Project Proposal  
ISE-244 Sec 80 – AI Tools Systm Eng

Uttej Kumar Reddy Gade  
016065543

**What is the problem the selected article is trying to solve and why it is important to study that problem?**

As part of the final project for the course, I am planning on implementing the algorithms in the paper, ‘Improving Recommendation Systems with User Personality Inferred from Product Reviews’. The authors of the paper try to increase the performance of the recommendation system by getting better hit rates by combining personality information with user information. As recommendation systems are a topic which I have not yet explored, I find this interesting to study and understand. Also, the problem of cold-start persists in recommendation system wherein the system takes some time to fine-tune its predictions to the users’ interests. However by inducing personality, which can be inferred from multiple sources such as reviews, browsing history etc. this problem can be addressed to a degree. The authors have in fact proved this by demonstrating an increase of 3-28% by this method. And better performance is needed out of these systems so that users can be directed to data which is more relevant and interesting to them out of the huge data out there.

**What is novel in this paper compared to what came before it. To accomplish that please select an article cited in the references and compare and contrast to it.**

The approach in the paper itself is simple. The authors combine Neural Collaborative Filtering (NCF) which is the first deep learning architecture to be used for recommendations and personality information and conduct experiments to check for performance increase. Out of the many citations in the paper, the relevant technical citations are ‘Improving Socially-Aware Recommendation Accuracy through Personality’ and ‘User Personality and User Satisfaction with Recommender Systems’. Both of these deal with inducing personality information into predictions but neither explored the possibility of utilizing NCF and exploring how personality can be integrated there. This is the major contribution in the paper.

**What has been the impact and/ or limits in the scope of the application of the approach proposed in the selected article. Select an article that cited this article and describe it in context of that. You can use scholar.google.com for getting a list of articles that cited the article.**

The experiments proved that an increase in performance can be indeed gained by this method. A 3-28% increase is significant in this field. However, there are also multiple limitations the authors detailed towards the end. The main ones are how the data is not always consistent w.r.t. personality thus indicating more research is needed on the connection between recommendation and personality for more accurate modelling. Another major limitation was the availability of personality datasets due to privacy and sensitivity concerns which lead the authors to generate their own datasets. This also limited their ability in performing experiments a bit. As this is a recent paper (published on 21 March 2023 according to <https://arxiv.org/pdf/2303.05039.pdf>) there are no citations for this paper.

**A few lines on what is the broader implications of the work?**

Coming to the implications of the paper, on the upside, there is the possibility of people getting better recommendations which would save their time significantly if utilized correctly. On the downside, there would be collection of personality data of users which increases privacy risk. This is a very significant risk, especially after the 2018 U.S. Presidential Candidate scandal, and this makes progress in this direction a more challenging task.

**Summary**

To summarize, my major focus in this project would be the implementation of the deep learning architectures outlined in the paper on the Personality 2018 dataset. I will be applying these architectures on a new dataset that I construct based on the Personality 2018 dataset. OCEAN model scores, which are a type of personality scores are used in the dataset and I plan to convert them to MBTI types and check if there is an increase in the performance.