



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Ausarbeitung Projekt 2

Bastian Probst

**Durchführung eines zweiten Votings und Erstellung eines
Vorhersagesystems**

*Fakultät Technik und Informatik
Studiendepartment Informatik*

*Faculty of Engineering and Computer Science
Department of Computer Science*

Bastian Probst

**Durchführung eines zweiten Votings und Erstellung eines
Vorhersagesystems**

Ausarbeitung Projekt 2

im Studiengang Master Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuer: Prof. Dr. Kai von Luck

Eingereicht am: 20. Oktober 2015

Inhaltsverzeichnis

1	Einleitung	1
2	Zweites Voting	1
2.1	Ziel	2
2.2	Durchführung	3
2.3	Aktueller Stand	4
3	Vorhersagesystem	4
3.1	Komponenten eines Vorhersagesystems	4
3.2	Auswahl der Daten	5
3.3	Vorverarbeitung	5
3.4	Transformation	6
3.5	Data Mining	6
3.6	Interpretation	7
4	Vorgehen	7
5	Fazit	8
5.1	Ergebnisse des ersten Experiments	8
5.2	Das zweite Experiment	9
5.3	Auswertung	10
5.4	Ausblick	11

1 Einleitung

Durch das beständige Wachstum an digitalen Medieninhalten ist es erforderlich, die Inhalte zu filtern. Eine Möglichkeit der Filterung ist die Personalisierung. [Fan u. a. \(2012\)](#) [Bellekens u. a. \(2009\)](#) [Weiß u. a. \(2008\)](#). Mittlerweile scheint der Trend von der Empfehlung hin zur Entscheidungsübernahme zu gehen.

In diesem Paper wird der zweite Teil eines Experiments beschrieben, welches im Rahmen meiner Masterarbeit durchgeführt wird. Im ersten Teil bewerteten die Teilnehmer 50 Filme und machten einige Angaben zur Person, wie Alter, Geschlecht und Konsumverhalten bei Filmen. Im zweiten Teil des Experiments wurden die Teilnehmer gebeten, 10 weitere Filme zu bewerten. Parallel wurde ein Vorhersagesystem entwickelt, welches die Bewertungen für diese Filme vorhersagt.

Um Benutzermodelle zu erstellen, existieren zwei unterschiedliche Methoden: die Selbstauskunft und die Analyse des Verhaltens der Benutzer [Weiß u. a. \(2008\)](#). Im Kontext dieser Arbeit wird die Methode der Selbstauskunft verwendet. Die Benutzer bewerten Filme und Kategorien explizit.

Die unterschiedlichen Arten von Benutzermodellen und deren Verwendung für Personalisierung wurden im Rahmen des ersten Vortrags des Masterseminars herausgearbeitet [Probst \(2014b\)](#). Im zweiten Vortrag wurde dargestellt, wie solche Benutzermodelle erstellt und für Vorschläge verwendet werden können [Probst \(2014a\)](#). Die theoretischen Grundlagen des Experiments wurden im Hauptseminar vorgestellt [Probst \(2015b\)](#). Der erste Teil des Projekts wurde im ersten Projektbericht dargestellt. [Probst \(2015a\)](#)

Im Kontext dieser Arbeit werden die im ersten Experiment gesammelten Daten verwendet, um die Vorhersagen für die Filme des zweiten Experiments zu generieren. Hierbei werden zwei unterschiedliche Methoden verwendet. Die eine Methode verwendet die Kategoriebewertungen, die andere die Bewertungen der Filme. Nach einer ersten Analyse der Ergebnisse stellte sich heraus, dass eine Reduzierung der Kategorien und die Verwendung des Herkunftslandes zu einer deutlichen Verbesserung führen.

Im [1.](#) Kapitel werden die Motivation und der Inhalt für diese Arbeit beschrieben. Im [2.](#) Kapitel werden die Ziele und die Durchführung des zweiten Experiments dargestellt. Eine Erklärung von Vorhersagesystemen im Kontext dieser Arbeit folgt in Kapitel [3](#). Die Methode der Vorhersage wird in Kapitel [4](#) erklärt. Im [5.](#) Kapitel werden die Ergebnisse der Experimente betrachtet und ein Ausblick auf mögliche Aufgaben im Rahmen der Masterarbeit gegeben.

