Journal of Informatics and Web Engineering
Vol. 2 No. 2 (September 2023) eISSN: 2821-370X

Multi-Label Classification with Deep Learning for Retail Recommendation

Zhi Yuan Poo[1*], Choo Yee Ting[2], Yuen Peng Loh[3], Khairil Imran Ghauth[4]
[1,2,3,4] Faculty of Computing and Informatics, Multimedia University, Malaysia.
corresponding author: (1181103059@student.mmu.edu.my, ORCiD: 0009-0005-6443-8794)

Abstract - Selecting the right retail business for a location is crucial for the success of a business because it determines the likelihood of favourable return on investment. One common approach used in retail recommendation is multi-class classification, where retail businesses are categorized into different classes or categories based on various features or attributes. Existing research in the field of retail recommendation has extensively proposed and evaluated different algorithms, techniques, and approaches for multi-class classification in the context of retail recommendation, however, limited work has been focusing on formulating retail recommendation as a multi-label problem. This is because in retail recommendation, one location can fit multiple retail businesses so that it can provide more options to recommend the most suitable business for the location. Therefore, multi-label classification will be attempted in this study. An analytical dataset will be constructed that provides comprehensive insights into the characteristics of the business area, and subsequently employ deep learning technique for multi-label classification. The analytical dataset is constructed based on the list of sites of interest data from YellowPages, population data from Humanitarian Data Exchange (HDX) and property data sourced from brickz.my. This work will be focusing on implement deep learning technique which is 1D convolutional neural network (CNN) model. The findings showed that the proposed model achieved 61.22% in terms of accuracy.

Keywords—multi-label classification, deep learning, retail recommendation, 1D convolutional neural network (CNN), YellowPages

Received: 6 July 2023; Accepted: 25 August 2023; Published: 16 September 2023

I. INTRODUCTION

A retail business refers to a business model that involves selling goods or services directly to consumers. It can span various sectors including, but not limited to, food and beverage establishments, clothing stores, electronics shops, and more. The retail business operates through various distribution channels especially physical storefronts [1]. They are often characterized by their accessibility to the public, offering a wide range of products or services catered to diverse customer needs and preferences. In this dynamic business landscape, making informed decisions about which retail business is suitable to be establish at a business location becomes crucially important.

Determining whether a location is suitable for a specific retail business is a crucial step in the process of opening a new store. This is because location is the main key factor to boost one's retail business performance which can create competitiveness in the surrounding retail neighbourhood [2, 3, 4]. However, due to the increased variety of retail types in urban areas, the placement of retail stores has become more complicated compared to the past [5]. Due to this circumstance, it is crucial to evaluate various factors such as demographics, geographical information,

Journal of Informatics and Web Engineering
https://doi.org/10.33093/jiwe.2023.2.2.16

Published by MMU Press. URL: https://journals.mmupress.com/jiwe

foot traffic, competition, accessibility, and an associated number of other factors to successfully recommend the right retailer for a certain business area [6]. These elements may be crucial to a retail company's performance and may have a big influence on its bottom line [3]. Retailers may decide if a site is good for their business by carefully evaluating these criteria and thoroughly reviewing data such as population demographics, foot traffic data, and local market research. This process can help ensure that the business is in the right location to attract customers, increase sales and ultimately grow the business.

Considering the factors as mentioned above, it becomes evident that a single location may possess attributes that can cater to multiple retail businesses simultaneously. As a result, the retail recommendation problem necessitates a more sophisticated approach than traditional single-label classification methods. Multi-label classification, which allows for the simultaneous prediction of multiple labels associated with a single instance, is particularly well-suited for retail recommendation tasks given the diverse range of retail business options that a target business location may possess. In the context of retail recommendation for a particular business location, a single location can simultaneously benefit from multiple viable recommendation choices, such as McDonald's, Starbucks, and Pizza Hut. For instance, a bustling urban area may be suitable for all three retailers, as their presence would cater to diverse customer preferences and capitalize on high foot traffic. By employing multi-label classification, an analytical dataset can be effectively constructed, where each location can be associated with one or more retail business. This approach not only captures the complex relationships among various retailers but also allows for a more nuanced understanding of the preferences and needs of a specific location, ultimately leading to more accurate and relevant recommendations.

The objective of this paper is to construct an analytical dataset to give a clear insight of the surrounding characteristics of the business area. Furthermore, this paper aimed to employ a deep learning model for multi-label classification problems which is scoped to the retail recommendation.

II. LITERATURE REVIEW

A. Variables for Retail Recommendation

Making a recommendation for a retail business at a specific location of interest is not a straightforward task as it involves a meticulous examination of demographic data, trade area characteristics, sales performance, traffic patterns, and environmental factors [7]. Certain variables from geography, demographics, trade area, and environment have been proven in studies to play a vital role in retail recommendation, due to the impact they have on retail performance. Although domain experts could extract relevant features, identifying the appropriate feature set can be a difficult and time-consuming task [8]. The features are a key consideration while choosing a relevant retail business for a specific location. There are several important factors that might influence the decision. [8] examined a variety of criteria to select a retail location, including points of interest (POI), surrounding property or neighbourhood, population, and education data. The features can be distinguished into geographical features and demographic features.

Geographical characteristics are variables that surround the retail site, such as points of interest (POI), surrounding properties or neighbourhoods, and transportation. All the location attractors are referred to as location features. [9, 10] mentioned that geographical factors play a significant role in shaping consumer behaviour in traditional brick-and-mortar stores. [7] conducts an experiment in which they use food and beverage data to train a computer algorithm to predict retail business in a specific location. The results of the paper showed that it has a high mean accuracy. As a result, it is highly possible that there are strong dependencies between the retail sector in an area [7]. [11] extract geographical features such as POI, check-in data, user reviews from diverse urban data sources and incorporated into various machine learning models to forecast the popularity of a potential retail store in the target area. The popularity of a store in urban areas is connected to the spatial features of its location. Thus, [10] consider the following features such as traffic convenience, POI, and surrounding neighbourhood to represent the geographic characteristics of a store and its surrounding area.

