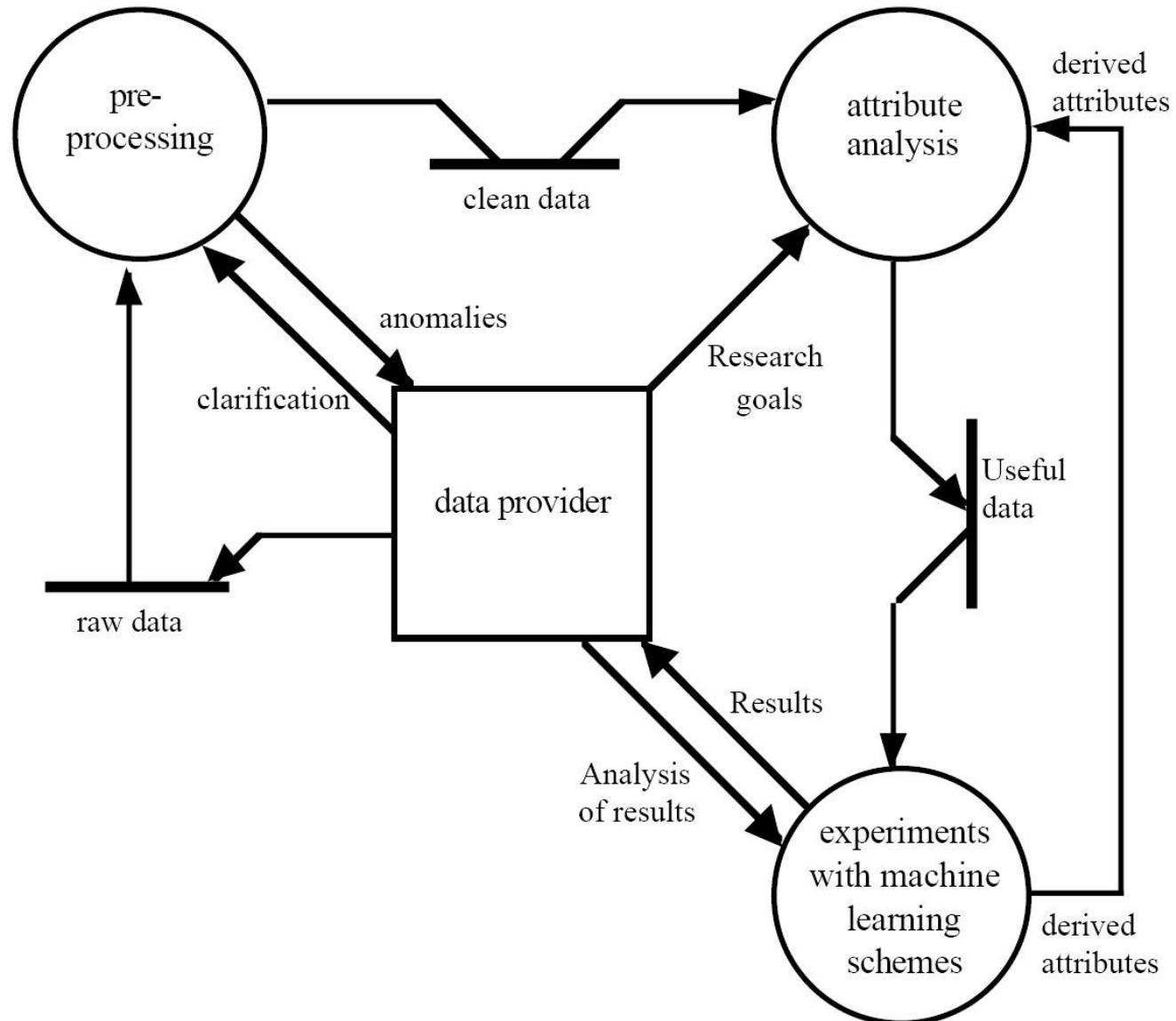
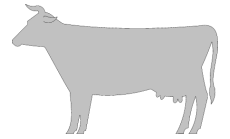


