

Entwicklung eines Lernsystems für Gruppenpräferenzen

Diplomarbeit im Fach Informatik

vorgelegt von

Sebastian von Mammen

geb. 08.02.1980 in Schwabach

angefertigt am

**Institut für Informatik
Lehrstuhl für Informatik 8
Künstliche Intelligenz
Friedrich-Alexander-Universität Erlangen–Nürnberg
(Prof. Dr. H. Stoyan)**

Betreuer: Dipl. Inf. Stefan Mandl, Prof. Dr. H. Stoyan

Beginn der Arbeit: 15.07.2005

Abgabe der Arbeit: 17.11.2005

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Der Universität Erlangen-Nürnberg, vertreten durch die Informatik 8 (Künstliche Intelligenz), wird für Zwecke der Forschung und Lehre ein einfaches, kostenloses, zeitlich und örtlich unbeschränktes Nutzungsrecht an den Arbeitsergebnissen der Diplomarbeit einschließlich etwaiger Schutzrechte und Urheberrechte eingeräumt.

Erlangen, den 15.11.2005

Sebastian von Mammen

Danksagungen

Ich möchte mich sehr herzlich bei meinem Betreuer Stefan Mandl bedanken. Er hat mir wichtige Impulse für die Inhalte dieser Arbeit gegeben, mich durchweg exzellent betreut und im höchstem Maße motiviert.

Weiterhin möchte ich folgenden Personen großen Dank aussprechen, ohne deren Leistungen im Vorfeld diese Arbeit so nicht hätte entstehen können. Zum einen hat Heike Ott große Vorarbeit für die Übersetzung von Texten in numerische Repräsentationen, sowie deren Vergleich aufgrund charakteristischer Vektoren geleistet. Von Bernd Ludwig wurde der Zugriff auf die Fernsehsendungen und deren Dornseiffgruppen ermöglicht. Daniel Loebenberger hat weitgehend dazu beigetragen, die algebraischen Eigenschaften des entwickelten Lernsystems zu formulieren.

Ich möchte mich außerdem bei Regina Engel und Christopher von Mammen bedanken, die sich als Versuchspersonen für Testläufe des Lernsystems zur Verfügung gestellt haben.

Schließlich gebührt meinen Eltern großer Dank, für den großartigen Zuspruch und die liebevolle Unterstützung während meiner Studienzeit.



Zusammenfassung

Entwicklung eines Lernsystems für Gruppenpräferenzen

Nach einer kurzen Einführung über existierende Vorschlagsysteme für einzelne Anwender, wird ein abstraktes Modell für ein Lernsystem für Gruppenpräferenzen vorgestellt. Es beachtet die grundlegende Dynamik einer Gruppe, sowie Kenntnisse über die Kompromisbildung innerhalb einer Gruppe. Durch unterschiedliche Implementierungen der einzelnen Komponenten des Lernsystems erhält man eine Vielzahl von Konfigurationen. Sie werden anhand einer konstruierten Lernsituation miteinander verglichen. Zwei der Konfigurationen, die beide die Erwartungen der Testsituation erfüllen und sich dabei möglichst stark voneinander unterscheiden, werden für Testläufe mit echten Benutzereingaben verwandt.

Für die spezielle Testsituation und die beiden Testläufe werden Fernsehsendungen als Beispieldaten verwandt. Im Zuge der Klassifikation von Fernsehsendungen wird eine Methode vorgestellt, die aus einem Text eine Menge von Vektoren bildet, die Gefühle widerspiegeln sollen.



Abstract

Design of a learning system for group preferences

After a brief introduction to existing recommendation systems for single users, an abstract model of a learning system for group preferences is given. It is built upon facts about group dynamics and group strategies to form compromises. A variety of configurations of the learning system is achieved by different implementations of its components. A test-situation is created to compare the capabilities and limitations of the configurations. The two most deviating configurations that comply with the expectations of the test-situation are chosen for test-runs with real user feedback.

For both, the test-situation and the test-runs, examples from television domain are used. The features to describe different programs are computed by a method that transforms text into a set of vectors that themselves are supposed to represent emotions.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Portabilität, Verfügbarkeit, Qualität, Funktionalität	1
1.2	Die Informationsflut	2
1.3	Lernsysteme	2
1.4	Gliederung der Arbeit	3
2	Problemstellung	5
2.1	Anforderungen an ein Lernsystem für Gruppenpräferenzen	5
2.1.1	Das Lernen von Präferenzen im Einzelfall	5
2.1.2	Das Lernen von Präferenzen in der Gruppe	6
2.2	Klassifikation von Multimediadaten	7
3	Stand der Forschung	9
3.1	Recommender Systems	9
3.1.1	Knowledge Discovery in Databases	10
3.1.2	Collaborative Filtering	10
3.2	Playlist Generation	10
3.2.1	Suche nach ähnlichen Titeln	11
3.2.2	Lernen für bestimmte Situationen	11
3.2.3	Constraintgesteuerte Suche	11
3.3	Vorschlagsysteme für Gruppen	12
3.3.1	Strategien zur Synthese individueller Präferenzen	12
3.3.2	Bestehende Systeme	12
3.3.3	Versuche mit Gruppen	13
3.4	Klassifikation von Multimediadaten	13
4	Abstraktes Lernsystem für Gruppen	15
4.1	Lernen von Bewertungen	15
4.2	Erzeugen einer Hypothese, 1. Fall	16

4.3	Einführen von Layern	16
4.4	Erzeugen einer Hypothese, allgemeiner Fall	17
4.5	Algebraische Eigenschaften des Modells	17
4.6	Zusammenfassung	18
5	Konkrete Realisierung des Lernsystems	19
5.1	Repräsentation der Instanzen	19
5.2	Lernvorgang	19
5.2.1	Bewertungen	20
5.2.2	Widerspruch zwischen individueller und Gruppenbewertung	20
5.2.3	Lazy Learner	20
5.2.4	Schrittweises Verfeinern der Präferenzen $\pi(A)$	22
5.2.5	Modifizieren der Layer $\lambda(A, B)$	23
5.3	Interpolation auf bewerteten Merkmalen	24
5.3.1	„0“ als Standardannahme	25
5.3.2	Voronoi Partitionierung	25
5.3.3	Inverse Distance Weighting	26
5.4	Hypothesenbildung	28
5.4.1	Repräsentation von Hypothesen	28
5.4.2	+ , * Operatoren	28
5.4.3	o Operator	29
5.5	Das Abstandsmaß $D(H, i)$	31
5.5.1	Gesamtabstand („Abst“)	32
5.5.2	Anteilmäßige Bewertung der Merkmale („Gewi“)	34
5.5.3	Kombination von „Abst“ und „Gewi“ („Komb“)	36
5.5.4	Differenz der Durchschnittsvektoren („AvgV“)	37
6	Anwendung und Evaluierung	39
6.1	Ziel der Anwendung	39
6.2	Eingangsdaten und deren Vorverarbeitung	39
6.2.1	Ein Film in Form emotionaler Komponenten	39
6.2.2	Aufbereitung und Auswahl der Beispieldaten	40
6.3	Testsituation für Konfigurationen des Lernsystems	51
6.4	Gewichtete Übernahme von Präferenzen und Layern	56
6.4.1	„0“ als Standardannahme	56
6.4.2	Voronoi Interpolation	62
6.4.3	Inverse Distance Weighted Interpolation	67
6.4.4	Differenz von Durchschnittsvektoren	73

6.5	Alternative Hypothesenbildung	73
6.5.1	„0“ als Standardannahme	73
6.5.2	Voronoi Interpolation	81
6.5.3	Inverse Distance Weighted Interpolation	87
6.6	Tabellarische Zusammenfassung der Ergebnisse	91
6.7	Testläufe mit Benutzereingaben	96
6.7.1	Testlauf „gew-0-Gewi.“ mit zwei Personen	97
6.7.2	Testlauf „alt-Inverse-AvgV“ mit drei Personen	111
7	Zusammenfassung und Ausblick	137
A	Anhang	141

Abbildungsverzeichnis

5.1	Bisher unbewertete Merkmale werden als neutral, in diesem Fall 0 gewertet. Die Pyramidenform um die Extrema ergibt sich nur aufgrund des verwendeten Visualisierungswerkzeugs und entspricht nicht den tatsächlichen Werten.	25
5.2	Für bisher unbewertete Merkmale wird die Bewertung des nächsten Nachbarn angenommen.	26
5.3	Für bisher unbewertete Merkmale werden alle bekannten Bewertungen verrechnet. Je weiter ein Merkmal vom Anfragepunkt entfernt ist, desto geringer ist der Einfluss seiner Bewertung.	27
5.4	Oben: Merkmale und deren Häufigkeiten der Instanz i_{Bsp} . Unten: Merkmale und deren Bewertungen der Hypothese H_{Bsp}	33
5.5	Die interpolierten Untermengen H_{pos} und H_{neg} der Hypothese H_{Bsp}	34
5.6	Die Abbildung zeigt die punktuellen Differenzen zwischen der Instanz i_{Bsp} und der Hypothese H_{Bsp} . Die gestrichelten blauen Linien mit rosanem „x“ kennzeichnen die Abstände zu negativen Hypothesenbewertungen. Die orangenen, durchgehenden Linien mit grünem „+“ kennzeichnen die Abstände zu positiven Werten der Hypothese.	35
5.7	Die Abbildung zeigt die Produkte der Merkmalsbewertungen aus H_{Bsp} und den Merkmalshäufigkeiten aus i_{Bsp}	36
5.8	Der Durchschnittsvektor der Instanz i_{Bsp} ist rot, der Durchschnittsvektor der Menge H_{pos} grün und der Durchschnittsvektor der Menge H_{neg} blau dargestellt. Die Spiegelung des Durchschnittsvektors H_{neg} an der 0-Ebene wird durch den gebogenen Pfeil symbolisiert.	38
6.1	Zwei Vorbereitungsschritte um die Extended Info in Activation/Evaluation Koordinaten zu übersetzen.	41
6.2	Activation/Evaluation Diagramm zur Sendung „Polizeiruf 110: Um Kopf und Kragen“.	44
6.3	Activation/Evaluation Diagramm zur Sendung „Sesamstraße“.	46
6.4	Activation/Evaluation Diagramm zur Sendung „Bianca - Wege zum Glück“.	48
6.5	Activation/Evaluation Diagramm zur Sendung „aspekte kompakt“.	48
6.6	Activation/Evaluation Diagramm zur Sendung „Bayern 3 - die MorningShow“.	51

6.7	Activation/Evaluation Diagramm zur Sendung „88acht GUTEN TAG BERLIN“	52
6.8	Activation/Evaluation Diagramm zur Sendung „nano“	55
6.9	Activation/Evaluation Diagramm zur Sendung „Lindenstraße“	56
6.10	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-0).	57
6.11	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-Voronoi).	63
6.12	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-Inverse).	68
6.13	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-0).	75
6.14	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-Voronoi).	82
6.15	Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-Inverse).	88
6.16	Christophers Präferenzen (nach 15 Bewertungen), Sebastians Präferenzen (nach 12 Bewertungen) und die Hypothese durch deren Übernahme. . . .	109
6.17	Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher und Sebastian, den Layer zwischen den beiden Benutzern und die gebildete Hypothese nach der letzten gemeinsamen Bewertung. .	110
6.18	Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher, Regina und Sebastian vor der Phase des Gruppenlernens. .	132
6.19	Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher, Regina und Sebastian nach der Phase des Gruppenlernens. .	133
6.20	Die Abbildung zeigt die Layer zwischen den paarweise verschiedenen Personen nach der Phase des Gruppenlernens in der Reihenfolge: Christopher ↔ Regina, Christopher ↔ Sebastian, Regina ↔ Sebastian.	134
6.21	Die Abbildung zeigt die Hypothese für Christopher, Regina und Sebastian nach der abgeschlossenen Gruppenlernphase.	135

Tabellenverzeichnis

6.1	Informationen zur Sendung „Polizeiruf 110: Um Kopf und Kragen“	43
6.2	Informationen zur Sendung „Sesamstraße“	46
6.3	Informationen zur Sendung „Bianca - Wege zum Glück“	47
6.4	Informationen zur Sendung „aspekte kompakt“	49
6.5	Informationen zur Sendung „Bayern 3 - die MorningShow“	50
6.6	Wörter der „Extended Info“ von „Bayern 3 - die MorningShow“ und korrespondierende Dornseiffgruppen.	50
6.7	Informationen zur Sendung „88acht GUTEN TAG BERLIN“	50
6.8	Informationen zur Sendung „nano“	54
6.9	Informationen zur Sendung „Lindenstraße“	55
6.10	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	58
6.11	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	60
6.12	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	61
6.13	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	64
6.14	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	65
6.15	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	66
6.16	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lerners (gew-Inverse-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	69

6.17	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Inverse-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	70
6.18	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Inverse-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	72
6.19	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.	74
6.20	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	77
6.21	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	78
6.22	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	79
6.23	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.	80
6.24	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	83
6.25	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	84
6.26	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	85
6.27	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.	86
6.28	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.	89
6.29	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.	90
6.30	Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.	92

6.31 Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.	93
6.32 Zusammenfassung der Testergebnisse. Die Benennungen der Spalten sind in Tabelle 6.33, die Abkürzungen innerhalb der verschiedenen Spalten in Tabelle 6.34 erklärt.	94
6.33 Legende der Spaltenkürzel	95
6.34 Legende der innerhalb der Spalten verwendeten Kürzel	95
6.35 Testlauf der Konfiguration „gew-0-Gewi“ mit bis zu zwei Personen. . . .	107
6.36 Testlauf der Konfiguration „alt-Inve-AvgV“ mit bis zu drei Personen. . .	131
A.1 Gefühle umschreibende Adjektive und deren Activation-/Evaluation-Koordinaten.	144
A.2 Die Liste derjenigen Programmbeispiele, die im Rahmen dieser Arbeit verwendet wurden.	155

1. Einleitung

Der immense technologische Fortschritt in der Unterhaltungsindustrie hält bereits seit mehreren Jahrzehnten an. Durch erst kürzlich etablierte „Schlüsseltechnologien“ in diesem Sektor, wie digitale Datenverarbeitung für die Speicherung und Wiedergabe multimedialer Inhalte oder Verbreitung von Musik und Film über das Internet, haben Funktionalität, Qualität (nicht unbedingt der Inhalte), Verfügbarkeit und Portabilität der Medien eine neue Dimension erreicht.

1.1. Portabilität, Verfügbarkeit, Qualität, Funktionalität

Die Verbreitung von standardisierten Verfahren der En- und Decodierung von multimedialen Inhalten sorgt dafür, dass sie auf den verschiedensten Endgeräten wiedergegeben werden können. Viele aktuelle Mobilfunktelefone, elektronische Terminplaner (sogenannte „Organizer“) und „Handheld“-Spielekonsolen können diverse Musik- und Filmdaten interpretieren und abspielen.

Die unglaubliche Menge an Möglichkeiten, multimediale Inhalte zu „konsumieren“, spiegelt sich in deren weitgestreuter Verfügbarkeit wieder. Im Internet stehen gewaltige Mengen an Filmen, Musikstücken und literarischen Texten zum Download bereit.

Die digitale Speicherung trägt außerdem dafür Sorge, dass der Informationsverlust, den jedes „analoge“ Medium, wie Tonband oder Langspielplatte auf lange Zeit und durch vielfaches Verwenden mit sich bringen, auf ein Minimum reduziert, wenn nicht sogar ganz vermieden wird. Zwar findet aufgrund der Umwandlung der Daten in digitale Form zwingend ein Informationsverlust statt, doch ist der Fortschritt diesbezüglich in vollem Gange. Die Ansprüche der Digitalisierungsstandards werden ständig erhöht (digitales Fernsehen, HD-TV, Blu-Ray Disc, etc.), wohingegen die Unterscheidungsfähigkeit des Menschen bezüglich Anzahl der Bildpunkte pro Flächeneinheit, Anzahl der Farbwerte, Anzahl der Tonfrequenzen et cetera bekanntlich begrenzt ist. In naher Zukunft darf man voraussichtlich mit der für den Menschen perfekten, gleichbleibenden Medienqualität rechnen.

Im gleichen Atemzug werden vielfältige neue Funktionen in Unterhaltungsendgeräte integriert. Ein Beispiel hierfür sind Mehrkanalmusiksysteme, die bei der Wiedergabe von Tonmaterial dem Benutzer ein räumliches Wahrnehmen ermöglichen.

1.2. Die Informationsflut

Zusammen genommen scheinen sich manche der oben genannten Vorteile des Fortschritts zu relativieren. In jeder Straßenecke, im Restaurant, im Hörsaal erklingen die schönsten Balladen aus den Lautsprechern der neusten Mobiltelefone. Die in gestochener Qualität abgespielten Musikstücke sind eigentlich dazu gedacht, den Besitzer des Telefons auf empfangene Informationen oder auf einen Anruf hinzuweisen. Nebenher ziehen sie allerdings gnadenlos die Aufmerksamkeit aller anderen Anwesenden auf sich. Das Mobilfunktelefon vereint die oben genannten fortschrittlichen Entwicklungen in sich und fällt trotz oder gerade wegen all der Fortschrittlichkeit allzu oft negativ auf. In dem erläuterten negativen Extremfall sind die Betroffenen gezwungen, den Lärm, die Desinformation wegzufiltern, um sich weiterhin auf wichtige Signale und Informationen in ihrer Umwelt konzentrieren zu können, die auch für sie bestimmt sind.

Die hohe Verfügbarkeit von multimedialen Inhalten/Informationen zwingt den Konsumenten auch im Normalfall zum Filtern. Stellt man sich ein Fernsehsystem mit mehreren hundert Kanälen vor, sieht man sofort, welche Probleme sich ergäben: Die Wahrscheinlichkeit, dass der Zuschauer ad hoc das Programm wählt, um die für ihn schönste Sendung zu sehen, fällt mit der steigenden Anzahl der Fernsehkanäle.

Könnte der Anwender über das Internet auf den gesamten Filmbestand der Menschheit zugreifen, wäre das Problem noch eindeutiger greifbar. Dann wäre keine Vorselektion durch die Wahl eines bestimmten Fernsehkanals möglich. Auch nach der Beschränkung auf ein bestimmtes Genre (was dem Anwenden eines Filters entspricht) stünden noch unvorstellbar viele Filme zur Auswahl. Nun könnte der Benutzer die Anzahl der vorgeschlagenen Sendungen reduzieren, indem er die Bedingungen an den Film (beziehungsweise die Kriterien des Filters) schrittweise verfeinert.

1.3. Lernsysteme

Ein guter Filter könnte mit Hilfe der Beurteilung des Anwenders auch automatisch erzeugt werden. Die Idee dahinter ist, dass sich ein Computerprogramm merkt, was der Anwender gerne und was er nicht gerne sieht, um ihm später auf dieser Grundlage weitere Vorschläge zu unterbreiten. Für einzelne Anwender gibt es bereits zahlreiche derartige Programme. Der Tatsache dass Fernsehsendungen, Spielfilme und Musik, genau wie Brettspiele, auch vor Gruppen ihren Unterhaltungswert beweisen sollen, wird bisher kaum Rechnung getragen.

Diese Lücke kann durch ein Lernsystem für Gruppen geschlossen werden. Im Rahmen dieser Arbeit soll ein derartiges System entwickelt werden. Es umfasst das Lernen im Einzelfall, das Lernen bei der Anwesenheit einer Gruppe, das Verarbeiten von widersprüchlichen Bewertungen und das Bilden von Kompromissen. Zum Evaluieren der Funktionalität des entwickelten Lernsystems wird es auf die bereits angesprochene Fernsehdomäne angewandt.

1.4. Gliederung der Arbeit

Im folgenden Kapitel werden die Herausforderungen eines Lernsystems für Gruppenpräferenzen, sowie für die Klassifikation multimedialer Inhalte dargelegt. Kapitel 3 trägt aktuelle Informationen über bestehende Vorschlagsysteme, über Verfahren zur Synthese von individuellen Präferenzen und über Klassifikationsmöglichkeiten von Multimedialdaten zusammen. Kapitel 4 gibt eine abstrakte Beschreibung eines Lernsystems für Gruppenpräferenzen. Es folgt die Darlegung einer konkreten Realisierung des abstrakten Lernsystems in Kapitel 5. Das realisierte Lernsystem wird in Kapitel 6 durch Anwendung auf Gruppenvorlieben beim Fernsehschauen evaluiert. Kapitel 7 fasst die Arbeit zusammen und geht auf weitere mögliche Ziele der Forschung ein.

2. Problemstellung

Die Problemstellung dieser Arbeit kann in zwei große Punkte gegliedert werden. Zum einen wird ein Lernsystem für Gruppenpräferenzen entwickelt. Zum anderen wird dieses Lernsystem auf die Wahl von Fernsehsendungen angewandt. Beide Punkte bedingen einige Anforderungen, die im Folgenden erläutert werden.

2.1. Anforderungen an ein Lernsystem für Gruppenpräferenzen

Um die Präferenzen von verschiedenen Personen zu vereinen, muss zunächst auf das Lernen im Einzelfall, anschließend auf das Lernen in der Gruppe eingegangen werden.

2.1.1. Das Lernen von Präferenzen im Einzelfall

Man stelle sich folgendes Problem vor: ein Lernsystem soll die Auswahl von Fernsehsendungen einer einzelnen Person erfassen und nachdem es ausreichend viele Trainingsdaten gesammelt hat, dem Anwender Fernsehsendungen vorschlagen, die er gerne sieht.

Unabhängig von der bestmöglichen konkreten Implementierung für dieses Problem, erkennt man sehr schnell, dass es sich um ein Standardlernproblem handelt, das mit konventioneller induktiver oder deduktiver Inferenz gelöst werden kann (eine Einführung verschiedener Lernalgorithmen gibt Mitchell in [16]).

Je mehr Trainingsdaten der Anwender zur Verfügung stellen würde (im Falle induktiver Verfahren) und/oder je konsistenter er in seiner Filmauswahl wäre (im Falle deduktiver Verfahren), desto schneller würde das Lernverfahren gegen eine gute Hypothese konvergieren und der Anwender wäre glücklich darüber, dass ihm adäquate Vorschläge unterbreitet würden.

Dies setzt voraus, dass die Präferenzen des Anwenders in etwa gleichbleiben und außerdem alle Klassifikationsmerkmale bezüglich der Situation und der Sendungen verfügbar sind, sofern sie die Entscheidungen des Anwenders beeinflussen. Diese Annahmen vereinfachen das Problem sehr stark.

Das gelegentlich notwendige Erweitern des Merkmalsraums der Trainingsdaten, oder das Erkennen von mit der Zeit wichtig/unwichtig gewordener Merkmale, könnte als zweiter Schritt der Entwicklung eines Lernsystems verstanden werden. Bruce Edmonds fasst wesentliche Aspekte dieser „Kontext“-Probleme in [5] zusammen. Im Rahmen dieser Arbeit wird nur noch kurz auf die Verarbeitung derartigen Kontextwissens eingegangen;

auch mit der Annahme gut klassifizierter Trainingsdaten und gleichbleibender Präferenzen der Anwender ergeben sich viele Konflikte, die ein Lernsystem für Gruppenpräferenzen zunächst meistern muss.

2.1.2. Das Lernen von Präferenzen in der Gruppe

Die Definition einer Personengruppe lautet wie folgt:

„Als *Gruppen* gelten soziale Gebilde, die als Ganze oder deren Glieder untereinander sich wechselseitig beeinflussen und steuern. [...]“ [9]

Weiterhin wird zwischen *Primären Gruppen*, *Sekundären Gruppen* und *Informellen Gruppen* unterschieden:

„*Primäre Gruppen* sind übersehbar; ihre Mitglieder treffen zusammen und kennen einander. *Sekundäre Gruppen* sind locker gefügte, größere Gebilde wie Stadtbevölkerung, Berufsgruppe, Verband, Staatsvolk; *informelle Gruppen* lockere Sondergruppierungen mit wechselnder oder unbestimmter Struktur.“ [9]

Im Rahmen dieser Arbeit wird versucht, ein Lernmodell für primäre Gruppen auszuarbeiten. Für die Meinungsbildung innerhalb einer Gruppe ergeben sich zunächst zwei Möglichkeiten.

1. Idealerweise haben alle Mitglieder der Gruppe die gleiche Meinung. Diese wird als Gruppenmeinung akzeptiert.
2. Die Meinungen der Gruppenmitglieder weichen voneinander ab. Es wird ein Kompromiss gefunden, der als Gruppenmeinung akzeptiert wird.

Man stellt sich dabei vor, dass jedes Mitglied der Gruppe seine individuelle Meinung in der Anwesenheit der anderen vertritt und sich durch Diskussion über widersprüchliche Präferenzen von Personen ein Kompromiss ergibt. Entsprechend müsste ein Lernsystems für Gruppenpräferenzen die verschiedenen individuellen Meinungen verrechnen, um den bestmöglichen Kompromiss für die anwesenden Personen zu finden.

Die oben genannten Situationen sagen jedoch nichts darüber aus, ob die in einer Gruppe reflektierte Meinung eines Individuums mit seinen eigentlichen Präferenzen korreliert oder nicht. Beispiele für widersprüchliche Situationen wären:

- Obwohl ein Vater niemals alleine eine Kindersendung sehen wollen würde, würde er sie gerne mit seinem Sohn zusammen anschauen.
- Obwohl der Vater sehr gerne bestimmte Erwachsenenfilme sieht, schließt er es aus, sie in Anwesenheit des Sohnes zu schauen.

Die Definition einer Gruppe (siehe oben) erfasst genau diese Relation zwischen den Gruppenmitgliedern, nämlich dass sie sich „untereinander beeinflussen und steuern [...]“. Widersprüchliche Meinungsäußerungen werden deshalb im Rahmen dieser Arbeit vereinfacht als Resultat der Anwesenheit anderer Personen betrachtet.

Es stellen sich die Fragen „Wie kann man aufgrund der individuellen Präferenzen einen Kompromiss für eine Gruppe finden?“ und „Wie kann man widersprüchliche Meinungsäußerungen verarbeiten?“. In Kapitel 4 wird ein umfassendes Modell vorgestellt, um den Problemen dieser Gruppenlernsituation zu begegnen.

2.2. Klassifikation von Multimediataten

Die Evaluierung des Lernsystems für Gruppenpräferenzen wird in der Fernsehdomäne erörtert. Dafür ist es notwendig, das vorhandene Material (Filme und sonstige Fernsehsendungen) klassifizieren zu können. Es liegt nahe, die direkt zur Verfügung gestellten Attribute zur Klassifikation von Filmmaterial zu verwenden. Die konventionellen *Metadaten* über Filme umfassen den Titel, das Genre, beteiligte Schauspieler und Regisseure, das Erscheinungsjahr, eine inhaltliche Zusammenfassung usw.

Verschiedene Genres wie „Thriller“, „Liebes-“ oder „Horrorfilm“ legen eine Klassifikation von Fernsehsendungen bezüglich ihrer emotionalen Komponenten nahe. In der vorliegenden Arbeit wird der Versuch unternommen, emotionale Komponenten aufgrund von inhaltlichen Zusammenfassungen von Fernsehsendungen zu erfassen und sie als Klassifikationsmerkmale zu verwenden.

Herkömmliche Attribute von Fernsehsendungen (wie die oben genannten) wurden nicht in den Klassifikationsprozess aufgenommen, um die Tauglichkeit der in Kapitel 6 vorgestellten Methode isoliert betrachten zu können.

3. Stand der Forschung

In den ersten beiden Unterkapiteln werden verschiedene Methoden vorgestellt, um die Vorlieben einzelner Anwender mit adäquaten Vorschlägen zu bedienen. Die dabei aufgeführten Informationen über bestehende Algorithmen sollen vor allem dazu dienen, einen groben Überblick über den momentanen Stand der Technik zu erhalten.

Um Vorschläge für mehrere Personen zu generieren, müssen die Strategien erfasst werden, wie innerhalb von Gruppen Kompromisse zustande kommen. Abschnitt 3.3 fasst diese Strategien allgemein zusammen, gibt Beispiele für Computersysteme, die bereits derartige Strategien anwenden, und legt dar, wie Gruppen tatsächlich verfahren.

Schließlich wird auf Vorgehensweisen zur Merkmalsgewinnung multimedialer Informationen eingegangen. Dieser letzte Abschnitt kategorisiert Metadaten von Filmen und Musikstücken und damit benennt die Schwierigkeiten der unterschiedlichen Kategorien. Es ist wichtig, sich der Probleme von Metadaten bewusst zu sein, weil in Kapitel 6 ein Klassifikationsverfahren vorgestellt wird, das auf bestimmten Metadaten über Fernsehsendungen aufbaut.

3.1. Recommender Systems

Gerade im letzten Jahrzehnt ist die massenhafte Kaufabwicklung über das Internet drastisch angestiegen. Da sowohl die Benutzerdaten, als auch die Transaktionen üblicherweise in großen Datenbanken gespeichert werden, liegt es nahe, diese Informationen systematisch auszuwerten. Die dafür entwickelten Algorithmen zielen darauf ab, dem Kunden passende, individuelle Angebote zu unterbreiten. Im Gegensatz zu der vorliegenden Arbeit, geht es bei den etablierten Methoden darum, nur jeweils einem Kunden Vorschläge zu unterbreiten.

Recommender Systems bedienen sich dabei der Informationen über Kunden und deren bisher getätigte Käufe oder Bewertungen. Die stochastische Verarbeitung der Datenbestände dient nur implizit dazu, Aussagen über die Präferenzen eines einzelnen Kunden zu tätigen. Es werden keine Präferenzen im Allgemeinen recherchiert, sondern Mengen an Produkten, die zu den Vorlieben des Kunden passen könnten. Die Produkte werden als repräsentativ für die Präferenzen eines Kunden angenommen.

3.1.1. Knowledge Discovery in Databases

Der traditionelle Data-Mining Ansatz „Knowledge Discovery in Databases“ (KDD) lässt bei der Recherche Benutzerprofile außen vor. Bei einer Anfrage wird innerhalb des gesamten Datenbestands mittels *Data-Mining* Algorithmen versucht, Überschneidungen mit anderen getätigten Einkäufen bezüglich erworbener Produkte zu ermitteln (wobei ein Einkauf natürlich eine Vielzahl von Produkten umfassen kann). Je öfter das vom aktiven Benutzer ausgewählte Produkt zusammen mit einem anderen gekauft wurde, desto dringlicher ist die Empfehlung des jeweils anderen Produkts (im Verhältnis zur Gesamtanzahl von Transaktionen, in denen das Produkt vorkommt).

3.1.2. Collaborative Filtering

Die Idee bei *Collaborative Filtering* (auch „CF“) ist, dass der aktive Benutzer diejenigen Produkte (Bücher, CDs, etc.) bevorzugen könnte, die ihm ähnliche Benutzer mögen, beziehungsweise, dass er vielleicht sogar mag, was ihm unähnliche Benutzer ablehnen (aus [23]). Bei einer Anfrage wird der gesamte Datenbestand nach *Nachbarn* des aktiven Benutzers durchsucht und gegebenenfalls Präferenzen abgeleitet. Nachbarn des aktiven Benutzers sind diejenigen Anwender, die entweder bisher verschiedene Produkte ähnlich bewertet oder ähnliche Produkte erworben haben (aus [25]).

Dabei unterscheidet man zwei Strategien. Zum einen gibt es *speicherbasierte* Algorithmen, die die Präferenzen sämtlicher Anwender in einer Datenbank bereithalten. Die Präferenzen können dadurch sehr leicht erweitert werden. Zum anderen gibt es *modellbasierte* Algorithmen, die die Präferenzen eines Anwenders zunächst in ein Modell übersetzen, das unter anderem eine Anwenderbeschreibung, gekaufte Güter und persönliche Bewertungen umfasst. Der zusätzliche Übersetzungsschritt stellt einen Nachteil gegenüber dem speicherbasierten Algorithmus dar. Vorteilhaft hingegen ist, dass das entworfene Benutzermodell die Wahl und/oder Bewertung des jeweiligen Benutzers verständlich macht und die produzierten Ergebnisse leichter auf ihre Plausibilität hin überprüft werden können.

Die beiden Strategien sind sich im Kern sehr ähnlich. Um den praktischen Nutzen beider zu vereinen, wurden bereits hybride Systeme vorgestellt (siehe wiederum [23]). Die Skalierung, das heißt die Anwendung dieser Systeme auf verschieden große Datenmengen, kann man mittels Dimensionsreduktion der verwendeten Benutzer×Produkte Matrizen in den Griff bekommen (siehe dazu [25]).

3.2. Playlist Generation

Im Musiksektor wurde bereits viel Energie darauf verwendet, automatisch Titel für bestimmte Situationen oder Vorlieben vorzuschlagen. Die Vorschläge werden in sogenannten Playlists gespeichert. Das sind Listen von Musikstücken, die wiedergegeben werden

sollen. Die Ideen dazu decken ein breites Spektrum ab und können in leicht abgewandelter Form auch in der Film- und Fernsehdomäne oder sonstigen Bereichen angewandt werden. In den folgenden Abschnitten werden verschiedene Ideen erläutert, um Playlists zu generieren.

3.2.1. Suche nach ähnlichen Titeln

Nach Vorgabe von einem oder mehreren Titeln kann bezüglich festgelegter Kriterien eine Ähnlichkeit zu anderen Titeln berechnet werden. Manche Systeme machen sich dies zunutze, um Playlisten aus wenigen, manuell ausgesuchten Musikstücken zu erstellen. In [11] erweist sich das Frequenzspektrum eines Titels in Kombination mit einer Genreangabe als dafür geeignetes Ähnlichkeitsmaß.

3.2.2. Lernen für bestimmte Situationen

PATS (Personalized Automatic Track Selection) lernt zuerst durch Benutzerbewertungen, welche Attribute von Musiktiteln für welche Situation wichtig sind (in [22]). Für bestimmte Situationen werden anschließend Gruppen von Titeln erstellt. Abhängig von der Situation kommen den Attributen eines Musikstücks verschiedene Gewichtungen zu. Durch die Verrechnung der Attribute und Gewichtungen qualifizieren sich manche Titel besser für eine Situation als andere.

3.2.3. Constraintgesteuerte Suche

In [2] wird ein System vorgestellt das effizient Playlisten erstellt, die vorgegebene Constraints erfüllen. Dafür wird in Kauf genommen, dass keine optimalen Lösungen gefunden werden. Die Constraints sind unter anderem:

- „All different“: Nur unterschiedliche Titel kommen in der Playlist vor
- „Duration“: Die Gesamtspielzeit der Playlist ist geringer als 76 Minuten
- „Continuity“: Die Genres der Titel sind einander ähnlich
- „Progression“: Das Tempo der Titel nimmt mit der Zeit zu

Das Nicht-/Erfüllen eines Constraints wird in Kosten überführt. Der Suchalgorithmus terminiert, sobald die Gesamtkosten der Playlittitel unter einem festgelegten Schwellwert liegen. Die Playlist wird optimiert, indem stets für den Titel mit den höchsten Kosten bezüglich eines Constraints die lokal beste Lösung gefunden wird. Mit einigen Hilfen, wie zum Beispiel durch „Tabu“-Markierungen von Titeln mit lokalen Maxima bezüglich einzelner Constraints, erweitern Aucouturier und Pachet die Skalierbarkeit dieses Algorithmus zusätzlich.

3.3. Vorschlagsysteme für Gruppen

Die bisher aufgeführten Systeme generieren Vorschläge für nur eine Person (sofern die Vorschläge nicht durch die Situation, wie bei PATS, oder durch allgemeine Constraints, wie im letzten Abschnitt, geprägt sind). Die Erläuterungen und experimentellen Ergebnisse über Gruppenpräferenzen, von denen in diesem Unterkapitel kurz berichtet wird, sind Grundlage des im folgenden Kapitel (Kapitel 4) vorgestellten abstrakten Lernsystems.

3.3.1. Strategien zur Synthese individueller Präferenzen

Masthoff fasst in [12] Möglichkeiten zusammen, wie man für mehrere Personen mit unterschiedlichen Meinungen Vorschläge generieren könnte. Man stelle sich dabei vor, dass jedes Mitglied einer Gruppe Bewertungen für alle zur Auswahl stehenden Instanzen vergibt. Es gibt mehrere Strategien, wie aus diesen persönlichen Bewertungen Gruppenbewertungen der Instanzen generiert werden können.

1. „Average Strategy“: Es wird der Mittelwert der Bewertungen einer Instanz gebildet
2. „Least Misery Strategy“: Die geringste Bewertung einer Instanz setzt sich durch
3. „Most Pleasure Strategy“: Die höchste Bewertung einer Instanz setzt sich durch
4. „Average Without Misery Strategy“: Der Mittelwert wird nur dann übernommen, wenn er oberhalb eines vorbestimmten Schwellwerts liegt
5. „Fairness Strategy“: Nacheinander wählt jede Person die Sendung, die sie am liebsten sehen würde
6. „Most Respected Person Strategy“: Die Präferenzen einer bestimmten Person geben vor, was geschaut wird

3.3.2. Bestehende Systeme

Masthoff geht außerdem auf zwei existierende Systeme ein, die Gruppenvorschläge aufgrund von individuellen Präferenzen erzeugen.

Zum einen *MusicFX*, das Radiosender als Quelle für Hintergrundmusik in Fitnesscentern, mittels der oben aufgeführten „Average Without Misery Strategy“, auswählt [13]. Die Bewertungen der Benutzer reichen dabei von -2 bis $+2$, was „wirklich schlecht“ (really hate this music) bis „wirklich gut“ (really love this music) entspricht. Die Bewertungen werden durch Addition von $+2$ in positive Zahlen umgewandelt und quadriert, um den Abstand zwischen positiven und negativen Bewertungen zu vergrößern. Um zu vermeiden, dass stets der gleiche Radiosender läuft, wird eine Zufallsauswahl bezüglich der m am besten bewerteten Sender durchgeführt.

