A Study on identification of dog breeds through multi-class classification using Deep Learning techniques

Dissertation submitted in part fulfilment of the requirements for the degree of

[**MSc in DATA ANALYTICS**]

at Dublin Business School

Mr. Renji Roy Isac

MSc in Data Analytics 2018-2019

Declaration: I, **Renji Roy Isac**, declare that this research is my original work and that it has never been presented to any institution or university for the award of Degree or Diploma. In addition, I have referenced correctly all literature and sources used in this work and this work is fully compliant with the Dublin Business School's academic honesty policy.

Signed: ______________________

Date: _____16/12/2018_________

ACKNOWLEDGEMENT

I would like to express my gratitude primarily towards Dublin Business School for giving me the opportunity and the platform to be able to conduct my dissertation. I would also like to extend my gratitude towards my mentor Ms. Terri Hoare who has been instrumental throughout the research process. Our constant brainstorming sessions have helped me maintain my focus throughout the project.

Finally, I must thank my peers, for their constant motivation and valuable inputs throughout, which enabled me to complete the research successfully.

Table of Contents

LIST OF ALGORITHMS 5
LIST OF FIGURES 6
ABSTRACT 7
CHAPTER – 1 – INTRODUCTION 8
1.1 Background 8
1.2: Exploring the Concepts 8
1.3 Current Research 10
CHAPTER-2 REVIEW OF LITERATURE 11
2.1: Background 11
2.2: Related Work 15
CHAPTER 3 – METHODOLOGY 18
3.1 Dataset 18
3.2 Data Pre-processing 19
3.3 Modelling 20
3.3.1 Primary Approach 23
3.3.2 Secondary Approach 25
CHAPTER – 4 ANALYSIS AND DISCUSSION 32
CHAPTER – 5 CONCLUSION 38
5.1 Future works: 39
5.2 Business Applications: 39
REFERENCES 41
Appendix A 43
Appendix B 45
Appendix C 46
GLOSSARY OF TERMS 47

LIST OF ALGORITHMS

1. Deep Learning H20
2. Support Vector Machine (Kernel -Linear)
3. Polynomial by Binomial Classification
4. Vote (for classification)
5. Simple Deep Learning
6. Deep Learning on Tensor
7. Support Vector Machine (Kernel- Radial)
8. Support Vector Machine (Kernel – Polynomial)

LIST OF FIGURES

Figure 2.1: Computer Vision	**11**
Figure 2.2: Working of a Neuron	**13**
Figure 2.3: Classification using support vectors for SVM	**14**
Figure 3.1: Snapshot of the Stanford Dog Dataset	**18**
Figure 3.2: Data pre-processing stages	**19**
Figure 3.3: Process Architecture	**23**
Figure 3.4: Deep Learning Process	**25**
Figure 3.5: Process for dogs with similar features- deep learning	**28**
Figure 3.6: Process for dogs with distinctive features- deep learning	**29**
Figure 3.7: Process for dogs with similar features- SVM	**30**
Figure 3.8: Process for dogs with distinctive features- SVM	**31**
Figure 4.1: Different shades for Labradors	**32**
Figure 4.2: Results for dogs with distinctive features process- deep learning	**33**
Figure 4.3: Results for dogs with distinctive features process- SVM	**34**
Figure 4.4: Results for dogs with similar features process- SVM	**34**
Figure 4.5: Results for dogs with similar features process- deep learning	**34**
Figure 4.6: Process for polynomial by binomial model	**36**
Figure 4.7: Results for the vote operator process	**37**

ABSTRACT

Image classification, a discipline which over the years has made tremendous advancements with new and improved techniques continuously being implemented to improve the accuracy. With an enormous amount of effort being put into the field, multiclass classification has proven to be particularly challenging. Hence, in the present study the researcher focuses on achieving multiclass classification on dog breed identification using state of the art deep learning techniques. Having previously collaborated with NGO's that managed dog adoptions, the researcher saw first-hand how wrongful breed classification affected the lives of these feeble beings by keeping them from getting the home they deserved. This in turn helped the researcher realize the urgency of such an identification system. Moreover, the researcher also focuses on a comparative study between deep learning and support vector machines which has proven to be particularly effective in binary classification. The results observed during the research inferred that 'multiclass classification' posed a problem even for deep learning however from the results observed from the comparative study, there was some light shed on how these problems could be tackled in the future. The results are presented.

CHAPTER – 1 – INTRODUCTION

1.1 Background

The evolution of human beings has been known since time immemorial with an awareness of various kinds of differences that exist between us. But beside us, another category that add onto the infinite number is the huge herd of animals that exist alongside. Like humans, there too exists a variety of breed, shapes and sizes and identifying each of them becomes a tough job, because often the breeds that are unknown or those that belong to mixed-lineages are identified with mere guesswork which brings the question of accuracy. Hence, the present research purports to identify dog breed through image classification using deep learning techniques. Having previously worked in the pet services sector has given the author immense amount of experience regarding issues pertaining to the pet population; be it pets being abandoned on the streets for not turning out to be the breed they were thought to be or not being adopted as they were assumed to belong to a certain breed. Such first-hand experience has been a strong source of motivation behind this research. Moreover, in an article published by Maddie's Fund, the author talks about how, when put to test even professionals who work with dogs, day in and day out found it difficult to identify the correct breed of dog. Furthermore, the author talks about a study that was conducted on 120 dogs and a total of 16 observers (who were dog shelter staff), where they were asked to classify the breed. Out of the 120 dogs 55 were classified as Pitbulls, when in reality the tests showed that out of the 55 only 25 were actually pitbulls, the rest just shared similar features (R. Olson and K. Levy, 2012). Wrongful classification of the staff has often led to the dogs not being adopted or in countries where pitbulls are considered a threat, has led to the dogs being euthanized (Greenwood, 2015). With the lives of these dogs literally being put on the line, there is an actual need for being able to perform multiclass classification as such identification tool is the need of the hour.

1.2: Exploring the Concepts

Image classification, a process through which information classes are abstracted from a multiband raster image has made tremendous strides over the years. Certain key changes have taken place regarding ways of dealing with the concept of image classification. Although, efforts in this field saw failures, they have also opened doors to solutions to various problems. One of the underlying assumptions regarding image classification is that the spectral response of any particular feature remains relatively consistent throughout the image. Moreover, there exists two different kinds of image classification namely; an unsupervised image classification, which recognizes those group of pixels which express a similar spectral response and then there is supervised image classification which categorize unknown pixels that represent regions of known composition. Present research is mainly focused on supervised classification.

"When programmable computers were first conceived, people wondered whether such machines might become intelligent, over a hundred years before one was built" (Lovelace, 1842,cited in Goodfellow et.a., 2016). However, in the present day, artificial intelligence is rapidly gaining momentum by making new discoveries as we keep moving forward. As Human beings, we are naturally curious by looking out for intelligent software that has the ability to ease our day to day activities and moreover, would enable us to understand speech and images much more accurately and efficiently. Artificial Intelligence found it easy to take on challenges that proved to be intellectually difficult for human beings, it was rather easily solved by implementing mathematical rules or certain high-speed calculations. It was observed that AI actually suffered where it found it really difficult to perform tasks which were quite easy for the humans but hard to describe formally for instance problems that we solve simply based on intuition like recognizing spoken words or faces in images or define right from wrong. In their book "Deep Learning" the authors (Goodfellow et.al, no date) have given a solution for the more intuitive problems. According to the authors, the solution was to enable the machines to learn from past experiences and understand the workings of the world based on the hierarchy of concepts wherein each concept defined through its relation to simpler concepts. Once the machines are able to gather the required knowledge from experience this approach would no longer need for human operators to formally specify all the knowledge the machine needs. A graphical representation of how these concepts are built on top of one another would result in the graph being deep with many layers. It is for this reason this approach is called **Deep Learning.** Deep learning is a part of the Machine learning family which is dependent on learning data representations, unlike task specific algorithms. A common myth that entails around deep learning is that the concept deep learning is relatively new which however is not the case. Even though the term deep learning was coined in the year 1986 by Detcher, whereas the first neural network was used in 1950's by Rosenbalt and the first multi layer neural network was published by Ivakhnenko and Lapa in 1965 (Foote, 2017). The reason it was overshadowed before was because the computational requirements for running deep learning were too high and at the same time more efficient methods like SVM were introduced which provided better results and in less time. Only recently with the emergence of new hardware enabling parallel computation for instance GPUs deep learning was re-discovered as computational time was no longer a major concern (RapidMiner Inc., 2016).

1.3 Current Research

From the above understanding, it will now be interesting to find out how effective deep learning techniques are in image classification. At the same time, considering the known accuracy and efficiency of Support Vector Machines, this study will explore existence of any differences and efficiency of performance between SVM and Deep learning techniques when compared with dog breeds with similar features and completely distinct dog breeds. Application of SVM model has been highly regarded by the author (LaRow et.al, 2016) for providing best results in dog breed classification. For this purpose, the researcher will be applying deep learning algorithm using H2O. According to the Gartner report in 2018, H2O is a popular open source platform in data science that enables the user to fit a vast array of potential models in order to uncover different patterns of the data (Gartner,2018). However, the present study uses H2O package as an extension in rapid miner.

CHAPTER-2 REVIEW OF LITERATURE

In this chapter the author has broken down the contents within into different sections. Section 2.1 sheds light on the concepts explored in the research and a brief understanding of how these concepts work. In section 2.2 the author explores and discusses in detail about the work done by previous researchers in the field and how their studies have helped the author to formulate his own study.

2.1: Background

A substantial increase in the volume of the images available online in the recent years, has managed to procure a lot of attention towards image classification in industry and academia (Lei Ma et.al, 2017). With digital images playing a quintessential role with respect to the multimedia content, the process for automation of the classification of images becomes an open research problem. A very general overview of what the term image classification means would be to simply assign labels to an input image. If we take any random image where we see an image as a whole, the computer just sees a grid of pixels. The main goal here is to basically see if the computer can actually look at the grid of pixels which contains an image for instance an elephant and actually assign a label to the image saying that this image is of an elephant. As simple as this process sounds it is indeed a daunting task for the computer to be able to predict. To get a better understanding of how this works we see figure 2.1.

Figure 2.1: Computer Vision (Sorokina, 2017)

To enable educators and researchers around the world who would like to work with images for academic research or educational purposes, ImageNet an easily accessible image dataset was introduced, a joint venture by Stanford, Princeton and Yahoo. The motto behind the development of ImageNet was considered to be “good research needs good resource” (Imagenet.org, no date). The ImageNet is an ongoing project with over 10 million images with over 21 thousand set of labels.

In the current research the author is rather interested in two specific algorithms which are Deep Learning and SVM.

Deep learning has been around as a part of the machine learning world since a long time. The actual lineage of deep learning according to the authors (Foote, 2017), can be tracked right back to 1943 where an attempt was made by Walter Pitts and Warren McCulloch to design a computer model which was based on the neural network of the human brain. However, “the earliest efforts in developing Deep Learning algorithms came from Alexey Grigoryevich Ivakhnenko (developed the Group Method of Data Handling) and Valentin Grigorʹevich Lapa (author of Cybernetics and Forecasting Techniques) in 1965” (Foote, 2017). Deep Learning is a part of machine learning which uses multi-layer neural networks for performing operations.

The deep learning architecture holds three main layers which would be the input layer, the hidden layer and the output layer. The hidden layer among the three layers is the most complex layer. A hidden layer is a layer that is placed in between the input and output layers. The hidden layer in itself has quite a few layers, out of which one of them would be the convolutional layer on which you apply a filter where you define the size of the filter with the help of which the layer creates an activation map. The function of an activation map is given below.

Activation Map = (N - F)/ Stride +1

Where,

N is the size of the convolutional layer

F being the size of the filter and

Stride is the step size that is used when moving from one region to another in a convolutional layer

Block :1 – Activation Map Formula

Each filter that is applied in the convolutional layer, one gets a different activation map. A convolutional layer would be a collection of the layer, the filter and the activation map.

Next part of the hidden layer which is often used in succession to the convolutional layer would be the pooling layer. The pooling layer also consists of a filter of a given size. Now there are basically two types of pooling operations that can be performed which are max pooling and average pooling. Max pooling is something that would basically take the maximum value from within the filter. Average pooling like the name suggests would sum up the values within the defined filter and divide it by the total number of values i.e. taking an average. Pooling layer is applied to reduce the size of the layer. If there is a convolutional layer before the pooling layer, pooling is applied to each of the activation map that is inside the convolutional layer.

The next type of layer would be the recurrent layer. This layer is basically used in case where the data we have is sequential which would not be of much use in the data of the present study. The last type of layer would be a dropout whose main functionality or rather purpose would be to reduce the complexity of the model. It is basically a method that is used for regularization.

Each layer inside the neural network consists of several neurons which are all connected together in such a way that they are able to communicate with each other. This neuron was devised with the hope that it would work in a similar way such as the neuron in a human brain would work. Here the neuron would try and calculate the weighted average of the values i.e. the input signal and an output signal that is transmitted by the connected neuron (Arora et al., 2015).