2 Zweites Voting

In diesem Kapitel wird das zweite Voting beschrieben. Zunächst wird das Ziel genauer betrachtet. Daraufhin werden die Durchführung und der aktuelle Stand dargestellt. Eine Analyse von Vorhersagesystemen im Kontext dieser Arbeit folgt in Kapitel 3. Die Methode zum Berechnen der Bewertungen wird in Kapitel 4 genauer beschrieben.

2.1 Ziel

Um weitere Bewertungen zu erfassen, wurde das zweite Voting durchgeführt. Dies hatte unter anderem das Ziel, Bewertungen mit einem zeitlichen Abstand zu erhalten. Auf diese Weise ist es möglich die Bewertungen eines Filmes beim ersten Voting mit der Bewertung des Filmes beim zweiten Voting zu vergleichen. Hierbei kann eine Tendenz des statistischen Rauschens annähernd ermittelt werden.

Ein weiterer Aspekt ist die Berechnung der Bewertungen von neuen Benutzern oder Filmen. Dieses Problem wird als Kaltstartproblem bezeichnet [Albanese u. a. \(2013\)](#) [Bu u. a. \(2010\)](#) [Lin u. a. \(2013\)](#). Hier müssen Bewertungen für Filme berechnet werden, welche noch nicht von Benutzern bewertet wurden. Oder es müssen Bewertungen für Benutzer berechnet werden, welche noch keine Filme bewertet haben.

Im Kontext dieser Arbeit existieren im Testdatensatz nur Bewertungen von bekannten Benutzern. Die Filme hingegen haben noch keine Bewertungen. Ein möglicher Ansatz hierfür ist der Vergleich mit ähnlichen Filmen.

Die meisten Algorithmen benötigen für ihre Vorhersagen einen Trainingsdatensatz [Said und Bellogín \(2014\)](#) [Bu u. a. \(2010\)](#) [Lakiotaki u. a. \(2008\)](#). Selbst bei dem Kaltstartproblem werden Trainingsdaten benötigt. Das Verhalten der unbekanntem Benutzer wird anhand von ähnlichen Benutzern berechnet. Bei den Filmen werden für die Berechnung Bewertungen ähnlicher Filme herangezogen. Eine weitere Option ist unter anderem die Matrixfaktorisierung.

Um die berechneten Bewertungen mit tatsächlichen vergleichen zu können, sind ein Trainings- und ein Testdatensatz erforderlich. Der Trainingsdatensatz wird von den Algorithmen zum Lernen verwendet. Der Testdatensatz wird für den Vergleich zwischen Berechnung und Realität benötigt [Said u. a. \(2014\)](#). Um die unterschiedlichen Algorithmen zu vergleichen, können Metriken wie der Abstand zwischen tatsächlicher und berechneter Bewertung verwendet werden. Weitere Vergleichsmöglichkeiten bieten die Betrachtung der „false positive“ und der „false negative“, „precision“, „recall“ oder „nDGC“ [Said und Bellogín \(2014\)](#). Eine ausführliche

Analyse der Metriken und deren Anwendung wird in der Masterarbeit durchgeführt werden. Somit haben die beiden Votings das Ziel, einen Trainings- und einen Testdatensatz zu erstellen. Die Daten des ersten Votings sind als Trainingsdatensatz zu betrachten. Beim zweiten Voting wurden die Testdaten erstellt. Die Durchführung des Votings wird im nächsten Kapitel beschrieben.

2.2 Durchführung

Das erste Voting ist in **Probst (2015a)** detailliert vorgestellt worden. Die Durchführung des zweiten Votings wird in diesem Kapitel genauer beschrieben und im Anschluss der aktuelle Stand dargestellt.

Bei dem zweiten Voting werden zehn weitere Filmbewertungen erfasst. Dieses sind jeweils zwei Filme aus den im ersten Voting verwendeten Kategorien (Aktion, Liebe, Sport, Science-Fiction und Biografien). Die Teilnehmer bei dem ersten Voting, welche alle Filme bewertet hatten, wurden um die Teilnahme am zweiten Voting gebeten. Dies sind ca. 100 Personen. Von diesen haben 42 das zweite Voting beendet. Am Login wurde nichts verändert, sodass die Teilnehmer die Logindaten des ersten Votings verwenden konnten. Für den Fall eines vergessenen Passwortes gab es wie im ersten Voting die Möglichkeit einer Passwortwiederherstellung. Das Ende des Votings wurde

Mad Max: Fury Road (2015)



Wie findest du den Film: "Mad Max: Fury Road"?

