National College of Ireland

Enhanced Online product similarity classification using description and Images.

MSc Research Project
MSc in Data Analytics

SURESHKUMAR DURAIRAJ
Student ID: x21178933

School of Computing
National College of Ireland

Supervisor: SEAN HEENEY

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	SURESHKUMAR DURAIRAJ
Student Id	21178933
Programme:	MSc in Data Analytics
Year:	2023
Module:	MSc Research Project – RIC – CA2
Supervisor:	SEAN HEENEY
Project Title:	Enhanced Online product similarity classification using description and Images
Submission Due Date:	16/04/2023
Word Count:	4684

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Sureshkumar Durairaj

Date: 16th April 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

For Office Use Only
Signature
Date
Penalty Applied (if applicable)

Contents

red1 Introduction **2**
red1.1 Research Background and Motivation 2
red1.1.1 Image based Classification Techniques 3
red1.1.2 Text based Classification Techniques 3
red1.2 Research Question 3
red1.3 Contribution 3

red2 Related Work **4**
red2.1 Analysis and discussion on the current work 4
red2.2 eCommerce based Similarity classification 4
red2.3 Evolution of Image based Similarity Classification 5
red2.4 Textual Description based Similarity Classification 5
red2.5 Final Discussion and Inference 6

red3 Research Methods & Specifications **6**
red3.1 Data Collection 7
red3.2 Deep Learning Models and its Evaluation 7
red3.2.1 Deep Neural Network 8
red3.2.2 Residual Neural Network 9
red3.3 Implementation 10

red4 Design Specification **10**
red4.1 Implementation 11
red4.2 Project Plan 11

red5 Conclusion and Future Work **11**

red6 Video Presentation **12**

Enhanced Online product similarity classification using description and Images

Sureshkumar Durairaj
21178933
MSc Data Analytics
National College of Ireland

Abstract

In the present-day modern era, digital transactions have become an essential commodity. We have customers across all age groups and these diverse user groups have varied choices, user engagement and user experience is a pivotal part of every e-commerce platform. The usage of similarity classification techniques to help in aiding a personalized recommendation based on the user order preferences to ensure customer engagement and enhance the user experience. These classification aids in availing the best offers matching the products based on image and textual search. Image based and textual description-based similarity classification techniques proved to be an optimal solution when used in conjunction with one another. The usage of Deep Learning image processing algorithms involving RNN for image classification, specifically ResNet-50v2 yields the maximum accuracy and faster response when used in combination with text-based classification such as BERT can perform at best when it is used in a context based scenario such as eComBERT which has proved to perform very effectively with an accuracy when used stand alone with Category Hard,Category Random,Batch Hard strategy . The proposed similarity classification system unlike the TF-IDF + ResNet-18 referred in the base paper aims to provide a better accuracy greater than 73.4% with a rapid response rate.

Keywords — **S**imilarity Classification,RNN,BERT,eComBERT,ResNet-50v2

1 Introduction

The culmination of the digital technologies catalysed by the essence of internet has resulted in the evolution of digital markets. The online commercial industry have fervidly increased to an extent that it has resulted in opening the lattices of the global market thereby empowering a even ground for businesses to compete across the world. When we are dealing with the e commerce platforms, user experience is the penultimate bourne of any type of e-commercial business . As described earlier because the market is plethorated on a global scale user have been equipped with opulent choices which therefore leads to conflict, when there are multiple choices, it leads to confusion as to which one would suit at each contextual point. With the increasing footprint of e-commerce, customers are facilitated with an overwhelming number of product choices. To help them narrow down through these abundant options, e-commerce companies are adopting various technologies to improve their product recommendation systems. These technologies are instrumental in providing personalized and relevant product recommendations to customers, leading to higher customer satisfaction and increased sales. A sample illustration of similar images classified for the data set considered is shown in the Figure 1 below,

1.1 Research Background and Motivation

The most common way of using an online commercial website is by typing the relevant text which directly associates to the product category and brand which produces the same. Text search is the simplest yet effective way for a naive user in an e-commerce ecosystem. As the business competency intensifies, organizations fund heavily on innovative technologies to enable them to distinguish from their competitors , Similarity classification is one such popular technique engineered to cater this purpose . This mainly involves in aiding the customers in yielding the list of products that matches the customer preferences to near reality. There are two main two categories in similarity search, an image-based and a text-based similarity classification. Image-based similarity classification uses computer vision AI

