

# An Intelligent And Effective Model To Recognize The Duplicate Products In Catalog Management System

P.Kalpana, M. Anitha, Dr.P.Tamije Selvy

**Abstract:** Nowadays, Electronic inventory management also provides consumers with more information on the items. The catalog management for e-commerce supports the e-commerce process of identifying, processing, retrieving and handling product information. In which digital database the item has two replicates so that consumer interactions are minimized and much more pertinent information on the items should not be included in the search recommendations. The following paper describes the approach for resolving the duplicates of the item and for through the search recommendations by RDMS and DBMS. It therefore generates an outstanding and systematic catalogue

**Index Terms:** Electronic Commerce, Catalog, E-Catalogs , Merchant, Eliminate Duplicate.

## 1 INTRODUCTION

Alist the executives framework gives the way to characterizing, putting away, and recovering e-catalogs, while supporting different sees and to straightforwardly bolster the different procedures of an e- trade application and it is very helpful for the vendor to transfer the items without the copy items. It builds the client encounters to see the items in a nifty gritty way. There are numerous basic issues in the plan and execution of an inventory the executives framework incorporate demonstrating and structure, institutionalization, interpretation, distributing, ordering, and form the board. Online business is the business trades with the help of web from wherever whenever. The energizing condition of electronic business is an expansive scope of communication forms among the different market participants; requesting the item, transport and conveyance, receipt and installment, and so on [1]. A run of the mill e-commerce exchange comprises of providers and customers to accumulate the data and investigating potential showcase accomplices for merchandise and enterprises, trading with the offers, development of products, installments, and other exchange exercises. Electronic lists hold a large portion of the data required in the flood of procedures and the modernized idea of e-catalogs[2][3] takes into account chances to robotize and streamline a significant number of the procedures in question. Therefore, e-catalogs structure the premise of an e-commerce exchange and index (or item data) the board is a capacity that is required in practically all e-commerce frameworks.

The inventory the executives in ecommerce site are the framework used to build the client experience and encourages for the client to find out about the items in detail. PET The vendor will in general view the items as far as the materials and procedures related with transferring the subtleties of the item with the assistance of the patterns. Also, the shipper would be progressively inspired by the item's estimate and weight. Contrast in perspectives can be unmistakably portrayed about the assortment of item characterization conspires that are being used these days. So here are numerous basic issues in the plan and usage of a list the board framework incorporate displaying and structure, institutionalization, interpretation, distributing, ordering, and form the board.

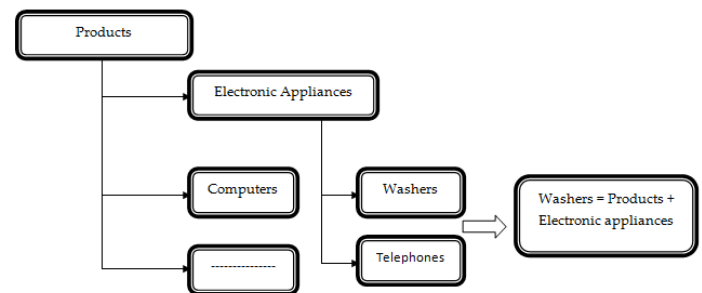


Fig.1. (a) Normal Product Group Definitio

## 2 DATABASE DESIGN

Typically, the index in e-commerce should be persevered in a database framework to be helpful. The test here is, it can't imagine all of the things it would have in the midst of its lifetime. Regardless of whether it occurs, It is an abuse of extra screening for items that may arrive after years to create and populate database tables.. This exhibits a test to plan the information model as these things and items would should be endured in a database, commonly utilising a social database the executives framework [4].

### 2.1 High available e-catalogs are necessary for e-commerce systems:

Any application of products to RDBMS or (DBMS) tables must have an executive framework from earlier data, all alone Fig 1

- P.Kalpana is currently working as Assistant Professor in Computer Science and Engineering in Sri Krishna College of Technology, India, PH -9600983981.

E-mail: kalpanapaulrajphd@gmail.com

- M.Anitha is currently pursuing PhD program in Computer Science and Engineering in Sri Krishna College of Technology, India, PH-9500821557. E-mail: anitha88.kool@gmsil.com
- P.Tamije Selvy is currently working as Professor in Computer Science and Engineering in Sri Krishna College of Technology, India, PH -9843598205. E-mail: p.tamijeselv@skct.edu.in

Therefore, the structure and information types in the social data base. The alteration of at least one of these database tables will normally require code changes to enter these tables, for example in order to define another element type. In these rows, it is impossible for new tables to be entered into the index server in the system runtime, or to alter the table structure. The program code should be revised and redeployed and must be removed.

## 2.2 Product attributes data need to be stored in atomic form

In order to ensure exact product searches and exams, these tasks should be done with regard to the assessment of nuclear information placed in segments [5]. Product attributes Data should be saved in atomic form. If all item data were removed in a large book field, extraction of this data would not be possible if necessary. For example, between 18c.f a customer needs a list of coolers. Therefore, 22c.f, with the highest quality for vitality. This means that customers must look at properties of the Refrigerator class "size" and "vitality utilization". If the data were inserted in a field with different properties in depiction contents, it would be a difficult task to remove them for this hunt.