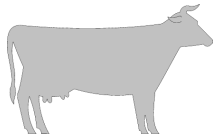
Livestock  
Corporation



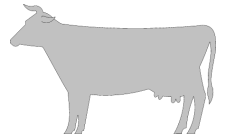


# Prozessmodell

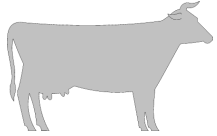




# Landwirtschaft Probleme

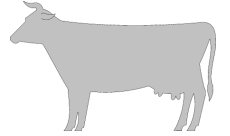


- 
- Konvertierung ins ARFF-Format
  - Rauschen in den Daten
  - Probleme beim Festlegen der Datentypen
  - Umgang mit fehlende Daten
  - “small disjunct”-Problem

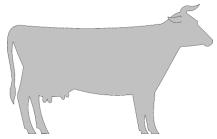


# Landwirtschaft Durchführung

---



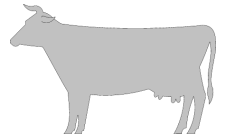
- Datenbasis:
  - Daten von 10 Herden über 6 Jahre
  - 19000 Datensätze
  - 705 Attribute
- Feature Selection: C4.5 und FOIL
- Classification-Algorithmus: C4.5



# Landwirtschaft

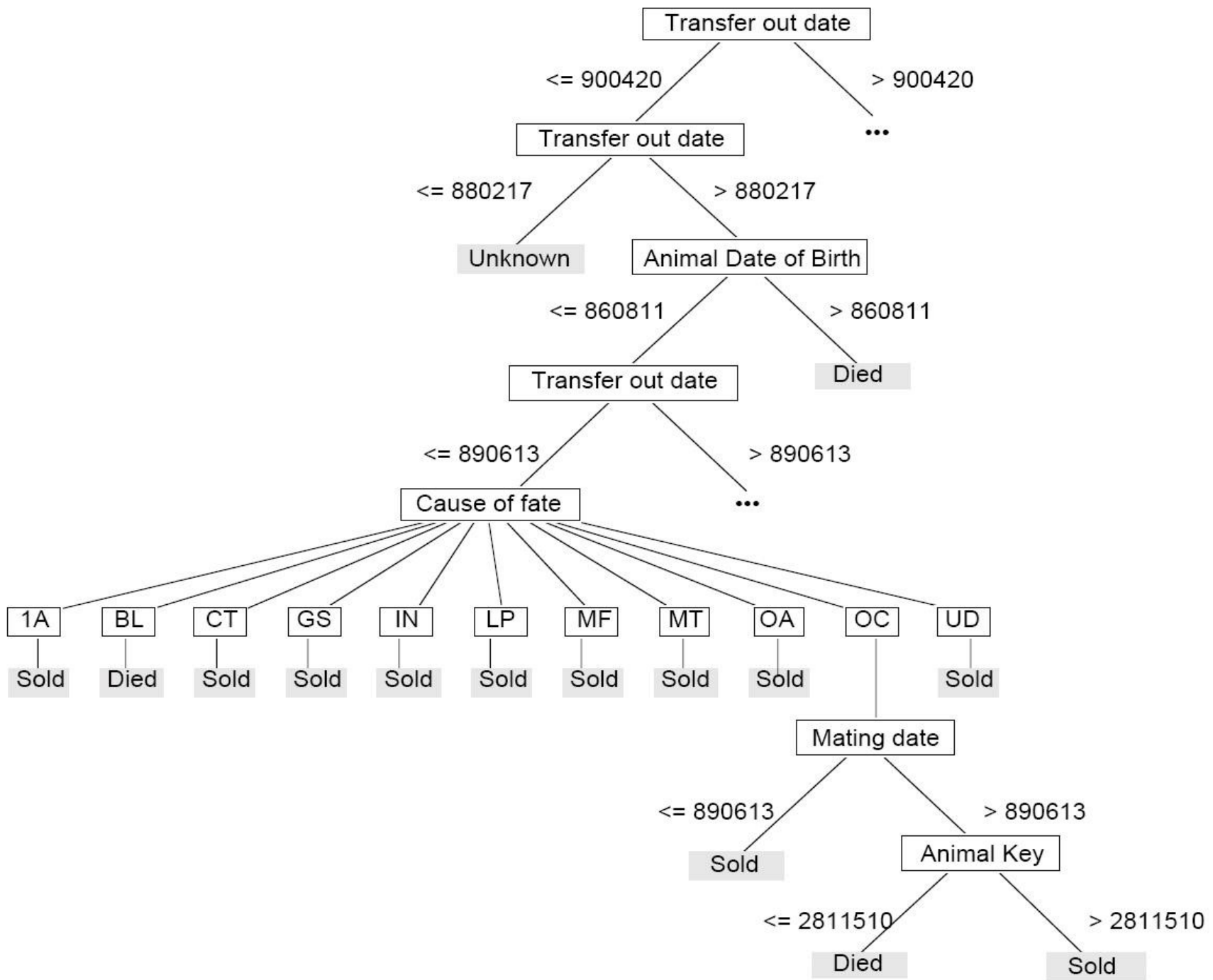
## Wichtige Attribute

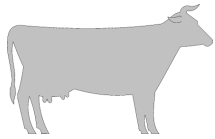
---



- breeding index (BI): Wert als Zuchttier
- production index (PI): Wert als Milchkuh
- transfer out date: Datum des Keulens
- fate code: sold, dead, lost, unknown
- cause of fate: Grund für den Tod

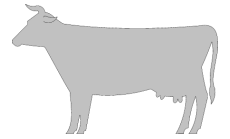
► Klassifikation nach “fate code”



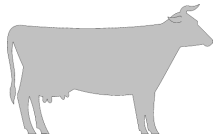


# Landwirtschaft Auswertung

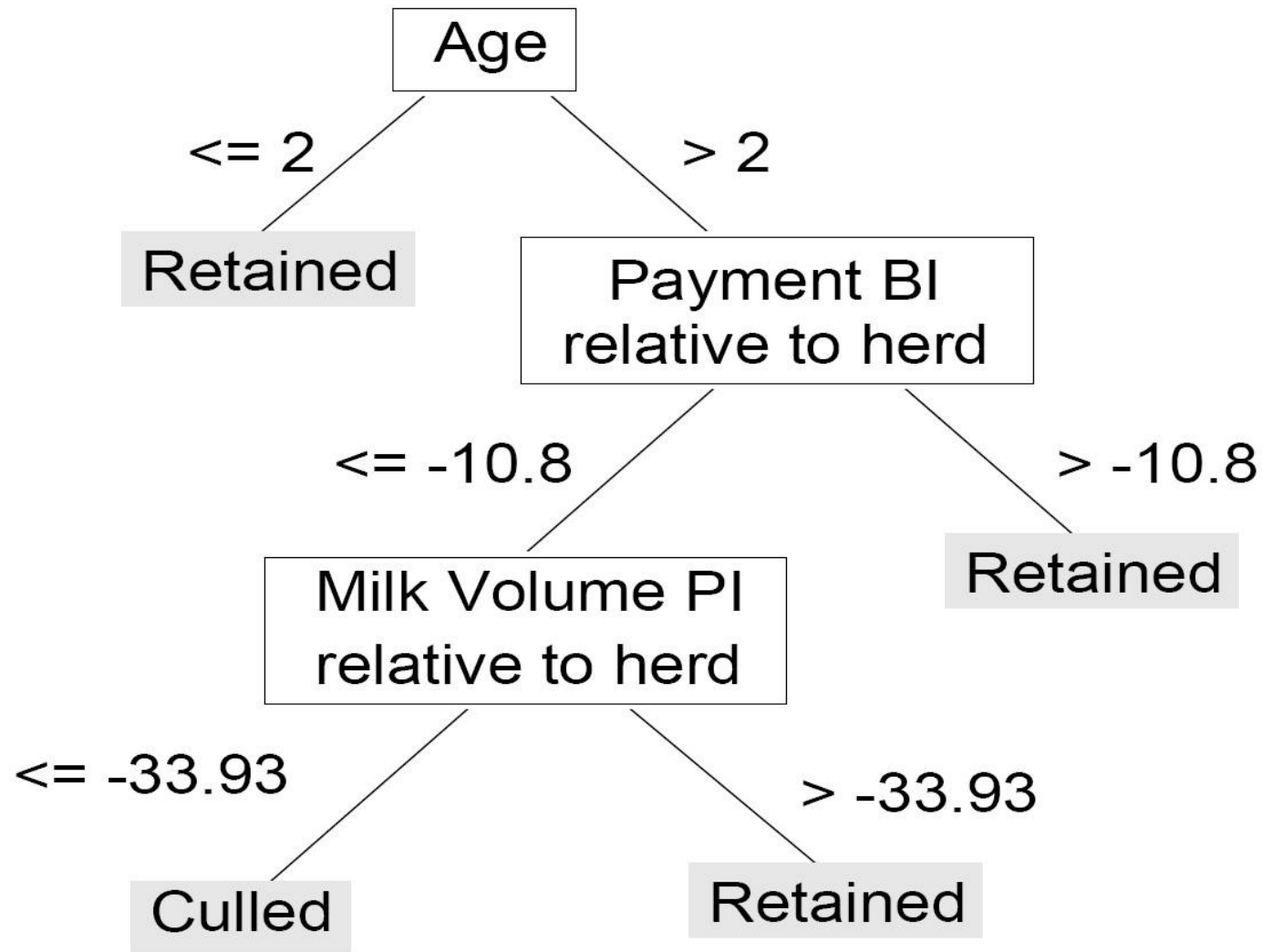
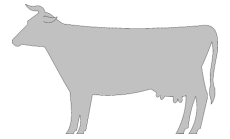
---

