

Association Rule Breaking Based Diversity for Recommender Systems

University of Southern Denmark

Ed Abel abel@sdu.dk

Andrew Goddard bossfansheff@gmail.com

Recommender Systems



• Have become omnipresent in our data lives



Recommender Systems Objectives



- Accuracy
- Looks to optimize user satisfaction
- Focusing on accuracy can recommend familiar items
- Notion of You like X so here is more X
- Danger "Toy Story 1" & "Toy Story 2" → "Toy Story 3"
- **Diversity:** Consensus that it's also an important consideration
- Promote the discovery of varied content
- Introduce new and varied less repetitive recommendations
- Can ultimately improve user experience and satisfaction
- Tension/Trade-off between Accuracy & Diversity

Diverse Recommendation Approaches SDU -

- Content-Based Filtering with Diversity Enhancement
- Adjusting item similarity considerations to encourage a broader item range
- E.g. movies with a mix of genres, even if a user watches 1 genre
- Re-Ranking with Diversity Considerations
- Generate an initial ranked list of recommendations based on relevance
- Then re-rank the list to increase diversity
- E.g. News articles list reordered to downweigh category repetitions
- Context-Aware Recommender Systems
- Considering contextual information such as time/location etc.
- Create diverse recommendations suited to different contexts
- E.g. Dining app suggesting breakfast places during AM, restaurants during PM
- Creating separate recommendation lists
- Creating distinct lists that are stratified by topics/feature
- E.g. Movie streaming site having separate categories/genre lists

Association Rule Mining



- Data Mining technique to find interesting relationships/associations
- Identifies relationships in the form of rules
- E.g. customer purchases data
- "if customer buys Product-X" they "are likely to also buy Product-Y"
- Within Association rule mining we have:
 - Items
 - Transactions
 - Derived Rules



- Rule form $A \Rightarrow B$
- Suggesting if A is/are present, B is/are likely to be present too

Transaction 1	🍎 🔰 😔 💊
Transaction 2	i 🔴 📔 😔
Transaction 3	i 💭 🔟
Transaction 4	Ö

Support, Confidence and Lift



• Measures to help quantify meaning and significance of association rules

Support

- measures how frequently the items in a rule appear together in the data
- It tells us how common or rare an item set is in the entire dataset

Confidence

- Rule reliability: shows how often the rule has been true
- **Predictive power:** indicates the likelihood that if the 1st item is present, the 2nd item will also be present

Lift

- Strength of connection: Indicating how much more likely the items are to appear together than by random chance
- **Comparative measure:** Telling us if the presence of one item increases the likelihood of the other item more than would be expected randomly
- A higher lift value indicates stronger and more significant rule

Association Rules In Rec Systems



- Work has considered users as transactions
- Each containing the set of items the user has consumed/liked
- Providing associations of the form
- If a person consumes/liked x they also consumes/liked y
- If a user has consumed/liked x but not y
- Y can be recommended



Association Rules In Rec Systems

- Work has considered users as transactions
- Each containing the set of items the user has consumed/liked
- In our approach we consider content as a transaction
- Made up of a set of item-parts



SDU

Jim Carrey Movies



- In our approach we consider association rules through:
- Items as transaction made up of a set of item-parts
- Movies with features regarding actors and genres etc.
- Jim Carrey ⇒ Comedy
- But not necessary the inverse (Comedy ⇒ Jim Carrey)



Our approach explores diversity through calculating diversity in terms of the level of association rule breaking or conformity

Jim Carrey Movies



- Many films with features of Jim Carrey and Comedy
- Jim Carrey ⇒ Comedy
- Instead of seeking diversity in terms of Not Jim Carrey & Not Comedy
- Interesting unusual case would be when this rule is broken
- Jim Carrey !⇒ Comedy



Recommendations both aligned to some user tastes and diverse





Historical Live Music Concerts Domain SDU -

- Historical live concert performance recordings •
- Becoming readily available
- Providing an unwieldly abundance of choice ٠



Live at Maple Leaf Gardens, Toront...

Live at Kent State University, USA -... 1973

Europe '72 Vol. 2 2011 • Album

Grateful Dead

Europe '72 Vol. 19:... 2011 • Album

Europe '72 Vol. 18:... 2011 · Album

2011 • Album

Europe '72 Vol. 16:... 2011 • Album

Europe '72 Vol 17: ... 2011 • Album

Europe '72 Vol. 2 2011 • Album

Live at University of Cincinnati, USA... PINK FLOYD 1973



Europe '72 Vol. 14:..

