

DATA SCIENCE AND DEEP LEARNING APPLICATIONS IN THE E-COMMERCE INDUSTRY: A SURVEY

I Kali Pradeep

Research Scholar, KLEF (Deemed to be University),
Vaddeswaram, Andhra Pradesh, India

M Jaya Bhaskar

Associate Professor, KLEF (Deemed to be University),
Vaddeswaram, Andhra Pradesh, India

B Satyanarayana

Professor, Sri Krishnadevaraya University, Andhra Pradesh, India

Abstract - The E-Commerce industry is a new way of shopping. Most manufacturers prefer to sell their products on E-commerce websites because of low maintenance and high returns. This paper highlights the application of data science and deep learning methodologies to drive the E-commerce business. The first part of the paper explains the role of data science to identify potential customers, items recommendations and fraud detections. Different stages of building E-commerce websites and the role of proper handling of data is then highlighted. The second part of the paper is about the use of deep learning approaches like Recurrent neural networks in analysing textual information is presented with a demonstration of three architectures. At last, the importance of convolutional neural networks in the E-commerce industry is presented. Fashion related images like dresses, shoes and bags are classified using two separate convolution neural network architectures.

Keywords: E-Commerce; Recurrent Neural Network; Convolution Neural Networks; Recommendations; Deep Learning; Data Science

1. Introduction

The E-commerce industry is changing the shopping experiences of people. They get a lot of choices to choose items from different websites. But the main challenge lies here. There are so many users and items to deal with. Many e-commerce companies have shut down because of bad engagement of customers. With the latest enhancements and cheap availability of computational power, a lot of data is generated [1]. Proper utilization of this data may ultimately lead to the profitability of an E-commerce business.

The data generated from an E-commerce website is both structured and unstructured [2]. The structured data includes customer age, location of the customer, set of items bought, rating information of customers on items and Date and time of browsing and purchase. The patterns generated from structured format is not sufficient to understand a person's shopping needs and there is a need to mine useful patterns from unstructured formats. They include reviews of customers on items [3], user click patterns, tweets, social media comments and items viewed by customers.

There are many types of E-commerce industries. Amazon, Flipkart, Paytm mall, Myntra, which sell products directly and are called direct E-commerce websites [4]. These items are sold and shipped to the customers as soon as the order is placed. On the other hand, Music streaming websites like Amazon music, Spotify and Video streaming websites like Netflix, Amazon prime video are also a type of E-commerce websites which engage the customers with good recommendations and charge customers on a prepaid basis [5]. Whatever may be the type of E-commerce industry, the ultimate goal is to engage customers and earn good profits.

2. Data Science in E-Commerce

In order to use and extract meaningful information from this diversified form of data, data science techniques and methodologies are used. Data science is used in the E-commerce industry in multiple ways. Some of the main aspects are highlighted.

2.1. Market basket analysis

This is one of the oldest approaches and it is also known as association analysis in terms of data mining [6]. The main concept is to find the item/items related to items liked/bought by the customer. This is performed by generating frequent association rules. Many parameters like support, confidence, lift, and leverage are used for this. Where support is to find out of all transactions, the number of transactions the item is present. Confidence

is about the probability of buying/liking an item along with another item known that customer already bought/like an item. Lift defines the ratio of confidence of item w.r.t to other item and support of that item.

2.2. Price optimization

It is another area where data science and machine learning approaches are heavily used. At the beginning of the E-commerce era, the price of an item is fixed based on parameters like profit margin, competitor's pricing and suggested price decided by the seller. By using machine learning models like univariate linear regression, the relationship between sales and the cost of items can be derived [7]. To have an in-depth understanding, multiple linear regression or polynomial regression can be used to decide the relation between cost of an item based on customer location, frequency of purchase, seasonality.

2.3. Recommendation Engines

This is one of the areas where machine learning is mainly used. Recommendations related items to the items liked by the user are made. There are many variations of recommendation systems in the literature. 3 main types of recommendation engines are popularity based, content-based and collaborative based [8].

In popularity-based recommenders, the interest of the user is not taken into consideration. The most popular items are sorted out w.r.t to the location and time and top N items are recommended to the user.

Many times, this simple recommendation system works well because a popular item is mostly liked by everyone. Popularity is generally calculated by using parameters like the number of items bought/consumed within a time interval.

The Second type of recommender is content-based filtering. If a user liked/bought an item, then similar items to the item interested are recommended to the user. The similarity between items is calculated using the metadata information. For example, if a user buys/likes books of an author, then other books by the same author are recommended to the user. There are many ways to know if a user liked an item. 1) Explicit rating given by a user to an item. 2) Analysing the review using natural language processing techniques. 3) Recommending items that the user viewed. 4) Analysing click logs, like the category of information user preferred. 5) Parameters like the price may also be taken into consideration. If a user tends to buy discounted items, then items having heavy discounts may be recommended to the user.

Collaborative filtering is another type of recommender system. Unlike content-based recommenders, the similarity between users is taken into consideration. In this approach, a similar set of users are found. Then, items purchased or liked by a user is recommended to similar users. Similarity measures of content and collaborative filtering as almost the same with small modifications.

2.4. Image analysis

This is a new concept and can be used to understand the user [9]. In the case of fashion E-commerce websites, following information about the user interest can be known. 1) Dress colour 2) Dress design 3) background of the image 4) lighting effect. 5) Model's presentation of the dress.

Image analysis can also be used to find similar items. In the traditional approach, similar items are searched based on item metadata and user information. To extend the concept further, similar images can be searched using deep learning. The convolutional neural network is a deep learning approach specially designed for handling image information.

2.5. Fraud protection

With a huge amount of data, it is impossible to track fraud purchases manually [10]. For example, A customer may order the product just to make duplicates out of it and return the product. In the same way, some products sold in E-commerce sites may not be for reselling. But people may buy discounted items with multiple nearby addresses and resell it to make money at the original cost. E-commerce websites may also identify customers who opt for cash on delivery and repeatedly return the product thereby making a loss to the companies. These types of transactions can be identified by using data analytics to improve the quality of business.

2.6. Searching the items

Simple keyword search is a traditional method of searching and is not suited for E-commerce websites [11]. Additional parameters like user context, location, season are helpful to uncover good patterns.