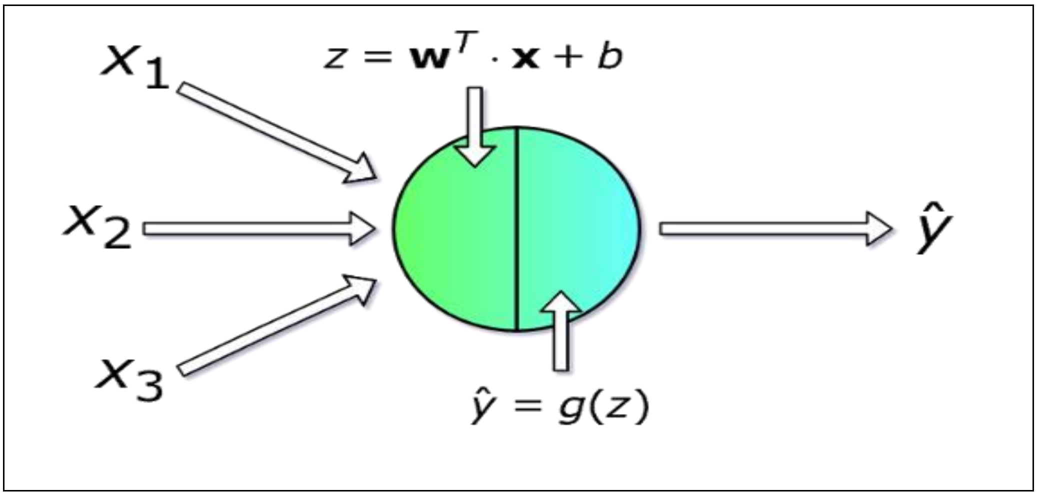


Figure 2.2: Working of a Neuron (Skalski, 2017)

In the paper submitted by the authors (Nokwon Jeong, Soosun Cho, 2017), the authors try image classification on the Instagram images where their focus hinges on evaluating the competitive power of deep learning for classification of real time social networking images. The authors (Nokwon Jeong, Soosun Cho, 2017) in their study also look at the performance of pre-existing CNN frameworks such as AlexNet and ResNet and how well they perform on the ImageNet dataset which showed outstanding capabilities.

The reason deep learning failed to make a mark even though it being discovered quite some time ago, was solely because, efficient models such as the support vector machine (SVM) were introduced which performed similar tasks but in a lot less time. The SVM algorithm falls under the supervised learning bracket which can be used for either classification or regression problems. However, it is considered to be quite popular for classification problems. SVM performs classification by dividing the classes with the help of a hyperplane. A hyperplane is the line or the decision boundary that basically segregates the two classes with the help of the support vectors which are the extremes in both the classes.

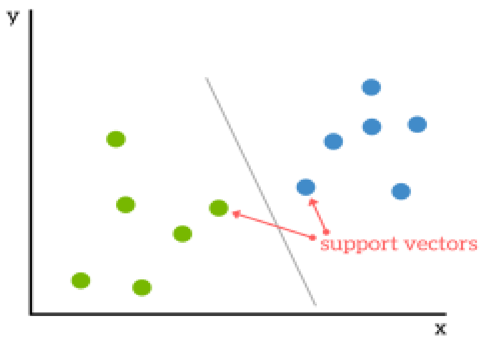


Figure 2.3: Classification using support vectors for SVM (Bambrick, 2016)

The SVM operates with the help of a kernel. “A function that takes the vectors as input vectors in the original space and returns a dot product of the vectors in the feature space is called a kernel function. With the help of the kernel function, we can apply the dot product between two vectors so that every point is mapped into a high dimension space via some transformation” (Augmented Startups, 2017). SVM provides us with the option to be able to choose our kernel. The few popular kernel types would be linear, radial, sigmoid and polynomial kernel. The support vector machines have numerous applications and is also considered to be a popular alternative to artificial neural networks (Foote, 2017). In a journal published by the Indian Academy of Sciences the authors (V. Karpagam and R. Rangarajan, 2013), talk about how the main focus of their study was to improve the accuracy for image retrieval and classification which they achieved by applying the SVM algorithm. The authors also discuss how their proposed system with SVM performed better as compared to the other methods.

2.2: Related Work

The researcher primarily bases his research on a fairly recent study, which was conducted by a group of scholars at Stanford University in the year 2016 wherein they attempted to identify the correct breed of the dog using image classification. The authors (LaRow et.al, 2016) took on the challenging task of image classification by treating it as fine-grained classification problem. Fine grained classification is performing classification operation on categories that are very similar or on categories that are divergent, but share a common part structure (Jacobs, no date). The authors (LaRow et.al, 2016) proceeded further into their research by using CNN on their data to identify the facial keypoints of the breeds of the dogs which then, they induced into a feature extraction system that enabled them to generate more meaningful features on the images. They believed it would prove to be instrumental in classification. A simple definition of the word feature would be a unique or distinctive quality of something and with respect to the images, it would simply be aspects that stand out such as a long nose or blue eyes. Upon achieving this feat of generating primary features, the authors then move on to the modelling stages wherein a variety of classification algorithms were applied out of which SVM, bag of words and k nearest neighbours were a few of them. According to the authors (LaRow et.al, 2016) SVM with a linear kernel proved to work best wherein with an accuracy of 90% the correct breed of dog came up among the top 10 predictions. In the future works of the study conducted by the authors (LaRow et.al, 2016), they talk about exploring the potential of neural networks even further, given the success of CNN for keypoint detection, they believe that using neural nets for dog breed prediction could lead to a higher success rate than what their current model had obtained. Keeping this in mind, the author decides to build upon on the bases of this research study all the while making modifications relevant to the purpose of the present study.

In the year 2012 the authors (Liu et.al, 2012) worked on a similar project that entailed the classification of different dog breeds for which they have used the part localization approach. Part localization, basically, is the ability wherein we are able to instinctively and precisely locate certain body parts in an image which could be an arm, an ear, a strand of hair which seems like an easy task for humans but is quite an excruciating task for the machine(Liang

Zhao, 2000). Although the results provided by the authors (Liu *et al.*, 2012) were impressive, considerable amount of work was put into the process of identifying the face of the dogs and its keypoints. In the current study the researcher deviates from putting too much work into dog face detection since in real world scenarios it is not always the case that the images will only contain the dog and nothing else. The model needs to be able to predict effectively, even if there are other factors included in the image.

Similar research in the field of visual classification was also done by a group of academicians (Girshick et.al, 2014) at University of California, Berkley where they worked on object detection. In their project they decided to use a pre- trained dataset which was trained using CNN on which they applied the models in order to attain image classification. The pretrained data set that the authors refer to in their project is the ImageNet data set. Using the ImageNet dataset the authors (Krizhevsky et.al, 2017) also take a shot at classification for which they use deep convolutional neural networks. Here the authors (Krizhevsky et.al, 2017) further elaborate how they were able to use deep neural networks on the ImageNet successfully and give a thorough explanation of the architecture used in their model, at the same time, also managing to reduce overfitting by introducing a recently developed regularization method named "dropout". They (Krizhevsky et.al, 2017) also manage to prove this theory by managing to enter a variant of their model for the ImageNet ILSVRC 2012 competition where they were able to achieve a top-5 error rate of 15.3% which was the best result as compared to others. However, concerns were also raised in the form of computational power where multiple GPU's had to be brought in to share the computational load.

Several other implementations in the field of image recognition have been carried out using deep neural networks at their very base for instance, the authors (Nguyen *et al.*, 2017) use deep CNN to be able to automate wildlife monitoring. Here the main focus behind automating the wildlife monitoring was to become more efficient by speeding up any research findings, establish systematic science based monitoring system and subsequent management decisions which the author (Nguyen *et al.*, 2017) believed would have the ability to leave a notable footprint on the world of trap image analysis which they wish to achieve by leveraging the advances that were being made in the field of deep learning by proposing a framework to build the automated animal identification system, were their experimental results managed to achieve an accuracy of 96.6%. With this they were able to successfully detect the images which contained an animal. On another similar project, authors (Trnovszky *et al.*, 2017) discuss how they use CNN to create an overall animal recognition system, here the authors talks about the challenges they faced while creating the systems and lists down different possibilities regarding ways to overcome the issue, at the same time they conducted a comparative study with previously well- known image recognition algorithms, to name a few the authors talk about support vector machines(SVM), principal component analysis (PCA), linear discriminant analysis(LDA) and Local Binary Pattern Histograms (LBPH), suggesting that the proposed model had better results on overall recognition.

On the basis of the articles and the journals that were reviewed, the current study takes on the task of multiclass classification on dog breeds using state of the art deep learning techniques. The above reviewed articles infer that time and again deep learning architectures have proved to be effective when it comes to complex tasks such as image classification, and what better way to test this theory by applying it on dog breed identification where the inter and intra breed variations are so diversified that it gives us the perfect platform to test the theory. The secondary approach carried out in this study would be to do a direct comparison at the very root level to analyse how well the deep learning and SVM models perform on a direct comparison with each other.

CHAPTER 3 – METHODOLOGY

3.1 Dataset

The data set used for this project is the Stanford Dogs Dataset, which was obtained from the Kaggle website. The dataset contains about 120 different dog breeds with atleast 150 images under each breed. The author had a dataset which accommodated a total of 20 thousand images of dogs to work with for this project.

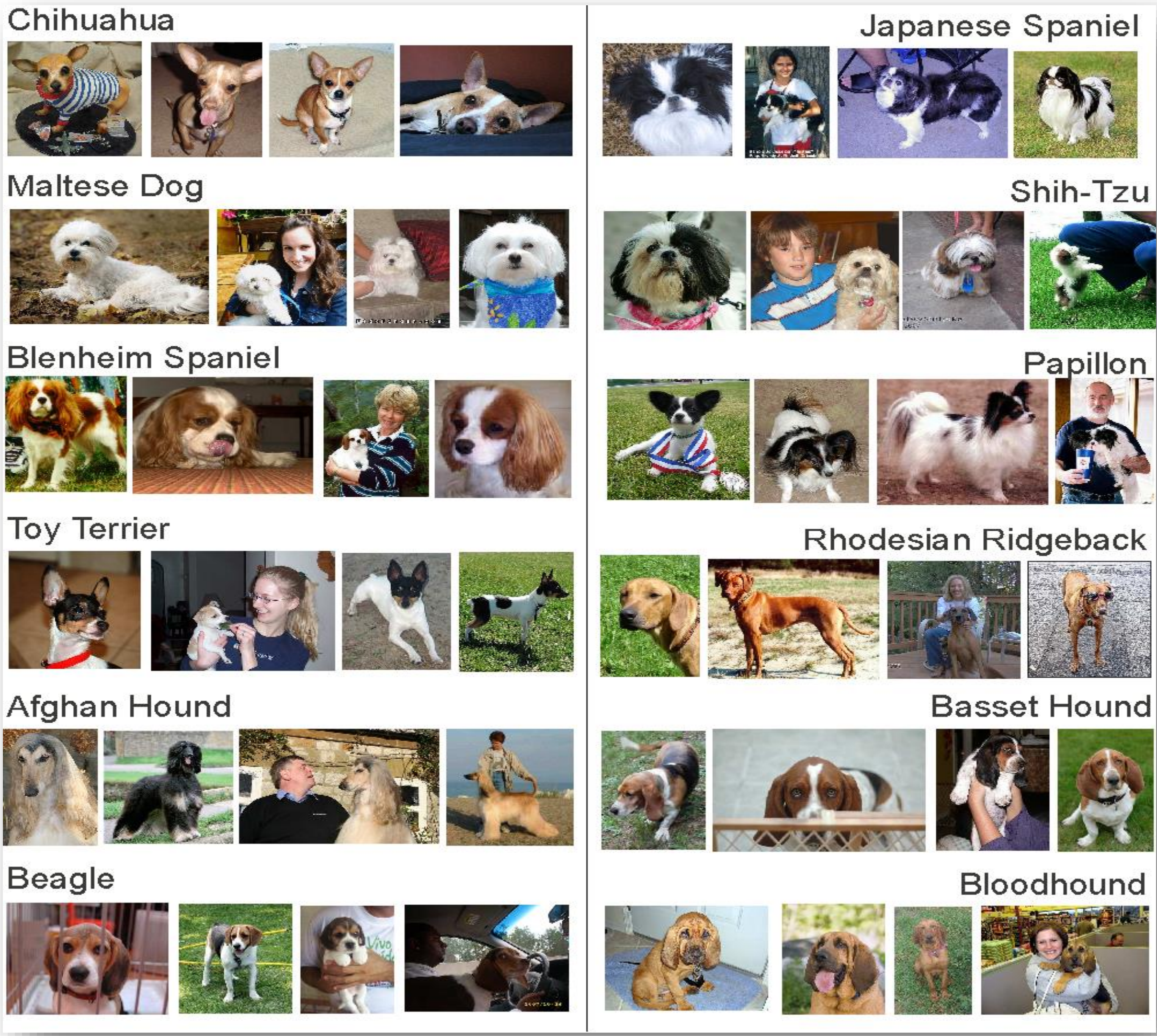


Figure 3.1: Snapshot of the Stanford Dog Data Set (Khosla *et al.*, 2012)

3.2 Data Pre-processing

The tool used in this project to work with the data specifically for pre-processing was R programming. R is an open source programming platform which is mainly used for statistical computing and graphics which was developed by John Chambers and colleagues at Bell laboratories which were formerly known as AT&T and now go by Lucent Technologies (*R: What is R?*, no date). R has a huge arsenal of statistical techniques such as linear and nonlinear modelling, classification, clustering and many more along with graphical techniques and is highly extensible (*R: What is R?*, no date).

