

|  |  |  |
| --- | --- | --- |
| Deep Learning and Image Classification | | |
| **Robert Wijntjes** | **C14356786** | **Oisin Creaner** |

**Table of Contents**

Deep Learning and Image Classification 1

1. Project Statement 4

2. Research 5

A Brief History 6

Data Mining 8

Data Modelling 11

Regression 14

Neural Networks 19

3. Description of Solution 24

4. Approach and Methodology 25

5. Design 31

6. Prototyping and Development 35

7. Testing 37

8. Issues and Risks 40

9. Plan and Future Work 44

10. Conclusions 47

11. Bibliography 49

[Figure 1 - Aristotle Logic Table 1st Iteration 6](#_Toc499471486)

[Figure 2 - Aristotle Logic Table 2nd Iteration 6](#_Toc499471487)

[Figure 3 - Four Pretences 7](#_Toc499471488)

[Figure 4 - Data Mining Hierarchy 8](#_Toc499471489)

[Figure 5 - Haar Wavelet Algorithm 13](#_Toc499471490)

[Figure 6 - Linear Regression Formula 15](#_Toc499471491)

[Figure 7 - Linear Regression Example 15](#_Toc499471492)

[Figure 8 - Linear Regression Chart 16](#_Toc499471493)

[Figure 9 - Logistic Regression Flow 17](#_Toc499471494)

[Figure 10 - Cross Entropy Formula 17](file:///C:\Users\rwijn\Desktop\Final%20Year%20College\FYP\RobertWijntjesC14356786_INTERIM.docx#_Toc499471495)

[Figure 11 - Kullback-Leibler Divergence Formula 18](#_Toc499471496)

[Figure 12 - Neural Network Low Level Diagram 19](#_Toc499471497)

[Figure 13 - Unsupervised Learning Graph 21](#_Toc499471498)

[Figure 14 - Supervised Learning Graph 22](#_Toc499471499)

[Figure 15 - Solution Breakdown Chart 24](#_Toc499471500)

[Figure 16 - Timeline breakdown 25](#_Toc499471501)

[Figure 17 - Gant Chart Workload Breakdown 25](#_Toc499471502)

[Figure 18 - Gant Chart Timeline 26](#_Toc499471503)

[Figure 19 - Scrum Board Iteration 26](#_Toc499471504)

[Figure 20 - ANN Use Case 31](#_Toc499471505)

[Figure 21 - Head Section of ANN 32](#_Toc499471506)

[Figure 22 - Tail Section of ANN 33](#_Toc499471507)

[Figure 23 – Low to Medium Level Prototype 36](#_Toc499471508)

[Figure 24 - Scrum Operation Flow 37](#_Toc499471509)

[Figure 25 - Rational Unified Process Flow 38](#_Toc499471510)

[Figure 26 - Bug Tracking Spreadsheet 39](#_Toc499471511)

[Figure 27 - Issues and Risks Tracking Spreadsheet 43](#_Toc499471512)

[Figure 28 - Current Gant Timeline 45](#_Toc499471513)

[Figure 29 - Current Gant Workload Breakdown 46](#_Toc499471514)

# Project Statement

The Objective of the project is to conduct research into the area of Image Classification.

However a main section of the project will be machine learnings implementation of image classification. Using the findings from varying research in the image classification spectrum I will then find comparisons and therefore draw conclusions based on research and implementation done.

The Project at this point can be broken simply down into three main areas:

* The Research into Deep Learning techniques, technologies and applications of it.
* The Research into Image Classification and the methods of improving accuracy and efficiency of such systems.
* Creating working Prototypes of such concepts in order to contrast and therefore assist the research that had been done.

# Research

Firstly I conducted initial research into the spectrum of environments that could be utilised to implement aspects of machine learning and image classification.

Python has many modules that are useful for image classification and/or machine learning. (1)

As python specializes often in data mining I could conclude it would be an important language to begin with. Listed below are a few of the many modules python has available for image classification and machine learning:

* ScikitLearn
* Tensorflow
* Opencv

Scikit learn showed that it was useful for a low to a high level complexity application as it has both the ability to create a low level classifier and a high level one.. Scikit learns documentation lists clearly its implementations it excels at, some of which are Classification, Regression, Clustering, Pre-processing, Modelling and Dimension Reduction

(2)

Tensorflow is another module that belongs to the python language which can be implemented when looking into the area of image classification and neural networks. From observation of the documentation however I would conclude it to be quite difficult to implement however it allows you to achieve a very high level of complexity due to the ability to select either your CPU or GPU to operate on.

(3)

Opencv is the module for python which is very strong at the transformation of datasets related to images as it was simple to implement while it maintained a very powerful capability to improve data.

Although one negative of the implementation of OpenCv is while there is much documentation often the documentation across updates is misquoted and therefore can cause issues with dependencies or syntax.(4)

In comparison to python there is also Java which possess the capabilities for machine learning and image classification. After observation of the documentation , generally I concluded that the documentation was worse and more complex than the python equivalent.(5) (6)

Therefore I concluded it would be more effective to focus on the python implementations as its variety is vaster.