3. Role of Data in Each Stage of Building Recommenders

This paper is mainly based on methods for making recommendations and engaging customers. There are mainly 5 stages in building and building a recommender system.

3.1. Knowing business requirements

Adding custom recommender systems to an E-commerce website is a costly affair. When there are very few items, then it is better to go with static web pages with manual recommendations by the expert [12].

3.2. Pre-processing the data

Like any other machine learning problem, data pre-processing is a challenging task. The data needed for the recommender system is both structured and unstructured. And, the timeline should also be taken into consideration while making recommendations [13]. In some categories, old items should not be recommended and in some cases, old items may be considered. The data is not only about what the user bought. But many other parameters like location, time of the day, friends, trend and reviews should also be taken into consideration. Integrating these data into a form understandable to the machine learning algorithm should be done almost in real-time.

3.3. Ranking prediction of items for users

The recommendations to a user can be made using basic statistics or by using complex deep learning approaches. But the output is generally the list of items to be recommended to a particular user. Sometimes making recommendations is not direct. Different algorithms are parallelly applied to the data and then the recommendations are aggregated to present to the user [14]. This concept can further be extended by using a weight-based hybrid recommender system. More importance is given to one algorithm and less importance is given to another algorithm. Final recommendations are made based on their weights.

3.4. Visualization of recommendations

As discussed earlier, the final output of a recommender system is to suggest products that may be of interest to the user [15]. These items can be shown to users using various web-based advertisements like google ads, by displaying in e-commerce websites or by sending promo messages.

3.5. Iteration and deployment of models

The recommendation engine should be updated regularly. The performance of a recommender system is directly related to the amount of data given for training. The performance of the system should be regularly monitored [16]. Because bad recommendations may lead to the chance of losing the customer.

4. Deep Learning in Recommender System

Traditional machine learning algorithms like linear regression [17], logistic regression [18], support vector machines [19] and decision tree models [20] try to model the data in a linear fashion. But many times, data in real-world is non-linear. And deep learning-based approaches are good at handling them. In recommender systems, there are many types of data like audio files, textual information, image data and video files. Deep learning-based methods can handle these kinds of data easily as compared to traditional machine learning-based systems [21]. There are many variants of neural networks to handle different kinds of data. Recurrent neural network systems are good at handling natural language processing and speech recognition. Implicit feedback like reviews are text and RNN can help up in this case. In the same way, to find similar images, viewed by the user, convolutional neural network architectures can be used. There are many open source frameworks like TensorFlow, Keras, PyTorch, Caffe, Theano for this purpose [22]. They make the implementation a lot easier. Let's discuss different architectures in detail:

4.1. Multilayer Neural Network

A multilayer neural network consists of many artificial neurons also known as perceptron. A network of perceptron tries to learn as a human brain does.

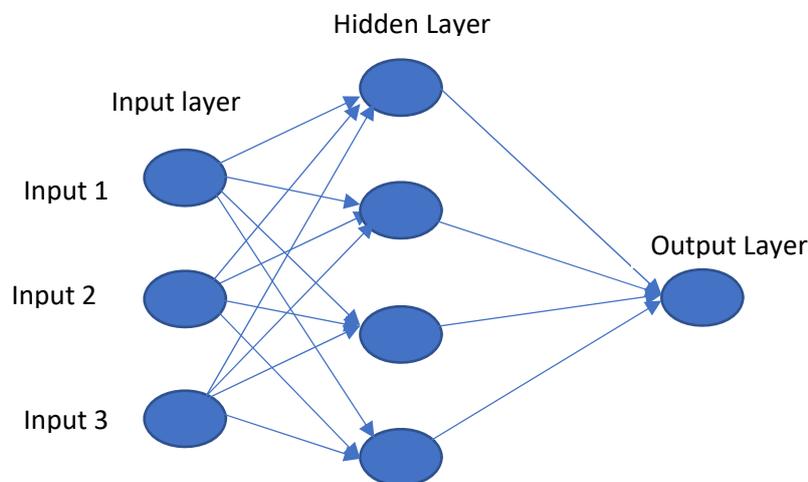


Figure 1: A multilayer perceptron

The goal of a multiple layer network is to map a complex function between input and output [23]. This type of mapping is possible by using multiple layers of computations also known as hidden layers. A multilayer perceptron consists of 3 types of layers: 1) Input layer 2) Hidden layer and 3) output layers.

There can be any type of output at the output layer. A sigmoid function is generally used for binary classification. Whereas, SoftMax function can be used for multiclass classification problems. There may be no activation function in the case of a regression problem.

4.2. Recurrent Neural Network

Simple feed-forward neural networks learn through training. But Recurrent neural networks also learn from prior computations for output prediction [23].

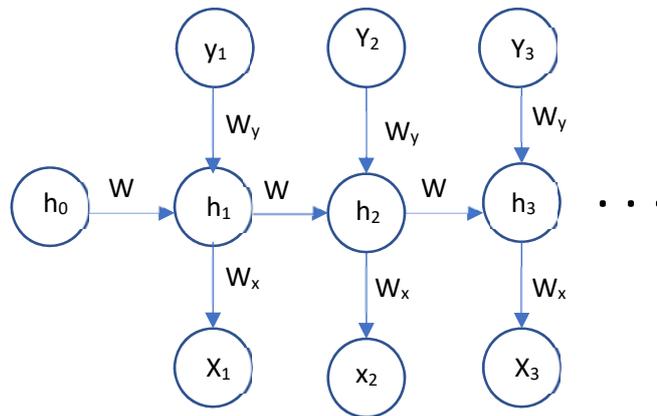


Figure 2: Recurrent neural network

As shown in the figure, the output y is dependent on the current input and past computations. RNN input parameters are shared. The hidden layer takes into consideration the relationship between neighbours. There are many variations of RNN 1) One-to-one 2) One-to-many 3) Many to one 4) Many-to-many. Vanishing of the gradient is one of the main drawbacks of RNN. This is because the increase in the number of layers is dependent on the increase or decrease in multiplicative gradient.

A bidirectional recurrent neural network is a variation of RNN. In B-RNN future is also taken into consideration from present computations [25].

