

A FASHION RECOMMENDATION SYSTEM USING AN
AUTOENCODER NETWORK WITH INTEGRATED
FEATURE EXTRACTION BY CNN MODELS

I. U. Kothalawala

(209346N)

Thesis/Dissertation submitted in fulfillment of the requirements for the
degree Master of Science in Computer Science

Department of Computer Science

University of Moratuwa

Sri Lanka

April 2023

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

.....

Signature

11/07/2023

.....

Date

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of the supervisor: Prof. Indika Perera

.....

Signature of the supervisor

11.07.2023

.....

Date

ABSTRACT

Due to the rapid rise in living standards, the fashion and textile industries have experienced remarkable growth in the fast fashion industry in recent years. On e-commerce websites, the users can find numerous types of fashion items and it can consume lot of time and effort for them to find exactly what they want. Due to this, there is a need for an effective recommendation system that allows users to easily search, sort, and receive pertinent information. In the present, most of the available recommendation systems provide recommendations for a user based on his or her previous activities and behaviors. The problem with this approach is that user preferences can change overtime and new trends can come and go very frequently. Therefore, we cannot expect that the user will like the same type of fashion items even for a short period of time because in the fashion industry designs move very quickly. The other popular approach is that suggesting items based on the information available on similar users. This is not successful either because the given user's preferences are not considered. This paper presents an automated fashion recommendation system that produce recommendations based on an image input by the user. This research also provides more insights on applicability of convolutional neural networks and autoencoder networks for feature extraction along with a comparison of the performance of each of the selected models. This work also suggests a hybrid approach for achieving better performance and evaluation of different ensemble techniques used for combining the outputs of individual models is also provided. The final outcome demonstrates how a hybrid strategy is helping to get results that are more accurate than individual models and selection of the most suitable ensemble technique is also important to further increase the accuracy of the overall system.

ACKNOWLEDGEMENTS

My sincere gratitude is extended to my supervisor, Dr. Indika Perera, for his knowledge, supervision and assistance in ensuring the success of the thesis.

My gratitude goes out to my family for their encouragement and support throughout my life.

I also want to express my gratitude to my MSc batch mates and my coworkers for their assistance and support in managing my research project.

TABLE OF CONTENTS

DECLARATION	I
ABSTRACT	II
TABLE OF CONTENTS	IV
LIST OF FIGURES	VIII
LIST OF TABLES	IX
1. INTRODUCTION	1
1.1 Introduction	1
1.2 Problem in Brief	2
1.3 Aim and Objectives	3
1.4 Scope	3
1.5 Outline	4
2. LITERATURE REVIEW.....	5
2.1 Introduction	5
2.2 Categories of Fashion Recommendation Systems	5
2.2.1 Fashion Image Retrieval	5
2.2.2 Personal Wardrobe recommendation.....	5
2.2.3 Fashion Pairing Recommendation Systems.....	6
2.2.4 Smart or Intelligent Recommendation	6
2.2.5 Social-network-based Recommendation	6
2.3 Algorithmic Models	7
2.3.1 Convolutional Neural Network (CNN).....	7
2.3.2 Recurrent Neural Network (RNN).....	7
2.3.3 Multilayer Perceptron (MLP)	8
2.3.4 Generative Adversarial Network (GAN).....	9
2.3.5 K-Nearest Neighbour (kNN)	9

2.3.6 Autoencoders (AE)	9
2.3.7 Bayesian Networks	10
2.3.8 Multi View Deep Neural Network (MV-DNN)	11
2.3.9 Neural Collaborative Filtering	11
2.3.10 Neural Autoregressive-based Recommendation.....	11
2.3.11 Neural Graph Filtering.....	12
2.3.12 Hybrid Model.....	12
2.3.13 Other Methodologies	13
2.4 Recommendations Filtering Techniques	13
2.4.1 Content-Based Filtering (CBF).....	13
2.4.2 Collaborative Filtering (CF)	14
2.4.3 Model-Based Collaborative Filtering	15
2.4.4 Memory-Based Collaborative Filtering	15
2.4.5 Hybrid Collaborative Filtering.....	16
2.4.6 Hybrid Filtering	16
2.4.7 Hyper personalization Technique	17
3. METHODOLOGY	19
3.1 Introduction.....	19
3.2 Dataset	20
3.3 Data Cleaning and Pre-processing.....	21
3.4 Convolutional Neural Networks	23
3.5 Ensemble Learning [67].....	24
3.5.1 Max Voting	24
3.5.2 Averaging.....	24
3.5.3 Weighted Average.....	24
3.5.4 Bagging	24
3.5.5 Stacking.....	25