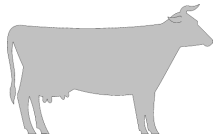


- Plausibilität der Regeln wurde geprüft
- Bedeutung von Attributen konkretisiert
- Hinzufügen neuer abgeleiteter Attribute:
  - verschiedene Indizes: Leistung relativ zur Herde in einem bestimmten Jahr
  - **status code**: erzeugt aus “**fate code**”, “**cause of fate**” und “**transfer out date**”
  - absolute Daten in relative Daten umwandeln
- Datensätze von Kühen mit unbeabsichtigtem Ableben entfernt

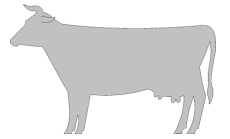


# Landwirtschaft





# Landwirtschaft Fazit



- 
- in der Praxis reicht es nicht aus, einfach einen Algorithmus auf Daten anzuwenden
  - Machine Learning ist ein iterativer Prozess
  - Zusammenarbeit von Domain- und ML-Experten ist sehr wichtig
  - Erfahrungen sind in die Entwicklung von WEKA eingeflossen



# Music Information Retrieval Definition



- 
- Gewinnung von Informationen aus Musik
    - basierend auf den Audiodaten
    - basierend auf externen Daten
    - mit Hilfe von Klassifikations-, Clustering-, ...  
Algorithmen
  - Anwendungen
    - Klassifizierung von Musik
    - Playlist-Generatoren
    - Empfehlung von Musik



# Music Information Retrieval Beispiel



- [www.sloud.com](http://www.sloud.com)
- Musikerkennung durch Summen

[2] Valeriy Lobaryev et al. Sloud Query-by-Humming Search Music Engine.