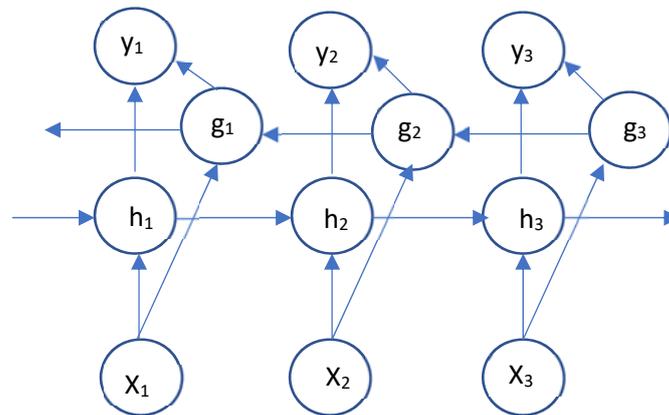


Figure 3: Bidirectional recurrent neural network

The main challenge in a bidirectional recurrent neural network is to find the range of future nodes to look into.

LSTM [26] is the next level of improvement over RNN. Selective remembering is the key aspect of LSTM. RNN is capable of handling only short contexts. On the other hand, LSTM is good at remembering the long context behind the sequence. The information flow modules are called gates. They use the sigmoid function, tanh function and pointwise computation as shown in the figure below.

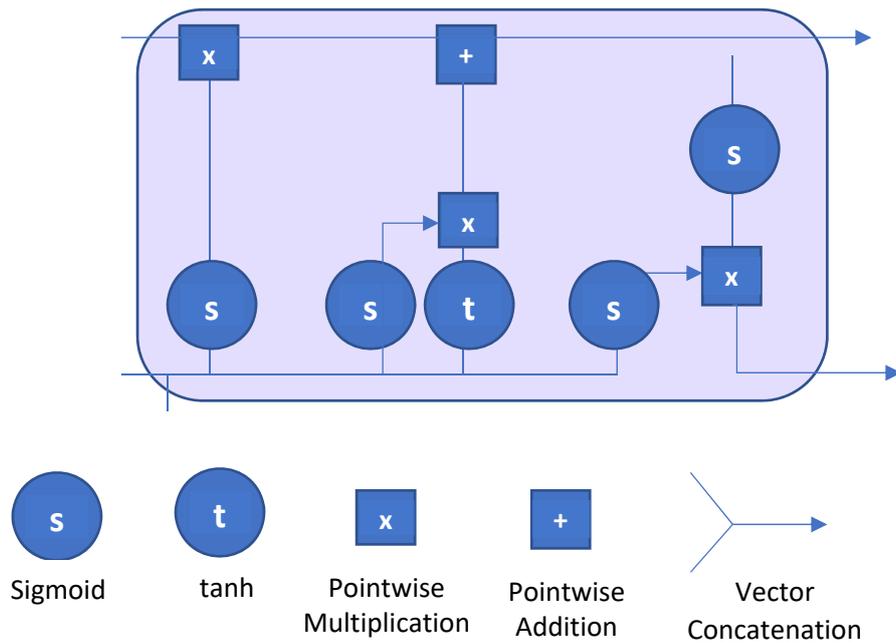


Figure 4: Long short term memory

These gates help throughout irrelevant information. A cell in LSTM can be thought of as a memory of networks. The information in the cells can be added or deleted. The capability of learning and forgetting is possible with the help of gates.

A sigmoid function takes only two values 0 and 1. The information to be remembered is multiplied by 1. And multiplied with zero it needs to be forgotten.

There are mainly 3 types of gates. 1) Forget gate 2) Input gate 3) Output gate

The forget gate: Initially, the current information and past information are the input for sigmoid function. The output is 0 or 1.

Input gate: The output from the previous hidden layer and current input are passed through sigmoid and tanh function. The output from both is then multiplied. Tanh function is for the regulation of the neural network.

Cell state: The cell state is computed in 2 stages. At first, the previous cell state and output from the forget gate gets multiplied. This output is then added with the output of the input gate.

Output gate: The new cell state is passed through a sigmoid function. This is then multiplied with the output of previous hidden state and current input passed through a sigmoid function.

Gated Recurrent unit is the new version of the recurrent neural network and its functionality is a bit related to long short-term memory [27]. In GRU, the hidden state is responsible for information transfer. GRU has 2 gates 1) Reset gates 2) Update gate.

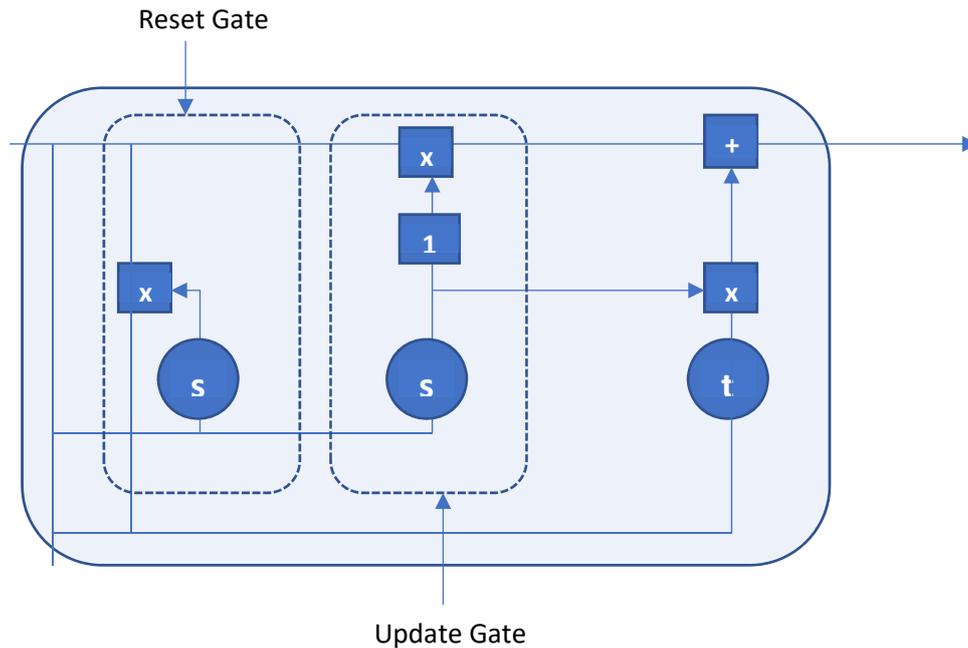


Figure 5: Gated recurrent units

Update Gate: The function of forget gate and input gate in LSTM is performed by Update gate. It makes the decision to keep or forget the information.