Demographic features include the area’s population, education data, economic data, and POI customer number. These are the facts that most researchers use while deciding on a retail location. It may be used to anticipate the popularity of an area and locate the best retail outlet. The problem of retail recommendation is markedly complex and multi-criteria, as it is influenced by a variety of various factors, one of which is demographic variables, which include population size, population density, and age profile [5]. To identify the potential customers in various communities near the retail store, [10] consider various demographic profiles and gather statistical data about the number of individuals with different profiles in each community, including the number of men or women and the number of people with varying levels of income. According to [12], any group of people in the vicinity of the retail industry can be potential consumers. If the assumption is made that each person has an equal chance of visiting the retail store, then a larger number of people in the area would lead to a greater customer flow and higher sales volume for the retail store. As a result, the population factor is a crucial aspect to consider when choosing the optimal location for the retail store. The Table 1 shows that the comparison between geographical features and demographic features used by other researchers.

Table 1. Comparison between geographical features and demographic features used by other researchers.

	Geographical			Demographic			
Authors	Point of Interest (POI)	Nearby Property or Neighbourhood	Transportation	Population	Education	Economy	POI Customer Number
[48]			✓	✓			
[8]	✓	✓		✓	✓	✓	
[6]	✓	✓	✓	✓			
[11]	✓	✓					✓
[28]	✓	✓					✓
[9]	✓	✓	✓	✓			✓
[47]	✓	✓	✓		✓		✓
[38]		✓		✓		✓	
[27]		✓	✓	✓			
[7]	✓	✓		✓	✓		
[35]	✓	✓	✓	✓			✓
[40]		✓	✓				
[10]	✓	✓	✓	✓			
[12]	✓	✓	✓		✓		
[2]		✓					
[39]		✓		✓			
[32]	✓	✓	✓				
[49]		✓					

B. Techniques for Retail Recommendation

Different techniques of retail recommendation will perform differently based on their desired output work. As a result, the literature review among the papers will help to identify the best technique, as each has advantages and drawbacks. Most of the researchers using four types of techniques to perform retail recommendation such as Machine Learning, Deep Learning, Analytic Hierarchy Process (AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS).

Machine Learning is a branch of study focused on comprehending and constructing methods that have the capability to learn, that is, methods that use data to enhance their performance in specific tasks. The algorithms of machine learning create a model based on the training data, and this model allows them to make predictions or decisions without explicit programming. Besides that, Deep learning is a branch of machine learning that utilizes artificial neural networks, which are algorithms based on the anatomy and operation of the human brain. While standard machine learning algorithms are linear, deep learning algorithms are designed in a hierarchy of increasing complexity and abstraction. Moreover, AHP is a well-organized method for managing and evaluating complex decision-making scenarios, it represents an accurate approach to quantifying the weights of decision criteria. Lastly, TOPSIS is a multi-criteria decision analysis approach that determines the best option by comparing the geometric distances between each candidate and the ideal solutions. The ideal choice is the one that is closest to the positive ideal solution (PIS) and farthest from the negative ideal solution (NIS). The Table 2 shows that the comparison between geographical features and demographic features used by other researchers.

Table 2. Comparison of Retail Recommendation Techniques with Retail Recommendation Work

	Retail Recommendation Techniques			
Authors	Machine Learning	Deep Learning	AHP	TOPSIS
[48]			✓	
[6]		✓		
[9]		✓		
[28]	✓			
[11]	✓			
[38]			✓	✓
[35]		✓		
[27]			✓	✓
[40]			✓	
[12]	✓			
[32]		✓		
[49]		✓		

C. Multi-Label Classification

One of the key areas of discussion in the field of machine learning is classification. The goal of classification is to create a model that, using a set of labelled examples, can correctly categorize fresh unlabelled data [13]. On the other hand, multi-label classification is a classification variation in which one instance may be simultaneously linked with several labels. Multi-label classification issues enable instances to have many valid labels, in contrast to typical classification problems where each instance is only linked with one label. It is beneficial because it enables the recording of intricate connections between occurrences and labels. This is crucial in real-world situations because instances may have a variety of qualities or traits. Multi-label classification can also be applied to issues when standard single-label categorization falls short in capturing the subtleties of the issue. Multi-label classification has been a prominent issue in data mining and machine learning research in recent years due to its broad variety of applications, including social network analysis [14] and photo or video annotation [15]. There are several difficulties that researchers must overcome while tackling multi-label classification assignments. Multi-label classification (MLC) datasets, for instance, have an unbalanced characteristic where samples and labels are not evenly distributed throughout the data space [17]. Many newly published research papers have focused on imbalanced MLC in recent years. Research on imbalanced MLC has been significant, indicating that the problem remains a valuable interest to researchers [18, 19]. Therefore, there are two main approaches that used by researchers to deal with the multi-label classification problem, which are Machine Learning and Deep Learning. The Table 3 shows that the techniques used by researchers for multi-label classification.