Once the dataset was available, being an image dataset the most important or rather the most crucial step before applying any sort of algorithms, was to perform pre-processing on the data so as to convert all the information into a machine-readable format. To tackle this problem the author formulates a strategy as shown in the image below.

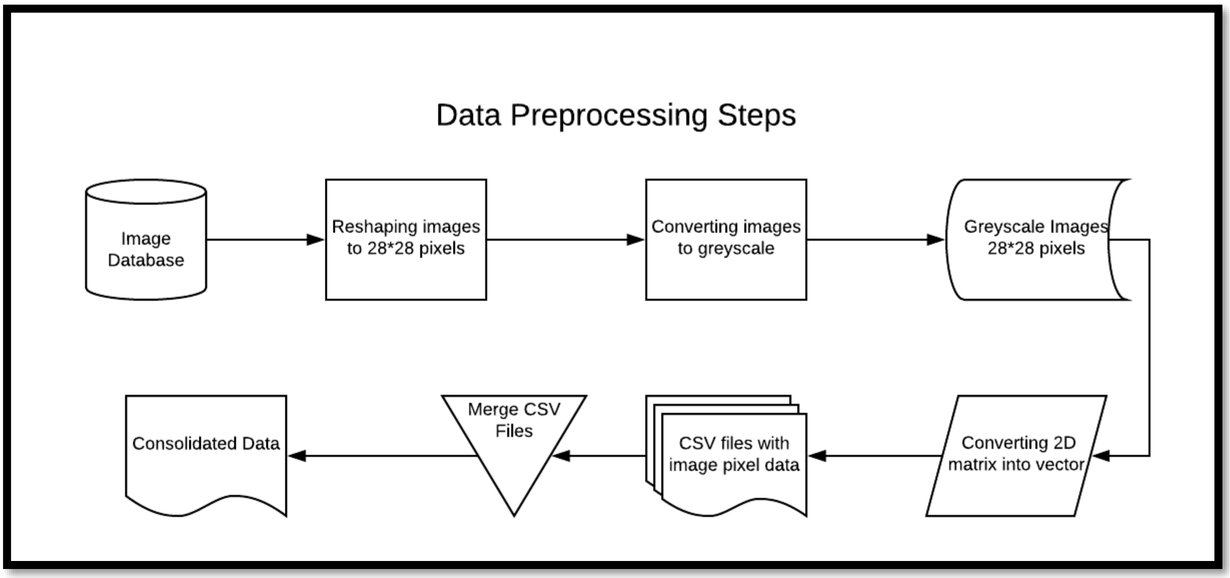


Figure 3.2: Data pre-processing stages

This was the initial challenge. There were 20 thousand images to be worked on and as expected all images had different sizes. To tackle this problem, the tool used was R programming. R programming had a nice package when it came to working with images, which was **EBImage**. EBImage is a predefined package that is provided in R to help work with images. With EBImage, the researcher had access to a lot of different tools that helped in the pre-processing by helping to manipulate the images as per the need. The package offers tools with the help of which changes can be made on a granular level, be it converting the image into pixel data or performing different operations such as linear scaling, transposing, applying various filters, combining different images, applying binary segmentation and many more. Scaling the images was necessary as it would give a standard size to work with and the data would be uniform. Using EBImage package in R, the researcher then decided to resize the images to 28*28 pixel. Once the researcher was successfully able to resize the images, the next thing to do was to convert the images into greyscale. With the EBImage package at the disposal, it made the task more doable. However, to get through 20 thousand images was a tedious and a time-consuming task even with the help of the R script.

On completing the pre-processing on the data, the images now needed to be stored in a format that would be suitable enough for the researcher to train a model on. The greyscale images which were now in a two- dimensional matrix format were then stored into a single dimension array or a vector, basically a flat file which would consist 784 attributes per image which was the product of converting the 28*28 grid into a single vector. Once the images were successfully converted into vectors with attributes the researcher then saves this data into a csv file which could later be directly called upon by the system and which was now in a machine readable format.

The author continues the same process for all the 120 breeds which again was time consuming but in the end managed to create a folder which contained csv files for all 120 dog breeds. Instead of calling all 120 files every time the researcher wanted to apply an algorithm, the researcher decides to merge all 120 csv files into one single excel file adding id and labels to each breed to make it easier to classify which saves a lot of time for the researcher instead of having to call 120 separate files at runtime.

Now that the data was in machine readable format the researcher then proceeds onto the next stage of the study which would be the modelling stage.

3.3 Modelling

Software Used

The modelling stages of this project were achieved by working on the tool known as RapidMiner. RapidMiner is basically a software platform in data science developed by RapidMiner Inc that acts as a one stop solution mainly catering to processes like data preparation, machine learning and model deployment all under one platform. According to the Gartner Magic Quadrant 2018 report RapidMiner is considered to be a leader for data science and Machine learning platforms (Idoine et al., 2018). Even though RapidMiner is not an open source platform, they do have a free edition which has a few limitations set in place namely with one Logical processor and 10,000 data rows. However, RapidMiner does provide a free license for one year with no major limitations set in place exclusively for students which was indeed a huge help for the researcher as it enabled him to work freely on the platform without any constraints. RapidMiner works to such a great extent that if need be it would also interact with other scripting languages in the backend and with perfect stability for instance python, R and java without any issues.

Data Upload

To start off with the process in RapidMiner, the researcher first created a local repository inside RapidMiner, where all the required datasets and the processes required for this study will be loaded into for easy accessibility. Once this is done, the next mode of action would be to import the datasets that are required from the system. The researcher now selects the import data option after which the system is pointed to the location in the local repository where the pre-processed image data is now saved in the form of an excel file. After successfully pointing to the location where the data is stored, the RapidMiner system reads in the data all the while checking for any issues that might require attention within the data for instance if there are any missing values or the data within a certain column/attribute is not readable. After conducting the basic checks and once the system has recorded no issues with the data, an option is also provided where the researcher can format or select the number of columns to read into. After which the researcher now points to the location of the repository, he had previously created for storing the data and processes required for the study, finally then the finish tab is selected, and the data is now saved within the repository. Depending on the size of the file, the system may take some time to load the data which is normal.

Defining the Process

On successfully importing the data, the researcher then moves on to define the process that would be used by him when trying to implement his approach. For the purpose of creating a process, RapidMiner provides a very user-friendly interface where, by the use of drag and drop function the required operators can be called upon to create the process. The initial step while creating the process is to get the data which was imported. Having the data in place, a set role operator is induced in the process. The set role operator is used to assign specific roles to any selected attribute. For instance, an attribute based on which the predictions will be made would be assigned the label operator. The defined role of an operator would determine how the other operators handle this attribute. The set role operator provides a set of options as the target role which can be selected according to preference. The next step in the process is implementing the cross-validation operator. The main function of the cross-validation operator is to estimate statistical performance of the learning model applied in the process. This feat is achieved by creating k number of folds of the data which is defining how many subsets the example data should be broken into and this also is the number of times the process would be carried out. The cross-validation operator is basically a nested operator i.e. it has two subprocesses within the operator which are the training sub process and the testing subprocess. Here the training sub process is used for training a model. In the testing subprocess, the trained model is applied on the unlabelled test example set to see how well the model performs. The next step in the process would be to implement a learning algorithm inside the training subprocess of the cross-validation operator. After implementing the learning algorithm, the next logical step is to move on to the testing subprocess section inside the cross-validation operator. Here, the researcher would add the apply model operator whose function is to apply the model which was trained using the learning algorithm on the test example set to see if the model can successfully perform the operation it was tasked to do. The final step in the process would be to add the performance operator in succession to the apply model. The performance operator basically evaluates the performance of the model on the test data and gives out an accuracy score along with a list of performance criteria values. These criteria for performance are automatically determined by RapidMiner in order to fit the learning task type.

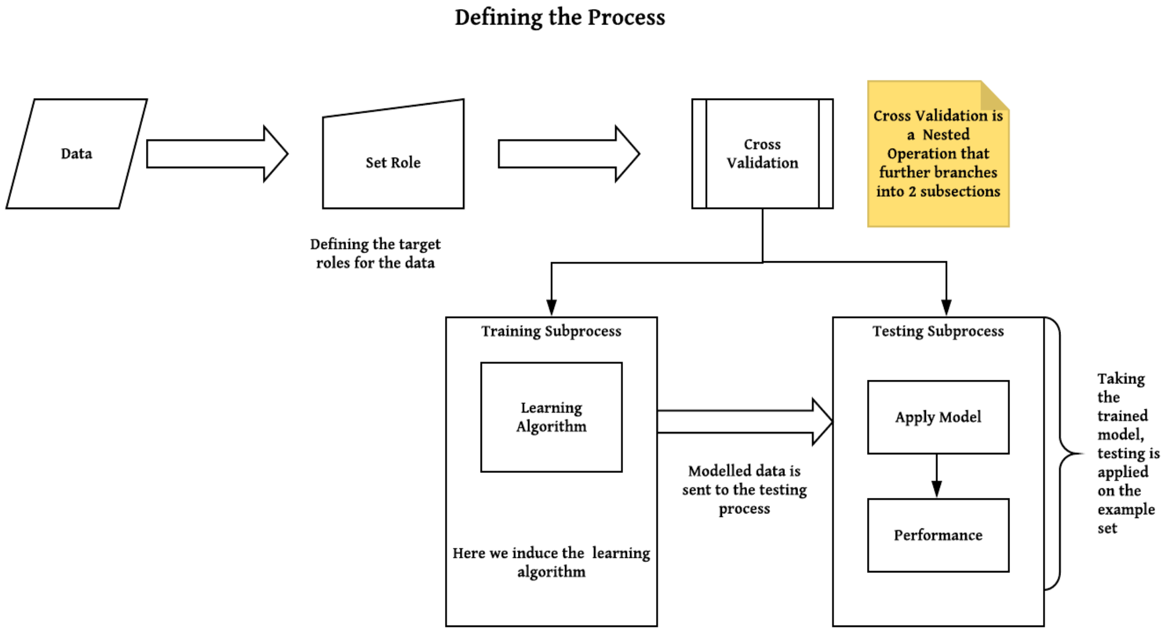


Figure 3.3: Process Architecture

3.3.1 Primary Approach

Deep Learning

In the primary approach the algorithm which the researcher has decided to implement after careful consideration is the Deep Learning algorithm using H20 version 3.8.2.6. “Deep Learning is based on a multi-layer feed-forward artificial neural network that is trained with stochastic gradient descent using back-propagation” (RapidMiner Inc, no date).

Now that the data has been imported and saved into the repository in RapidMiner, it is ready to be used and the entire process outline has been defined. The researcher then moves on to create the process for the specific algorithm. First, the data set required is brought into the panel. Next, the data is connected to the set role operator in the process. Here the researcher assigns the attribute ID the target role as “id”. This field consists of the name of the dog breed, since it is in character format the researcher defines this field as id and not label because, not defining it as id the system would fire an error saying classification is not possible since the field is not nominal. Then the label attribute in the data set is assigned the target role “label” since this is the field based on which model would be trained on. Not setting the target attribute also results into an error which says; “Input example must have a special attribute ‘label’”.

Then the researcher continues the connection from set role onto the cross-validation operator. Here, within the cross-validation operator the researcher sets the number of folds to 10 which tells the operator to subset the data into 10 folds and the number of iterations to run into. The researcher also defines to sample the data into stratified sampling type which ensures the classes are almost equally represented and contains roughly the same proportion of two classes. Other sampling types such as linear and shuffled do not really fit in with the data the researcher works on since the data is multinomial and it is vital that all classes are represented equally. After the researcher has defined the parameters for the cross- validation operator, he then moves on to the training subprocess inside the cross-validation operator, here he induces the deep learning algorithm which is the most crucial part for this process. The deep learning operator here accepts the training data inside the training section which the researcher connects to the input training port of learning algorithm. After connecting the algorithm to the process, the researcher then moves his attention to the parameters inside the deep learning operator. The initial parameter which is activation function is changed to "Rectifier" as it gives the best result and is the most followed approach. The activation function also known as the non-linearity function is what is used by the neurons inside the hidden layers. Next is the number and size of the hidden layers to be included into the data set. Since the researcher wants to avoid a case of overfitting, he set the number of hidden layers to three and size of the layers is changed to 100, 200, 100 respectively. Since the researcher is using H20's deep learning algorithm, the model takes care of all the additional sections inside the hidden layers all on its own during runtime. By this the researcher means that after we have defined the hidden layers, the model integrates the pooling layers and the fully connected layers all on its own as it deems necessary. The next parameter inside deep learning is the epochs function where the researcher sets the number of epochs to 10 which means the training data will be run on a loop 10 times. Now since the data is for multinomial classification problem, the researcher then changes the loss function to cross entropy which is only applicable for the classification problems and also the distribution function is changed to multinomial given that their multiple classes in the data is provided by the researcher. To make sure that the results produced by the sampled data set is consistent, the reproducible parameter is checked of by the researcher. Successful application of the required parameters marks the end for the training section of the process. Next attention is now focused on the testing subprocess where the apply model operator is included. This apply model operator is now given the model which was trained on the data in the model port. Next the data that was reserved for testing is given to the apply model through the unlabelled port. The model here tries to see if it enables to learn from the trained data and correctly perform classification function on the test data. The final thing the researcher does here is to pass the apply model results to the performance operator which is the last step in the process which after evaluating how well the model performed gives out accuracy scores along with a list of performance criteria. This with respect to the data provided was multinomial classification gives the overall accuracy and the kappa statistics for the model.