- -3 (sehr schlecht)
- -2
- -1
- 0 (neutral)
- 1
- 2
- 3 (sehr gut)

VOTE

Abbildung 2.1: Ansicht der Filmbewertung für einen Benutzer

auf den 18. Juli 2015 gesetzt. Im zweiten Voting wurden neben den Filmbewertungen keine Fragen zur Person gestellt.

2.3 Aktueller Stand

Das zweite Voting ist beendet. 42 der 100 Teilnehmer haben alle der 10 Filme bewertet. Ein erstes Vorhersagesystem ist entwickelt und zwei unterschiedliche Algorithmen sind dabei getestet worden. Eine Reduzierung der Kategorien hatte hier zu verbesserten Ergebnissen geführt. Die einzelnen Schritte zum Erstellen einer Vorhersage werden im nächsten Kapitel näher beschrieben. Die Algorithmen für die Vorhersage werden in Kapitel 4 genauer betrachtet.

3 Vorhersagesystem

In diesem Kapitel werden die grundsätzlichen Elemente von Knowledge Discovery in Databases beschrieben. Dieses wird unter dem Gesichtspunkt eines Vorhersagesystems betrachtet. Das im Kontext dieses Projektes entwickelte Vorhersagesystem wird genauer beschrieben. Die konkreten Methoden der Vorhersage und deren Evaluierung folgen im nächsten Kapitel. Eine erste Analyse der Ergebnisse und ein Ausblick für weitere Versuche und Möglichkeiten werden im nächsten und letzten Kapitel dargestellt.

3.1 Komponenten eines Vorhersagesystems

Ein Vorhersagesystem besteht im Wesentlichen aus folgenden Komponenten [Fayyad u. a. \(1996\)](#):

- einer **Auswahl** der Daten, die für den Prozess verwendet werden sollen,
- einer **Vorverarbeitung** der Daten,
- einer **Transformation** der Daten in eine gewünschte Struktur,
- einem **Datamining** Prozess und
- einer **Interpretation** der Ergebnisse.

Die Einzelnen Schritte des KDD-Prozesses sind in [Abbildung 3.1](#) dargestellt.

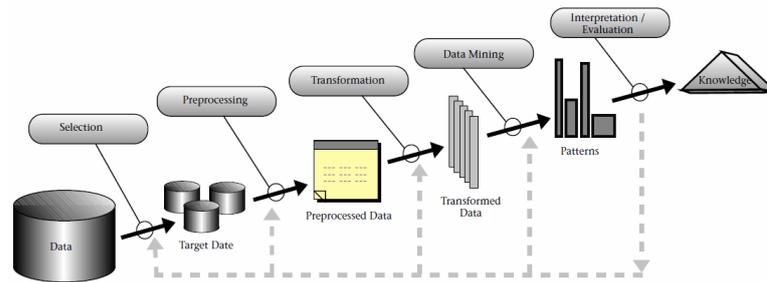


Abbildung 3.1: Darstellung des KDD Prozesses [Fayyad u. a. \(1996\)](#)

3.2 Auswahl der Daten

Bei der Auswahl der Daten werden die für die Vorhersagen relevanten Daten extrahiert. Dies sind im Kontext dieser Arbeit die Benutzer, welche beide Votings erfolgreich durchgeführt haben. Die Benutzer, die nicht am zweiten Voting teilgenommen haben oder nur einen Teil der Filme bewerteten, werden nicht weiter berücksichtigt. Ebenfalls ist es sinnvoll, die Benutzer mit einem auffälligen Verhalten zu entfernen. Einerseits werden Teilnehmer entfernt, welche sich nicht genügend Zeit fürs Voting genommen haben. Andererseits werden Benutzer entfernt, die immer das Gleiche voten oder bei denen andere Auffälligkeiten sind, wie zum Beispiel eine alternierende Folge $(-3,3,-3,3,-3,3,\dots)$. Die ausgewählten Daten werden vorverarbeitet. Dieser Schritt wird im nächsten Kapitel beschrieben.

3.3 Vorverarbeitung

Bei der Vorverarbeitung werden die ausgewählten Daten so verarbeitet, dass diese für die Vorhersagen verwendet werden können. Dies beinhaltet zum einen die Erstellung der Benutzermodelle und zum anderen das Zusammenfassen der Filmdaten.