2011 - Album



Europe '72 Vol. 13:... 2011 - Album



2011 • Album



FINEST QUALITYO

2010 • Album



Crimson, White 2010 • Album



Live at Chicago International Amphi... 1973



6

.... 🗢 🕞

Q Search

f

Home

|||Your Library

Concert setlists



Setlist

Queen News of the World Bingley Hall, Stafford, England

We Will Rock You Brighton Rock somebody to Love Death on Two Legs (Dedicated to ...) Killer Queen Good Old-Fashioned Lover Boy I'm in Love With My Car Get Down, Make Love The Millionaire Waltz You're My Best Friend Spread Your Wings It's Late Now I'm Here Love of My Life '39 My Melancholy Blues White Man The Prophet's Song Guitar Solo The Prophet's Sona Liar Bohemian Rhapsody Keep Yourself Alive Tie Your Mother Down

Encore

We Will Rock You We Are the Champions Sheer Heart Attack Jailhouse Rock (Elvis Presley cover) God Save the Queen ([traditional] cover)

setlist.fm

Setlist

Queen News of the World Madison Square Garden, New York, NY, USA

We Will Rock You Brighton Rock somebody to Love Death on Two Legs (Dedicated to...) Killer Queen Good Old-Fashioned Lover Boy I'm in Love With My Car Get Down, Make Love The Millionaire Waltz You're My Best Friend Doing All Right Liar Love of My Life ' 39 White Man The Prophet's Song Guitar Solo The Prophet's Song Now I'm Here stone Cold Crazy Bohemian Rhapsody Keep Yourself Alive Tie Your Mother Down Encore We Will Rock You We Are the Champions

We Are the Champions Sheer Heart Attack Jailhouse Rock (Elvis Presley cover) Saturday Night's Alright for Fighting (Elton John cover) Stupid Cupid (Connie Francis cover) Be-Bop-A-Lula (Gene Vincent & His Blue Caps cover) God Save the Queen ([traditional] cover)

setlist.fm

Setlist

Queen News of the World Maple Leaf Gardens, Toronto, ON, Canada

We Will Rock You Brighton Rock Somebody to Love It's Late Death on Two Leas (Dedicated to ...) Killer Queen Good Old-Fashioned Lover Boy I'm in Love With My Car Get Down, Make Love The Millionaire Waltz You're My Best Friend Spread Your Wings Liar Love of My Life '39 My Melancholy Blues White Man The Prophet's Sona Guitar Solo The Prophet's Song Keep Yourself Alive stone Cold Crazy Now I'm Here Bohemian Rhapsody Tie Your Mother Down

Encore We Will Rock You We Are the Champions Sheer Heart Attack Jailhouse Rock (Elvis Presley cover) God Save the Queen ([traditional] cover)

setlist.fm

We contain patterns of the same songs appearing together

Historical Live Music Concerts Domain SDU -

- Association Rules Elements in this domain
- *Items*: Individual songs
- *Transactions:* Concert Setlists, each containing a set of songs
- Rules: such as A ⇒ B, would indicate that when song(s) A are played then song(s) B will likely also be played
- The notion of rules being broken represent:
 - Song A is played but Song B is not
 - Contrary to how this is often the case

Our Approach





Calculating association rules



- Taking an artist's entire setlist career history
- Each show a (transaction) set of songs (items)
- Find rules that have single song of the Left-Hand Side (LHS)
- And a single song the Right-Hand Side (RHS)



• Set of Association Rules and Each's Lift Value

Calculating ARBD Diversity Scores

- Given example rule of: $C \Rightarrow A$
- And Example Transactions of
 - {A,C,D} - - The Rule Holds
 - {B,C,D} + - The Rule is Broken

- Given rule of
- Song A \Rightarrow Song B



SDU合

Association Rule Breaking Diversity (ARBD) SDU -

• For a transaction (Concert) – iterative through every Rule



- Calculate normalised ARBD value for transaction the higher the more Diverse
- So ARBD values for the set of all transactions defines:
 - "most diverse" (in terms of ruling breaking)
 - "most conforming" (in terms of rules holding)

Calculating Recommendations



- So, stages of going from user preferences to recommendations
- User preferences allow different "types" of fans
- Different fan types may have very different diversity preferences
- Causal Fans \rightarrow Looking for accessible recommendations
- Super Fans \rightarrow Looking for something more diverse







