

## Literature Review

### Product Recommendation System

A recommendation system is an essential component of a scientific library's operations. It enables users to move beyond catalog searches. Recommender systems are both beneficial to both user and service provider as it eases work for both the parties. Recent theoretical developments have revealed that recommendation systems are being used widely on every platform gaming, e-commerce and entertainment industry. Because of the standpoint of E-commerce as a framework that helps consumers search through records of information connected to their interests and preferences, these techniques have been important in the area. One of the major topics to be investigated in this field is collaborative filtering as it promotes product by detecting other users with common likes to the active user and depending on their suggestions. Collaborative recommender systems have been utilized in a wide range of applications. (F.O. Isinkaye, 2015) Knowledge-dependent or knowledge-poor recommendation techniques are both possible. Knowledge dependent refers to the use of ontological descriptions of users or objects, or restrictions, or social relationships and activities of users, whereas knowledge deficient refers to the use of simple and basic facts, such as user ratings/evaluations for products. (Mehrbakhsh Nilashi, April 30, 2013) In order to accomplish the core job of picking useful products, a recommender system must appraise items that are beneficial for presenting to the target user. The system must be able to predict how some of them will be used, or at the very least evaluate how specific products will be used and then decide which things to recommend based on that judgment. (Mehrbakhsh Nilashi, April 30, 2013)

Explicit evaluations are those in which a user is stimulated to provide an estimation on a specific item whereas implicit ratings are those that can be deduced from a user's activities. These could be scalar rating, binary ratings or unary ratings as well. Each row contains a user, each cell a distinct merchandise, and the number there at confluence of a row and a column reflects the participant's grade. (J. Ben Schafer, 2007)

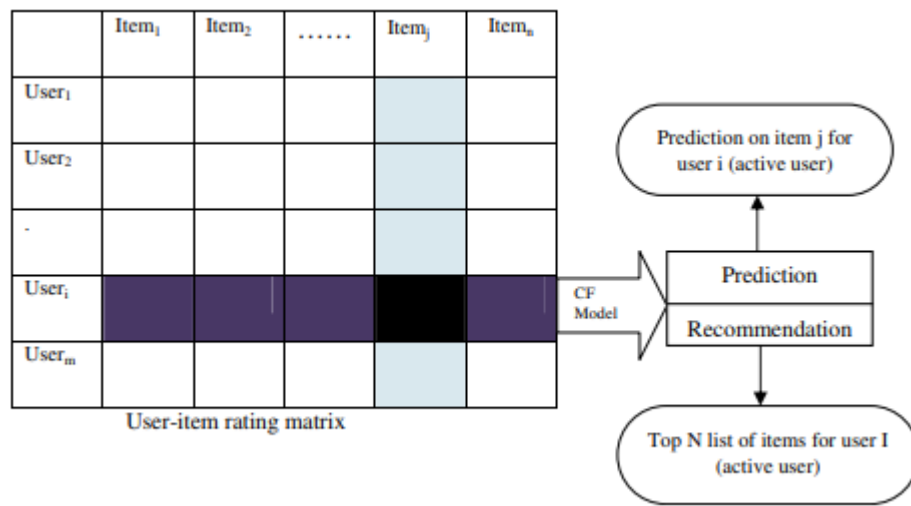


Figure 1 Collaborative Filtering

Data in a memory-based collaborative filtering system is kept in a "User x Item" matrix, where the rows represent the U users and the columns represent the items. The memory-based technique makes advantage of user input on objects. Model-based tactics have been implemented into recommendation systems to enhance and fix faults with memory-based approaches. (Abdelakder Grota, 2021) The weighted average of the active users' ratings on the connected items k is used to construct the prediction. To calculate item/user similarity, a number of similarity metrics are utilized. Correlation-based and cosine-based resemblance measurements are by far the most common. The Pearson's correlation coefficient is a measurement of how closely two factors are connected statistically. It is defined as:

$$s(a, u) = \frac{\sum_{i=1}^n (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i=1}^n (r_{a,i} - \bar{r}_a)^2} \sqrt{\sum_{i=1}^n (r_{u,i} - \bar{r}_u)^2}}$$