Table 3. Comparison of Multi-Label Classification Techniques

	Multi-Label Classification Techniques	
Authors	Machine Learning	Deep Learning
[21]		✓
[46]	✓	
[33]		✓
[24]		✓
[25]		✓
[26]		✓
[37]		✓
[29]		✓
[23]		✓
[45]	✓	
[42]	✓	
[50]		✓
[44]		✓
[43]		✓
[36]		✓
[22]		✓
[16]	✓	
[34]	✓	
[30]	✓	
[13]		✓
[41]		✓

III. ANALYTICAL DATASET

This section discusses the dataset and its source, all the data will be used to construct an analytical dataset to perform the prediction in this paper.

A. Model Performance (Refer to Table 4)

- *Point of Interest Data (D_{poi}):* Point of Interest (POI) data is a collection of retail business listings in Malaysia up until 2022. The information was sourced from the online business directory, Yellow Pages. The POI dataset includes 443,041 entries, each consisting of the retail business's address, category, latitude, longitude, name, and state.
- *Base Data (D_{base}):* The base data is a portion of the POI data that includes retail businesses of interest, such as restaurants and cafes. The base dataset includes information such as the retail business's address, name, latitude, and longitude.
- *Residential Property Data (D_{ppt}):* The residential property data is a collection of properties obtained from Brickz.my and the Value and Property Services Department (JPPH). The data covers the period from 2000 to 2019 and includes 491,721 entries with details such as the name of the residence, the type and tenure of the property, the property's price, and its latitude and longitude.
- *Population Data (D_{pop}):* The Humanitarian Data Exchange (HDX) provided the population data, which is a compilation of demographic data for Malaysia. The dataset contains information on the population density overall, the number of men and women, women in reproductive age (15–49 years), elderly people (60+ years old), kids (0–5 years old), and young people (15–24 years old).
- *Relative Wealth Index Data (D_{rwi}):* The Relative Wealth Index Data contains a prediction of the standard of living in Malaysia based on de-identified connectivity data, satellite imagery, and other unconventional data sources. This dataset includes information such as the latitude, longitude, relative wealth index for each region, and an error measure.
- *Geospatial Data:* The Geospatial data is a collection of geographical information about Malaysia, primarily used for geocoding and reverse geocoding purposes. The dataset displays details about Malaysia, ranging from the state level to the district level, as well as the coordinates for each state and district.

- *Competitor Data (D_{cpt}):* The competitor data is similar to the POI data, containing a list of retail businesses in Malaysia. The dataset includes information such as the business's address, category, latitude, longitude, name, and state.

Table 4. Overview of Data

Dataset	Description	Dimension
Base Data (D_{base})	The retail business data, which include different categories of retail data.	1714, 4
Point of Interest Data (D_{poi})	The nearby location features.	443041, 6
Competitor Data (D_{cpt})	The nearby competitors	443041, 6
Residential Property Data (D_{ppt})	The nearby residence and property	491721, 6
Population Data (D_{pop})	The population of men, women, women of reproductive age, children, youth and elderly	3855823, 8
Relative Wealth Index Data (D_{rwi})	The relative wealth index of nearby areas	18148, 4
Geospatial Data	The geographical details of Malaysia	144, 17

B. Data Pre-processing

Data pre-processing is crucial in ensuring that the results are of good quality. This process transforms raw data into a format that is both usable and efficient. Datasets often contain meaningless or missing data, and the cleaning process is necessary to address these issues. In this project, some missing coordinate values in the D_{poi} and D_{ppt} were dropped because coordinates play an important role in determining location. Besides that, the reverse geocoding has been implemented in the D_{base} by using the Geospatial Data. As a result, the City and State of the coordinates will be generated. Lastly, there are some unnecessary features in there are some features in D_{ppt} and D_{rwi} has been filtered out.

C. Construction of Analytical Dataset

Algorithm 1 outlines the process for creating an analytical dataset. This algorithm takes D_{base}, D_{poi}, D_{ppt}, D_{pop}, and D_{rwi} as inputs and produces $D_{analytical}$, the analytical data, as the output. D_{base} serves as the foundational data for merging with other datasets. Initially at line 4, the `get-nearby-poi` function retrieves nearby Points of Interest (POI) from the base data using its latitude and longitude, generating a list of POIs within 2.0 km and selecting the top POI. Furthermore, at line 5, the `get-nearby-category` function produces the top 5 POI categories within 2.0 km. Subsequently at line 6, the `get-nearby-property` function generates nearby properties, and the top 3 nearby neighbourhoods are determined using the base data's latitude and longitude. This step reveals all relevant details about these neighbourhoods. At line 7, the `get-nearby-population` function calculates population density, providing the density for each neighbourhood. Additionally at line 8, competitors within 2.0 km are identified from the POI data using `get-nearby-competitor` function. At line 9, the `get-nearby-relative-wealth-index` function calculates the cumulative relative wealth index within 2.0 km of the nearby neighbourhoods. Finally at line 10, D_{base}, D_{nearby_poi}, $D_{nearby\_category}$, D_{nearby_ppt}, D_{nearby_pop}, D_{nearby_cpt}, and D_{nearby_rwi} are merged based on latitude and longitude, resulting in the analytical dataset as the output.