## 2.3 Organize the types of products (categories), in a hierarchy of parent-infant relationships (use patterns of inheritance)

For example, a young person would be given ascripts normal to each of his children. This composition is to refer to the items or things in the list and to the fundamental qualities in the chain of importance. For example, for all items in a store, a standard property would be price, quantity, supplier, limits and product title. This would allow new classifications that would inherit properties from those already existing to be distinguished by marginal exercise.

## 2.4 Allow user in the system administrator to set a class when necessary

- Identify category names,
- Parent category (multi heritance categories),
- Attribute names and
- Types of the category.

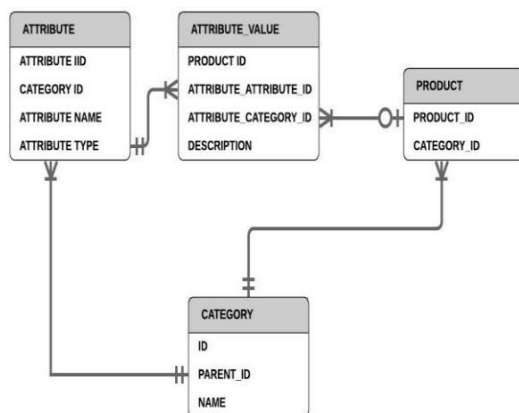


Fig .2. Catalog Data Model

## 2.4.1 CATEGORY

For this category, the ID [6]section retains exceptional Id and is the key. The parent classification I d contains PARENT ID. This is a remote key that corresponds to the ID category, because any parent class is a class itself. This section would include a NULL opportunity to identify the chain of value, since there is no parent. The class name is included in NAME. If this class should not contain any solid elements, ABSTRACT IND is a flag to be set. This table is connected to the R1 connection as it can have different classes for children.

## 2.4.2 ATTRIBUTE Table

Includes a category attributes. There is an attribute in each row. The ATTRIBUTE ID belongs to the class specified by Class ID and is a unique I d for the attribute. The attribute name is included in ATT NAME. The attribute type is given by ATT TYPE. The attributes defined in the ATTRIBUTE table include ATTRIBUTE VALUE table values. It is stored in the column VALUE. This table will store ATTRIBUTE ATTRIBUTE ID values for specific products that belong to ATTRIBUTE CATEGORY ID specified by PRODUCT ID.

## 2.4.3 PRODUCT Table

This table associates the real instance of the product to a certain category. It contains the information about which product (PRODUCT\_ID) is of which product type (CATEGORY\_ID). Relationship R/2 means that there are many attributes of a category and every attribute is one category. The R/3 attribute indicates that a lot of values can be stored, and that each value is linked to a particular specific product. Relationship R/4 indicates that zero, one or more items can be placed in the inventory of this sort of classification. One modification of the implementation would be that common to each product attributes would be specified as columns in the Table PRODUCT rather than identified and stored as attributes in tables ATTRIBUTE and ATTRIBUT VALUE.

## 3 PRODUCT CLASSIFICATION

The fundamental advantages of utilizing this inventory is to evacuate the copies of the item and kept up at the scale. Aside from this, a large number of dollars are spent creating programming that keep up data about items, purchasing history of clients for items, and so forth. Be that as it may, as the list measure and number of providers continue developing the issue of keeping up the list precisely develops exponentially [7]. The fundamental misclassification of items can prompt poor client experience. Unimportant list items and proposals. In this item arrangement, there are numerous difficulties incorporates off base labels sway item disclosure, absence of real-time deceivability into stock and failure to dissect and respond to request. CrowdANALYTIX approach is utilized to streamline (wiping out the copies) the index the executives in e-commerce framework and lessens the speed, expands the exhibition as fig 2. The procedure included the accompanying key advances:

### Primary keys:

- Pre-processing[8],
- Transformation what's more,
- Accuracy of the interior item list information as

referenced in (Fig 3).

- Reliable information: Once the models were handled, guarantee that the steady and precise information to keeping up the nature of the item.

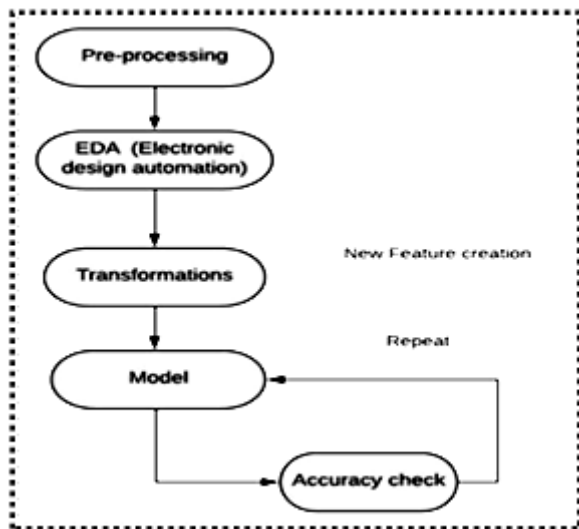


Fig. 3. Step by step flow of approaches

### 3.1 Approach 1

Information section publicly supporting stages: Crowdsourcing stages like Mechanical Turk gives access to enormous number of items over the world. It gives a Quality of information (Since people enter the qualities, information quality is sensibly great, anyway human information passage blunders do sneak in and should be looked out for). Its primary favorable position is, it gives a Speed of updates (Distributed Workforces can deal with all things considered 40,000 SKUs in a month) [9]. This isn't constantly perfect since items on an index should be on boarded each day and any deferral in streamlining the list can bring about lost deals because of poor client experience.