Figure 2

This equivalence denotes the correspondence between two user (a and u) and (r.a,i) is the assessment given to an item. It finds the mean rating of an item.

$$p(a, i) = \bar{r}_a + \frac{\sum_{i=1}^n (r_{u,i} - \bar{r}_u) \times s(a, u)}{\sum_{i=1}^n s(a, u)}$$

Figure 3

Cosine similarity, unlike the Pearson-based metric, is a vector-space model based on linear algebra rather than a statistical technique. It calculates how similar two n-dimensional vectors are by measuring the angle between them. In the domains of data reclamation and script withdrawal, a cosine-based metric is commonly used to compare two text booklets, which in this case are characterized as vectors of words. The similarity of two items, u and v, may be expressed as (F.O. Isinkaye, 2015)

$$s(\vec{u}, \vec{v}) = \frac{\vec{u} \cdot \vec{v}}{|\vec{u}| * |\vec{v}|} = \frac{\sum_i r_{u,i} r_{v,i}}{\sqrt{\sum_i r_{u,i}^2} \times \sqrt{\sum_i r_{v,i}^2}}$$

Figure 4

Model-based algorithms are also based on earlier user evaluations (profiles), but instead of making predictions right away, this approach splits people into groups or develops models from their data. (Abdelakder Grota, 2021) To increase the efficacy of Collaborative filtering Methods, this technique uses historical ratings to learn a model. Machine learning and statistical techniques may be used to create models. Supervised learning have also transformed the way suggestions are made, shifting the focus from advising clients on what to consume to advising clients on when to take a commodity. As a result, it is vital to examine the various learning approaches used in model-based recommender systems. (F.O. Isinkaye, 2015) The focus of this research will be matrix factorization, often known as matrix decomposition. It requires breaking a matrix down into several smaller matrices. Making the merchandise of these mediums will result in the original matrix. Matrix factorization has achieved positive outcomes in recommender systems.

Association rule: Association rules withdrawal methods abstract instructions that anticipate the arrival of a certain item based on the existence of other items in a transaction. A B pattern, for example, is applied to a collection of transactions by an association rule, where A A and B are two sets of objects.

Clustering: Classification technique, picture processing, data presentation, and knowledge extraction are all examples of applications that use clustering techniques. The goal of the clustering is to divide a collection of data into comment thread in order to find meaningful groupings within them. (F.O. Isinkaye, 2015)

Decision tree: The tree graph technique is used to construct a decision tree, which is constructed by reviewing a set of training samples for which the class labels are known. They are then utilized to classify previously unidentified instances. If trained on very high quality data, they can provide highly accurate predictions. (F.O. Isinkaye, 2015)

Artificial Neural network: An artificial neural network (ANN) is characterized by a large collection of linked neurons (nodes) organized in layers in a uniform way. Neuronal interconnections are given weights based on how much impact one synapse has on someone else. In some issue circumstances, neural nets have a number of advantages. (F.O. Isinkaye, 2015)

Link analysis: The technique of creating networks of connected objects in order to analyze patterns and trends, it has showed a lot of promise in terms of improving the efficiency of web search. In link analysis, the PageRank and HITS algorithms are utilized. Most link analysis tools approach a web page as a single node in the web graph. (Tirthankar Ghosh, 16.05.2018)

Matrix completion techniques: The basic purpose of the matrix completion approach is to forecast unknown values in user-item matrices. The rating matrix is frequently vast and sparse since users do not score the bulk of the objects. In reality, several low rank model modifications have been used for matrix completion, notably in collaborative filtering. (Tirthankar Ghosh, 16.05.2018)

