Visualization of Bias of Machine Learning for Content Recommendation

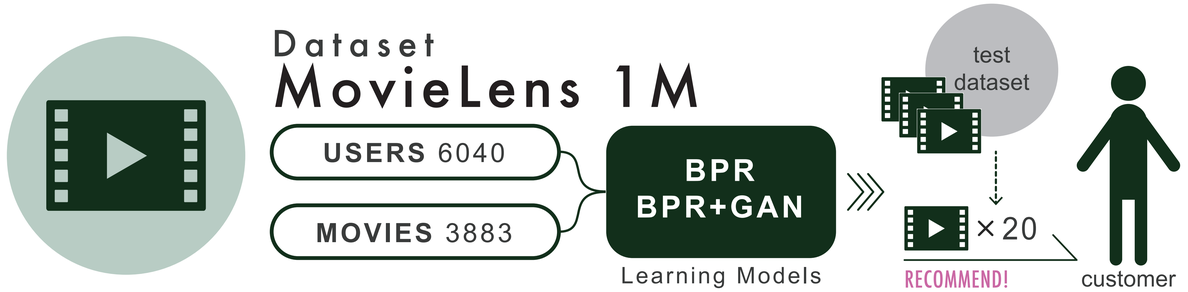
Ami Tochigi[1], Takayuki Itoh[1] and Xiting Wang[2] [1]Ochanomizu University and [2]Microsoft Research Asia

Introduction

- ✓ Excessively personalized recommendation systems may cause biased recommendation results
- ✓ Machine learning techniques have been applied to recommendation engines in recent years

This study

Propose a visualization system to assist the comparison between statistics of appreciation and recommendation of the particular contents/customers and the discovery of the biases in learning results.



VISUALIZATION SYSTEM

VIS 01 SCATTERPLOT

displays the clusters of contents and customers

Color coding 1 **Similarity**

Cosine similarity of feature vectors between the specified customer/content node and other customer/content nodes

customers / contents: similar ⟷ different

Color coding 2 **Difference**

Cosine similarity between feature vectors of each content or between feature vectors of each customer

customers / contents: similar ⟷ different

- **Visualize the patterns of contents and customers**
- **Find the clusters that contain diverse or outlier contents/customers that may cause biased recommendations**

VIS 02 CLUSTER BAR

displays the distribution of features of a user-specified cluster

- **Find specific reasons for biased recommendations**

◆ select **customer** cluster

- Attribute distribution of customers in the cluster
- Attribute distribution of contents selected by customers in the cluster

◆ select **content** cluster

- Attribute distribution of contents in the cluster
- Attribute distribution of customers who selected contents in the cluster

VIS 03 ATTRIBUTES BAR

displays features of a user-specified content or customer

- **Observe the detailed information of contents or customers**

VIS 04 DATA TABLE

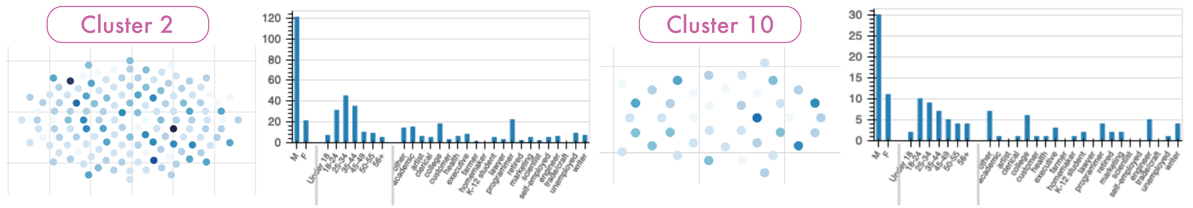
displays the detailed information of a user-specified content or customer

- **Deeply discuss the cause of biased recommendations**

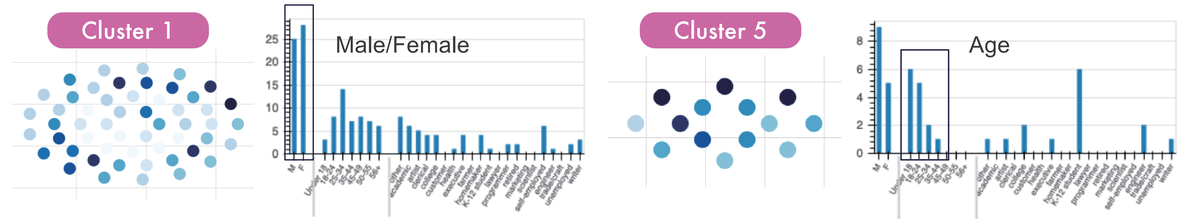
CASE STUDY

VIS01 **SCATTERPLOT** Selected particular cluster

VIS02 **CLUSTER BAR** Customer distribution inthe selected cluster



Clusters with many movies have **small** differences between appreciations and recommendations
(The bar graph shows the customer attribute distribution in each cluster)

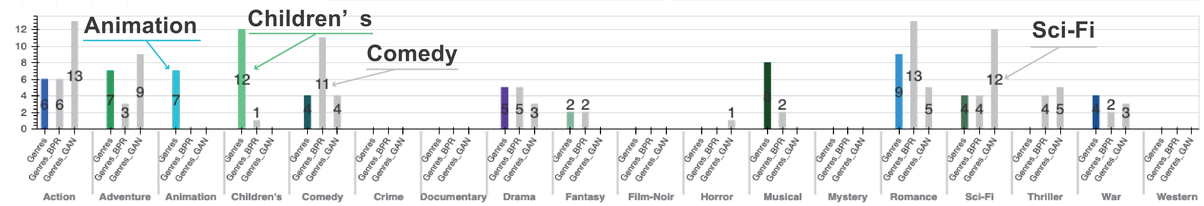


Clusters with many movies have **large** differences between appreciations and recommendations
(The bar graph shows the customer attribute distribution in each cluster)

F, 18-24, college/grad student

VIS01 **SCATTERPLOT** Selected particular node

VIS03 **ATTRIBUTES BAR** Movie distribution watched by customer nodes



Attribute distribution of node with a **large** differences between appreciations and recommendations

Conclusion

This study presented a visualization tool aiming at the discovery of bias of machine learning results of the recommendation systems. Results of case study suggests that the difference in recommendation results among the clusters is caused by the differences in attributes of customers.

Future Work

- ✓ Improvement of calculation method of similarity between nodes
- ✓ Comparison between learning models
- ✓ Selection of quantitative evaluation values

VIS 2021