Zum anderen *PolyLens*, das Vorschläge für Filme, abhängig von den Benutzerbewertungen und anderen Filtern, generiert [17]. Dabei werden nach der „Least Misery Strategy“ für die anwesende Gruppe Vorschläge unterbreitet.

3.3.3. Versuche mit Gruppen

Neben der Zusammenfassung verschiedener Verrechnungsstrategien für Gruppenvorschläge und den Referenzen auf zuvor erschienene Werke in diesem Sektor, glänzt Masthoffs Arbeit mit Experimenten, die untersuchen, nach welchen Strategien sich Gruppen tatsächlich verhalten. Eine einheitliche Strategie konnte allerdings nicht festgestellt werden. Oft wird nach der „Average“, der „Average Without Misery“ und der „Least Misery“ Strategie verfahren.

„Fairness plays a role, but our human subjects did not have a clear strategy for applying it.”

Alles in allem wurde augenscheinig, dass sich die Meinungen der einzelnen Versuchspersonen von vielerlei Faktoren beeinflussen lassen. So spielen die Autorität verschiedener Personen (zum Beispiel des Alters wegen) oder der Wille zur Berücksichtigung von Präferenzen, die bisher noch keine Geltung fanden, eine Rolle.

3.4. Klassifikation von Multimediatdaten

Die Verarbeitung von multimedialen Informationen wird von vielen Seiten beleuchtet, zum Beispiel in Bezug auf Speicherung (siehe [28]), Wiedergewinnen (z.B. in [8] oder [14]) und visuelle Aufbereitung (wie in [1], [26] oder [20]).

Im Rahmen dieser Arbeit wird ein Katalog von Fernsehsendungen durchforstet, der mit *EPG*-Daten für die einzelnen Titel aufwarten kann. *EPG* steht für *Electronic Program Guide* und umfasst den Titel der Sendung, eine Kurzbeschreibung, eine ausführlichere Beschreibung, das Erscheinungsjahr, die Dauer der Sendung et cetera.

Es steht also bereits eine Anzahl an Merkmalen zur Verfügung, durch die die Klassifikation der Sendungen bezüglich des Publikumswunsches passieren kann. Man könnte allerdings die Gewinnung weiterer Merkmale anstreben, indem man das Bild- und Tonmaterial automatisch auf gewisse Eigenheiten hin untersuchen ließe. Unter anderem in [29], [10], [21] und [15] wurden beispielsweise die Tonspuren von Fernsehsendungen analysiert, um mit charakteristischen akustischen Merkmalen verschiedene Typen von Sendungen zu unterscheiden. Leicht zugängliche Merkmale sind zum Beispiel das Tempo, die zugrundeliegende Frequenz oder die Pausen innerhalb eines Audiofragments.

Pachet trägt in [19] Arten von Metadaten über Musik zusammen und geht dabei auf ihre Schwachstellen ein. Seine Einteilung und auch die Hinweise auf Probleme können weitestgehend auf Film- und Fernsehsendungen übertragen werden.

- „Editorial Metadata“: Alle Daten die durch einen oder mehrere Editoren angegeben werden. Neben dem Titel, Künstlern, Erscheinungsjahr, Erscheinungsort können auch weiterführende Informationen über das Stück angegeben werden (historischer Hintergrund, Künstlerbiographien, usw.). Die für die Klassifikation wichtige Zuordnung zu einem Genre ist für gewöhnlich subjektiver Natur. Die Herausforderungen sind die strenge Einhaltung einer Taxonomie für Genres, mögliche Erweiterungen (wegen der Neu- und Weiterentwicklung von Genres) und eine sorgfältige Trennung zwischen der Verwaltung von Metadaten und deren Eingabe.
- „Cultural Metadata“: Dahinter verbergen sich Informationen, die wegen der Umgebung oder Kultur mit dem jeweiligen Titel in Zusammenhang gebracht werden. Würde man die in Abschnitt 3.1.1 vorgestellte Data-Mining Methode KDD auf einen Schauspieler (bspw. Bruce Willis) anwenden, so könnte man eine Liste an Genres ermitteln, die mit dem Schauspieler assoziiert werden (bspw. „Actionfilm“ oder „Science Fiction“). In der Mehrzahl der Fälle stimmen „Editorial“ und „Cultural“ Metadaten nicht überein.
- „Acoustic Metadata“: Diese Metadaten werden durch Analyse der Audiodaten erzeugt. Wie oben erwähnt, kann man verschiedene Charakteristika (wie Tempo oder Frequenzbereich) für Audiofragmente automatisch ermitteln. Die Rohdaten von Filmen könnte man auf ähnliche Weise auf Muster durchsuchen. Probleme verursacht, dass sich einzelne Kenngrößen auf ein gesamtes Musikstück (oder auf einen Datenstrom bestimmter Länge) beziehen. Im Gegensatz zu derartigen *unären* Beschreibungswerten, gibt es auch *nicht unäre*, die den Verlauf eines Musikstücks (bzw. Datenstroms) dokumentieren. Das größte Problem ist, dass abgesehen von trivialen Kennwerten noch kein Vorrat musikalisch fundierter Größen ausgemacht werden konnte.

Die Einteilung von Titeln in verschiedene Genres bringt noch weitere Probleme mit sich (siehe [18]). Es hat sich erwiesen, dass die Grenzen zwischen verschiedenen Genres fließend sind, weshalb eine klare Zuordnung zu einem Genre sehr schwierig sein kann. In Musikalben treffen sogar vollkommen verschiedene Genres aufeinander, was wiederum keine direkte Zuteilung des Albums zu einem Genre erlaubt. Die komplexe Terminologie, mit der Experten in der Lage sind, konsistente Klassifikationen vorzunehmen, ist nicht unbedingt für den potentiellen Konsumenten/Kunden eines Mediums nachvollziehbar. Sofern die entsprechenden Begriffe bekannt sind, weichen die Vorstellungen über deren Bedeutung oft deutlich voneinander ab.

4. Abstraktes Lernsystem für Gruppen

In diesem Kapitel wird ein abstraktes Modell für ein Lernsystem für Gruppen entworfen. Es zieht sowohl die in Kapitel 2 dargelegten „Vater und Sohn“-Probleme, als auch die in Kapitel 3.3 erläuterten Strategien zur Bildung einer Gruppenmeinung in Betracht zieht.

4.1. Lernen von Bewertungen

Das Lernsystem muss Bewertungen für beliebige Instanzen der jeweiligen Domäne für beliebig viele Personen entgegennehmen. Instanzen können unter anderem Fernsehsendungen, literarische Texte, Kochrezepte, Musikstücke, etc. sein, die wiederum durch unterschiedliche Merkmale repräsentiert werden. Ein Merkmal wird verstanden als ein Attribut und ein dazugehöriger Wert, beispielsweise „Farbe = gelb“. Die Bewertung einer Instanz i resultiert in einer Menge von Merkmalen und einer jeweils assoziierten Bewertung.

Sei $Merm$ die Menge aller Merkmale und Bew die Menge aller Bewertungen. Eine Instanz i wird beschrieben durch eine Teilmenge aller möglichen Merkmale, ist demnach Element der Potenzmenge \mathcal{P} von $Merm$.

$$i \in \mathcal{P}(Merm)$$

Es sei $Bmerm$ die Menge aller Mengen bewerteter Merkmale.

$$Bmerm = \mathcal{P}(Merm \times Bew)$$

Die Bewertung von i führt zu einer Menge bewerteter Merkmale ($Bmerm_i$), also einem Element von $Bmerm$.

$$Bmerm_i \in Bmerm$$

Die Funktion $\pi(A)$ entspricht den Präferenzen einer Person A und stellt eine Erweiterung der bekannten Bewertungen einzelner Merkmale auf den gesamten Merkmalsraum dar.

$$\pi(A) \in Bmerm$$

Die Funktion $\pi(A)$ wird aufgrund der von A bewerteten Menge von Instanzen, I_A , gelernt.

$$Bmerm_i \Rightarrow \pi(A) \quad , \forall i \in I_A$$

4.2. Erzeugen einer Hypothese, 1. Fall

Gegeben sei folgender Verknüpfungsoperator:

$$+ : Bmerm \times Bmerm \mapsto Bmerm$$

Mit diesem Operator und den gelernten Präferenzen π der anwesenden Personen (zusammengenommen die Gruppe G) lässt sich nun eine Hypothese H_G generieren. Dabei könnte für den $+$ Operator beispielsweise eine der Verrechnungsmethoden aus Kapitel 3.3 Anwendung finden.

$$H_G = \pi(0) + \cdots + \pi(p) \quad , \text{ mit } 0, \dots, k \in G \quad (4.1)$$

Die Adäquatheit einer Instanz i für eine Hypothese H wird durch ein Distanzmaß $D(H, i)$ bestimmt. Bei einer gegebenen Hypothese H kann man somit eine partielle Ordnung auf der Menge aller zur Verfügung stehenden Instanzen definieren.

4.3. Einführen von Layern

Es wurde bereits eingeräumt, dass die erreichte Ordnung auf den Instanzen, selbst bei sorgfältiger Auswahl des $+$ Operators und des Distanzmaßes D , im Einzelfall nicht der Gruppenmeinung entspricht (aufgrund der „Vater und Sohn“-Probleme, siehe Kapitel 2, beziehungsweise aufgrund der experimentell nachgewiesenen Gruppendynamik, siehe Kapitel 3.3).

Die Gruppenmeinung hängt von der Gruppenkonstellation ab. Es kann vorkommen, dass Person A in der Gegenwart von Person B eine andere Meinung preisgibt, als durch ihre individuellen Präferenzen reflektiert wird. Dadurch würde sich die erreichte Ordnung mit der Gruppenbewertung widersprechen. Um den gegenseitigen Einfluss zwischen paarweise verschiedenen Mitgliedern zu erfassen, wird ein sogenannter Layer zwischen allen paarweise verschiedenen Personen A und B modifiziert.

Ein Layer zwischen A und B wird als zweistellige Funktion, $\lambda(A, B)$, verstanden, die Merkmalsmengen von A und B auf Bewertungen abbildet.

$$\lambda(A, B) \in Bmerm$$

Sei I_G^W die Menge aller Instanzen, die die anwesende Gruppe G widersprüchlich zur berechneten Ordnung bewertet hat. Und sei I_{AB}^W die Menge aller Elemente aus I_G^W , zu denen $\pi(A)$ oder $\pi(B)$ in Widerspruch stehen. So lernt die Funktion $\lambda(A, B)$ für alle paarweise verschiedenen Personen A und B aus G , die Gruppenbewertungen aller Elemente von I_{AB}^W .

$$Bmerm_i \Rightarrow \lambda(A, B) \quad , \forall i \in I_{AB}^W$$

4.4. Erzeugen einer Hypothese, allgemeiner Fall

Die Berechnung der Hypothese nach Gleichung 4.1 beachtet keine Layer. Um sowohl Präferenzen als auch Layer einer Gruppe in die Hypothese einfließen zu lassen, müssen diese zunächst separat verarbeitet werden.

Analog zu Definition 4.1 wird folgende Verknüpfung für die Verrechnung der Layer (λ Funktionen) eingeführt.

$$* : Bmerm \times Bmerm \mapsto Bmerm$$

Sämtliche Präferenzen werden mittels des $+$ Operators zur Gesamtpräferenz Π_G , sämtliche Layer mittels des $*$ Operators zum Gesamtlayer Λ_G der Gruppe G verrechnet. Die Gruppe G umfasst n Personen P .

$$\begin{aligned} \Pi_G &= \pi(P_0) + \dots + \pi(P_{n-1}) \\ \Lambda_G &= \lambda(P_0, P_1) * \dots * \lambda(P_i, P_j) \quad , \text{mit } 0 \leq i < j \leq n \end{aligned}$$

Der Gesamtlayer schlägt sich auf andere Weise in der Hypothese nieder, als die Gesamtpräferenz. Deshalb wird ein weiterer Operator definiert, um beide zu einer Hypothese zu verknüpfen.

$$\circ : Bmerm \times Bmerm \mapsto Bmerm$$

Die Hypothese für die Gruppe G lässt sich nun folgendermaßen berechnen:

$$H_G = \Pi_G \circ \Lambda_G$$

4.5. Algebraische Eigenschaften des Modells

Die Operatoren $+$ und $*$ müssen kommutativ sein, da man von Mengen anwesender Personen ausgeht, wobei keine Reihenfolge vorgegeben ist.

$$\begin{aligned} \pi(A) + \pi(B) &= \pi(B) + \pi(A) \\ \lambda(A, B) * \lambda(C, D) &= \lambda(C, D) * \lambda(A, B) \end{aligned}$$

Außerdem lassen sich die Operatoren assoziativ anwenden.

$$\begin{aligned} \pi(A) + (\pi(B) + \pi(C)) &= (\pi(A) + \pi(B)) + \pi(C) \\ \lambda(A, B) * (\lambda(C, D) * \lambda(E, F)) &= (\lambda(A, B) * \lambda(C, D)) * \lambda(E, F) \end{aligned}$$

Sowohl für $+$ als auch für $*$ existiert ein neutrales Element. Sind die vorliegenden Präferenzen oder Layer leer, beeinflussen sie das Ergebnis nicht.

$$\begin{aligned} \exists e_p \forall \pi \quad e_p + \pi &= \pi + e_p = \pi \\ \exists e_l \forall \lambda \quad e_l * \lambda &= \lambda * e_l = \lambda \end{aligned}$$

Da ein Layer stets für zwei Personen A und B gelernt wird, und er außerdem nur bei Anwesenheit beider Personen modifiziert wird, ist es sinnvoll λ symmetrisch in den Argumenten zu gestalten. So stellen wir eine weitere Bedingung an λ :

$$\lambda(A, B) = \lambda(B, A)$$

$+$ und $*$ definieren beide einen kommutativen Monoid auf der Menge von Mengen bewerteter Merkmale.

Bei der Verknüpfung von Präferenzen und Layern zur aktuellen Hypothese behandelt die \circ Verknüpfung die beiden Operanden unterschiedlich. Die Reihenfolge der Operanden ist daher wichtig. Da außerdem die Reduktion der Präferenzen und Layer auf jeweils ein Objekt möglich ist, bestehen für den \circ Operator keine Ansprüche auf Kommutativität oder Assoziativität. Sowohl für Präferenzen, als auch für Layer muss es ein neutrales Element für \circ geben, da weder leere Präferenzen noch leere Layer das Ergebnis nicht beeinflussen dürfen.

$$\begin{array}{ll} \exists e_l \forall \Pi & \Pi \circ e_l = \Pi \\ \exists e_p \forall \Lambda & e_p \circ \Lambda = \Lambda \end{array}$$

4.6. Zusammenfassung

Um eine Implementierung des abstrakten Lernsystems für Gruppen zu realisieren, muss man konkrete Konzepte folgender Teile des Modells entwickeln:

- Repräsentation der Instanzen
- Wertebereich für Bewertungen
- Definition eines Widerspruchs zwischen individuellen und Gruppenbewertungen
- $\pi(A)$, $\lambda(A, B)$, $D(H, i)$
- $+$, $*$ und \circ Operator

Im folgenden Kapitel werden konkrete Implementierungsmöglichkeiten vorgestellt.

5. Konkrete Realisierung des Lernsystems

Dieses Kapitel beschreibt mögliche Implementierungen aller Teile des in Kapitel 4 dargestellten Lernsystems für Gruppenpräferenzen.

Die Bestandteile des Lernsystems, die in Abschnitt 5.1 und 5.2 erläutert werden, liegen jeder Ausführung des Lernsystems zugrunde. Für die nachfolgenden Komponenten des Lernsystems, werden jeweils mehrere Möglichkeiten angegeben, deren Kombinationen im Kapitel 6 bezüglich einer konkreten Lernsituation getestet werden. Die Kombination verschiedener Bestandteile wird *Konfiguration* genannt.

Die Vielzahl der Konfigurationsmöglichkeiten des Lernsystems macht es teilweise notwendig, auf Abkürzungen zurückzugreifen. Für jede Implementierung einer vorgestellten Komponente wird deshalb ein Kürzel eingeführt.

5.1. Repräsentation der Instanzen

Zunächst muss eine Repräsentation für Instanzen der Domäne gefunden werden. Im vorliegenden Fall wird eine Instanz durch eine Menge dreidimensionaler Vektoren repräsentiert. Im Kapitel 6 wird aufgezeigt, was diese Vektormengen bedeuten (können). Vorerst ist nur wichtig, dass die ersten beiden Vektorelemente ein Merkmal beschreiben, die dritte Komponente die Häufigkeit des Merkmals auszeichnet. Abgesehen davon sollte man die Vektormengen als allgemeine Art einer Instanzenrepräsentation verstehen, die praktischerweise sehr anschaulich ist. Verknüpfungsoptoren, Interpolationsfunktionen und Distanzmaße lassen sich nämlich mittels 3D-Vektormengen einfach visualisieren.

5.2. Lernvorgang

Bevor auf die Vorlieben des anwesenden Publikums reagiert werden kann, müssen die einzelnen Präferenzen und Layer gelernt werden. Das setzt wiederum die Realisierung eines Bewertungsschemas sowie die klare Definition eines Widerspruchs zwischen Bewertungen voraus. Danach ist die Grundlage geschaffen, das Lernsystem mit Trainingsbeispielen zu versorgen. Wie das System die Trainingsbeispiele in Präferenzen und Layer überführt, wird in Kapitel 4 nicht erwähnt. In diesem Kapitel wird dafür eine einheitliche Vorgehensweise erläutert.

5.2.1. Bewertungen

Es wird jede Zahl zwischen -1 und $+1$ als Bewertung einer Instanz zugelassen. Durch das Vorzeichen alleine kann bereits eine grundlegende Wertung vollzogen werden. Ein negativer Wert steht für die Ablehnung, ein positiver Wert für die Akzeptanz der bewerteten Instanz. Die Festlegung auf einen bestimmten Betrag gibt dem Anwender die Möglichkeit, diese generelle Bewertung zu differenzieren. 0 als Mittelpunkt des Intervalls kann intuitiv als „neutrale“ Bewertung verstanden werden. Der maximale Betrag 1 stellt keine allgemeine Limitierung dar, sondern wurde als oberste Grenze festgelegt, um möglichst einheitliche Bewertungen zu erhalten.

In [12] wird darauf hingewiesen, dass die Erhebung von Bewertungen vielerlei Tücken mit sich bringt. Einerseits steht nicht fest, auf welche Weise zwischen der Bedeutung einer Bewertung und dem Betrag einer Bewertung abgebildet werden kann; es ist nicht klar, ob die Bewertung $0,9$ gegenüber $0,8$ um genauso viel besser ist wie $0,2$ gegenüber $0,1$. Desweiteren kann man von verschiedenen Anwendern keine einheitliche Strategie für die Zuordnung von Bewertungen erwarten.

Würde man sich aufgrund dieser Probleme auf ein anderes Intervall beziehen wollen, könnte man dieses problemlos auf den Bereich $[-1, +1]$ abbilden. In die Abbildungsfunktion könnte außerdem Hintergrundwissen über das oben erläuterte Verhältnis zwischen Bedeutung einer Bewertung und dem Betrag einer Bewertung einfließen.

5.2.2. Widerspruch zwischen individueller und Gruppenbewertung

Es liegt nahe, auf dem Intervall $[-1, +1]$ einen Widerspruch festzustellen, wenn zwei Werte verschiedene Vorzeichen aufweisen. Das ist eine starke Vereinfachung, die nur zwischen Akzeptanz und Ablehnung einer Instanz unterscheidet. Allerdings zeichnet sich diese Methode einen Widerspruch festzustellen dadurch aus, dass sie eine klare Unterscheidung ermöglicht, und dass sie stets leicht nachvollziehbare Ergebnisse liefert.

5.2.3. Lazy Learner

Die Trainingsbeispiele, um die sich der Anwender im Laufe der Zeit durch Bewertungen von Instanzen bemüht, sind ein rares Gut. Nur wenn ausreichend große Mengen an Trainingsdaten zur Verfügung stehen, die außerdem den Merkmalsraum möglichst gut abdecken, können Lernsysteme prinzipiell leisten, was sie versprechen (eine allgemeine Einführung in dieses Thema vermittelt Mitchell in [16]).

„Lazy Learner“ beschreibt eine Gattung von maschinellen Lernalgorithmen, die während der Lernphase (in unserem Fall das Bewerten von Instanzen) nichts weiter verrichten, als die Trainingsbeispiele entgegenzunehmen. Die Archivierung sämtlicher Trainingsdaten wird dadurch bereits sichergestellt.

Erst bei einer Klassifizierungsanfrage wird für die aktuelle Situation aufgrund der Trainingsdaten eine Hypothese gebildet. Auch das deckt sich mit den Ansprüchen eines

Lernsystems für Gruppen. Da sich die Konstellation der anwesenden Personen ständig ändern kann, ist es sinnvoll, erst dann eine Hypothese zu bilden, wenn Vorschläge für eine bestimmte Gruppe generiert werden müssen. Neben der Konstellation des Publikums ändern sich außerdem die individuellen und die Gruppenpräferenzen im Lauf der Zeit.

Der Veränderung von Präferenzen in Abhängigkeit der Zeit könnte Rechnung getragen werden, indem man beim Überführen der Trainingsdaten in π , beziehungsweise in λ , die Wichtigkeit der Daten von ihrem Alter abhängig macht. Also ein sogenanntes *Zeitfenster*, wie zum Beispiel in [27], benutzt. Es wird stets der Zeitpunkt einer Bewertung festgehalten. Würde man schlichtweg die Bewertung durch die Anzahl der nach dem Bewertungszeitpunkt verstrichenen Tage oder Wochen teilen, würde sie gegen den neutralen Wert 0 konvergieren und würde an Einfluss verlieren. Das Lernsystem ist nicht darauf ausgelegt, einzelne Merkmale von Instanzen unterschiedlich zu gewichten, es führt keine zusätzlichen Merkmale ein und erkennt keine Interdependenzen zwischen den Merkmalen. Das Zeitfenster ermöglicht aber zumindest wechselnde Akzentuierungen der Präferenzen. Auch für ein Zeitfenster bietet sich ein „Lazy Learner“ an: für die Trainingsdaten wird zum Anfragezeitpunkt die Aktualität bestimmt und darauf aufbauend wird eine Hypothese für den Moment generiert. Für die in Kapitel 6 aufgeführten, sehr kurzen Probeläufe der Anwendung wird kein Zeitfenster berücksichtigt. Bewertete Daten werden in der vorliegenden Anwendung jedoch bereits mit Zeitstempel versehen.

Schließlich kann „Lazy Learning“ mit einem weiteren großen Vorteil aufwarten. Durch Separieren des Lernens vom Hypothesenbilden kann man die Strategie zur Bildung der Hypothese nach Belieben wechseln, ohne das System per se verändern zu müssen. Auch könnten Routinen vor das Hypothesenbilden geschaltet werden, die die mit Rauschen behafteten Trainingsdaten glätten etc.

Im vorliegenden Fall wurde „Lazy Learning“ verwirklicht, indem die Instanzen mit der Bewertung verknüpft und dann für Präferenzen von Personen oder für Layer zwischen Personen gespeichert werden. Die Verknüpfung greift bereits auf die im nächsten Kapitel (Kapitel 6) erläuterte Anwendung vor. Sie besteht darin, dass die Häufigkeit eines Merkmals mit der Bewertung multipliziert wird. Die Häufigkeit sagt in diesem Fall aus, inwiefern das Merkmal zur Gesamtbeschreibung der Instanz beiträgt - entsprechend fällt die Bewertung anteilig auf dieses Merkmal zurück. Für den allgemeinen Fall bietet sich an, die Bewertung als zusätzliche Vektorkomponente zu speichern.

Gleich nach der Bewertung einer Instanz wird vom Lernsystem erörtert, ob das vorliegende Trainingsbeispiel (ein Trainingsbeispiel entspricht einer Instanz und ihrer Bewertung) für individuelle Präferenzen einer Person oder für Layer zwischen zwei Personen gespeichert wird. Sobald eine Hypothese gebildet werden muss, um Vorschläge zu generieren, werden die für jede Person und jedes Personenpaar gespeicherten Daten in die Funktionen π für individuelle Präferenzen und λ für Layer zwischen zwei Personen überführt.

5.2.4. Schrittweises Verfeinern der Präferenzen $\pi(A)$

Falls sich die abgegebene Bewertung der anwesenden Gruppe G nicht mit den individuellen Präferenzen von Person A (mit $A \in G$) widerspricht (Widerspruch über Vorzeichen, siehe Abschnitt 5.2.2), oder ausschließlich A anwesend ist, wird die bewertete Instanz für die Präferenzen von A übernommen.

Genau genommen wird diese Überlegung für jedes einzelne Merkmal eines neu erschlossenen Trainingsbeispiels durchgeführt. Das heißt, jedes Merkmal des Trainingsbeispiels, für das die Funktion $\pi(A)$ nicht im Widerspruch mit der momentanen Gruppenbewertung steht, bekräftigt die Präferenzen von A . Diejenigen individuellen Merkmalsbewertungen werden also bestärkt, die sich in der Gruppe durchgesetzt haben.

Die für Person A gelernten bewerteten Merkmalsmengen werden wie folgt in $\pi(A)$ überführt:

1. Addieren sämtlicher Bewertungen bezüglich eines Merkmals.
2. Normieren der Gesamtbewertung auf 1.
3. Für nicht explizit gelernte Bewertungen bezüglich eines Merkmals wird entweder ein neutraler Wert (0) zurückgeliefert, oder aufgrund der vorhandenen Stützpunkte interpoliert (in Abschnitt 5.3 wird näher darauf eingegangen).

Die einzelnen Schritte ergeben sich aus den folgenden Überlegungen:

1. Durch Addition von Bewertungen aus dem Intervall $[-1, +1]$ verstärken sich die Werte, falls viele Beispiele mit dem entsprechenden Merkmal einheitlich (also „nicht widersprüchlich“) bewertet wurden. Falls ein Merkmal keine Rückschlüsse auf die Bewertungen von Instanzen zulässt - dann wenn manche Instanzen, die das Merkmal aufweisen, positiv, manche negativ bewertet wurden - gleicht sich die summierte Bewertung in entsprechendem Grad aus. Der Wert verringert sich, tendiert unter Umständen sogar gegen die neutrale 0.
2. Beim Verrechnen der Präferenzen aller Personen zu einer gemeinsamen Hypothese, soll jeder Person der gleiche Stellenwert eingeräumt werden. Bloße Addition treibt die Merkmalsbewertungen einer Person unter Umständen stark in die Höhe. Eine Normierung aller Merkmalsbewertungen auf eine Summe von 1 für jede Person erhält die individuellen Bewertungen und macht sie untereinander vergleichbar, unabhängig von der Anzahl bewerteter Instanzen pro Person.
3. π muss den gesamten Merkmalsraum abdecken, damit man die Präferenzen zweier Personen bezüglich beliebiger Merkmale vergleichen kann. Man kann nicht davon ausgehen, dass für eine Person die Bewertungen aller möglichen Merkmale bekannt sind. Deshalb greift man auf ein Verfahren zurück, das auch nicht explizit gelernte Werte zurückgibt.

5.2.5. Modifizieren der Layer $\lambda(A, B)$

Falls sich die abgegebene Bewertung der anwesenden Gruppe G nicht mit den individuellen Präferenzen der paarweise verschiedenen Personen A und B (mit $A, B \in G$) deckt, wird die bewertete Instanz in den Layer von A und B übernommen.

Im Detail geschieht dies für jedes Merkmal der bewerteten Instanz. Es wird für jedes Merkmal der Instanz überprüft, ob seine aktuelle (Gruppen-)Bewertung den individuellen Merkmalsbewertungen von A und denen von B gleicht. Wenn mindestens einem Widerspruch vorliegt (zwischen der Gruppe und A oder der Gruppe und B), wird das bewertete Merkmal in den Layer von A und B übernommen.

Für Layer werden also diejenigen Merkmale samt Bewertungen gelernt, bezüglich derer mindestens eine der beiden Personen A und B eine andere Meinung hat als die Gruppe. Das heißt, in Layern wird gespeichert, wie die Gruppe bei Anwesenheit von A und B entscheidet, sofern sich die individuellen Präferenzen von A oder B von der Gruppenmeinung unterscheiden. Man kann sagen, die Information der Layer überlagert die individuellen Präferenzen.

Das Überführen von bewerteten Merkmalsmengen zwischen zwei Personen A und B in $\lambda(A, B)$ gleicht dem Überführen der bewerteten Merkmalsmengen einer Person A in $\pi(A)$. Im Folgenden werden die einzelnen Schritte nochmals stichpunktartig aufgeführt.

1. Addieren sämtlicher Bewertungen bezüglich eines Merkmals.
2. Normieren der Gesamtbewertung auf 1.
3. Generieren von Werten für nicht gelernte Merkmale.

Die Argumente zur Verarbeitung der bewerteten Merkmalsmengen für Layer sind in ihren Grundaussagen die gleichen wie die für das Verarbeitungsschema der Präferenzen. In Layern schlagen sich allerdings die Gruppenmeinungen nieder, weswegen die Verrechnung der bewerteten Merkmale erneut und von einem anderen Blickwinkel beleuchtet werden muss.

1. Addition der Merkmalsbewertungen bekräftigt die Gruppenmeinung für die Personen A und B , unabhängig davon, ob sie einer oder beider Personen widerspricht (im Fall, dass die Präferenzen von A und B der Gruppenmeinung entsprechen, wird nicht für den Layer gelernt). Andererseits sorgt die Addition auch dafür, dass sich der Einfluss von Merkmalsbewertungen automatisch reduziert, falls die Gruppe kein konstant bleibendes Urteil bezüglich des Merkmals fällt. Wenn sich A und B uneins über ein Merkmal sind, sich also manchmal die eine, manchmal die andere der beiden Präferenzen in der Gruppe durchsetzt, relativiert sich die Bewertung des betrachteten Merkmals im Layer des Personenpaars A und B . Ein Layer bezieht Position für die Vorlieben einer Person, widerspricht den Vorlieben beider Personen, oder reflektiert, dass ein Merkmal bezüglich der Gruppenmeinung nicht relevant ist (Merkmalsbewertung nahe der 0). Durch die Speicherung einer Merkmalsbewertung im Layer wird den Gruppenbewertungen Rechnung getragen.

2. Die Normierung der Gesamtbewertung eines Layers wird durchgeführt, um mehreren Layern den jeweils gleichen Stellenwert bei der Hypothesenbildung einzuräumen.
3. Layer müssen unter Umständen untereinander verglichen werden. Das heißt, es müssen Bewertungen für sämtliche Merkmale abfragbar sein. Genau wie für Präferenzen kann man sich verschiedener Methoden bedienen, um für jede Anfrage einen Wert zurückzuliefern. In Abschnitt 5.3 wird auf drei verschiedene Funktionen für diese Aufgabe eingegangen.

5.3. Interpolation auf bewerteten Merkmalen

Es kommt in mehreren Schritten des Lernsystems vor, dass Bewertungen für den gesamten Merkmalsraum zu Verfügung stehen müssen, obwohl nur eine geringe Anzahl bewerteter Merkmale vorliegt. Dieses Problem kann, abhängig von der Konfiguration des Lernsystems, eintreten, wenn Präferenzen oder Layer gelernt und verarbeitet, oder wenn Instanzen aufgrund einer Hypothese bewertet werden.

Eine Lösung des Problems wäre, für ein Merkmal eine neutrale Bewertung anzunehmen, solange noch nicht darüber geurteilt wurde. Besitzt man Hintergrundwissen über die Merkmale von Instanzen einer bestimmten Domäne, so kann man von bereits vorhandenen Merkmalsbewertungen auf bisher unbekannte Bewertungen schließen.

Bezüglich der Anwendung, die in Kapitel 6 vorgestellt wird, wissen wir, dass sich die Verwandtschaft zwischen den Merkmalen in ihrer geometrischen Nähe widerspiegelt. Deshalb werden neben einer Funktion, die neutrale Werte für noch unbewertete Merkmale zurückgibt, zusätzlich zwei Funktionen angegeben, die diese Beziehung zwischen den Merkmalen berücksichtigt. Alle drei Funktionen werden im Folgenden anhand einer Beispielmenge bereits bewerteter Merkmale, M_{Bsp} , erläutert.

Die Menge M_{Bsp} setzt sich zusammen aus Vektoren, die in den ersten beiden Elementen (x_{Bsp} und y_{Bsp}) ein Merkmal codieren und in der dritten Komponente (z_{Bsp}) die dazugehörige Bewertung bereithalten. Analog dazu seien (x', y') alle Elemente des Merkmalsraums, für die keine Bewertung in M_{Bsp} vorkommt, für die also gilt:

$$x', y' \in [0, 10] \wedge \begin{pmatrix} x' \\ y' \\ \cdot \end{pmatrix} \notin M_{Bsp}.$$

M_{Bsp} wird nun zur Menge M^* erweitert, die Bewertungen für den gesamten Merkmalsraum enthält. Das heißt, M^* hält für jedes Tupel $(x^*, y^*) \in [0, 10] \times [0, 10]$ eine Bewertung $z^* \in [-1, +1]$ bereit.

$$M_{Bsp} = \left\{ \begin{pmatrix} 1 \\ 3 \\ -0,125 \end{pmatrix}, \begin{pmatrix} 6 \\ 1 \\ -0,25 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \\ 0,5 \end{pmatrix}, \begin{pmatrix} 9 \\ 4 \\ 0,125 \end{pmatrix} \right\}$$

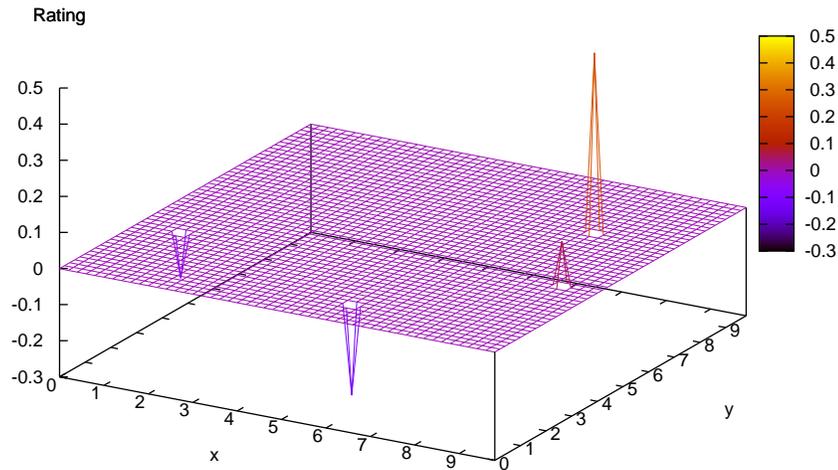


Abbildung 5.1.: Bisher unbewertete Merkmale werden als neutral, in diesem Fall 0 gewertet. Die Pyramidenform um die Extrema ergibt sich nur aufgrund des verwendeten Visualisierungswerkzeugs und entspricht nicht den tatsächlichen Werten.

5.3.1. „0“ als Standardannahme

Abbildung 5.1 zeigt die Menge $M_{B_{sp}}$ aller bewerteten Merkmale und 0 für alle Merkmale (x^*, y^*) , die nicht in $M_{B_{sp}}$ enthalten sind. Diese Funktion wird, im entsprechenden Kontext, oftmals nur mit „0“ referenziert. Die zurückgelieferten Werte lassen sich folgendermaßen beschreiben.

$$M^* = M_{B_{sp}} \cup \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix}$$

5.3.2. Voronoi Partitionierung

Hier wird für jedes Tupel (x^*, y^*) die Bewertung desjenigen Vektors $\vec{v}_i \in M_{B_{sp}}$ übernommen, der zu (x^*, y^*) am nächsten in der xy-Ebene liegt, also den geringsten euklidischen Abstand hat (siehe Abbildung 5.2). Diese Vorgehensweise partitioniert den Merkmalsraum bezüglich der xy-Ebene als Voronoi Diagramm. Entsprechend wird sie im Folgenden mit „Voronoi“ abgekürzt.

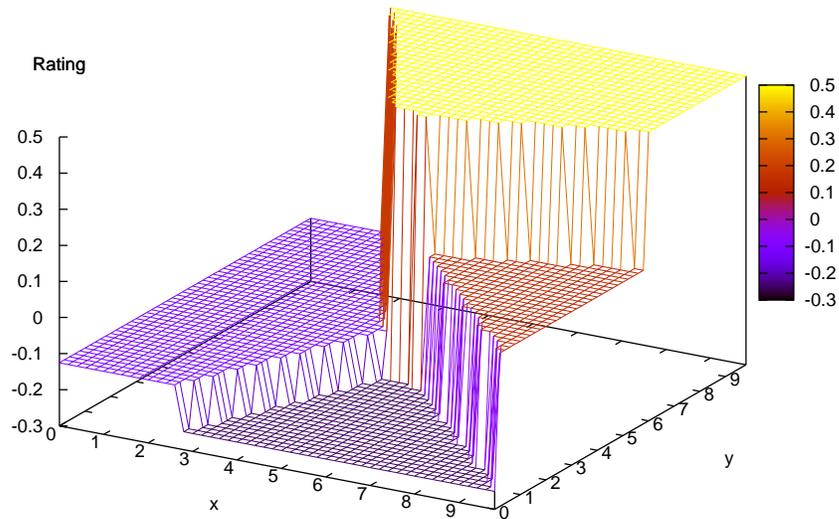


Abbildung 5.2.: Für bisher unbewertete Merkmale wird die Bewertung des nächsten Nachbarn angenommen.

