International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 04 | Apr 2023 www.irjet.net p-ISSN: 2395-0072

Pet Care Application

Ratnangsu Chatterjee, Apoorvi Singh

Apoorvi Singh, Student SRM Institute of Science and Technology, Tamil Nadu, India
Ratnangsu Chatterjee, Student & SRM Institute of Science and Technology, Tamil Nadu, India
Mrs. Saveetha D, Professor, Dept. of Networking and Communications, SRM Institute of Science and Technology, Tamil Nadu, India

---

Abstract - *Recently there has been a major uptick in the number of pet owners in India. In the past few years, the number of pet owners has increased. Thus, it has become increasingly important for proper pet care applications and personnel. Knowing the correct breed of a pet is crucial for providing adequate care. The pet's age, height, and weight all have a significant impact on how well they respond to a healthy diet, regular exercise, and other measures of wellness. These considerations determine dietary choices. Having more pets increases the risk of their deaths, as it is difficult to properly care for any animal. As a result, it is crucial to know the correct disease in the early stages as it can save the lives of many pets. The ability to predict breed more effectively is a strength made possible by deep learning, a model of algorithms capable of tackling informational challenges. It is possible to categories and forecast based just on the occurrence of raw data as a source of information. An example, Convolution Neural Networks (CNNs) provide a common framework for picture Sorting and spotting. In this effort, we use a convolution neural network (CNN) for detecting canines in randomly generated photos thereby demonstrating carelessness with regard to attribution of a rare and unique canine breed. The analysis was validated using commonly used metrics; hence the diagram display verifies that the algorithm (CNN) produces accurate evaluation precision across all data- sets. One area where machine learning has found use is in disease prediction. More cutting-edge medical technology is needed to give patients the best care possible. Decision Tree Classifier in Na¨ıve Bayes algorithm were chosen and applied to the data to achieve the best possible outcomes. Machine learning's potential in healthcare has been demonstrated, and it has the potential to significantly improve patients' health. Therefore, the purpose of this study is to investigate the feasibility of integrating machine learning into an existing veterinary healthcare infrastructure for pets. The entire therapy procedure can be optimised if the disease is forecast-ed in advance using specific machine learning algorithms rather than being performed directly on the patient. The failure to perform or carry out early diagnosis of an illness may also occur. As a result, anticipating the course of a disease is crucial. As the old adage goes, "Prevention is better than cure," so accurate disease forecasting would undoubtedly result in earlydisease prevention.*

Key Words: Convolution Neural Network(CNN), Decision Tree, Machine Learning, Deep learning, Breed Prediction, Dis- ease Prediction

1.INTRODUCTION

In India, over 19.5 million dogs were maintained as companion animals in 2018, according to recent estimates, it is anticipated that there will be more than 31 million people living with their pets in the India by 2023. Thus, it is important for pet owners to properly understand their pets' breed and its particular needs.

1.1 Breed Prediction using CNN

Images of dogs will be analysed in an effort to determine their breed. Since all dog breeds are similar in terms of their physical traits and overall structure, distinguishing between them is a challenging problem, making this a fine-grained classification problem. Furthermore, there is minimal inter- breed and high intra-breed variation; that is, there are relatively few differences between breeds and there are quite substantial changes within breeds, including size, shape, and colour. The canine species is the most genetically and physically varied on Earth. Photographs showing dogs of the same breed in differ- ent lighting and poses add to the challenges of breed identifica- tion brought on by the dataset's diversity. This problem is not only difficult, but the solution is also applicable to other fine- grained classification problems. The methods used to solve this problem, for example, would aid in the identification of cat and horse breeds, as well as bird and plant species - and even car models. As a fine-grained classification problem, any set of classes with relatively little variation within it can be solved. In the real world, such an identifier could be used in biodiversity studies, saving scientists time and resources when conducting research on the health and abundance of specific species populations. These studies are critical for assessing the state of ecosystems, and their accuracy is especially important because of their influence on policy changes. Breed prediction may also aid veterinarians in treating breed-specific ailments in stray, unidentified dogs in need of medical attention. We eventually decided that dogs were the most interesting

class to experiment with because of their enormous diversity, loving nature, and abundance in photographs, but we also hope to broaden our understanding of the fine-grained classification problem and provide a useful tool for scientists across disciplines.