## A Brief History

The true beginning of Artificial Intelligence begins with Aristotle. An inheritor of categorical syllogisms as he has been once quoted on saying

“All men are mortal. Socrates is a man. Therefore Socrates is a mortal”.(7)

In this raw way, it is the beginning of all machine learning principles as it can be translate to almost any scenario.

In this way it can be used on a low level to classify objects by the usage of four main principles, which are:

1. **Universal Affirmatives -** All Humans are Mortal.
2. **Universal Negatives -** No Humans are perfect.
3. **Particular Affirmatives -** Some Humans like to Dance.
4. **Particular Negatives -** Some Humans cannot Dance.

(8)

This can be more simply broken down into a logic equation:

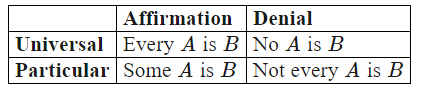
**

Figure 1 - Aristotle Logic Table 1st Iteration

In the latter Aristotle changes his method of expressing these predications by adopting a more artificial way of expression:

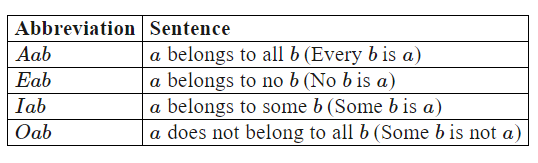


Figure 2 - Aristotle Logic Table 2nd Iteration

Instead of saying that “X is predicated of Y” he says “X belongs to Y”. It is an important change in technicalities as the term predicates only can give a true or false. The term ‘belongs’ in its original usage can mean both “Begin” and “Exist”. Therefore one can say that “X exists within Y” or “Some of X exists within Y”.(8)

These Following Diagrams would should be used as an aid to get a better understanding of the syllogistic logic:

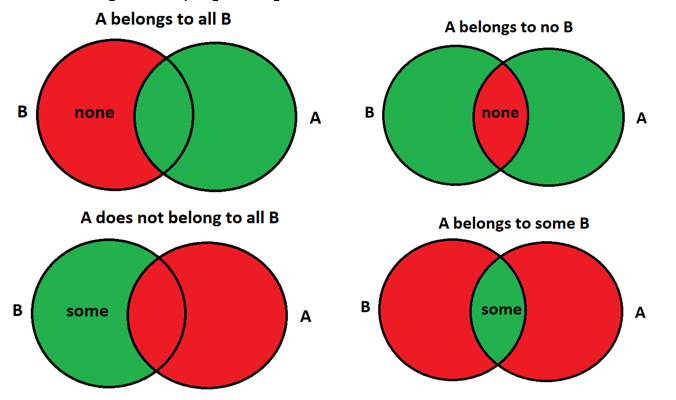


Figure 3 - Four Pretences

These four are the main pretences that build the foundation level of any machine learning system as it learns how to identify entities. While the diagrams do not show the data types, how it learns or the system itself, the fundamentals of it are here.(9)

It can quite simply be tested by applying these pretences whether or not your systems methods are of sound logic as if one of these pretences is not in conjunction with one of your methods then we can easily find the errors of your logic and quickly apply the correct logic fix and therefore improve the accuracy of the system as a whole.

## Data Mining

Data mining is the analytical process designed to explore data. This is often in relation to market or business data.(10)

Data mining’s two main principles are focused towards the development of scalable and efficient data mining tools. Data mining can be viewed as the natural evolution of information technology. Data mining specifically refers to the extracting or mining knowledge from large amounts of data. Although Data mining is the most widely used term for this activity it can also be referred to as KDD or Knowledge Discovery from Data. Alternatively others view data mining as an essential step in a process known as Knowledge Discovery. (11)

The figure shown below shows the flow:

Figure 4 - Data Mining Hierarchy

Below are the terms observed in the figure and given are their specific usage and definition:

1. **Data Cleaning:** Use to remove noise and Inconsistent data.
2. **Data Integration:** Combination of Data Sources.
3. **Data Selection:** Relevant Data in Relation to operation is retrieved.
4. **Data Transformation:** Data is transformed into the format accepted by forms appropriate for data mining.
5. **Data Mining:** Intelligent methods are applied in order to extract data patterns。
6. **Pattern Evaluation:** Identify interesting patterns representing knowledge on interesting measures.
7. **Knowledge Presentation:** Visualisation and knowledge representation techniques are applied.

(11,12)

Steps 1 – 4 are forms of data pre-processing where the data are prepared for mining by transforming the data into a format that will be more easily processed. Sometimes both data transformation and consolidation are performed before data selection in the case of data warehousing.

Steps 5-6 are data processing techniques which is where the intention is to extract useful or meaningful patterns from a dataset.