Der euklidische Abstand in der xy -Ebene zwischen einem Tupel (x, y) und einem Vektor \vec{v} wird im Folgenden geschrieben als Funktion $euklid((x, y), \vec{v})$. Es ergibt sich:

$$M^* = M_{Bsp} \cup \left\{ \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \mid z' = z_k, \quad k = \min_i (euklid((x', y'), \vec{v}_i)) \right\}$$

5.3.3. Inverse Distance Weighting

Hier wird für jedes Tupel (x', y') ein Wert aus allen verfügbaren Vektoren bewerteter Merkmale, $\vec{v} \in M_{Bsp}$, errechnet. M_{Bsp} wird als Menge von Stützpunkten für eine Interpolation auf dem gesamten Merkmalsraum verstanden (siehe Abbildung 5.3). Dabei kommt jedem Vektor \vec{v}_i eine Bedeutung zu, die sich aus der Entfernung zum Anfragepunkt ergibt. Je geringer der euklidische Abstand in der xy -Ebene zwischen dem Tupel (x', y') und einem Vektor \vec{v}_i ist, desto größer ist der Einfluss der Bewertung des Vektors \vec{v}_i auf die interpolierte Bewertung am Anfragepunkt (x', y') . Falls bereits eine Bewertung für ein Merkmal gelernt wurde, wird diese zurückgegeben. In diesem Fall findet keine Interpolation statt. Die Menge M^* ergibt sich aus der Menge M_{Bsp} , der Größe n , wie folgt.

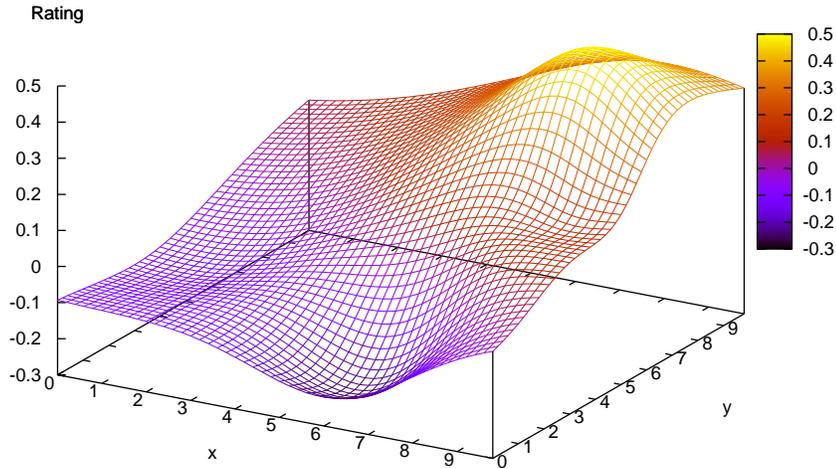


Abbildung 5.3.: Für bisher unbewertete Merkmale werden alle bekannten Bewertungen verrechnet. Je weiter ein Merkmal vom Anfragepunkt entfernt ist, desto geringer ist der Einfluss seiner Bewertung.

$$M^* = M_{Bsp} \cup \left\{ \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \mid z' = \frac{\sum_{i=1}^n \frac{z_i}{\text{euklid}((x',y'), \vec{v}_i)^\beta}}{\sum_{i=1}^n \frac{1}{\text{euklid}((x',y'), \vec{v}_i)^\beta}} \right\}$$

Die Interpolationsfunktion lässt sich wie folgt erklären. Es wird zunächst die Summe aller Bewertungen der Stützpunktvektoren \vec{v}_i , relativ zu deren Abständen zum Anfragepunkt (x', y') , gebildet. Der resultierende Wert ist normiert auf den gesamten Abstand aller Vektoren \vec{v}_i vom Anfragepunkt (x', y') . Durch die Abbildung des bisherigen Ergebnisses auf das Intervall der vorhandenen Bewertungen wird diese Normierung rückgängig gemacht. Man erhält einen invers zu den Abständen gewichteten Wert z' für den Anfragepunkt (x', y') .

β gewichtet den Einfluss der Stützpunkte. Im Beispiel (Abbildung 5.3) und auch in den kommenden Abwendungen dieser Interpolationsfunktion wurde $\beta = 2$ gewählt. Testläufe mit $\beta = 1$ haben ergeben, dass eine Veränderung dieses Parameters die Ergebnisse des Lernsystem kaum beeinflusst.

Die vorgestellte Methode wird als „Inverse Distance Weighting“ oder auch als „Inverse Distance Weighted Interpolation“ bezeichnet. „Inverse“ wird im Folgenden als Abkürzung dieser Funktion gewählt.

5.4. Hypothesenbildung

Es stellt sich die Frage, wie man Präferenzen π zusammen mit Layern λ verarbeitet, um Hypothesen für Gruppen zu bilden. In diesem Kapitel werden dafür zwei Wege aufgezeigt. Die Grundidee der „Gewichteten Übernahme der Präferenzen und Layer“ entspricht der in Kapitel 3 vorgestellten „Average Strategy“. Die „alternative Hypothesenbildung“ ähnelt am ehesten der erwähnten „Average Without Misery Strategy“.

5.4.1. Repräsentation von Hypothesen

Eine Hypothese ergibt sich durch das Verknüpfen von Präferenzen (π Funktionen) und Layern (λ Funktionen). Sie beschreibt ebenso eine Menge bewerteter Merkmale. Ob und wie auf den bewerteten Merkmalen interpoliert wird, ergibt sich aufgrund der Konfiguration des Lernsystems. Gleiches gilt für die Interpolationsmechanismen für π und λ . Innerhalb einer Konfiguration des Lernsystems wird stets die gleiche Interpolationsmethode verwendet, um die Anzahl der Konfigurationen des Lernsystems einzuschränken und die Wirkungsweisen der Interpolationsmechanismen klar voneinander zu trennen.

Eine Bewertung ist definiert als Element der Menge aller reellen Zahlen des Intervalls $[-1, +1]$. π und λ liefern entsprechende Werte zurück. Eine Hypothese, die ebenfalls als Menge bewerteter Merkmale anzusehen ist, muss dieser Einschränkung allerdings nicht unbedingt unterliegen. Falls ein Verarbeitungsschritt die Normalisierung der Hypothese voraussetzt, wird das Vorgehen im Detail erläutert.

5.4.2. +, * Operatoren

Die Unterscheidung der Verknüpfungen $+$ und $*$ wird in Kapitel 4 getroffen, um auf abstrakter Ebene die Chance einzuräumen, Präferenzen grundsätzlich anders zu verarbeiten als Layer. Im vorliegenden Fall bietet sich jedoch an, die beiden sehr ähnlichen Funktionen π und λ mit ein und demselben Operator, $+$, zu verknüpfen. Für die verfolgten Strategien der Hypothesenbildung wurden zwei Varianten des $+$ Operators (für π und λ) implementiert.

Direkte Übernahme der Bewertungen mit $+_1$

Seien B_1 und B_2 Mengen bewerteter Merkmale, so resultiert durch die Verknüpfung der beiden Mengen mit dem Operator $+_1$ eine weitere Menge bewerteter Merkmale, B_{Res} , die sämtliche bewertete Merkmale aus B_1 und B_2 übernimmt. Warten B_1 und B_2 mit Bewertungen für die gleichen Merkmale auf, werden beide Bewertungen addiert und zu B_{Res} hinzugefügt. Die Verknüpfung zieht nur diejenigen Merkmale aus den Operandenmengen in Betracht, die explizit bekannt sind, das heißt ohne Zuhilfenahme von Interpolationsmechanismen feststehen.

Vor der Übernahme beziehungsweise Addition der Merkmalsbewertungen werden B_1 und B_2 normalisiert. Für beide Mengen wird jeweils die Summe aller Bewertungen, die Gesamtbewertung, berechnet. Division der einzelnen Bewertungen durch die Gesamtbewertung normiert die Gesamtbewertung von den Mengen B_1 und B_2 jeweils auf den Wert 1. Eine neutrale Bewertung, 0, wird nicht verändert, bei einer Gesamtbewertung von 0 findet entsprechend auch keine Division statt. Diese Normalisierung gewährleistet, dass beiden Operanden gleichwertig sind.

Addition der Bewertungen wirkt wiederum, genau wie beim Lernen der Präferenzen und Layer (siehe Abschnitt 5.2.4 und folgende), als ausgleichend für widersprüchliche Werte und verstärkend für Bewertungen, die in die gleiche Richtung (positiv oder negativ) weisen.

Gemeinsames übernehmen, sich Widersprechendes ausschließen mit $+_2$

Ein guter Kompromiss für eine Gruppe könnte herbeigeführt werden, indem nur vorge schlagen wird, was alle Mitglieder der Gruppe für positiv erachten. Eine entsprechende Verknüpfung lässt sich erreichen, indem sich diejenigen positiven Merkmalsbewertungen aus B_1 und B_2 in B_{Res} niederschlagen, über die Einigkeit herrscht.

Gleichzeitig sollten negative Bewertungen immer berücksichtigt werden. Schließlich sollte ein Kompromiss vermeiden, dass für ein Publikum Vorschläge unterbreitet werden, die einzelnen Personen zusetzen würden.

Vereint man die beiden Ziele, so erhält man alle Merkmalsbewertungen, die sich nicht widersprechen (wiederum festgestellt über das Vorzeichen) und außerdem auch negative Merkmalsbewertungen, die vielleicht von einer der beiden Personen als positiv erachtet werden.

Wie schon für den Operator $+_1$ wird zunächst eine Normalisierung der Mengen B_1 und B_2 erwirkt, die ihre jeweilige Gesamtbewertung auf 1 reduziert. Merkmalsbewertungen, über die Einigkeit herrscht, werden durch Addition verstärkt.

5.4.3. ◦ Operator

Der in Abschnitt 5.4.2 begonnene Weg wird fortgesetzt, indem nun analog zu den beiden Varianten des $+$ Operators zwei \circ Verknüpfungen zur Hypothesenbildung vorgestellt werden.

Entsprechend der Gleichungen in Kapitel 4.4 verrechnet man die Präferenzen und Layer der anwesenden Personen zu Π_G und Λ_G . Der als nächstes erläuterte Ansatz zur Hypothesenbildung erwartet mittels $+_1$ gebildete Gesamtpräferenzen und Gesamtlayer. Der zweite Ansatz dagegen basiert auf mittels $+_2$ erzeugten Π_G und Λ_G .

Gewichtete Übernahme der Präferenzen und Layer mit \circ_1

Der direkte Weg, um aus Π_G und Λ_G eine Hypothese zu generieren und dabei noch der Sonderstellung der des Gesamtlayers Rechnung zu tragen, ist beide Operanden unterschiedlich gewichtet zu übernehmen. Die Bewertungen der Operandenmengen werden mit einem Faktor multipliziert, anschließend werden die Operanden genau wie mit dem $+_1$ Operator zu einer Hypothese verschmolzen.

Es entspricht einem einheitlichen Design, bereits die Operanden Π_G und Λ_G im Vorfeld mittels des $+_1$ Operators zu berechnen.

Da die Layer eingeführt wurden, um individuelle Präferenzen bei Anwesenheit anderer zu überschatten, wird der Gesamtlayer Λ_G stärker gewichtet. In den Experimenten aus Kapitel 6 wird Λ_G dreifach stärker gewichtet als Π_G . Im Weiteren wird für diese Art der Hypothesenbildung das Kürzel „gew“, für „gewichtet“, verwendet.

Alternative Hypothesenbildung mit \circ_2

Wenn man die Layer als Constraints versteht, die die Hypothese H_G erfüllen muss, ergibt sich eine gänzlich andere Sichtweise auf das Problem der Hypothesenbildung.

Vor der Anwendung des Operators \circ_2 verbinde man jeweils die Gesamtpräferenzen und den Gesamtlayer mittels der oben vorgestellten Verknüpfung $+_2$. Die Vektoren des resultierenden Λ_G werden als Menge von Einzelbedingungen betrachtet, die H_G erfüllen muss. Grundlage für H_G bilden die Gruppenpräferenzen Π_G .

Übernimmt man Π_G eins zu eins in H_G . Dann wäre es ein wünschenswertes Ziel, H_G minimal so zu verändern, dass es die Bedingungen Λ_G erfüllt. Zu ermitteln, wie diese Transformation im Einzelnen aussähe, sprengt den Rahmen dieser Arbeit. Alternativ kann H_G angepasst werden, indem man Vielfache der bewerteten Merkmale aus Λ_G zu H_G hinzunimmt.

Für alle Merkmalsbewertungen, die explizit in Λ_G gespeichert sind, wird festgestellt, ob sie mit den (eventuell interpolierten) Bewertungen von Π_G in Widerspruch stehen. Sofern ein Widerspruch vorliegt, wird berechnet, wie oft das bewertete Merkmal aus Λ_G zu Π_G hinzugenommen werden müsste, damit sich der Widerspruch auflöst. Die Bewertung des Gesamtlayers wird nun um dieses ermittelte Vielfache, γ , multipliziert und zur Hypothese H_G hinzugenommen.

Umfasst Λ_G zum Beispiel nur einen Vektor, \vec{v}_Λ , wird die Bewertung z_Π des Merkmals (x_Λ, y_Λ) in Π_G nachgesehen. Falls sich die Bewertungen z_Π und z_Λ widersprechen, wird berechnet, wie oft Π_G um \vec{v}_Λ erweitert werden muss, damit Π_G schließlich die von Λ_G erwartete Bewertung für das Merkmal (x_Λ, y_Λ) ausreichend approximiert. Die resultierende Hypothese H_G würde sich aus Π_G unter der γ -fachen Hinzunahme von \vec{v}_Λ ergeben.

Als Widerspruch zwischen Merkmalsbewertungen von Π_G und Λ_G könnte man das gleiche Kriterium ansetzen, wie für Widersprüche zwischen individuellen und Gruppenbewertungen. Um sicherzugehen, dass die Constraints aus Λ_G sich auch tatsächlich in

der formulierten Hypothese widerspiegeln, wird Widerspruch in diesem Zusammenhang anders definiert.

Es existiert ein Widerspruch, falls die Differenz der Bewertungen aus Π_G und Λ_G für ein Merkmal (x, y) größer ist als eine sehr kleine, vorgegebene Konstante ϵ (für die Berechnungen in Kapitel 6 wird $\epsilon = 0.001$ verwendet).

Der Faktor γ , um den der Vektor \vec{v}_Λ zu Π_G hinzugenommen werden muss, wird durch Intervallschachtelung bestimmt. γ wird zuerst im Intervall von $i_a = 0$ bis $i_b = 2$ gesucht. Ist i_b kein ausreichend großer Wert - d.h. die i_b -fache Erweiterung der Menge Π_G um den Vektor \vec{v}_Λ erfüllt noch nicht die durch Λ_G gestellte Bedingung -, so wird der Suchbereich verändert. Die ehemals obere Grenze des Suchintervalls (i_b) wird als unterer Ausgangspunkt des neuen Intervalls verwendet ($i'_a = i_b$), gleichzeitig ergibt sich die neue obere Grenze des Suchintervalls aus dem Quadrat der vorherigen oberen Grenze: $i'_b = i_b^2$.

Auf diese Weise wächst das Suchintervall exponentiell schnell, um sich dem Faktor γ anzunähern. Gleichzeitig wird der Bereich, in dem γ nicht vorkommt von der Suche ausgeschlossen.

Steht das Suchintervall schließlich fest (i_a reicht für γ noch nicht aus, i_b schon), wird darin mittels binärer Suche der minimale Wert für γ gesucht.

Zusammenfassend resultieren die erwähnten Punkte in einer Hypothese H_G , die alle Präferenzen Π_G in sich vereint und außerdem sicherstellt, dass sämtliche Bedingungen (gespeichert in Λ_G) erfüllt werden. Dabei wird darauf geachtet, dass die Verknüpfung von Präferenzen bzw. Layern mit dem $+_2$ Operator durchgeführt wird, also alle negativen Bewertungen, sowie die Schnittmenge positiver Bewertungen des Publikums übernommen werden.

Zur Unterscheidung mit der im vorhergehenden Abschnitt vorgestellten Methode zur Hypothesenbildung „gew“, wird diese wesentlich komplexere Vorgehensweise mit „alt“, für „alternative Methode zur Hypothesenerzeugung“, abgekürzt.

5.5. Das Abstandsmaß $D(H, i)$

Die Aufgabe der Funktion $D(H, i)$ ist es, eine zur Auswahl stehende Instanz i auf der Grundlage einer Hypothese H zu bewerten. Man erhält eine *Fitness* einer Instanz, die umso größer ist, je weniger sich i und H voneinander unterscheiden. Mit Hilfe der Fitness einer Instanz lässt sich eine partielle Ordnung auf der Menge zur Auswahl stehender Instanzen definieren. Für das Publikum können dann die k Instanzen mit den höchsten Fitnesswerten vorgeschlagen werden.

Eine Hypothese umfasst Merkmale (x_H, y_H) und deren Bewertungen z_H . Da eine Hypothese Endprodukt des Verarbeitungsprozesses von Präferenzen und Layern ist, muss sie bezüglich ihrer Bewertungen nicht mehr normalisiert werden. Bewertungen können daher beliebige Werte aus der Menge der reellen Zahlen annehmen.

Eine Instanz wird im vorliegenden Fall durch eine Menge von Merkmalen (x_i, y_i) und deren Häufigkeiten z_i beschrieben. Die Häufigkeit eines Merkmals kann als Maß

verstanden werden, zu welchem Anteil das Merkmal die jeweilige Instanz beschreibt. Es ergeben sich für z_i Werte aus dem Intervall $[0, 1]$.

Es folgen verschiedene Implementierungen des Abstandsmaßes $D(H, i)$ für die beschriebenen Repräsentationen von H und i . Für einfache Referenzierung werden in den Überschriften der Abschnitte Abkürzungen für die Verfahren angegeben.

Um die verschiedenen Ansätze zu veranschaulichen, wird einerseits eine Hypothese H_{Bsp} und andererseits eine Instanzbeschreibung i_{Bsp} aus den Testläufen des kommenden Kapitels herangezogen. H_{Bsp} und i_{Bsp} werden wegen der besseren Übersicht als Matrizen dargestellt, die ersten beiden Spalten beinhalten das Merkmal (x, y) , die dritte Spalte enthält für i_{Bsp} die relativen Häufigkeiten des jeweiligen Merkmals, im Fall von H_{Bsp} die Bewertungen des Merkmals. Die aufgezeigten Werte sind die explizit gelernten bzw. festgestellten Werte. Für Anfragen über Merkmale, die nicht in den Matrizen aufgelistet sind, werden die in Abschnitt 5.3 vorgestellten Interpolationsmethoden verwendet. Für die visualisierten Beispiele wird bei Bedarf mittels Inverse Distance Weighting interpoliert. In Abbildung 5.4 werden i_{Bsp} und H_{Bsp} graphisch dargestellt.

$$i_{Bsp} = \begin{pmatrix} 4.4 & 3.1 & 0.195652 \\ 4.9 & 3.4 & 0.130435 \\ 4.2 & 2.7 & 0.0434783 \\ 6 & 2.3 & 0.23913 \\ 5.9 & 5.5 & 0.0434783 \\ 5.9 & 4.7 & 0.0434783 \\ 3.8 & 2.8 & 0.0434783 \\ 4.7 & 5.4 & 0.130435 \\ 5.3 & 4.3 & 0.0869565 \\ 4.3 & 5.2 & 0.0434783 \end{pmatrix} \quad H_{Bsp} = \begin{pmatrix} 5 & 5.1 & -0.0114635 \\ 4.9 & 5 & 0.0147388 \\ 5.9 & 2.9 & 0.496639 \\ 4.7 & 5.4 & 0.0406826 \\ 6 & 2.3 & 0.0275814 \\ 3.1 & 3.8 & 0.205827 \\ 4.4 & 3.1 & -0.0692984 \\ 2.4 & 3.2 & 0.0289605 \\ 5.3 & 4.3 & -0.10481 \end{pmatrix}$$

5.5.1. Gesamtabstand („Abst“)

Die Grundidee dieser Funktion ist es, die Differenz zwischen Frequenz und Bewertung für jedes Merkmal aufzusummieren, über das handfeste Informationen zur Verfügung stehen, das also mindestens in einem der beiden Operanden vorkommt. Dabei kann man den Abstand einer Merkmalshäufigkeit zu einer negativen Bewertung als positiv und umgekehrt, den Abstand zu einer positiven Bewertung als negativ verstehen. Die Hypothese H wird zunächst in zwei Untermengen aufgeteilt: H_{pos} umfasst alle positiven, H_{neg} alle negativen Bewertungen aus H .

Damit die Instanz i in jedem Merkmal mit H_{pos} und H_{neg} verglichen werden kann, müssen beide Hypothesenuntermengen bezüglich ihrer Bewertungen auf den Betrag 1 normiert werden. Außerdem muss man alle Bewertungen in H_{neg} mit positiven Vorzeichen versehen, die Frequenzen der Merkmale in i kommen schließlich nur im Intervall $[0, 1]$ vor.

Alle Abstände zwischen Merkmalen aus i und den interpolierten Werten aus H_{neg} , sowie die Abstände von allen Merkmalen aus H_{neg} und den interpolierten Werten aus i

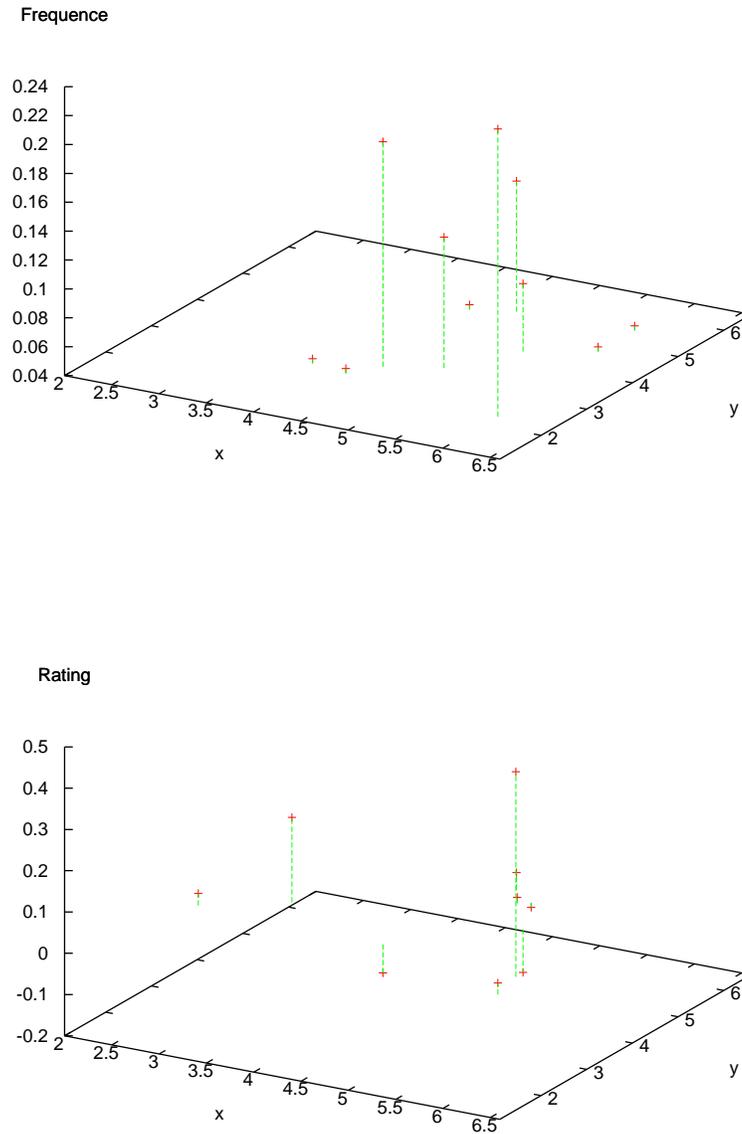


Abbildung 5.4.: Oben: Merkmale und deren Häufigkeiten der Instanz i_{Bsp} . Unten: Merkmale und deren Bewertungen der Hypothese H_{Bsp} .

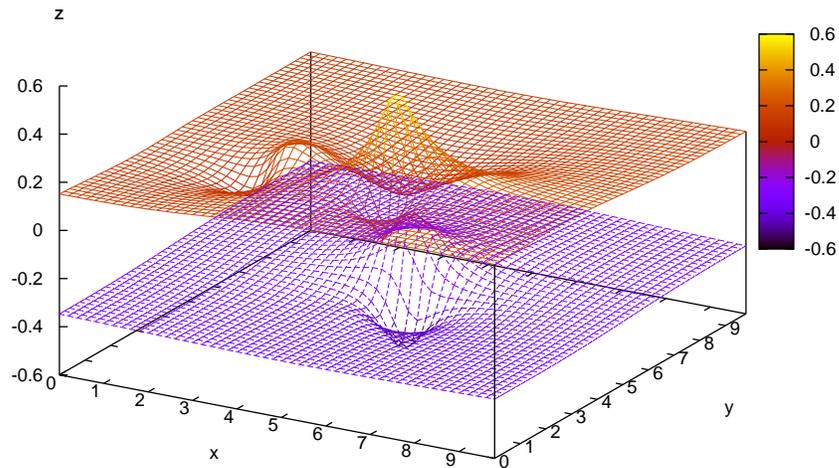


Abbildung 5.5.: Die interpolierten Untermengen H_{pos} und H_{neg} der Hypothese H_{Bsp} .

werden aufsummiert. Von dieser Summe werden alle Abstände zwischen Merkmalen aus i und den interpolierten Werten aus H_{pos} , sowie die Abstände von allen Merkmalen aus H_{pos} und den interpolierten Werten aus i abgezogen.

Je größer das Ergebnis, desto näher sind die Werte der Instanz i an den positiven Bewertungen aus H und desto weiter entfernt sind sie von den negativen Bewertungen aus H .

Abbildung 5.5 zeigt die beiden Untermengen H_{pos} und H_{neg} der Hypothese H_{Bsp} . In Abbildung 5.6 sieht man die punktuellen Differenzen zwischen Instanz i_{Bsp} und Hypothese H_{Bsp} . Die Fitness wäre im vorliegenden Fall 1.17518.

5.5.2. Anteilsmäßige Bewertung der Merkmale (“Gewi”)

Ohne Vorverarbeitung der Hypothese oder der Instanz werden für jedes bekannte Merkmal aus H die Bewertung mit der interpolierten Frequenz aus i , und umgekehrt für jedes bekannte Merkmal aus i die Frequenz mit der interpolierten Bewertung aus H , multipliziert. Je größer die Summe dieser Produkte ist, desto besser: Positiv bewertete Merkmale mit einem hohen Anteil an der Instanzbeschreibung treiben das Ergebnis in die Höhe, negativ bewertete Merkmale mit großem Häufigkeitswert hingegen schmälern die Summe, können sie ins Negative gleiten lassen.

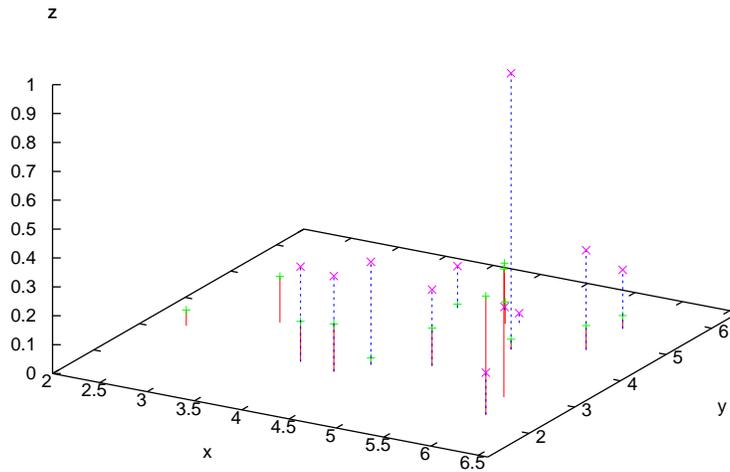


Abbildung 5.6.: Die Abbildung zeigt die punktuellen Differenzen zwischen der Instanz i_{Bsp} und der Hypothese H_{Bsp} . Die gestrichelten blauen Linien mit rosanem „x“ kennzeichnen die Abstände zu negativen Hypothesenbewertungen. Die orangenen, durchgehenden Linien mit grünem „+“ kennzeichnen die Abstände zu positiven Werten der Hypothese.

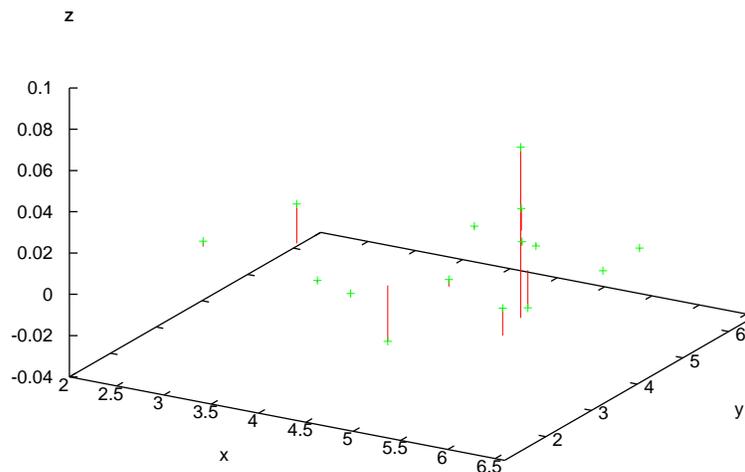


Abbildung 5.7.: Die Abbildung zeigt die Produkte der Merkmalsbewertungen aus H_{Bsp} und den Merkmalshäufigkeiten aus i_{Bsp} .

Abbildung 5.7 zeigt die Produkte der Merkmalsbewertungen aus H_{Bsp} und den Merkmalshäufigkeiten aus i_{Bsp} . Der ermittelte Fitnesswert beträgt in diesem Fall 0.0885947.

5.5.3. Kombination von „Abst“ und „Gewi“ („Komb“)

„Abst“ errechnet die Fitness durch einen Vergleich der Ebenen auf Grundlage der vorhandenen Stützpunkte. Negativ zu bewerten ist, dass „Abst“ Häufigkeiten mit Bewertungen vergleicht. Zwar sind die Wertebereiche aneinander angepasst und es ist auch sinnvoll, da eine starke Bewertung eines Merkmals auch ein starkes Vorkommen des Merkmals erwarten lässt. Auf der anderen Seite ist es jedoch ein Verlust, dass sich die konkreten Merkmalsbewertungen nicht direkt im Fitnesswert einer Instanz niederschlagen.

Bei „Gewi“ resultiert das Ergebnis aus den Bewertungen und den jeweiligen Frequenzen (und umgekehrt). Allerdings fehlt bei dieser Methode der direkte Vergleich der beiden Mengen.

Die Kombination der beiden Methoden durch einfache Addition der beiden Fitnesswerte liegt deshalb nahe. Der Fitnesswert für i_{Bsp} liegt entsprechend bei 1.26378.

5.5.4. Differenz der Durchschnittsvektoren (“AvgV”)

In der hier vorgeschlagenen Methode werden charakteristische Vektoren für die Instanz i und für die Hypothese H berechnet. Das Maß an Ähnlichkeit wird durch den Abstand zwischen Vektoren ermittelt.

Die z -Komponente der Durchschnittsvektoren enthält einerseits einen gemittelten Wert für Merkmalshäufigkeiten (für i), andererseits den Durchschnitt von Bewertungen (für H). Der Vergleich von Vektoren birgt daher indirekt den einen Vergleich von Häufigkeiten und Bewertungen. Es wird deshalb ähnlich verfahren wie bei „Abst“. Für die Untermengen H_{pos} und H_{neg} der Hypothese H werden die Durchschnittsvektoren $\vec{\phi}_{pos}$ und $\vec{\phi}_{neg}$ gebildet. Außerdem wird der Durchschnittsvektor $\vec{\phi}_i$ für die Instanz i ermittelt.

Alle Vektoren werden auf die Länge 1 normiert. Für das Ergebnis wird nun der Betrag der Differenz zwischen $\vec{\phi}_{pos}$ und $\vec{\phi}_i$ vom Betrag der Differenz zwischen $\vec{\phi}_{neg}$ und $\vec{\phi}_i$ abgezogen.

Mit der vorgestellten Funktion erhält i_{Bsp} die Fitness 0.00631208. Abbildung 5.8 zeigt die jeweils berechneten Vektoren, $\vec{\phi}_i$ wird rot, $\vec{\phi}_{pos}$ grün und $\vec{\phi}_{neg}$ blau dargestellt. Die Spiegelung von $\vec{\phi}_{neg}$ an der 0-Ebene wird durch den gebogenen Pfeil symbolisiert.

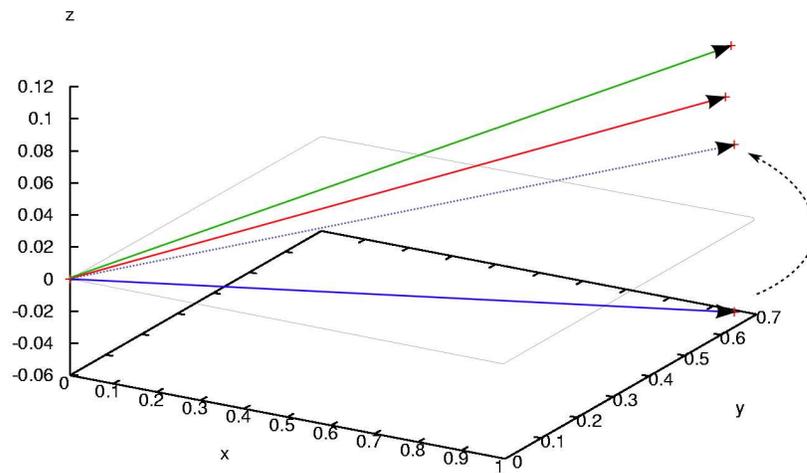


Abbildung 5.8.: Der Durchschnittsvektor der Instanz i_{Bsp} ist rot, der Durchschnittsvektor der Menge H_{pos} grün und der Durchschnittsvektor der Menge H_{neg} blau dargestellt. Die Spiegelung des Durchschnittsvektors H_{neg} an der 0-Ebene wird durch den gebogenen Pfeil symbolisiert.

6. Anwendung und Evaluierung

6.1. Ziel der Anwendung

Ziel der Anwendung ist es, einer Gruppe von Personen eine bestimmte Anzahl an Fernsehsendungen vorzuschlagen. Die Vorschläge sollen aufgrund von erlernten Präferenzen der Anwender und aufgrund von erlernten Layern zwischen Personenpaaren generiert werden (siehe Kapitel 4.3). Sofern noch keine Präferenzen des Publikums bekannt sind, gibt es keine Titel, die bevorzugt vorgeschlagen werden können. Trainingsbeispiele für Präferenzen und Layer werden durch die Bewertung von Sendungen von dem anwesenden Publikum zur Verfügung gestellt.

6.2. Eingangsdaten und deren Vorverarbeitung

Moderne Fernsehsysteme unterstützen die Verwaltung und Darstellung von Metadaten über die zur Auswahl stehenden Sendungen. Unter anderem wird eine sogenannte „Extended Info“ bereitgestellt, die den Inhalt der Sendung mit einigen Worten beschreiben soll. Im folgenden wird eine Methode vorgeschlagen, um die „Extended Info“ in Zahlen zu überführen, die emotionale Komponenten und ihre Quantität repräsentieren. Damit wird einerseits versucht, die Vielseitigkeit der Inhalte einzufangen, und andererseits, eine emotionale Beschreibung der Titel zu ermitteln.

6.2.1. Ein Film in Form emotionaler Komponenten

Die Übertragung der Filmbeschreibung in eine numerische Repräsentation emotionaler Komponenten, wird in zwei Schritten vorbereitet, die in Abbildung 6.1 zu sehen und im Folgenden erklärt sind.

- In [3] werden verschiedene Ansätze zusammengefasst, wie man Gefühle quantitativ erfassen kann. Unter anderem wird dabei aus [24] eine Liste Gefühle umschreibender Wörter und assoziierter Activation/Evaluation Koordinaten zitiert (siehe Tabelle A.1 im Anhang).

Für die systematisch übersetzten Wörter¹ werden sogenannte *Dornseiffgruppen* ermittelt[6]. Eine Dornseiffgruppe gruppiert Wörter, die thematisch zusammengehören.

Durchläuft man diesen Vorgang rückwärts, so erhält man bei Eingabe einer Dornseiffgruppe die dazugehörenden Activation/Evaluation Koordinaten (wobei es sich hierbei um mehrere Koordinatenpaare handeln kann).

- Jedes Wort (jedes Token t_1 bis t_n) der Beschreibung einer Fernsehsendung (*Extended Info*), wird mittels einer Morphologiekomponente in sämtliche Grundformen übersetzt. Für diese Grundformen werden wiederum die Dornseiffgruppen ermittelt (siehe Algorithmus 6.1). Fasst man diesen Vorgang zusammen, so wird die Filmbeschreibung auf eine Multimenge an Dornseiffgruppen abgebildet.

Durch die Verknüpfung beider Schritte kann man einen beliebigen Text, im vorliegenden Fall die Beschreibungen von Fernsehsendungen, in eine Multimenge an Activation/Evaluation Koordinatenpaaren überführen.