Reset gate: This gate is responsible to decide the amount of information to throw out.

4.3. Convolutional Neural Network

Deep learning concepts for computer vision gave an idea of convolutional neural networks. Its main area of application includes image and video analysis [28].

Flattened image pixels, when used with a feed-forward network, give a bad performance with less accuracy. The number of parameters used in the convolutional neural network is less as compared to a normal feed-forward neural network.

A grayscale image is easy to handle. But a colour image has 3 to 4 layers. Each layer corresponds to each colour. The common colour formats are red, green, blue (RGB) and Cyan, Magenta, Yellow, Black (CMYK).

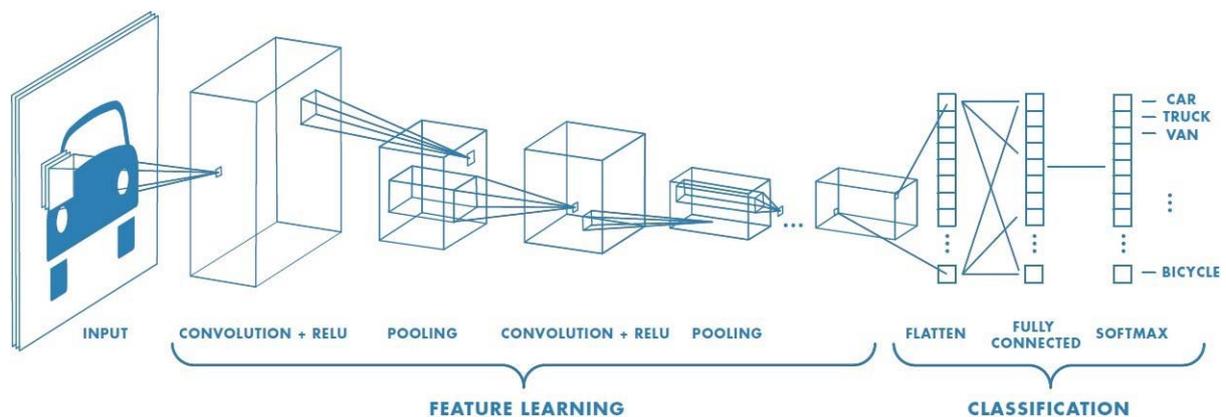


Figure 6: A basic skeleton of a convolutional neural network

CNN converts the image to a form good for analysis. There are many types of convolutional networks available in the market. But almost every architecture has few things in common. 1) Convolutional layer 2) Pooling layer 3) ReLu 4) Fully connected layers 5) SoftMax

4.3.1) Convolution Layer

It is a matrix multiplication operation between a filter and a part of the image to produce a feature map. A filter is a relatively small matrix w.r.t to the size of the image. The slide shifts the kernel filter throughout the image to perform convolution operation on the whole image.

The depth of the kernel filter is the same as that of the image. In the case of RGB, the depth is three and in the case of CMYK, the depth is four. The convolution operation is performed to extract some features in the image. In reality, multiple convolution layers are used to extract a different kind of information like shape, colour, size, patterns from the image.

The size of the image is reduced when the convolution operation is directly performed. Padding is used to increase/retain the size of an image. Padding is a process of adding dummy pixel information to the image. If convolution operation is performed after padding, then the size of the feature map is either the size of the image or its size is increased.

4.3.2) Pooling Layer

The number of pixels for an image has led to large size images. Hence leading to large feature maps. Processing this information needs huge computational power. Pooling layer is responsible for decreasing the size of the convolution features. This should be done with minimal loss of information. Like a convolution layer, we fix the size of the kernel. There are mainly two types of pooling layers 1) Max pooling 2) Average pooling.

In max-pooling, the maximum pixel value from part of an image covered by the kernel is returned. On the other hand, average pooling returns the average pixel value. Compared to both, max-pooling performs better at reducing noise from an image.

4.3.3) Activation function

The mathematical computations like convolutions may introduce some non-linearity. Rectified linear unit (ReLU) is a simple activation function with less computational complexity [29]. The ReLU replaces the negative values with zero and keeps positive values as it is. Sigmoid and tanh are other activation functions.

The sigmoid activation function is mainly used for binary classification problems. It outputs only 2 values 0 or 1. Tanh is another activation function which takes the value between -1 and 1. They are mainly employed in feedforward networks.

4.4.4) Dropout

It is a technique used in neural networks to reduce overfitting. The concept of reducing the chance of overfitting is called regularization. In dropout, some neurons are not activated during forward and backward propagation. Drop connect is another variation to reduce overfitting. Unlike dropout, here weights of links between the neurons are set to zero. Generally, activation functions are applied just after convolution operation.

After applying multiple convolutions and pooling operations, the values are finally given as input to a feed-forward layer and multiple epochs are performed. The output is then passed through SoftMax activation for classification purposes.

5. Convolution Neural Network in The E-Commerce Industry

With the increase in computation power, it is possible to extract features from images using CNN. In an E-commerce website (Except audio recommender like Spotify), Every item is represented as an image.

By using CNN, similar images to an image liked or viewed by the user may be revealed.

There are many convolutional neural network architectures in the market. Some of them are 1) AlexNet 2) VGG 3) Resnet 4) Inception

AlexNet was the first to use Rectified Linear units (ReLU). The architecture has five convolution layers, three max-pooling layers and three fully connected layers. It was invented in 2012[30].

Visual Geometry Group (VGG-16) is one of the most used networks to date [31]. This architecture has 13 convolution layers, 2 max-pooling layers, and 3 fully connected layers. The size of max-pooling layers is small as compared to Alexnet.

Inception-V1 has 22 layers in its architecture [32]. Dimensionality reduction is performed using 1×1 convolution. Inception uses a network inside the network architecture. A 1×1 convolution layer is also helpful for removing bottlenecks and also helps in adding non-linearity. Convolution filters of multiple sizes are used to avoid overfitting. There are Inception versions v2, v3 and v4 which are better at handling the images. The inception architectures are computationally expensive. Another name for inception is google net.