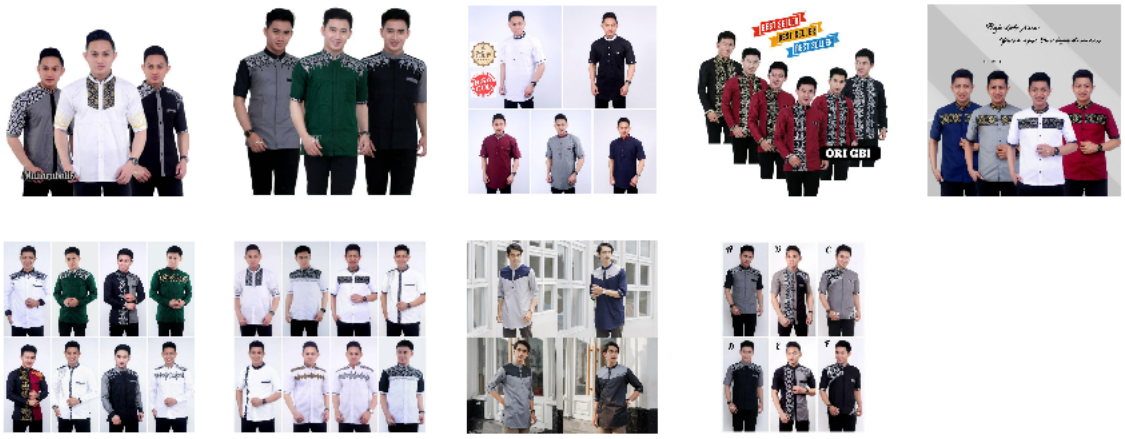

Figure 1: An image of a Similarity Classification

algorithms to analyse product images and find similar products, while text-based similarity classification uses NLP techniques to evaluate product descriptions and therefore identify the product similarities.

1.1.1 Image based Classification Techniques

Thorpe et al. (2018) discusses the boundaries of using a deep learning based Residual network specifically , ResNet-50 for Image based similarity classification . The experimental analysis aids in proving that although there are certain differential limits of RNN when used in an ordinary differential equation , the es sense of regularised result can be achieved by hyper-tuning the training parameter for the ResNet, also the rate of convergence could be effectively handled by our proposed system since we are dealing with the latest of these algorithms which is ResNet-50v2 which aims to eliminate these stochastic limitations.

1.1.2 Text based Classification Techniques

Whereas, Tracz et al.(2020) discusses on the effectiveness of using BERT algorithm for text based similarity learning for product matching especially in the e-commerce platform , which shows that the usage of specific scenario based algorithm such as eComBERT performs better with higher accuracy and results when used with a varied pool of strategies.

These techniques are aimed towards ensuring an enriched user experience by enhancing the customer engagement thereby the e-commerce business can improve the accuracy of the recommendations to their most priceless asset, the user.

1.2 Research Question

The Research question has been evolved from analyzing and spooling through a diverse set of articles involving the similarity classification on eCommerce Platforms. The resource knowledge acquired helps to perform a comparative analysis for online commerce based similarity classification techniques and identify a resolute and refined combination of Image-based and text-based similarity classification techniques,

RQ: Can the proposed system with the combination of ResNet-50V2(Residual Network) - Image based classification approach and BERT - eComBERT - text based classification yield better accuracy on the similarity classification and prediction on the Shopee data set in consideration with a faster response rate.

This research question is focused towards enabling the user with better experience and customer engagement thereby improving the business production and increase the customer retention.

1.3 Contribution

The main intent of this project is to design and develop a deep learning model which will provide a refined similarity classification which aids the customer to search the right product with best offers at the right time enabling the highest level of user using the modernized image and text based classification

approach involving deep learning neural networks. This ultimately yields towards better business for the eCommerce industries .

The upcoming sections contains the following , Section 2 illustrates the study , analysis and inference of previous works which have been performed on the same data set with similar context ,Section 3 contains the evaluation process and section 4 describes the Design and methodology of the proposed mechanism.

2 Related Work

The basic idea of the research has been derived from the (Hari Krishnan 2021) , where the same data set considered for the research is explored . The core concept of this work is to identify not only the matching products , provided it is also equally important to derive the non-Matching products as well. From being a frequent online shopper myself , have come across several instances over multiple online e-commerce platforms , have faced scarcity of offers availability at the right time . The primary modification made to this current research is to include the product item meta heuristics as the prime facet in discovering the match similarity along with the Textual description and Image data set . Also , we have derived certain techniques and methodologies from the works of (Sanaullah Shariff 2022) . Unlike the former (Hari Krishnan 2021) , here they are considering the similarity algorithms such as cosine and jaccard to determine the similarities and dis-similarities .

2.1 Analysis and discussion on the current work

The analysis of the current research is initiated by analyzing the work of (ibid.) , where they have considered the data set from Shopee.com provisioned in the Kaggle available for public use for academic purposes . Here they have used CNN models such as TF-IDF (Term Frequency Inverse Document Frequency) Vectorizer which is a statistic based modelling technique applied to determine the significance of key aspects retrieval , with an attempt to ideate the focus of a specific word in a document, here in the text data set integrated into a common entity called the corpus . This is done by identifying the frequency of word occurrence in the document in our case it is the product title. ResNet 18 a Residual Neural Network which is 18 layers deeper intended to solve complex image processing tasks . ResNet-50 , further enhance RNN which is even 50 layers deep which is provisioned with a pre-trained network from Image-Net comprising of millions of images . For the research considered there are several type of image objects such as computer accessories, Fashion , Electronics , Lifestyle etc., accounting up to 1000 objects .