Es gibt zwei Gründe für die Wahl der verwendeten Liste emotionaler Adjektive und assoziierter Activation/Evaluation Koordinaten.

Zum einen postuliert die Zuordnung von Emotionen zu Activation/Evaluation Koordinaten, dass eine örtliche Beziehung zwischen den Gefühlen numerisch widergespiegelt wird. Die geometrische Nähe verwandter Gefühle kann man sich zunutze machen, wenn man annimmt, dass sich auch die Präferenzen bezüglich zweier Gefühlskomponenten in deren Abstand widerspiegelt.

Zum anderen umfasst die vorliegende Liste sehr viele Wörter und entsprechende Koordinatenpaare. Dies ist wichtig, wenn man für ein breites Spektrum an Wörtern, die in den Beschreibungen für Fernsehsendungen vorkommen können, Paare von Activation/Evaluation Koordinaten zur erhalten möchte.

6.2.2. Aufbereitung und Auswahl der Beispieldaten

Programmbeispiele wurden von einer Datenbank bezogen, die aktuelle, über ein Satellitensystem empfangene, Fernsehinformationen verwaltet. Von insgesamt 3308 Einträgen wurden zunächst 1677 gefunden, für die Dornseiffgruppen angegeben werden können, für den Rest lagen keine Zusatzinformationen in Form eines „Extended Infos“ vor. Von diesen wiederum ließen sich für 1262 Programme Activation/Evaluation Diagramme mit mindestens einem Eintrag erstellen. Dem Titel nach mehrfach vorkommende Programmbeispiele wurden eliminiert und es blieben schließlich Informationen über 415 Sendungen erhalten (siehe Tabelle A.2 im Anhang).

¹Die Wörter wurden mittels des Onlinewörterbuches der Informatik der Technischen Universität München übersetzt, wobei sämtliche Ergebnisse eins zu eins übernommen wurden[4]. Für die Wörter „loathful“ und „unaffectionate“ waren keine direkten Übersetzungen vorhanden, weswegen Übersetzungen der bei dem Webservice „The Free Dictionary, by Farlex“ gefundenen Synonyme übernommen wurden[7].

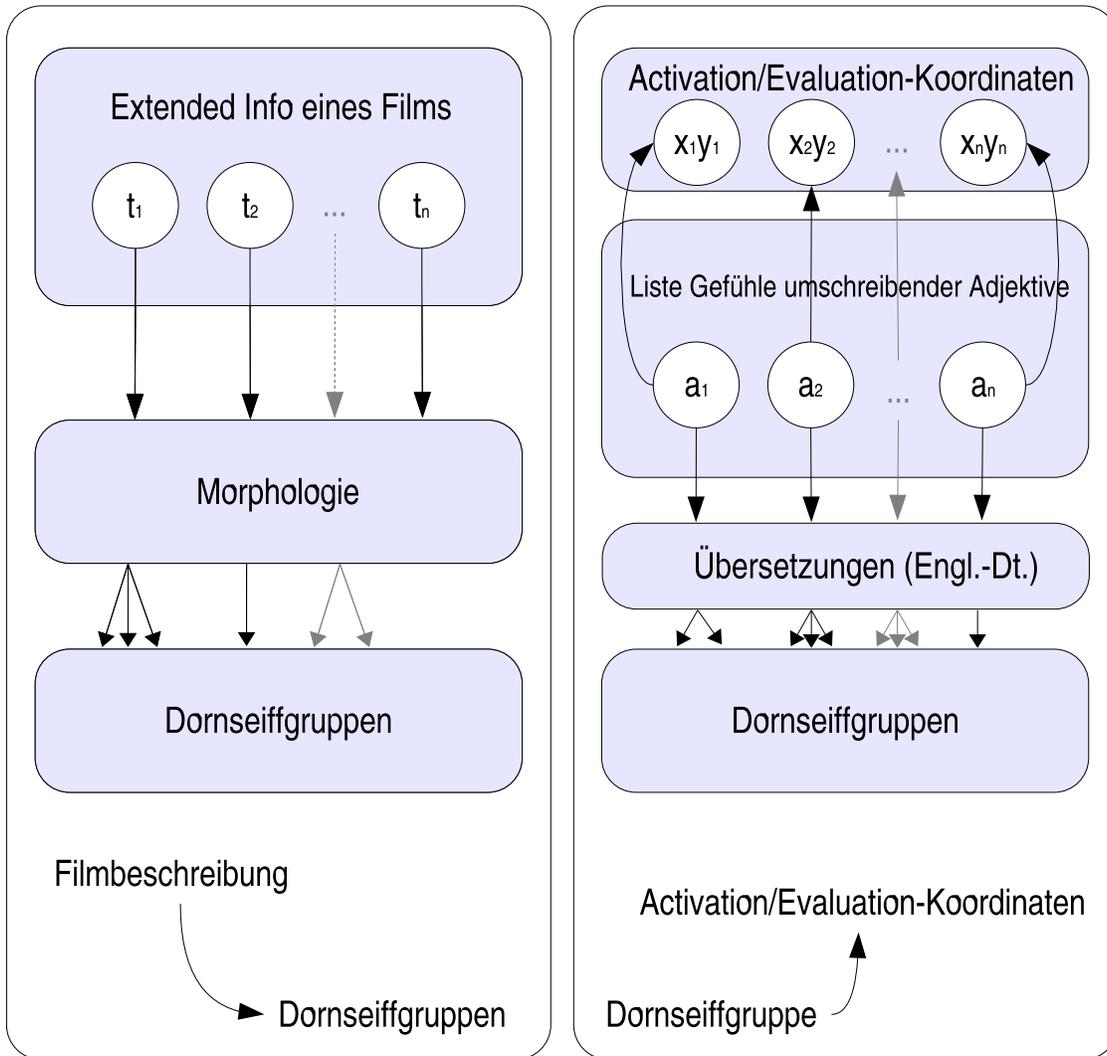


Abbildung 6.1.: Zwei Vorbereitungsschritte um die Extended Info in Activation/Evaluation Koordinaten zu übersetzen.

Algorithm 6.1 Algorithmus zur Berechnung der Dornseiffgruppen

Gegeben:*ExtendedInfo*: Eine textuelle Filmbeschreibung*getBaseForms(t)*: Ermittelt die morphologischen Grundformen des Tokens *t**DSG(g)*: Ermittelt die Dornseiffgruppen eines Wortes *g***Rückgabe:** Dornseiffgruppen in *ExtendedInfo*

```
1:  $DSG(ExtendedInfo) \leftarrow \{\}$ 
2: for all Token  $t_i \in ExtendedInfo$  do
3:    $DSG(t_i) \leftarrow \{\}$ 
4:    $G \leftarrow getBaseForms(t_i)$ 
5:   for all  $g \in G$ : do
6:      $DSG(t_i) \leftarrow DSG(t_i) \cup DSG(g)$ 
7:   end for
8:  $DSG(ExtendedInfo) \leftarrow DSG(ExtendedInfo) \cup DSG(t_i)$ 
9: end for
```

Für Testzwecke wurden pauschal diejenigen zehn Sendungen betrachtet, deren Activation/Evaluation Diagramme am meisten Speicherplatz belegen, was als Indikator für ein reichhaltiges Informationsvorkommen gewertet wurde.

Von diesen zehn Beispielen werden vier Stück vorgestellt. Zu den einzelnen Sendungen sind die ermittelten Dornseiffgruppen, deren Vorkommen und daraus abgeleitete Activation/Evaluation Koordinaten (*a* und *e*) samt normalisierter Häufigkeit (*f*) tabellarisch aufgelistet. Um die verschiedenen Sendungen besser miteinander vergleichen zu können, und auch um die von dem implementierten Lernsystem generierten Daten leichter nachvollziehen zu können, werden jeweils die Koordinaten *a* und *e* und die normalisierte Häufigkeit *f* in einem Diagramm eingetragen.

Die ausgewählten Sendungen sind „Polizeiruf 110: Um Kopf und Kragen“ (siehe Tabelle 6.1 und Abbildung 6.2), „Sesamstraße“ (siehe Tabelle 6.2 und Abbildung 6.3), „Bianca - Wege zum Glück“ (siehe Tabelle 6.3 und Abbildung 6.4), „aspekte kompakt“ (siehe Tabelle 6.4 und Abbildung 6.5).

Beispielhaft für die Sendungen, von denen nur wenige Activation/Evaluation Informationen ergattert werden können, sind „Bayern 3 - die MorningShow“ (siehe Tabelle 6.5 und Abbildung 6.6) und „88acht GUTEN TAG BERLIN“ (siehe Tabelle 6.7 und Abbildung 6.7).

An den wenigen Wörtern der „Extended Info“ von „Bayern 3 - die MorningShow“ kann man sehr schnell ersehen, wie die Dornseiffgruppen für diese Sendung zustande gekommen sind. Tabelle 6.6 entlarvt die verschiedenen ausschlaggebenden Wörter und die Bedeutungen der Dornseiffgruppen.

Extended Info		
Diesmal wird es brenzlig für die Kriminalhauptkommissare Obermaier und Tauber: Mord in den eigenen Reihen. Die junge Streifenpolizistin Mary wird erschossen in ihrem Auto gefunden. Sie war in ihrer Dienststelle nicht besonders beliebt.		
Dornseiffgruppen und deren Häufigkeiten (falls größer 1)		
11.56, 20.3 (2), 3.3 (4), 20.7, 19.24, 16.15, 16.21, 7.40 (2), 21.31, 4.17, 21.27, 4.3 (3), 21.34, 21.9, 4.33 (2), 4.28, 4.34, 4.36 (2), 4.50 (2), 4.51, 4.47, 10.16, 4.48, 5.1 (2), 10.52, 10.55, 18.10 (2), 9.23, 6.11 (3), 9.29, 9.40, 6.20 (2), 15.10, 6.31, 15.23, 15.33, 3.18 (3), 9.72, 11.8 (2), 15.62, 5.14, 5.20, 5.31 (2), 5.26 (2), 5.29, 2.14, 2.15 (2), 2.22 (2), 11.18, 11.23, 2.40, 2.41 (2), 11.33		
<i>a, e</i>	<i>f</i>	Übersetzungen
5.3, 4.8	0.0454545	fröhlich, schadenfroh
4.7, 5.4	0.272727	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan
4.3, 5.2	0.0454545	angenehm, annehmlich, einverstanden, gefällig, liebenswürdig
5.2, 5.5	0.0454545	ekstatisch, verzückt
5.1, 4.8	0.0909091	begeistert, enthusiastisch, passioniert, schwärmerisch
5.9, 2.9	0.0909091	aggressiv, angreifend, angriffslustig, ausfällig, draufgängerisch, dynamisch, energisch, feindlich, kampfbetont, offensiv, streitlustig, streitsüchtig
4.9, 3.4	0.0454545	ängstlich, bange, besorgt, fürchtend, erfolgshungrig, gefräßig, gierig, habgierig
6, 2.3	0.136364	ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt
5.3, 4.3	0.0454545	achtsam, aufmerksam
4.2, 5.9	0.0909091	abenteuerlich, gewagt, unternehmenslustig, unternehmungslustig
3, 4.3	0.0909091	apathisch, teilnahmslos, bescheiden, demütig, gütig, kleinlaut, mild, sanft, zahm

Tabelle 6.1.: Informationen zur Sendung „Polizeiruf 110: Um Kopf und Kragen“.

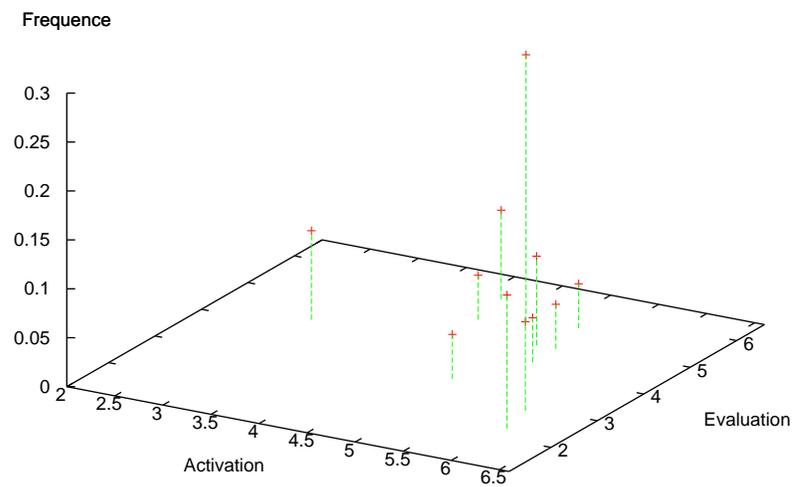


Abbildung 6.2.: Activation/Evaluation Diagramm zur Sendung „Polizeiruf 110: Um Kopf und Kragen“.

Extended Info

Auf Finchens Laubhaufen landet ein trauriger, müder Schmetterling. Gerne würde Finchen sich mit ihm anfreunden, aber zuallererst muss er aufgemuntert werden. Von Ella erfährt Finchen, dass Schmetterlinge Blüten gern haben, und schon geht sie daran, ihren Gast mit möglichst vielen Blumen glücklich zu machen. Auch die Plastinots betätigen sich als Gärtner und geben sich viel Mühe mit einer Distel. Ernie, Bert und Lulatsch treten im Schultheater auf und führen vor, wie ein Schmetterling entsteht. Ein trauriges Meerschweinchen sucht das Meer.

Dornseiffgruppen und deren Häufigkeiten (falls größer 1)

11.46, 11.51, 11.47, 15.1, 20.1 (4), 15.3, 20.3 (2), 3.3 (3), 16.20, 7.27 (3), 21.24, 4.20 (2), 4.22 (2), 4.3 (2), 16.4 (2), 7.50 (3), 21.9 (4), 4.33 (7), 4.28 (8), 4.35, 4.37 (3), 1.19 (2), 4.50 (2), 10.13 (3), 4.51 (2), 10.14 (3), 10.16, 10.25 (5), 10.31 (3), 10.32 (3), 5.2 (2), 10.35, 5.6, 5.8, 5.9, 9.22, 10.9, 9.19 (2), 6.10 (2), 9.27, 6.11, 6.12, 6.20 (2), 18.32, 9.38, 6.17 (2), 9.39 (2), 9.50, 6.32, 15.16, 6.28 (2), 20.18, 9.62 (2), 9.58 (3), 15.33, 3.25, 3.32 (2), 3.28, 9.75 (2), 9.80 (3), 9.76 (3), 3.41, 15.46 (2), 3.39, 3.44, 15.55, 15.62, 12.2, 8.3 (2), 8.4, 8.5 (2), 8.6 (2), 8.9, 8.10, 8.12, 8.14 (2), 8.15 (2), 17.10, 8.21, 8.23 (3), 17.13, 17.15, 17.20 (3), 8.32, 17.18, 5.13, 17.24, 8.40 (2), 5.20, 5.17, 17.27, 9.3, 9.4, 5.23, 5.25 (3), 5.31 (2), 5.26 (3), 5.34, 14.23 (2), 2.13, 2.14, 2.15, 5.39, 5.44, 2.19 (2), 5.46 (2), 2.30 (2), 2.1, 2.2, 11.22, 2.3 (2), 11.18, 2.4 (2), 11.30 (2), 2.39, 2.8, 11.34

<i>a, e</i>	<i>f</i>	Übersetzungen
4.3, 5.2	0.0208333	angenehm, annehmlich, einverstanden, gefällig, lebenswürdig
5, 5.1	0.0625	begierig, eifrig
5.3, 4.8	0.0833333	fröhlich, schadenfroh
4.7, 5.4	0.0833333	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan
4.9, 3.4	0.0625	ängstlich, bange, besorgt, fürchtend, erfolgshungrig, gefräßig, gierig, habgierig
4.2, 2.7	0.125	ärgerlich, aufgebracht, böse, düster, entzündet, erbost, erzürnt, schmerzhaft, stürmisch, ungehalten, verärgert, wütend, zornig
6, 2.3	0.1875	ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt
5.6, 3.7	0.125	aufgebracht, erbost, grimmig, wütend
4.9, 5	0.0208333	belustigt, erfreut, vergnügt
5.9, 2.9	0.104167	aggressiv, angreifend, angriffslustig, ausfällig, draufgängerisch, dynamisch, energisch, feindlich, kampfbetont, offensiv, streitlustig, streitsüchtig

5.3, 2.5	0.0416667	antagonistisch, entgegenwirkend, feindlich, gegnerisch, widerstreitend
4.4, 3.1	0.0208333	betreten, gehindert, verlegen, verwirrt, in Verlegenheit, in Verlegenheit gebracht
5.3, 4.3	0.0208333	achtsam, aufmerksam
3.1, 3.8	0.0416667	abgearbeitet, leer, nichtssagend, unbeladen

Tabelle 6.2.: Informationen zur Sendung „Sesamstraße“.

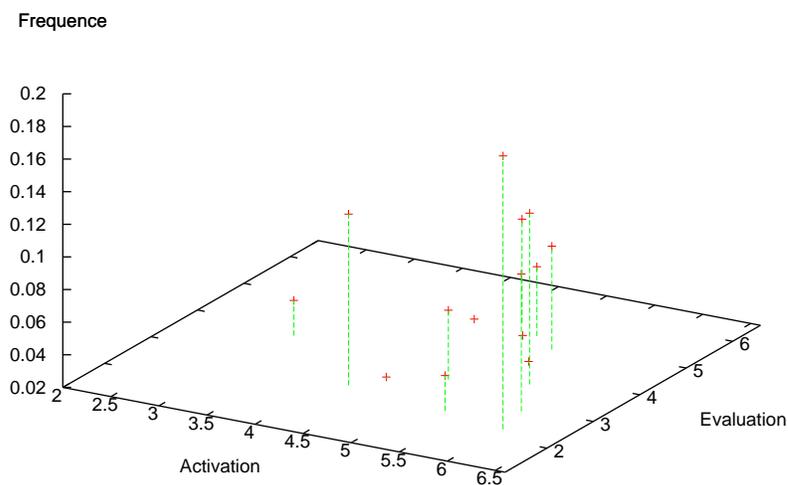


Abbildung 6.3.: Activation/Evaluation Diagramm zur Sendung „Sesamstraße“.

Es gibt Fälle, in denen das Übersetzen von Extended Infos in Activation/Evaluation Koordinaten und Frequenzen nicht zu einer adäquaten emotionalen Repräsentation einer Sendung führt. Zunächst fallen zum Beispiel folgende Probleme auf:

- Das Fehlen der Extended Info einer Sendung (die Hälfte der zufällig erhobenen Beispielsendungen war ohne Extended Info).
- Mehrdeutigkeiten der Sprache: Deutlich bei der emotionalen Deutung der Extended Info von „88acht GUTEN TAG BERLIN“, siehe Tabelle 6.7.

Extended Info		
Bianca macht sich Sorgen, weil sie fühlt, dass Eddie versucht, sich selbst über seine Trauer hinweg zu täuschen. Heiko kann seine Wut immer noch kaum zügeln. Er macht Bianca Vorwürfe und hält ihr vor, dass sie mit Oliver den falschen Mann heiraten wird. Deutschland, 2005		
Dornseiffgruppen und deren Häufigkeiten		
11.52, 15.1 (5), 3.3 (3), 3.4, 15.6, 21.21, 21.19 (2), 4.3, 21.8 (2), 21.9, 4.32, 4.33 (2), 4.28 (3), 4.37, 4.50, 10.13 (2), 4.51, 10.14, 4.52, 10.30, 10.31, 10.32, 5.1 (4), 10.41, 5.5, 10.57, 10.1, 9.12, 9.17, 9.22, 18.12 (4), 6.10 (2), 6.5, 18.17 (2), 6.6, 6.7, 9.40, 6.20, 18.30 (2), 20.10, 18.33, 9.50, 6.26, 9.49, 15.24, 3.14, 3.15, 9.57, 3.18, 9.71, 3.25, 9.68, 3.32, 9.80, 12.19, 3.35, 15.53, 12.28, 3.44, 12.29, 12.40 (2), 15.70 (2), 12.50 (3), 12.45, 17.11, 8.23 (2), 5.10, 5.13, 17.19, 5.19, 5.26, 14.20, 5.31 (2), 2.13, 2.14 (2), 5.35 (3), 5.36 (3), 2.16 (2), 5.38, 11.13, 5.47, 11.21, 11.17, 11.23, 11.25 (3)		
<i>a, e</i>	<i>f</i>	Übersetzungen
3.1, 3.8	0.015873	abgearbeitet, leer, nichtssagend, unbeladen
6, 2.3	0.238095	ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt
2.4, 3.2	0.047619	bedrückt, dunkel, düster, finster
5.3, 4.8	0.015873	fröhlich, schadenfroh
4.7, 5.4	0.015873	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan
4.9, 3.4	0.0793651	ängstlich, bange, besorgt, fürchtend, erfolgshungrig, gefräßig, gierig, habgierig
4.2, 2.7	0.126984	ärgerlich, aufgebracht, böse, düster, entzündet, erbost, erzürnt, schmerzhaft, stürmisch, ungehalten, verärgert, wütend, zornig
5.6, 3.7	0.190476	aufgebracht, erbost, grimmig, wütend
5.9, 2.9	0.111111	aggressiv, angreifend, angriffslustig, ausfällig, draufgängerisch, dynamisch, energisch, feindlich, kampfbetont, offensiv, streitlustig, streitsüchtig
3.8, 2.8	0.031746	argwöhnisch, misstrauisch
4.4, 3.1	0.047619	betreten, embarrassed, gehindert, verlegen, verwirrt, in Verlegenheit, in Verlegenheit gebracht
5.3, 2	0.015873	abgünstig, neidig, neidisch, scheelsüchtig
5.3, 4.3	0.015873	achtsam, aufmerksam
5, 5.1	0.047619	begierig, eifrig

Tabelle 6.3.: Informationen zur Sendung „Bianca - Wege zum Glück“.

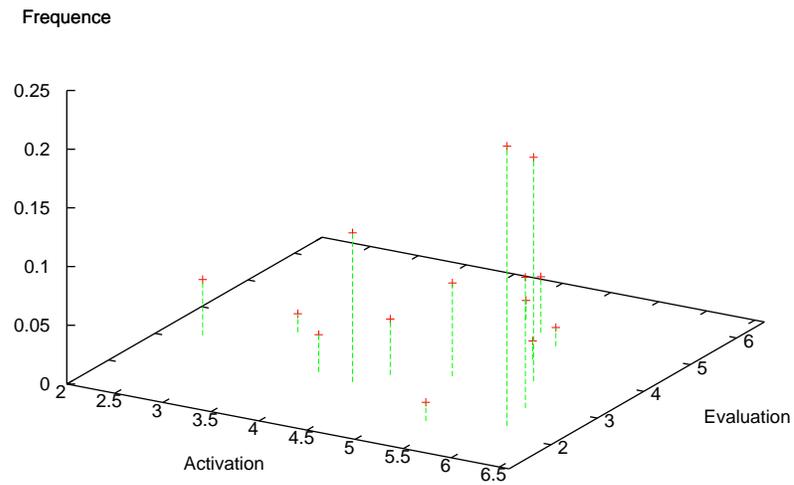


Abbildung 6.4.: Activation/Evaluation Diagramm zur Sendung „Bianca - Wege zum Glück“.

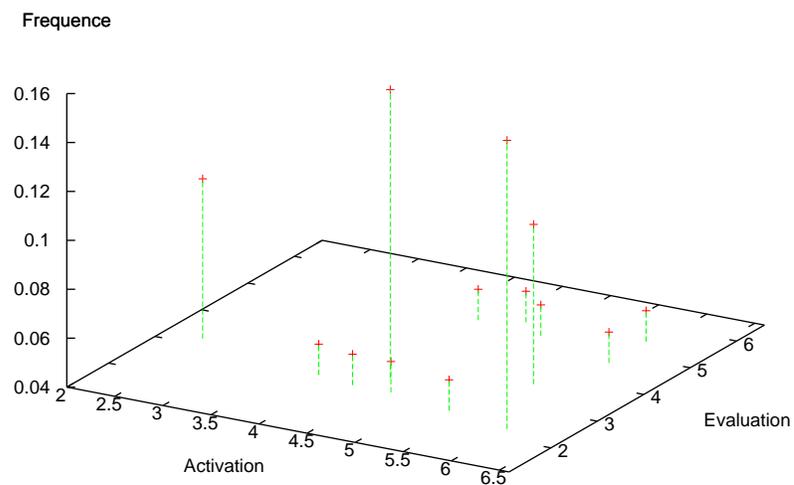


Abbildung 6.5.: Activation/Evaluation Diagramm zur Sendung „aspekte kompakt“.

Extended Info

Wagner-Marathon in Tirol Ein Rund-um-die-Uhr-Spektakel: Richard Wagners Trilogie „Der Ring des Nibelungen“ wird erstmalig innerhalb von 24 Stunden aufgeführt.

„Extrem laut und unglaublich nah“ Wie lässt sich der 11. September aus Sicht eines kleinen Jungen erzählen? Der amerikanische Autor Jonathan Safran Foer versucht es in seinem neuen Roman.

Ein Hauch vom Orient Die Kunsthalle Krems in Österreich zeigt in der Ausstellung „Harem - Geheimnis des Orients“ Kunstwerke aus dem 19. Jahrhundert.

Musiktip: Tanzmagie aus Brasilien Umwerfende Schönheit der Bewegung - das „Bahia Ballett“ ist berühmt für seine ausdrucksreiche Tanzsprache. 2005

Dornseiffgruppen und deren Häufigkeiten

11.49 (2), 20.1, 15.1 (3), 15.4, 3.3 (7), 15.6, 3.8, 7.11, 7.23, 7.30 (2), 4.10, 7.28 (2), 4.17, 7.38, 4.3, 4.4, 13.12 (3), 7.50, 4.24, 4.7, 13.15, 4.28, 4.33 (3), 1.13, 4.38, 4.39, 4.50, 10.20, 5.1 (6), 10.29, 5.3, 5.5, 5.8, 10.6 (2), 6.1 (2), 9.30, 6.2, 9.31 (3), 6.11 (4), 6.14, 6.20, 6.21, 6.9 (2), 15.11 (2), 6.23, 15.25, 3.16 (2), 3.23 (4), 3.18 (6), 12.13, 15.36, 12.21, 7.2, 9.80 (3), 15.51 (2), 3.37, 15.48, 19.9 (2), 15.62, 12.50, 12.51 (2), 12.2, 12.4, 12.6, 12.58, 12.7, 17.10 (3), 8.23 (2), 5.13, 1.7, 17.19 (2), 8.41, 17.25, 5.21, 17.26, 5.17, 14.11, 8.39 (2), 5.19, 5.31 (5), 5.26, 14.21 (3), 14.17, 5.29, 14.23 (3), 2.15 (5), 5.36, 5.39, 14.1, 11.21, 2.2, 11.23, 11.31, 11.33, 11.36

<i>a, e</i>	<i>f</i>	Übersetzungen
4.4, 3.1	0.157895	betreten, gehindert, verlegen, verwirrt, in Verlegenheit, in Verlegenheit gebracht
4.7, 5.4	0.0526316	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan
5.9, 5.5	0.0526316	erstaunt, überrascht, verblüfft
5.9, 4.7	0.0526316	erstaunt, überrascht, verwundert, in Erstaunen versetzt
6, 2.3	0.157895	ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt
4.3, 5.2	0.0526316	angenehm, annehmlich, einverstanden, gefällig, lebenswürdig
4.6, 2.7	0.0526316	frustriert, unzufrieden
5.3, 2.5	0.0526316	antagonistisch, entgegenwirkend, feindlich, gegnerisch, widerstreitend
2.4, 3.2	0.105263	bedrückt, dunkel, düster, finster
5, 5.1	0.0526316	begierig, eifrig
4.2, 2.7	0.0526316	ärgerlich, aufgebracht, böse, düster, entzündet, erbost, erzürnt, schmerzhaft, stürmisch, ungehalten, verärgert, wütend, zornig
5.6, 3.7	0.105263	aufgebracht, erbost, grimmig, wütend
3.8, 2.8	0.0526316	argwöhnisch, misstrauisch

Tabelle 6.4.: Informationen zur Sendung „aspekte kompakt“.

Extended Info		
Wettercheck - jede halbe Stunde		
Lothar & Franz		
Keksi & Bussi		
Bundesliga-Flashback		
Dornseiffgruppen und deren Häufigkeiten (falls größer 1)		
15.1 (2), 4.45, 17.1, 17.4, 6.1, 9.32, 9.7, 11.31		
<i>a, e</i>	<i>f</i>	Übersetzungen
3, 4.3	1	apathisch, teilnahmslos, bescheiden, demütig, gütig, kleinlaut, mild, sanft, zahm

Tabelle 6.5.: Informationen zur Sendung „Bayern 3 - die MorningShow“.

<i>Wort/Wörter</i>	<i>Dornseiffgruppe</i>
Franz, Lothar	15.1 Vornamen (15. Menschliches Zusammenleben)
halbe	4.45 Bruchteil (4. Größe, Menge, Zahl)
halbe	9.32 Unvollendet lassen (9. Wollen und Handeln)
halbe	9.7 Unentschlossen (9. Wollen und Handeln)
Bundesliga	17.1 Sport (17. Sport und Freizeit)
Bundesliga	17.4 Wettkampf, Turnier (17. Sport und Freizeit)
Stunde	6.1 Zeitraum (6. Zeit)
Stunde	11.31 Lehren (11. Das Denken)

Tabelle 6.6.: Wörter der „Extended Info“ von „Bayern 3 - die MorningShow“ und korrespondierende Dornseiffgruppen.

Extended Info		
15:30 88acht-Report, Wetter, Verkehr		
16:00 Nachrichten, Wetter, Verkehr		
16:30 88acht-Report, Wetter, Verkehr		
17:00 Nachrichten, Wetter, Verkehr		
17:30 88acht-Report, Wetter, Verkehr		
Dornseiffgruppen und deren Häufigkeiten		
4.39 (3), 15.21 (3), 15.33 (4), 12.45 (3), 12.2 (5), 12.57 (3), 12.6 (5), 12.7 (5), 8.6 (4), 1.5 (5), 1.6 (5), 14.21 (3), 2.25 (4)		
<i>a, e</i>	<i>f</i>	Übersetzungen
4.7, 5.4	1	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan

Tabelle 6.7.: Informationen zur Sendung „88acht GUTEN TAG BERLIN“.

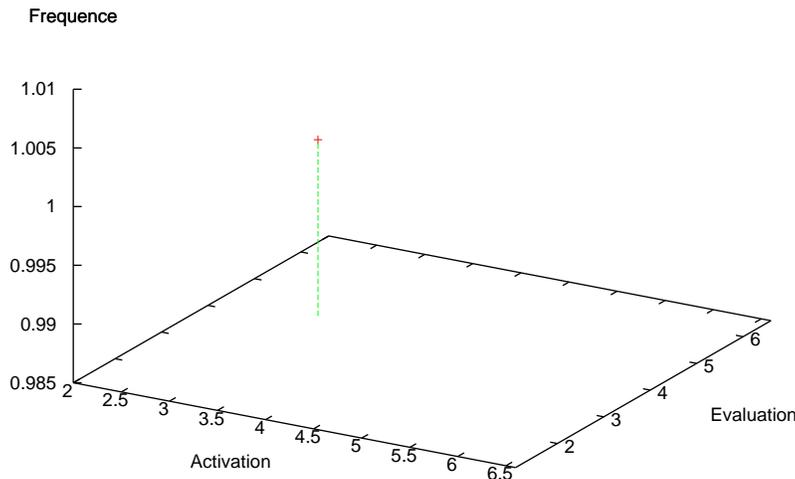


Abbildung 6.6.: Activation/Evaluation Diagramm zur Sendung „Bayern 3 - die MorningShow“.

6.3. Testsituation für Konfigurationen des Lernsystems

Um zu sehen, inwiefern die Resultate (gelernte Präferenzen, generierte Hypothesen und berechnete Fitnesswerte für die verfügbaren Filme) verschiedener Konfigurationen des Lernsystems voneinander abweichen, untersuchen wir für jede ein und dieselbe Situation. Die Testsituation deckt nur Fälle ab, bei denen keine Layer gelernt werden müssen. Es geht zunächst ausschließlich darum zu sehen, inwiefern verschiedene Konfigurationen adäquate Vorschläge für einzelne Anwender und Gruppen generiert. Es werden keine Bewertungen von einer Gruppe entgegengenommen, weswegen keine Widersprüche zwischen individuellen Präferenzen und Gruppenpräferenzen auftreten können und auch keine Daten über Ausnahmen bezüglich der Bewertungen in Layer übernommen werden müssen.

Die Bildung von Layern wird anhand zweier Testläufe mit Benutzereingaben in Kapitel 6.7.1 und 6.7.2 veranschaulicht.

In der Testsituation passiert Folgendes.

- Die drei Fernsehsendungen „nano“, „Lindenstraße“ und „aspekte kompakt“ stehen zur Auswahl
- Benutzer Emil bewertet die Fernsehsendung „nano“ mit +1

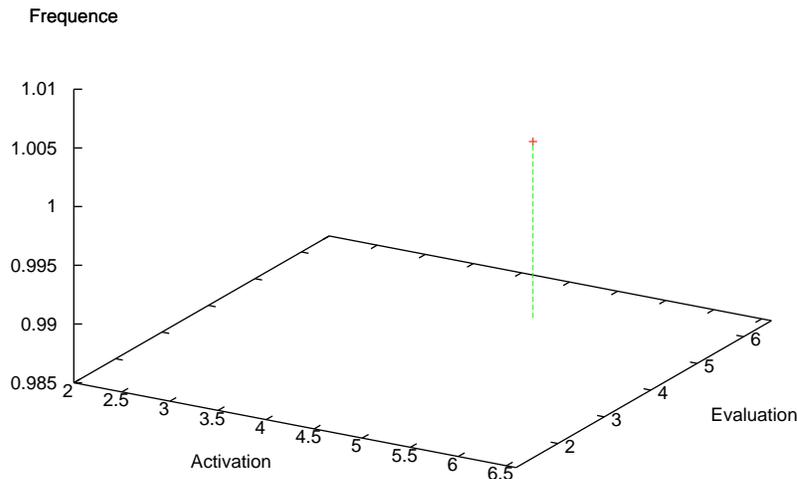


Abbildung 6.7.: Activation/Evaluation Diagramm zur Sendung „88acht GUTEN TAG BERLIN“.

- Es werden die Ergebnisse des Lernsystems basierend auf Emils Präferenzen generiert
- Benutzer Bert bewertet die Fernsehserie „Lindenstraße“ mit +1 und die Serie „aspekte kompakt“ mit -0.5
- Nun werden die Ergebnisse des Lernsystems basierend auf Berts Präferenzen generiert
- Anschließend werden die Ergebnisse des Lernsystems basierend auf Emils und Berts Präferenzen generiert

„nano“, „Lindenstraße“ und „aspekte kompakt“ werden in den Tabellen 6.8, 6.9 und 6.4 beschrieben, ihre emotionalen Repräsentationen sind in den Abbildungen 6.8, 6.9 und 6.5 visualisiert.

„Lindenstraße“ hat nur Einträge an Activation/Evaluation Koordinaten, die nicht in „nano“ vorkommen. Eine Bewertung von „nano“ und die daraus resultierende Hypothese sollten die Beurteilung von „Lindenstraße“ daher nur wenig beeinflussen. „aspekte kompakt“ teilt sich 8 Einträge an Activation-/Evaluation Koordinaten mit „Nano“ (insgesamt 10 Einträge). Darüberhinaus beinhaltet die Repräsentation von „aspekte kompakt“

noch Einträge an 5 weiteren Activation/Evaluation Koordinatenpaaren. „aspekte kompakt“ und „nano“ lassen sich teilweise mit den gleichen Merkmalen beschreiben, sind sich also zu einem bestimmten Grad ähnlich. Für die eine Sendung gelernte Informationen sollten entsprechend auf die andere Sendung anwendbar sein. Ähnliche Instanzen sollten ähnliche Fitnesswerte zugeteilt bekommen.

„aspekte kompakt“ hält auch eine Merkmalsausprägung an einem Koordinatentupel bereit, die in „Lindenstraße“ vorkommt (A/E-Koordinaten 5 und 5.1). Für „aspekte kompakt“ spielt dieser Eintrag allerdings kaum eine Rolle, er trägt zur Beschreibung der Sendung nur zu circa 1/19 bei.

Nach Emils Wertung ist folgende Ordnung der Sendungen nach Fitnesswerten wünschenswert: „nano“ (höchste Wertung), „aspekte kompakt“ (wegen hoher Ähnlichkeit zu „nano“ gleich danach), „Lindenstraße“ (geringste Fitness, da keine Überschneidungen mit „nano“).

Nach Berts Wertung ist eine Rangordnung der Filme wünschenswert, die mit der „Lindenstraße“ an höchster Position beginnt. Da „Lindenstraße“ und „aspekte kompakt“ sich in einem Vektor überschneiden, kann nicht eindeutig Stellung bezogen werden für eine fixe Ordnung auf den Rängen zwei und drei. Für „nano“ an zweiter Stelle spricht, dass „nano“ selbst nicht negativ bewertet worden ist. Für „aspekte kompakt“ an zweiter Stelle spricht, dass es zwar negativ bewertet wurde, jedoch eine Überschneidung mit der positiv bewerteten „Lindenstraße“ aufweist, was für „nano“ nicht zutrifft.