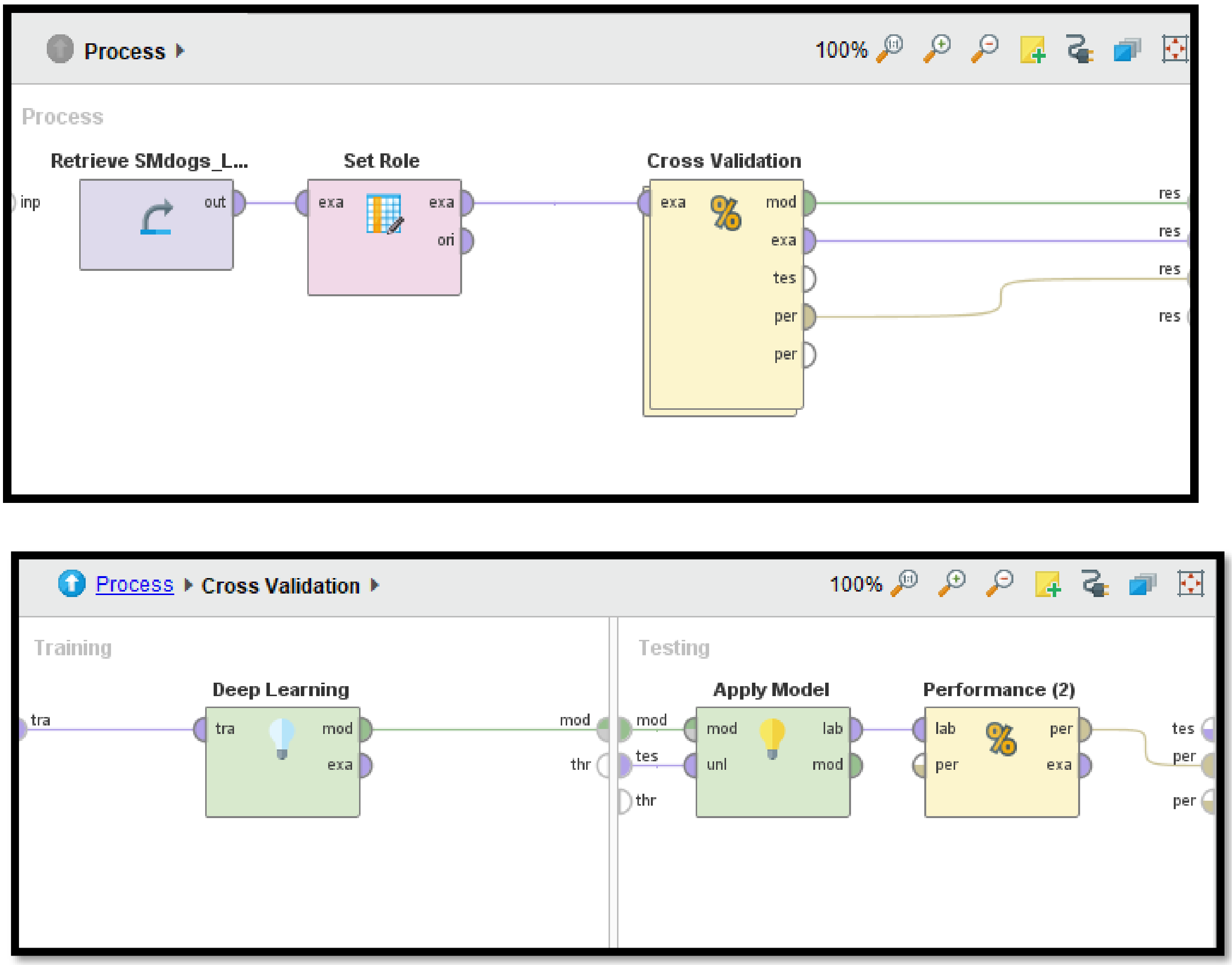


Figure 3.4: Deep Learning Process

3.3.2 Secondary Approach

In the second approach the researcher does a comparative study between the two learning algorithms i.e. deep learning and support vector machines. The reason the researcher has decided to take on this study is to simply gauge how well deep learning which is an advanced and continuously evolving technology performs against SVM which is a tried and tested method that was proven to be the best performer among the other classification models for image classification according to the author (LaRow et.al, 2016) on a very granular level. For this study, the researcher first takes on two classes of different dog breeds which share very similar features, making it quite difficult to tell them apart, which is a quite daunting task even for professionals. Secondly, two more classes of different dog breeds are considered this time the ones which share quite distinctive features as compared to each other, hoping to make it slightly better for the models to classify. The goal is to analyse when taken to the grass root level, how well do the models perform in either of the cases; Does it make a difference in the accuracy and does it help one model to outperform the other.

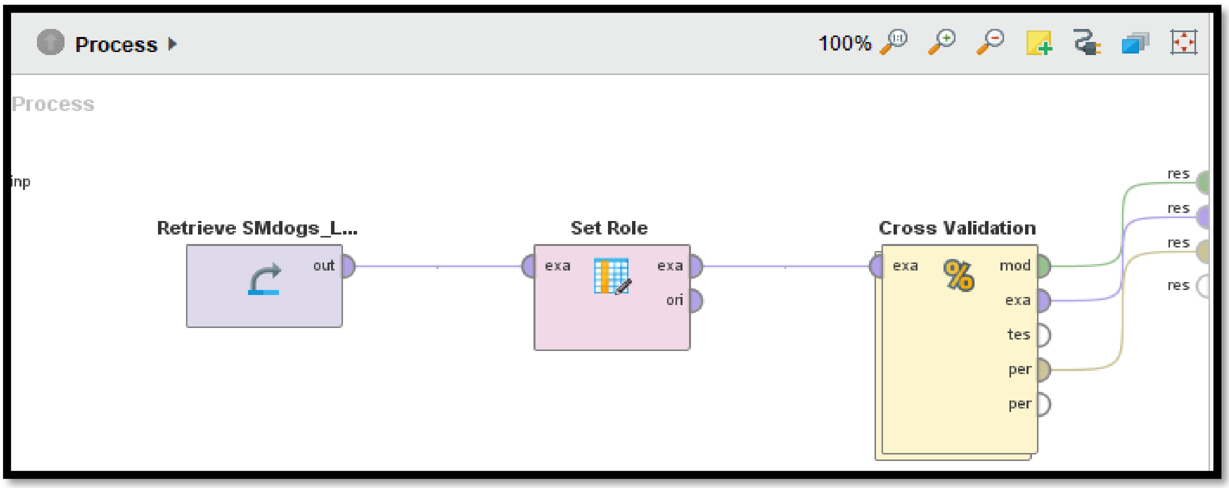
For the dogs sharing similar features, the researcher decides to narrow down on two breeds that are indeed quite difficult to distinguish between, which were Lhasa Apso and Shih Tzu. For the dogs that share quite distinctive features, when compared to each other making which made it easier for the researcher to tell them apart, Siberian Husky and Chihuahua were the two breeds that were selected.

After narrowing down on what breeds to use for this study, the researcher then consolidates the data for these breeds into two separate excel files naming them similar and dissimilar. The data was then imported into RapidMiner using the above-mentioned process since the process for the execution of the model was already defined above by the researcher. All that had to be done now was to implement the desired data into the pre-defined process and set the parameters as per the requirement of that data and then implement the learning algorithm into the process. After the parameters were set and the algorithm were implemented the researcher could just as well use the same process for similar and dissimilar dogs by only having to replace the main data being read into the process.

Process 1: Deep Learning

After reading in the desired data set for the process the researcher then follows similar procedure for implementing the process by making changes on them. After reading in the data, the next thing to be done is to pass the data through the set role operator. Here in the set role operator the target role for the ID attribute is set as ‘id’ whereas the target role which is assigned to the Label attribute as the name suggests is ‘label’. After defining the target roles for the data then the data is passed on to the cross-validation operator. Here the number of folds is set to 10 which is also the number of subsets i.e. k the data is broken into where among the k number of subsets, a single subset is retained for testing whereas the remaining k-1 number of subsets are used as the data to be trained on. After defining the number of folds the researcher moves his attention to the type of sampling that should be performed. Since the data is now binomial and the researcher wants to ensure that the sample subsets are split in such a way that both the classes are equally represented wherein, they contain roughly the same proportion of the two class labels, he defines the sampling type as stratified. Once the parameters for cross validation are set, attention is then turned towards the subprocesses within the operator where majority of the heavy lifting is done in the process. Since deep learning is being applied on the data first, the deep learning algorithm is implemented into the training subprocess where after making the necessary connections for the data flow, the author now zooms in on the parameter selection for the deep learning operator. The initial parameter which is activation function is changed to "Rectifier" as it gives the best result and is the most followed approach as also suggested by the RapidMiner community. Next is the number and size of the hidden layers to be included into the data set. Since the researcher wants to avoid a case of overfitting, he set the number of hidden layers to three and size of the layers is changed to 100, 100, 100 respectively. Since the researcher is using H20's deep learning algorithm, the model takes care of all the additional sections inside the hidden layers all on its own during runtime. By this the researcher means that after we have defined the hidden layers, the model integrates the pooling layers and the fully connected layers all on its own as it deems necessary. The next parameter inside deep learning is the epochs function where the researcher sets the number of epochs to 10 which means the training data will be run on a loop 10 times. Since the data is now suitable for a binomial classification problem the loss function is changed to cross entropy which is only applicable for the classification problems. On being a binomial classification problem, the distribution function parameter is now being changed to Bernoulli given that there are only two classes to classify from in the data provided by the researcher. To make sure that the results produced by the sampled data set is consistent, the reproducible parameter is checked of by the researcher. Successful application of the required parameters for the deep learning algorithm suggests that the preparations are now complete, and the data can be now trained. For the next step in the modelled data based on the training subsets is passed onto the testing subprocess where it is given in the form of input to the apply model operator which now tries to apply this trained model on the single subset of data which was retained for testing and thereby tries to see how effectively the model was able to classify the breed of dogs based on the test data . To evaluate the performance of the model on the test data set information in the form of labelled data is passed onto the performance operator which is the final piece of the puzzle which evaluates the performance of the model and provides output in the form of performance criteria. Which in the case where data was binomial would provide the researcher with a list of evaluations such as the overall accuracy, precision of the model, recall and also the AUC for general, optimistic and pessimistic.