In den Benutzerprofilen werden die Bewertungen der Kategorien, die Filmbewertungen und die Eigenschaften der Benutzer gespeichert. Zu den Eigenschaften zählen das Alter, das Geschlecht und das Konsumverhalten. Zu den Filmdaten gehören die Kategorien, die Bewertungen und weitere Eigenschaften des Filmes. Bei dem ersten Versuch wurden die Kategorien der IMDB verwendet. Beim zweiten wurde eine Hauptkategorie und das Herkunftsland für die Vorhersagen verwendet. Die verarbeiteten Daten werden für das Datamining transformiert.

3.4 Transformation

Bei der Transformation werden die Filme zu Clustern zusammengefasst. Im ersten Versuch werden die Cluster anhand der Kategorien gebildet. Die Cluster unterscheiden sich durch die Anzahl der zugehörigen Kategorien. Bei dem zweiten Versuch werden die Cluster anhand des Herkunftslandes und der Hauptkategorie gebildet. Die gebildeten Cluster werden für das Datamining verwendet. Ähnlich der Methode von [Gantner u. a. \(2010\)](#).

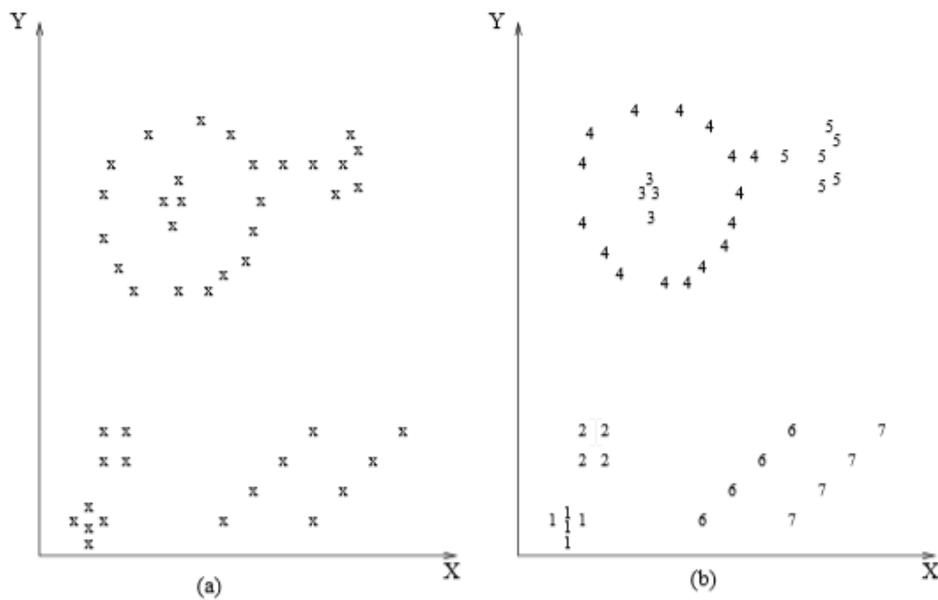


Abbildung 3.2: Darstellung des Data Clusterings [Jain u. a. \(1999\)](#)

3.5 Data Mining

Um Vorhersagen zu generieren werden Methoden des Dataminings angewendet. Im Kontext dieser Arbeit werden zwei unterschiedliche Methoden angewendet und miteinander verglichen. Bei der ersten Methode werden die Bewertungen mithilfe der Kategorienbewertungen berechnet. Für die zweite Methode werden die Bewertung der Filme verwendet.

Für die Vorhersagen mithilfe der Kategorienbewertungen wird die durchschnittliche Bewertung der beteiligten Kategorien verwendet. Daraus ergibt sich folgende Formel:

$$\frac{\sum_{i=1}^n \text{Kategoriebewertung}_n}{i}$$

Bei den Vorhersagen unter Verwendung der Filmbewertungen werden die gebildeten Clusterbewertungen betrachtet. Die Clusterbewertungen berechnen sich anhand der abgegebenen Filmbewertungen. Inwieweit eine Clusterbewertung für eine Filmbewertung herangezogen wird, hängt von der Übereinstimmung der Kategorien ab. Je mehr Kategorien übereinstimmen, desto stärker wird die Clusterbewertung für die Filmbewertung mit einbezogen. Die Methoden der Berechnung werden in Kapitel 4 genauer dargestellt.