Algorithm 1 generate-Analytical-Dataset

Require: $D_{base\_1}, D_{base\_2}, D_{base\_3}, D_{poi}, D_{ppt}, D_{pop}, D_{rwi}$
Ensure: $D_{analytical}$
1: **for** *data* in ($D_{base\_1}, D_{base\_2}, D_{base\_3}$) **do**
2: $L \leftarrow$ `get-latlng`(*data*)
3: $D_{base} \leftarrow$ `get-base-data`
4: $D_{nearby\_poi} \leftarrow$ `get-nearby-poi`(D_{base}, L)
5: $D_{nearby\_category} \leftarrow$ `get-nearby-category` (D_{base}, L)
6: $D_{nearby\_ppt} \leftarrow$ `get-nearby-property`(D_{ppt}, L)
7: $D_{nearby\_pop} \leftarrow$ `get-nearby-population` (D_{pop}, D_{nearby_ppt})
8: $D_{nearby\_cpt} \leftarrow$ `get-nearby-competitor`(D_{poi}, L)
9: $D_{nearby\_rwi} \leftarrow$ `get-nearby-relative-wealth-index`(D_{rwi}, L)
10: $D_{analytical} \leftarrow D_{base} \bowtie D_{nearby\_poi} \bowtie D_{nearby\_category} \bowtie D_{nearby\_ppt} \bowtie D_{nearby\_pop} \bowtie D_{nearby\_cpt} \bowtie D_{nearby\_rwi}$
11: **end for**

Algorithm 2 outlines the steps for generating nearby POI data within a 2.0 km radius of the base data. At line 2, the haversine formula is employed to compute the distance between two pairs of coordinates. Once distances within 2.0 km are determined, the $D_{poi\_within\_2km}$ is derived from D_{poi}, and the counts of nearby POIs are obtained by tallying their frequencies. Subsequently at line 4, a new column is created using the nearby POI data within 2.0 km, which includes the frequency counts for each entry. The top 5 POIs with the highest frequency counts are chosen for this project.

Algorithm 2 get-nearby-poi

Require: D_{poi}, L
Ensure: D_{nearby_poi}
1: **for** *lat* and *lng* in *L* **do**
2: $D_{poi\_within\_2km} \leftarrow$ (`haversine`($lat, lng, lat_{poi}, lng_{poi}$) $*$ 1000) ≤ 2000
3: **end for**
4: $D_{nearby\_poi} \leftarrow D_{poi} \cap D_{poi\_within\_2km}$
5: **return** D_{nearby_poi}

The method for producing neighbouring POI category data within a 2.0 km radius of the base data is shown in Algorithm 3. At line 2, the distance between two sets of coordinates within 2.0 km is estimated using the haversine formula. Following the extraction of $D_{cateogry\_within\_2km}$ from $D_{category}$, the counts of neighbouring POI categories are calculated by adding up their frequencies. The subsequent step involves creating a new column that includes the 2.0 km away POI types and their related frequency counts at line 4. The top 5 POI categories with the greatest frequency counts are chosen for this project.

Algorithm 3 get-nearby-category

Require: D_{poi}, L

Ensure: $D_{nearby\_category}$

1: **for** *lat* and *lng* in *L* **do**

2: $D_{cateogry\_within\_2km} \leftarrow$ (`haversine`($lat, lng, lat_{poi}, lng_{poi}$) $*$ 1000) $\leq$ 2000

3: **end for**

4: $D_{nearby\_category} \leftarrow D_{category} \cap D_{category\_within\_2km}$

5: **return** $D_{nearby\_category}$

The stages in Algorithm 4 are described for creating neighbourhoods within a 2.0 km radius of the basis data. L and D_{ppt} serve as the algorithm's inputs, while D_{nearby_ppt} serves as the algorithm's output. Local neighbourhoods that are close to the base dataset are found by applying the haversine formula at line 2. The top three adjacent neighbourhoods have been selected for this project, and line 4 of the illustration gives detailed information for each of them.

Algorithm 4 get-nearby-property

Require: D_{ppt}, L

Ensure: D_{nearby_ppt}

1: **for** *lat* and *lng* in *L* **do**

2: $D_{ppt\_within\_2km} \leftarrow$ (`haversine`($lat, lng, lat_{ppt}, lng_{ppt}$) $*$ 1000) $\leq$ 2000

3: **end for**

4: $D_{nearby\_ppt} \leftarrow D_{poi\_within\_2km}$ [:3]

5: **return** D_{nearby_ppt}

The process for calculating each neighbourhood's population density is shown in Algorithm 5. This function aggregates the population density for areas within a 2.0 km radius, and detailed information on population density will be displayed.

Algorithm 5 get-nearby-population

Require: D_{pop}, D_{nearby_ppt}

Ensure: D_{nearby_pop}

1: **for** *lat* and *lng* in *L* **do**

2: $D_{pop\_neighbourhood} \leftarrow$ (`haversine`($lat, lng, lat_{ppt}, lng_{ppt}$) $*$ 1000) $\leq$ 2000

3: $D_{nearby\_pop} \leftarrow$ `sum`($D_{pop\_neighbourhood}$)

4: **end for**

5: **return** D_{nearby_pop}

Algorithm 6 presents the process for computing the relative wealth index of each neighbourhood. This function aggregates the relative wealth index for neighbourhoods situated within a 2.0 km radius and displays the calculated index for each individual neighbourhood.

Algorithm 6 get-nearby-relative-wealth-index

Require: D_{rwi}, D_{nearby_ppt}

Ensure: D_{nearby_rwi}

1: **for** lat and lng in D_{nearby_ppt} **do**

2: $D_{rwi\_neighbourhood} \leftarrow$ (haversine($lat, lng, lat_{rwi}, lng_{rwi}$) $*$ 1000) ≤ 2000

3: $D_{nearby\_rwi} \leftarrow$ sum($D_{rwi\_neighbourhood}$)

4: **end for**

5: **return** D_{nearby_rwi}

Algorithm 7 presents the method for identifying nearby competitors within a 2.0 km radius of the base data. This function takes D_{poi} and L as inputs and produces $D_{nearby\_competitor}$ as the output. The haversine formula is employed to generate nearby POI within 2.0 km at line 2. From $D_{poi\_within\_2km}$, users can select specific competitors of interest at line 4.