Additionally they have used Siamese ResNet-50 to help understand the contrasting loss in order to measure the similarity scale between the two objects in consideration , ideally the final layer of the two identical neural network is provisioned into this method , like the previous method this is again 50 layers deep. These experimental comparative analysis helps in understanding the performance criterion between different combination of the Research Methods . Among the various iteration of the models Resnet-18 + TFIDF proved to be the optimal model with a Cross validation score of about 0.73 . Additionally , it is also very critical to derive the analysis from the article (Sanaullah Shariff 2022) . They employ three different deep learning methods such as MobileNet-V2,VGG-19,REsNet-50 which are being trained prior to the experimental analysis and these are run over the given image and text description data set similar to our research context.Apart from finding an efficient , This research (ibid.) also focuses on the performance improvement using the Cosine Similarity ,Levenshtien distance and custom metric score. The Custom metric score achieved for the above said DL methods are 0.7969 , 0.7719 and 0.7518 which shows us that MoblieNet has the maximum performance for this Research Context. However ,these experiments have been made in an isolated fashion thereby not taking advantage of the both the algorithms , hence we have propose to use multiple combinations of ResNet for processing the image data set and multiple combinations of BERT for processing the Text description data set from Shopee.com to design an effective model for identifying the similarity match between the identical / similar products .

2.2 eCommerce based Similarity classification

The essence of internet has paved way for millions of possibilities . The online users need an interactive system which enables them to efficiently utilize their time on any e-commerce platforms , perhaps the most effective way of doing the same is by building an efficient recommendation systems . This article (Abluton 2022) discusses on the usage of over the top / visual recommendations by leveraging the these

Multi-modal capabilities . The experiment is based on an user interaction , where the user gets to upload an imagery of the product he intends to hover around and depending on the search the user will be provisioned with similar products . Here they use visual recommendation and Visual search approach using the image processing algorithm like Yolo , although this is effective in simpler use case these models can be limited to only small / medium scale e commerce platforms because of the model's limitation in performance . However , The visual recommendation approach proposed in this article is an interesting technique which uses K-NN to search the given image with all the other similar images in the data set .

The majority of the online commercial platform provides sufficient loads of multi-modal capabilities which are hardly put to use .A similar modal approach is discussed in the article (Mehta et al. 2022) ,which signifies the importance of usage of fashion BERT and also emphasizes the necessity of using a CNN based model. This particular e-commerce based journal (Hendriksen 2022) cites the usage of these capabilities by considering the primary factor of mismatch between textual description of the product and the corresponding picture of the product category . Therefore , emphasizing on the textual category to image retrieval , image category is extracted based on the actual description of the product by using the CTIR (Category-to Image Retrieval) and FGTIR (Fine-Grained Text-to-Image Retrieval) using the BERT for text classification and match detection. Here they have identified the accomplishment of task with an eye on the ability to reproduce . Using the above mentioned considerations they have employed a CNN-RNN (which is based on BM25,MPNEt,CLIP) model as well a Transformer based model and therefore perform a comparative measure across multiple categories in online commerce . Hence we use the BERT for our research context as it has enriched feature of accessing the multi-modal capabilities and performs efficient when used in particular context as mentioned in the work of (Tracz et al. 2020). Also , the article (Le and Hinneburg 2022) discusses on the product similarity match in e-commerce by considering the facts that the item considered should be popular among users , similar and near-similar products should be flagged as non-matching products . This is done by employing a classification model for entity match and an image processing method such as Siamese CNN ,Ditto model and a combination of both using which they have achieved the F1 score of about 87 .

The product matching is a key factor in our research context , therefore the study is further extended to the article (Peeters and Bizer 2022) , where in a pre-trained transformed based model is considered in a supervised learning environment . This article showcases that supervised techniques when use along with source-aware sampling is proved to improve the performance of the by a huge margin ,the results are measured using F1 score which shows that AptBuy reached about 94% ,Amazon-Google reached about 79% respectively . This shows that further iterative experiments on increasing the data volume has a hit on the performance however this is effective in small volume of training data. Though these are applicable only for smaller data-sets , we can resort to use the F1-score as the performance measure from this article.

2.3 Evolution of Image based Similarity Classification

The Residual Network is considered to be one of the effective deep learning methodology when it comes to product similarity match detection . This article on the Residual Network (Hanif and Bilal 2020) entails further broader understanding on the same . Considering the shortcomings of the ResNet , one of the most cons is the presence of the deep layer stochastic gradient , when there are no explicit regularization . The modernization and recent advancements in the Residual Network has effected in the convergence of gradient flow of the parameters. Also , (ibid.) states that the Residual Network is highly perfromant when compared to AlexNet,GoogleNet,VggNet and MobileNet respectively . Since these ResNets are basically RNN with unique identity mappings. It uses a custom based CoRN on CIFAR data-set for image processing . This shows that out of all the methodologies ResNet with about 110 layers deep neural network on CIFAR-10 datset produces only 1.7% error rate. From the above study and analysis and also considering (Hari Krishnan 2021) , the best way is to go forward with the combinative comparison approach by using ResNEt50 versions along with BERT for Vectorization discusses in the section 2.4 below .