Es ist offenkundig, dass es für Emil und Bert ein guter Kompromiss wäre, „Lindenstraße“ zu schauen: Bert möchte „aspekte kompakt“ nicht (das dazu ähnliche „nano“ wäre ein gewagter Vorschlag) und Emil hat sich noch nicht wertend zu „Lindenstraße“ geäußert.

In den folgenden Abschnitten werden die Ergebnisse der verschiedenen Konfigurationen bezüglich des oben geschilderten Lernverlaufs zusammengetragen.

Die Ergebnisse des Lernsystems hängen von verschiedenen Komponenten ab (siehe Kapitel 5). Einerseits spielt die Art der Hypothesenbildung eine große Rolle. Andererseits ist die Wahl der Interpolationsfunktion für die alternative Hypothesenbildung, sowie für die meisten Fitnessfunktionen ebenso relevant.

Deshalb wird am Anfang jedes Abschnitts für die Kombination aus Interpolationsfunktion und Hypothesenbildungsmethode gezeigt, wie man sich die interpolierten Präferenzen von Emil und Bert, sowie die daraus generierte, interpolierte Hypothesenmenge vorstellen muss. Anschließend werden die Ergebnisse der verschiedenen Fitnessfunktionen betrachtet.

Einzige Ausnahme ist die Konfiguration aus Hypothesenbildung durch „gewichtete Übernahme der Präferenzen und Layer“ und Fitnessberechnung mittels „Differenz der Durchschnittsvektoren“. Hier werden keine Interpolationsmechanismen verwendet, weswegen nur ein Fall (statt dreier Fälle) betrachtet werden muss.

Extended Info		
Themen u.a.:		
* Gutes Essen gerade in der Schule		
* Dämmer Schlaf gegen die krankhafte Angst vorm Zahnarzt		
* Chip mustert Bakterien		
* Windkraft vom Hochhaus		
Dornseiffgruppen und deren Häufigkeiten		
20.1 (2), 20.3 (2), 3.3, 3.8, 7.11, 19.16, 19.19 , 21.10, 16.12 (2), 7.27, 21.3 (2), 4.4, 7.44, 16.5 (2), 4.36 (2), 1.16 (2), 10.13 (2), 7.68, 10.29 (2), 10.41 (2), 10.51 (2), 9.12 (2), 9.14 (2), 10.6, 9.24, 6.11, 18.30, 6.15 (2), 6.18 (2), 9.45 (2), 9.54 (2), 3.18, 20.29 (2), 11.7, 3.31, 19.1 (2), 19.2, 9.77 (4), 22.11, 13.4 (2), 13.5 (2), 13.6 (2), 5.23, 5.18, 9.8 (2), 5.29, 2.22, 5.39 (2), 5.47 (2), 11.20, 2.33, 14.4, 2.35, 2.4 (2), 11.30		
a, e	f	Übersetzungen
4.4, 3.1	0.195652	betreten, embarrassed, gehindert, verlegen, verwirrt, in Verlegenheit, in Verlegenheit gebracht
4.9, 3.4	0.130435	ängstlich, bange, besorgt, fürchtend, erfolgshungrig, gefräßig, gierig, habgierig
4.2, 2.7	0.0434783	ärgerlich, aufgebracht, böse, düster, entzündet, erbost, erzürnt, schmerzhaft, stürmisch, ungehalten, verärgert, wütend, zornig
6, 2.3	0.23913	ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt
5.9, 5.5	0.0434783	erstaunt, überrascht, verblüfft
5.9, 4.7	0.0434783	erstaunt, überrascht, verwundert, in Erstaunen versetzt
3.8, 2.8	0.0434783	argwöhnisch, misstrauisch
4.7, 5.4	0.130435	anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan
5.3, 4.3	0.0869565	achtsam, aufmerksam
4.3, 5.2	0.0434783	angenehm, annehmlich, einverstanden, gefällig, liebenswürdig

Tabelle 6.8.: Informationen zur Sendung „nano“.

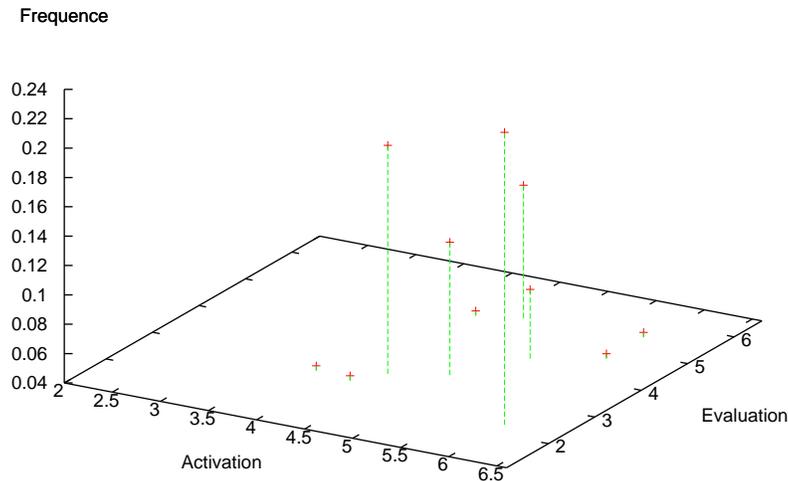


Abbildung 6.8.: Activation/Evaluation Diagramm zur Sendung „nano“.

 Extended Info

Franziska plant eine Modenschau und ist auf der Suche nach Freiwilligen, die für sie über den Laufsteg gehen. Das wird jedoch schnell zur Nebensache, als sie zufällig ein Gespräch zwischen Urszula und Christian mithört.

 Dornseiffgruppen und deren Häufigkeiten

15.1 (2), 3.3 (2), 3.9, 19.15, 4.3, 4.33 (2), 4.28 (2), 5.1 (2), 5.2, 10.35, 10.36, 5.7, 9.11, 9.14, 9.16, 9.21, 6.11 (2), 6.12, 9.36, 6.20, 9.37, 6.17, 9.43, 11.1, 3.14, 3.18, 3.24, 3.26, 12.21, 3.32 (2), 9.75, 12.30, 12.47, 12.7, 8.3, 8.6, 8.23 (4), 8.19, 17.13, 13.2, 5.13, 8.40 (2), 5.20, 9.2, 8.38, 5.31, 5.26, 5.33, 5.29, 2.14 (2), 5.45

a, e	f	Übersetzungen
5, 5.1	0.571429	begierig, eifrig
5.3, 4.3	0.285714	achtsam, aufmerksam
3.1, 3.8	0.142857	abgearbeitet, leer, nichtssagend, unbeladen

Tabelle 6.9.: Informationen zur Sendung „Lindenstraße“.

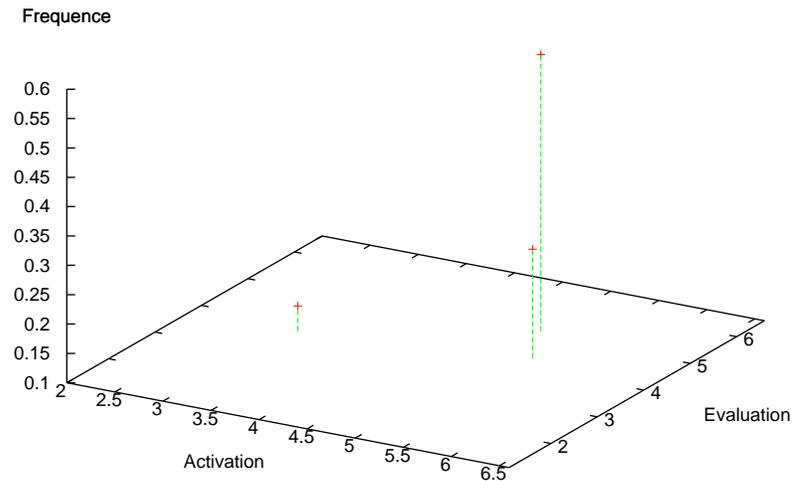


Abbildung 6.9.: Activation/Evaluation Diagramm zur Sendung „Lindenstraße“.

6.4. Gewichtete Übernahme von Präferenzen und Layern

Der direkte Weg, um aus Präferenzen und Layern eine Hypothese zu generieren, ist, deren Einträge mit einer festgelegten Gewichtung zu übernehmen. Im Folgenden wird dieses Verfahren zur Hypothesenbildung zusammen mit verschiedenen Interpolations- und Fitnessfunktionen bezüglich der oben beschriebenen Lernsituation getestet.

6.4.1. „0“ als Standardannahme

Die Präferenzen und auch die Hypothese sind in Abbildung 6.10 als Punktemengen dargestellt, da keine tatsächliche Interpolation zwischen den gelernten Vektoren stattfindet. Für unbesetzte Stellen im A/E-Raum wird 0 zurückgeliefert. Sofern noch keine konkreten Daten gelernt wurden, wird keine implizite Annahme darüber gemacht, ob nahegelegene Koordiantenpaare die gleichen oder ähnliche Bewertungen vom Anwender bekämen.

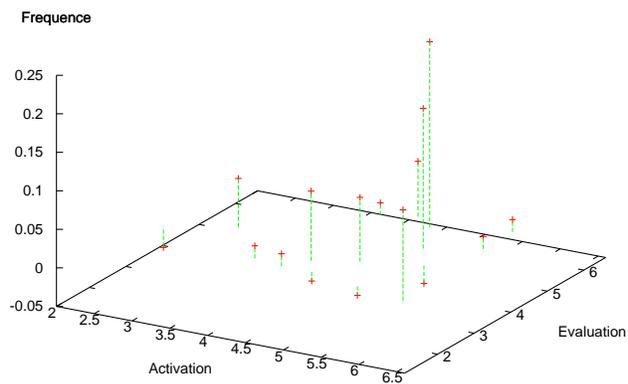
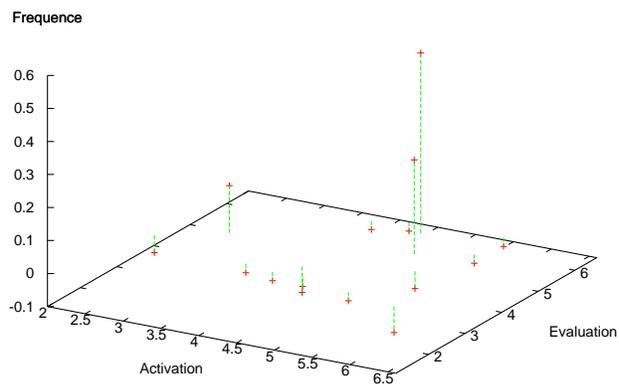
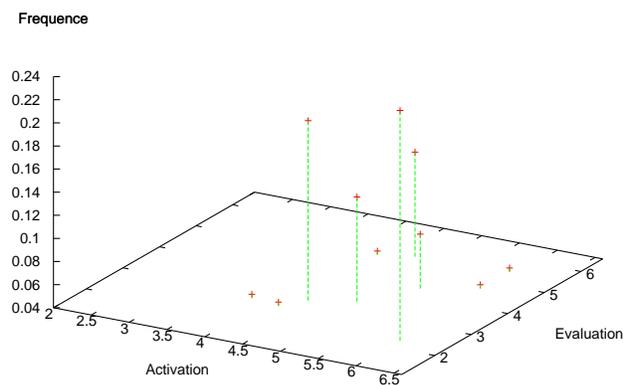


Abbildung 6.10.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-0).

<i>gew-0-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-1.02484
aspekte kompakt	-0.0709358
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-1.00841
Lindenstraße	1.95367
aspekte kompakt	-2.24406
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.585481
Lindenstraße	0.566848
aspekte kompakt	0.385979

Tabelle 6.10.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.

Fitness anhand des Gesamtabstands

Der bloße Vergleich von Hypothese und Filmrepräsentationen durch die Berechnung des Gesamtabstands funktioniert tadellos (siehe Tabelle 6.10). Einerseits wird entsprechend Emils Bewertungen „nano“ an die erste Stelle der Rangordnung gesetzt, „aspekte kompakt“ folgt auf den Schritt und an letzter Stelle steht „Lindenstraße“. Die Bewertungen von Bert werden ebenso in den für ihn erzeugten Fitnesswerten der Sendungen widergespiegelt. „Lindenstraße“ bekommt eine bessere Fitness vom Lernsystem zugeschrieben als „aspekte kompakt“ (an 2. Stelle) oder „nano“. Ebenfalls positiv zu bemerken ist, dass die Bewertungen von „nano“ und „aspekte kompakt“ näher beisammen liegen, als sie von „Lindenstraße“ entfernt sind. Die Ähnlichkeit der beiden Sendungen schlägt sich also im Ergebnis nieder.

Nach der Zusammenführung von Emils und Berts Präferenzen bekommt „nano“ die höchste Fitness zugeordnet. Berts negative Bewertung der ähnlichen Sendung „aspekte kompakt“ beeinflusst die Fitness von „nano“ offensichtlich nicht stark genug. „Lindenstraße“ erreicht den zweithöchsten Fitnesswert. „aspekte kompakt“ wurde als einzige Sendung negativ bewertet und fällt dementsprechend auf den dritten Platz. Der Wunschkandidat für einen Kompromiss „Lindenstraße“ kann sich mit der vorliegenden Fitnessfunktion nicht durchsetzen.

Fitness anhand anteilmäßiger Bewertung der Merkmale

Die durch die Präferenzen gewichtete Summe der Instanzmerkmale (siehe Tabelle 6.11) scheint besser zu funktionieren. Für Emil bekommen „nano“ und „aspekte kompakt“ (in dieser Reihenfolge) eine verhältnismäßig hohe Fitness, die ursprüngliche Fitness von „Lindenstraße“ wird kaum durch Emils Bewertung beeinträchtigt.

Bert wäre ebenfalls mit der vom Lernsystem generierten Reihenfolge zufrieden. „Lindenstraße“ hat mit Abstand die größte Fitness. „nano“ bekommt eine leicht schlechtere Fitness als „aspekte kompakt“ zugestanden. Dieses Ergebnis ist insofern plausibel, da „aspekte kompakt“ noch eine, wenn auch sehr geringe, Überschneidung mit der positiv bewerteten Sendung „Lindenstraße“ aufweist.

Die vom Lernsystem vorgeschlagene Rangordnung für beide Anwender sieht „Lindenstraße“ an erster Stelle vor. „nano“ erreicht eine fast doppelt so hohe Fitness wie „aspekte kompakt“.

Fitness anhand der Kombination von „Abst“ und „Gewi“

Die nicht plausible Rangordnung der Vorschläge für beide Anwender, die von der Fitness des Gesamtabstands herrührt, schlägt sich nicht bei der Kombination der beiden vorhergehenden Methoden nieder. Tabelle 6.12 wartet entsprechend mit Werten auf, die sowohl im Einzelfall gute Ergebnisse liefern, als auch für beide Anwender zugleich vernünftige Werte berechnet („Lindenstraße“ auf dem ersten Platz, „aspekte kompakt“ durch einmalig negative Bewertung an letzter Stelle).

<i>gew-0-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.0496894
aspekte kompakt	0.173913
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0257482
Lindenstraße	0.571429
aspekte kompakt	-0.0254272
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.17092
Lindenstraße	0.397227
aspekte kompakt	0.094962

Tabelle 6.11.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>gew-0-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-0.975155
aspekte kompakt	0.102977
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-1.03416
Lindenstraße	2.5251
aspekte kompakt	-2.26949
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.756402
Lindenstraße	0.964075
aspekte kompakt	0.480941

Tabelle 6.12.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

6.4.2. Voronoi Interpolation

Die in Abbildung 6.11 dargelegten Ergebnisse für Emils und Berts Präferenzen und der Hypothese für beide basieren auf der vereinfachenden Annahme, dass man für bislang unbesetzte A/E-Koordinatenpaare die Bewertung des nächsten Nachbarpunktes in Betracht zieht. Diese Interpolation partitioniert die Präferenzen und Hypothesen bezüglich der A/E-Ebene als Voronoi Diagramm.

Fitness anhand des Gesamtabstands

In Tabelle 6.13 sind die vom Lernsystem erzeugten Fitnesswerte in Emils Fall passend („nano“ und „aspekte kompakt“ haben positive Fitnesswerte, „Lindenstraße“ hat einen weitaus schlechteren, sogar negativen Fitnesswert), auch Berts Präferenzen werden von dem Lernsystem adäquat bedient: „Lindenstraße“ bekommt einen sehr hohen, positiven Wert zugeordnet, „nano“ und „aspekte kompakt“ fallen mit hohen negativen Fitnesswerten weit zurück.

Die Verrechnung der individuellen Präferenzen auf der Hypothese für beide Anwender resultiert allerdings in einer ungünstigen Fitnessrangordnung der Titel. „aspekte kompakt“ erreicht in diesem Fall die beste Fitness, dicht gefolgt von „nano“. Die Sendung „Lindenstraße“ fällt auf den dritten Platz der Fitnessrangordnung, ihre Bewertung ist wesentlich kleiner als die der anderen beiden.

Fitness anhand anteilmäßiger Bewertung der Merkmale

Die Werte in Tabelle 6.14 entsprechen den Erwartungen für den Anwender Bert und für beide Benutzer. Die Fitnessberechnungen für Emil hingegen liefern disputable Werte. Nach der positiven Bewertung von „nano“, erreicht „Lindenstraße“ die höchste Fitness. Erwartungsgemäß müsste nach einer sehr positiven Bewertung der Sendung „nano“ diese Sendung auch den höchsten Fitnesswert erhalten. „Lindenstraße“ und „nano“ haben keine gemeinsamen Merkmale, weswegen man von einer positiven Bewertung von „nano“ nicht auf eine übermäßig hohe Fitness von „Lindenstraße“ schließen kann. „Lindenstraße“ wurde von Emil nicht bewertet, und dennoch wird dieser Sendung der höchste Fitnesswert zugeordnet.

Es fällt auf, dass in allen drei Szenarien die Fitnesswerte von „nano“ und „aspekte kompakt“ sehr ähnlich sind und sich deutlich von den Werten für „Lindenstraße“ abheben.

Fitness anhand der Kombination von „Abst“ und „Gewi“

Der Versuch aus den beiden vorhergehenden Fitnessfunktionen adäquate Fitnesswerte zu berechnen, schlägt im Falle der Anwesenheit beider Benutzer fehl (siehe Tabelle 6.15). „Lindenstraße“ bekommt den schlechtesten Fitnesswert zugeteilt, was eindeutig Berts

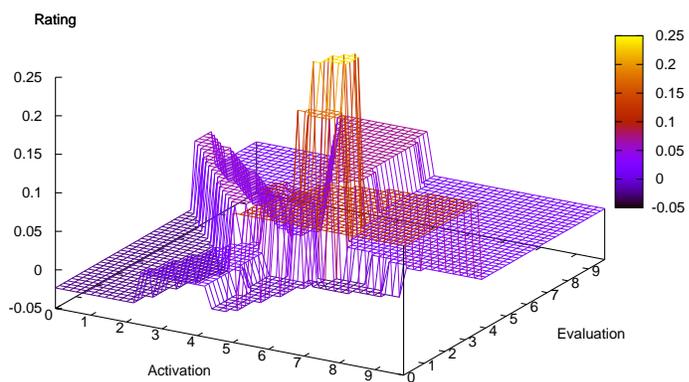
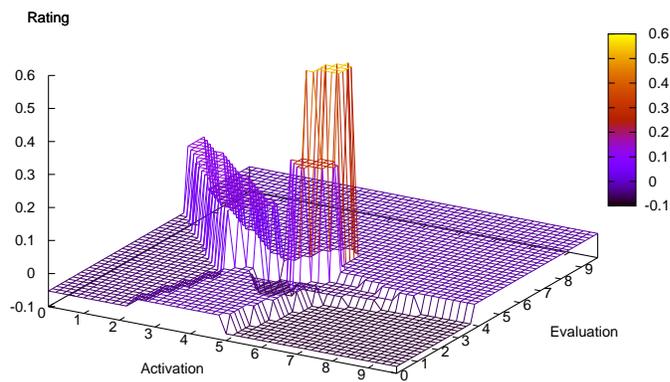
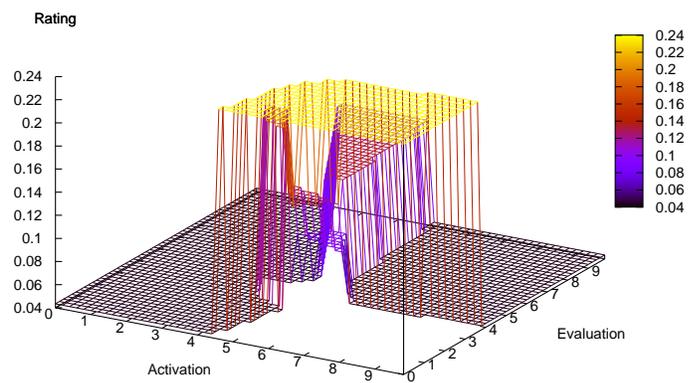


Abbildung 6.11.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-Voronoi).

<i>gew-Voronoi-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-2.13043
aspekte kompakt	0.115564
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.30918
Lindenstraße	3.4378
aspekte kompakt	-3.94445
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	1.36084
Lindenstraße	-2.52534
aspekte kompakt	1.5867

Tabelle 6.13.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>gew-Voronoi-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.413044
aspekte kompakt	0.239131
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	0.010672
Lindenstraße	0.48052
aspekte kompakt	0.00574152
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.198762
Lindenstraße	0.519826
aspekte kompakt	0.133921

Tabelle 6.14.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>gew-Voronoi-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-1.71739
aspekte kompakt	0.354694
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.29851
Lindenstraße	3.91832
aspekte kompakt	-3.9387
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	1.5596
Lindenstraße	-2.00551
aspekte kompakt	1.72062

Tabelle 6.15.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Voronoi-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

positiver Bewertung des Titels widerspricht. „nano“ und „aspekte kompakt“ haben beide ähnliche und insgesamt wesentlich höhere Fitnesswerte als „Lindenstraße“.

Betrachtet man nur die Fälle für einzelne Anwender, so können Emil und Bert jeweils für sich mit den Ergebnissen zufrieden sein: „nano“ erreicht für Emil den höchsten Fitnesswert, gefolgt von „aspekte kompakt“. „Lindenstraße“ fällt auf den letzten Platz der Rangliste für Emil. Dieses Ergebnis entspricht den Erwartungen nach einer positiven Bewertung von „nano“.

In Berts Fall hebt sich die Fitness von „Lindenstraße“ deutlich von den anderen beiden Sendungen ab. Entsprechend der negativen Bewertung von „aspekte kompakt“ erhält diese Sendung den schlechtesten Fitnesswert.

6.4.3. Inverse Distance Weighted Interpolation

Abbildung 6.12 zeigt die Bewertungen sämtlicher A/E-Koordinaten der Präferenzen und der gemeinsamen Hypothese von Emil und Bert. Die Graphen ergeben sich aus den gelernten Merkmalsbewertungen und zwischen diesen interpolierten Werten. Die Interpolation passiert durch Gewichtung der Bewertungen an den Stützpunkten (hier gleichzusetzen mit den Merkmalsbewertungen) invers zu deren Abständen zum Anfragepunkt.

Fitness anhand des Gesamtabstands

Wie bereits durch Voronoi Interpolation im vorhergehenden Abschnitt, erzeugt die Fitnessfunktion anhand des Gesamtabstands auch in Kombination mit Inverse Distance Weighted Interpolation keine adäquaten Werte (siehe Tabelle 6.16). Die Fitnessberechnungen aufgrund der individuellen Bewertungen sind akzeptabel. Allerdings entspricht die Reihenfolge der Fitnesswerte für beide Anwender nicht den Erwartungen. Dem von Bert negativ bewerteten „aspekte kompakt“ würde der höchste Fitnesswert zugesprochen werden, die von ihm positiv bewertete Sendung „Lindenstraße“, zu der sich Emil nicht wertend geäußert hat, bekommt die geringste Fitness zugeteilt.

Die Werte unterscheiden sich im Einzelnen nur geringfügig von denen der vorhergehenden Tests mit Voronoi Interpolation (in Tabelle 6.13).

Fitness anhand anteilmäßiger Bewertung der Merkmale

Wieder werden aufgrund der Interpolationsmethode ähnlich schlechte Werte produziert wie im Falle der Voronoi Interpolation (siehe diesmal Tabelle 6.17). Obwohl Emil „Lindenstraße“ nicht bewertet, bekommt diese Sendung die höchste Fitness zugeteilt.

Die für Bert generierten Fitnesswerte der Sendungen, sowie die Rangordnung der Fernsehsendungen für beide Anwender wären hingegen akzeptabel. Genau wie bereits bei der Verwendung der Voronoi Interpolation sind auch hier die Fitnesswerte der beiden ähnlichen Sendungen „nano“ und „aspekte kompakt“ in jedem Fall sehr nah beieinander.

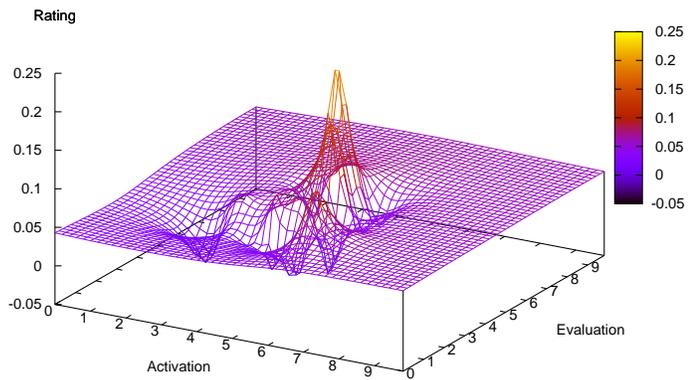
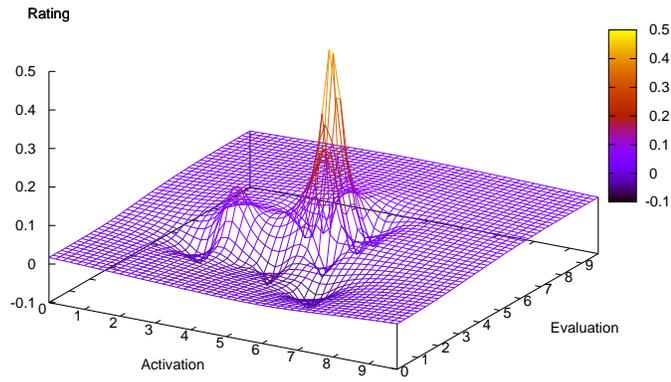
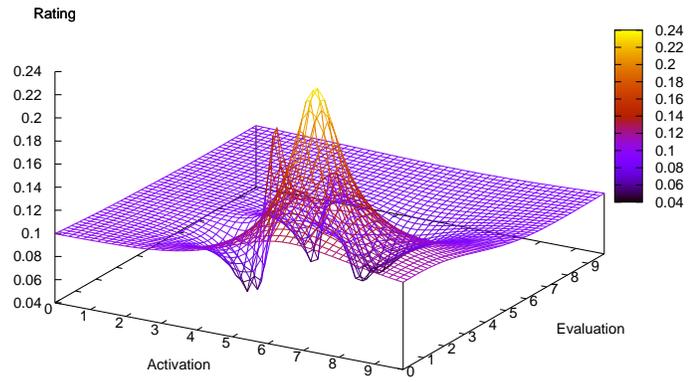


Abbildung 6.12.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration gew-Inverse).

<i>gew-Inverse-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-2.27886
aspekte kompakt	0.254128
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.59765
Lindenstraße	3.78917
aspekte kompakt	-4.30605
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.77007
Lindenstraße	-2.84414
aspekte kompakt	1.59087

Tabelle 6.16.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernalgorithmus (gew-Inverse-Abst) für verschiedene Eingaben und verschiedene Publikumsgruppen produziert.

<i>gew-Inverse-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.441802
aspekte kompakt	0.231562
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	0.00795057
Lindenstraße	0.465914
aspekte kompakt	-0.00274103
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.192397
Lindenstraße	0.536944
aspekte kompakt	0.121364

Tabelle 6.17.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Inverse-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

Fitness anhand der Kombination von „Abst“ und „Gewi“

Kombiniert man die beiden vorhergehenden, untauglichen Verfahren, so erhält man ebenso unzufriedenstellende Fitnesswerte (siehe Tabelle 6.18). Zwar wären die Rangordnungen für Emil und Bert im Einzelfall passend. In Anwesenheit beider generiert das Lernsystem allerdings unzufriedenstellende Fitnesswerte. Für beide gemeinsam wäre dementsprechend „aspekte kompakt“ der beste Kompromiss, „nano“ hätte die zweithöchste Fitness und „Lindenstraße“ wäre verhältnismäßig weit abgeschlagen.

<i>gew-Inverse-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-1.83706
aspekte kompakt	0.48569
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.5897
Lindenstraße	4.25509
aspekte kompakt	-4.30879
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.962468
Lindenstraße	-2.3072
aspekte kompakt	1.71224

Tabelle 6.18.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-Inverse-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

6.4.4. Differenz von Durchschnittsvektoren

Da die Fitnessfunktion aufgrund von Durchschnittsvektoren keine Interpolation voraussetzt, wird sie in Kombination mit gewichteter Übernahme von Präferenzen und Layern in diesem Abschnitt gesondert behandelt.

Wie die Werte in Tabelle 6.19 zeigen, kann die Differenz von Durchschnittsvektoren einer Vektormenge als Ähnlichkeitsmaß verwendet werden. Emils positive Bewertung von „nano“ bewirkt, dass die ähnlich repräsentierte Sendung „aspekte kompakt“ eine ähnliche Bewertung bekommt. Die wesentlich davon abweichende Sendung „Lindenstraße“ hingegen ist weiter entfernt von dem charakteristischen Vektor von „nano“ und erreicht daher den geringsten Fitnesswert.

Allerdings muss man bemerken, dass die Fitness von „Lindenstraße“ vom ursprünglich neutralen Wert 0 sehr stark angestiegen ist (auf 1.04122), obwohl diese Sendung nicht einmal implizit bewertet wurde.

Die für Bert erzeugten Fitnesswerte entsprechen auch den Erwartungen. So bekommt „Lindenstraße“ einen hohen Wert zugeteilt, „nano“ fällt weit zurück auf den zweiten Platz und die Sendung „aspekte kompakt“ bekommt den letzten Platz. Die Fitnesswerte von „nano“ und „aspekte kompakt“ sind, verglichen mit dem Abstand zur Fitness von „Lindenstraße“, sehr nah beisammen, was die Ähnlichkeit der beiden Sendungen unterstreicht.

Das Lernsystem errechnet für beide Anwender eine adäquate Rangordnung der Filme: „Lindenstraße“ als bester Kompromiss, gefolgt von „nano“ und an letzter Stelle „aspekte kompakt“. Es fällt auf, dass die Fitnesswerte im Falle der Anwesenheit beider Benutzer sehr nah beieinander liegen.

6.5. Alternative Hypothesenbildung

Die Idee hinter der alternativen Hypothesenbildungsmethode ist, Gemeinsames zu übernehmen und sich Widersprechendes auszuschließen. Im Folgenden werden die verschiedenen Interpolationsverfahren gemeinsam mit der alternativen Methode der Hypothesenbildung getestet.

6.5.1. „0“ als Standardannahme

Wie in Kapitel 5 erläutert, spielt die Art der Interpolation bei der alternativen Hypothesenbildung eine Rolle, bereits bevor verschiedene Fitnessfunktionen auf Hypothese und Filmkandidaten angewendet werden. Bei der alternativen Hypothesenbildung werden nur diejenigen positiven Einträge beider Präferenzen von Emil und Bert übernommen, für die die Interpolation der jeweils anderen Präferenz an der entsprechenden Stelle (A/E-Koordinatenpaar) ebenfalls ein positives Vorzeichen vorweist.

Im Gegensatz zu der zuerst diskutierten Hypothesenbildung mit 0 als Standardannahme (siehe Abbildung 6.10) kommen in der alternativen Hypothesenbildung wesentlich

<i>gew-0-AvgV</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.17467
Lindenstraße	1.04122
aspekte kompakt	1.15061
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0611978
Lindenstraße	0.175714
aspekte kompakt	-0.0995236
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.0664996
Lindenstraße	0.0788877
aspekte kompakt	0.0527214

Tabelle 6.19.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (gew-0-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.

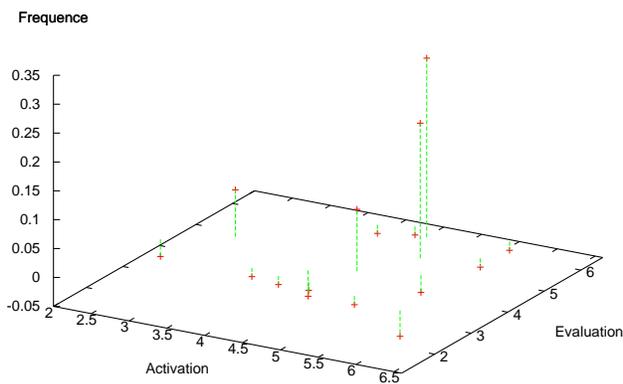
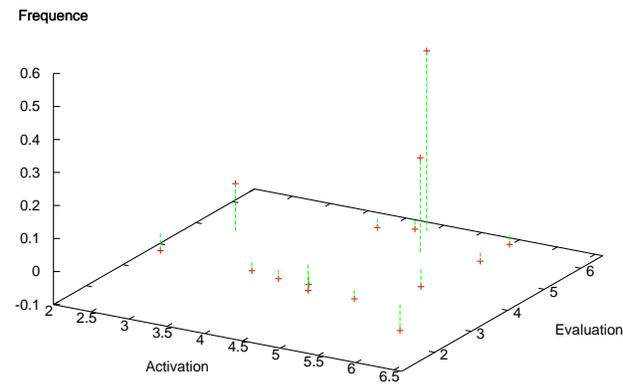
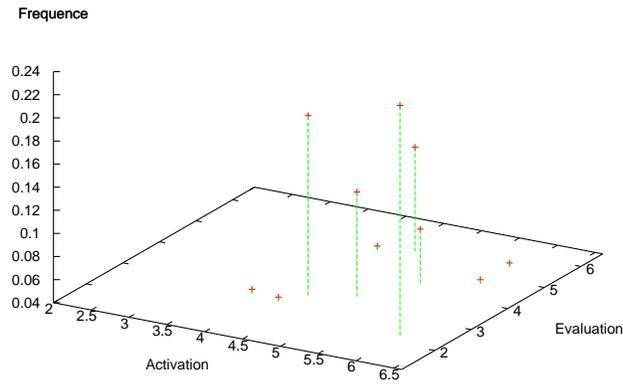


Abbildung 6.13.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-0).

mehr negative Werte vor (siehe Abbildung 6.13). Das liegt daran, dass keine negativen Werte durch Hinzuaddieren von positiven Werten der anderen Präferenz ausgeglichen werden, sondern bei derartigen Widersprüchen nur der negative Wert zur Geltung kommt.

Fitness anhand des Gesamtabstands

Die vorliegende Konfiguration generiert die gleichen Fitnesswerte für jeweils Emils und Berts Präferenzen wie die Kombination aus Hypothesenbildung durch direkten Übertrag der Präferenzen zusammen mit 0 als Defaultrückgabewert und der Ermittlung der Gesamtabstand zur Fitnessberechnung (siehe 6.10).

Der Mangel der anderen Konfiguration, „nano“ als bestmöglichen Kompromiss für beide Anwender zu werten, wird diesmal überwunden (die Werte der Konfiguration mit der alternativen Hypothesenbildung können in Tabelle 6.20 nachgesehen werden). „Lindenstraße“ erhält in der Anwesenheit beider Benutzer den weitaus besten Wert. „nano“ folgt an zweiter Stelle, „aspekte kompakt“ bildet das Schlusslicht.

Fitness anhand anteilmäßiger Bewertung der Merkmale

Wie in Tabelle 6.21 dokumentiert, produziert die gewichtete Summe von Präferenzwerten und Emotionshäufigkeiten zusammen mit 0 als Defaultrückgabewert für Emil und Bert gute Ergebnisse. Mit der alternativen Hypothesenbildung gelingt es, die erwünschten Fitnesswerte der Filmtitel im Vergleich zur gewichteten Präferenzübernahme (siehe Tabelle 6.11) noch weiter zu differenzieren. „nano“ und „aspekte kompakt“ rücken auf der Fitnessskala näher zusammen und gleichzeitig ein wenig weiter weg von „Lindenstraße“.

Fitness anhand der Kombination von „Abst“ und „Gewi“

Die Verrechnung der beiden vorherigen Fitnessfunktionen produziert gute Ergebnisse (siehe Tabelle 6.22). Genau wie die einfache Hypothesenbildung durch Übernahme gewichteter Präferenzen wird selbst für beide Anwender die adäquate Rangordnung der Sendungen berechnet.

Fitness anhand der Differenz von Durchschnittsvektoren

Für die Fitnessberechnung mittels Durchschnittsvektoren selbst ist keine Interpolation notwendig. Das Ergebnis der alternativen Hypothesenbildung hängt jedoch stark von der verwendeten Interpolationsfunktion ab, weswegen die Fitnessberechnung mittels Durchschnittsvektoren für jede Kombination von alternativer Hypothesenbildung und Interpolationsmethode erörtert werden muss.