Algorithm 7 get-nearby-competitor

Require: D_{poi}, L

Ensure: $D_{nearby\_competitor}$

1: **for** lat and lng in L **do**

2: $D_{poi\_within\_2km} \leftarrow$ (haversine($lat, lng, lat_{poi}, lng_{poi}$) $*$ 1000) ≤ 2000

3: **end for**

4: $D_{nearby\_competitor} \leftarrow D_{poi\_within\_2km}$ [$selected\ competitor$]

5: **return** $D_{nearby\_competitor}$

D. Construction of Multi-Label Data

Algorithm 8 computes the nearest businesses within a 1.0 km radius for each base data point in the constructed analytical dataset. Initially at line 2, the function calculates the haversine distance between the businesses in $D_{business}$, the list of businesses in Malaysia and each point name in $D_{analytical}$. If there are one or more businesses within the 1 km radius, the nearest and the second nearest businesses will be selected and store in $D_{nearest\_business}$ and $D_{2^{nd}\_nearest\_business}$ at line 4 and 5 respectively. The features and description of analytical dataset is shown in Table 5.

Algorithm 8 get-nearby-business

Require: $D_{business}, D_{analytical}$

Ensure: $D_{nearest\_business}, D_{2^{nd}\_nearest\_business}$

1: **for** each row in $D_{business}$ **do**

2: $D_{business\_within\_1km} \leftarrow$ haversine($lat, lng, lat_{business}, lng_{business}$) < 1

3: **end for**

4: $D_{nearest\_business} \leftarrow D_{business\_within\_1km}$ [$selected\ nearest\ business$]

5: $D_{2^{nd}\_nearest\_business} \leftarrow D_{business\_within\_1km}$ [$selected\ 2nd\ nearest\ business$]

6: **return** $D_{nearest\_business}, D_{2^{nd}\_nearest\_business}$

Table 5. Features and Description of Analytical Dataset with Multi-label Data

Feature	**Description**
address	The address of the retail store
lat	Latitude of the retail store
lng	Longitude of the retail store
point_name	Name of the retail store
point_name2	Nearest business to point_name
point_name3	Second nearest business to point_name
city	City of the retail store
state	State of the retail store
Top 5 Point of Interest	Top 5 point of interest for each base data
Top 3 Nearby Neighbourhoods	Top 3 nearby neighbourhoods' details of the retail store, which includes name of neighbourhood, latitude, longitude, tenure of neighbourhood, type of neighbourhood and median price of neighbourhood.
Population of Each Neighbourhoods	The population density of each neighbourhood, which includes total density population, men population, women population, women of reproductive age population (ages 15-49), elderly population (ages 60+), children population (ages 0-5), and youth population (ages 15- 24)
Relative Wealth Index	The relative wealth index of each neighbourhood
Nearby Competitor	Nearby competitor for each base data

IV. METHOD

A. Model Construction and Evaluation

In this work, 1D Convolutional Neural Network (CNN) model is utilized for multi-label classification, an advanced technique that extends beyond traditional single-label classification. Unlike single-label classification, which restricts each instance to one class, multi-label classification allows multiple labels per instance, thereby more accurately reflecting real-world complexities. The concept of multi-label classification is notably more challenging due to intricate relationships among labels. An instance can belong to multiple classes simultaneously, with these classes not being mutually exclusive. In the context of this work, multi-label classification presents a unique advantage by predicting a multitude of suitable businesses for a given location, thus mirroring the reality of diverse business ecosystems co-existing within the same locale.

Each step in the process, from data pre-processing to performance evaluation, is adapted to this multi-label framework. In the data pre-processing phase, input features are extracted and standardized, and one-hot encoding is employed to transform multi-label target variables into a binary format. This crucial step redefines the problem into predicting a 'presence' or 'absence' for each business, thereby enabling the model to handle multiple labels. The architecture of the model, built using Keras Sequential API, comprises Conv1D and MaxPooling1D layers, followed by a Flatten layer, a Dense layer with 300 neurons, a Dropout layer to mitigate overfitting, and finally, a Dense layer with 1313 neurons which is corresponding to unique labels. The final layer uses a SoftMax activation function, outputting probabilities indicating the suitability of each business for a given location.

During training, the Adam optimizer and the categorical cross-entropy loss function are used, facilitating multi-label classification. Performance is monitored using accuracy, precision, and recall metrics. Additionally, Early Stopping and a custom Callback are used, which reset the model's weights and decrease the learning rate if performance does not improve, ensuring optimal learning. Upon completion, the model's performance is evaluated on the test set across accuracy, precision, recall, and f1-score metrics, refer to Equation (1-4). The emphasis on accuracy in this study helps in understanding the overall model performance across all labels. More importantly, it signifies the model's potential to recommend a diverse range of suitable businesses for a selected location, which is a fundamental aspect of this study.

$$Accuracy = \frac{TN + TP}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ Score = 2\left(\frac{Precision *\ Recall}{Precison + Recall}\right) \tag{4}$$

B. Experiment Setting

In the experimental setting of this study, an analytical dataset is utilized to comprise three primary variables: *point_name*, *point_name2*, and *point_name3* which will be used to develop the proposed 1D CNN model. The *point_name* variable represents specific retailers of interest, including well-known franchises such as McDonald's, Starbucks, and Pizza Hut. In contrast, *point_name2* and *point_name3* denote the businesses in closest proximity to the retailers in *point_name*. Specifically, *point_name2* refers to the nearest business, while *point_name3* signifies the second nearest business to the retailers in *point_name*. This experimental setup is designed to provide insights into the relationships and potential interactions between the focal retailers and their neighbouring businesses. The Table 6 shows the combination of three primary variables that used to form an analytical dataset.