Processes for Deep learning:



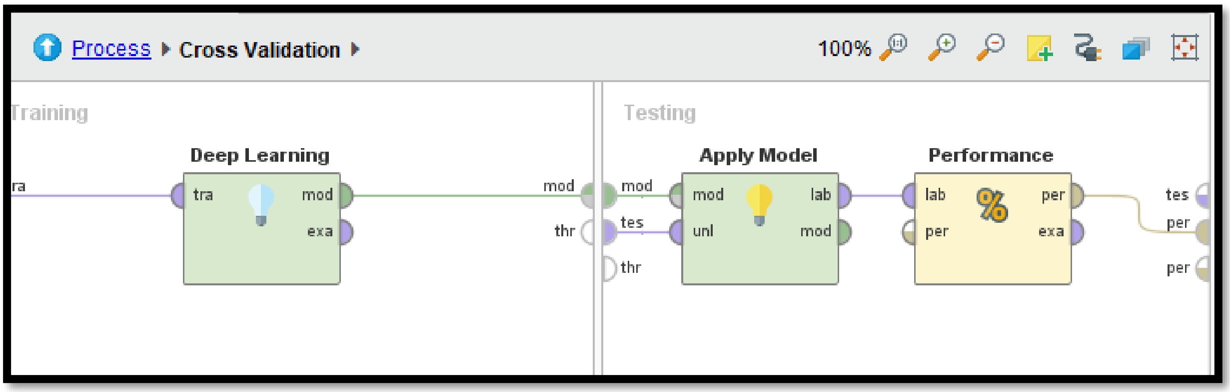
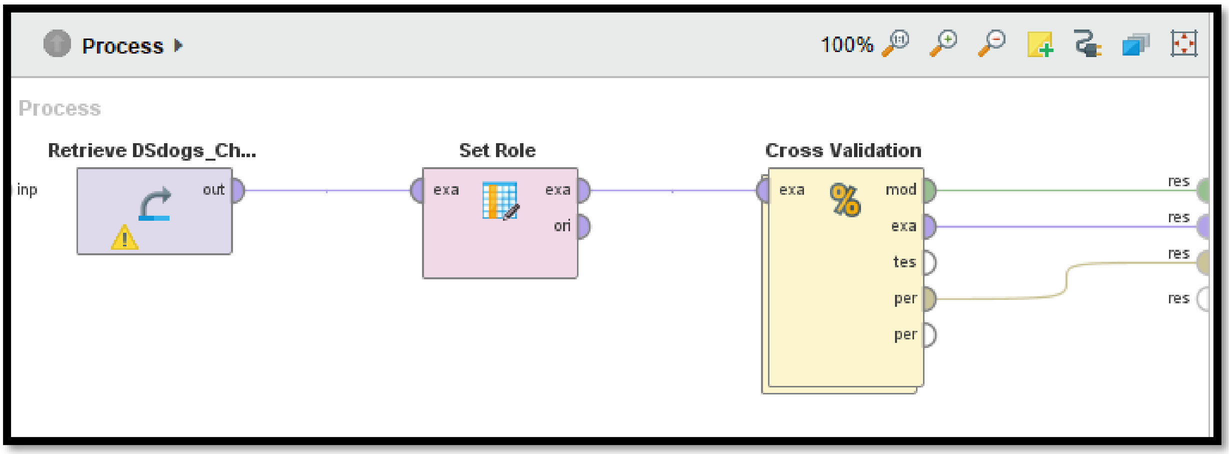


Figure 3.5: Dogs with similar features



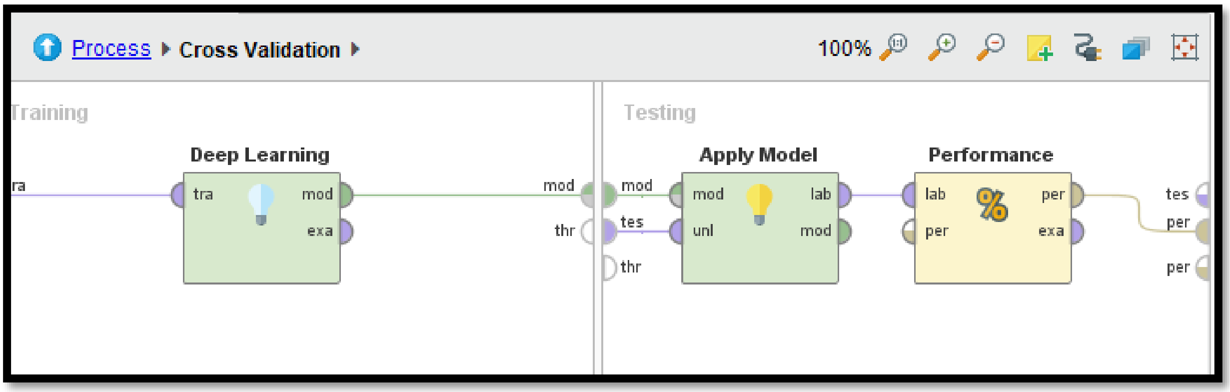


Figure 3.6: Dogs with distinctive features

Process 2: Support Vector Machine

This brings the researcher to second part of the comparison study on the data. Here the data for similar and dissimilar dog breeds would be same but the learning algorithm which will be applied would be the support vector machine algorithm. After reading in the desired data set which in this case could either be the data set for the similar dog data set or the dog set for dogs with distinctive features. The processes for both the data set would be the same once the correct algorithm has been implemented into the pre-defined process and the parameters specific to support vector machine has been successfully applied. After having selected the dataset when the data is being passed into the set role operator as a rule the special attributes in the dataset are assigned their specific target roles which helps the model perform accordingly. The ID attribute containing the name of the dog breed is assigned the target role 'id' and the attribute label is assigned the target role 'label' which holds the nominal value defining the breed of the dog. The connections are then made to the cross-validation operator. Here, similar to the previous process the parameters are set to the same values to ensure uniformity for the comparison study. Hence the value for the number of folds is set to 10 where 1 subset would be used for testing and k-1 i.e. 9 subsets are used to train the model. The sampling type that is similar to the previous process is set to stratified sampling to ensure equal representation of labels for both classes. On having defined the parameters for the cross validation, the next step which was the crucial step was to implement the learning algorithm which was support vector machine into the training subprocess of the validation operator. Like any other operator it was crucial on what parameters to select for the SVM operator. Hence the main focus of the researcher was to determine what kernel function to use. Now, RapidMiner does provide quite a few options for the kernel function namely linear, polynomial, radial, neural, annova, multiquadratic etc. Since the main purpose behind carrying out this process was to conduct a comparative study between deep learning and SVM. Hence, the researcher then decided to keep the kernel as linear which is considered to be the best performer according to the research conducted by the author (LaRow et.al, 2016), on image classification. After selecting the kernel as linear, the researcher then assigns the training data from the training subprocess as input to the SVM operator. Here, training is performed on the k-1 subsets of data and then passed on to the testing subprocess as a model. This model from the testing process is then given to the apply model operator which applies this trained model on the testing data which is passed to the apply model operator through the unlabelled data input port. After applying the model on the test data, the model tries to see how successful it was in performing the classification task on the test data. To evaluate the performance of the model on the test data set information in the form of labelled data is then passed along onto the performance operator which is the last step of the process which evaluates the performance of the model and provides output in the form of performance criteria.

Processes for SVM:

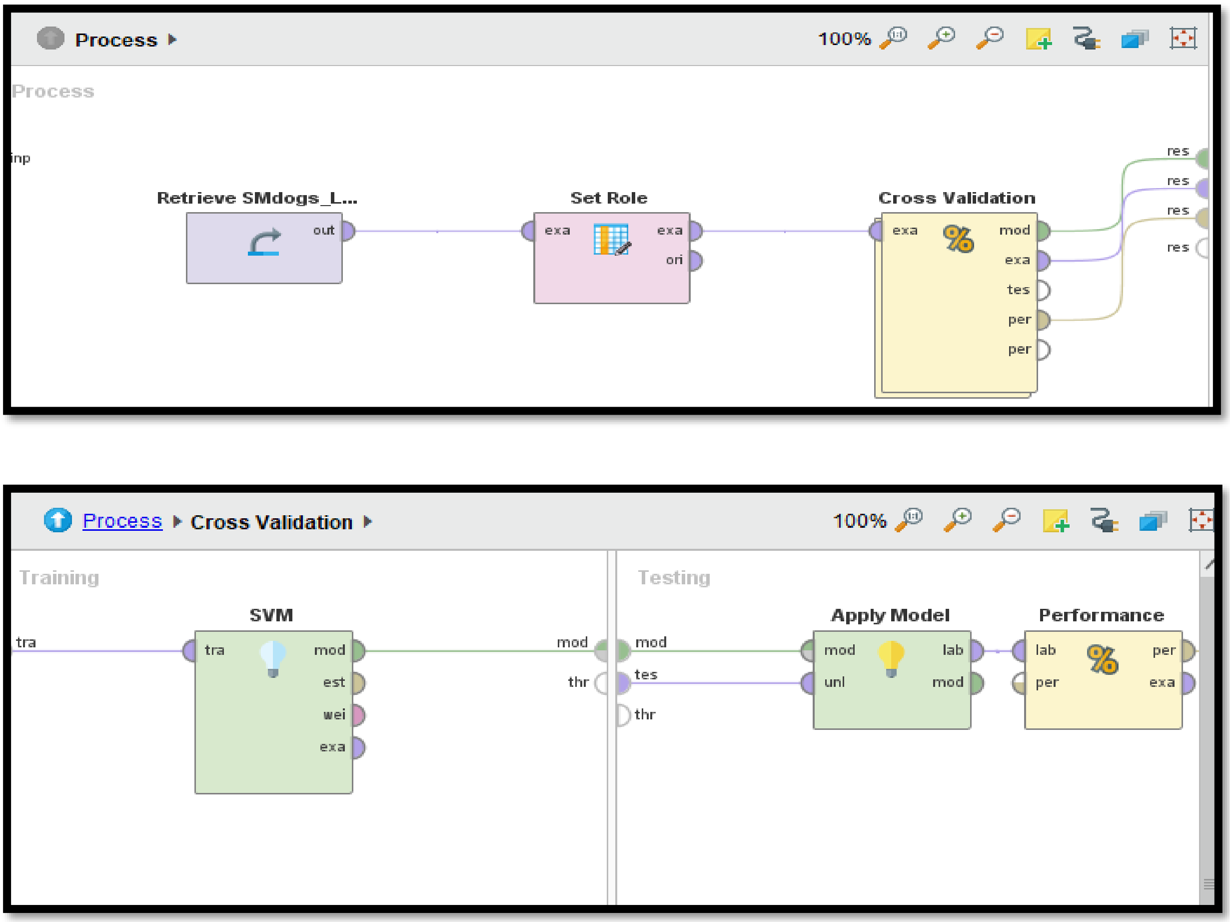


Figure 3.7: Dogs with similar features

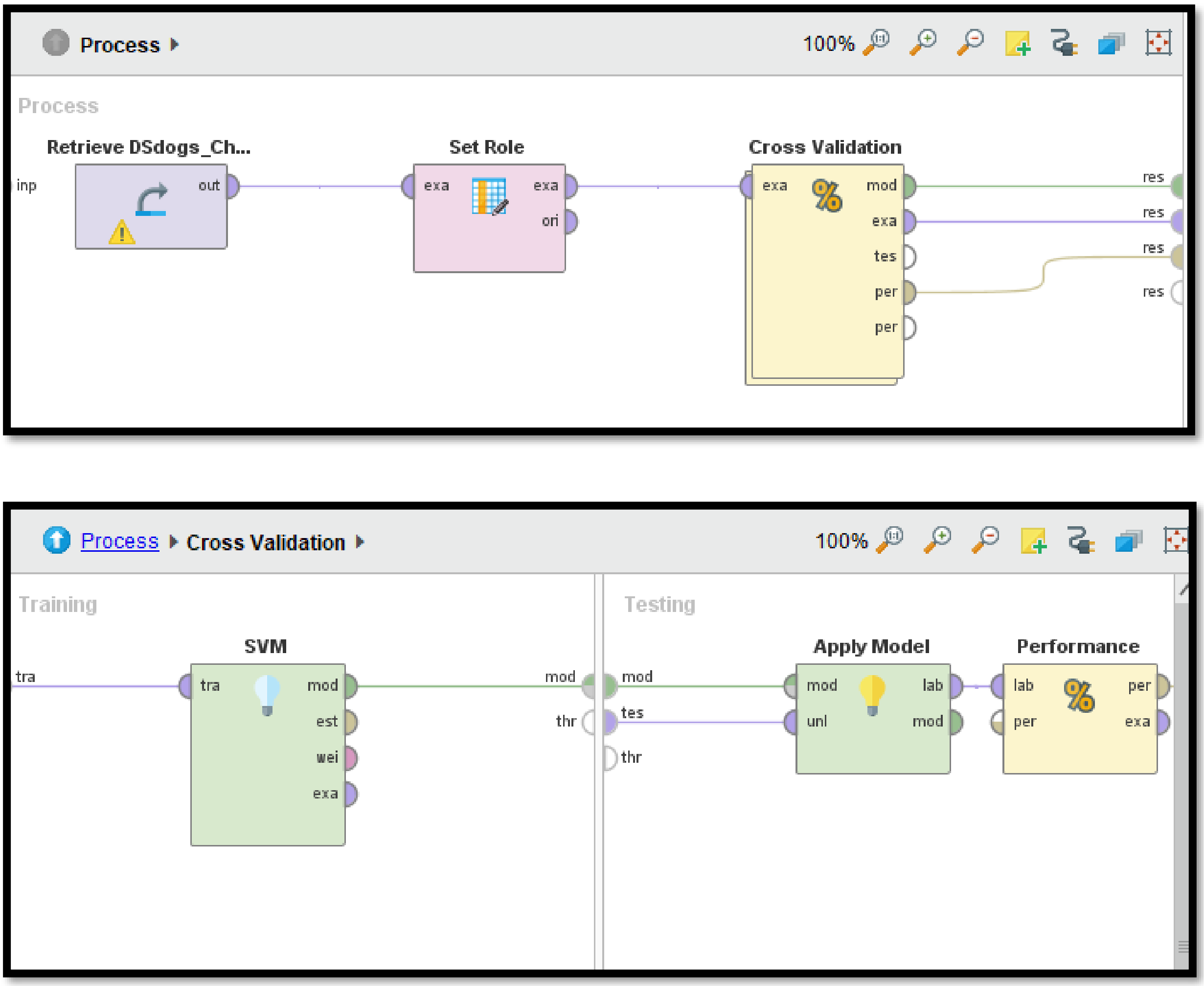


Figure 3.8: Dogs with distinctive features

CHAPTER – 4 ANALYSIS AND DISCUSSION

During the image pre- processing stages, all the images were first resized into 28*28 pixel size. This process of resizing the images were performed to ensure uniformity in the data and to make sure that no images were being stretched and to ensure that all the data could be used.