Im vorliegenden Fall: Alternative Hypothesenbildung, „0“ als Standardannahme und Fitness anhand der Differenz von Durchschnittsvektoren, unterscheiden sich die Fitnesswerte der Sendungen in Anwesenheit beider Benutzer wesentlich deutlicher (siehe

<i>alt-0-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-1.02484
aspekte kompakt	-0.0709358
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-1.00841
Lindenstraße	1.95367
aspekte kompakt	-2.24406
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-0.789672
Lindenstraße	1.4254
aspekte kompakt	-2.10733

Tabelle 6.20.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-0-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.0496894
aspekte kompakt	0.173913
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0257482
Lindenstraße	0.571429
aspekte kompakt	-0.0254272
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.0192224
Lindenstraße	0.510204
aspekte kompakt	-0.0208867

Tabelle 6.21.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-0-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-0.975155
aspekte kompakt	0.102977
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-1.03416
Lindenstraße	2.5251
aspekte kompakt	-2.26949
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-0.770449
Lindenstraße	1.9356
aspekte kompakt	-2.12821

Tabelle 6.22.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-0-AvgV</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.17467
Lindenstraße	1.04122
aspekte kompakt	1.15061
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0611978
Lindenstraße	0.175714
aspekte kompakt	-0.0995236
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0144096
Lindenstraße	0.144971
aspekte kompakt	-0.0531382

Tabelle 6.23.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-0-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.

Tabelle 6.23), als bei der einfachen Hypothesenbildung (siehe Tabelle 6.19). Die positiv zu bewertenden Rangordnungen in den drei Fällen bleiben erhalten.

6.5.2. Voronoi Interpolation

In Abbildung 6.14 sieht man die Voronoipartitionierung bezüglich der Activation/Evaluation Ebene der Präferenzen von Emil und Bert und die Hypothese für beide, die durch die alternative Hypothesenbildung zustande kommt. Die Hypothese ähnelt Berts Präferenzen sehr. Durch die alternative Hypothesenbildung ergänzen sich die positiven Bewertungen beider Benutzer. Die negativen Einträge von Berts Präferenzen werden übernommen. Durch die Normalisierung der Hypothese verringern sich Berts negative Bewertungen leicht.

Fitness anhand des Gesamtabstands

Nach Emils Bewertung erfährt „nano“ die höchste Fitness. „aspekte kompakt“ wird bei weitem nicht so gut eingestuft. Die Sendung „Lindenstraße“ bekommt die schlechteste Wertung - nochmals weit entfernt von „aspekte kompakt“ (siehe Tabelle 6.24). In Berts Fall bekommt „Lindenstraße“ einen extrem positiven, die beiden anderen Sendungen extrem schlechte Fitnesswerte („nano“ Platz 2, „aspekte kompakt“ Platz 3). Die Rangordnung der Titel im Falle der Anwesenheit beider Benutzer ist akzeptabel, „Lindenstraße“ belegt den ersten Platz und lässt die Sendungen „nano“ und „aspekte kompakt“ weit hinter sich (in dieser Reihenfolge). Es fällt auf, dass sich der Fitnesswert von „aspekte kompakt“ nochmals verschlechtert, sobald Emil Bert Gesellschaft leistet.

Fitness anhand anteilmäßiger Bewertung der Merkmale

Wie bereits bei den vorhergehenden Tests mit „gew“-Hypothesenbildung (siehe Tabelle 6.14) berechnet die Fitnessfunktion mittels Addition gewichteter Häufigkeiten gepaart mit Voronoi Interpolation nur teilweise gute Werte. In Emils Anwesenheit bekommt „Lindenstraße“ den höchsten Fitnesswert, obwohl er diese Sendung nicht bewertet hat. Die sonstigen Ergebnisse, aufgelistet in Tabelle 6.25, sind plausibel. Die Ähnlichkeit der Sendungen „nano“ und „aspekte kompakt“ schlägt sich in deren Fitnesswerten nieder.

Fitness anhand der Kombination von „Abst“ und „Gew“

Die Ermittlung der Fitness durch Kombination der beiden vorhergehenden Fitnessfunktionen erzielt gute Werte (siehe Tabelle 6.26). Bei der Kompromissuche für beide Anwender würde „Lindenstraße“ klar bevorzugt, „nano“ käme an zweiter und „aspekte kompakt“ an letzter Stelle.

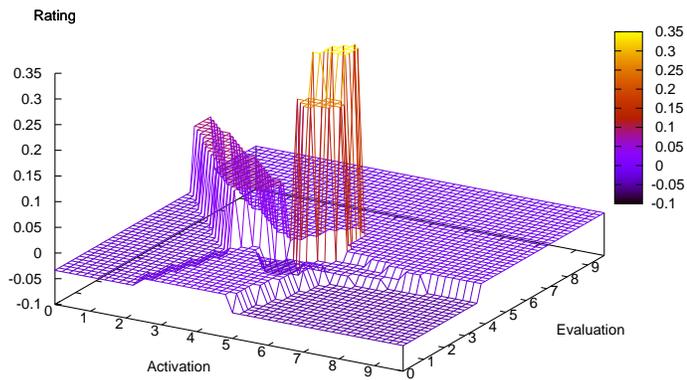
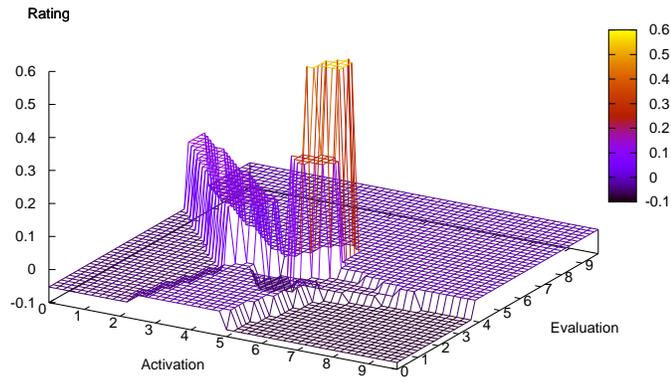
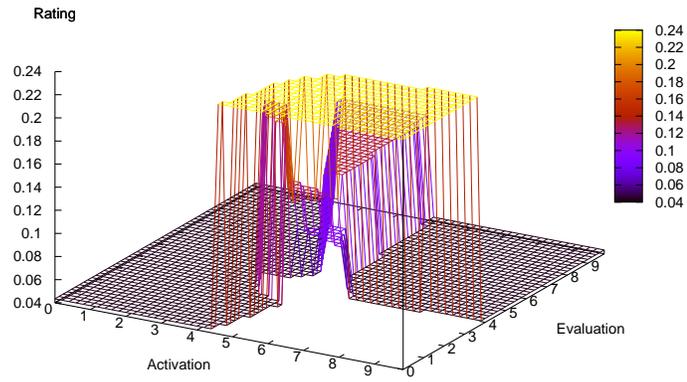


Abbildung 6.14.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-Voronoi).

<i>alt-Voronoi-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-2.13043
aspekte kompakt	0.115564
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.30918
Lindenstraße	3.4378
aspekte kompakt	-3.94445
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-2.42365
Lindenstraße	3.12973
aspekte kompakt	-4.05891

Tabelle 6.24.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-Voronoi-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.413044
aspekte kompakt	0.239131
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	0.010672
Lindenstraße	0.48052
aspekte kompakt	0.00574152
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.0237312
Lindenstraße	0.487792
aspekte kompakt	0.0137032

Tabelle 6.25.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-Voronoi-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-1.71739
aspekte kompakt	0.354694
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.29851
Lindenstraße	3.91832
aspekte kompakt	-3.9387
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-2.39991
Lindenstraße	3.61753
aspekte kompakt	-4.04521

Tabelle 6.26.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-Voronoi-AvgV</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.17467
Lindenstraße	1.04122
aspekte kompakt	1.15061
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0611978
Lindenstraße	0.175714
aspekte kompakt	-0.0995236
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0647717
Lindenstraße	0.184321
aspekte kompakt	-0.103321

Tabelle 6.27.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Voronoi-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.

Fitness anhand der Differenz von Durchschnittsvektoren

Tabelle 6.27 zeigt die Fitnesswerte, die durch die Differenz der Durchschnittsvektoren ermittelt werden. In allen Fällen werden Rangordnungen generiert, die den Erwartungen entsprechen. Die Fitnesswerte der Titel bei Anwesenheit beider Benutzer ist von besonderem Interesse. Die Ergebnisse der individuellen Fitnessermittlungen hängen nur von der Fitnessfunktion durch Differenz der Durchschnittsvektoren ab und sind entsprechend identisch mit denen aus Tabelle 6.19 und Tabelle 6.23. Die ermittelte Rangordnung der Filme für beide Benutzer fällt ebenso positiv aus. „Lindenstraße“ wird als bestmöglicher Kompromiss erkannt und lässt die beiden Titel „nano“ und „aspekte kompakt“ (in dieser Reihenfolge) weit hinter sich. Die beiden Sendungen „nano“ und „aspekte kompakt“ haben in allen Fällen ähnliche Fitnesswerte erhalten.

6.5.3. Inverse Distance Weighted Interpolation

Abbildung 6.15 zeigt die Bewertungen sämtlicher A/E-Koordinaten der Präferenzen und der gemeinsamen Hypothese von Emil und Bert. Genau wie in Abbildung 6.12 des vorigen Unterkapitels vereinen die Graphen die Bewertungen der Merkmale, sowie interpolierte Werte zwischen den bekannten Bewertungen. Die Interpolation am Anfragepunkt erfolgt durch Gewichtung der Stützpunktwerte invers zu den Abständen.

Der Hypothesengraph ergibt sich diesmal aufgrund aller Einträge in Emils und Berts Präferenzen, die sich nicht widersprechen. Die Fitnesswerte für jeweils Emils und Berts Präferenzen gleichen den Ergebnissen aus Unterkapitel 6.4.3. Deshalb werden im Folgenden ausschließlich die Fitnesswerte und die sich daraus ergebenden Rangordnungen der Sendungen im Falle von Emils und Berts Anwesenheit betrachtet.

Fitness anhand des Gesamtabstands

Die Fitnessfunktion durch Berechnung des Gesamtabstands ermittelt, basierend auf der Hypothese für Emil und Bert, akzeptable Werte (siehe Tabelle 6.28). „Lindenstraße“ kommt an erster Stelle und setzt sich deutlich von den Fitnessrängen von „nano“ und „aspekte kompakt“ ab. „nano“ kommt auf den zweiten Platz. „aspekte kompakt“ erhält den schlechtesten Fitnesswert. Durch die bloße gewichtete Übernahme der Präferenzen wurde diese positiv zu bewertende Reihenfolge der Sendungen nicht erreicht (siehe 6.16).

Fitness anhand anteilmäßiger Bewertung der Merkmale

Wie bereits im Unterkapitel 6.4.3 gesehen, produziert die gewichtete Summe der Emotionshäufigkeiten, in Kombination mit interpolierten Werten, keine konsistenten Daten in allen Fällen. In Anwesenheit von Emil und Bert produziert diese Kombination allerdings eine akzeptable Rangordnung der Filmtitel. Durch die alternative Hypothesenbildung hebt sich die Fitness von „Lindenstraße“ als bester Kompromiss noch mehr von den

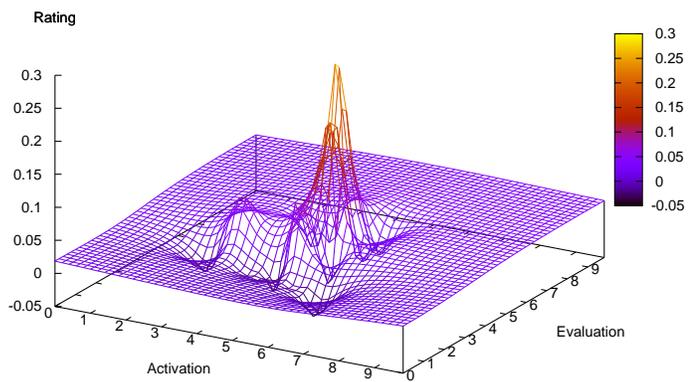
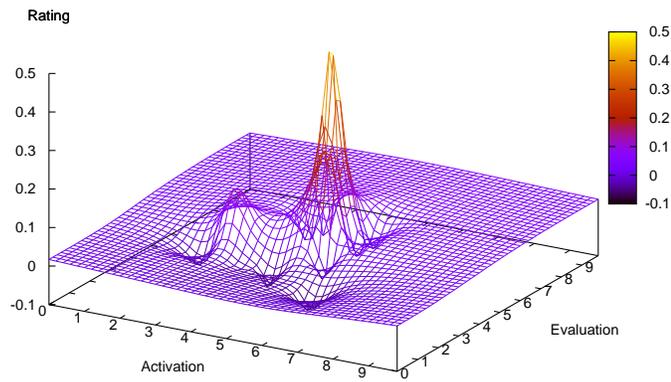
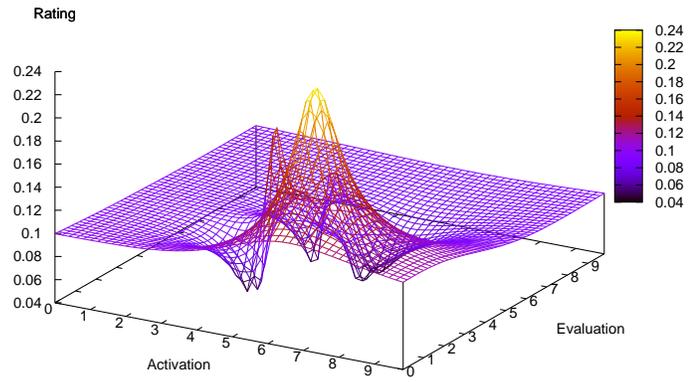


Abbildung 6.15.: Emils und Berts individuelle Präferenzen und die Hypothese, die für beide gebildet wurde (Konfiguration alt-Inverse).

<i>alt-Inverse-Abst</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1
Lindenstraße	-2.27886
aspekte kompakt	0.254128
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.59765
Lindenstraße	3.78917
aspekte kompakt	-4.30605
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-1.62975
Lindenstraße	3.23443
aspekte kompakt	-2.98231

Tabelle 6.28.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Abst) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-Inverse-Gewi</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	0.293005
Lindenstraße	0.441802
aspekte kompakt	0.231562
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	0.00795057
Lindenstraße	0.465914
aspekte kompakt	-0.00274103
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0.046808
Lindenstraße	0.457769
aspekte kompakt	0.0130242

Tabelle 6.29.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Gewi) für verschiedene Eingaben und verschiedenes Publikum produziert.

anderen Titeln ab. Die 2. Stelle belegt „nano“ und den geringsten Fitnesswert erhält „aspekte kompakt“ (siehe 6.29).

Fitness anhand der Kombination von „Abst“ und „Gewi“

Im völligen Gegensatz zu der Kombination aus Hypothesenbildung in Unterkapitel 6.4.3, resultiert die nun erzeugte, alternative Hypothese in guten Werten bei der Kompromiss-suche für Emil und Bert. „nano“ bekommt eine doppelt so hohe Fitness zugeteilt wie „aspekte kompakt“. Beide Titel sind weit von der Fitness von „Lindenstraße“ entfernt (siehe Tabelle 6.30).

Fitness anhand der Differenz von Durchschnittsvektoren

Bereits bei der gewichteten Übernahme der Präferenzen für die Hypothesenbildung ergab sich die gewünschte Fitnessreihenfolge der drei zur Auswahl stehenden Titel mit Fitnessberechnung durch Differenz der Durchschnittsvektoren (siehe Tabelle 6.19). Bei der Anwendung auf die alternative Hypothese ergibt sich die gleiche Reihenfolge: „Lindenstraße“, „nano“ und „aspekte kompakt“. Allerdings unterscheiden sich die errechneten Fitnesswerte nun wesentlich stärker voneinander (siehe Tabelle 6.31).

6.6. Tabellarische Zusammenfassung der Ergebnisse

Die verschiedenen Konfigurationen des Lernsystems haben teilweise sehr unterschiedliche Ergebnisse für die oben beschriebene Situation produziert (siehe Abschnitt 6.3). In Tabelle 6.32 werden die Ergebnisse bezüglich verschiedener Kriterien zusammengefasst. Die Abkürzungen für die Benennung der Spalten werden in Tabelle 6.33 aufgelöst, Tabelle 6.34 erklärt die Bedeutungen der innerhalb der Spalten verwendeten Kürzel.

Im Einzelnen sind die Kriterien:

- Die Rangordnung (oder Reihenfolge) der drei verschiedenen Titel bezüglich ihrer Fitnesswerte
- Die Größe des Intervalls, in dem die Fitnesswerte angesiedelt sind (Wertebereich)
- Die Ähnlichkeit der Fitnesswerte der Sendungen „nano“ und „aspekte kompakt“

Ergibt sich jeweils für die Personen Emil, Bert und für beide gemeinsam die in Abschnitt 6.3 formulierte Rangordnung bezüglich der Fitnesswerte, so wurde das entsprechende Feld (Rf) mit „+“, ansonsten mit „-“ besetzt. Sofern „Lindenstraße“ die höchste Fitness im Falle Emils oder beider Benutzer Anwesenheit erreicht, wird „+“ in die Liste eingetragen. Nimmt dabei „nano“ den zweiten Platz ein, wird „n“ hinzugefügt, wird „aspekte kompakt“ auf an zweiter Stelle der Rangordnung positioniert, wird „a“ hinzugefügt.

<i>alt-Inverse-Komb</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.29301
Lindenstraße	-1.83706
aspekte kompakt	0.48569
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-2.5897
Lindenstraße	4.25509
aspekte kompakt	-4.30879
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-1.58294
Lindenstraße	3.69219
aspekte kompakt	-2.96928

Tabelle 6.30.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-Komb) für verschiedene Eingaben und verschiedenes Publikum produziert.

<i>alt-Inverse-AvgV</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	0
Lindenstraße	0
aspekte kompakt	0
<i>Emil</i>	
<i>Sendung</i>	<i>Bewertung</i>
nano	1
<i>Sendung</i>	<i>Fitness</i>
nano	1.17467
Lindenstraße	1.04122
aspekte kompakt	1.15061
<i>Bert</i>	
<i>Sendung</i>	<i>Bewertung</i>
Lindenstraße	1
aspekte kompakt	-0.5
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0611978
Lindenstraße	0.175714
aspekte kompakt	-0.0995236
<i>Emil & Bert</i>	
<i>Sendung</i>	<i>Fitness</i>
nano	-0.0144096
Lindenstraße	0.144971
aspekte kompakt	-0.0531382

Tabelle 6.31.: Die einzelnen Schritte zeigen, welche Ergebnisse die Konfiguration des Lernsystems (alt-Inverse-AvgV) für verschiedene Eingaben und verschiedenes Publikum produziert.

			Emil			Bert			Emil & Bert		
Hyp.	Int.	Fit.	Rf	Wb	Äk	Rf	Wb	Äk	Rf	Wb	Äk
gew.	„0”	Abst.	+	2.02484	-	+n	4.19773	++	-	0.199502	-
gew.	„0”	Gewi.	+	0.2433156	+	+a	0.5971772	++	+n	0.302265	++
gew.	„0”	Komb.	+	2.268165	-	+n	4.79459	++	+n	0.483134	-
gew.	Voro.	Abst.	+	3.13043	+	+n	7.38225	++	-	4.11204	++
gew.	Voro.	Gewi.	-	0.173913	+	+n	0.47477848	++	+n	0.385905	++
gew.	Voro.	Komb.	+	3.0104	++	+n	7.85702	++	-	3.72613	++
gew.	Inve.	Abst.	+	3.27886	++	+n	8.09522	++	-	4.43501	++
gew.	Inve.	Gewi.	-	0.21024	++	+n	0.46865503	++	+n	0.41558	++
gew.	Inve.	Komb.	+	3.13007	++	+n	8.56388	++	-	4.01944	++
gew.		AvgV.	+	0.13345	++	+n	0.2752376	++	+n	0.0261663	-
alt.	„0”	Abst.	+	2.02484	+	+n	4.19773	++	+n	3.53273	+
alt.	„0”	Gewi.	+	0.2433156	+	+a	0.5971772	++	+n	0.5310907	++
alt.	„0”	Komb.	+	2.268165	+	+n	4.79459	++	+n	4.06381	+
alt.	„0”	AvgV.	+	0.13345	++	+n	0.2752376	++	+n	0.1981092	++
alt.	Voro.	Abst.	+	3.13043	+	+n	7.38225	++	+n	7.18864	++
alt.	Voro.	Gewi.	-	0.173913	+	+n	0.47477848	++	+n	0.4740888	++
alt.	Voro.	Komb.	+	3.0104	++	+n	7.85702	++	+n	7.66274	++
alt.	Voro.	AvgV.	+	0.13345	++	+n	0.2752376	++	+n	0.287642	++
alt.	Inve.	Abst.	+	3.27886	++	+n	8.09522	++	+n	6.21674	++
alt.	Inve.	Gewi.	-	0.21024	++	+n	0.46865503	++	+n	0.4447448	++
alt.	Inve.	Komb.	+	3.13007	++	+n	8.56388	++	+n	6.66147	++
alt.	Inve.	AvgV.	+	0.13345	++	+n	0.2752376	++	+n	0.1981092	++

Tabelle 6.32.: Zusammenfassung der Testergebnisse. Die Benennungen der Spalten sind in Tabelle 6.33, die Abkürzungen innerhalb der verschiedenen Spalten in Tabelle 6.34 erklärt.

Um einen Überblick über die Streuung der Fitnesswerte zu gewinnen, ist die Größe des Intervalls angegeben, das durch den kleinsten und größten Fitnesswert definiert wird.

Die Repräsentationen von „nano” und „aspekte kompakt” haben viele Überschneidungen (siehe Abschnitt 6.3). Da sich die Fitnessberechnungen auf diese Repräsentationen beziehen, sollten ihre Fitnesswerte ebenfalls ähnlich sein. Befinden sich die Fitnesswerte im gleichen Drittel des Wertebereichs, wird „++” in der Spalte „Äk” eingetragen. Befinden sie sich in der gleichen Hälfte des Wertebereichs, wird das mit „+” gekennzeichnet. Falls die Fitnesswerte von „nano” und „aspekte kompakt” noch weiter auseinander liegen (in diesem Fall muss man davon ausgehen, dass die Fitness von mindestens einer der beiden Sendungen näher an der Fitness von „Lindenstraße” liegt), wird „-” in das entsprechende „Äk”-Feld eingetragen.

Kürzel	Bedeutung
Hyp.	Methode der Hypothesenbildung
Int.	Interpolationsfunktion
Fit.	Fitnessfunktion
Rf	Reihenfolge
Äk	Ähnlichkeit
Wb	Wertebereich
Emil	Zusammenfassung der Fitnessbewertungen für Emils Präferenzen
Bert	Zusammenfassung der Fitnessbewertungen für Berts Präferenzen
Emil & Bert	Zusammenfassung der Fitnessbewertungen basierend auf der Hypothese für Emil und Bert

Tabelle 6.33.: Legende der Spaltenkürzel

Kürzel	Spalte	Bedeutung
gew. alt.	Hyp.	gewichtete Übernahme von Präferenzen und Layern alternative Hypothesenbildung
„0” Voro. Inve.	Int.	„0” als Standardrückgabe Voronoi Interpolation der Merkmalsausprägungen Inverse Distance Weighted Interpolation der Merkmalsausprägungen
Abst. Gewi. Komb. AvgV.	Fit.	Gesamtabstand von Hypothese und Filmbeschreibung Anteilmäßige Bewertung der Merkmale Kombination von „Abst” und „Gewi” Differenz der Durchschnittsvektoren
+	Rf	Die Reihenfolge entspricht den Erwartungen
+n		Die Reihenfolge entspricht den Erwartungen, „nano” erhält den zweithöchsten Fitnesswert
+a		Die Reihenfolge entspricht den Erwartungen, „aspekte kompakt” erhält den zweithöchsten Fitnesswert
-		Die Reihenfolge entspricht den Erwartungen nicht
++	Äk	„nano” und „aspekte kompakt” erhalten Fitnesswerte innerhalb des gleichen Drittels des Wertebereichs
+		„nano” und „aspekte kompakt” erhalten Fitnesswerte innerhalb der gleichen Hälfte des Wertebereichs
-		„nano” und „aspekte kompakt” erhalten Fitnesswerte, die die Ähnlichkeit der Sendungen nicht reflektieren

Tabelle 6.34.: Legende der innerhalb der Spalten verwendeten Kürzel

6.7. Testläufe mit Benutzereingaben

Anhand von zwei der vorgestellten Konfigurationen wird gezeigt, welche Vorschläge das Lernsystem für Anwender mit komplexen Präferenzen generiert. Hierbei werden für jeden der Anwender, unabhängig voneinander Filmbewertungen gelernt. Es wird sichergestellt, dass die Benutzer die Bewertungen der jeweils anderen während dieser Phase nicht erfahren. Im Anschluss werden Vorschläge des System für die gesamte Gruppe erwartet, die entsprechend von allen Personen gemeinsam bewertet werden. In dieser Lernphase darf reger Meinungs austausch stattfinden, denn bei der gemeinsamen Bewertung von Titeln muss sich eine Gruppenbewertung manifestieren. Da es zu Widersprüchen der individuell gelernten Präferenzen und der Gruppenbewertungen kommen kann, werden entsprechende Layer konstruiert und im Verlauf des Lernprozesses modifiziert werden.

Das Lernsystem schlägt stets die 5 Titel vor, die die höchste Fitness aufweisen. Solange noch keine Eingaben verarbeitet sind und alle Titel somit die gleiche Fitness haben, sind die Vorschläge des Lernsystem durch deren Reihenfolge im Datenbestand determiniert.

Nach der Bewertung eines Titels wird dieser Titel nicht mehr vorgeschlagen, bis sich die Liste der Anwesenden ändert.

Aus Tabelle 6.32 wurden folgende Konfigurationen für die Testläufe ausgewählt.

gew.	„0”	Gewi.	+	0.2433156	+	+a	0.5971772	++	+n	0.302265	++
alt.	Inve.	AvgV.	+	0.13345	++	+n	0.2752376	++	+n	0.1981092	++

Beide Konfigurationen unterscheiden sich in einem hohen Grad. Außerdem liefern sie in allen erprobten Fällen erwünschte Resultate und eine ähnliche Klassifikation der überschneidenden Repräsentationen von „nano” und „aspekte kompakt”. Bei der ersten Konfiguration wird die Hypothese durch gewichtete Übernahme der Präferenzen und Layer gebildet, im zweiten Fall wird Gemeinsames übernommen und Widersprüchliches ignoriert.

Um die Fitnesswerte im ersten Fall zu berechnen, wird die Häufigkeit eines A/E-Tupels mit der entsprechenden Merkmalsausprägung der Hypothese multipliziert. Ist an der entsprechenden Stelle in der Hypothese oder im Filmdiagramm nichts eingetragen, wird „0” angenommen, weswegen der entsprechende Wert überhaupt nicht in den Fitnesswert der Fernsehsendung hineinspielt. „0” als Standardannahme scheint somit keine große Hilfe zu sein. In den vorhergehenden Tests hat es sich dennoch als einzige Methode zusammen mit der gewichteten Übernahme der Präferenzen zur Hypothesenbildung erwiesen, die die erwünschten Resultate bringt.

Bei der zweiten Konfiguration werden Widersprüche in den Präferenzen durch Vorzeichenvergleich des durch Inverse Distance Weighted Interpolation zurückgelieferten Wertes ermittelt. Alle widerspruchsfreien Einträge und alle negativen Einträge werden übernommen und bilden die neue Hypothese. Die Fitnessfunktion ermittelt durch Vergleich von Durchschnittsvektoren inwiefern eine Fernsehsendung zur gebildeten Hypothese passt.

Alle Verfahren sind detailliert in Kapitel 5 erläutert.

6.7.1. Testlauf „gew-0-Gewi.“ mit zwei Personen

<i>Interaktiver Lernmodus</i>	
Christopher	
Sendung	Fitness
...so alt wie man sich fühlt	0
25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
Sendung	Bewertung
...so alt wie man sich fühlt	-1
Christopher	
Sendung	Fitness
25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
88acht GUTEN TAG BERLIN	0
Sendung	Bewertung
88acht GUTEN TAG BERLIN	-1
Christopher	
Sendung	Fitness
25 Tage in Europa	0
ARD-NACHTKONZERT (IV)	0
ARTE Kultur	0
Bayern 3 - die MorningShow	0
Bilderbuch Deutschland	0
Sendung	Bewertung
ARD-NACHTKONZERT (IV)	1
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	0.666667
Die Sprechstunde	0.333333
Kulturzeit	0.270833
MusicHall	0.275
Verbotene Liebe	0.227778
Sendung	Bewertung

Die Sprechstunde	1
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	0.75
Der Elefant, sein Mahout und der Moloch	0.333333
Kulturzeit	0.390625
MusicHall	0.40625
Verbotene Liebe	0.320833
Sendung	Bewertung
Kulturzeit	1
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	0.842105
Der Elefant, sein Mahout und der Moloch	0.385965
Doppel-Kopf	0.381579
MusicHall	0.484211
Verbotene Liebe	0.378947
Sendung	Bewertung
MusicHall	1
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	0.934579
Der Elefant, sein Mahout und der Moloch	0.442368
Doppel-Kopf	0.441589
Eine Zigeunerkapelle geht um die Welt	0.370717
Verbotene Liebe	0.438629
Sendung	Bewertung
Eine Zigeunerkapelle geht um die Welt	0
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	0.934579
Der Elefant, sein Mahout und der Moloch	0.442368
Doppel-Kopf	0.441589
Verbotene Liebe	0.438629
Waldo	0.369381
Sendung	Bewertung
Verbotene Liebe	-1
Christopher	
Sendung	Fitness

Bilderbuch Deutschland	0.815534
Der Elefant, sein Mahout und der Moloch	0.364617
Die Deutsche Alleenstraße	0.298813
Doppel-Kopf	0.381068
Waldo	0.306056
Sendung	Bewertung
Bilderbuch Deutschland	-1
Christopher	
Sendung	Fitness
Der Elefant, sein Mahout und der Moloch	0.291834
Die Deutsche Alleenstraße	0.210174
Doppel-Kopf	0.35241
Waldo	0.196213
Wildbach	0.161044
Sendung	Bewertung
Der Elefant, sein Mahout und der Moloch	-1
Christopher	
Sendung	Fitness
Die Deutsche Alleenstraße	0.0975255
Doppel-Kopf	0.252183
Sachsen-Anhalt heute	0.089607
Waldo	0.113537
Wildbach	0.0994178
Sendung	Bewertung
Wildbach	0
Christopher	
Sendung	Fitness
Der Stellvertreter	0.0819617
Die Deutsche Alleenstraße	0.0975255
Doppel-Kopf	0.252183
Sachsen-Anhalt heute	0.089607
Waldo	0.113537
Sendung	Bewertung
Sachsen-Anhalt heute	1
Christopher	
Sendung	Fitness
Der Stellvertreter	0.12294
Die Deutsche Alleenstraße	0.133185
Doppel-Kopf	0.276316
Mir geht es gut!	0.121032

Waldo	0.153129
Sendung	Bewertung
Mir geht es gut!	0
Christopher	
Sendung	Fitness
Bianca - Wege zum Glück	0.0980032
Der Stellvertreter	0.12294
Die Deutsche Alleenstraße	0.133185
Doppel-Kopf	0.276316
Waldo	0.153129
Sendung	Bewertung
Bianca - Wege zum Glück	-0
Christopher	
Sendung	Fitness
Der Stellvertreter	0.12294
Die Deutsche Alleenstraße	0.133185
Doppel-Kopf	0.276316
Lebenslinien	0.0898443
Waldo	0.153129
Sendung	Bewertung
Lebenslinien	1
Sebastian	
Sendung	Fitness
Café Trend	0.208825
Doppel-Kopf	0.268956
Dryja & Rosch	0.208825
KULTURRADIO AM MITTAG	0.208825
Radiofeuilleton	0.208825
Sendung	Bewertung
Sebastian	
Sendung	Fitness
...so alt wie man sich fühlt	0
25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
Sendung	Bewertung
...so alt wie man sich fühlt	0.3
Sebastian	
Sendung	Fitness

Geschichten aus dem Zauberhaus	1.5
Glanz der Erde	1.02083
Profile	1
Raketen, nein danke	1.08333
Täter - Opfer - Polizei	1.5
Sendung	Bewertung
Glanz der Erde	0.4
Sebastian	
Sendung	Fitness
Geschichten aus dem Zauberhaus	1.40476
Landesschau Rheinland-Pfalz	0.936508
Profile	0.936508
Raketen, nein danke	1.00397
Täter - Opfer - Polizei	1.40476
Sendung	Bewertung
Geschichten aus dem Zauberhaus	0.5
Sebastian	
Sendung	Fitness
Dreiländernacht	1.10185
Landesschau Rheinland-Pfalz	1.10185
Profile	1.10185
Raketen, nein danke	1.1412
Täter - Opfer - Polizei	1.65278
Sendung	Bewertung
Raketen, nein danke	0.1
Sebastian	
Sendung	Fitness
Dreiländernacht	1.08547
Landesschau Rheinland-Pfalz	1.08547
Profile	1.08547
Täter - Opfer - Polizei	1.62821
Unser blauer Planet	0.950742
Sendung	Bewertung
Unser blauer Planet	0.7
Sebastian	
Sendung	Fitness
Blickpunkt Sport	0.76864
Dreiländernacht	0.975731
Landesschau Rheinland-Pfalz	0.975731
Profile	0.975731

Täter - Opfer - Polizei	1.4636
Sendung	Bewertung
Blickpunkt Sport	-0.5
Sebastian	
Sendung	Fitness
Dreiländernacht	0.873422
Landesschau Rheinland-Pfalz	0.873422
Profile	0.873422
Täter - Opfer - Polizei	1.31013
c't magazin	0.694839
Sendung	Bewertung
c't magazin	0.9
Sebastian	
Sendung	Fitness
Dreiländernacht	0.805821
Landesschau Rheinland-Pfalz	0.805821
Profile	0.805821
Sportplatz	0.61698
Täter - Opfer - Polizei	1.20873
Sendung	Bewertung
Sportplatz	-0.8
Sebastian	
Sendung	Fitness
Dreiländernacht	0.662185
Krakaus junge Helden	0.496639
Landesschau Rheinland-Pfalz	0.592312
Profile	0.592312
Täter - Opfer - Polizei	0.993277
Sendung	Bewertung
Dreiländernacht	0.1
Sebastian	
Sendung	Fitness
Krakaus junge Helden	0.503069
Landesschau Rheinland-Pfalz	0.603528
Profile	0.603528
SWR2 Musik Spezial	0.503069
Täter - Opfer - Polizei	1.00614
Sendung	Bewertung
SWR2 Musik Spezial	-1
Sebastian	

Sendung	Fitness
Fernsehen verboten!	0.396086
Joe und Max - Rivalen im Ring	0.396086
N-JOY MORNINGSHOW	0.396086
SWR2 Wissen	0.363244
Täter - Opfer - Polizei	0.627965
Sendung	Bewertung
SWR2 Wissen	0.6
Sebastian Christopher	
Sendung	Fitness
Fernsehen verboten!	0.544761
Joe und Max - Rivalen im Ring	0.544761
LexiTV - Wissen für alle	0.357181
N-JOY MORNINGSHOW	0.544761
Täter - Opfer - Polizei	0.585805
Sendung	Bewertung
Sebastian Christopher	
Sendung	Fitness
Fernsehen verboten!	0.324742
Joe und Max - Rivalen im Ring	0.324742
N-JOY MORNINGSHOW	0.324742
SWR2 Wissen	0.252537
c't magazin	0.275659
Sendung	Bewertung
SWR2 Wissen	1
Sebastian Christopher	
Sendung	Fitness
Fernsehen verboten!	1.14069
Joe und Max - Rivalen im Ring	1.14069
LexiTV - Wissen für alle	0.758309
N-JOY MORNINGSHOW	1.14069
c't magazin	0.76734
Sendung	Bewertung
LexiTV - Wissen für alle	0.8
Sebastian Christopher	
Sendung	Fitness
Fernsehen verboten!	0.964687
Guten Abend Sachsen-Anhalt	0.503152
Joe und Max - Rivalen im Ring	0.964687
N-JOY MORNINGSHOW	0.964687

c't magazin	0.634921
Sendung	Bewertung
Fernsehen verboten!	0
Sebastian Christopher	
Sendung	Fitness
Guten Abend Sachsen-Anhalt	0.503152
Joe und Max - Rivalen im Ring	0.964687
N-JOY MORNINGSHOW	0.964687
Spielräume	0.482343
c't magazin	0.634921
Sendung	Bewertung
c't magazin	0.5
Sebastian Christopher	
Sendung	Fitness
Geschichten aus dem Zauberhaus	0.511197
Guten Abend Sachsen-Anhalt	0.440977
Joe und Max - Rivalen im Ring	0.842318
N-JOY MORNINGSHOW	0.842318
Täter - Opfer - Polizei	0.511197
Sendung	Bewertung
Täter - Opfer - Polizei	-1
Sebastian Christopher	
Sendung	Fitness
Guten Abend Sachsen-Anhalt	0.447874
Joe und Max - Rivalen im Ring	0.857426
N-JOY MORNINGSHOW	0.857426
SWR2 Thema Musik	0.428713
Spielräume	0.428713
Sendung	Bewertung
N-JOY MORNINGSHOW	-1
Sebastian Christopher	
Sendung	Fitness
DAS! ab 2	0.407971
Dir gehört mein Leben	0.407971
Lebensunwert	0.407971
Total Genial	0.407971
höchstpersönlich	0.407971
Sendung	Bewertung
Total Genial	-1
Sebastian Christopher	

Sendung	Fitness
DAS! ab 2	0.259644
Dir gehört mein Leben	0.259644
Lebensunwert	0.259644
Tagesgespräch	0.131655
höchstpersönlich	0.259644
Sendung	Bewertung
Tagesgespräch	0.7
Sebastian Christopher	
Sendung	Fitness
DAS! ab 2	0.46467
Der Letzte seines Standes?	0.428995
Dir gehört mein Leben	0.46467
Lebensunwert	0.46467
höchstpersönlich	0.46467
Sendung	Bewertung
DAS! ab 2	-1
Sebastian Christopher	
Sendung	Fitness
Bayern 3 - die MorningShow	0.453644
Bremen Eins am Abend	0.453644
Der Letzte seines Standes?	0.453644
Dir gehört mein Leben	0.358807
Lebensunwert	0.358807
Sendung	Bewertung
Bremen Eins am Abend	0
Sebastian Christopher	
Sendung	Fitness
Bayern 3 - die MorningShow	0.453644
Der Letzte seines Standes?	0.453644
Dir gehört mein Leben	0.358807
Lebensunwert	0.358807
höchstpersönlich	0.358807
Sendung	Bewertung
Bayern 3 - die MorningShow	-0.3
Sebastian Christopher	
Sendung	Fitness
Dir gehört mein Leben	0.487357
Lebensunwert	0.487357
Lindenstraße	0.241057

Pavillon	0.253712
höchstpersönlich	0.487357
Sendung	Bewertung
Lindenstraße	-1
Sebastian Christopher	
Sendung	Fitness
Der Letzte seines Standes?	0.110126
Dir gehört mein Leben	0.260607
Lebensunwert	0.260607
Nachtmagazin	0.107025
höchstpersönlich	0.260607
Sendung	Bewertung
Nachtmagazin	0.5
Sebastian Christopher	
Sendung	Fitness
Bremen Vier: Rocken, poppen, dancen	0.269976
Das Forum der Europäer	0.217118
Dir gehört mein Leben	0.362972
Lebensunwert	0.362972
höchstpersönlich	0.362972
Sendung	Bewertung
Bremen Vier: Rocken, poppen, dancen	-1
Sebastian Christopher	
Sendung	Fitness
Das Forum der Europäer	0.181661
Dir gehört mein Leben	0.350107
Lebensunwert	0.350107
Teletubbies	0.189646
höchstpersönlich	0.350107
Sendung	Bewertung
Teletubbies	1
Sebastian Christopher	
Sendung	Fitness
Das Forum der Europäer	0.284751
Der Wunschkuss	0.270399
Dir gehört mein Leben	0.486044
Lebensunwert	0.486044
höchstpersönlich	0.486044
Sendung	Bewertung
Der Wunschkuss	0.6

Sebastian Christopher	
Sendung	Fitness
Das Forum der Europäer	0.339623
Dir gehört mein Leben	0.568507
Lebensunwert	0.568507
Pavillon	0.30513
höchstpersönlich	0.568507
Sendung	Bewertung
Dir gehört mein Leben	-1

Sebastian Christopher	
Sendung	Fitness
Das Forum der Europäer	0.326119
Großstadtrevier	0.310457
Lebensunwert	0.490562
N-JOY	0.310457
höchstpersönlich	0.490562
Sendung	Bewertung
Großstadtrevier	-0.8

Sebastian Christopher	
Sendung	Fitness
Das Forum der Europäer	0.296694
Lebensunwert	0.548928
Länder - Menschen - Abenteuer	0.280844
Pavillon	0.295696
höchstpersönlich	0.548928
Sendung	Bewertung
Länder - Menschen - Abenteuer	1

Sebastian Christopher	
Sendung	Fitness
88acht HEY MUSIC	0.454072
88acht NACHTISCH	0.454072
88acht WECKER	0.454072
Lebensunwert	0.536614
höchstpersönlich	0.536614
Sendung	Bewertung

Tabelle 6.35.: Testlauf der Konfiguration „gew-0-Gewi“ mit bis zu zwei Personen.