ResNet is the answer for inception architecture. Most of the architectures discussed above increased the number of layers to get good accuracy [33]. But deep networks have a problem of vanishing gradients. By simply increasing the layers leads to less performance of the model. But in ResNet, the complexity is increased by adding skip connections. Resnet uses batch normalization which is novel in this architecture. ResNet uses 152 layers but still generalizes better. This architecture is from Microsoft.

6. Experimental Study

Let's consider a data set which has fashion-related items like dresses, footwear, bags. The name of the data set is fashion_mnist. There are sixty thousand 28x28 grayscale images for training and ten thousand 28x28 testing. In total there are 10 class labels. This dataset is provided by Zalando and is taken from the Kaggle website. Since the dataset has grayscale images, the pixel values of images range from zero to two hundred twenty-five. The pixel value of 0 indicates the black colour and 1 indicates white colour. A darker pixel value is nearer to 0 and the lighter pixel value is more near to 255. Each image has 784 pixels (28x28). Both the training and test data have 785 columns. Where the first column is the class label and the rest of the columns have values for each pixel. So, each row represents an image.

In this paper, we are going to apply 2 of the above architectures on this data set.

First, let's apply VGG architecture. We implement this in python. Packages like NumPy, pandas, sklearn, and Keras are initially imported. Both training and test data are then taken into the environment. Initially, the data is divided into train-features, test-features, train-label and test label. Where train and test labels are the class label columns in our dataset. Now, reshape the images into 3 channels and resize the 28x28x3 size image into 48x48x3. The purpose of resizing is to make input images in a suitable format for the architecture. The pixel values are normalized and converted into float32. Data is changed into one-hot encoding format and given as input for the network. The output is the extracted features after 43 epochs. These features are flattened and given to a dense layer and the model is built. The SoftMax activation function is used in discrete class label identification.

In VGG models, 3x3 models were used instead of 11x11 models as in AlexNet. A 2 3x3 filters are used to create 5x5 filter. Using multiple small filters will help in reducing the number of parameters. Overfitting can be avoided and faster convergence can be achieved by using less parameters.

VGG-11, 13, 16 and 19 versions are compared in this paper. In VGG -11, one 3x3 64 filters and one 3x3 128 filters, four 3x3 512 filters are used. In VGG 13, an extra 3x3 64 filters, and one 3x3 128 filters are used. On the other hand the new addition in VGG 16 were the addition of 1x1 convolution of 256 filters and two 1x1 convolution with 512 layers. The VGG 19 version has extra layers of 3x3 256 filters and 2 3x3 512 filters. This addition has not much impacted the performance. In Every version of VGG, the number of layers were increased hence increasing the dept. While applying the model, SGD (Stochastic Gradient Descent) optimizer was used for better optimizing.

The second model is AlexNet, this model is applied to the fashion_mnist dataset using PyTorch. Now, At the time of the first convolution, the number of input channels were 1 and output channels were 96. Kernel size of 11 is used and stride is kept as 4 with no padding. ReLu activation function is used with 3x3 kernel size with a stride value of 2. At the time of the second convolution, input channels were 96 and output channels were 256 with a kernel size of 5 and with stride value of 2 is used. Here the padding of size 2 is used to avoid compression of the image. ReLu and max-pooling is applied as above. At the 3rd convolution stage, the input channel of 256 and the output channel of 384 is used. Kernel size of 3 is applied with stride and padding value of 1. In the 4th convolution stage, Both the input channel and output channel of 384 is used with the same channels, stride and padding as in the previous layer. In the last stage, the input channel is 384 and the output channel is 256 with other parameters equal to the preceding layer.

The final architecture applied is Resnet-18. The architecture is used in Pytorch framework-Torch vision library. Initially it was designed for mnist dataset, where the task was to identify one from the 10 handwritten numbers. In Resnet architecture, the input and output channels are initially the same with two 3x3 convolutions for residual. ReLu and batch normalization is applied after each residual. The google net (Inception) and Resnet have the same type of layers in the first 2 steps. Convolution layers of size 7x7 with 64 output channels are used with slide value as 2. There are 4 models with residual blocks. The channel size is doubled in each residual block for the consequent module. 10 Epochs are applied and the test accuracy is 90.4%.

Table I) Table comparing accuracy of models

Name of architecture	Number of epochs	Maximum Test Accuracy	Number of parameters
AlexNet	16	86.45	57,029,322
VGG 11	13	93.22	4,508,298
VGG 13	12	93.83	4,693,194
VGG 16	10	93.68	7,644,618
VGG 19	17	93.55	10,596,042
Resnet 18	16	93.00	11,172,810
Custom model		93.49	372,032