# Musik Information Retrieval Übersicht



- 
- Ziel der Untersuchung:
    - Einteilung der Musik in 6 verschiedene Wahrnehmungskategorien
    - Erprobung von 3 verschiedenen Feature Sets
    - und 12 verschiedenen ML-Algorithmen
  - Durchführung:
    - alle Kombinationen wurden auf die Vorhersagekraft für die einzelnen Kategorien getestet

[3] Time Pohle et al. Evaluation of Frequently Used Audio Features for Classification of Music Into Perceptual Categories



# Musik Information Retrieval Kategorien

---



## 6 Kategorien:

- Mood (happy, neutral, sad)
- Perceived tempo  
(very slow, slow, medium, fast, very fast, varying)
- Complexity (low, medium, high)
- Emotion (soft, neutral, aggressive)
- Focus (vocals, both, instruments)
- Genre (blues, classical, electronical, folk, jazz, new age, noise, rock, world)



# Musik Information Retrieval Datenbasis



- 
- 834 Lieder im MP3 Format
  - alle Lieder wurden von einem Mann klassifiziert
  - Lieder wurden ins WAV-Format konvertiert (11025 kHz mono)
  - Audioausschnitte aus der Mitte jedes Liedes (30 Sekunden)
  - jeder Ausschnitt wurde in 256 Teile geteilt

Categorization	Classes (# of songs in class)
mood	happy (29%), neutral (50%), sad (21%)
perceived tempo	very slow (4%), slow (20%), medium (43%), fast (24%), very fast (5%), varying (4%)
complexity	low (18%), medium (56%), high (7%)
emotion	soft (29%), neutral (44%), aggressive (26%)
focus	vocals (6%), both (69%), instruments (26%)
genre	blues (1%), classical (5%), electronica (13%), folk (2%), jazz (1%), new age (5%), noise (0.1%), rock (60%), world (10%)



# Music Information Retrieval Feature Set I



- 
- besteht aus 30 Attributen pro Lied
  - der Schwerpunkt liegt auf der Struktur und dem Rythmus der Musik
  - Feature Set hat drei Unterbereiche :

## 1) Klangfarbe (19 Werte)

- Short Time Fourier Transformation
- Eigenschaften des Spektrums (z.B. spectral flux)



# Music Information Retrieval Feature Set I



- besteht aus Grund- und Obertönen
- der Schwerepunkt liegt auf dem Rythmus
- Feature Set

**Ton:** besteht aus Grund- und Obertönen

**Klangfarbe:** zeitlicher Verlauf von Grund- und Oberfrequenzen

## 1) Klangfarbe (19 Werte)

- Short Time Fourier Transformation
- Eigenschaften des Spektrums (z.B. spectral flux)



# Music Information Retrieval Feature Set I



- besteht aus 30 Attributen pro Lied
  - der Schwerpunkt liegt auf der Struktur und dem Rythmus der Musik
  - Menge aller Frequenzen, aus denen ein Ton besteht
- 1) Klangfarbe (19 W)
- Short Time Fourier Transformation
  - Eigenschaften des Spektrums (z.B. spectral flux)



# Musik Information Retrieval

## Feature Set I

---



### 2) Beat Histogram (6 Features)

- Wavelet Transformation
- X-Achse: Periodizität in bpm
- Y-Achse: Stärke/Häufigkeit

### 3) Pitch Histogram (5 Features)

- X-Achse: Pitch
- Y-Achse: Häufigkeit

=> verschiedene Kennwerte (z.B. max. Amplitude, ...)



# Musik Information Retrieval

## Feature Set I



### 2) Beat Histogram (6 Features)

- wahrgenommene Grundfrequenz eines Tones
- X-Achse: Periodizität in bpm
- Y-Achse: Stärke/Häufigkeit

### 3) Pitch Histogram (5 Features)

- X-Achse: Pitch
- Y-Achse: Häufigkeit

=> verschiedene Kennwerte (z.B. max. Amplitude, ...)