2.4 Textual Description based Similarity Classification

Since we are dealing with online merchant based similarity match detection based on Image as well as textual description . It is important to begin with BERT classification - Bidirectional Encoder Representations , This article (Tracz et al. 2020) deals withe the implementation of BERT based methodology for product matching and offer matching for an e-commerce data-set by applying a transformer based deep

learning architecture with appropriate sampling technique which significantly improves the performance for a varied list of e-commerce. The product matching problem is address by using the zero-shot strategy in order to avoid the heftiness of retraining the model . The consideration of offer matching vs matching and non-matching product is carried out . This is improved by adjusting the network parameters to minimize the triplet loss objectives . The usage of context specific BERT such as eCOmBERT when used in combination with using Category Random(BH) , Category Hard(CH) and Batch hard(BH) strategies respectively yielding an accuracy of 93% when used with CH . The article (Abolghasemi, Verberne and Azzopardi 2022) uses QBD in which the seed document is used as the query and the objective is to retrieve the documents relevant to the search seed. The multi-task optimization is proved to develop the performance thereby increasing the ranking performance . These experimental analysis are intended to effect an early risk prediction of the user in an attempt to enhance the user experience which can be included as our future scope.

2.5 Final Discussion and Inference

From the above analysis on various research articles on the existing methods , it can be inferred that for our research we can use ResNet methodologies for image processing based similarity match detection and BERT for text based vectorised approach and use the performance measure such as Cross Validation score , F1-Score , Cosine Similarity , Levenshtien distance and few other metrics .

3 Research Methods & Specifications

This paraphrase primarily describes and focuses on the emphasis of the methodology,Models and possible assumptions used for arriving at an affluent algorithm for determining the similarity classification and also it elucidates the details of the framework that is being considered for the exploration of this study. For simplicity , we are considering to carry on a five phased KDD (Knowledge Discovery Database) approach .This framework is quintessential in ensuring the problem to be approached in a refined and thorough manner and orchestrates a road map for the process to be carried out in order to accomplish the tasks effectively. The frame work below is specifically designed to arrive at finite set of process at a detailed level on the extraction of data so as to solve the problem effectively .Therefore this gives the outline for carrying out the process starting from data extraction to the finalized similarity classifying model and therefore resolve the task in hand successfully by evaluating the results.

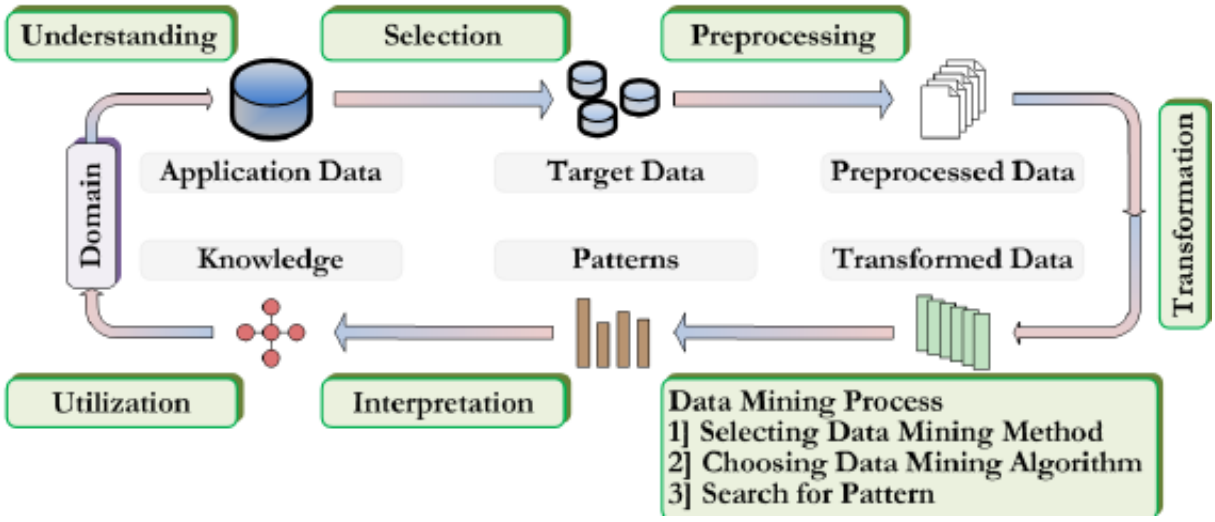

Figure 2: Processing Steps in Knowledge Discovery Database

The proposed mechanism above primarily involved two stages of research , to build the similarity classification model and evaluate the result for at a finite iterations for effective results. The first stage is primarily to use the ResNet-50V2 (RNN based Deep Learning Neural Network for Image based similarity classification) which when processed will be held in KB (Knowledge Base) . The second stage involves the result and further extend the classification and analysis by subjecting the same to Textual Description based similarity classification algorithm eComBERT(BERT) .Finally the similarities arrived

from both the classification strategies are compiled and integrated together to produce a refined Model which can help the customer to identify the matching and non-matching products and offers for a given product. The details of each of the steps involved are described in the following sections and therefore the outcomes of the proposed system .