Table 6. Combination of variables in Analytical Dataset

No.	point_name	point_name2	point_name3
1	McDonalds	Nearest business to McDonalds	Second nearest business to McDonalds
2	Starbucks	Nearest business to Starbucks	Second nearest business to Starbucks
3	Pizza hut	Nearest business to Pizza hut	Second nearest business to Pizza hut

V. FINDINGS

With an emphasis on assessing the effectiveness of the employed model, this part will dive into the findings and insights from the research. The goal is to highlight the model's benefits and drawbacks, giving people a thorough knowledge of how well it works in the context of the research.

A. Model Performance

In this study, a 1-dimensional Convolutional Neural Network (CNN) model is developed that is intended to effectively handle a multi-label classification issue for suggesting retail establishments for certain business locations. The model's effectiveness is demonstrated, with a focus on the accuracy measure, which is of utmost significance in the context of this purpose.

The model that was tested and trained has an accuracy rating of 61.22%. In the retail business recommendation work, accuracy is crucial since it quantifies the percentage of all predictions made by the model that were properly detected. High accuracy makes the model's recommendations more likely to be accurate, resulting in more effective and trustworthy retail business recommendations. An accuracy of 61.22% means that on the validation set, the model can accurately predict the labels for a given occurrence roughly 61.22% of the time. This high level of accuracy gives the model's suggestions confidence.

The model was assessed on additional measures, including as precision, recall, and the F1 score, in addition to accuracy. The precision score stood at 68.75%, suggesting that when the model predicted an instance to belong to a certain class, it was correct approximately 68.75% of the time. This result reinforces the reliability of the model's recommendations. However, despite promising accuracy and precision scores, the model demonstrated a lower recall of 3.74% and an F1 score of 7.1%. The relatively low recall and F1 score signify that there are opportunities for further improvements in the model, including reducing the number of false negatives and improving the balance between precision and recall.

Ultimately, the proposed 1D CNN model shows considerable promise in terms of accuracy, making it a useful tool for retail business recommendations. The model's accuracy-driven approach ensures a higher probability of successful business establishment at recommended locations. Nevertheless, the study also identifies areas for enhancement, particularly in improving recall and F1 scores. Consequently, future efforts could be directed towards refining these aspects, to deliver an even more robust tool for retail business location recommendation.

Table 7. Performance of 1D CNN Model

Classification Model	**Accuracy Score**	**Precision**	**Recall**	**F1 Score**
1D Convolutional Neural Network	61.22%	68.75%	3.74%	7.1%

B. Multi-Label Classification Outcome

In the context of multi-label classification, the developed 1-dimensional Convolutional Neural Network (CNN) model demonstrates its capability of making multiple recommendations concurrently for a given business location. The outcome provided by the model offers a set of retail business recommendations, enabling a comprehensive view of the possible business ventures that can be established in a particular location.

The results of the model's predictions are "Pizza Hut," which has a likely of 38.24%, "Starbucks," which has a likelihood of 6.86%, and "ATM HSBC Bank Pavilion Kuala Lumpur," which has a chance of 0.37%. This example shows how the model can predict numerous labels that include different types of companies while also giving each possible company a probability. There are several business opportunities presented by such a multi-label classification conclusion, offering more than just one sort of retail business. This feature is essential since it enables the model to consider the intricate, nuanced nature of a business site decision. It acknowledges the possibility for a place that holds many business types, offering a more complex and accurate portrayal of actual business decisions. This decision-making is aided by the accompanying probabilities, which provide an estimation of each business type's appropriateness and likelihood of success in the chosen location.

Moreover, the differing probabilities for each label indicate the model's nuanced understanding of the data. Higher probabilities suggest a more suitable match between the retail business and the location, while lower probabilities indicate fewer ideal matches. In the given instance, the model suggests that the location could be an excellent fit for a 'Pizza Hut', a reasonable fit for a 'Starbucks', and a less suitable location for an 'ATM HSBC Bank Pavilion Kuala Lumpur'. In conclusion, the multi-label prediction capability enhances the practicality of the model by offering diverse, probability-weighed recommendations, which can be used more effectively for real-world selection of retail businesses for a location.

Table 8. An example of multi-label output of 1D CNN Model

Retail Business	Pizza Hut	Starbucks	ATM HSBC Bank Pavilion Kuala Lumpur
Probability	38.24%	6.86%	0.37%

VI. CONCLUSION

The implementation of a deep learning approach, more specifically a 1D convolutional neural network (CNN) model, for multi-label classification in the context of retail recommendation, was examined in this research. The goal of the study was to create an analytical dataset and a model that could produce precise and pertinent retail suggestions based on the available information. The outcomes illustrated the 1D CNN model's potential for overcoming the difficulties of multi-label classification for retail recommendation. Accuracy, precision, recall, and F1 score were just a few of the criteria used to assess the model's performance, giving researchers a thorough grasp of how successful it is in this field.

This study has laid a strong framework for using deep learning techniques, including 1D CNN, for retail recommendation. Future study should concentrate on including further characteristics and variables that might further improve the model's performance, though, as the retail environment is extremely complicated and impacted by a variety of factors. In this context, the adoption of data augmentation methods, like to those employed in the recognition of plant disease leaf images [51], may provide fresh opportunities for enhancing model performance. To give a completer and more accurate picture of the retail market dynamics, it is important to take into account a variety of characteristics, including geographic, demographic, and socioeconomic ones. Researchers and practitioners can continue to improve and develop more sophisticated deep learning models for multi-label classification in retail recommendation by building on the findings and methodologies presented in this study. This will ultimately help the retail industry make decisions that are more individualized and successful.