Step 7 is a technique called post-processing where the output of the previous step is taken and transformed into an entity which helps to more easily convey the data’s meaning or usefulness. This step is quite important as patterns may be difficult for a human to perceive if it’s still in its data formation but when it is transformed with the aid of a process to give a visual representation, the data can be more easily understood.(11)

The main body of an average data mining system involves these following main components:

**Database Warehouse:** This is one or a set of databases, data warehouses, spreadsheets or any other kind of file format that is a kind of data repository.

**Database Server:** The database or a data warehouse server is responsible for the requests of the user in terms of transmitting data

**Knowledge Base:** This component is used to evaluate the interestingness of a resulting pattern. An example of this could be concept hierarchies, used to organize attributes or attributes values on different levels of extraction.

In addition it could also be defined as the interestingness constraints,thresholds or metadata.

**Data Mining Engine:** This is the functional side of the data mining system and consists of a set of nodes for task characterization, association, correlation, classification, prediction, cluster analyse, outlier analysis and evolution analysis.

**Pattern Evaluation Module:** This component primarily interacts with the data mining modules to aid in the focusing of the analytics towards interesting data patterns.

**User Interface:** This component deals with the bridge between the Data Mining System and the User. In case the user interface is the front-end while the data mining system is the backend.

(11)

## Data Modelling

Data mining functionalities can be classified into 3 main categories. These 3 categories aim to give an insight into events and therefore give a clearer picture of the course of action.

**Descriptive:** It is used for learning from the past behaviours and understanding how they might affect future outcomes. Should be used when what is going on at an aggregate level.(11)

**Predictive:** It is used for understanding the future. It provides estimates of the probability of an outcome reoccurring in the future. Although until now it cannot be 100% accurate it can be used as guidelines to make intelligent decisions. Should be used when you want to know future details or fill in information you don’t currently have.(11)

**Prescriptive:** It is used in giving advice on possible outcomes of the future. The user inputs different actions in order to quantify the effect of future decisions in order to give advice on the action before execution. In summary it aids in predicting outcomes but also giving advice on those outcomes.(13)

Although these functionalities may lead to a pattern that shows the transformed data in a way that seems suitable one must also access the output with a level of trustworthiness.

Through the usage of these functionalities a data mining system may derive many kinds of patterns in relation to a dataset, however each kind of pattern can be categorised with a beneficial function.

**Content Class Description**

The Association of data with classes. Is used in building class Descriptions.

**Characterization:** Data Characterization involves the summarization of general characteristics of a target class. This can be done by simple database queries on fields within a class and the output can be Pie Charts, Bars Charts or characteristic rules.(11)

**Discrimination:** Data Discrimination involves the comparison of a target class with a single or set of comparative classes. It compares the general features of the couple using simple database queries. The outputs are similar to characterization however they include distinguishing descriptors and instead of characteristic rules, it outputs discrimination rules.(11)

**Frequent Patterns**

These are patterns that occur frequently within data. Some types of frequent patterns are item sets, subsequence and substructures.

**Association and Correlation Analysis:** This involves the frequency in which two or more entities show a correlation and the probability in which if one entity is there that another will correlate to it. These are known as the Support and Confidence rules. Minimum Support Threshold is use to find all Frequent Patterns in a database. Minimum confidence threshold is applied to all frequent patterns in order to help form constraints.(11)

**Classification and Prediction**

Classification is the process of finding a model that describes and distinguishes data classes or concepts. The model is derived from the analysis of training data. Below are 3 types of implementations that can be used.

**Decision Tree:** A Decision tree is a flow chat in a linear direction which forms the Tree Pattern. Stem represents a category and each branch represents an outcome and the lowest level represents an entity. Decision trees can very easily be converted into classification rules as the diagram itself is very easily comprehended.

(11) (14)

Input: shape( X , “Round” ) AND colour( X , “Orange” ) 🡪 Class( X , “Mandarin” )

**Neural Network:** A Neural Network is a collection of weighted processing nodes. A (ANN) Artificial Neural Network consists of 3 parts. The Input layer, the hidden layer and the output layer. The Hidden layer usually incorporates a kind of learning process. It often uses a backwards propagation of error rule. The process follows through to output and then repeats itself as many times until it becomes relative to the model. (11) (15) (12)

While it is very effective for classification many argue because of the processing power needed to operate, it isn’t too realistic for everyday usage.

**Haar Scale:** Uses the Haar Wavelet to detect objects. Specifically done through the usage of square shaped functions. (16) .This technique was adapted for image classification by the focusing of this algorithm on pixel intensities.

**formula**

Figure 5 - Haar Wavelet Algorithm

The algorithm above Takes adjacent squares at a specified region, sums up the pixel intensities in each region and calculates the difference between these summations of regions. (16)

The main feature of the Haar Cascade Classification is its processing speed however its accuracy is lacking as its algorithm for classification’s complexity is not high and although it is less accurate than that of the neural network.(17)

## Regression

When performing these classification functions it is equally important to classify objects that are not part of the dataset in which you want to classify as that which are as removing the possibility of false data making its way into the training sets are important as it often degrades your accuracy. For example in a dataset of flowers, it would be important to remove the pictures that have over the threshold of noise or are not relevant such as a picture of a car in a dataset of flowers. This process of identifying attributes that do not contribute positively to the classification or prediction is done by implementing a Relevance analysis technique.