3.1 Data Collection

The framework described above for performing the Product Similarity Match Detection using the eCommerce based images and the textual description is performed and analyzed using the Shopee data set which has been published by the Shopee online commercial business organization . The data set considered is made available for academic research in public domain , Kaggle for competition by 2020. The data set is not provided directly as a image and text set , rather they provisioned the researchers with a link to fetch the data directly from the Shopee Website. The data set in consideration consists of about 32,400 images and textual description as a text file of the online commercial products from a varied list of product categories . The analysis performed on these data sets is yield the result set containing the identical products among the list of products listed by subjecting them to the considered and possible iterations of training and testing . The Figure below describes the key attributes of the data set such as product title , product image and image phash . However few other attributes such as Label Group and PostingId should also be taken into consideration . A sample of the data set describing how the input data set looks like is as shown in the Figure . The upcoming step is to analyse the detailed traits of the image and text data set separately . Further down the data sets should be classified into two main categorical types such as Match Detected and No-Match Detected by associating them into different groups.

posting_id	image	image_phash	title	label_group
train_129225211	../input/shopee-product-matching/train_images/...	94974f937d4c2433	Paper Bag Victoria Secret	249114794
train_3386243561	../input/shopee-product-matching/train_images/...	af3f9460c2838f0f	Double Tape 3M VHB 12 mm x 4,5 m ORIGINAL / DO...	2937985045
train_2288590299	../input/shopee-product-matching/train_images/...	b94cb00ed3e50f78	Maling TTS Canned Pork Luncheon Meat 397 gr	2395904891
train_2406599165	../input/shopee-product-matching/train_images/...	8514fc58eafea283	Daster Batik Lengan pendek - Motif Acak / Camp...	4093212188

Figure 3: An image of a Similarity Classification

3.2 Deep Learning Models and its Evaluation

There has been n number of instances where we could observe e-commerce business faces a deficit on the refined product Recommendations since the existing system mainly involves the discovery and comparison of multiple products using text based methods . Hence we need to employ by conducting multiple experimental analysis to solve the above mentioned issue , the following are the different kinds of strategies which can be considered for the research .

Convolutional Neural Networks (CNNs): CNNs are commonly used for image classification and object detection tasks. They are designed to automatically learn and extract important features from images by convolving a set of learnable filters over the input image.
Recurrent Neural Networks (RNNs): RNNs are commonly used for sequential data such as time series data and natural language processing tasks. They have a "memory" that allows them to process and

analyze data that has a temporal relationship.
Transfer Learning: Transfer learning is a technique that involves using pre-trained neural networks as a starting point for a new task. By leveraging the pre-trained weights, the model can be trained on a new task with fewer data samples and in less time.

Data Augmentation: Data augmentation is a technique that involves creating new training data from existing data by applying random transformations such as flipping, rotating, and cropping. This can increase the amount of training data and reduce over fitting.

Regularization: Regularization techniques such as L1 and L2 regularization can help prevent over fitting by adding a penalty term to the loss function that encourages the model to have smaller weights.

Dropout: Dropout is a regularization technique that randomly drops out neurons during training to prevent the model from relying too much on any one feature.

Batch Normalization: Batch normalization is a technique that normalizes the inputs to each layer, which can help speed up training and improve generalization.

Hyper parameter Tuning: Deep learning models often have many hyper-parameters that need to be tuned, such as learning rate, batch size, and number of layers. Hyper-parameter tuning involves systematically searching for the best combination of hyper-parameters that result in the best model performance.

There are multiple deep learning methodologies which are encompassed within the following three categories , Supervised (CNN, Multi Layer Perceptron and RNN), Unsupervised(Auto-Encoder, Restricted Boltzmann Machine,Deep Belief Network) and hybrid learning models (Deep Transfer Learning and Deep Reinforcement Learning). The proposed system is aimed towards developing a model with training the model from the root level and analysing the accuracy of the iterative and varying hyper tuned model and arrive a best fit model will be opted for finalized model. The figure 4 shows the process involved in the deep learning framework for similarity identification as described in the fore-coming sections.