3.6 Interpretation

Bei der Interpretation werden die Ergebnisse analysiert. Es wird betrachtet, inwieweit die Methoden des Datamining die gewünschten Ergebnisse erzielen. Im Kontext dieser Arbeit wird der Abstand zwischen tatsächlichen und berechneten Bewertungen betrachtet („mean absolute error“ [Biancalana u. a. \(2011\)](#) [Baltrunas \(2008\)](#) [Lakiotaki u. a. \(2008\)](#))

Die aus der Analyse gewonnenen Erkenntnisse werden verwendet, um Verbesserungen für die vorherigen Schritte zu erarbeiten. Im Kontext dieser Arbeit stellte sich heraus, dass eine Reduzierung der Kategorien und das Heranziehen des Herkunftslandes eine Verbesserung der Ergebnisse bewirkte. Die in dieser Arbeit verwendeten Algorithmen für die Vorhersage werden im nächsten Kapitel genauer beschrieben. In Kapitel 5 werden die Ergebnisse analysiert und Fragestellungen für die Masterarbeit herausgearbeitet.

4 Vorgehen

Im Rahmen dieser Arbeit wurde folgendes Experiment durchgeführt. Es wurde die Bewertung für die Filme anhand der Genrebewertungen berechnet (wie in Kapitel 3.5 beschrieben).

Ebenfalls wurden die Bewertungen der Filme anhand der vorherigen Filmbewertungen berechnet. Hierfür wurden Cluster erstellt. Ein Film wird einem Cluster zugeordnet, wenn die Kategorien des Filmes mit denen des Clusters übereinstimmen. Die Bewertung eines neuen Films wird anhand der Cluster mit ähnlichen Kategorien berechnet. Hierfür wird betrachtet, inwieweit ein Film in einen Cluster gehört. Die Zugehörigkeit eines Filmes zu einem Cluster berechnet sich wie folgt:

$$\text{Abhängigkeit} = \frac{\text{Kategorien}_{\text{Film}} \cap \text{Kategorien}_{\text{Cluster}}}{\text{Kategorien}_{\text{Cluster}}} \quad (4.1)$$

Dies bedeutet, dass die Abhängigkeit in Relation dazu steht, wie viele der Kategorien eines Filmes mit den Kategorien des Clusters übereinstimmen. Die Anzahl der übereinstimmenden

Kategorien werden durch die Anzahl der Kategorien des Clusters geteilt. Dies bedeutet zum Beispiel für einen Film mit den Kategorien: „Sport, Aktion, Liebe“ und ein Cluster mit den Kategorien: „Sport, Aktion, Biografie“:

$$\begin{aligned}
 \text{Abhängigkeit} &= \frac{\text{Kategorien}_{\text{Film}} \cap \text{Kategorien}_{\text{Cluster}}}{\text{Kategorien}_{\text{Cluster}}} \\
 &= \frac{\text{Sport, Aktion, Liebe} \cap \text{Sport, Aktion, Biografie}}{\text{Sport, Aktion, Biografie}} \\
 &= \frac{\text{Sport, Aktion}}{\text{Sport, Aktion, Biografie}} = \frac{2}{3}
 \end{aligned} \tag{4.2}$$

Die Cluster werden unterschiedlich gewichtet. Ein Cluster mit einer höheren Anzahl an Genres hat eine höhere Gewichtung. Die Gewichtung berechnet sich somit wie folgt:

$$\text{Gewichtung}_{\text{Cluster}} = \text{Anzahl}_{\text{Genres}} \tag{4.3}$$

Die Gewichtung eines Clusters entspricht der Anzahl der zugehörigen Kategorien. Unter Einbeziehung der Clusterzugehörigkeit und der Gewichtungen ergibt sich folgende Gesamtformel für die berechnete Bewertung eines Filmes:

$$\text{Filmbewertung}_{\text{Film}_x} = \frac{\sum_{i=1}^n \text{Clusterbewertung}_i * \text{Abhängigkeit}_i^2 * \text{Gewichtung}_i}{\sum_{i=1}^n \text{Abhängigkeit}_i^2 * \text{Gewichtung}_i}$$

Mit n = Anzahl der Clusterzugehörigkeiten für Film X, i = Clusternummer

Die berechnete Bewertung eines Filmes ist somit die Summe der Clusterbewertungen multipliziert mit der jeweiligen Abhängigkeit und der quadrierten Zugehörigkeit geteilt durch die Summe der Abhängigkeiten multipliziert mit den quadrierten Zugehörigkeiten.