Regression: Whenever multiple variables are assumed to be systematically associated by a linear connection, regression was used to analyze. It is a versatile and effective method for identifying the associative links between one and perhaps more completely reliant and study variables. Curve matching, forecasting, and testing fundamental claims about variable connections are all examples of regressive applications. A curve, whether linear, parabolic, or of a different kind, can be used to identify a trend in a dataset. (F.O. Isinkaye, 2015)

## Pros and Cons within collaborative filtering

The CF technique may make surprising predictions, which means it may propose items relevant to the user even if the content is not in the user's profile. Despite their efficacy, the widespread use of CF techniques has indicated a few potential concerns. Some of the problems we face with CF is a cold-start problem happens when a recommender does not know enough about a user or an item to make accurate recommendations. (J. Ben Schafer, 2007) Lack of data to process is also another problem that causes the CF to not function to its prime potential. Furthermore, a lack of data invariably leads to coverage concerns. Recommendation systems must be able to handle large numbers of users and objects in a database. Scalability is another challenge for recommendation systems, because computation increases linearly with the size of the database. As a result, it is vital to employ recommendation algorithms that can properly scale up as the number of datasets grows. Synonymy

refers to the tendency for extremely comparable commodities to have unique names or listings. Most recommender systems have difficulty discriminating between similarly related items, such as baby clothes and baby cloth. Collaborative filtering systems typically find no match between the two sentences when calculating their similarity. (F.O. Isinkaye, 2015)

## Uses of collaborative filtering

Recommendation algorithms have changed the way websites and users connect in recent years. The recommendation engine sifts through massive amounts of data to locate users' areas of interest and makes information retrieval easier. Collaborative filtering helps filtering for information or patterns using techniques that require collaboration across diverse players, perspectives, and data sources. Here we don't take past data or preference of group or an individual user. In collaborative filtering there may include problems that forces us of predicting unrated items, for such similarities between items and users are calculated using different approach. (J. Ben Schafer, 2007)

### User Functionality

- Assist me in finding new items that I would appreciate
- Advise me on a specific item
- Assist me in locating a person (or a group of users) that I might like
- Assist our group in finding something new that we would like
- Assist me in finding a combination of "new" and "old" products

### System Functionality

- Recommend items
- Predict for a given item
- Commend from a set of substances

## Recommendation System Using Deep Learning

A recommender system is a tool that helps consumers search knowledge data based on their interests and choices. It can employ collaborative filtering, content-based filtering, or a mixture of the two. The most developed and extensively used technique is collaborative filtering. Recommender systems assist consumers in dealing with information overload by proposing personalized, one-of-a-kind content and services. Because of collaborative and content-based methods to recommender systems, the accuracy and performance of search engine suggestions have increased. By integrating two or more filtering algorithms in various ways, hybrid filtering has been proposed to address some of the shortcomings of both techniques. Based on their activities, they are classified as weighted hybrid, mixed hybrid, switching hybrid, feature-combination hybrid, cascade hybrid, and meta-level fusion. Ziegler et al. introduced a hybrid collaborative filtering technique for using bulk taxonomic information for exact product categorization. (Tirthankar Ghosh, 16.05.2018)

To avoid defects and maximize performance, the qualities of both contained and explicit feedback can be combined in a hybrid system. This can be achieved by leveraging implicit data to validate explicit ratings or by allowing the user to submit explicit input only when he indicates explicit interest. For example we can use two algorithms to filter the datasets one for popularity based and another to

evaluate precision and precision recall. Then we can filter out unique users and unique products so that we can build a recommender based on personalized model for unique user. (Tirthankar Ghosh, 16.05.2018)