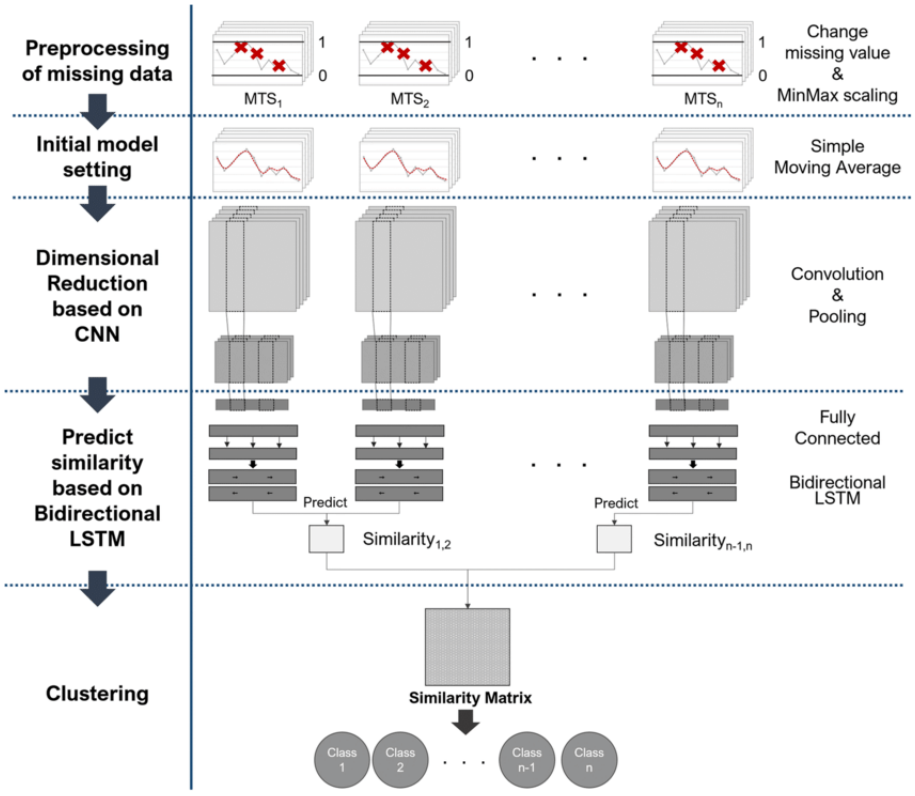

Figure 4: Process of Deep Bidirectional Learning for Similarity Learning Model

3.2.1 Deep Neural Network

Since we are starting the research from the scratch, the analysis is initiated by following the below steps for methodology for performing the similarity detection on the e-Commerce data as shown in the Figure 5 . The below flowchart in the Figure 5 explains the workflow of the proposed system for analyzing

the given image and text data set and produce Similarity match detection . Get Data : This is the

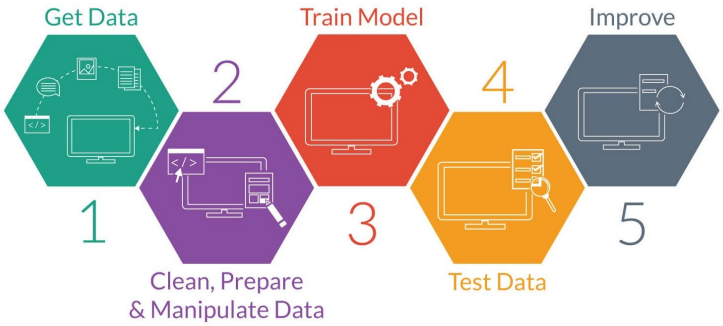

Figure 5: Methodology for Similarity Detection

initial step , where we will bet fecthing the Image and text data set using python snippets from the links provided in Kaggle for directly accessing these data sets . Understand the metadata and characteristics , as mentioned in the above sections we are considering only certain columns which are the potential independent variables for developing the model . When the data is loaded , the processing steps which are required for the Similarity match detecting model like resizing the data set , encoding the data and modifying the pixelation etc.could be performed. Once split, a Data Loader is facilitated with collection of tuples along with the block size and data randomization for every epoch could be determined .

Clean Prepare and Manipulate Data : This is also a critical step since for this research the data set considered is huge and real-world data set it is likely to contain noisy or missing data which has to be cleaned . It is highly essential to spend ample time on this phase as this forms the baseline for our research . Hence we need to make sure to process only valid data and minimize the processing time and resource utilization on invalid records. They can then be sliced or resized into marts for further processing to be collaborated finally in yielding the model . For simplicity we can split them into test and training set ideally in the 80:20 ratio.

Train Model This is where we apply the data set against multiple versions of ResNet algorithms to compare the efficiency and measuring the accuracy of each of the individual models . (Hanif and Bilal 2020) states that in comparison with the high end Image based Neural Networks such as AlexNet, VggNEt, GoogleNet the ResNet has proved to yield much better performance , particularly in image classification and object detection which exactly matches our research context. Hence we are considering to use the latest version of the ResNet i.e., ResNet-50V2 . Similarly we use BERT algorithm for the similarity match detection through textual description . The journal (Abolghasemi, Verberne and Azzopardi 2022) cites that BERT when used in specific context is expected to achieve maximum performance , hence eComBERT and along with the standard BERT, HerBERT with varied strategies should also be considered for comparative analysis . These two algorithms ResNet-50V2 and eComBERT is expected to arrive at a refined model for similarity match detection as per the discussion and analysis above .

Test Data The results obtained in the model arrived using both of the above mentioned algorithms should be evaluated iterative subjecting to different sample subsets to arrive at a refined model with higher accuracy .

Improve Depending upon the results and analysis from the intermediate models; appropriate slight changes in the hyper-tuning , hyper parameters and subset swapping are performed until we arrive at a maximum performance . With the tool-set available , we can make use of them to automate these improvements iteratively by auto setting values .

3.2.2 Residual Neural Network

Residual Neural Networks (ResNets) is one of the classification type of Convolutional Neural Network (CNN). ResNets were originally introduced for image classification tasks and extended to varied image processing tasks by applying numerous computer vision tasks , such as Similarity match detection and classification been applied to various computer vision tasks, such as object detection and segmentation.

ResNets follows a pattern wise similarity to the structure of other types of CNNs, with convolutional layers and pooling layers used to extract features from the input image. However, the key aspect of these type of CNN is the use of residual connections, which allow the network to learn residual mappings and enable the training of much deeper networks than was previously possible as described in the Figure 4 mentioned above.

