**Abstract No: SI-06**

## Hybrid recommender system for categorized Sinhala news articles

T. M. S. A. Tennakoon* and G. R. N. A. Gamlath

Department of Information Technology, Sri Lanka Institute of Information Technology, Sri Lanka
sajinitennakoon@gmail.com*

Despite the existence of popular news recommendation systems like Yahoo, there exists no such site for the Sinhala language which provides personalized news recommendations. Hence, there is a need to implement a similar platform which recommends categorized Sinhala news articles according to user preferences. Though, there are several recommender systems that are widely used, the proposed solution is focused on two famous recommender system methods, named Collaborative Filtering and Content-Based Filtering. Collaborative filtering is a method of making automatic predictions about the taste of users by processing information about interactions between many users and items, whereas content-based filtering uses item features to suggest other items similar to what the user likes, based on their previous behaviour. There are many weaknesses of using these recommender systems individually, such as making recommendations for novel users/articles and limited content analysis in item profiles. Therefore, the main objective of this study is to propose a hybrid recommender system that combines both approaches to eliminate the mentioned weaknesses. Furthermore, it has been implemented for recommending Sinhala News Articles specifically, combining Multi-Layer Perceptron with Skip-Gram Architecture, which is novel and not seen previously. The study implements Collaborative Filtering using Multi-Layer Perceptron, which is a deep neural network, owing to its performance on imbalanced data with infrequent users and items. It also uses a deep neural network architecture named Skip-Gram Architecture due to the fact that it predicts similar articles from context words for given target words to implement the Content-Based Filtering. Both of these deep neural networks are combined to create a hybrid recommender system to recommend Sinhala news articles as the methodology of this proposed work. The experiments were conducted using five data sets, which had 17,350 Sinhala news articles. Datasets were prepared by using scrapped Sinhala news articles, categorized under Political, Religion, Crime, Cricket, Football, Rugby, and Entertainment and combined with MovieLens users' and items' profile dataset. For training purposes, several optimization algorithms were used in order to reduce the loss value and a few activation functions were used in order to determine the output of the deep neural network. Few loss functions were used including Mean Absolute Error, to measure the loss in the hybrid model. One hundred and fifty epochs were conducted with 0.2 as the test size for the proposed model. Best performance, that is, the least loss of the hybrid model was given by the Rectified Linear Unit activation function along with Adam optimizer which contains 0.01 as the learning rate accompanied by Mean Absolute Error. The study revealed that the hybrid recommender system used. performed well to eliminate the above mentioned weaknesses found with the other two recommender models, and with higher accuracy. However, the loss decreases with the number of documents in one dataset, and the higher the number of articles, the lesser the loss value. Other than that, a higher number of epochs also helps minimize loss, thereby increasing accuracy.

**Keywords:** Collaborative filtering, Content-based filtering, Hybrid recommender system, Recommender system