**Summary of Movie Contextual Recommender System Using Ontology & Probabilistic Graphical Model**

Recommender systems aim at handling the issue of information overload in the World Wide Web by filtering relevant information and providing personalized recommendations to users based on their profile and feedback taking into consideration user’s specific needs and interests. Examples of such applications are Amazon.com, MovieLens, NetFlix, Jumia, Facebook etc.

Recommender systems are usually classified into the following categories, based on how recommendations are made (Adomavicius, 2007):

* Content-based recommendations: the user is recommended items similar to the ones the user preferred in the past;
* Collaborative recommendations: the user is recommended items that people with similar tastes and preferences liked in the past;
* Hybrid approaches: these methods combine collaborative and content-based methods.

All these approaches are two-dimensional and do not consider the context in which the item is used or rated. Two dimensional recommendation requires further improvements to make recommendation methods more effective and applicable to an even broader range of real-life applications which include incorporation of various contextual information into the recommendation process (Adomavicius and Tuzhilin, 2005).

Context Aware Recommender System is multidimensional which aim to more accurately identify user preferences by exploiting additional dimensions (e.g location, time, season etc) to user and item dimensions.(Adomavicius et. al., 2011). Three main types of models exist for Context Aware Recommender System; Contextual Prefiltering, Contextual Postfiltering and Contextual Model.(Adomavicius and Tuzhlin, 2008). Context is any information that can be used to characterize the situation of an entity. (Dey et al, 1999)

There are three main types of contexts; Explicit context, Implicit context and Infered Context.

Explicit context might be difficult to obtain as it involves the users giving the reason for using, buying or rating the item. Implicit context on the other hand might misinterprete user’s goal, since it is not giving directly by the user but obtained from available data stream.

In this research work we want to see the possibility of inferring user’s context in Contextual Recommendation. On the other hand, Recommender System preciseness can be increased by semantic knowledge gained from Ontology (Uzun and Rack, 2013). Semantic web is an effective infrastructure to enhance visibility of knowledge on the web. It can give well defined meaning which can be understood and processed by machine. The core of the semantic web is Ontology; which is used to explicitly represent our conceptualizations.

The aim of this research work therefore is to develop a model for inferring contextual knowledge from two dimensional data using ontology and making contextual recommendation using probabilistic graphical models. Probabilistic graphical models allow us to represent and visualise the relationships between many variables. Graphical Models help to capture complex dependencies among random variables, building large-scale statistical models and designing efficient algorithms for Inference.(Yunshu, 2013). Towards achieving our aim we have the following objectives;

1. To extract relevant features classes for contextual inferences using One Vs All Classifier.
2. To develop domain ontology for hierarchical conceptualisation and contextual inferences.
3. To develop Structural Parametric Learning model (SPL) for probabilistic recommendation
4. To evaluate the performance of the model with existing methods.

Semantic Knowledge from InternetMovie database(imdb)

Domain Ontology

SPL

Structure

SPL

Learning

SPL

Inference

p1

p2

.

.

.

.

.

.

.

.

.

pn

Query

Genre Classification Model

**Recommendations**

Data Preprocessing

Two dimensional input data from Movielens

Semantic dataset

U \*I R

Preprocesseddata

data

Genre

Classes

Contextualised

data

Figure 1: Research Model

**Research Model**

Our research model is as shown above in figure 1.

1. To achieve objective 1, we follow the model in Figure 2

Genres

[[0 1 1 ... 0 0 0]

[0 0 0 ... 0 0 0]

...

[0 0 0 ... 0 0 0]

[0 0 0 ... 0 0 0]]

One Vs All Classifier

*Movie url*

Imdb Combined Dataset

Movie Genres Data

Multilabel Binarizer

Improved

Genre Classes

Movie Title

(Movielens)

Movie Plot Data

Preprocessing & Feature Extraction

Tf- idf vectorizer

omdb api

Plots

Extracted

Features

Terms Vectors

Figure 2:

1. To achieve objective 2, we followed the model below;

Purpose: Movie Domain Contextual Inferences

Terms Enumeration from Data

Conceptualization

Formalization using owl and Population using cellfie interface

Consistency checking and Inferences

Evaluation using Competency Questions

Figure 3: Ontology Domain Model

1. For objective 3, the model below is used.

Item Data

User Data

Contextual Ontology Data

Rating Data

Probabilistic Recommendation

Structural Parametric

Learning(SPL)

Figure 3: Structural Parametric Learning Model

BICHillClimbing Algorithm

Dtrain

Parameter Learning P(θ|G,D)

Structure Learning

P(G|D)

P(θx1|πx1,D)--------------P(θxi|πxi,D )

Candidate

Ordering(**α**)

Structured Parametric Learning(SPL)

S=(V,E,θ)

Approximate Inference Algorithm

Dval

Query

Specification(e)

Recommendation

Network Structure

p(Xi---n)=

Bayesian Estimator

Figure 4: Structural Parametric Learning Model

We have obtained result for Objectives 1 and 2. These results however serve as inputs to our objective 3. The major challenge is implementing objective 3 which has three parts, structure learning, parameter learning and inference. Below are algorithms to obtain best structure and inference. However, these are not rigid, other better algorithm can be used.

**HillClimbing Search Algorithm to Obtain best Structure**

Input: Dataset D

Output: S’(Structure with maximum score)

Method S = BIChillclimb(D)

1. E 0

2. T ConditionalProbabilityTables(E,D)

3. S (V, E, T)

4. BICScore -∞

5. maxscore BICScore

6. for each attribute pair (P,Q) do:

7. for each E’ {E {P },{Q }{P } }

8. T’ ConditionalProbabilityTables(E’,D)

9. S’ (V, E’, T’)

10. end for

11. end for

12. newBICScore BICScore (S’,D)

13. If newBICScore > BICScore then

14. S S’

15. BICScore newBICScore

16. while BICScore > maxscore

17. Return S

**Approximate Inference Algorithm**

Input: SPL(Variables(xi), cpt)

Output: Sample P(xi|e) given multiple observations e1---------en

1. Sort the nodes xi ------------xn  topologically

2. Initialise w = 1

3. for i = 1--------n:

a) let ui be the current assignment to the parents of xi

b) if xi  e: sample xi from P(xi |ui)

c) If xi e:

i) set xi to ei

ii) multiply w by P(ei|ui)

4. Return (xi,--------,xn), w