Traditionally, deep neural networks suffer from the problem of vanishing gradients, where the gradients become very small as they are propagated through many layers, making it difficult for the network to learn useful features. ResNet addresses this problem by introducing residual connections that allow the network to skip over layers and propagate the input directly to the output of a later layer. This allows the network to learn residual mappings, which are the difference between the input and the output of a layer, instead of trying to learn the entire mapping from scratch . The usage of RNN where the image data set is divided between multiple blocks of residual layer and arrive at a max pool architecture towards determining the similarity between them is illustrated in the figure 6 as mentioned below . Further more as discussed in the study of (Thorpe and Gennip 2018) , it is evident that Residual Network could address the stochastic limit thereby resolving the conflict of gradient descent in case there are no presence of explicit regularization and hence ensuring the convergence of gradient flow of attributes.

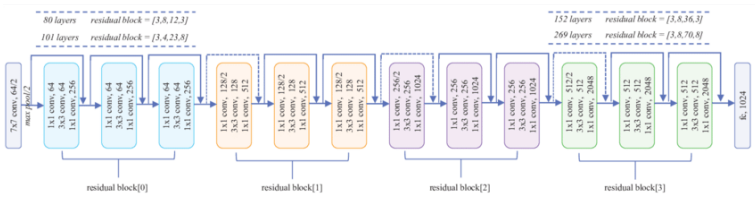

Figure 6: Process of Deep Bidirectional Learning for Similarity Learning Model

3.3 Implementation

The implementation of this research work is obtained is such a fashion to determine and bring out which of the set of business products image and the corresponding text matches the same e-commerce item.The above could be hindered by few of the implementation difficulties such as ,

* To identify the similar duplicate product level items specifically instead of only the pictures.

* To understand and evaluate the significance of removing the atmosphere and background of the given product items specifically

* By utilizing the textual description of the product items or possibly the labelled titling of the product item itself

4 Design Specification

The matching pattern of the images and text and the match prediction result depend heavily on the deep learning methodologies as described in the Section 3.2.1 . A multiple combination of pre-trained deep learning models are therefore used in the indication of matched product items . From the previous discussions in section 3.2.1 we can then utilize transfer learning which uses the value that has already been computed from one model and function to achieve accurate results . As per the discussion in the sections 3.2.1 ,3.2.2,3.2 the final frame work will consists of the implementation of the models using the following combination listed below in the figure 7 ,

Data Collection	Pre-Processing	Modelling	Evaluation
Shopee DataSet	Vectorization	BERT/ecomBERT	Validation Loss
Sourced at Shopee website	Embedding	ResNet-18	CV Score
Product Title & Images	Augmentation		Experiments
JPG and CSV format	Remove Duplicates	ResNet-50 / ResNEt-50V2	Model Comparison

Figure 7: Functional Layers

4.1 Implementation

The details of the technique to be used for accomplishing the research and how they will be implemented has been discussed in detail in the section 3.2.1 above. Henceforth for accomplishing the implementation of the above mentioned tasks , the design specification flow framed is illustrated in the Figure 8 below , which clearly illustrates the difference between the two layers (presentation Layer and Business Logic Layer) .

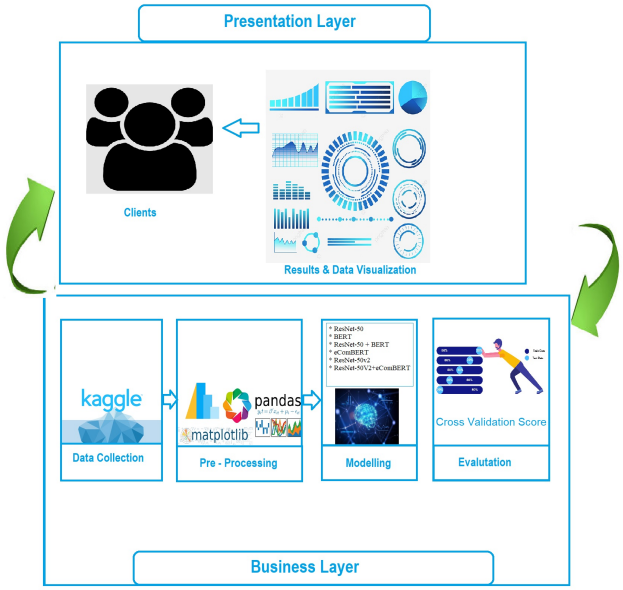

Figure 8: Design Specification Framework

Presentation Layer :

This is where the user will get to visualize the similarity match detection results as an outcome of the research yielded by various versions of the above mentioned methodologies and presented as a single file showing the similarity of products vs another product/offers.

Business Logic Layer :

This is where the actual Engine lies upon , the core of the work is performed here , beginning from inferring , interpreting and enchantment of data to be redily processed by the DNN , RNN and other Deep learning techniques using python as the tool base.