Die Ergebnisse des Experiments und die daraus resultierenden Fragestellungen für die Masterarbeit werden im nächsten Kapitel beschrieben.

5 Fazit

5.1 Ergebnisse des ersten Experiments

Bei der Verwendung der IMDB Kategorien stellte sich heraus,

- dass die durchschnittliche Filmbewertung 0.8226 beträgt,

- dass die durchschnittliche berechnete Filmbewertung anhand der Kategorien 1.0288 ist - mit einer Abweichung von 1.3994,
- dass die durchschnittliche berechnete Filmbewertung anhand der bisherigen Filmbewertungen einen Wert von 0.1675 hat - mit einer Abweichung von 1.5813

Bei genauerer Analyse der Ergebnisse stellte sich heraus, dass die größte Abweichung bei Filmen aus dem indischen Raum auftraten. Diese Filme wurden von den Teilnehmer signifikant schlechter bewertet, als die berechneten Werte. Die Bewertungen anhand der Filmbewertungen haben wenig starke Ausprägungen und nähern sich im Mittel der 0 an. Dies scheint seinen Ursprung in der hohen Anzahl an Clustern zu haben, welche für die Berechnung eines Filmes herangezogen werden.

Weitere Auffälligkeiten sind:

- Der gleiche Film wurde in unterschiedlichen Kategorien unterschiedlich bewertet. Dies deutet daraufhin, dass Filme unter dem Aspekt der Kategorie bewertet wurden. Somit wird der selbe Film von einigen Benutzern in unterschiedlichen Kategorien auch unterschiedlich bewertet. Die Filme des zweiten Votings waren keinen Kategorien zugeordnet. Welchen Effekt dies auf die Bewertungen hat, ist nicht eindeutig zu identifizieren.
- Der selbe Film wurde bei unterschiedlichen Votings unterschiedlich bewertet. Eine Möglichkeit für die veränderte Bewertung ist, dass der Teilnehmer den Film jetzt gesehen hat und ihn besser bewerten kann. Eine weitere ist, dass der Teilnehmer das zweite Voting in einer anderen Stimmung gemacht hat und sich dieses auf die Bewertungen ausgewirkt hat. Letztlich ist auch eine Änderung der Lebensumstände eine Möglichkeit für eine andere Bewertung. Grundsätzlich ist es nicht erforderlich die näheren Umstände der Abweichungen zu erfassen. Wichtig ist, dass hier festgehalten wird, dass es ein Grundrauschen von 0.22 gibt.
- Filme aus dem indischen Raum wurden schlecht bewertet.
- Filme mit einem geringen Budget wurden schlechter bewertet als Filme mit einem großen Budget.

5.2 Das zweite Experiment

Diese ersten Erkenntnisse führten zu einer Modifizierung des Experiments. Die Anzahl der Kategorien wurde auf 1 reduziert. Diese Kategorie wurde anhand der größten Signifikanz

aus den IMDB Kategorien des Filmes ausgewählt. Weiterhin wurde das Herkunftsland als eine weitere Kategorie hinzugefügt. Da von den Benutzern keine expliziten Bewertungen der Herkunftsländer existieren, wurde diese bestimmt. Filme aus Amerika wurden mit +3 und Filme aus Indien mit -3 bewertet.

Durch das Reduzieren der Kategorien und das Hinzufügen von den Herkunftsländern verbessert sich die Genauigkeit.

- Die durchschnittliche berechnete Filmbewertung anhand der Kategorien beträgt 1.5063 mit Abweichung von 1.2989. Dies entspricht einer Verbesserung von 0.1.
- Die durchschnittliche berechnete Filmbewertung anhand der vorherigen Filmbewertungen ist 0.8923 mit einer Abweichung von 1.1148. Dies ist eine Verbesserung von 0.47.

5.3 Auswertung

Durch das Gegenüberstellen der Ergebnisse des ersten und des zweiten Votings können folgende Thesen aufgestellt werden:

- Eine Reduzierung der Kategorien und das Hinzufügen des Herkunftslandes verbessert die Ergebnisse anhand der Filmbewertungen erheblich und die der Kategoriebewertungen etwas.
- Das Herkunftsland spielt eine größere Rolle als die Kategorien.

Einige der Vorhandenen Informationen wurden noch nicht für die Vorhersagen verwendet. Zu diesen gehören:

- das **Geschlecht** des Benutzers,
- das **Alter** des Benutzer
- und das **Konsumverhalten** des Benutzers.
- die **Schauspieler**,
- die **Drehorte**,
- das **Budget**,
- der **Regisseur** und
- die **Handlung** der Filme.