3.5.6 Boosting	26
3.6 Early stopping [68]	26
3.6.1 Model Performance	27
3.6.2 Trigger to Stop Training	27
3.6.3 Model Choice	28
3.7 Feature Extraction.....	28
3.7.1 Autoencoder [69]	28
3.8 Cosine Similarity [70].....	30
3.9 Learning Rate Schedules [71].....	31
4. ARCHITECTURE AND IMPLEMENTATION.....	33
4.1 Image Preprocessing	33
4.2 DenseNet161 [72]	33
4.3 ResNet152 [73]	35
4.4 VGG19 [74]	37
4.5 Ensemble Techniques	38
4.6 Training Ensemble Members.....	40
4.6.1 Hyperparameters	40
4.7 Early stopping	41
4.8 Feature Extraction.....	42
4.9 Autoencoder.....	43
4.10 Cosine Similarity	43
4.11 User interface	44
4.12 Overall System Architecture.....	45
5. EVALUATION.....	46
5.1 Introduction.....	46
5.2 Selecting the best classification models.....	46
5.3 Selecting the best ensemble technique.....	46

5.4 Performance of the hybrid models.....	47
5.5 Autoencoder Integration	47
5.6 Choosing the similarity measure.....	50
5.7 Response time	51
5.8 Subjective Analysis.....	51
6. CONCLUSION	57
6.1 Introduction.....	57
7.2 Research contribution	57
7.3 Limitations and Future Work.....	58
REFERENCES.....	60

LIST OF FIGURES

Figure 3.1: Steps of Image Preprocessing.....	22
Figure 3.2: How a variational autoencoder would learn from an input image [69]...	30
Figure 3.3: Cosine Similarity Calculation [70]	31
Figure 3.4: Performance comparison of Different Learning Rate Schedules [71]	31
Figure 4.1: DenseNet Structure [72]	34
Figure 4.2: DenseNet Architecture [72].....	34
Figure 4.3: Multiple Dense Blocks [72].....	35
Figure 4.4: ResNet Architecture [73].....	35
Figure 4.5: Residual Learning – Building block [73]	36
Figure 4.6: Network architecture of VGG-19 model [74]	37
Figure 4.7: VGG-19 structure [74].....	37
Figure 4.8: Architecture of the Autoencoder	43
Figure 4.9: User Interface	44
Figure 4.10: Overall System architecture.....	45
Figure 5.1: Recommendation results without the Autoencoder example 1	48
Figure 5.2: Recommendation results with the Autoencoder: example 1	48
Figure 5.3: Recommendation results without the Autoencoder: example 2	49
Figure 5.4: Recommendation results with the Autoencoder: example 2	49
Figure 5.5: Results of Cosine Similarity	50
Figure 5.6: Results of Euclidean distance	51
Figure 5.7: Recommendations example 1.....	53
Figure 5.8: Recommendations example 2.....	54
Figure 5.9: Recommendations example 3.....	54
Figure 5.10: Recommendations example 4.....	55
Figure 5.11: Recommendations example 5.....	55
Figure 5.12: Recommendations example 6.....	56

LIST OF TABLES

Table 5.1: Performance analysis of the classification models	46
Table 5.2: Accuracies of the ensemble techniques	47
Table 5.3: Performance analysis of the hybrid models	47
Table 5.4: Response time of the system.....	51
Table 5.5: Results of the subjective analysis	53