## K-Nearest Neighbor (KNN)

The K-NN approach, which is based on the Supervised Knowledge methodology, is one of the most essential Machine Learning algorithms. Although it may be applied to both regression and classification issues, it is more typically employed for classification problems. When fresh data is generated, the K-NN technique may be utilized to quickly put it into an appropriate category. The KNN technique considers that new and old instances are comparable and assigns them to the category that is most similar to the existing categories. It executes an action on the dataset during classification. (javatpoint, 2021) The K-nearest neighbor (KNN) strategy is widely used to train an accurate model from a small set of data. Traditional KNN labels new data based on previously labeled data points' labels. The calculation of point distance is based on non-weighted feature values, which is insufficient for calculating network flow. This paper introduces the concept of feature weight, and a weighted feature KNN (WKNN-Selfada) approach is developed to account for the varied contributions of multiple features. To get the optimal feature and feature weight set, a feature selection and feature adaption technique for WKNN is proposed. (Chencheng Ma, 2020) Undertake there are two classes, A and B, and that we get a new data point  $x_1$ . Which of the following categories will this data point come under? A K-NN method is required to address this sort of problem. We can simply discover the category or class of a dataset using K-NN:

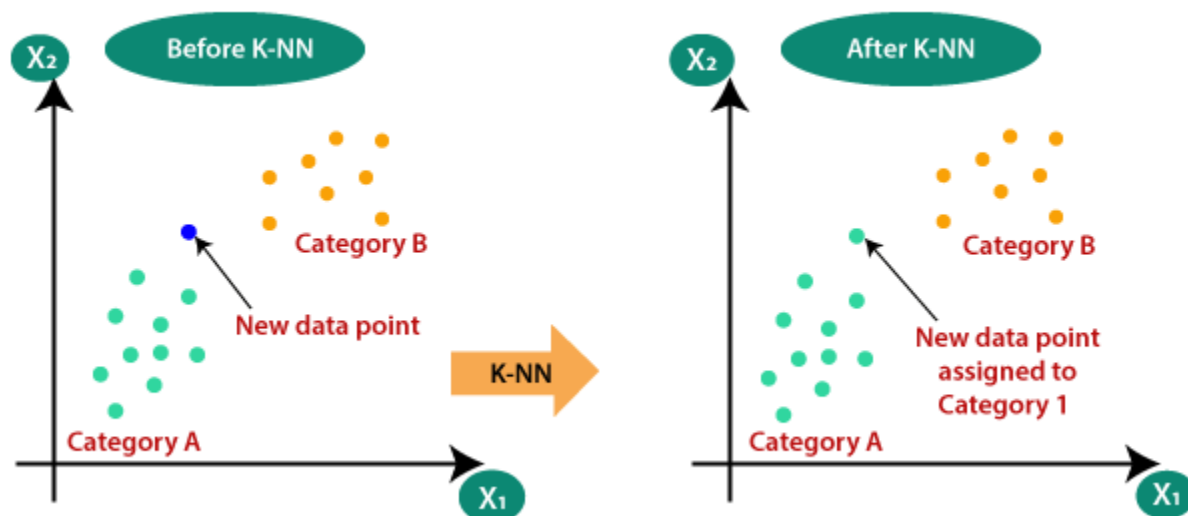


Figure 5 KNN working Steps

The KNN algorithm picks  $K$ , which is the number of neighbors. Calculate the Distance matrix between  $K$  of your closest neighbors. Count how many data points there are in each class among the closest neighbor. Allocate the sets of data to the subcategory with the most adjacent pieces of data. Our prototype is finished.

## Long Short Term Memory (LSTM)

LSTM networks are a sort of recurrent neural network that can learn order dependence and may be used to solve sequence prediction problems. This is necessary in a variety of difficult issue areas, including machine translation, speech recognition, and others. The LSTMs of deep learning are a difficult issue. It may be challenging to comprehend the notion of LSTMs and how terms like bidirectional and sequence-to-sequence apply to the field. The LSTM system must be able to hold data indefinitely and be noise resistant (i.e. fluctuations of the inputs that are random or irrelevant to predicting a correct output). A recurrent neural network system's parameters can be altered (in reasonable time).