5.4 Ausblick

Somit gilt es in der Masterarbeit folgende Fragen zu klären:

- Ist es möglich, Gewichtungen für die Kategorien einzuführen?
- Wie müssen die Kategorien gewichtet werden?
- Können Attribute verwendet werden wie Liebe, Spannung, Aktion, ähnlich der Verwendung in Fernsehzeitschriften (Sternchen)?
- Welchen Effekt hat die Verwendung von weiteren Attributen wie zum Beispiel Budget, Schauspieler, etc.?
- Können andere Genres von anderen Datenbanken verwendet werden?
- Kann eine Distanz zwischen Filmen definiert werden?
- Kann diese Distanz für die Bewertung anhand der dichtesten Filme verwendet werden?
- Kann eine Distanz zwischen Benutzern gebildet werden?
- Kann diese Distanz für die Bewertung anhand der dichtesten Benutzer verwendet werden?
- Kann eine Matrixfaktorisierung angewendet werden?
- Können die Benutzerattribute wie Alter, Geschlecht und Konsumverhalten für die Vorhersage verwendet werden?

Im Rahmen der Masterarbeit wird versucht, diese Fragen zu klären. Es wird versucht werden, die Algorithmen und Schritte so zu verbessern, um genauere Ergebnisse zu erzielen. Genaue Ergebnisse sind gerade im Hinblick auf den Trend der Entscheidungsübernahme erforderlich. Ein mögliches Anwendungsszenario ist eine personalisierte Set-Top-Box. Diese kann für den Benutzer Filme aufnehmen, die ihn interessieren.

Literaturverzeichnis

- [Albanese u. a. 2013] ALBANESE, Massimiliano ; D'ACIERNO, Antonio ; MOSCATO, Vincenzo ; PERSIA, Fabio ; PICARIELLO, Antonio: A Multimedia Recommender System. In: *ACM Trans. Internet Technol.* 13 (2013), November, Nr. 1, S. 3:1–3:32. – URL <http://doi.acm.org/10.1145/2532640>. – ISSN 1533-5399
- [Baltrunas 2008] BALTRUNAS, Linas: Exploiting Contextual Information in Recommender Systems. In: *Proceedings of the 2008 ACM Conference on Recommender Systems*. New York, NY, USA : ACM, 2008 (RecSys '08), S. 295–298. – URL <http://doi.acm.org/10.1145/1454008.1454056>. – ISBN 978-1-60558-093-7
- [Bellekens u. a. 2009] BELLEKENS, Pieter ; HOUBEN, Geert-Jan ; AROYO, Lora ; SCHAAP, Krijn ; KAPTEIN, Annelies: User Model Elicitation and Enrichment for Context-sensitive Personalization in a Multiplatform Tv Environment. In: *Proceedings of the Seventh European Conference on European Interactive Television Conference*. New York, NY, USA : ACM, 2009 (EuroITV '09), S. 119–128. – URL <http://doi.acm.org/10.1145/1542084.1542106>. – ISBN 978-1-60558-340-2
- [Biancalana u. a. 2011] BIANCALANA, Claudio ; GASPARETTI, Fabio ; MICARELLI, Alessandro ; MIOLA, Alfonso ; SANSONETTI, Giuseppe: Context-aware Movie Recommendation Based on Signal Processing and Machine Learning. In: *Proceedings of the 2Nd Challenge on Context-Aware Movie Recommendation*. New York, NY, USA : ACM, 2011 (CAMRa '11), S. 5–10. – URL <http://doi.acm.org/10.1145/2096112.2096114>. – ISBN 978-1-4503-0825-0
- [Bu u. a. 2010] BU, Jiajun ; TAN, Shulong ; CHEN, Chun ; WANG, Can ; WU, Hao ; ZHANG, Lijun ; HE, Xiaofei: Music Recommendation by Unified Hypergraph: Combining Social Media Information and Music Content. In: *Proceedings of the International Conference on Multimedia*. New York, NY, USA : ACM, 2010 (MM '10), S. 391–400. – URL <http://doi.acm.org/10.1145/1873951.1874005>. – ISBN 978-1-60558-933-6
- [Fan u. a. 2012] FAN, Xiangyu ; MOSTAFA, Javed ; MANE, Ketan ; SUGIMOTO, Cassidy: Personalization is Not a Panacea: Balancing Serendipity and Personalization in Medical News Content Delivery. In: *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium*. New York, NY, USA : ACM, 2012 (IHI '12), S. 709–714. –