1.2 Disease Prediction Using Decision Tree Classifier

Early and accurate disease prediction is also a necessary factor in helping new pet owners. The healthcare industry generates and uses a large amount of data that can be used to extract information about a specific disease in a pet. This healthcare information is be used to provide the most effective and best treatment for the health of your pet. This area also requires some improvement through the use of informative data in healthcare sciences. However, because there is so much data, extracting information from it can be difficult, so data mining and machine learning techniques are used. The expected outcome of this project is to predict the disease in advance so that the risk of death can be avoided at an early stage, saving the lives of pets and reducing the cost of treatment to a certain extent. The main goal is to improve pet care by incorporating the concept of machine learning into healthcare. Machine learning has already made identifying and forecasting various diseases much easier. Predictive disease analysis using many machine learning algorithms allows us to predict the disease and treat the pets effectively. Disease prediction using machine learning also makes use of the pet's history and health data by employing various concepts such as data mining and machine learning techniques, as well as some algorithms. Deep learning research in disparate areas of machine learning has led to a shift toward machine learning models that can learn and understand hierarchical representations of raw data with some pre-processing.

2. OBJECTIVE

Our mission is to provide assistance to pet owners so that they can better understand and care for their animals.

- The ability to accurately identify the early stages of any disease can save the lives of a great number of pets.
- Deepening our understanding of AI and machine learning across a range of fields and putting that understanding to use. Utilizing knowledge of a variety of machine learning algorithms and artificial intelligence designs.

3. RELATED WORKS

Deep learning has been a part of the machine learning world for quite some time. According to the authors Foote, 2017, the origins of deep learning can be traced back to 1943, when Walter Pitts and Warren McCulloch attempted to design a computer based on the neural network of the human intellect. Foote, 2017 states that "the earliest efforts in developing Deep Learning algorithms came from Alexey Grigoryevich Ivakhnenko (developed the Group Method of Data Handling) and Valentin Grigoryevich Lapa (author of Cybernetics and Forecasting Techniques) in 1965." Deep Learning is a subset of machine learning that employs multi-layer neural networks to carry out operations. Each layer of the neural network is made up of several neurons that are all linked in such a way that they can communicate with one another. This neuron was created in the hope that it would function similarly to the neuron in human intelligence. In this case, the neuron would attempt to calculate the weighted average of the values, i.e. the input signal and the output signal transmitted by the connected neuron. Arora et al., 2015. In Nokwon Jeong and Soosun Cho's 2017 paper, the authors attempt image classification on Instagram images, with the goal of evaluating the competitive power of deep learning for classification of real-time social networking images. In their study, the authors Nokwon Jeong and Soosun Cho (2017) look at the performance of pre-existing CNN frameworks such as AlexNet and ResNet and how well they perform on the ImageNet dataset, which demonstrated outstanding capabilities.

4. SYSTEM DESIGN

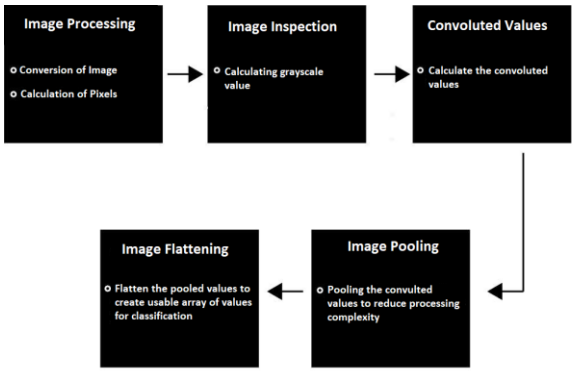

Fig-1 Block Diagram Breed Prediction

4.1.1 Breed Prediction

For breed prediction we use image recognition. We create a Simple image recognition agent that uses a pre-trained CNN to accurately recognise the breed of a dog using image processing. We use a accuracy metric to test the accuracy of 5 different pre- trained CNN namely, VGG16, VGG19, RESNET50, Inception and Xception. Out of these 5 Exception was found out to be the most accurate at 83.78% accuracy. A data-set of 133 dog breed was taken from Kaggle to train these CNNs and a pre-trained human

and dog face detector was taken from the github link of OpenCV.