**Relevance analysis** is an important part of each analysis step, Recognition of the most important variables aid in reducing redundant data and also reduces the levels of the uncertainty model. There are two key features that relevance analysis best implements.

* Recognition of the most important variables
* To understand the relation and logic:
  + Predictor to the target data
  + Predictor to the target data from the target data’s perspective.

Therefore this aids to remove irrelevant data and in this projects case this could be images with heavy noise or not of the classification.(18) (11)

**Regression Analysis** is a set of statistical processes for calculating the relationship among variables, or often the calculation of a standard linear or non-linear line in order to aid in the relationship between data. The two main variables that are examined are the Dependant variable and then one or many independent variables. More specifically regression analysis helps to show how a dependent variable changes when an independent variable changes while other remain the same.(19) (11)

Regression analysis is widely used for prediction and forecasting where it is often used in machine learning however the rule of correlation and causation must be adhered to. This is when two variables are found to be correlated as they it is assumed that one must cause the other. This rule is untrue as variables may correlate but may not be a causation relationship. (20)

For example:

*If the rate of people who have been in an accident correlates with people who are now in a wheelchair.*

One could derive that using a wheelchair is dangerous as it is incorrectly compared. There are many types of regression models that can be implemented but just for an example sake below are three that are commonly used.

**Linear Regression** attempts to model a relationship between two types of data on a scale. The first type of data is a dependant variable and the second is the independent variable. These two variables are combined to find a single case. The objective in Linear Regression is to find the regression line that most suitably fits the data currently available. This means trying to find a line with the minimum amount of errors.(11) (12) (21)

Linear Regression can be done by the following formula: Y = a + Bx

Y is equal to the dependant variable and x is equal to the independent variable.

(22)

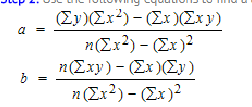


Figure 6 - Linear Regression Formula

Equation 2 shows the required formula to complete a linear regression calculation where both the x and y, their power 2 and their multiplication are needed in order to successfully carry it out.

This will give you both an ‘a’ and a ‘b’ so that you may map these numbers to the graph in order to plot out the appropriate model.

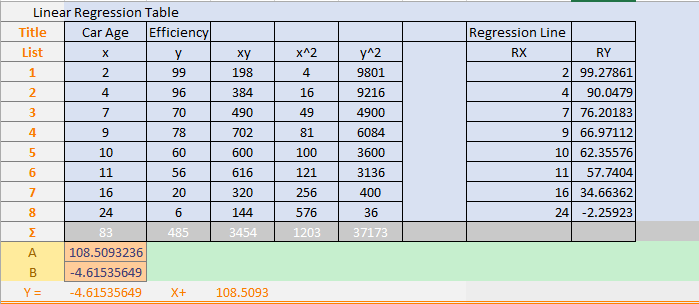


Figure 7 - Linear Regression Example

In Figure 8 I have done the calculation of a simple Linear Regression Model. The theory is that with this model an individual could predict some data. A younger car age is usually equal to a higher efficiency rating.

When doing any kind of regression by hand I would suggest using Excel to work it out in the way I have done so that you can truly understand the underlying calculations.

In figure 8. I have taken the output of the equation and mapped it into a graph to show the linear model in correlation with the original data fed in.

Figure 8 - Linear Regression Chart

While linear regression is quite easily implemented it does have a few negatives to its usage.

It assumes the relationship is linear and doesn’t adapt to varying data. However it is decent for seeing the statistical slope in the data’s relationship. This means its usage for Image Classification would not be the suitable as often the data derived from a set of images is greatly varying and would need a regression technique that would be able to adapt to such a style.

However there is a kind of regression called **Softmax Regression** which is specifically useful in the area of image classification.(21)

Softmax regression is quite effective for image classification as it is a form of Logistic Regression however it has one benefit as unlike logistic regression it isn’t binary and therefore can take in many classes as input.

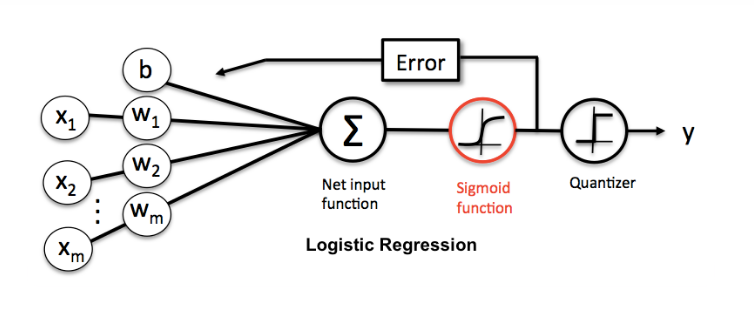


Figure 9 - Logistic Regression Flow

It takes a standard logistic regression classifier and then instead of the sigmoid function it uses the Softmax function. The sigmoid function is used for determining the curve.(23)

The benefit of using a Softmax function over a Sigma function is the sigma functions are more effective when each output are independent. Therefore a new class every time. However in my case the output will always be one of a set of image descriptors then Softmax would be more suitable as it excels when each output can be linked to exactly one class.(24)

In conclusion Sigma is good for binary classification and Softmax is good for multi-classification. Softmax Regression doesn’t assume a linear relationship between Independent variables and dependent variables and therefore can handle non-linear effects. It is also far more robust as variables do not need to be normally distributed.