# Music Information Retrieval Feature Set II



- 
- basierend auf MPEG 7
  - Low Level Descriptors (LLDs)
  - Informationen zur Bandbreite
  - Audio-Spektrum (ähnlich zu Feature Set I)
    - z.B. audio spectrum flatness
  - Harmonie (Zusammenklang)
    - Grundfrequenz (ähnlich “main pitch”)
    - harmonische Obergrenze (Frequenz)



# Music Information Retrieval Feature Set II



- basierend auf MPEG 7
- Low Level Descriptors (LLDs)
- Informationen zur Bandbreite
- Audio-Spektrum (ähnlich zu Feature Set I)
  - z.B. audio spectrum flatness
- Harmonic Range (Differenz zwischen höchster und niedrigster Frequenz des Audiosignals)
  - "C" (C1) ("C" = "C1")
  - harmonische Obergrenze (Frequenz)



# Music Information Retrieval Feature Set II



- 
- basierend auf MPEG 7
  - Low Level Descriptors (LLDs)
  - Informationen zur Bandbreite
  - Audio-Spektrum (ähnlich zu Feature Set I)
    - z.B. audio spectrum flatness
  - Harmonie (Zusammenklang)
    - Grundfrequenz (ähnlich “main pitch”)
    - harmonische Obergrenze (Frequenz)



# Music Information Retrieval Feature Set II



- basierend auf MPEG 7
- Low Level Descriptors (LLDs)
- Informationen zur Bandbreite
- Audio-Spektrum (ähnlich zu Feature Set I)
  - z.B. audio spectrum flatness
- Harmonie (Zusammenklang)
  - Grundfrequenz (ä “main pitch”)
  - wie stark weicht das Signal von einer glatten Kuve ab (frequenz)



# Music Information Retrieval Feature Set II



- 
- basierend auf MPEG 7
  - Low Level Descriptors (LLDs)
  - Informationen zur Bandbreite
  - Audio-Spektrum (ähnlich zu Feature Set I)
    - z.B. audio spectrum flatness
  - Harmonie (Zusammenklang)
    - Grundfrequenz (ähnlich “main pitch”)
    - harmonische Obergrenze (Frequenz)



# Music Information Retrieval Feature Set III



- 
- Summe der ersten beiden Feature Sets
  - Zusätzliche Features:
    - spectral power
    - wie “percussive” ein Stück ist
    - weitere statistische Werte zu den Daten aus den beiden ersten Feature Sets (z.B. Varianz)
  - insgesamt 146 Features



# Music Information Retrieval Feature Set III



- Summe der ersten beiden Feature Sets
- Zusätzliche Features:
  - spectral power
  - wie “percussive” ein Stück ist
  - weitere statistische Werte zu den Daten aus den beiden Feature Sets (z.B. Varianz)
- insgesamt **Amplituden der einzelnen Teilfrequenzen**



# Music Information Retrieval Feature Set III



- 
- Summe der ersten beiden Feature Sets
  - Zusätzliche Features:
    - spectral power
    - wie “percussive” ein Stück ist
    - weitere statistische Werte zu den Daten aus den beiden ersten Feature Sets (z.B. Varianz)
  - insgesamt 146 Features



# Music Information Retrieval Algorithmen I



- 
- K-Nearest Neighbors
  - Naïve Bayes
  - C4.5
  - Support Vector Machine
  - AdaBoost mit C4.5 und Decision Stumps
  - Klassifikation über Regression  
(M5 und lineare Regression)
- => Evaluation über 10-fold Cross Validation



# Music Information Retrieval Algorithmen II

---



- zusätzlich eine weitere Kombination aus Feature Set und Algorithmus
- Sony Computer Science Laboratory Paris
- Feature Set
  - Schwerpunkt auf Erkennung der Klangfarbe
  - Erkennung von Ähnlichkeiten zwischen Liedern
- Algorithmus: KNN



# Ergebnisse

## Perceived Tempo I

---



- Klassifikationsgenauigkeit:

Baseline	: 42,53%
Feature Set I	: 42,53%
Feature Set II	: 43,13%
Feature Set III	: 44,67%
Sony	: 48,67%
- Baseline wurde in den meisten Experimenten nicht erreicht
- Naïve Bayes liefert die besten Ergebnisse



# Ergebnisse

## Perceived Tempo I



- Klassifikationsgenauigkeit:

Baseline : 42,53%

Feature Set I : 42,53%

Feature Set II : 43,13%

Feature Set III : 43,13%

Sony

Klassifikationsgenauigkeit,  
wenn immer in die größte  
Klasse klassifiziert wird