4.1.2 Disease Prediction

Disease prediction was done by using Decision Tree Classifier and Naive Bayes algo- rithm. A data-set of dog disease was taken from kaggle. The data-set was divided into three columns based on disease name, frequency of occurrence and symptoms. Using these data a decision tree was created by cleaning and arranging the data.This decision tree was then used to train the agent using Naive Bayes algorithm. Naive Bayes algorithm was used as it is a slow yet accurate method of predicting in conditions where probability is present.

5. METHODOLOGY

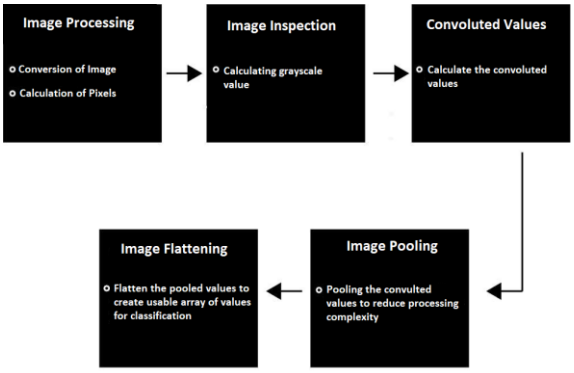

Fig-2 Block Diagram Disease Prediction

5.1 Breed Prediction using CNN

A convolution network consumes such images as three distinct colour strata stacked one on top of the other. A standard colour image is viewed as a rectangular box whose width and height are determined by the number of pixels in those dimensions. Channels are the depth layers in the three layers of colours (RGB) interpreted by CNNs.Such images are consumed by a convolution network as three colour layers, one on top of the other. A standard colour image is perceived as a box, with width and height determined by the amount of pixels used. Channels refer to the depth levels in the three layers of colours (RGB) used by CNNs for their interpretation. The CONVOLUTION LAYER is the core building block of a CNN network and does the majority of the computational heavy lifting. Filters or kernels are used to convolve data or images. Filters are small units that we apply to data via a sliding window. The depth of the image is the same as the depth of the input; for a colour image with an RGB depth of 4, a depth 4 filter would also be applied to it. This procedure entails taking the element-wise product of the image's filters and then summing those specific values for each sliding action. A 2d matrix would be the output of a convolution with a 3d filter and colour.Now, imagine a flashlight shining over the top left corner of the image to explain a convolution layer. To understand how this works, imagine a flashlight shining its light over a 5 x 5 area. Now imagine this flashlight moving across all of the areas of the input image. This flashlight is known as a filter (also known as a neuron or a kernel), and the region it illuminates is known as the receptive field. This filter is also a number array (the numbers are called weights or parameters).The second layer is the ACTIVATION LAYER, which uses the ReLu (Rectified Linear Unit). In this phase, we use the rectifier function to increase non-linearity in the CNN. Images are made up of various objects that are not linearly related to one another. The third layer is the POOLING LAYER, which incorporates feature down- sampling. It is applied to each layer of the 3D volume. This layer typically contains the following hyper- parameters. A typical POOLING LAYER employs a non-overlapping 2 cross 2 max filter with a stride of A max filter would return the maximum value in the region's features. When there is a volume of 26 across 32, the volume can be decreased to 13 crosses, 32 feature map by utilising a max pool layer with 2 cross 2 filters and an astride of 2. Finally, then comes the FULLY CONNECTED LAYER, which requires flattening. The entire pooling feature map matrix is transformed into a single column, which is then supplied to the neural network for pro- messing. We created a model by combining these features using fully connected layers. Finally, to classify the output, we have an activation function such as soft-max or sigmoid.