The figure x shown above is the standard format for the structure in the hidden layer of a neural network. It takes x amount of inputs and in our case it would be the datasets related to images. It then runs a function over them and checks the output.

Using the Softmax Algorithm I have provided a visual aid to display the process in which it is done.

Another technique that is very useful when implementing a system with many classes is called **Cross Entropy**. Cross Entropy is most often used to quantify the difference between two sets of probabilities or in many cases two vectors. In the case of an image classifying neural network it would take the probability exported from your classes and then the hot encoded vector that corresponds to ones labels. It is represented by the formula below.(25)

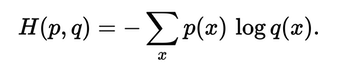


Figure 10 - Cross Entropy Formula

H is the distance between the p, q variables. The P variable is representative of the vector results from the Softmax Regression. Q variable is representative of the hot encoded labels. p(x) is representative of the iteration through the Softmax results vector and q(x) is representative of the iteration through the q labels vectors. E means to get the inverse of the average over the whole set of vectors.(26)

The aim of this calculation is to get two different types of data. One where it has a low distance for the correct class and a high distance for the incorrect class. The objective is to achieve a training loss which is small as this would prove that the operations do a good job of classifying data however if the training loss is big then one could prove that the operations do a bad job of classifying the data.

Another kind of Entropy that is often used is one called **Kullback-Leibler Divergence** or **KL divergence** for short. It is represented by the formula shown below.(11)

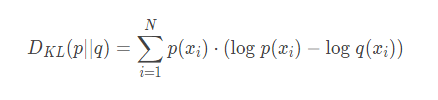


Figure 11 - Kullback-Leibler Divergence Formula

In this case the target is fixed and therefore there would be no difference from running cross entropy compared to Kullback-Leibler divergence but Cross Entropy is the more widely used in image classification.

## Neural Networks

Neural networks are one deep learning technique for image classification. There are many different types of neural networks. Listed below are a few:

* Feedforward
* Regulatory feedback
* Radial basis function
* Recurrent neural network
* Modular
* Physical

(27)

There are four main types of neural networks under the Feedforward category.

* Auto encoder
* Probabilistic
* Time Delay
* Convolutional

The specific nature of these types of neural networks are their architecture.

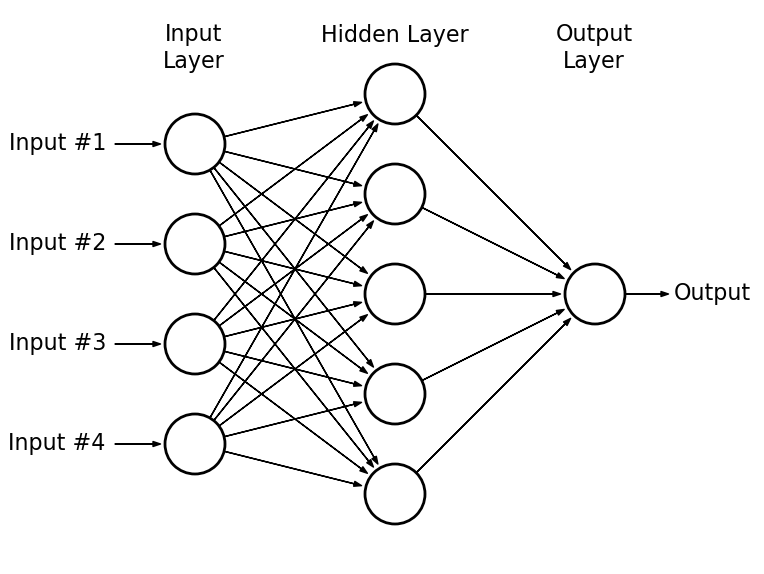


Figure 12 - Neural Network Low Level Diagram

The input layer is a passive layer that doesn’t modify the data while the hidden layer and the output layer are active layers and may modify the data.(28)

However this architecture can be altered to become either **supervised** or **unsupervised**. The difference on a low level between supervised learning and unsupervised learning is the inputs in which it takes in. (11) (12)

In supervised learning the overall objective is to map input to an output whose correct values are given by a ‘supervisor’. In unsupervised learning this is not the case. There is no supervisor and therefore only a dataset for input is given.

One method of density estimation would be Clustering. The overall objective of this is to find groupings for different perceived sets of data.