### 3.2 Approach 2

Item APIs: Platforms like Wisier, Index, Semantics3 and other web-based information stages that guarantee to have pre-built information storehouses that can be coordinated with existing lists. Reconciliation are simple and it is cost proficient while utilizing this methodology [10].

### 3.3 Approach 3

Some refined retailers assemble a lot of NLP[9][8] and ML calculations to extricate estimations of information qualities from picture, title and depiction of items. Since the picture, title and depiction gave by providers are commonly reliable, this technique gives exact trait esteems if the models are worked at an accuracy of 90% or more. Appropriateness for little retailers (There is an one-time exertion in building models yet post that the expense is truly reasonable)

### 3.4 Approach 4

dataXTM : utilizes ML calculations to naturally refresh item inventory information. A large number of models are worked by a horde of information researchers, one model comparing to every item quality. The models are then made available

through an API for close real-time refreshes. It limiting blunders and decreasing expenses

## 4 IMPLEMENTATION

Presently the outcomes from the above item grouping methodology is going to use in the RDMS tables to recover and show in the pages to the end clients. Create the table with the name ATTRIBUTE, ATTRIBUTE VALUE, PRODUCT, CATEGORY and having fields has discussed earlier in the database design module. Now the table along with the sample data looks like in the Table 1 and 2.

Table-1: Attribute Table

ATTRIBUTE ID	CATEGORY ID	ATTRIBUTE NAME	ATTRIBUTE VALUE
1	101	BRAND	1
2	102	MODEL	2
3	103	COLOR	3

Table 2: Attribute Value

PRODUCT ID	ATTRIBUTE_ ATTRIBUTE ID	ATTRIBUTE_ CATEGORY ID	VALUE
1	1	101	Samsung
	2		Galaxy J5
	3		Black
2	1	101	Samsung
	2		Galaxy J5
	3		White

## 5 CONCLUSION

The proposed index in e-commerce presents a strategy how to wipe out the copies in the items that is shown to the clients and expands the inquiry suggestions by utilizing the CrowdANALYTIX approach in the productive manner. With the assistance of this inventory the board bunches of clients can purchase their needs as per their desire and by decreasing the copies in the items dependent on the classification diminishing the presentation issues.

## 6 REFERENCES

- [1] Anant Jhingran, "Moving up the food chain: Supporting E-Commerce Applications on Databases", SIGMOD Record, vol. 29, no. 4.2.
- [2] Systems" in Managing E-Commerce and Mobile Computing technologies in IRM .908 publishing Eduardo B. Fernandez, Yi Liu, and RouYi Pan, Patterns for Internet shops, in PLOP 2001.
- [3] Dongkyu Kim, Sang-goo Lee, Jonghoon Chun, Sangwook Park, Jaeyoung Oh, Catalog Management in E-Commerce Systems, proc. of Computer Science and Technology (CST 2003).
- [4] Kiryong Kim, Dongkyu Kim, Jeuk Kim, Sangwook Park, Ighoon Lee, Sang-goo Lee, Jong-hoon Chun, "An Evaluation of Dynamic Electronic Catalog Models in Relational Database system
- [5] (RDBMS).
- [6] D. Kim, et al, Catalog integration for electronic commerce through category-hierarchy merging technique, Proc. 12th IEEE Workshop on Research Issues on Data Engineering, Engineering E-Commerce/E-Business Systems (RIDE-2EC).
- [7] Ms. P. Kalpana, R. Kirubakaran, Dr.P.Tamije selvy,

- “Hybrid SVD Model for Document Representation”, International Journal of Engineering and Advanced Technology,, ISSN:2249 – 8958, Volume-8 Issue-2, sep 2019.
- [8] Siamala Devi S, Shanmugam A., “An Integrated Harmony Search Method for Text Clustering using a Constraint based Approach”, Indian Journal of Science and Technology, Vol 8(29), 73986, 2015.
- [9] Ms. P. Kalpana,Dr.P.Tamije selvy “A Probe on Document clustering methodologies and it Performance Metrics” in International Journal of Recent Technology and engineering(IJRTE) ISSN: 2277-3878, Volume-8 Issue-2, July 2019.
- [10] Siamala Devi S, Dr. A. Shanmugam, “Hybridization of K-Means and Harmony Search Based On Optimized Kernel Matrix and Unsupervised Constraints”, Asian Journal of Information Technology, Vol.15, No.22, pp.4512-4521, 2016.
- [11] Jhingran, Moving up the food chain: Supporting e-commerce applications on databases, Technical Report, IBM Almaden Research Center.