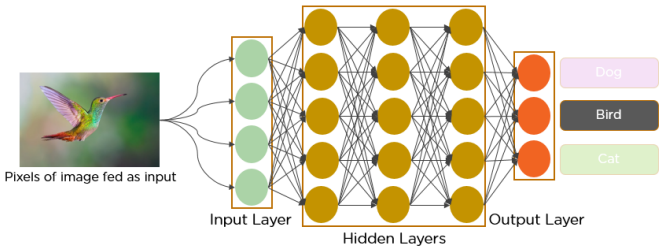

Fig. 3. Layers of a CNN

5.2 Disease Prediction using Naive Bayes Algorithm

The Naive Bayes method is a supervised learning technique that uses the Bayes theorem to solve classification issues. It is mostly utilised in text classification with a large training data-set. The Nave Bayes Classifier is a simple and effective Classification method that aids in the development of fast machine learning models capable of making quick predictions. It is a probabilistic classifier, which means it predicts based on an object's likelihood.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \quad (1)$$

P(A|B) - Probability of A on event B
P(B|A) - Probability of the evidence given that the probability of a hypothesis is true
P(A) - Probability of hypothesis before observing the evidence
P(B) - Probability of Evidence

Fig-3 Formula Used

6. MATERIALS AND METHOD

6.1 Breed Prediction

6.1.1 Statistics and Datasets

The dataset used to train the image recognition soft- ware was taken from kaggle. It includes 3 folders of labelled pictures of different dog breeds divided into three sets train,test and valid. The train folder contains pictures of 133 dog breed along with its name and are well-labelled. The pictures for the dog breed was sorted in alphabetical order for ease of reading and access. Another dataset of human faces was also used that was required to train the human/dog face detectors to distinguish between human dog faces.

6.1.2 Data Preprocessing and Visualization

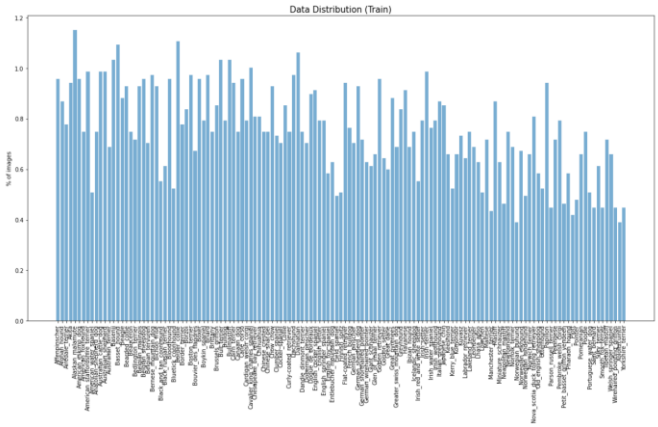

Fig-4 Distribution of Data in Dataset

The acquired dataset was put into the system, where it was transformed and read. It is divided into three categories: test, train, and valid datasets. Each image was measured and added to a list based on its size. The size was then plotted along a graph, followed by a feature pixel. Pixels from all three datasets were combined and plotted together. Finally, the dataset was partitioned into available data. Following the division of the dataset, the animals' breeds were enumerated and plotted on a graph.

6.1.3 Model and Approach

The dataset was imported and the data in it was read, sorted and visualised. A pre-trained face recognition agent was download from OpenCV github and trained using the human dataset and dog dataset to detect humans and dogs. after the agents were trained some random images where imported and the face detector was tested. After a satisfying level of accuracy was reached the face detector was imported in CNN and 5 different pre-trained CNNs were tested using an ac- curacy matrix and the one with highest accuracy was finally used as our CNN.the CNNs used where VGG16, VGG19, RESNET50, Inception and Xception.

6.1.3.1 VGG16

VGG16 is a convolutional neural network model proposed in the publication "Very Deep Convolutional Networks for Large-Scale Image Recognition" by K. Simonyan and A. Zisserman of the University of Oxford. In ImageNet, a dataset of over 14 million images classified into 1000 classes, the model achieves 92.7% top-5 test accuracy. It was one of the well-known models submitted to the ILSVRC-2014.