ACKNOWLEDGEMENT

A version of this paper was presented at the third International Conference on Computer, Information Technology and Intelligent Computing, CITIC 2023, held in Malaysia on 26th-28th July 2023.

REFERENCES

[1] D. Chaffey, & F. Ellis-Chadwick, "Digital marketing." Pearson uk, 2019.

[2] V. Ahedo García, J.I. Santos Martín, J.M. Galán Ordax, et al. "Knowledge transfer in commercial feature extraction for the retail store location problem." IEEE Access, vol. 9, pp. 132967-132979, 2021.

[3] E.T. Bradlow, M. Gangwar, P. Kopalle, & S. Voleti, "The role of big data and predictive analytics in retailing." Journal of Retailing, 93(1), pp. 79–95, 2017.

[4] D. Grewal, A.L. Roggeveen, & J. Nordfält, "The future of retailing." Journal of Retailing, 93(1), pp. 1-6, 2017. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022435916300872 (The Future of Retailing) doi: https://doi.org/10.1016/j.jretai.2016.12.008.

[5] G. Lin, X. Chen, & Y. Liang, "The location of retail stores and street centrality in Guangzhou, China." Applied geography, vol. 100, pp. 12–20, 2018.

[6] L. Wang, H. Fan, & Y. Wang, "Site selection of retail shops based on spatial accessibility and hybrid bp neural network." ISPRS International Journal of Geo-Information, vol. 7, no. 6, p. 202, 2018.

[7] C.-Y. Ting, C.C. Ho, & H.-J. Yee, "Geospatial insights for retail recommendation using similarity measures." Big data, vol. 8, no. 6, pp. 519–527, 2020.

[8] C.-Y. Ting, C.C. Ho, & H.-J. Yee, & W.R. Matsah, "Geospatial analytics in retail site selection and sales prediction." Big data, vol. 6, no. 1, pp. 42–52, 2018.

[9] Y. Liu, B. Guo, B., N. Li, J. Zhang, J. Chen, D. Zhang, L. Yao, "Deepstore: An interaction-aware wide & deep model for store site recommendation with attentional spatial embeddings." IEEE Internet of Things Journal, vol. 6, no. 4, pp. 7319–7333, 2019.

[10] Y. Liu, B. Guo, D. Zhang, D. Zeghlache, J. Chen, K. Hu, Z. Yu, "Knowledge transfer with weighted adversarial network for cold-start store site recommendation." ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 15, no. 3, pp. 1–27, 2021.

[11] J. Zeng, & B. Tang, "Mining heterogeneous urban data for retail store placement." In Proceedings of the ACM Turing Celebration Conference-China, pp. 1–5, 2019.

[12] W. Si, & X. Yang, "Medicine retail terminal layout and site selection problems based on machine learning research: Take S enterprise as an example." In Proceedings of the 2021 6th International Conference on Cloud Computing and Internet of Things, pp. 22–28, 2021.

[13] H. Ebrahimi, K. Majidzadeh, & F. Soleimanian Gharehchopogh, "Integration of deep learning model and feature selection for multi-label classification." International Journal of Nonlinear Analysis and Applications, vol. 13, no. 1, pp. 2871–2883, 2022.

[14] M.M. Keikha, M. Rahgozar, & M. Asadpour, "Community aware random walk for network embedding." Knowledge-Based Systems, vol. 148, pp. 47–54, 2018.

[15] F. Markatopoulou, V. Mezaris, & I. Patras, "Implicit and explicit concept relations in deep neural networks for multi-label video/image annotation." IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 6, pp. 1631–1644, 2018.

[16] W. Weng, D.-H. Wang, C.-L. Chen, J. Wen, & S.-X. Wu, "Label specific features-based classifier chains for multi-label classification." IEEE Access, vol. 8, pp. 51265–51275, 2020.

[17] A.N. Tarekegn, M. Giacobini, & K. Michalak, "A review of methods for imbalanced multi-label classification." Pattern Recognition, vol. 118, p. 107965, 2021.

[18] P. Prajapati, & A. Thakkar, "Performance improvement of extreme multi-label classification using k-way tree construction with parallel clustering algorithm." Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 8, pp. 6354–6364, 2022.

[19] Y. Zhong, B. Du, B., & C. Xu, "Learning to reweight examples in multi-label classification." Neural Networks, vol. 142, pp. 428–436, 2021.

[20] K.L. Ailawadi, & P.W. Farris, "Managing multi-and omni-channel distribution: metrics and research directions." Journal of Retailing, vol. 93, no. 1, pp. 120–135, 2017.

[21] A. Aipe, N. Mukuntha, A. Ekbal, & S. Kurohashi, "Deep learning approach towards multi-label classification of crisis related tweets." In Proceedings of the 15th ISCRAM Conference, 2018.

[22] M. Bello, G. Nápoles, R. Sánchez, R. Bello, & K. Vanhoof, "Deep neural network to extract high-level features and labels in multi-label classification problems." Neurocomputing, vol. 413, pp. 259–270, 2020.

[23] J. Cai, W. Sun, J. Guan, & I. You, "Multi-ECGNet for ECG arrhythmia multi-label classification." IEEE Access, vol. 8, pp. 110848–110858, 2020.

[24] Z.-M. Chen, X.-S. Wei, P. Wang, & Y. Guo, "Multi-label image recognition with graph convolutional networks." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5177–5186, 2019.

[25] J. Du, Q. Chen, Y. Peng, Y. Xiang, C. Tao, & Z. Lu, "ML-Net: Multi-label classification of biomedical texts with deep neural networks." Journal of the American Medical Informatics Association, vol. 26, no. 11, pp. 1279–1285, 2019.