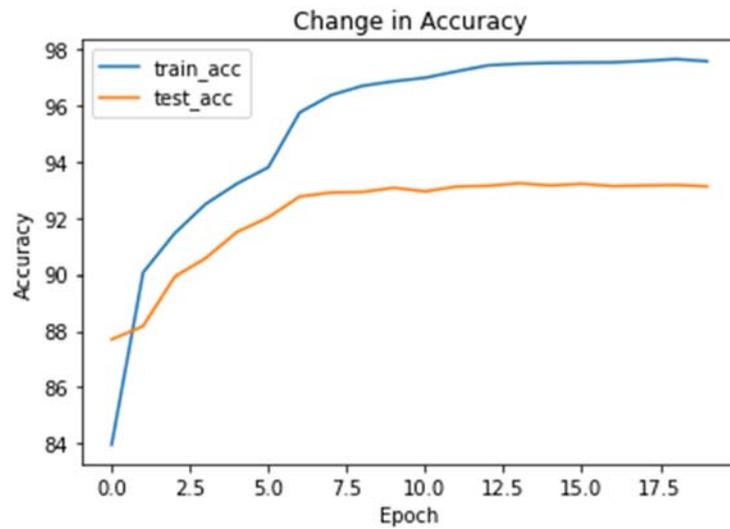


Figure 7: Train Vs Test accuracy of VGG 11

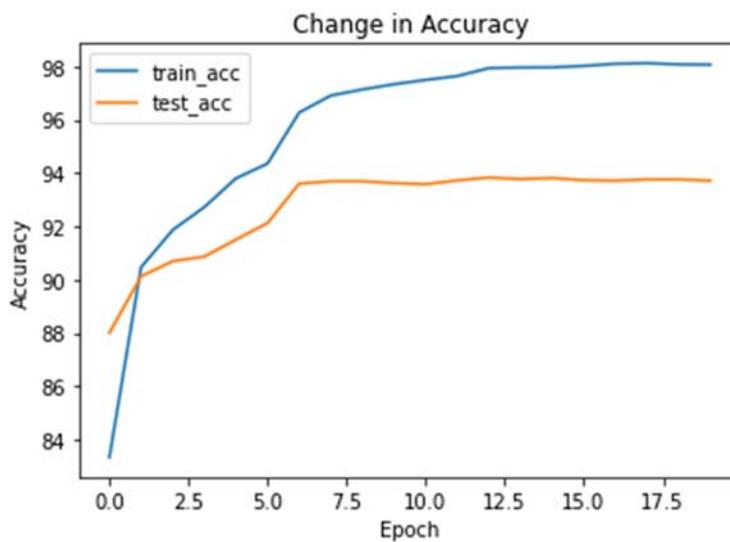


Figure 8: Train Vs Test accuracy of VGG 13

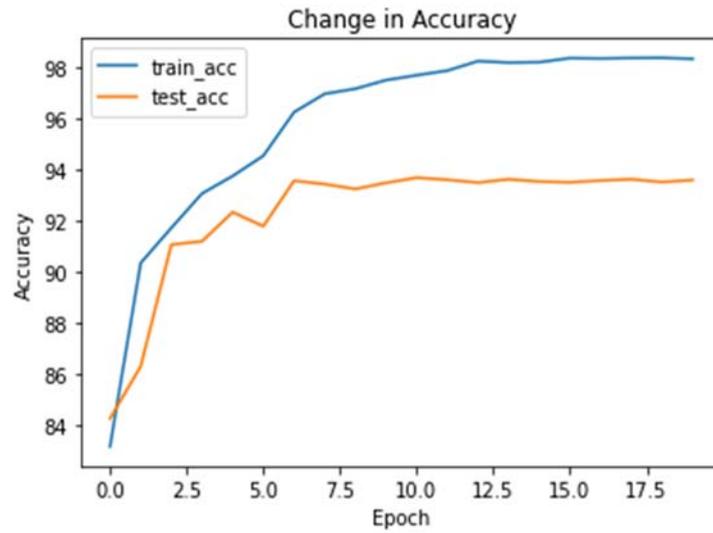


Figure 9: Train Vs Test accuracy of VGG 16

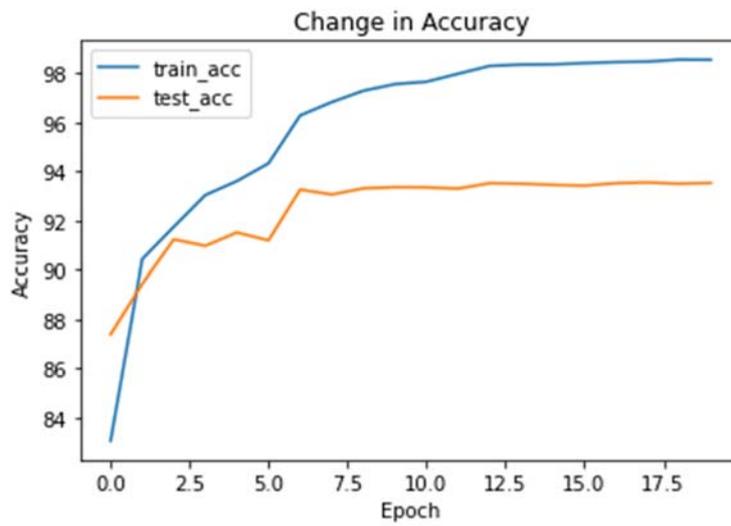


Figure 10: Train Vs Test accuracy of VGG 19

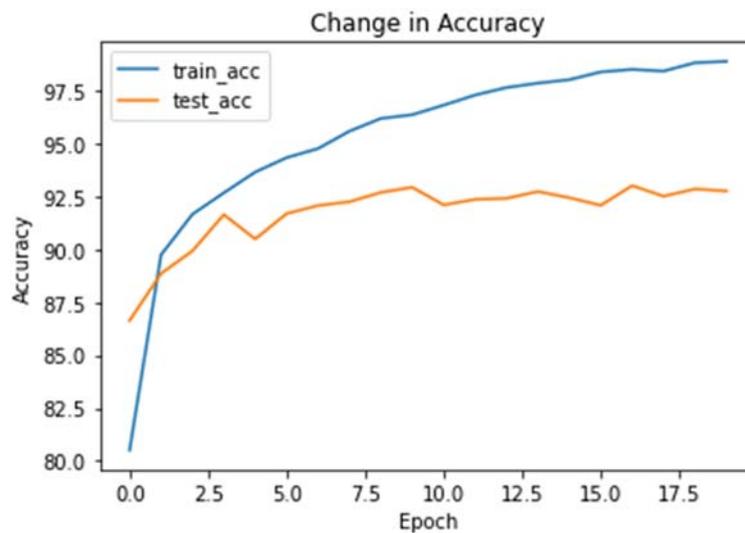


Figure 11: Train Vs Test accuracy of ResNet 18

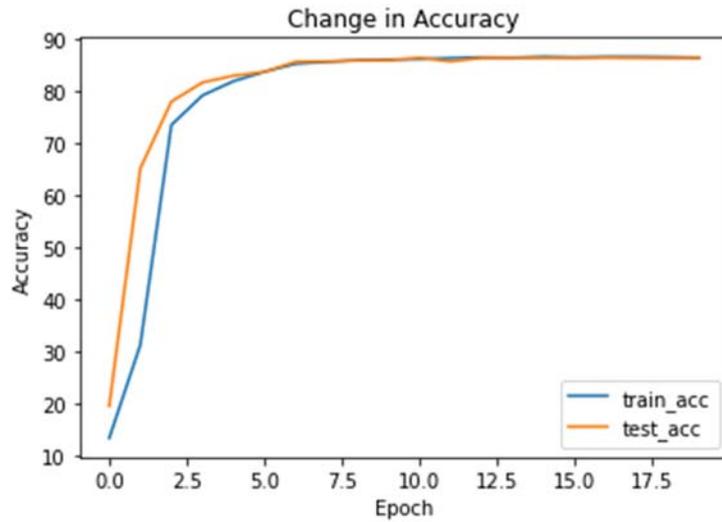


Figure 2: Train Vs Test accuracy of AlexNet.