For our test CGG16 achieved an accuracy of 49.76%.

6.1.3.2 VGG19

VGG-19 is a 19-layer deep convolutional neural network. A pretrained version of the net- work trained on over a million photos from the Im- ageNet database can be loaded [1]. The pretrained network can categorise photos into 1000 different object categories, including keyboards, mice, pen- cils, and various animals. As a result, the network has learned detailed feature representations for a diverse set of images.

For our test CGG16 achieved an accuracy of 50.00%.

6.1.3.3 RESNET50

Residual Networks, or ResNets, learn residual functions with reference to the layer inputs rather than learning unreferenced functions. Instead of hoping that each few stacked layers directly fit a desired underlying mapping, residual nets allow these layers to fit a residual mapping. They stack residual blocks on top of each other to form networks, such as a ResNet-50, which has fifty layers.

For our test CGG16 achieved an accuracy of 79.98%.

6.1.3.4 Inception

An inception network is a deep neural network with an architectural design made up of re- peating components known as Inception modules. For our test CGG16 achieved an accuracy of 77.75%.

6.1.3.5 Xceptionn

An inception network is a deep neural network with an architectural design made up of re- peating components known as Inception modules. For our test CGG16 achieved an accuracy of 77.75%..

For our test CGG16 achieved an accuracy of 83.25%.

Thus, Xcpetion has the best accuracy in identifying the dog breed through image and hence it was used to design our agent.

Architecture	Trainable Parameters	Train Accuracy	Valid Accuracy	Test Accuracy
Custom	24,831,909	97.6%	5.03%	4.19%
VGG16	68,229	63.58%	49.7%	49.76%
VGG19	68,229	61.74%	48.14%	50.0%
ResNet-50	272,517	99.82%	82.75%	79.9%
Inception	272,517	98.88%	86.23%	77.75%
Xception	272,517	97.98%	84.91%	**83.25%**

Fig-5 Evaluation Metrics

7. DISEASE PREDICTION

7.1. Statics and Data

The dataset used to analyse dog diseases came from the website kaggle. The dataset includes a variety of human diseases, each of which was broken down into it's own column and labelled with the disease's name, the count of disease occurrences, and the symptoms.

7.2. Data Preprocessing and Visualisation

A dataset that had not been cleaned up was obtained from Kaggle. The raw dataset was brought into the pro- gram where it was read after being imported there. The dataset was cleaned up by first removing all of the values that were NaN. After that, the tables were rearranged so that the disease occurrences were listed in descending order. Following that, the table was partitioned using lambda in accordance with the symptoms. After that, this dataset was input into an AI system that employs decision trees based on the Naive Bayes algorithm.

7.3. Model and Approach

We used a number of different analytical data mining techniques in order to estimate the most accurate illness that could be associated only with the patient's condi- tion. Additionally, we use an algorithm called Naive Bayes in order to map the symptoms with potential diseases based on a database that contains multiple disease symptoms records. Patients not only benefit from this system be- cause it makes the doctors' jobs easier but also because it ensures that patients receive the care they require as quickly as possible. The accuracy score on the test and train dataset was compared with the X-axis and Y-axis to get the accuracy number. Additionally for train dataset we use gradient boosting classifier. By cross- validating the mean was calculated to be 100%. We checked the discrep- ancies between the actual values and the predicted values so as to prevent wrong prediction results. then k-fold was imported and multiple different algorithms were tested with different values of k-fold. The best outcome was gained when the k-value was set to 2. so the model was build on the basis of k-value being 2.

8. RESULTS AND CONCLUSION

8.1. Result

With the increase int he number of pet owner in the coun- try it is very essential to get proper attention and care for the pets. hence our project aims at streamlining the pro- cess of getting a better understanding of our pets and tak- ing better care of there health by using image prediction software to understand their breed and disease prediction to get an idea of what they might be suffering from the conclusion regarding the prediction accuracies gained through the application of the CNN procedure to a variety of standard datasets. The results have not changed thanks to the estimate accuracy in proportion that was calculated separately for test data and inside train information. In addition to the prediction accuracy percentage values, a graph displaying the MSE, or mean squared error, is also provided. The graphs display the variation of MSE as a function of regard along the path to the training epochs. The MSE metric is the simplest and most widely utilised of all the quality metrics. Thus, a very accurate method of identifying the breed of a dog was reached by using Xcep- tion neural Network, which helped us in getting a high percentage of accuracy in determining the breed of the dog.

