1 Parameter sensitivity

Fig. 1 and Fig. 2 illustrate the performance of ViGCN at different SRD thresholds φ from 1 to 10. As the SRD threshold φ increases, the performance gradually rises to the highest value and then declines gradually. A possible reason is that when φ is too small, the model cannot capture enough information from the local context. On the contrary, when it is too large, noise may be introduced into the global information preserved by model.

Fig. 3 and Fig. 4 illustrate the experimental results of different numbers of ViGCN layers from 1 to 8. The model achieves the best performance with a ViGCN layer of 2 on all the three datasets. Hence, when the number of layers is 1, the model cannot capture enough sentence information, whereas when this number is too high, information redundancy and over-smoothing will deteriorate the model's performance.

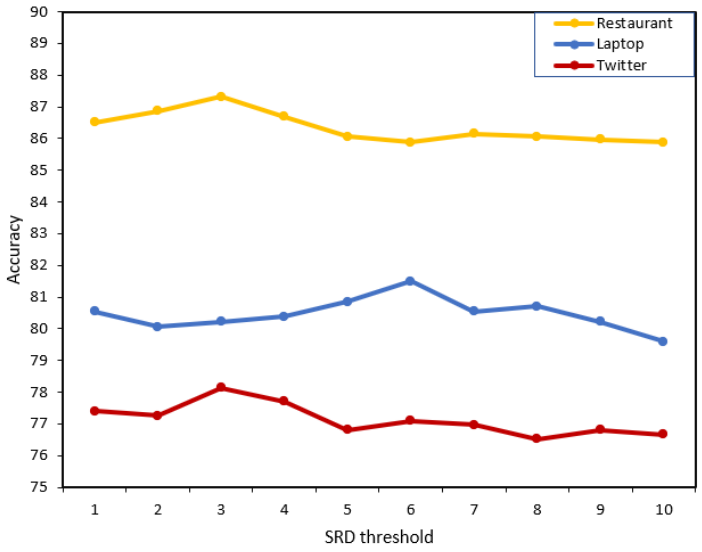

Fig. 1. Impact of the SRD threshold φ on accuracy.

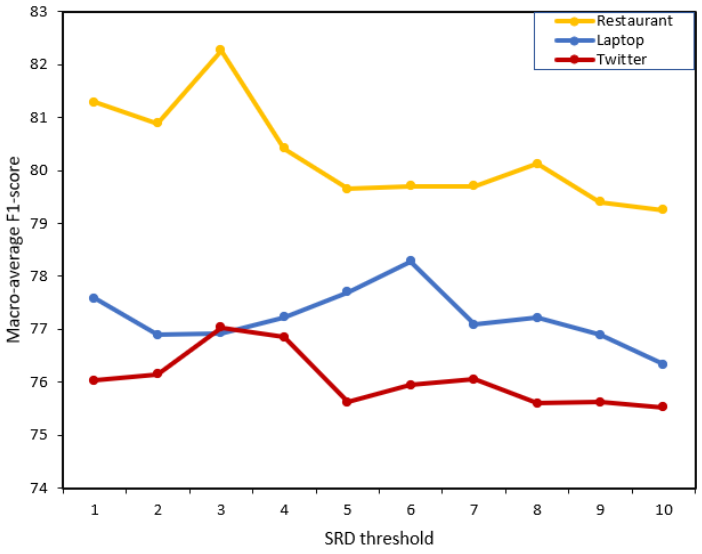

Fig. 2. Impact of the SRD threshold φ on Macro-average F1-score.

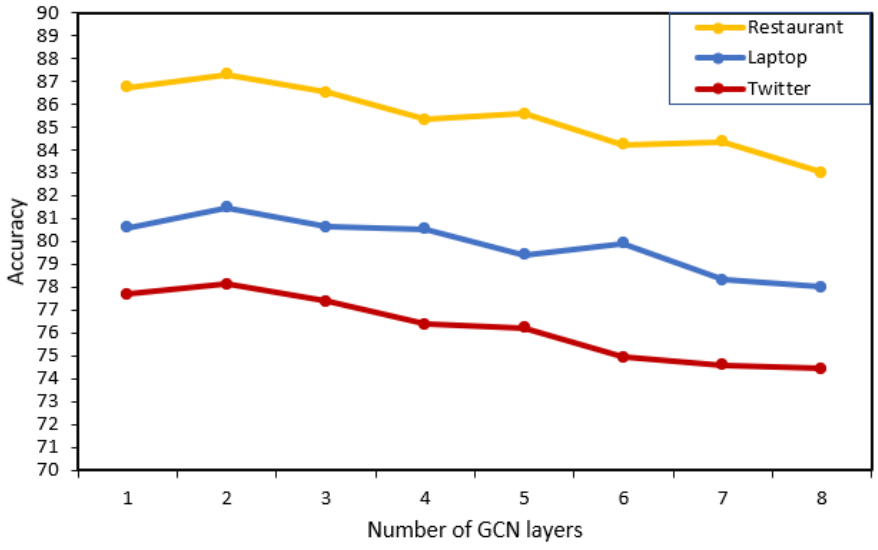

Fig. 3. Effect of the number of ViGCN layers on accuracy.

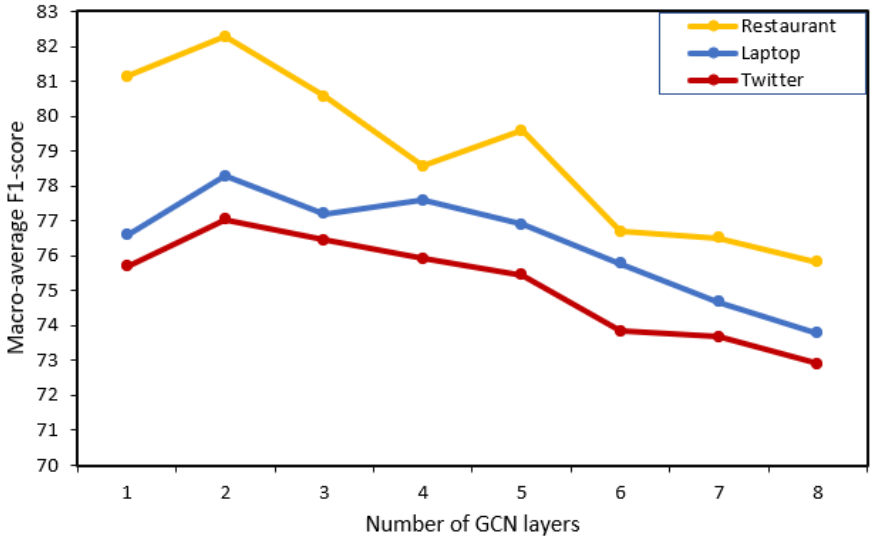

Fig. 4. Effect of the number of ViGCN layers on Macro-average F1-score.