

# 整合數據品質於兩階段深度學習推薦系統

研究生：李力為

指導教授：蔡介元 博士

元智大學 工業工程與管理研究所

## 摘要

Research regarding recommender systems has grown fast in the past years. Specifically, for a collaborative filtering approach, investigations regarding hybrid strategies, deep learning, and the use of text analysis, have allowed enhancing current challenges such as sparsity, cold-start, and accuracy. Still, less attention has been paid to an essential problem that is highly correlated to accuracy: the quality of the ratings as input data. Indeed, some review-rating pairs that a user gave to an item are not consistent, expressing a different opinion. If this phenomenon is ignored, the recommendations may be biased, affecting both users and e-enterprises negatively. For previous reasons, a two-phase framework is proposed. Firstly, a sentiment predictor of textual reviews is created, serving as the quality inspector that cleans and improves the input for a recommender. Numerous deep learning-based algorithms are tested as sentiment predictors and their results are compared between them to determine the best. Secondly, the application of deep learning to model the main recommendation process is experimented, comparing their outcomes against traditional machine learning methods. Thus, the overall scope is to implement a two-step deep learning-based recommender to determine if the application of a raw data quality inspector truly enhances the accuracy of recommender systems. The experiments on one real-world dataset showed the application of a data quality inspector, based on textual reviews, improves the overall performance of a collaborative filtering recommender system, demonstrating the importance of data cleaning to address the accuracy boost challenge. Also, empirical evidence indicates the deep learning approach is the most suitable for modeling the sentiment predictor and the core recommendation process, outperforming the traditional machine learning methods.

關鍵詞：深度學習、推薦系統、協同過濾、文本分類、資料淨化