The next step the author performed was converting the images to greyscale. The reason the author decided against using the colour for the data processing is because the author strongly feels that using the colour as a feature especially for dog breed classification would result in confusing the model since there being huge intra species variation wherein for instance a Labradors most commonly are cream in colour, however, there also Labradors that are black and brown in colour. In such an instance if the model refuses to classify the dog as Labrador because it is black in colour then the accuracy of the model will be affected.

Figure 4.1: Labradors in different colours (Lovable Labradors, no date)

Hence, instead of focusing on the colour if the model solely focuses on the features, classification should become simpler.

On applying the deep learning H20 algorithm on the processed data set, we see that the results were not what was expected. The algorithm fails to make any significant mark and the results were fairly poor. On examining the results, the model was not able to perform multiclass classification efficiently. The author tried to make changes in the parameters and also increased the number of hidden layers at play within the architecture but none of this worked as the accuracy was still quite poor and was still not good enough to be presented. Another reason which the author believed could bring about a change in the results was to alter the size of the images which was a possibility that could give the model more features to work with and in turn help increase the overall accuracy of the model. So, in order to test the theory, the author took a sample of the images from the dogs dataset and performed data processing on both but this time creating two files out of which one would hold the images of size 28 * 28 and the other would hold images of size 64 * 64 and by pushing both the files into deep learning model. Thereafter, what the author found was that the accuracy of the model actually reduced for the 64* 64 as compared to the 28 * 28 proving the theory of improving size for slightly better accuracy void. After exhausting the options to improve the accuracy of the model, the author had to concur that the current approach was not working.

For the second approach, where we are comparing the deep learning model to the SVM model at the very grass route level to see which of the two models perform better when given a case of binomial classification. For this purpose, two sets of dogs were taken where one set contained a couple of dog breeds with similar features and another set which contained a couple of dog breeds which shared quite distinctive features.

On comparing both the models on the two different sets of data, it was found that both the models do perform slightly better when it came to the dataset with the dogs with distinctive features as compared to the dog with similar features.

Results for Dogs with distinctive features:

accuracy: 62.77% +/- 8.99% (micro average: 62.73%)

	true 0	true 95	class precision
pred. 0	114	82	58.16%
pred. 95	38	88	69.84%
class recall	75.00%	51.76%	

Figure 4.2: Deep learning

accuracy: 61.23% +/- 7.73% (micro average: 61.18%)

	true 0	true 95	class precision
pred. 0	96	69	58.18%
pred. 95	56	101	64.33%
class recall	63.16%	59.41%	

Figure 4.3 Support Vector Machine (SVM)

Here when examining the results on the dataset with dogs with distinctive dog features the author infer that the deep learning model performs slightly better than the SVM model showing that the deep learning model has the potential to outperform the SVM model.

Results for Dogs with similar features:

accuracy: 60.50% +/- 7.40% (micro average: 60.50%)

	true 3	true 52	class precision
pred. 3	152	96 true 52	61.29%
pred. 52	62	90	59.21%
class recall	71.03%	48.39%	

Figure 4.4: Support Vector Machine (SVM)

accuracy: 55.25% +/- 6.47% (micro average: 55.25%)

	true 3	true 52	class precision
pred. 3	141	106	57.09%
pred. 52	73	80	52.29%
class recall	65.89%	43.01%	

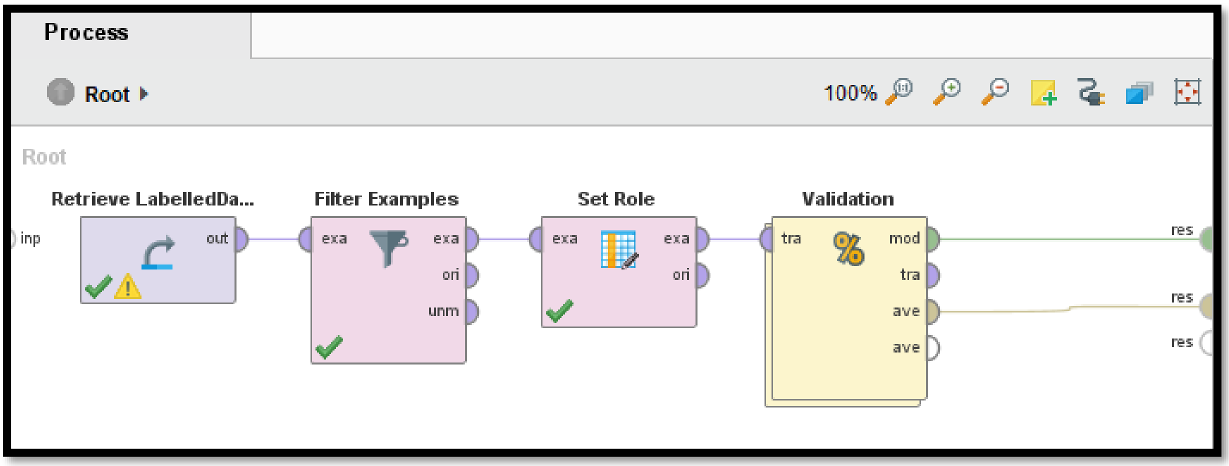
Figure 4.5: Deep Learning

Interpreting the results for the data set for the dogs with similar features, performance of the deep learning model goes down in the present study. From this section, the author infers that

SVM results are better when compared to deep learning, justifying why over the years why SVM gained its popularity for classification problems.

Overall, the research shows that both the models work well when given a binomial classification problem which point to a thought that if there would be a way to leverage this into improving the overall accuracy of the model when applied on the entire dataset. Keeping that in mind the researcher proposes an alternative approach which could be used to increase the accuracy which would be to transform the multiclass classification to polynomial by binomial classification problem. The polynomial by binomial operator is a nested operator that consists of another process inside the main process which would hold the binomial classification learner. The way this operator functions is by using a binary classifier which basically produces classification models that are binary for all the different classes in your data and then performs an aggregate of all the responses for the binomial models and then presents the results for the polynomial label.

The present study also tried implementing this approach by applying the technique on the current data, the researcher was aware that it would take a lot of time and computational power to train the data using that approach which we estimated we would still be able to manage. However, the computational power required to carry out this operation was simply too great, and the system fired a notification which read the task could not be performed.



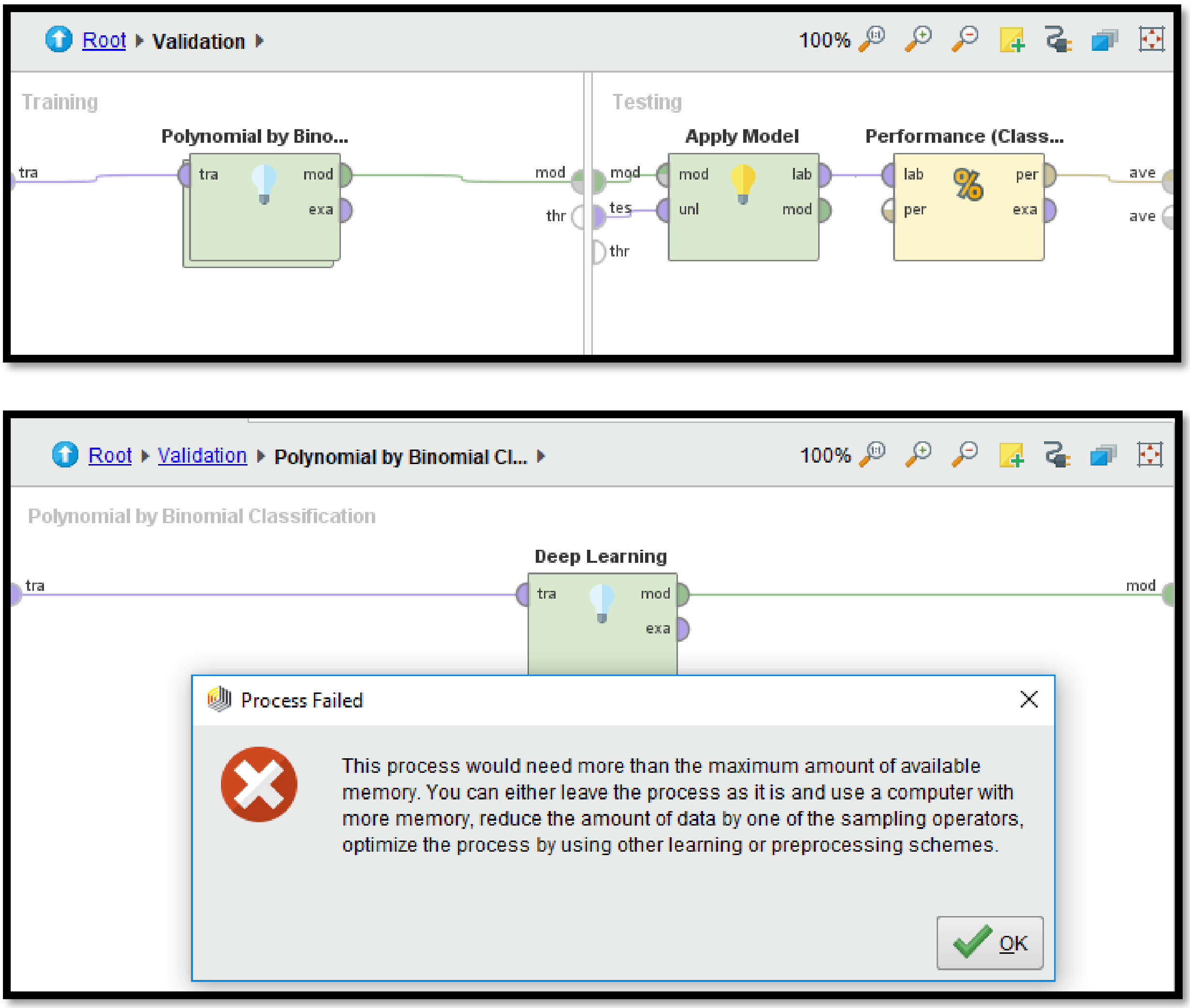


Figure 4.6: Process for polynomial by binomial

The researcher even tried cutting down the total number of classes in the dataset by applying filters on the dataset to see if it could manage to get a prototype working. For this purpose, the researcher reduced the data set to so far as 30 classes out of the total 120 and yet it proved to be more than what the system could handle. Since the researcher was unable to perform this operation, putting it up as a proposed solution or rather future works that could be explored further on the existing system was taken into consideration.

During the course of study, the researcher also looked at a technique introduced in machine learning called ensemble models. Ensemble models are a platform that combine the results obtained from different models which in turn helps to obtain better predictive performance. The present study also applies ensemble techniques for the sample data set to see how well this technique would do. Here, the researcher used the vote operator which held the two learning algorithms i.e. deep learning and SVM, inside as sub processes on the sample data wherein the operator combined the results and gave the best accuracy score.

Table View   Plot View

accuracy: 63.05% +/- 5.59% (micro average: 63.04%)

	true 0	true 95	class precision
pred. 0	121	88	57.89%
pred. 95	31	82	72.57%
class recall	79.61%	48.24%	

Figure 4.7: Results with vote operator

Here, it is observed that the results in general are better than what both the models i.e. deep learning and SVM performed individually [Referring to figure 4.2 and 4.3 - the results obtained for dissimilar dog set for deep learning and SVM]. The reason why the researcher did not apply this method on the entire data set was because inducing the entire data would lead to a more complex architecture were nested ensemble models would be used because SVM would not perform unless the polynomial to binomial transformation on the data since SVM can only work on binary classification. As seen in the above process [Referring to the polynomial to binomial process] there was not enough computational power to support the process. Leveraging the fact that voting works better on the models, this approach would work better on the models that can be taken for future considerations.

CHAPTER – 5 CONCLUSION

The main takeaway from working intrinsically on this project for the researcher was to be able to get a better understanding of the workings of the deep learning model, and also to acknowledge the fact on how effective the RapidMiner framework is to tackle machine learning problems all under one roof while the backend interacts with different programming languages with ease. With respect to the main research question being based on whether deep learning would be able to improve the accuracy for dog breed classification, the author infers that when it came down to a multiclass classification problem even the deep learning algorithm faced certain hinderances in classifying the dog breeds. The reason being, the model was not able to generate and store the patterns necessary during training which would help the algorithm to distinguish the breed of the dog from amongst the other 119 breeds included in the dataset. On examining the results for this approach, the author comes to an understanding about two possible reasons that could have played a role in the poor performance of the model. The first reason being the algorithm had insufficient data to work with, we know that deep learning model performs feature extraction all on its own but this feat is achieved by the model by training on an immensely large data which was a factor that was lacking in our project. The second reason the author realised affected the performance was since the dataset was not massive the model could have performed slightly better with the data being pre-trained before being fed into the model.