(Brownlee, 2017) Every automatic learning approach requires data pre-processing. Pre-processing is done to optimize data for the learning model. The categorizing operation will be undertaken out on preprocessed data. Deep learning algorithms might be used to categorize data. Deep learning is a way of expressing data that employs a computer model comprised of several layers of processes. Deep learning categorization may be used in conjunction with a wide range of development approaches. One of them is the Long Short-Term Memory (LSTM) method. (Ahmad Hanif Asyhar, 2020) The fact that LSTMs were one of the first solutions to tackle technical challenges and deliver on the promise of machine learning algorithms is a crucial factor in their success.

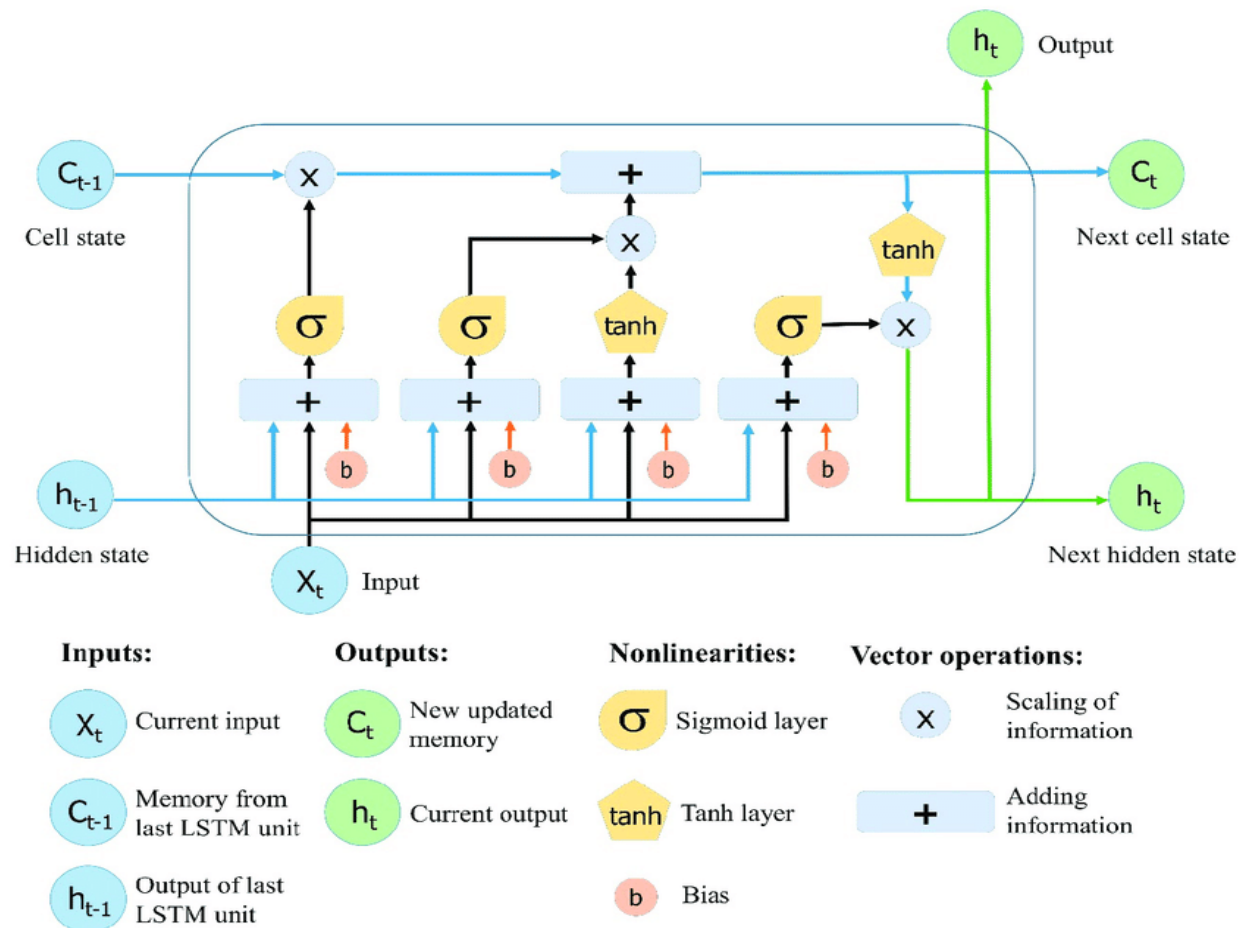


Figure 6 LSTM Working steps

How the LSTM works is that the forget gate decides whether bits of long-term memory should now be forgotten based on the previous hidden state and the current data point in the sequence (have less weight). To do this, a neural network is trained to produce values close to 0 when a component of the input is judged irrelevant, and values closer to 1 when the component is deemed relevant. This network generates a vector for each member in the  $[0, 1]$  range (ensured by sigmoid activation). This network is then transmitted up and pointwise multiplied with the cell state that came before it (inside the forget gate). Consider each component of this vector to be a filter/sieve through which more information can flow as the value approaches one. The new memory network and the input gate are used in the next phase. This is a neural network that learns how to integrate the prior hidden state with incoming input data to produce a 'new memory update vector.' Given the new input, this vector informs us how much to update each component of the network's long-term memory (cell state). The input gate is a sigmoid activated network that functions as a filter, determining which components of the final combined vector should be retained. Now that our modifications to the network's long-term memory are complete, we can proceed to the final stage, the output gate, which determines the new hidden state. This is accomplished by updating the cell state, prior concealed state, and new input data. Then, to acquire the new concealed state, we design a filter, the output gate. We put the freshly updated cell state via a tanh function before applying the filter to obtain the compressed cell state. This ensures that just the information required is output (saved to the new state). (Dolphin, Oct 21, 2020)