Supervised learning uses a training set, a validation set and a test set. The training set is the basis for constructing a model in which can then be used to predict data in the later stages. The validation set is used to provide an unbiased evaluation of the model found in the training set while tuning model hyper parameters. The test set is used to provide an unbiased evaluation of the final model. This is one of the key differences between the two types of learning. The structure of neural network implemented later is an MLP – Multi Layer Perceptron. (29)

A MLP structure is broken down into 3 main characteristics.

* One of more Layers of Hidden neurons that are not a part of the input output layers although it adds to the systems capabilities to solve complex calculations and issues.
* Non-Linear flow of data is differentiable.
* Interconnected model that has a high level of connectivity.

(29) (23)

Ultimately the main difference between unsupervised learning and supervised is the type of data in which Unsupervised takes as input.

Figure 13 - Unsupervised Learning Graph

In unsupervised learning the data which is given has not been set a dataset where example data has been associated with being a positive or a negative of being associated with a class. It uses self-organising neurons to identify hidden patterns. Although this paradigm doesn’t provide a learning signal and therefore there can be a lack of direction. (11,12,29)

While in some cases this may be negative as the neural network may learn to find patterns that are not useful it can also be used to discover patterns that had been previously overlooked or missed and therefore this technique is good for the discover of new unprecedented patterns.

This lack of direction however can be fixed in later stages using different kinds of analysis such as Regression or Relevance analysis.

As we can see in figure 13. , In this case of unsupervised learning when the inputs are first read in they are all perceived as anonymous classes as there has been no training set provided and therefore there is no possibility of previous data existing to aid in the prediction process. All data is the same class for the time being.

Figure 14 - Supervised Learning Graph

In the figure 14. We can see an example of a type of dataset that would be an input for supervised learning. This is because the given data has already been assigned a positive and a negative of being a class or in this case it has already given example of what a typical piece of data from a class might look like statistically.(11,12,15)

In a case study done by R. Sathya of Jyoti Nivas College they made a comparison of two of the paradigms of artificial neural networks.(29)

The Supervised learning technique they implemented was the MLP type which is the same as the one I have implemented. The Unsupervised learning technique that they implemented was the SOM – Self Organising Maps technique.

The purpose of the study was to demonstrate the effects of these two learning paradigms in a pattern classification scenario.

In the case of the supervised technique it was given an 11-4-3 design with back propagation. The ANN was trained with a 300 dataset and 50 were used to train it.

In the case of the unsupervised technique it was given a 10 input and 3 output.

The dataset used to originally train the supervised ANN was used to train this model.

Their intended dataset was a set of students and their skillsets and group them based on their performance. In their results they found that:

* The Observation through these paradigms facilitates new theories and knowledge that is imbedded in patterns.
* Learning behaviour of the neural network enhances the classification properties.

Although the main conclusion in which they took was that the supervised learning algorithm is very efficient for non-linear real time problems while the unsupervised model is not.

Therefore in this case where it is linear, the unsupervised method was more efficient than the supervised method.

Although in the case of image classification as it is non-linear it would be best to implement a supervised method.

(29)

# Description of Solution

The Solution in which I would like to implement is showing the effectiveness of algorithms for image classification however I would also like to specifically delve into the spectrum of Data mining and deep learning techniques to show the advantages and disadvantages of such paradigms. Therefore I would be able to show a clear comparative between image classification with deep learning and without and why certain techniques are more useful in certain circumstances.

A few solutions in which I already have given in the research conducted up to this point however in the future there will be more to divulge:

* Benefits of using Supervised Learning and Unsupervised and how they are beneficial in certain areas.
* The different types of analysis available for the data mining part of the system and therefore which would be most useful in this project.
* The difference between certain analysation techniques and where in my system they are used effectively.
* Performance Cost of running different type of techniques to show their pros and cons.

Figure 15 - Solution Breakdown Chart

Therefore the solutions given above have been thoroughly research however I still believe in the months to come there will be research added to the sections when I begin to more deeply look into Image Classification.

# Approach and Methodology

At the beginning of the project I believed it was most important to create a kind of plan for the coming months.

Through past experience I was knowledgeable in this area as I had previously done modules on software development. Therefore I decided it was key to first build a base chart for what I would like to accomplish in time. Below is the first draft of the chart.

|  |  |  |  |
| --- | --- | --- | --- |
| Start Date | End Date | Description | Duration ( Days ) |
| 02-Oct | 31-Oct | Work on Background Research and Begin Origins Work. Design Timelines and Project Design. | 30 |
| 01-Nov | 30-Nov | Begin Heavy Research into Data Mining and Deep Learning | 30 |
| 01-Dec | 31-Dec | Implementation of Deep Learning Technique and Continue Research of such Area. Begin Research into Classification | 31 |
| 01-Jan | 31-Jan | Continue Research and Implementation of Deep Learning. Begin Research of Image Classification. | 31 |
| 01-Feb | 28-Feb | Begin Heavy Research Into Image Classification. Continue Research Deep Learning. Implement Image Classification. | 28 |
| 01-Mar | 30-Mar | Apply finishing touches to Image Classicisation and Deep learning. | 30 |
| 01-Apr | 30-Apr | Plan to have the Project Complete and all minor problems solved and documented. | 30 |

Figure 16 - Timeline breakdown

The next iteration of this chart was when I began modelling the timelines and used the Microsoft project software to help visualize this.