In the second approach, a comparative study was implemented between deep learning, and SVM, both of which are considered as two of the best learning algorithms for image classification according to previous research. To put this theory to test, the datasets were broken down into two different sets of binomial classification problem. The two sets were separated by selecting dog breeds that looked quite similar to each other and dog breeds that had obvious differences on how they appeared. On carefully evaluating the results for both the cases the author sees that SVM model gave better results where the dogs were similar looking as compared to deep learning, justifying why the model achieved its status of being among the best classifiers. On the other hand, the dataset where the dogs had clearly distinctive features saw a better performance from the deep learning model affirming the researcher's intuition that the model had the potential given the suggested approach was followed.

5.1 Future works:

Delving into the performance of the deep learning model for binomial classification, the author affirms his suspicion that the model was infact successful in creating the required patterns to classify the dog breeds when compared one on one, it only faltered when it was given 120 breeds to classify from.

There are several ways in which this study can be taken forward. One aspect would be to leverage on the fact that deep learning algorithm performs well for binomial classification. By keeping this in mind in the proposed future work, the author suggests to perform polynomial by binomial ensemble technique on the dataset where the model would simply convert the labels into binomial groups where one class would be compared to all the other classes and the resulting score would be aggregated and presented. Talking about the ensemble techniques there was also the voting operation that the author conducted which showed great promise [Referring to Figure 4.7]. Building on top of that concept another interesting approach in the future could be to find the perfect blend between supervised and unsupervised learning into the same architecture and see how well that performs. To further conceptualize this chain of thought the researcher suggests a pseudo approach where clustering is applied on the data first, once the data is split into different bins we apply the voting operator on it with some of the best classifiers (which in our case would be deep learning and SVM), but with the advancements made in the field everyday there could be another approach to be tried on, which of course can be included inside the voting operator. The voting operator would simply perform an aggregate/majority operation and give us an improved result. This is only a suggested prototype and there may very well be issues to work out on but it could lead to something promising [demo model assembled in Appendix C]. However, a very important and perhaps the most important aspect going forward is to have additional dedicated GPU support for your processing which always poses a problem when it comes to image classification tasks which really takes a toll on your computer performance.

5.2 Business Applications:

Once refined and perfected, the proposed system can have a whole spectrum of uses, both as a service and as multiple business applications. Though it may be a niche group, there definitely is a sect of people who dedicate themselves for welfare of the pack. This application aims to help these people in whatever little way possible to further their mission, one match at a time. Keeping this in mind, one area where this business implementation would make the most difference could be in animal shelters where no longer would any dog be abandoned because of wrongful classification.

Another interesting business application for dog breed identification could be where a surveillance system is created which could help locate lost dogs, where the system would be specifically programmed with a task to look for a specific breed of dog.

Apart from the above, this application will also prove to be of immense help to those individuals who come forward to help by adopting orphaned or abandoned dogs. It is known that a dog's requirements, behaviour and needs differ based on the family of breed it belongs to. Barring a few specific sets of dog lovers who know exactly what dog they are looking for, often there are people who take a pup home just because they had an instant connect. In such cases knowing what lineage their companion belongs too would make their journey easier where they can specifically cater to the needs of their pup specific to their breeds and they would also find it helpful to identify breed specific diseases in dogs like Labradors often being plagued with ACL tears and Dachshunds falling prey to slipped discs allowing them to immediately seek the correct medical attention based on the diagnosis.

REFERENCES

1. Arora, A., Candel, A., Lanford, J., LeDell, E. and Parmar, V. (2015). Deep Learning with H2O. 3rd ed. [ebook] Available at: https://h2o-release.s3.amazonaws.com/h2o/master/3190/docs-website/h2o-docs/booklets/DeepLearning_Vignette.pdf
2. Augmented Startups (2017). Support Vector Machine (SVM) - Fun and Easy Machine Learning. [video] Available at: https://www.youtube.com/watch?v=Y6RRHw9uN9o
3. Bambrick, N. (2016). *Support Vector Machines: A Simple Explanation*. [online] Kdnuggets.com. Available at: https://www.kdnuggets.com/2016/07/support-vector-machines-simple-explanation.html [Accessed 22 Nov. 2018].
4. Foote, K. (2017). A Brief History of Deep Learning - DATAVERSITY. [online] DATAVERSITY. Available at: http://www.dataversity.net/brief-history-deep-learning/
5. Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep learning in the broader field of AI. [image] Available at: (HTTP://WWW.DEEPLEARNINGBOOK.ORG)
6. Greenwood, A. (2015). *New Study Proves That It's Extremely Difficult To Visually Identify Pit Bulls*. [online] barkpost.com. Available at: https://barkpost.com/good/study-proves-difficult-visually-identify-pit-bulls/ (Accessed: 10 September 2018).
7. He, K. *et al.* (2015) 'Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification', *arXiv:1502.01852 [cs]*. Available at: http://arxiv.org/abs/1502.01852 (Accessed: 25 November 2018).
8. Idoine, C., Krensky, P., Brethenoux, E., Hare, J., Sicular, S. and Vashisth, S. (2018). Magic Quadrant for Data Science and Machine Learning Platforms. Gartner. [online] Available at: https://RapidMiner.com/resource/read-gartner-magic-quadrant-data-science-platforms/ [Accessed: 13 Oct. 2018].
9. Jacobs, D. (no date) 'Fine-Grained Classification', p. 30. Available at: http://www.cs.umd.edu/~djacobs/CMSC733/FineGrainedClassification.pdf [Accessed 28 Oct. 2018].
10. Khosla, A. *et al.* (2012) 'Novel Dataset for Fine-Grained Image Categorization: Stanford Dogs', p. 2.
11. Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2017) 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM*, 60(6), pp. 84–90. doi: 10.1145/3065386.
12. Liu, J. *et al.* (2012) 'Dog Breed Classification Using Part Localization', in Fitzgibbon, A. et al. (eds) *Computer Vision – ECCV 2012*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 172–185. doi: 10.1007/978-3-642-33718-5_13.

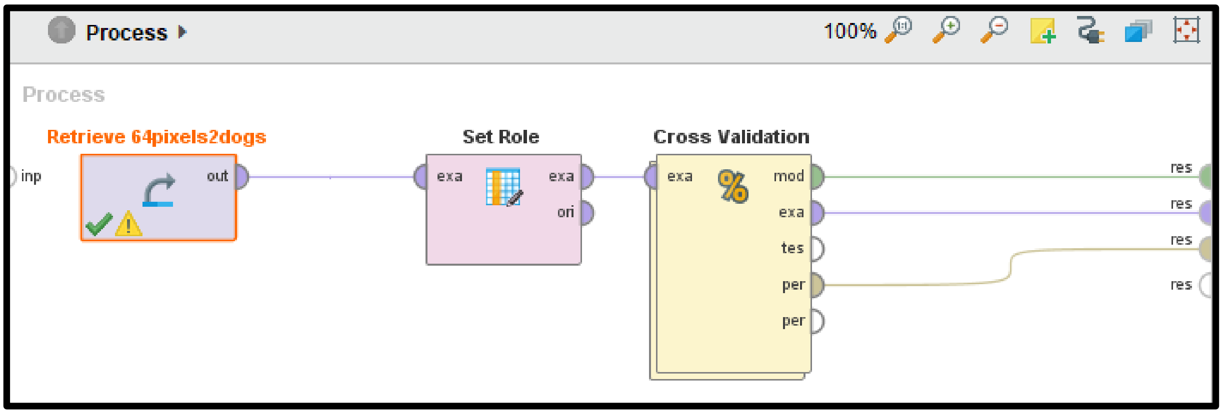
13. Lovable Labradors (n.d.). [image] Available at: http://www.lovablelabradors.com/labrador-information/how-big-do-labrador-retrievers-get-25/
14. Nokwon Jeong, Soosun Cho (2017) 'Instagram image classification with Deep Learning'.
15. R. Olson, K. and K. Levy, J. (2012). *Incorrect Breed Identification*. [online] Maddie's Fund. Available at: https://www.maddiesfund.org/incorrect-breed-identification.htm [Accessed 12 Nov. 2018].
16. *R: What is R?* (no date). Available at: https://www.r-project.org/about.html (Accessed: 4 December 2018).
17. RapidMiner Inc. (2016). Deep Learning: The Promise Behind the Hype. [image] Available at: https://www.youtube.com/watch?v=JLuekxdQkMM
18. Ross Girshick, Jeff Donahue, Trevor Darrell, JitendraMalik (2014) 'R-CNN for Object Detection', p. 34. [Accessed: 2 Dec 2018]
19. Skalski, P. (2017) *Deep Dive into Math Behind Deep Networks*, *Towards Data Science*. Available at: https://towardsdatascience.com/https-medium-com-piotr-skalski92-deep-dive-into-deep-networks-math-17660bc376ba (Accessed: 8 December 2018).
20. Sorokina, K. (2017). [image] Available at: https://medium.com/@ksusorokina/image-classification-with-convolutional-neural-networks-496815db12a8
21. The Startup (2018). [image] Available at: https://medium.com/swlh/ill-tell-you-why-deep-learning-is-so-popular-and-in-demand-5aca72628780deepl Learnng
22. V. Karpagam and R. Rangarajan (2013) 'Improved content-based classification and retrieval of images using support vector machine', *Current Science*, (9), p. 1267.
23. Yadav, I. S. (2018) 'EVOLUTIONARY REVOLUTION OF DEEP LEARNING MODEL IN DIABETIC RETINOPATHY', 7(11), p. 3.
24. Zhang, N. *et al.* (2013) 'Deformable Part Descriptors for Fine-Grained Recognition and Attribute Prediction', in *2013 IEEE International Conference on Computer Vision*. *2013 IEEE International Conference on Computer Vision (ICCV)*, Sydney, Australia: IEEE, pp. 729–736. doi: 10.1109/ICCV.2013.96.

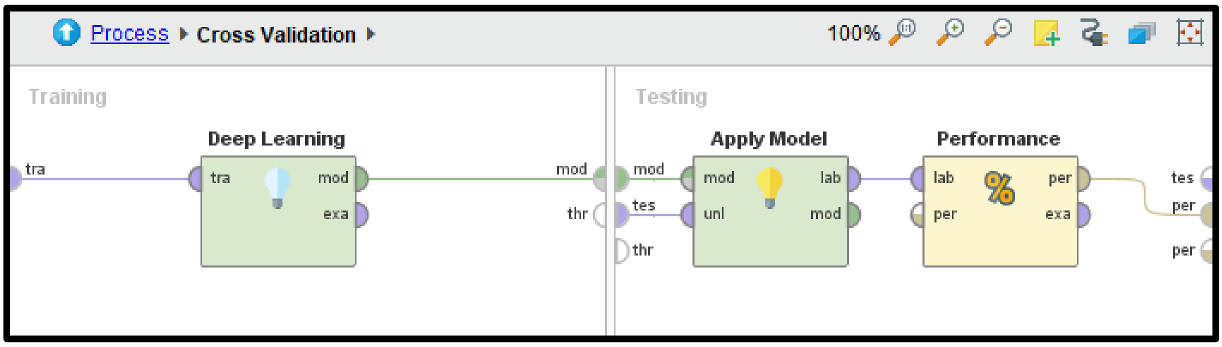
Appendix A

Given below is the process and results conducted to check if by increasing the size of the image does the accuracy of the model increase.

When we converted the images to 64* 64 size

Process





Result:

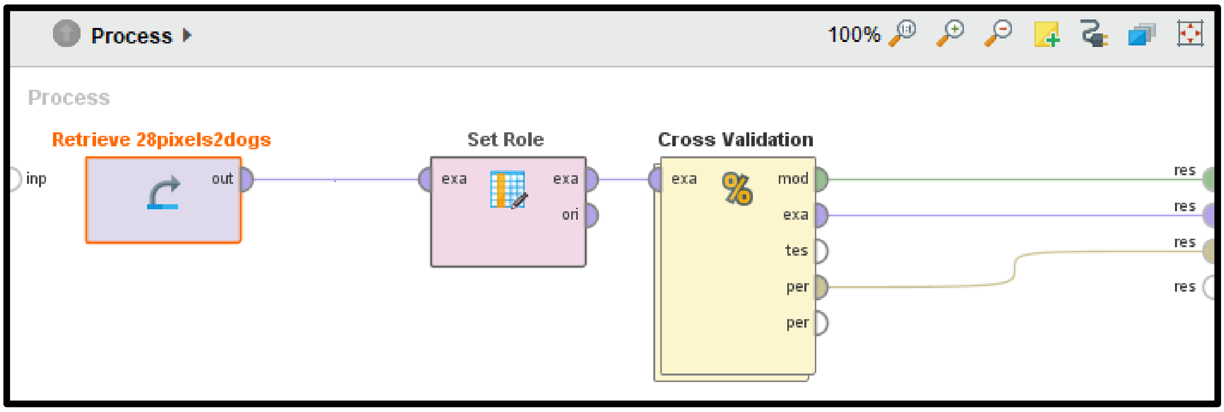
Table View
Plot View

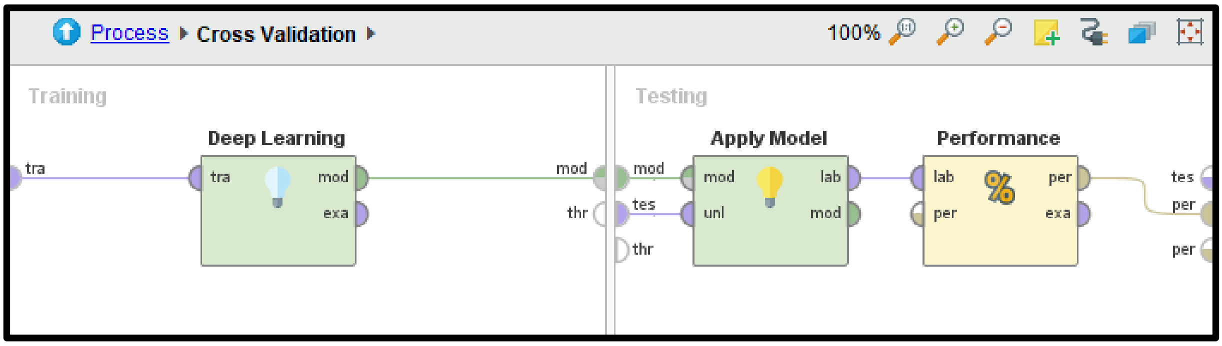
accuracy: 70.81% +/- 5.42% (micro average: 70.79%)

	true 2	true 1	class precision
pred. 2	216	82	72.48%
pred. 1	36	70	66.04%
class recall	85.71%	46.05%	