4.2 Project Plan

The following Gannt Chart in the below 9 shows the list of completed and task and pending task which are scheduled and lined up to be accomplished in the future,

5 Conclusion and Future Work

The product similarity match detection is very essential for any e-Commerce user to get hold of the offers for the right product at the right in a robust manner . The match detection using the given image and text description data set implemented by various Deep learning and RNN algorithm is the crucial part of the research. The implementation of the previously mentioned steps will be worked up on and delivered in the future in the upcoming weeks as per the schedule provided in the figure 9

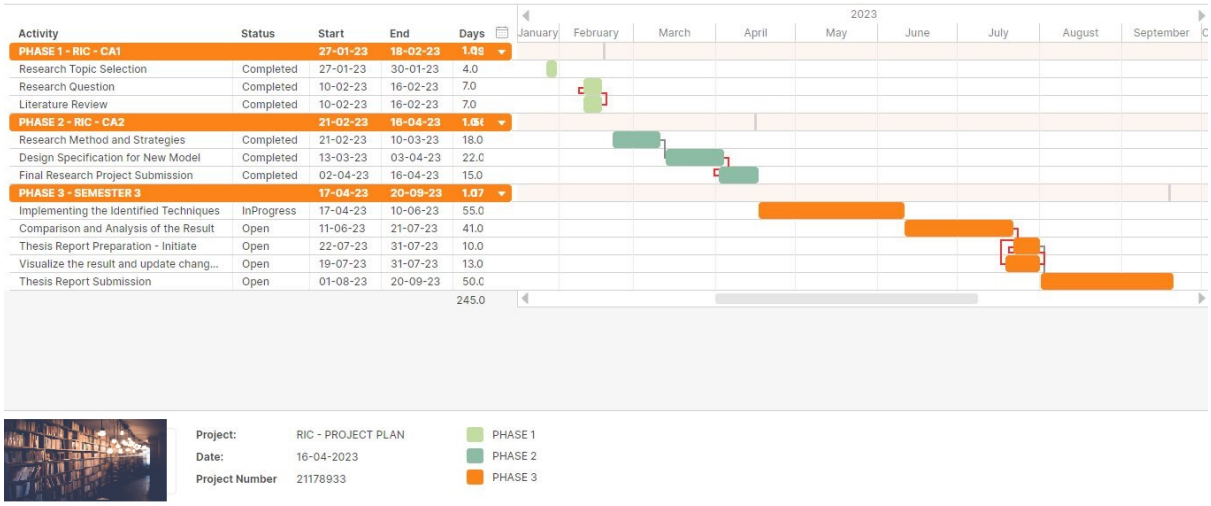

Figure 9: Project Plan for Research Work

6 Video Presentation

The Video presentation containing the clear scope and flow of the project , including the techniques and Data-sets / Pre-processing steps are available in the URL below .

Video Presentation Link

References

Abluton, Alessandro (2022). 'Visual Recommendation and Visual Search for Fashion E-Commerce'. In: *Similarity Search and Applications: 15th International Conference, SISAP 2022, Bologna, Italy, October 5–7, 2022, Proceedings*. Springer, pp. 299–304.

Abolghasemi, Amin, Suzan Verberne and Leif Azzopardi (2022). 'Improving BERT-based query-by-document retrieval with multi-task optimization'. In: *Advances in Information Retrieval: 44th European Conference on IR Research, ECIR 2022, Stavanger, Norway, April 10–14, 2022, Proceedings, Part II*. Springer, pp. 3–12.

Hanif, Muhammad Shehzad and Muhammad Bilal (2020). 'Competitive residual neural network for image classification'. In: *ICT Express* 6.1, pp. 28–37.

Hari Krishnan, Kannan (2021). 'E-commerce Product Similarity Match Detection using Product Text and Images'. PhD thesis. Dublin, National College of Ireland.

Hendriksen, Mariya (2022). 'Multimodal Retrieval in E-Commerce: From Categories to Images, Text, and Back'. In: *Advances in Information Retrieval: 44th European Conference on IR Research, ECIR 2022, Stavanger, Norway, April 10–14, 2022, Proceedings, Part II*. Springer, pp. 505–512.

Le, My Hong and Alexander Hinneburg (2022). 'An Application of Learned Multi-modal Product Similarity to E-Commerce'. In: *Similarity Search and Applications: 15th International Conference, SISAP 2022, Bologna, Italy, October 5–7, 2022, Proceedings*. Springer, pp. 25–39.

Mehta, Karan et al. (2022). 'Multimodal Classification in E-Commerce: A Systematic'. In.

Peeters, Ralph and Christian Bizer (2022). 'Supervised contrastive learning for product matching'. In: *arXiv preprint arXiv:2202.02098*.

Sanaullah Shariff, Zahra Fathima (2022). 'Product Matching for E-commerce Platform based on Text and Image Similarity using Deep Neural Network Architecture'. PhD thesis. Dublin, National College of Ireland.

Thorpe, Matthew and Yves van Gennip (2018). 'Deep limits of residual neural networks'. In: *arXiv preprint arXiv:1810.11741*.

Tracz, Janusz et al. (2020). 'BERT-based similarity learning for product matching'. In: *Proceedings of Workshop on Natural Language Processing in E-Commerce*, pp. 66–75.