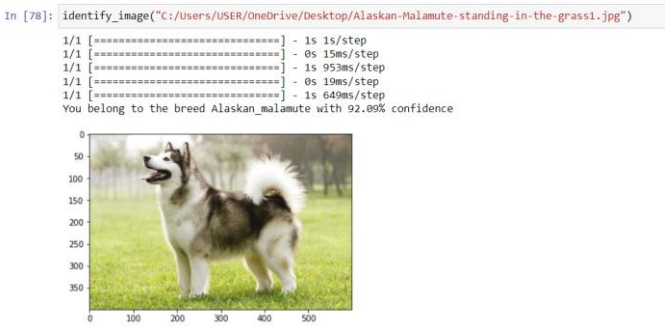

Fig-6 Alaskan Malamute

Our primary purpose of was to comprehend and enhance the method of disease prediction, as well as to carry out a comparative analysis of algorithms in order to locate

the algorithm that was most ideally suited. The accuracy scores of the algorithms were compared, and in addition to that, data visualisation was done in order to gain a more in-depth understanding of the data and the trends within it. In the end, when all of the results obtained by the various algorithms used for disease prediction from hospital data were compared to one another, it was found that the CART model, also known as a decision tree, gave the highest performance.

Fig-7 Disease Prediction Result

8.2. Conclusions

In addition to its use as a method for data analysis and prediction, convolutional neural networks have recently seen a surge in popularity as a solution for problems involving the classification of images. The goal of this deep learning method for predicting dog breeds, which was developed with the help of a convolutional neural network, is to determine of one hundred images by using their images as input. Utilize transfer learning as a means toward building a model that can produce output and around hundreds of different dog breeds. The results for the images that were shown to the model were satisfactory overall. The algorithm was very accurate when it came to determining the breeds of dogs. Transfer learning requires a great deal of flexibility in the future in terms of combining a prebuilt model with the model that we developed. It is possible to achieve comparatively higher levels of ac- curacy using the system that we have proposed. After that, researchers, doctors, and other medical professionals will use this in order to provide patients with the most effective treatment and medical care possible. Therefore, the application of machine learning in the medical field can result in an effective treatment, while also ensuring that the patient receives adequate care. In this section, we make an attempt to incorporate some of the machine learning in healthcare functions that are available into our system. When a disease is predicted for a patient, machine learning is implemented instead of direct diagnosis. Certain machine learning algorithms are used during this process, and as a result, healthcare can be made more intelligent and effective. The Logistic Regression algorithm and the KNN algorithm have the highest accuracy when compared to the other algorithms used for disease prediction based on our dataset and the output we anticipate. This is the case when we analyse both the input data and the expected output.

9. FUTURE WORKS

Future research should look into the potential of convolutional neural networks in predicting dog breeds. Given the success of our keypoint detection network, this technique looks promising for future projects. However, neural networks take a long time to train, and due to time constraints, we were unable to perform many iterations on our technique. We recommend further research into neural networks for keypoint detection, specifically training networks with a different architecture and batch iterator to see which approaches may be more suc- cessful. Furthermore, given our success with neural networks and keypoint detection, we recommend implementing a neural network for breed classification as well, as this has not been done previously. Finally, neural networks take time to train and iterate on, which should be taken into account for future efforts; however, neural networks are formidable classifiers that will improve prediction accuracy over more traditional techniques.