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 32, 28, 28]	320
ReLU-2	[-1, 32, 28, 28]	0
BatchNorm2d-3	[-1, 32, 28, 28]	64
Dropout-4	[-1, 32, 28, 28]	0
Conv2d-5	[-1, 32, 28, 28]	9,248
ReLU-6	[-1, 32, 28, 28]	0
BatchNorm2d-7	[-1, 32, 28, 28]	64
Dropout-8	[-1, 32, 28, 28]	0
Conv2d-9	[-1, 32, 28, 28]	9,248
ReLU-10	[-1, 32, 28, 28]	0
BatchNorm2d-11	[-1, 32, 28, 28]	64
Dropout-12	[-1, 32, 28, 28]	0
MaxPool2d-13	[-1, 32, 14, 14]	0
Conv2d-14	[-1, 64, 14, 14]	18,496
ReLU-15	[-1, 64, 14, 14]	0
BatchNorm2d-16	[-1, 64, 14, 14]	128
Dropout-17	[-1, 64, 14, 14]	0
Conv2d-18	[-1, 64, 14, 14]	36,928
ReLU-19	[-1, 64, 14, 14]	0
BatchNorm2d-20	[-1, 64, 14, 14]	128
Dropout-21	[-1, 64, 14, 14]	0
MaxPool2d-22	[-1, 64, 7, 7]	0
Conv2d-23	[-1, 128, 7, 7]	73,856
ReLU-24	[-1, 128, 7, 7]	0
BatchNorm2d-25	[-1, 128, 7, 7]	256
Dropout-26	[-1, 128, 7, 7]	0
Conv2d-27	[-1, 128, 7, 7]	73,856
ReLU-28	[-1, 128, 7, 7]	0

BatchNorm2d-29	[-1, 128, 7, 7]	256
Dropout-30	[-1, 128, 7, 7]	0
Conv2d-31	[-1, 128, 7, 7]	147,584
ReLU-32	[-1, 128, 7, 7]	0
BatchNorm2d-33	[-1, 128, 7, 7]	256
Dropout-34	[-1, 128, 7, 7]	0
AvgPool2d-35	[-1, 128, 1, 1]	0
Conv2d-36	[-1, 10, 1, 1]	1,280

=====

7. Conclusion and Future Scope

The E-commerce industry has been evolving very rapidly. With more people moving towards online shopping, the amount of data being generated is enormous. The proper study of data will help to better engage customers. With the evolution of deep learning, this data can be better used to extract good information. Recurrent neural networks should be used to understand a customer's sentiments towards the products by analysing his reviews and comments. In the same way, convolutional neural networks should be used to produce better recommendations and show related content.

Deep learning frameworks are generally slow and time-consuming. Whereas engaging a customer in an E-commerce platform is real-time. An efficient neural network should be designed to get good recommendations in near real-time. It should be noted that a good recommendation doesn't mean the obvious one. They should be noval and acceptable.

References

- [1] J. Liu, L. Sun, R. Higgs, Y. Zhang and Y. Huang, "The Electronic Commerce in the Era of Internet of Things and Big Data," 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, 2017, pp. 360-363, doi: 10.1109/CSE-EUC.2017.252.
- [2] L. Li, "E-Commerce Data Analysis Based on Big Data and Artificial Intelligence," 2019 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 2019, pp. 133-138, doi: 10.1109/ICCNEA.2019.00034.
- [3] S. Vanaja and M. Belwal, "Aspect-Level Sentiment Analysis on E-Commerce Data," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, 2018, pp. 1275-1279, doi: 10.1109/ICIRCA.2018.8597286.
- [4] M. Tabibian and H. Amindavar, "E-commerce: technical and market approach," EUROCON'2001. International Conference on Trends in Communications. Technical Program, Proceedings (Cat. No.01EX439), Bratislava, Slovakia, 2001, pp. 149-152 vol.1, doi: 10.1109/EURCON.2001.937785.
- [5] Gera, P., Sabbisetty, V. B., Devarasetty, T., Nukala, M., & Vittamsetty, N. (2018). A fuzzy preference tree-based recommender system for medical database. International Journal of Engineering and Technology(UAE), 7(1.1), 319-321.
- [6] Blattberg R.C., Kim BD., Neslin S.A. (2008) Market Basket Analysis. In: Database Marketing. International Series in Quantitative Marketing, vol 18. Springer, New York, NY
- [7] Rupa Radhika Jahnvi, S., Kiran Kumar, K., & Sai Hareesh, T. (2018). A semantic web based filtering techniques through web service recommendation. International Journal of Engineering and Technology(UAE), 7(2), 41-43. doi:10.14419/ijet.v7i2.7.10254
- [8] M. H. Mohamed, M. H. Khafagy and M. H. Ibrahim, "Recommender Systems Challenges and Solutions Survey," 2019 International Conference on Innovative Trends in Computer Engineering (ITCE), Aswan, Egypt, 2019, pp. 149-155, doi: 10.1109/ITCE.2019.8646645.
- [9] Bhavesh Shri Kumar, Naren.J, Dr. G. Vithya and K. Prahathish, "A Novel Architecture based on Deep Learning for Scene Image Recognition", DOI: 10.37200/IJPR/V23I1/PR190251, Pages: 400-404, International Journal of Psychosocial Rehabilitation, March 2019.
- [10] Sara Ristagno"Types of e-commerce fraud you should lookout for and how to protect yourself" Oct 29, 2018, <https://www.ecommerce-nation.com/>
- [11] Artem Vovk, Dmitrii Tochilkin, Pradyumna Narayana, Kazoo Sone, and Sugato Basu. 2019. Product Phrase Extraction from e-Commerce Pages. In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, 393–397. DOI:<https://doi.org/10.1145/3308560.3316608>
- [12] M. Al-Jaberi, N. Mohamed and J. Al-Jaroodi, "e-commerce cloud: Opportunities and challenges," 2015 International Conference on Industrial Engineering and Operations Management (IEOM), Dubai, 2015, pp. 1-6, doi: 10.1109/IEOM.2015.7093867.
- [13] K. N. Asha and R. Rajkumar, "Pre-processing of user behaviour for e-commerce," 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore, 2017, pp. 715-719, doi: 10.1109/SmartTechCon.2017.8358464.
- [14] Shubhra Kanti Karmaker Santu, Parikshit Sondhi, and ChengXiang Zhai. 2017. On Application of Learning to Rank for E-Commerce Search. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17). Association for Computing Machinery, New York, NY, USA, 475–484. DOI:<https://doi.org/10.1145/3077136.3080838>
- [15] A. Marwade, N. Kumar, S. Mundada and J. Aghav, "Augmenting e-commerce product recommendations by analyzing customer personality," 2017 9th International Conference on Computational Intelligence and Communication Networks (CICN), Gime, 2017, pp. 174-180, doi: 10.1109/CICN.2017.8319380.
- [16] Alwahaishi, Saleh and Amine Nehari-Talet. "Opportunities and Challenges for the Electronic Commerce Deployment in Developing Countries Discussed in a Case of Saudi Arabia." Encyclopedia of Information Science and Technology, Third Edition, edited by Mehdi Khosrow-Pour, D.B.A., IGI Global, 2015, pp. 2339-2350. <http://doi:10.4018/978-1-4666-5888-2.ch227>