- Baseline wurde in den meisten Experimenten nicht erreicht
- Naïve Bayes liefert die besten Ergebnisse



# Ergebnisse

## Perceived Tempo I

---



- Klassifikationsgenauigkeit:

Baseline	: 42,53%
Feature Set I	: 42,53%
Feature Set II	: 43,13%
Feature Set III	: 44,67%
Sony	: 48,67%
- Baseline wurde in den meisten Experimenten nicht erreicht
- Naïve Bayes liefert die besten Ergebnisse



# Ergebnisse

## Perceived Tempo II



Tempo, NB, All Descriptors

vsl	48.5	27.3	6.1	6.1	9.1	3.0
sl	15.0	29.9	15.0	6.0	6.6	27.5
med	14.7	18.4	14.4	8.8	8.2	35.4
fst	11.1	13.1	14.6	12.1	14.6	34.3
vfs	9.1	11.4	13.6	11.4	31.8	22.7
var	17.1	8.6	14.3		14.3	45.7
	vsl	sl	med	fst	vfs	var

Tempo, NB, Set from [TC02]

vsl	42.4	30.3	6.1	12.1	3.0	6.1
sl	16.2	28.7	26.9	8.4	9.0	10.8
med	9.9	25.8	18.7	18.4	13.9	13.3
fst	8.1	19.7	21.7	22.7	21.7	6.1
vfs	4.5	6.8	25.0	13.6	40.9	9.1
var	5.7	22.9	25.7	28.6	2.9	14.3
	vsl	sl	med	fst	vfs	var



# Ergebnisse Emotion



- 
- Klassifikationsgenauigkeit:

Baseline	: 44,46%
Feature Set I	: 45,06%
Feature Set II	: 46,75%
Feature Set III	: 47,47%
Sony	: 57,95%
  - Vermutung: die Klangfarbe ist für die Klassifizierung von “emotion” wichtig



# Music Information Retrieval

## Zusammenfassung I

---



- Ergebnisse kaum besser als die Baseline
- die Kombination Sony war in allen Tests am besten
- Genre-Klassifikation funktioniert am besten
- **complexity** konnte nicht vorhergesagt werden
- für **focus**, **perceived tempo** und **mood** waren die Vorhersagen meist unter der Baseline
- **emotion** hängt wahrscheinlich von der Klangfarbe ab



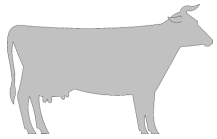
# Music Information Retrieval

## Zusammenfassung II

---



- Mögliche Ursachen:
  - inkonsistente Annotation der Trainingsdaten
  - die meisten Features beziehen sich auf einen sehr kurzen Zeitraum
- Verbesserungsmöglichkeiten:
  - Aufbau der Musik ist wichtig (Events)
  - möglicherweise existiert eine “unsichtbare” Obergrenze für nur auf Audiodaten gestützte Klassifikation
  - weitere Informationsquellen: Texte, kulturelle



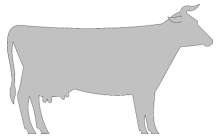
# WEKA Anwendungsbeispiele

## Referenzen I

---



- [1] Stephen R. Garner et al. Applying a Machine Learning Workbench: Experience with Agricultural Databases. 1995
- [2] Valeriy Lobaryev et al. Sloud Query-by-Humming Search Music Engine
- [3] Tim Pohle et al. Evaluation of Frequently Used Audio Features for Classification of Music Into Perceptual Categories. 2005



# WEKA Anwendungsbeispiele

## Referenzen II

---



- [4] Georg Tzanetakis et al. Automatic Musical Genre Classification Of Audio Signals. 2001
- [5] Georg Tzanetakis. Tempo Extraction using Beat Histograms. 2005
- [6] Jean-Julien Aucouturier et al. Music Similarity Measures: What's the Use?. Proceedings of the 3<sup>rd</sup> International Symposium on Music Information Retrieval, 2002