

# Dataminig - Versuch 2

Lukas Spieß, Mathis Hoffmann

17. November 2013

## 2 Durchführung Teil 1: Fiktive Filmbewertung

### 2.2 Ähnlichkeitsbestimmung

1. *Welche Bedeutung hat der Übergabeparameter `normed` in der Funktion `sim_euclid`?*

Wird bei der Überhabe "`normed=True`" übergeben, erfolgt eine Normalisierung des Rückgabewertes der Funktion. Dies geschieht, um verlässlichere Werte für relative Ähnlichkeit zu erhalten, da die Euklidische Distanz hier normalerweise nur begrenzt geeignet ist. Dabei wird das Ergebnis zusätzlich durch die Anzahl der verglichenen Elemente dividiert.

2. *Vergleichen Sie die beiden Ähnlichkeitsmaße. Welches Ähnlichkeitsmaß erscheint Ihnen für diesen Anwendungsfall sinnvoller und warum?*

Für diese Aufgabe erscheint die Verwendung der Pearson Korrelation sinnvoller, da hier eine variierende Anzahl an Bewertungen vorliegt, für die die Euklidische Distanz nicht besonders gut geeignet ist.