- User Preferences
- 1. Indication of how diverse the recommendations should be
- 2. How big a set of recommendations to be shown
- Select window of setlists from list that is ordered by ARBD
- Aligned to the level of diversity desired

[Less Rule Breaking Diverse]

[More Rule Breaking Diverse]

• The desired size of transactions shown to the user

- User can inspect the recommendations and fine tune
- Level of Diversity sought/no of results

Results and Experiments



- Bruce Springsteen
- 5 very diverse recommendations requested



Results and Experiments



- Bruce Springsteen
- 10 quite typical recommendations requested

Select Artist:			Concert Date	Name of Gig's Tour
Bruce Springsteen	•	R1	17 May 2016	The Ties That Bind Tour
		R2	8 Dec 1980	The River Tour
Typical Vs Diverse	1000	R3	25 Aug 2016	The Ties That Bind Tour
0% 20% 100%	R4	26 Oct 1984	Born in the USA Tour	
0 10 20 30 40 50 60 70	80 90 100	R5	21 May 2016	The Ties That Bind Tour
No. to Recommend:		R6	2 Oct 1985	Born in the USA Tour
	R7	25 Feb 2017	The Ties That Bind Tour	
10		R8	17 May 2012	Wrecking Ball World Tour
		R9	13 Sep 1981	The River Tour
GO		R10	30 Jul 1984	Born in the USA Tour

Varying Variety Recommendation Rows SDU 👉

• Could use various "preferences" as part of a multi-row interface



Summary



- Recommender systems omnipresent in our data lives
- Diversity important aspect when curating recommendations
- Proposed an approach to make recommendations
- Utilising Association Rule Mining
- Measuring diversity in terms of association rule breaking
- Exploring example domain of music live concerts
- Live concerts (transactions) made up of songs (item)
- Finding association rules between songs (items)
- User flexibility to define how much diversify they seek
- Casual fan vs Superfan can get very different recommendations



Association Rule Breaking Based Diversity for Recommender Systems

University of Southern Denmark

Ed Abel abel@sdu.dk

Andrew Goddard bossfansheff@gmail.com

Support, Confidence and Lift

- Measures to help quantify significance of association rules
- $A \Rightarrow B$
- **Support** frequency an item(set) appears in the dataset
- A appears in 7 out of 10 transactions
- Support 7/10
- B appears in 6 out of 10 transactions
- Support 6/10
- A and B appear (together) in 4 out of 7
- Support 4/10

Concert	Song X	Song Y
1	Yes	Yes
2	Yes	Yes
3	Yes	Yes
4	Yes	Yes
5	Yes	No
6	Yes	No
7	Yes	No
8	No	Yes
9	No	Yes
10	No	No



Support, Confidence and Lift



- Measures to help quantify significance of association rules
- $A \Rightarrow B$
- **Support** frequency an item(set) appears in the dataset
- A appears in 7 out of 10 transactions = Support 7/10
- B appears in 6 out of 10 transactions = Support 6/10
- A and B appear (together) in 4 out of 7 = Support 4/10
- **Confidence**: Likelihood a RHS of rule given the LHS
- "A and B" appear together in 4 out of the 7 transactions that A
- Confidence = 4/7 ≈ 0.57
- Lift evaluates the strength of a rule
- Comparing confidence to expected confidence if items were independent
- Lift = Confidence/RHS support
- Lift = 0.57 / 0.6 ≈ 0.95
- A higher lift value indicates stronger and more significant rule

ARDB Calculation



- For each concert (*t*) calculate **ARDB**
- Association Rule Breaking Diversity (ARDB)

$$ARBD_t = \frac{1}{m} (\sum_{1}^{m} x)$$

• Where *m* is the number of rules found and *x* is calculated via:

$$x = Rule_m(A \Rightarrow B, T) \begin{cases} +L_m \text{ if } A \subseteq T \text{ and } B \subseteq T \\ -L_m \text{ if } A \subseteq T \text{ and } B \notin T \\ 0 \text{ otherwise} \end{cases}$$

- Where A is the left-hand side of Rule m
- *B* is the right-hand side of Rule *m*
- *T* is the set of items in transaction *t*,
- L_m is the Lift value of $Rule_m$.
- Then normalise so
- "most diverse" (in terms of ruling breaking) has value of 1
- "most conforming" (in terms of rules holding) has value of 0