Zunächst bewertet Christopher 15 verschiedene Titel. Danach ist Sebastian anwesend, der wiederum 12 Sendungen bewertet. Im Anschluss generiert das Lernsystem Vor-

schläge, die beiden Benutzern gefallen könnten. Die Gruppe, bestehend aus Christopher und Sebastian, bewertet anschließend 20 Sendungen. Die vom System generierten Vorschläge samt Fitnesswerte und die Bewertungen der Benutzer können im Einzelnen in Tabelle 6.35 eingesehen werden.

Folgende vorläufige Bemerkungen müssen festgehalten werden:

- Für den Testlauf konnten sich die Anwesenden weder aufgrund der betitelten Fernsehsendung, noch aufgrund einer Beschreibung ein Urteil bilden. Ihre Bewertungen wurden vom Titel der Sendungen motiviert.
- Christopher hat ausschließlich von den Bewertungen -1 , 0 und $+1$ Gebrauch gemacht. Das Mapping eines bekannten Bewertungsschemas (zum Beispiel nach dem Schulnotensystem) ist sicherlich hilfreich, um eine bessere Differenzierung bezüglich der Beurteilungen zu erzielen.
- Sebastian beschränkt sich bezüglich der Differenzierung seiner Beurteilungen auf die erste Nachkommastelle und beurteilt im Gegensatz zu Christopher keine einzige Sendung mit dem neutralen Wert 0 .
- In den Gruppenbewertungen spiegeln sich beide Bewertungsschemata. Es kommen oft die Extrema -1 und 1 , der neutrale Wert 0 und einige Male Bewertungen vor, die in der ersten Nachkommastelle ihre Geltung erlangen.

Abbildungen 6.16 zeigt die individuellen Präferenzen der Benutzer Christopher und Sebastian, sowie die durch deren Übernahme erzeugte Hypothese.

In Tabelle 6.35 werden die Layer nicht erwähnt, die zwischenzeitlich vom Lernsystem erzeugt wurden, um Widersprüche zwischen individuellen und Gruppenpräferenzen auszugleichen. Nach der letzten Gruppenbewertung ergibt sich die Gesamtsituation aus Abbildung 6.17.

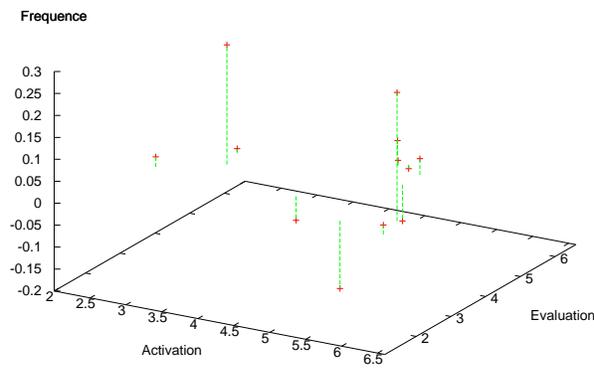
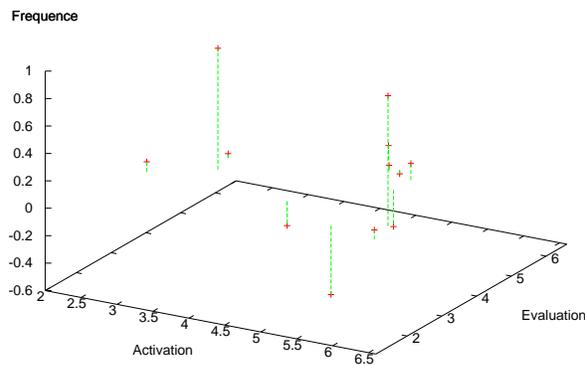
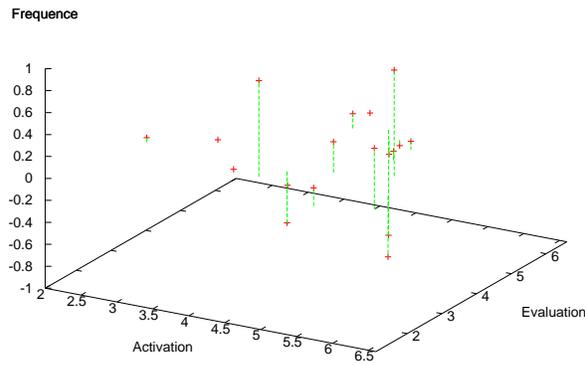


Abbildung 6.16.: Christophers Präferenzen (nach 15 Bewertungen), Sebastians Präferenzen (nach 12 Bewertungen) und die Hypothese durch deren Übernahme.

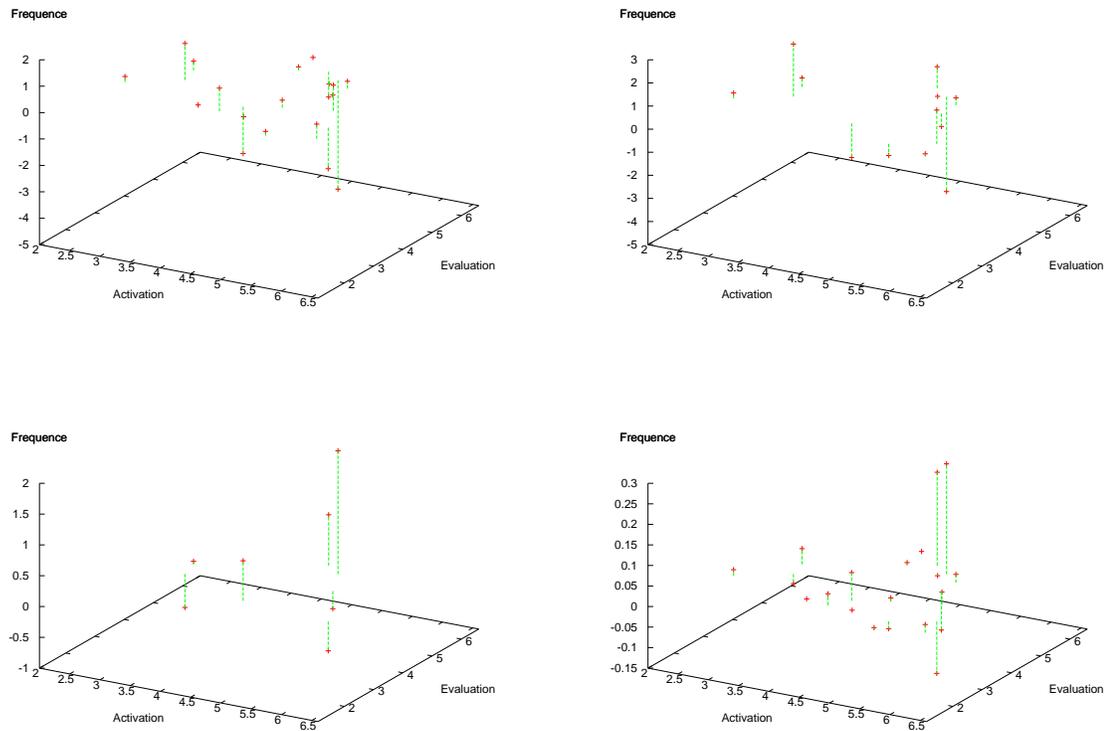


Abbildung 6.17.: Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher und Sebastian, den Layer zwischen den beiden Benutzern und die gebildete Hypothese nach der letzten gemeinsamen Bewertung.

6.7.2. Testlauf „alt-Inverse-AvgV“ mit drei Personen

<i>Interaktiver Lernmodus</i>	
Christopher	
Sendung	Fitness
...so alt wie man sich fühlt	0
25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
Sendung	Bewertung
88acht GUTEN ABEND BERLIN	-1
Christopher	
Sendung	Fitness
Café Trend	-0.690319
Dryja & Rosch	-0.690319
KULTURRADIO AM MITTAG	-0.690319
Radiofeuilleton	-0.690319
Resonanzen	-0.690319
Sendung	Bewertung
KULTURRADIO AM MITTAG	1
Christopher	
Sendung	Fitness
Café Trend	-0.690319
Dryja & Rosch	-0.690319
Radiofeuilleton	-0.690319
Resonanzen	-0.690319
SWR2 RadioART: Essay	-0.690319
Sendung	Bewertung
Radiofeuilleton	1
Christopher	
Sendung	Fitness
Café Trend	-0.690319
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
SWR2 RadioART: Essay	-0.690319
Serenade	-0.690319
Sendung	Bewertung
Café Trend	-1
Christopher	

Sendung	Fitness
Dryja & Rosch	-0.690319
Geschichten aus dem Zauberhaus	-0.767715
Resonanzen	-0.690319
SWR2 RadioART: Essay	-0.690319
Serenade	-0.690319
Sendung	Bewertung
Geschichten aus dem Zauberhaus	0
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
SWR2 RadioART: Essay	-0.690319
Serenade	-0.690319
Täter - Opfer - Polizei	-0.767715
Sendung	Bewertung
Täter - Opfer - Polizei	-0.5
Christopher	
Sendung	Fitness
Bilderbuch Deutschland	-0.829626
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
Bilderbuch Deutschland	0
Christopher	
Sendung	Fitness
ARD-NACHTKONZERT (IV)	-0.829626
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
ARD-NACHTKONZERT (IV)	1
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
NDR Sinfonieorchester	-0.833513
Resonanzen	-0.679486

SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
NDR Sinfonieorchester	1
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Radiomarkt	-0.833513
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
Radiomarkt	0.5
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
N-JOY	-0.833513
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
N-JOY	-0.75
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Großstadtrevier	-0.833513
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Sendung	Bewertung
Großstadtrevier	-1
Christopher	
Sendung	Fitness
Die Sendung mit der Maus	-0.797752
Dryja & Rosch	-0.653005
Resonanzen	-0.653005
SWR2 RadioART: Essay	-0.653005
Serenade	-0.653005
Sendung	Bewertung
Die Sendung mit der Maus	1

Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
Tibetisch heilen im Himalaya	-0.833513
Sendung	Bewertung
Tibetisch heilen im Himalaya	0

Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
Serenade	-0.679486
VOR ORT	-0.833513
Sendung	Bewertung
Serenade	0

Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 RadioART: Essay	-0.679486
VOR ORT	-0.833513
Viertel vor ...	-0.833513
Sendung	Bewertung
SWR2 RadioART: Essay	0

Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
VOR ORT	-0.833513
Viertel vor ...	-0.833513
WDR 3 Konzert	-0.833513
Sendung	Bewertung
WDR 3 Konzert	1

Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486

VOR ORT	-0.833513
Viertel vor ...	-0.833513
WDR 3 PHON	-0.833513
Sendung	Bewertung
WDR 3 PHON	1
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
VOR ORT	-0.833513
Viertel vor ...	-0.833513
rbb regional	-0.833513
Sendung	Bewertung
VOR ORT	-0.33
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
Viertel vor ...	-0.833513
rbb regional	-0.833513
rbb um 6	-0.833513
Sendung	Bewertung
rbb um 6	-1
Christopher	
Sendung	Fitness
Dryja & Rosch	-0.679486
Resonanzen	-0.679486
SWR2 Musik Spezial	-0.849344
Viertel vor ...	-0.833513
rbb regional	-0.833513
Sendung	Bewertung
SWR2 Musik Spezial	1
Christopher	
Sendung	Fitness
Daten der Woche	-0.834288
Der Stellvertreter	-0.865453
Die Deutsche Alleenstraße	-0.834885
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
Sendung	Bewertung

Der Stellvertreter	1
Christopher	
Sendung	Fitness
Daten der Woche	-0.834288
Die Deutsche Alleenstraße	-0.834885
Dryja & Rosch	-0.690319
FAKT	-0.866337
Resonanzen	-0.690319
Sendung	Bewertung
FAKT	0
Christopher	
Sendung	Fitness
Daten der Woche	-0.834288
Die Deutsche Alleenstraße	-0.834885
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
Treffpunkt Sport	-0.866337
Sendung	Bewertung
Treffpunkt Sport	-0.5
Christopher	
Sendung	Fitness
Bus 174 - Geiseldrama in Rio	-0.870644
Daten der Woche	-0.834288
Die Deutsche Alleenstraße	-0.834885
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
Sendung	Bewertung
Bus 174 - Geiseldrama in Rio	1
Regina	
Sendung	Fitness
Daten der Woche	-0.834288
Die Deutsche Alleenstraße	-0.834885
Dryja & Rosch	-0.690319
Resonanzen	-0.690319
Viertel vor ...	-0.872373
Sendung	Bewertung
Regina	
Sendung	Fitness
...so alt wie man sich fühlt	0

25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
Sendung	Bewertung
88acht GUTEN MORGEN BERLIN	0.5
Regina	
Sendung	Fitness
88acht GUTEN TAG BERLIN	1.37603
88acht HEY MUSIC	1.37603
88acht NACHTISCH	1.37603
88acht WECKER	1.37603
Auf ein Wort	1.37603
Sendung	Bewertung
Auf ein Wort	-0.7
Regina	
Sendung	Fitness
Der Stellvertreter	-0.55934
Die Deutsche Alleenstraße	-0.552242
Hol den Doc	-0.551253
Hund oder Katze - Wer ist klüger?	-0.555631
Sachsen-Anhalt heute	-0.568463
Sendung	Bewertung
Hol den Doc	0.5
Regina	
Sendung	Fitness
Bianca - Wege zum Glück	-0.570146
Der Stellvertreter	-0.55934
Die Deutsche Alleenstraße	-0.552242
Hund oder Katze - Wer ist klüger?	-0.555631
Sachsen-Anhalt heute	-0.568463
Sendung	Bewertung
Bianca - Wege zum Glück	-1
Regina	
Sendung	Fitness
Café Trend	-0.591946
Dryja & Rosch	-0.591946
KULTURRADIO AM MITTAG	-0.591946
Radiofeuilleton	-0.591946
Resonanzen	-0.591946

Sendung	Bewertung
KULTURRADIO AM MITTAG	0.3
Regina	
Sendung	Fitness
Café Trend	-0.588508
Dryja & Rosch	-0.588508
Radiofeuilleton	-0.588508
Resonanzen	-0.588508
SWR2 RadioART: Essay	-0.588508
Sendung	Bewertung
SWR2 RadioART: Essay	-0.1
Regina	
Sendung	Fitness
Café Trend	-0.588508
Dryja & Rosch	-0.588508
Radiofeuilleton	-0.588508
Resonanzen	-0.588508
Serenade	-0.588508
Sendung	Bewertung
Serenade	0
Regina	
Sendung	Fitness
Café Trend	-0.588508
Dryja & Rosch	-0.588508
Geschichten aus dem Zauberhaus	-0.646875
Radiofeuilleton	-0.588508
Resonanzen	-0.588508
Sendung	Bewertung
Café Trend	-0.2
Regina	
Sendung	Fitness
Dryja & Rosch	-0.591946
Geschichten aus dem Zauberhaus	-0.647372
Radiofeuilleton	-0.591946
Resonanzen	-0.591946
Täter - Opfer - Polizei	-0.647372
Sendung	Bewertung
Täter - Opfer - Polizei	0.8
Regina	

Sendung	Fitness
Bilderbuch Deutschland	-0.718702
Dryja & Rosch	-0.591728
Geschichten aus dem Zauberhaus	-0.649327
Radiofeuilleton	-0.591728
Resonanzen	-0.591728
Sendung	Bewertung
Bilderbuch Deutschland	0.4
Regina	
Sendung	Fitness
ARD-NACHTKONZERT (IV)	-0.726452
Dryja & Rosch	-0.596281
Geschichten aus dem Zauberhaus	-0.655219
Radiofeuilleton	-0.596281
Resonanzen	-0.596281
Sendung	Bewertung
ARD-NACHTKONZERT (IV)	0
Regina	
Sendung	Fitness
Dryja & Rosch	-0.596281
Geschichten aus dem Zauberhaus	-0.655219
N-JOY	-0.728486
Radiofeuilleton	-0.596281
Resonanzen	-0.596281
Sendung	Bewertung
N-JOY	-0.1
Regina	
Sendung	Fitness
Dryja & Rosch	-0.596281
Geschichten aus dem Zauberhaus	-0.655219
NDR Sinfonieorchester	-0.728486
Radiofeuilleton	-0.596281
Resonanzen	-0.596281
Sendung	Bewertung
NDR Sinfonieorchester	0
Regina	
Sendung	Fitness
Dryja & Rosch	-0.596281
Geschichten aus dem Zauberhaus	-0.655219
Großstadtrevier	-0.728486

Radiofeuilleton	-0.596281
Resonanzen	-0.596281
Sendung	Bewertung
Geschichten aus dem Zauberhaus	0.3
Regina	
Sendung	Fitness
Dryja & Rosch	-0.596281
Großstadtrevier	-0.728486
Radiofeuilleton	-0.596281
Radiomarkt	-0.728486
Resonanzen	-0.596281
Sendung	Bewertung
Dryja & Rosch	-1
Regina	
Sendung	Fitness
Die Sendung mit der Maus	-0.728486
Großstadtrevier	-0.728486
Radiofeuilleton	-0.596281
Radiomarkt	-0.728486
Resonanzen	-0.596281
Sendung	Bewertung
Die Sendung mit der Maus	1
Sebastian	
Sendung	Fitness
Großstadtrevier	-0.739054
Radiofeuilleton	-0.603869
Radiomarkt	-0.739054
Resonanzen	-0.603869
Tibetisch heilen im Himalaya	-0.739054
Sendung	Bewertung
Sebastian	
Sendung	Fitness
...so alt wie man sich fühlt	0
25 Tage in Europa	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
Sendung	Bewertung
25 Tage in Europa	0.8
Sebastian	

Sendung	Fitness
ARTE Kultur	1.32988
Fazit	1.32988
Fofteihn	1.32988
Journal am Vormittag	1.32988
KULTURRADIO AM MORGEN	1.32988
Sendung	Bewertung
ARTE Kultur	0.7
Sebastian	
Sendung	Fitness
Fazit	1.32988
Fofteihn	1.32988
Journal am Vormittag	1.32988
KULTURRADIO AM MORGEN	1.32988
Klassisch in den Tag	1.32988
Sendung	Bewertung
Fazit	0.5
Sebastian	
Sendung	Fitness
Fofteihn	1.32988
Journal am Vormittag	1.32988
KULTURRADIO AM MORGEN	1.32988
Klassisch in den Tag	1.32988
Klassisch unterwegs	1.32988
Sendung	Bewertung
Fofteihn	0
Sebastian	
Sendung	Fitness
Journal am Vormittag	1.32988
KULTURRADIO AM MORGEN	1.32988
Klassisch in den Tag	1.32988
Klassisch unterwegs	1.32988
Marc Pircher: Mein Zillertal	1.32988
Sendung	Bewertung
Marc Pircher: Mein Zillertal	-1
Sebastian	
Sendung	Fitness
Journal am Vormittag	1.32988
KULTURRADIO AM MORGEN	1.32988
Klassisch in den Tag	1.32988

Klassisch unterwegs	1.32988
Matinee	1.32988
Sendung	Bewertung
Matinee	-1
Sebastian	
Sendung	Fitness
...so alt wie man sich fühlt	0
88acht BEROLINA	0
88acht GUTEN ABEND BERLIN	0
88acht GUTEN MORGEN BERLIN	0
88acht GUTEN TAG BERLIN	0
Sendung	Bewertung
...so alt wie man sich fühlt	0.1
Sebastian	
Sendung	Fitness
SWR2 Wissen	1.19294
Service: Auto und Verkehr	1.2024
Service: Natur und Technik	1.20376
Von Binnenland und Waterkant	1.19166
der TAG	1.19155
Sendung	Bewertung
Service: Natur und Technik	1
Sebastian	
Sendung	Fitness
Gottes verlassene Häuser	1.18258
Im Tal der Wetterpropheten	1.18209
Ninotschka	1.18636
Videolexikon	1.1719
Zapping International	1.17616
Sendung	Bewertung
Im Tal der Wetterpropheten	1
Sebastian	
Sendung	Fitness
Die Abendschau	1.17691
Flüssiges Gold	1.18384
LILIPUZ - Radio für Kinder	1.18423
Planet Wissen	1.19294
Polizeiruf 110: Um Kopf und Kragen	1.18571
Sendung	Bewertung
Planet Wissen	1

Sebastian	
Sendung	Fitness
Brisant	1.16529
Flüssiges Gold	1.16018
LILIPUZ - Radio für Kinder	1.16639
Lola	1.18447
Polizeiruf 110: Um Kopf und Kragen	1.1803
Sendung	Bewertung
Brisant	0.3

Sebastian	
Sendung	Fitness
Flüssiges Gold	1.1499
LILIPUZ - Radio für Kinder	1.15558
Lola	1.17904
Polizeiruf 110: Um Kopf und Kragen	1.16964
Sturm der Liebe - Wie alles begann	1.14897
Sendung	Bewertung
Sturm der Liebe - Wie alles begann	-0.3

Sebastian	
Sendung	Fitness
Flüssiges Gold	1.1499
LILIPUZ - Radio für Kinder	1.15558
Lola	1.17904
Polizeiruf 110: Um Kopf und Kragen	1.16964
Servicezeit: Essen & Trinken	1.14789
Sendung	Bewertung
Flüssiges Gold	0.1

Sebastian	
Sendung	Fitness
Grüße und Musik	1.14716
LILIPUZ - Radio für Kinder	1.15558
Lola	1.17904
Polizeiruf 110: Um Kopf und Kragen	1.16964
Servicezeit: Essen & Trinken	1.14789
Sendung	Bewertung
Lola	0.1

Sebastian	
Sendung	Fitness
Die Abendschau	1.13633
Grüße und Musik	1.14002

LILIPUZ - Radio für Kinder	1.14689
Polizeiruf 110: Um Kopf und Kragen	1.16108
Servicezeit: Essen & Trinken	1.13971
Sendung	Bewertung
LILIPUZ - Radio für Kinder	-1
Sebastian	
Sendung	Fitness
Bianca - Wege zum Glück	0.084892
Björn Bär	0.0845997
Hol den Doc	0.0848769
Sachsen-Anhalt heute	0.0846707
Sesamstraße	0.0842381
Sendung	Bewertung
Sachsen-Anhalt heute	-1
Sebastian	
Sendung	Fitness
Grüße und Musik	0.0537215
Herbst in der Rhön	0.0548305
Hier ab vier	0.0498338
Polizeiruf 110: Um Kopf und Kragen	0.0497694
Willis Quiz Quark Club	0.0538565
Sendung	Bewertung
Willis Quiz Quark Club	0.1
Christopher Regina Sebastian	
Sendung	Fitness
Grüße und Musik	0.0552182
Herbst in der Rhön	0.0562869
Hier ab vier	0.0514769
Polizeiruf 110: Um Kopf und Kragen	0.0513434
Thema Europa	0.05105
Sendung	Bewertung
Christopher Regina Sebastian	
Sendung	Fitness
Café Trend	-0.562098
Dryja & Rosch	-0.562098
KULTURRADIO AM MITTAG	-0.562098
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
Sendung	Bewertung
Dryja & Rosch	-0.1

Christopher Regina Sebastian	
Sendung	Fitness
Café Trend	-0.562098
KULTURRADIO AM MITTAG	-0.562098
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Sendung	Bewertung
Café Trend	-0.8

Christopher Regina Sebastian	
Sendung	Fitness
KULTURRADIO AM MITTAG	-0.562098
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Sendung	Bewertung
KULTURRADIO AM MITTAG	0.8279

Christopher Regina Sebastian	
Sendung	Fitness
Geschichten aus dem Zauberhaus	-0.619486
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Sendung	Bewertung
Geschichten aus dem Zauberhaus	0.8738

Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Täter - Opfer - Polizei	-0.619486
Sendung	Bewertung
Täter - Opfer - Polizei	35

Christopher Regina Sebastian	
Sendung	Fitness
Bilderbuch Deutschland	-0.689239
Radiofeuilleton	-0.562098

Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Sendung	Bewertung
Bilderbuch Deutschland	0.5
Christopher Regina Sebastian	
Sendung	Fitness
ARD-NACHTKONZERT (IV)	-0.689239
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Sendung	Bewertung
ARD-NACHTKONZERT (IV)	0.7
Christopher Regina Sebastian	
Sendung	Fitness
NDR Sinfonieorchester	-0.691081
Radiofeuilleton	-0.562098
Resonanzen	-0.562098
SWR2 RadioART: Essay	-0.562098
Serenade	-0.562098
Sendung	Bewertung
NDR Sinfonieorchester	-1
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
SWR2 RadioART: Essay	-0.560094
Serenade	-0.560094
Sendung	Bewertung
SWR2 RadioART: Essay	0.4
Christopher Regina Sebastian	
Sendung	Fitness
N-JOY	-0.687937
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung

N-JOY	-0.9
Christopher Regina Sebastian	
Sendung	Fitness
Großstadtrevier	-0.687937
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Großstadtrevier	0
Christopher Regina Sebastian	
Sendung	Fitness
Die Sendung mit der Maus	-0.687937
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Die Sendung mit der Maus	1
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
Serenade	-0.560094
Tibetisch heilen im Himalaya	-0.687937
Sendung	Bewertung
Tibetisch heilen im Himalaya	0.6
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Radiomarkt	-0.687937
Resonanzen	-0.560094
Serenade	-0.560094
VOR ORT	-0.687937
Sendung	Bewertung
Radiomarkt	0.3
Christopher Regina Sebastian	
Sendung	Fitness

Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
VOR ORT	-0.687937
Viertel vor ...	-0.687937
Sendung	Bewertung
VOR ORT	-1
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Viertel vor ...	-0.687937
WDR 3 Konzert	-0.687937
Sendung	Bewertung
Viertel vor ...	-1
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
WDR 3 Konzert	-0.687937
WDR 3 PHON	-0.687937
Sendung	Bewertung
WDR 3 PHON	0.7
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
WDR 3 Konzert	-0.687937
rbb regional	-0.687937
Sendung	Bewertung
rbb regional	-3
Christopher Regina Sebastian	
Sendung	Fitness
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
WDR 3 Konzert	-0.687937

rbb um 6	-0.687937
Sendung	Bewertung
rbb um 6	-1
Christopher Regina Sebastian	
Sendung	Fitness
Goldene Hits	-0.721187
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
WDR 3 Konzert	-0.687937
Sendung	Bewertung
Goldene Hits	-1
Christopher Regina Sebastian	
Sendung	Fitness
Journal am Vormittag	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
WDR 3 Konzert	-0.687937
Sendung	Bewertung
WDR 3 Konzert	0.5
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
Journal am Vormittag	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Journal am Vormittag	-0.3
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
KULTURRADIO AM MORGEN	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
KULTURRADIO AM MORGEN	1
Christopher Regina Sebastian	

Sendung	Fitness
Fofteihn	-0.729406
Klassisch in den Tag	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Klassisch in den Tag	0
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
Klassisch unterwegs	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Klassisch unterwegs	0.1
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
Marc Pircher: Mein Zillertal	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Marc Pircher: Mein Zillertal	-1
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
Matinee	-0.729406
Radiofeuilleton	-0.560094
Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung
Matinee	0.7
Christopher Regina Sebastian	
Sendung	Fitness
Fofteihn	-0.729406
Mikado	-0.729406
Radiofeuilleton	-0.560094

Resonanzen	-0.560094
Serenade	-0.560094
Sendung	Bewertung

Tabelle 6.36.: Testlauf der Konfiguration „alt-Inve-AvgV“ mit bis zu drei Personen.

Tabelle 6.36 zeigt den Testlauf für die zweite ausgewählte Konfiguration: alternative Hypothesenbildung, Inverse Distance Weighted Interpolation und Fitnessberechnung durch Differenz der Durchschnittsvektoren. Zunächst formen die Benutzer ihre individuellen Präferenzen ohne Anwesenheit der anderen. Christopher bewertet als erster (25 Bewertungen), es folgen Regina (17 Bewertungen) und Sebastian (18 Bewertungen).

Die Abbildung 6.18 zeigt die individuellen Präferenzen der drei Personen Christopher, Regina und Sebastian, bevor sie sich gemeinsam von dem Lernsystem Vorschläge unterbreiten lassen, beziehungsweise bevor sie dem Lernsystem ihre Bewertungen als Gruppe mitteilen. Reginas Präferenzen scheinen nicht den Grad an Differenziertheit aufweisen, wie die von Christopher oder Sebastian. Die wesentlich homogener wirkenden Diagramme ihrer Präferenzen ergeben sich allerdings sicherlich auch deshalb, weil Regina vergleichsweise oft eine neutrale Wertung (0) vergeben hat.

Abbildung 6.19 zeigt die Präferenzen der drei Anwender nach dem abgeschlossenen Lernprozess. Es fällt auf, dass die Differenziertheit der ursprünglichen Präferenzen verloren gegangen ist. Dadurch, dass das Lernsystem die individuellen Präferenzen nur erweitert, falls das Gruppenurteil ihnen nicht widerspricht, werden sie während der Gruppenlernphase einseitig weiterentwickelt.

Während der Gruppenlernphase wurden drei Layer geschaffen, und zwar für jeweils zwei paarweise verschiedene Personen. Es entstanden die Layer aus Abbildung 6.20. Die mit der Zeit immer ähnlicher gewordenen Präferenzen der drei Anwender bedingen, dass sich auch die Layer einander sehr ähneln.

Die Aus den Präferenzen und Layern generierte Hypothese ist in Abbildung 6.21 zu sehen. Während des Testlaufs wurde häufig eine negative Bewertung vergeben, was auch insgesamt zu einer einseitig negativen Hypothese führt.

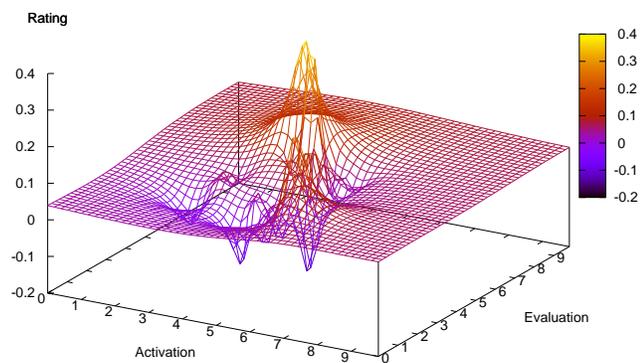
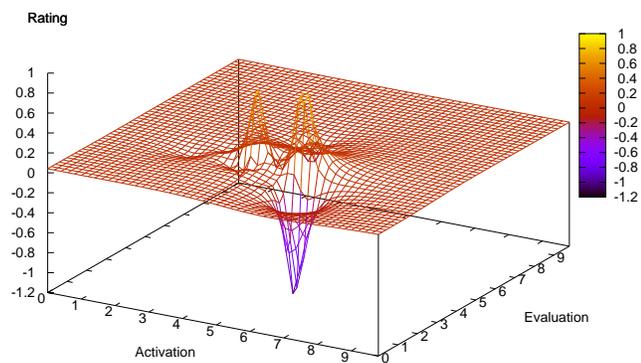
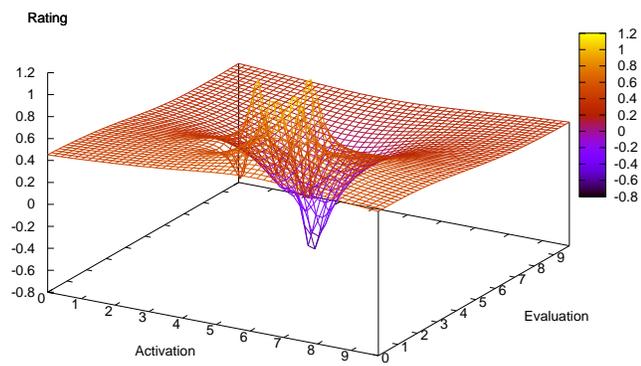


Abbildung 6.18.: Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher, Regina und Sebastian vor der Phase des Gruppenlernens.

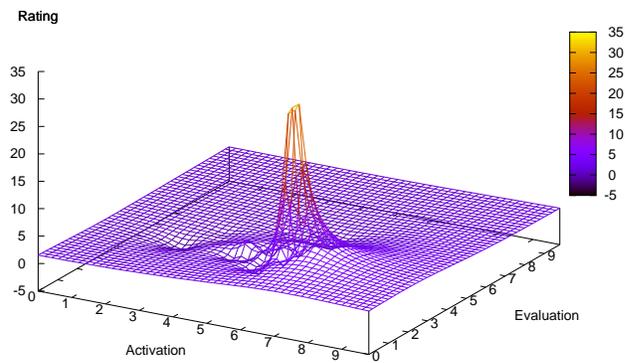
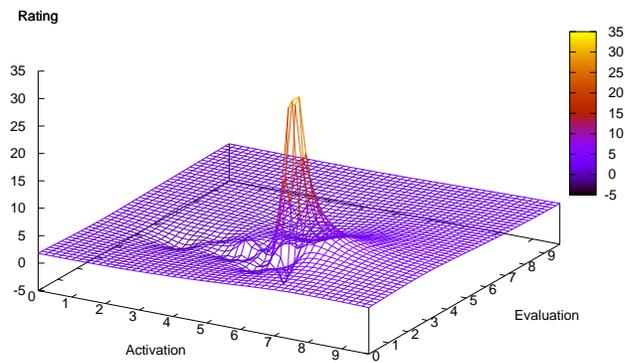
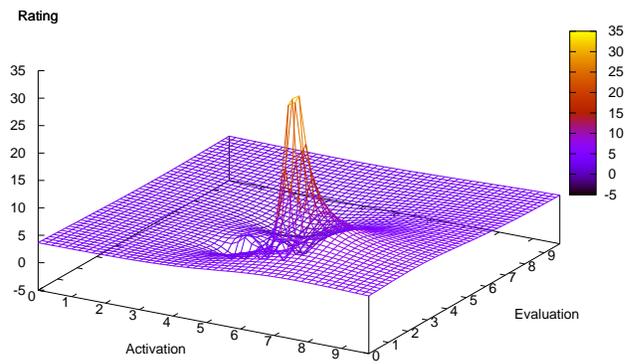


Abbildung 6.19.: Die Abbildung zeigt in dieser Reihenfolge: die Präferenzen für jeweils Christopher, Regina und Sebastian nach der Phase des Gruppenlernens.