## Collaborative Filtering Algorithms: Theory and Practice

We can disregard product tags and demographic information in order to focus on the users and their ratings. We'll put the hybrid model through its paces to determine if combining a model-based (LSTM) and a memory-based (KNN) method yields better results than either strategy alone. The combined model is quite precise that suggests the majority of the suggested things are relevant. However, the model has a relatively low recall, which means that only a tiny fraction of relevant items are recommended. More generally, these basic findings are consistent with research showing that LSTM algorithm is more useful for building a recommendation model than using a KNN algorithm as KNN is more suitable for small scale data's and whereas LSTM is more reliable and consistent even with large data sets being an improved version of recurrent neural network. LSTM is more suitable for using as a recommender system using Keras LSTM for product purchases as time-series data where purchases are considered as time-series data and 2 new products are recommended based on last 3 purchases.

## Bibliography

Abdelakder Grotta, M. E. (2021). *Building Recommendation Systems Using the Algorithms* . Casablanca Morocco: Advanced Smart Systems (ASS) Hassan II University of Casablanca.

Ahmad Hanif Asyhar, A. Z. (2020). *Implementation LSTM Algorithm for Cervical*. Indonesia: UIN Sunan Ampel Surabaya.

Brownlee, J. (2017, May 24). A Gentle Introduction to Long Short-Term Memory Networks by the Experts.

Chencheng Ma, X. D. (2020, January 8). Improved KNN Algorithm for Fine-Grained.

Dolphin, R. (Oct 21, 2020). A Comprehensive Introduction to LSTMs. *LSTM Networks / A Detailed Explanation*.



- F.O. Isinkaye, Y. F. (2015, June 30). Recommendation systems: Principles, methods and.
- J. Ben Schafer, D. F. (2007). *Collaborative Filtering Recommender Systems*. Minneapolis, MN 55455: Department of Computer Science .
- J. Bobadilla . Ortega, A. H. (2013). Recommender systems survey. 109 -136.
- javatpoint. (2021). *K-Nearest Neighbor(KNN) Algorithm for Machine Learning*. Retrieved from javatpoint.
- Mehrbakhsh Nilashi, K. B. (April 30, 2013). Collaborative Filtering Recommender Systems.
- Tirthankar Ghosh, S. M. ( 16.05.2018). *Recommendation System Using Deep Learning*. Beliaghata, Kolkata: RCC Institute of Information Technology.