- [17] X. Ge, J. Liu, Q. Qi and Z. Chen, "A new prediction approach based on linear regression for collaborative filtering," 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shanghai, 2011, pp. 2586-2590, doi: 10.1109/FSKD.2011.6020007.
- [18] H. Tian, H. Cai, J. Wen, S. Li and Y. Li, "A Music Recommendation System Based on logistic regression and eXtreme Gradient Boosting," 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 2019, pp. 1-6, doi: 10.1109/IJCNN.2019.8852094.
- [19] Min SH., Han I. (2005) Recommender Systems Using Support Vector Machines. In: Lowe D., Gaedke M. (eds) Web Engineering. ICWE 2005. Lecture Notes in Computer Science, vol 3579. Springer, Berlin, Heidelberg
- [20] Eyal Shulman and Lior Wolf. 2020. Meta Decision Trees for Explainable Recommendation Systems. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20). Association for Computing Machinery, New York, NY, USA, 365–371. DOI:https://doi.org/10.1145/3375627.3375876
- [21] Srivats. S. Ramanujam, Sivaneshwar.P. Naren.J, Madhumitha.S, Dr. G. Vithya, "A Study on Hybrid Recommender System with Deep Learning and Deployment in Big Data", Vol 81, Page No. 1869-1875, TEST Engineering and Management, Dec 2019..
- [22] Srivats. S. Ramanujam, Sivaneshwar.P. Naren.J, Madhumitha.S, Dr.G.Vithya, "A Novel Architecture for the development of an e-Learning Hybrid Recommender System with Deep Learning Techniques in Big Data Environment", Pages: 2504-2510, Volume 11 | 08-Special Issue, Journal of Advanced Research in Dynamical and Control Systems – JARDCS, Dec 2019.
- [23] Marius-Constantin Popescu, Valentina E. Balas, Liliana Perescu-Popescu, and Nikos Mastorakis. 2009. Multilayer perceptron and neural networks. WSEAS Trans. Cir. and Sys. 8, 7 (July 2009), 579–588.
- [24] Marhon S.A., Cameron C.J.F., Kremer S.C. (2013) Recurrent Neural Networks. In: Bianchini M., Maggini M., Jain L. (eds) Handbook on Neural Information Processing. Intelligent Systems Reference Library, vol 49. Springer, Berlin, Heidelberg.
- [25] M. Schuster and K.K. Paliwal. 1997. Bidirectional recurrent neural networks. Trans. Sig. Proc. 45, 11 (November 1997), 2673–2681. DOI:https://doi.org/10.1109/78.650093
- [26] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. Neural Comput. 9, 8 (November 15, 1997), 1735–1780. DOI:https://doi.org/10.1162/neco.1997.9.8.1735
- [27] R. Dey and F. M. Salem, "Gate-variants of Gated Recurrent Unit (GRU) neural networks," 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, 2017, pp. 1597-1600, doi: 10.1109/MWSCAS.2017.8053243.
- [28] Hari Kishore, K., Noorbasha, F., Sandeep, K., Bhupesh, D. N. V., Khadar Imran, S. K., & Sowmya, K. (2018). Linear convolution using UT vedic multiplier. International Journal of Engineering and Technology(UAE), 7(2.8 Special Issue 8), 409-412.
- [29] Ismail, M., Prakash, K. B., & Rao, M. N. (2018). Collaborative filtering-based recommendation of online social voting. International Journal of Engineering and Technology(UAE), 7(3), 1504-1507.
- [30] Kishore, P. V. V., Anantha Rao, G., Kiran Kumar, E., Teja Kiran Kumar, M., & Anil Kumar, D. (2018). Selfie sign language recognition with convolutional neural networks. International Journal of Intelligent Systems and Applications, 10(10), 63-71. doi:10.5815/ijisa.2018.10.07
- [31] Kishore, P. V. V., Kumar, K. V. V., Kiran Kumar, E., Sastry, A. S. C. S., Teja Kiran, M., Anil Kumar, D., & Prasad, M. V. D. (2018). Indian classical dance action identification and classification with convolutional neural networks. Advances in Multimedia, 2018 doi:10.1155/2018/5141402
- [32] Narasinga Rao, M. R., Venkatesh Prasad, V., Sai Teja, P., Zindavali, M., & Phanindra Reddy, O. (2018). A survey on prevention of overfitting in convolution neural networks using machine learning techniques. International Journal of Engineering and Technology (UAE), 7(2.32 Special Issue 32), 177-180.
- [33] Prasad, M. V. D., Lakshamma, B. J., Chandana, A. H., Komali, K., Manoja, M. V. N., Kumar, P. R. Kiran, P. S. (2018). An efficient classification of flower images with convolutional neural networks. International Journal of Engineering and Technology (UAE), 7(1.1), 384-391.