When we converted the images to 28 * 28 size

Process





Result

Table View Plot View

accuracy: 70.28% +/- 6.93% (micro average: 70.30%)

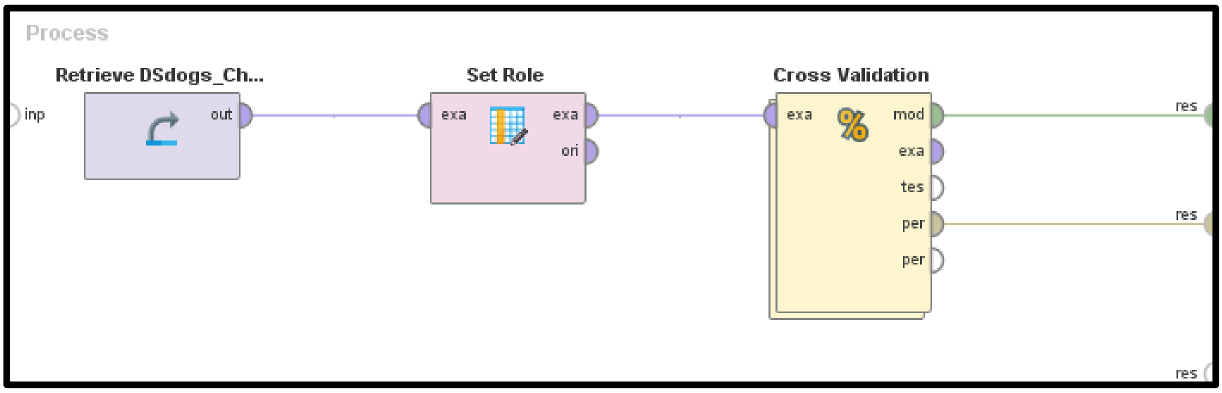
	true 0	true 2	class precision
pred. 0	71	39	64.55%
pred. 2	81	213	72.45%
class recall	46.71%	84.52%	

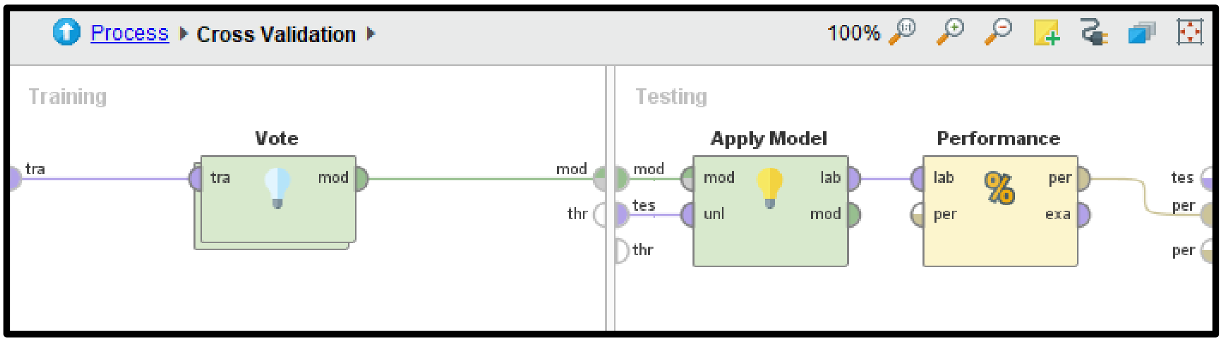
When we compare both the results we see that there is not much difference in the accuracy which indicates the size was not an issue.

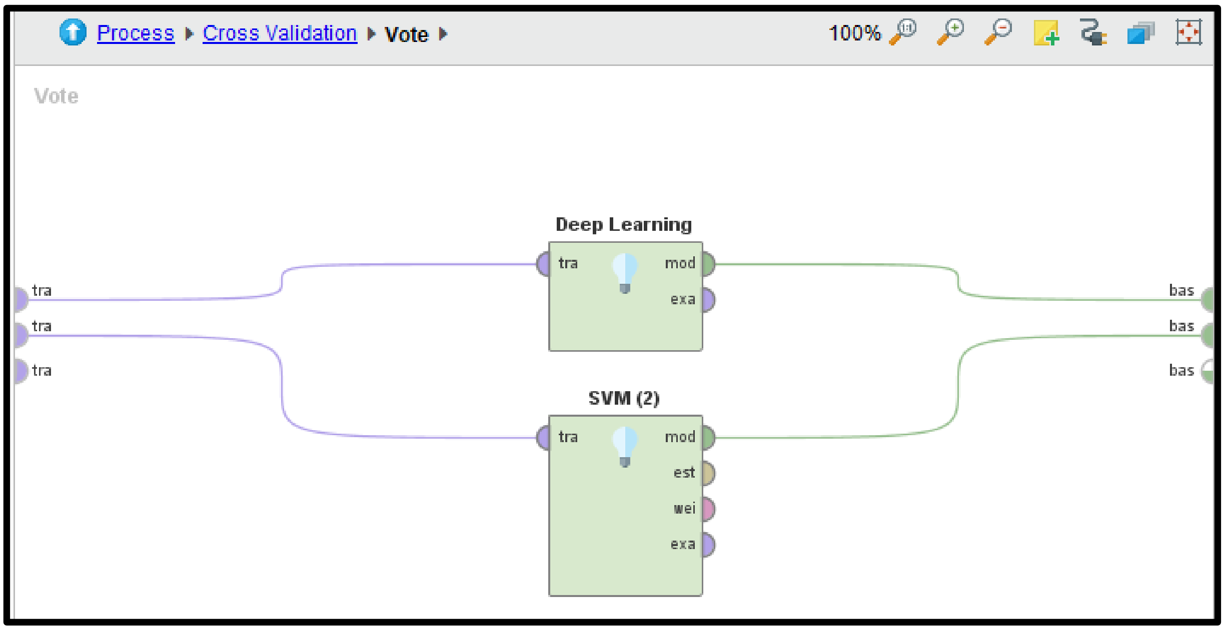
Appendix B

Referring to the results discussed in the **Figure 4.7** in Discussion and Analysis section**.**

Here is the process for the vote operation that was applied on the dataset

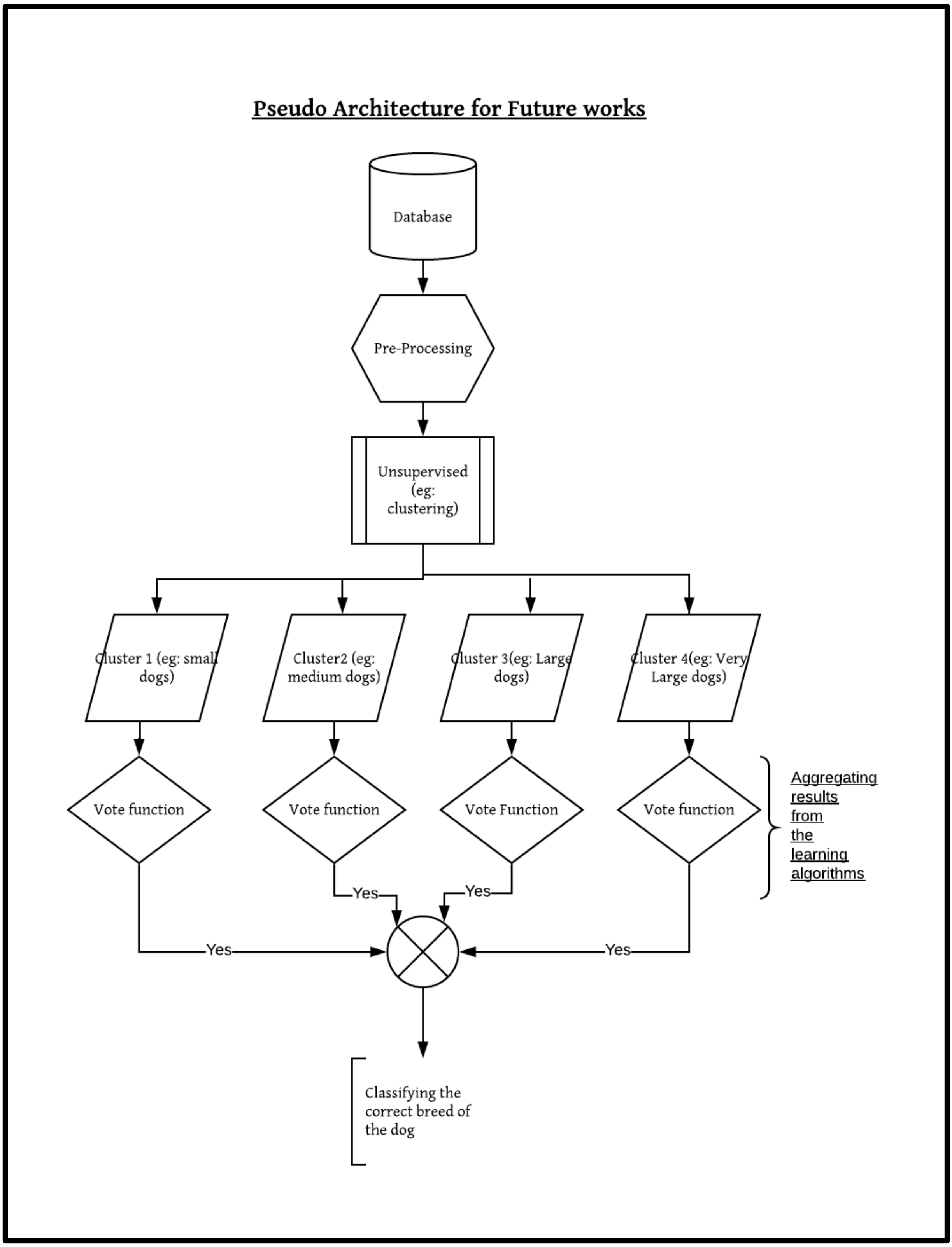






Appendix C

In reference to what the author discusses in the future works section of the document.



GLOSSARY OF TERMS

1. **Algorithm**: A process or a set of rules to be followed in calculations or other problem-solving operations, especially by a computer.
2. **Binomial:** Is an algebraic expression of the sum or the difference of two terms.
3. **Classification**: The action or process of classifying something.
4. **Cross validation**: Sometimes called rotation estimation, or out-of-sample testing is any of various similar model validation techniques for assessing how the results of a statistical analysis will generalize to an independent data set.
5. **Ensemble**: In statistics and machine learning, ensemble methods use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone.
6. **Fine grained**: Involving great attention to detail.
7. **GPU (Graphic Processing Unit)**: specialized electronic circuit designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to a display device.
8. **Hidden Layers**: A hidden layer in an artificial neural network is a layer in between input layers and output layers
9. **Kernel**: The most integral part of something
10. **Localization:** Is the adaptation of a product or service to meet the needs of a particular language, culture or desired population's "look-and-feel."
11. **Multiclass**: In machine learning, multiclass or multinomial classification is the problem of classifying instances into one of three or more classes.
12. **Multinomial**: In machine learning, multiclass or multinomial classification is the problem of classifying instances into one of three or more classes.
13. **Parameter**: A limit or boundary which defines the scope of a particular process or activity.
14. **Polynomial**: Consisting of several terms.
15. **Pre-Processing**: Subject to preliminary processing.
16. **Subset**: A set of which all the elements are contained in another set.
17. **Support Vectors**: The data points nearest to the hyperplane, the points of a data set that, if removed, would alter the position of the dividing hyperplane.

18. **Testing**: Investigation conducted to provide stakeholders with information about the quality
19. **Training**: Training is teaching, or developing in oneself or others, any skills and knowledge that relate to specific useful competencies.
20. **Process**: a series of actions or steps taken in order to achieve a particular end.
21. **Pseudo**: The prefix pseudo- is used to mark something that superficially appears to be one thing but is something else.
22. **Validation**: The action of checking or proving the validity or accuracy of something.