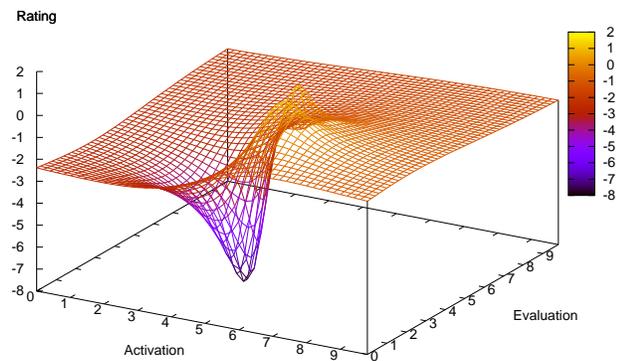
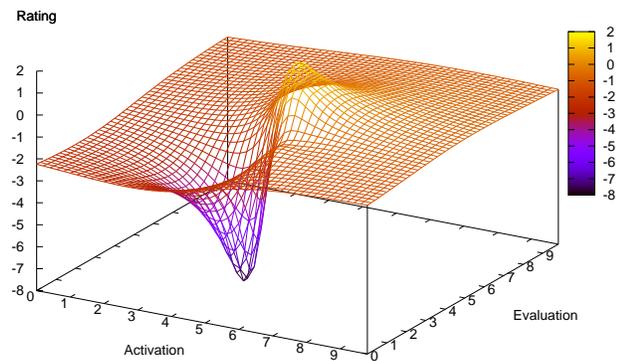
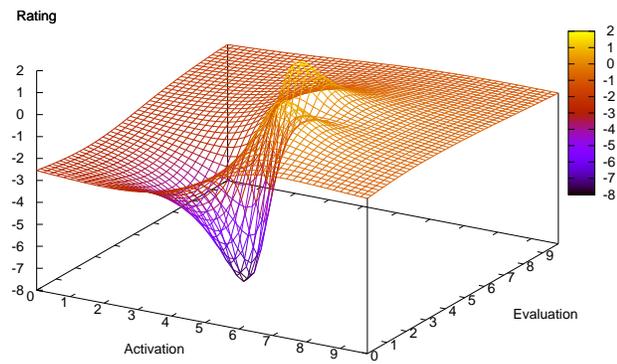


Abbildung 6.20.: Die Abbildung zeigt die Layer zwischen den paarweise verschiedenen Personen nach der Phase des Gruppenlernens in der Reihenfolge: Christopher↔Regina, Christopher↔Sebastian, Regina↔Sebastian.

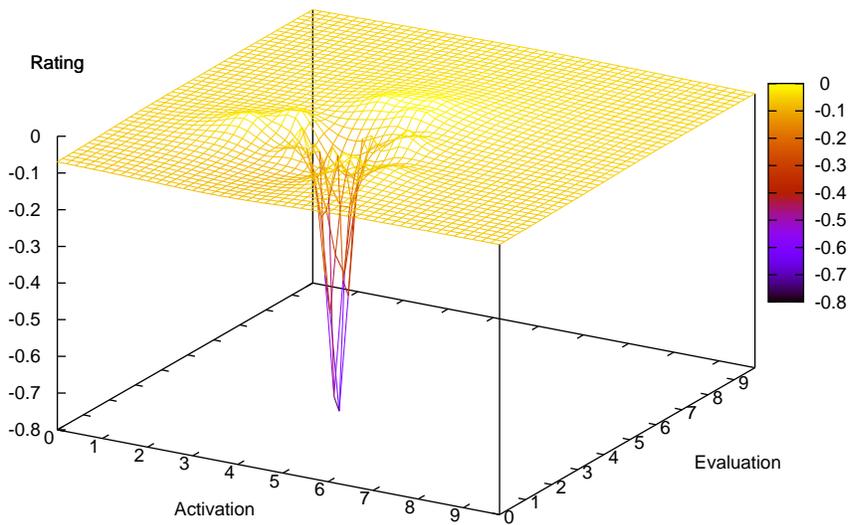


Abbildung 6.21.: Die Abbildung zeigt die Hypothese für Christopher, Regina und Sebastian nach der abgeschlossenen Gruppenlernphase.

7. Zusammenfassung und Ausblick

Aufgrund der immensen Datenmassen, die uns das Internetzeitalter beschert, und die man nicht mehr im Detail nach den erwünschten Informationen durchsuchen kann, müssen Methoden gefunden werden, automatisch die Wünsche der Anwender oder Konsumenten zu bedienen. Es gibt bereits Anwendungen, die versuchen, aus Verhaltensprotokollen oder persönliche Angaben, einem einzelnen Benutzer Produkt-, oder besser, Konsumvorschläge zu unterbreiten. Für Personengruppen gibt es bisher nur wenige, scheinbar kaum ausgereifte Applikationen, die diesem Anspruch gerecht zu werden versuchen.

Das liegt vor allem daran, dass die Wünsche von Personengruppen nicht direkt aus den einzelnen Präferenzen abgelesen werden können. Es muss vielmehr zunächst ermittelt werden, welche Strategie der Kompromissbildung in einer Gruppe verfolgt wird. Es gibt dazu noch keine umfassende Theorie, doch vereinzelte Experimente haben bereits zu ein paar grundlegenden Einsichten geführt: Rücksicht und Fairness scheinen bei der Kompromissbildung in einer Gruppe eine große Rolle zu spielen. Die Entscheidung einer Gruppe lässt sich jedoch im Einzelfall nicht voraussehen, dafür scheint es zu viele mögliche Einflussfaktoren zu geben.

Einen eindeutigen Faktor kann man allerdings ausmachen: oft beeinflusst die bloße Anwesenheit einer Person die verlautbarte Meinung einer anderen Person. Das entspricht der Definition einer Personengruppe, deckt sich mit den Resultaten entsprechender Versuche und lässt sich überdies hinaus noch leicht erklären. Als stereotypisches Beispiel kann man an einen Vater denken, der zwar alleine gerne Filme für Erwachsene ansieht, dies allerdings niemals in Gegenwart seines Sohnes tun würde. Analog zu dieser Kaschierung von individuellen Präferenzen gibt es auch den Fall, dass man Vorlieben äußert, die man eigentlich gar nicht hat. Denkt man diesen Weg noch weiter, müsste man meinen, dass man aufgrund individueller Präferenzen der Gruppenmitglieder niemals auf die Gruppenmeinung schließen könnte.

Oder aber, man nimmt an, dass die Anwesenheit einer Person eine andere Person in ihrer Meinungsäußerung nur im Hinblick auf sehr spezielle Punkte beeinflusst. Dann sollte man festhalten, welcher Einfluss im Detail verübt wird. Ziel sollte sein, diese Relation zwischen allen paarweise verschiedenen Mitgliedern einer Gruppe zu kennen. Zusammen mit den eigentlichen Vorlieben der Personen würde diese Kenntnis dazu gereichen, die Interessen einer Gruppe zu erfassen.

In der vorliegenden Arbeit wird ein abstraktes Modell für ein Lernsystem für Gruppenpräferenzen vorgestellt, das individuelle Präferenzen mit den Abhängigkeiten zwischen Personenpaaren in Einklang bringt. Darüberhinaus werden viele konkrete Möglichkeiten

zur Implementierung vorgestellt und schließlich anhand einer Beispielanwendung getestet. Durch Vorgabe einer konstruierten Lernsituation, die bestimmte Erwartungen an die Ergebnisse der Anwendung mit sich bringt, werden einige Konfigurationen des Lernsystems miteinander verglichen. Schließlich werden die Resultate zweier Testläufe mit tatsächlichen Benutzereingaben dargelegt. Für die Testläufe werden Konfigurationen der Anwendung verwendet, die in der einfachen, gestellten Lernsituation alle Erwartungen erfüllen, und die sich andererseits möglichst stark voneinander unterscheiden.

Für die Klassifikation der Instanzen der Testanwendung werden neue Wege beschritten. Bei den Instanzen handelt es sich um Fernsehsendungen, über die einige Metadaten, unter anderem eine kurze Beschreibung, zur Verfügung stehen. Die Wörter dieser Beschreibung werden zunächst oberflächlich analysiert, dann semantisch gruppiert und zuletzt auf Koordinatenpaare abgebildet, die Emotionen repräsentieren. Für einige Instanzen werden die Ergebnisse dieser Verarbeitung zusammengefasst und in Diagrammen veranschaulicht.

Als problematisch für diese Methode wurden bereits die Mehrdeutigkeit von Wörtern und das Fehlen der Metadaten entlarvt. Ein großer Vorteil der Methode ist hingegen, dass sie keinesfalls auf die Fernsehdomäne beschränkt, sondern universell einsetzbar ist. Genauso wie der Anwendungsbereich, sind auch die einzelnen Module austauschbar, wie zum Beispiel die Übersetzungstabelle für Emotionskoordinaten.

Sicherlich ist die emotionale Klassifizierung von Instanzen in vielen verschiedenen Bereichen sinnvoll. Man könnte gut Internetseiten, Filmkritiken, Beschreibungen von Büchern oder Spielen in eine derartige emotionale Klassifikation überführen. Dabei wäre natürlich interessant, inwiefern die emotionalen Präferenzen von Anwendern unabhängig von der jeweiligen Domäne sind.

Das gewählte Schema zur Repräsentation emotionaler Präferenzen versucht, Emotionen auf zwei Dimensionen zu reduzieren. Es gibt mehrere Wissenschaftler der Psychologie, die diesen Ansatz verfolgen. Ein Vergleich der vorliegenden Wörter-Koordinaten Liste mit einer alternativen Abbildung, von Gefühle beschreibenden Wörter auf Koordinatentupel, wäre vor allem in Hinblick auf gute Ergebnisse der Interpolation zwischen bekannten Bewertungen emotionaler Komponenten sinnvoll.

Es ist schwierig, die Inhalte eines Informationsstroms auf fixe Werte in zwei Dimensionen abzubilden. Um möglichst zutreffende Merkmalsausprägungen von Filmen, Musikstücken, etc. zu erhalten, scheint die Berücksichtigung der Zeit sehr wichtig zu sein. Es gibt bereits Verfahren, mittels derer man den Verlauf emotionaler Regungen von Personen (quasi ein Fluss emotionaler Information) speichern und vergleichen kann. Einerseits könnte man einen derartigen Informationsfluss durch das „Emotionsmapping“ von Texten sehr einfach generieren, wenn man die entsprechenden Wörter, wie bei einem Buch, vorliegen hätte. Andererseits könnte man diese multidimensionalen Klassifikationsmerkmale auch mit dem Lernsystem für Gruppenpräferenzen verbinden.

Einen großen Spielraum für weitere Untersuchungen bietet auch das Lernsystem für Gruppenpräferenzen an sich, beziehungsweise jede seiner Komponenten. Solange noch keine Einigkeit über eine allgemeine Strategie zur Kompromissfindung in Gruppen herr-

scht, wäre es sinnvoll, diejenige Strategie ausfindig zu machen, die am öftesten zutrifft und gleichzeitig einfach mit Ausnahmen vereinbar ist. Man könnte auch auf konventionelle Lernalgorithmen zurückgreifen, um die Operatoren $+$, $*$ und \circ zu implementieren. Dies hätte eine Repräsentation der Gruppenhypothese als neuronales Netz, Entscheidungsbaum, Regelmenge oder Ähnlichem zur Folge. Für jede mögliche Repräsentation von Präferenzen und Layern sollte die minimale Transformation angegeben werden, die die Präferenzen mit den Constraints der Layer abgleicht.

Im vorgeschlagenen System haben Layer die gleiche Repräsentation wie Präferenzen. Es könnten auch hier, je nach Problemstellung der Anwendung, vollkommen verschiedene Repräsentationen gewählt werden, solange man sie mit dem \circ Operator zu einer Hypothese verarbeiten kann.

Das Problem der automatischen Generierung von Vorschlägen für Gruppen wird voraussichtlich immer stärker an Bedeutung gewinnen. Im Rahmen dieser Arbeit werden einige Ideen formuliert, um ihm zu begegnen.

A. Anhang

<i>Gefühle umschreibende Wörter</i>	<i>a, e</i>
abgestoßen, angeekelt, angewidert, empört, überdrüssig, abenteuerlich, gewagt, unternehmenslustig, unternehmungslustig	4.2, 5.9
desinteressiert, uneigennützig, unparteiisch, unvoreingenommen	2.1, 2.4
anhänglich, gütig, herzlich, liebevoll, teilnehmend, zärtlich, zugeneigt, zugetan	4.7, 5.4
ängstlich, bange, besorgt, fürchtend	4.9, 3.4
aggressiv, angreifend, angriffslustig, ausfällig, draufgängerisch, dynamisch, energisch, feindlich, kampfbetont, offensiv, streitlustig, streitsüchtig	5.9, 2.9
frustriert, unzufrieden	4.6, 2.7
angenehm, annehmlich, einverstanden, gefällig, liebenswürdig	4.3, 5.2
argwöhnisch, misstrauisch	3.8, 2.8
erstaunt, überrascht, verblüfft	5.9, 5.5
begierig, eifrig	5, 5.1
ambivalent, doppelwertig, zwiespältig	3.2, 4.2
ekstatisch, verzückt	5.2, 5.5
belustigt, erfreut, vergnügt	4.9, 5
ärgerlich, aufgebracht, böse, düster, entzündet, erbost, erzürnt, schmerzhaft, stürmisch, ungehalten, verärgert, wütend, zornig	4.2, 2.7
betreten, embarrassed, gehindert, verlegen, verwirrt, in Verlegenheit, in Verlegenheit gebracht	4.4, 3.1
ärgerlich, unmutig, verärgert, verdrießlich	4.4, 2.5
abgearbeitet, leer, nichtssagend, unbeladen	3.1, 3.8
antagonistisch, entgegenwirkend, feindlich, gegnerisch, widerstreitend	5.3, 2.5
begeistert, enthusiastisch, passioniert, schwärmerisch	5.1, 4.8
erwartend, vorausschauend, vorgehend, vorwegnehmend	3.9, 4.7
abgünstig, neidig, neidisch, scheelsüchtig	5.3, 2
ängstlich, bange, bedenklich, bemüht, besorgt, bestrebt, unruhig, darauf bedacht, ängstlich über, begierig auf, besorgt über, unruhig um, unruhig wegen, begierig nach, um etw. bestrebt	6, 2.3
apathisch, teilnahmslos	3, 4.3
beschämt	3.2, 2.3

<i>Gefühle umschreibende Wörter</i>	<i>a, e</i>
aufgebracht, erbost, grimmig, wütend	5.6, 3.7
erstaunt, überrascht, verwundert, in Erstaunen versetzt	5.9, 4.7
achtsam, aufmerksam	5.3, 4.3
fröhlich, schadenfroh	5.3, 4.8
bedrückt, dunkel, düster, finster	2.4, 3.2
schamhaft, schüchtern, verlegen, verschämt	2, 2.7
erfolgshungrig, gefräßig, gierig, habgierig	4.9, 3.4
bestürzt, fassungslos	3.1, 2.3
bitter, bitterkalt, hart, herb, rau, schmerzlich	6.6, 4
grantig, griesgrämig, nörglerisch	4.4, 2.9
prahlerisch, überheblich, sich rühmend	3.7, 3
schuldbewusst, schuldig	4, 1.1
angebohrt, ausgebohrt, gelangweilt	2.7, 3.2
erfreut, freudvoll, fröhlich, glücklich	5.3, 5.3
gelassen, ruhig, still, windstill	2.5, 5.5
hilflos, ratlos	3.5, 2.8
behutsam, umsichtig, verhalten, vorsichtig, auf der Hut	3.3, 4.9
aufgekratzt, aufgeräumt, fröhlich, heiter	5.2, 5
erwartungsvoll, hoffnungsvoll	4.7, 5.2
durcheinander, konfus, verwirrt	4.8, 3
aussichtslos, ausweglos, hoffnungslos, unverbesserlich, verzweifelt	4, 3.1
geringschätzig, verachtend, verächtlich, verachtungsvoll	3.8, 2.4
ablehnend, feindlich, feindselig, lebensfeindlich	4, 1.7
zufrieden	4.8, 5.5
entgegengesetzt, konträr, umgekehrt, unbeständig, zuwider, entgegen, gegen, zuwider	2.9, 3.7
ungeduldig	3.4, 3.2
entgegenkommend, genossenschaftlich	3.1, 5.1
gefühlsbeherrscht, impulsiv, stoßartig, treibend	3.1, 4.8
bedenklich, entscheidend, kritisch	4.9, 2.8
unentschlossen	3.4, 2.7
ausgefallen, eigenartig, gespannt, merkwürdig, neugierig, seltsam, sonderbar	5.2, 4.2
gewagt, wagemutig, waghalsig	5.3, 4.4
aufsässig, trotzig	4.4, 2.8
entzückt, erfreut	4.2, 6.4
intolerant, unduldsam	3.1, 2.7
abfordernd, anspruchsvoll, bedrückend, drückend, erhöht, fordernd, genau, hoch, lastend, schwierig, streng, verlangend	5.3, 4

<i>Gefühle umschreibende Wörter</i>	<i>a, e</i>
ärgerlich, genervt, irritiert	5.5, 3.3
bedrückt, deprimiert, flau, gedrückt, niedergeschlagen, rückläufig, schleppend	4.2, 3.1
scheelsüchtig, eifersüchtig auf, neidisch auf, abgünstig	6.1, 3.4
verzweifelt	4.1, 2
erfreulich, freudig, freudvoll	5.4, 6.1
lästig, unangenehm, unliebenswürdig, unliebsam, unsympathisch, verdrießlich, widerlich, widerwärtig, übel gelaunt	5,3.7
loathful	3.5, 2.9
enttäuscht	5.2, 2.4
abgeschlossen, einsam	3.9, 3.3
entmutigt, mutlos	4.2, 2.9
bescheiden, demütig, gütig, kleinlaut, mild, sanft, zahm	3, 4.3
aufgeregt, gereizt, kräftig, nervös, reizbar	5.9, 3.1
ergeben, folgsam, gehorsam, lenkbar, obrigkeitshörig, unterwürfig	3.1, 4.7
automatisch, beherrscht, selbsterregt	4.4, 5.5
diensteifrig, entgegenkommend, gefällig, höflich, kulant, pflichtgemäß, verbindlich, zuvorkommend	2.7, 3
gelassen, heiter, ruhig	4.3, 4.4
schockiert	4.3, 3.2
beunruhigend	5.4, 3.6
gesellig, kontaktfreudig	4.8, 5.3
beharrlich, geduldig, langmütig	3.3, 3.8
betrübt, kummervoll, sorgenvoll, traurig	4.5, 3.1
gedankenvoll, nachdenklich	3.2, 5
eigensinnig, halsstarrig, störrisch, unbeugsam, verbissen, widerspenstig	4.9, 3.1
devot, gehorsam, unterwürfig	3.4, 3.1
überrascht, erstaunt über, verwundert über	6.5, 5.2
befriedigt, erfreut, zufrieden	5.3, 5.1
argwöhnisch, suspekt, verdächtig, misstrauisch gegenüber jmdm.	4.4, 3
besitzanzeigend, besitzend, besitzergreifend, besitzgierig, habgierig, possessiv	4.7, 2.8
mitfühlend, sympathisch, verständnisvoll	3.6, 3.2
hochmütig, stolz	4.7, 5.3
in Panik	6.3, 3.4
verblüfft, verdutzt, verwirrt, verwundert	2.6, 3.8
händelsüchtig, streitsüchtig, zänkisch	4.6, 2.6
vertrauend, vertrauensvoll, zutraulich	3.4, 5.2
unaffectionate	3.6, 2.1

<i>Gefühle umschreibende Wörter</i>	<i>a, e</i>
auführerisch, aufsässig, aufständisch, rebellisch	5.2, 4
abgelehnt	5, 2.9
unfreundlich, unsympathisch	4.3, 1.6
reuevoll, reuig, reumütig	3.1, 2.2
abgünstig, aufgebracht, grollend, nachtragend, nachtragenden, übelnehmerisch, aufgebracht gegen jmdn./etw.	5.1, 3
betrübt, kummervoll, schlimm, traurig, trist	3.8, 2.4
beißend, höhnisch, sarkastisch	4.8, 2.7
befriedigt, zufrieden, zufrieden gestellt	4.1, 4.9
staunend, wundernd	3.3, 5.2
höhnisch, verächtlich, erachtungsvoll, voll Verachtung	5.4, 4.9
abgequält, besorgt, beunruhigt	3.9, 2.9

Tabelle A.1.: Gefühle umschreibende Adjektive und deren Activation-/Evaluation-Koordinaten.

<i>Beispielsendungen</i>
...so alt wie man sich fühlt
25 Tage in Europa
88acht BEROLINA
88acht GUTEN ABEND BERLIN
88acht GUTEN MORGEN BERLIN
88acht GUTEN TAG BERLIN
88acht HEY MUSIC
88acht NACHTISCH
88acht WECKER
ARD MHP-Testkanal
ARD-Buffer
ARD-NACHTKONZERT (IV)
ARD-Nachtexpress mit Radiowecker
ARD-Nachtkonzert
ARD-Ratgeber: Heim & Garten
ARTE Kultur
Abendschau
Abenteuer Labor
Abenteuer Wildnis
Aktuelle Stunde
Aktueller Bericht
Albanisch
Alfredissimo!
Allegro
Arsen und Spitzenhäubchen
Auf ein Wort
BR Verkehr
Bayern 1 - Café
Bayern 1 - Feierabend
Bayern 1 - Musikjournal
Bayern 1 - Radiotreff
Bayern 1 am Abend
Bayern 3 - die MorningShow
Bayern 3 - die RadioShow
Bayern Mobil - Radio, das sich sehen lassen kann
Beckmann
Bergauf-Bergab
Bianca - Wege zum Glück
Bilderbuch Deutschland

<i>Beispielsendungen</i>
Bill Lishman - Der mit den Kranichen fliegt
Bistro
Björn Bär
Blickpunkt Sport
Blinder Hass
Bosnisch
Brandenburg aktuell
Brauchen wir sie noch?
Bremen Eins am Abend
Bremen Vier: Dein Ding
Bremen Vier: Planet Arnie / Planet Malte
Bremen Vier: Rocken, poppen, dancen
Brisant
Bubu, der Hund mit dem Hausschuh
Bulgarisch
Bus 174 - Geiseldrama in Rio
Café Global
Café Trend
Christa Ludwig singt Lieder von Johannes Brahms
Clubnight
Coma
Cosmo
Cosmo TV
DAS!
DAS! ab 2
DW Musik
Damals in der DDR
Das Forum der Europäer
Das Herbstfest der Volksmusik
Das Tagesgespräch
Das Wichtige wissen - im Viertelstundentakt
Das Wunder vom Dachstein
Das malvenfarbene Taxi
Daten der Woche
Den Mörder trifft man am Buffet
Der Dicke und der Dünne
Der Elefant, sein Mahout und der Moloch
Der Fahnder
Der Fischer vom Heiligensee

<i>Beispielsendungen</i>
Der Kleine Bär
Der Letzte seines Standes?
Der Nachmittag
Der Stellvertreter
Der Teufel soll die Dichterei beim Hemdenwaschen holen
Der Vormittag
Der Wunschkuss
Der schwarze Mustang
Der wandernde Maestro
Die Abendschau
Die Bamberger Symphoniker
Die Deutsche Alleenstraße
Die Erben der Pharaonen
Die Fallers - Eine Schwarzwaldfamilie
Die Fallers - eine Schwarzwaldfamilie
Die Gruselschule
Die JUMP Morningshow
Die Lesung
Die Pfefferkorn Familie
Die Sendung mit der Maus
Die Sprechstunde
Die Stadt frisst ihre Kinder
Die Thelon-Expedition
Die schöne Tochter des Rheins - eine kulinarische Moselreise
Dinner for two
Diplomatie im Mauerschatten
Dir gehört mein Leben
Dok 5 - Das Feature
Doppel-Kopf
Dreiländernacht
Dryja & Rosch
Ein Herz voll Musik
Ein Tag „auf den Spuren der Gedanken“
Ein neues Herz aus Leipzig
Eine Couch für alle Felle
Eine Zigeunerkapelle geht um die Welt
Enteignet für die Einheit?
Euro-Blick
FAKT

<i>Beispielsendungen</i>
FIGARO AM ABEND
Faszination Kreuzfahrt
Fazit
Felix und die wilden Tiere
Fernsehen verboten!
Flüssiges Gold
Fofteihn
France Info
Frisch gekocht
FrühStück
Funkjournal
Für uns in Mecklenburg-Vorpommern
Gartenduell
Geschichten aus dem Zauberhaus
Gioacchino Rossini
Glanz der Erde
Goldene Hits
Gottes verlassene Häuser
Griechisch
Großstadtrevier
Grüße und Musik
Guten Abend Sachsen-Anhalt
Guten Morgen Sachsen
Guten Morgen Sachsen-Anhalt
HEIMATJOURNAL
Hallo Niedersachsen
Hamburg Journal
Hamburg nach Sieben
Hanni und Nanni
Heidi und Erni
Heiraten macht mich nervös
Herbst in der Rhön
Hessen à la carte
Hessenschau
Hexe Lilli
Hier ab vier
Hier und Heute
Hierzuland
Hierzulande

<i>Beispielsendungen</i>
Hol den Doc
Horizonte
Hund oder Katze - Wer ist klüger?
Im Hause Chanel
Im Jahreskreis
Im Tal der Wetterpropheten
In aller Freundschaft
JUMP - Der neue Sound im Radio
JUMP am Nachmittag mit Mitri
Joe und Max - Rivalen im Ring
Journal am Vormittag
Jungbrunnen für den alten Kontinent
KULTURRADIO AM MITTAG
KULTURRADIO AM MORGEN
KULTURRADIO AM NACHMITTAG
Kaffee oder Tee?
Klassik-Lunch
KlassikPlus
Klassisch in den Tag
Klassisch unterwegs
Konzert
Krakaus junge Helden
Kriegsende an Rhein, Ruhr und Weser
Kroatisch
Kultur Journal
KulturReport
Kulturjournal
Kulturweltspiegel
Kulturzeit
LILIPUZ - Radio für Kinder
Landesschau Baden-Württemberg
Landesschau Rheinland-Pfalz
Landesschau unterwegs
Landpartie zum Nachbarn
LateLounge: Club
Lebenslinien
Lebensunwert
Leonardo - Wissenschaft und mehr
Leporello

<i>Beispielsendungen</i>
LeseZeichen
LexiTV - Wissen für alle
Liebe an der Macht
Lindenstraße
Lokalzeit Bergisches Land
Lokalzeit Münsterland
Lokalzeit OWL aktuell
Lokalzeit Ruhr
Lokalzeit Südwestfalen
Lokalzeit aus Aachen
Lokalzeit aus Dortmund
Lokalzeit aus Düsseldorf
Lokalzeit aus Köln
Lola
Länder - Menschen - Abenteuer
Ländersache - Sommerreise
Länderzeit
MDR um zwölf
MEX. das marktmagazin
Mach Dich ran
Magdeburger Gespräch
Marc Pircher: Mein Zillertal
Marienhof
Matinee
Mazedonisch
Melilla - Europas Vorposten in Afrika
Mensch, Umwelt, Technik
Mikado
Mir geht es gut!
Mit hr4 in den Tag
Mord an der roten Rita
Morgenmelodie
MusicHall
Musik liegt in der Luft
Musik zum Träumen
MusikThema
Musikzeit
Mutter und Kind - Gebet für die Zukunft
N-JOY

<i>Beispielsendungen</i>
N-JOY MORNINGSHOW
N-JOY@WORK
NDR 1 Radio MV - Die Mittagsschau
NDR 1 Radio MV - Die Nachmittagsschau
NDR 1 Welle Nord - Der Nachmittag
NDR 1-Nacht
NDR Sinfonieorchester
NDR Talk Show
Nachrichten - 10 min
Nachrichten - 5 min
Nachrichten, Wetter, Verkehr
Nachrichtenradio vom BR
Nachrichtenradio vom rbb
Nachtkultur
Nachtmagazin
Neugier genügt
Ninotschka
Nordbilder
Nordmagazin
Nordwestradio Journal
Notizbuch
Ortszeit
Overland Pacific
Pavillon
Planet Wissen
Polizeiruf 110: Um Kopf und Kragen
Polnisch
Profile
Radio Colonia
Radiofeuilleton
Radiomarkt
Raketen, nein danke
Reise in die Antarktis
Resonanzen
Rothkirchs Krieg
SPUTNIK - Der Tag - Back to the Music
SPUTNIK - Fett geweckt
SPUTNIK Ungedatet
SWR2 Eckpunkt

<i>Beispielsendungen</i>
SWR2 Journal am Morgen
SWR2 Musik Spezial
SWR2 RadioART: Essay
SWR2 Thema Musik
SWR2 Treffpunkt Klassik
SWR2 Wissen
SWR2 am Morgen: Musik zum Tagesbeginn
Sabine Christiansen
Sachsen-Anhalt heute
SachsenSpiegel
Scala - Aktuelles aus der Kultur
Schlager
Schlaglicht
Schlemmerreise Altbayern und Schwaben
Schlemmerreise Frankreich
Schleswig-Holstein Magazin
Schloss Einstein
Schätze der Welt
Schätze der Welt - Erbe der Menschheit
Serbisch
Serenade
Service: Auto und Verkehr
Service: Gesundheit
Service: Natur und Technik
Service: Recht und Finanzen
Service: Reise und Freizeit
Servicezeit: Essen & Trinken
Servicezeit: Gesundheit
Sesamstraße
Spielräume
Sport am Montag
Sport kompakt
Sportplatz
Stadtgeschichten: Brückenschläge
Standpunkt
Stefan Aust
Straße der Träume
Streifzüge
Sturm der Liebe

<i>Beispielsendungen</i>
Sturm der Liebe (1)
Sturm der Liebe - Wie alles begann
Sängerporträt
THE VOICE
TagesZeichen
Tagesgespräch
Tagesschau
Tagesthemen
Tatort: Requiem
Tele-Gym
Teletubbies
Thema
Thema Europa
Tibetisch heilen im Himalaya
Total Genial
Tote ohne Begräbnis
Treffpunkt Sport
Tuerkisch
Turbo, Spoiler, breite Schlappen
Täter - Opfer - Polizei
Unser blauer Planet
Unsere Schule ist die Beste
Unsere Tierklinik
Unterwegs im Land
VOR ORT
Verbotene Liebe
Videolexikon
Viertel vor ...
Von Binnenland und Waterkant
WDR 2 Klassik Mittagsmagazin
WDR 2 Klassik MonTalk
WDR 2 Mittagsmagazin
WDR 2 MonTalk
WDR 2 Soundfiles
WDR 2 Westzeit
WDR 3 Konzert
WDR 3 PHON
WDR 3 open: pop drei
WDR WELTWEIT

<i>Beispielsendungen</i>
WDR aktuell
WISO
Waldo
Welt der Tiere
Weltspiegel
Weltwunder der Technik
Westblick
Westpol
Westpol nachgefragt
Wiesn exotisch
Wildbach
Willis Quiz Quark Club
Windrose
Wir gratulieren
Wir in Bayern
Wir testen die Besten
World Music Night
You FM-Vormittag
ZDF-Mittagsmagazin
ZDF-Morgenmagazin
Zapping International
ZeitZeichen
Zeitfragen
Zum 60. Todestag von Béla Bartók am 26.09.2005
Zum 75. Geburtstag von Fritz Wunderlich am 26.09.2005
Zwei Wochen Argentinien
Zwei nach Hawaii
aspekte kompakt
betrifft
buten un binnen
c't magazin
daheim & unterwegs
der TAG
die story
hr 3 Pop & Weck TV
hr-klassik.lunch
hr1 - LOUNGE
hr1 - MERIDIAN
hr1 - METRO

<i>Beispielsendungen</i>
hr1 - MOBIL
hr1 - PRISMA
hr1 - START
hr1 - VITA
hr3-Extra
hr3-Life
hr3-Madhouse
hr3-Pop & Weck
hr3-Puls
hr4 - Bunt gemischt
hr4 - Radioladen
höchstpersönlich
maintower
markt
nano
neues spezial
neuneinhalb
offen gesagt
praxis kompakt
radioWissen
rbb regional
rbb um 6
rundungesund
service: familie
zibb

Tabelle A.2.: Die Liste derjenigen Programmbeispiele, die im Rahmen dieser Arbeit verwendet wurden.

Literaturverzeichnis

- [1] R. St. Amant, J. E. Blair, P. Barry, Y. Bentor, and C. G. Healey. A visual interface to a music database. <http://citeseer.csail.mit.edu/489637.html>.
- [2] J. Aucouturier. Scaling up music playlist generation, 2002.
- [3] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J.G. Taylor. Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18:32 – 80, January 2001.
- [4] Informatik der Technischen Universität München. Onlinewörterbuch, 2005. <http://dict.leo.org>.
- [5] B. Edmonds. The pragmatic roots of context. In Paolo Bouquet, Luigi Serafini, Patrick Brézillon, Massimo Benerecetti, and Francesca Castellani, editors, *Modeling and Using Contexts: Proceedings of the Second International and Interdisciplinary Conference, CONTEXT'99*, pages 119–134. Springer-Verlag, Berlin, 1999.
- [6] F. Dornseiff et al. *Der deutsche Wortschatz nach Sachgruppen*. Walter de Gruyter GmbH, Berlin, 2004.
- [7] Farlex. The free dictionary, 2005. <http://www.thefreedictionary.com>.
- [8] M. Goto and K. Hirata. Recent studies on music information processing. <http://staff.aist.go.jp/m.goto/PAPER/JAS200411goto/>.
- [9] W. Hehlmann. *Wörterbuch der Psychologie*. Alfred Kröner Verlag, Stuttgart, 1974.
- [10] Z. Liu, J. Huang, and Y. Wang. Classification of audio events in broadcast news. *Proceedings of the 2nd IEEE Workshop of Multimedia Signal Processing*, pages 364 – 369, Dezember 1998.
- [11] Experiments B. Logan. Content-based playlist generation: Exploratory. <http://citeseer.ist.psu.edu/580794.html>.
- [12] J. Masthoff. Modeling a group of television viewers. In *Proceedings of the Future tv: Adaptive instruction in your living room workshop.*, 2002.

- [13] J. McCarthy and T. Anagnost. Musicfx: An arbiter of group preferences for computer supported collaborative workouts. *Proceedings of the ACM 1998 Conference on CSCW, Seattle, WA*, pages 363 – 372, 1998.
- [14] C. Meghini and F. Sebastiani. Conceptual modelling in multimedia information seeking. In *Selected Papers from the Symposium on Conceptual Modeling, Current Issues and Future Directions*, volume 1565, pages 281 – 286, 1999.
- [15] K. Minami, A. Akutsu, and H. Hamada et al. Video handling with music and speech detection. *IEEE Multimedia*, pages 17 – 25, Fall 1998.
- [16] T. M. Mitchell. *Machine Learning*. McGraw Hill, Boston, Massachusetts, 1997.
- [17] M. O’Conner, D. Cosley, J.A. Konstan, and J. Riedl. PolyLens: A recommender system for groups of users. *Proceedings of ECSW 2001, Bonn*, pages 199 – 218.
- [18] F. Pachet. Content management for electronic music distribution. *Commun. ACM*, 46(4):71–75, 2003.
- [19] F. Pachet. *Knowledge Management and Musical Metadata*. Idea Group, 2005.
- [20] F. Pachet, A. Laburthe, and J.-J. Aucouturier. The cuidado music browser: an end-to-end electronic music distribution system. In INRIA, editor, *Proceedings of CBMI 03*. IRISA, 2003.
- [21] N. Patel and I. Sethi. Audio characterization for video indexing. *Proceedings of the SPIE Conference on Storage Retrieval Still Image Video Databases, San Jose, CA*, 2670:373 – 384, 1996.
- [22] S. Pauws and B. Eggen. Pats: Realization and user evaluation of an automatic playlist generator. In *International Symposium on Music Information Retrieval*, pages 222–230, 2002.
- [23] D. Pennock, E. Horvitz, S. Lawrence, and C. L. Giles. Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, UAI 2000*, pages 473–480, Stanford, CA, 2000.
- [24] R. Plutchik. *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, New York, 1980.
- [25] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl. Analysis of recommendation algorithms for e-commerce. In *ACM Conference on Electronic Commerce*, pages 158–167, 2000.

- [26] M. Torrens, P. Hertzog, and J.-L. Arcos. Visualizing and exploring personal music libraries. In *5th International Conference on Music Information Retrieval (ISMIR 2004)*, Barcelona, Catalonia, Spain, October 2004.
- [27] G. Widmer. Tracking context changes through meta-learning. In *Machine Learning*, volume 27(3), pages 259–286, 1997.
- [28] O. Wüst and O. Celma. An mpeg-7 database system and application for content-based management and retrieval of music. In *Proceedings of Fifth International Conference on Music Information Retrieval*, Barcelona, 2004.
- [29] T. Zhang and C.-C. J. Kuo. Audio content analysis for online audiovisual data segmentation and classification. In *Transactions on Speech and Audio Processing*, volume 9(4), pages 441–457, May 2001.