[26] T. Durand, N. Mehrasa, & G. Mori, "Learning a deep convnet for multi-label classification with partial labels." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 647–657, 2019.

[27] N. Ghorui, A. Ghosh, E.A. Algehyne, S.P. Mondal, & A.K. Saha, "AHP-TOPSIS inspired shopping mall site selection problem with fuzzy data." Mathematics, vol. 8, no. 8, p. 1380, 2020.

[28] H.-P. Hsieh, F. Lin, C.-T. Li, I.E.-H. Yen, & H.-Y. Chen, "Temporal popularity prediction of locations for geographical placement of retail stores." Knowledge and Information Systems, vol. 60, pp. 247–273, 2019.

[29] M. Jabreel, & A. Moreno, "A deep learning-based approach for multi-label emotion classification in tweets." Applied Sciences, vol. 9, no. 6, p. 1123, 2019.

[30] P.K. Jain, R. Pamula, & E.A. Yekun, "A multi-label ensemble predicting model for service recommendation from social media contents." The Journal of Supercomputing, pp. 1–18, 2022.

[31] S. Kohli, B. Timelin, V. Fabius, & S.M. Veranen, "How COVID-19 is changing consumer behavior – now and forever." McKinsey & Company, pp. 1–2, 2020.

[32] T. Lan, H. Cheng, Y. Wang, & B. Wen, "Site selection via learning graph convolutional neural networks: A case study of Singapore." Remote Sensing, vol. 14, no. 15, p. 3579, 2022.

[33] C.-W. Lee, W. Fang, C.-K. Yeh, & Y.-C.F. Wang, "Multi-label zero-shot learning with structured knowledge graphs." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1576–1585, 2018.

[34] Y. Omozaki, N. Masuyama, Y. Nojima, & H. Ishibuchi, "Evolutionary multi-objective multi-tasking for fuzzy genetics-based machine learning in multi-label classification." In 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–8, 2022.

[35] J. Ouyang, H. Fan, L. Wang, M. Yang, & Y. Ma, "Site selection improvement of retailers based on spatial competition strategy and a double-channel convolutional neural network." ISPRS International Journal of Geo-Information, vol. 9, no. 6, p. 357, 2020.

[36] X. Pan, K. Jin, J. Cao, Z. Liu, J. Wu, K. You, et al. "Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning." Graefe's Archive for Clinical and Experimental Ophthalmology, vol. 258, pp. 779–785, 2020.

[37] M.A. Parwez, M. Abulaish, et al. "Multi-label classification of microblogging texts using convolution neural network." IEEE Access, vol. 7, pp. 68678–68691, 2019.

[38] T. Sahin, S. Ocak, & M. Top, "Analytic hierarchy process for hospital site selection." Health Policy and Technology, vol. 8, no. 1, pp. 42–50, 2019.

[39] R.M. Sanchez-Saiz, V. Ahedo, J.I. Santos, S. Gomez, & J.M. Galan, "Identification of robust retailing location patterns with complex network approaches." Complex & Intelligent Systems, vol. 8, no. 1, pp. 83–106, 2022.

[40] S.A. Shaikh, M.A. Memon, M. Prokop, & K.-s. Kim, "An AHP/TOPSIS-based approach for an optimal site selection of a commercial opening utilizing geospatial data." In 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 295–302, 2020.

[41] D. Song, A. Vold, K. Madan, & F. Schilder, "Multi-label legal document classification: A deep learning-based approach with label-attention and domain-specific pre-training." Information Systems, vol. 106, p. 101718, 2022.

[42] D. Wang, & S. Zhang, "Unsupervised person re-identification via multi-label classification." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10981–10990, 2020.

[43] J. Wang, L. Yang, Z. Huo, W. He, & J. Luo, "Multi-label classification of fundus images with EfficientNet." IEEE Access, vol. 8, pp. 212499–212508, 2020.

[44] Y. Wang, D. He, F. Li, X. Long, Z. Zhou, J. Ma, & S. Wen, "Multi-label classification with label graph superimposing." In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 12265–12272, 2020.

[45] Z. Wang, T. Wang, B. Wan, & M. Han, "Partial classifier chains with feature selection by exploiting label correlation in multi-label classification." Entropy, vol. 22, no. 10, p. 1143, 2020.

[46] P. Yang, X. Sun, W. Li, S. Ma, W. Wu, & H. Wang, "SGM: Sequence generation model for multi-label classification." arXiv preprint arXiv:1806.04822, 2018.

[47] Y. Yao, P. Liu, Y. Hong, Z. Liang, R. Wang, Q. Guan, & J. Chen, "Fine-scale intra-and inter-city commercial store site recommendations using knowledge transfer." Transactions in GIS, vol. 23, no. 5, pp. 1029–1047, 2019.

[48] J.Y. Yap, C.C. Ho, & C.-Y. Ting, "Analytic hierarchy process (AHP) for business site selection." In AIP Conference Proceedings, vol. 2016, p. 020151, 2018.

[49] N. Zaheer, S.-U. Hassan, M. Ali, & M. Shabbir, "Optimal school site selection in urban areas using deep neural networks." Journal of Ambient Intelligence and Humanized Computing, pp. 1–15, 2022.

[50] Y. Zhong, & M. Zhao, "Research on deep learning in apple leaf disease recognition." Computers and Electronics in Agriculture, vol. 168, p. 105146, 2020.

[51] M. Xin, L.W. Ang, S. Palaniappan, "A Data Augmented Method for Plant Disease Leaf Image Recognition based on Enhanced GAN Model Network," Journal of Informatics and Web Engineering, vol. 2, no. 1, pp. 1-12, 2023.