As we can clearly see today, computers and technology are being used to consider a massive amount of data, computers are being used to perform various complex tasks with commendable accuracy rates. Machine learning (ML) is a collection of techniques and algorithms that allow computers to perform such complex tasks in a simplified manner. We can say that we have grown in the fields of big data, machine learning, and data sciences, among other things, and that we have been a part of one of those industries that have been able to collect such data and staff to transform their goods and services in the desired manner. The learning methods developed for these industries and researches have tremendous potential to improve medical research and clinical care for patients in the best way possible. Machine learning employs mathematical algorithms and procedures to describe the relationship between model variables and others. Our paper will describe the process of training the model and learning an appropriate algorithm to predict the presence of a specific disease from a tissue sample based on its features. Though these algorithms work in different and distinct ways depending on how they are developed and used by the researchers. One approach is to consider their ultimate goals. The purpose of our paper and statistical methods is to reach a conclusion about the data collected from a wide range of samples drawn from our population. Although many techniques, such as linear and logistic regression, can predict diseases.

ACKNOWLEDGMENT

We express our humble gratitude to Dr C. Muthamizhchelvan, Vice-Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support. We extend our sincere thanks to

Dean-CET, SRM Institute of Science and Technology, Dr T.V.Gopal, for his valuable support.

We wish to thank Dr Revathi Venkataraman, Professor and Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work. We are incredibly grateful to our Head of the Department, Dr. Annapurani Panaiyappan.K Professor, Department of Net- working and Communications, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our Panel Head, Vinoth Kumar S, Associate Profess, Department of Networking and Communications, SRM Institute of Science and Technology, for their inputs during the project reviews and support. We register our immeasurable thanks to our Faculty Advisor, Dr.P Supraja, Associate professor, Department of Networking and Communication, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to my guide, Mrs. Eliza- beth Jesi, Associate professor, Department of Networking and Communication, SRM Institute of Science and Technology, for providing us with an opportunity to pursue our project under her mentorship. She provided us with the freedom and support to explore the research topics of our interest.

We sincerely thank the Networking and Communications Department staff and students, SRM Institute of Science and Technology, for their help during our project.

REFERENCES

1. Kaitlyn Mulligan and Pablo Rivas- Dog Breed Identification with a Neural Network over Learned Representations from The Xception CNN
2. J Sreenand Manoj, Rakshith S, Kanchana V- Identification of Cattle Breed using the Convolutional Neural Network
3. Bickey Kumar shah ,Aman Kumar,Amrit Kumar-Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks
4. Dr. D. Durga Bhavani, Mir Habeebullah Shah Quadri, Y. Ram Reddy- Dog Breed Identification Using Convolutional Neural Networks on Android
5. Yu Han LIU-Feature Extraction and Image Recognition with Convolu- tional Neural Networks
6. Brankica Bratic,Vladimir Kurbalija, Mirjana Ivanovic,Iztok Oder, Zoran Bosnic, Machine Learning for Predicting Cognitive Diseases: Methods, Data Sources and Risk Factors ,2018
7. K. Gomathi, D. Shanmuga Priya-Multi Disease Prediction using Data Mining
8. S.Vijiyarani,S.Sudha-Disease Prediction in Data Mining Technique,January 2013
9. Emily Jones , John Alawneh , Mary Thompson , Chiara Palmieri , Karen Jackson and Rachel Allavena-Predicting Diagnosis of Australian Canine and Feline Urinary Bladder Disease Based on Histologic Features,November 2020
10. Sneha I. Kadari, Shubhada S. Kulkarni, Sharada G. Kulkarni-Dog Breed Prediction using Convolutional Neural Network,June 2020
11. Kriti Gandhi1, Mansi Mittal2, Neha Gupta3, Shafali Dhall-Disease Prediction using Machine Learning,June 2020
12. Kriti Gandhi1, Mansi Mittal2, Neha Gupta3, Shafali Dhall-Disease Prediction using Machine Learning,June 2020
13. Keith D. Foote, (2017). A Brief History of Deep Learning - DATAVER- SITY.
14. Arora, A., Candel, A., Lanford, J., LeDell, E. and Parmar, V. (2015).Deep Learning with H2O. 3rd ed.
15. NokwonJeong, Soosun Cho (2017)' Instagram image classification with Deep Learning