Figure 17 - Gant Chart Workload Breakdown

Above is the second iteration of the Gant chart in which I would like to model later. There were some changes to the Task Names and the Start/End dates.

Below is the Model of the Gant chart in which I designed to set out the general tasks aimed to be completed in their associated months.

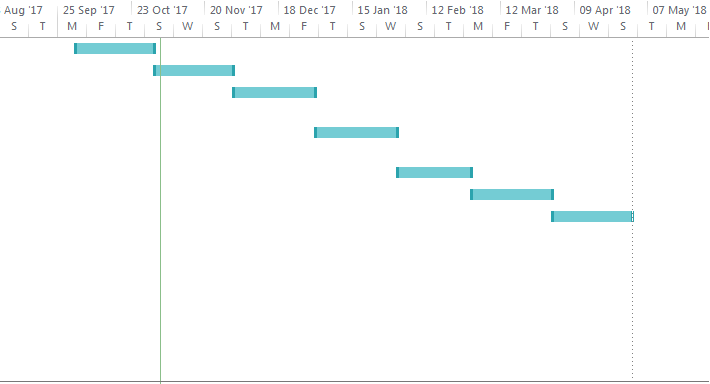


Figure 18 - Gant Chart Timeline

(30)

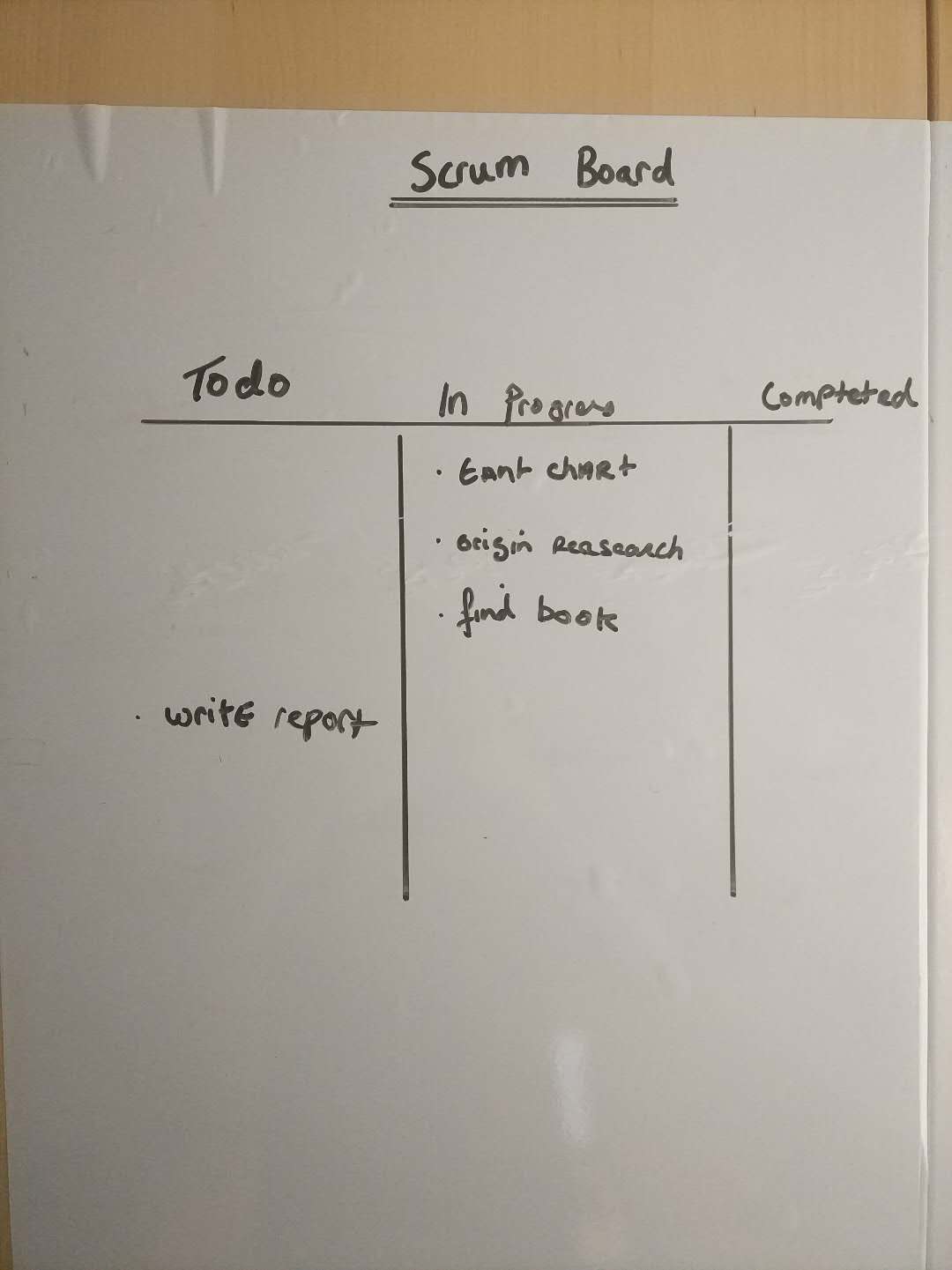


Figure 19 - Scrum Board Iteration

(31)

Since the Gant chart that would illustrate the general outline of the months until submission was designed, the next step in which I took was to implement the Scrum methodology in where I would set a sub-monthly objectives.