- URL <http://doi.acm.org/10.1145/2110363.2110445>. – ISBN 978-1-4503-0781-9
- [Fayyad u. a. 1996] FAYYAD, Usama ; PIATETSKY-SHAPIRO, Gregory ; SMYTH, Padhraic: From Data Mining to Knowledge Discovery in Databases. In: *AI Magazine* 17 (1996), S. 37–54
- [Gantner u. a. 2010] GANTNER, Zeno ; DRUMOND, Lucas ; FREUDENTHALER, Christoph ; RENDLE, Steffen ; SCHMIDT-THIEME, Lars: Learning Attribute-to-Feature Mappings for Cold-Start Recommendations. In: *Proceedings of the 2010 IEEE International Conference on Data Mining*. Washington, DC, USA : IEEE Computer Society, 2010 (ICDM '10), S. 176–185. – URL <http://dx.doi.org/10.1109/ICDM.2010.129>. – ISBN 978-0-7695-4256-0
- [Jain u. a. 1999] JAIN, A. K. ; MURTY, M. N. ; FLYNN, P. J.: Data Clustering: A Review. In: *ACM Comput. Surv.* 31 (1999), September, Nr. 3, S. 264–323. – URL <http://doi.acm.org/10.1145/331499.331504>. – ISSN 0360-0300
- [Lakiotaki u. a. 2008] LAKIOTAKI, Kleanthi ; TSAFARAKIS, Stelios ; MATSATSINIS, Nikolaos: UTA-Rec: A Recommender System Based on Multiple Criteria Analysis. In: *Proceedings of the 2008 ACM Conference on Recommender Systems*. New York, NY, USA : ACM, 2008 (RecSys '08), S. 219–226. – URL <http://doi.acm.org/10.1145/1454008.1454043>. – ISBN 978-1-60558-093-7
- [Lin u. a. 2013] LIN, Jovian ; SUGIYAMA, Kazunari ; KAN, Min-Yen ; CHUA, Tat-Seng: Addressing Cold-start in App Recommendation: Latent User Models Constructed from Twitter Followers. In: *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York, NY, USA : ACM, 2013 (SIGIR '13), S. 283–292. – URL <http://doi.acm.org/10.1145/2484028.2484035>. – ISBN 978-1-4503-2034-4
- [Probst 2014a] PROBST, Bastian: Personalisierte Vorschläge für Medieninhalte. (2014)
- [Probst 2014b] PROBST, Bastian: Personalisierung von interaktiven Second Screen Anwendungen mithilfe von Benutzermodellen. (2014)
- [Probst 2015a] PROBST, Bastian: Durchführung eines Votings zur Erstellung von Benutzerprofilen. (2015)
- [Probst 2015b] PROBST, Bastian: Personalisierte Vorschläge für Medieninhalte. (2015)

- [Said und Bellogín 2014] SAID, Alan ; BELLOGÍN, Alejandro: Comparative Recommender System Evaluation: Benchmarking Recommendation Frameworks. In: *Proceedings of the 8th ACM Conference on Recommender Systems*. New York, NY, USA : ACM, 2014 (RecSys '14), S. 129–136. – URL <http://doi.acm.org/10.1145/2645710.2645746>. – ISBN 978-1-4503-2668-1
- [Said u. a. 2014] SAID, Alan ; DOOMS, Simon ; LONI, Babak ; TIKK, Domonkos: Recommender Systems Challenge 2014. In: *Proceedings of the 8th ACM Conference on Recommender Systems*. New York, NY, USA : ACM, 2014 (RecSys '14), S. 387–388. – URL <http://doi.acm.org/10.1145/2645710.2645779>. – ISBN 978-1-4503-2668-1
- [Weiß u. a. 2008] WEISS, Diana ; SCHEUERER, Johannes ; WENLEDER, Michael ; ERK, Alexander ; GÜLBAHAR, Mark ; LINNHOF-POPIEN, Claudia: A User Profile-based Personalization System for Digital Multimedia Content. In: *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*. New York, NY, USA : ACM, 2008 (DIMEA '08), S. 281–288. – URL <http://doi.acm.org/10.1145/1413634.1413687>. – ISBN 978-1-60558-248-1