In figure 19. Shows an example of one of the weekly iterations where I would need to do some objectives that had been set within the first month. The usefulness of the scrum board quickly became very clear as it allowed for changes in plans and therefore added a factor of dynamism. (32)

Since this project is a research based structure, the priority of information analysis is very high. The implementation of Research Vetting therefore is key to a solid information background.

Research Vetting is the process of running one or many background checks on data that is used when writing an article of choice.

There are 3 main characteristics when looking for useful information on a subject. These are named below:

* Relevant
* Reliability
* Point of View

(33)

The meaning of relevance is that is the data found up to date and in conjunction with the most modern understand and descriptors.

In a case when I was looking for data on the best methods of regression for image classification I found a document stating that an old outdated method was best , although it may have been best at that time , it is not now.

The meaning of Reliability is that is the source in which the data originates from reputable and is it accurate. For example is the data you have collected perhaps second hand data that was misquoted.

When reading through papers of comparisons of feedforward neural networks I would often find information that would not specifically mention that the type of feedforward is special as it implements some other techniques which change the structure.

The meaning of Point of view is also an important characteristic as it is the deciding factor if the information may be up-to-date and published by a reliable source but unfortunately this time round there is a shade of bias. A balance of left and right helps to weigh the pros and cons of a situation and therefore are important.

When researching I never found any biased information as it was mostly factual information in which I searched for such as algorithms. However the conclusions may be biased as it is in my informed opinion which I draw them from.

The importance of correct citation will lead to:

* Credibility to topics explored.
* Gives a rabbit hole of information for the reader.
* Protects against plagiarism.
* Upholds the institution of studies name.

(34)

Therefore for Reference management I used a combination of two software’s. The first being ‘CiteThisForMe’. CiteThisForMe is a very usefull software when wanting to quick reference documents where there would not need a huge level of storage of such references. However it has a separate function for searching for articles of information by such headers as title, author or ISBN number. Therefore I found this software to be very useful when writing the references for articles.(35)

The second software I used was ‘Zotero’. Zotero’s benefit over CiteThisForMe is its cross platform capabilities and functionality involved when processing references. Although it doesn’t have a functionality to search a database of articles for writing up your references, it does however have advanced functionality to store them. It can be easily implemented into long documents with many references needed as it incorporates an automatic updating system for references.

In Conclusion I used ‘Zotero’ for Storage and implementation of references however often used ‘CiteThisForMe’s’ Article search capabilities. (36)

Requirements capture is one the initial steps in software development or event project development as it aims to give a set of core aims for the project.

However requirements capture can be broken down into 3 main areas:

**Elicitation Requirements.**

**Analysing Requirements.**

**Recording Requirements.**

(37)

The elicitation of requirements can be also known as the requirements discovery phase as its core idea is to find a set of initial requirements which can be followed. This may be broken down into a few main components of which this should be carried out:

Identify the Problems.

Identify that the Problems are real.

Show the value of Meeting the Problems.

Define what’s needed to meet these requirements.

Evaluate and Deploy the Documentation.

(38,39)

During the first stage of the project I aimed to complete a bulk of the initial requirements capture. This involved meeting with my mentor, discussing it with fellow students, discussing the idea with lectures and doing some research into the area.

Identifying the problem that I wanted to solve was quite difficult as I knew I wanted to use a certain technology but unsure of how it helps.

Finding what was needed to solve or evaluate the problem wasn’t too difficult as there was much existing documentation and it solely involved selecting which route I would like to take to evaluate the problem.

The challenges involved in the capture of requirements may also be broken down into a 3 main areas:

**Project Scope** – Boundaries defined to encompass the project.

**Understanding** – Comprehend the true direction of the project.

**Volatility** – The ability to adapt to changes in the requirements.

Project scope involves the boundaries which will constrain the area of study in which will be conducted. This area is quite important when it comes to management of time as setting your scope too big will limit the quality of your work and too small there will not be enough quantity to your work.

Initially I had issues with the scope as I wanted to do many different things but I decided to slim the scope to make the project more feasible.

Understand of the project meaning is that you have a high grasp of what the requirements set actually mean. This also means you understand the constraints involved that may limit how you carry it out so that certain standards are implemented.

Once the Initial issues were sorted out with scope then understanding the project didn’t have many issues however I still would get advice on topics aimed to research to make sure my direction was still okay. I deemed it best to get advice from a different angle/

Volatility is often one of the topics in which people do not inherit as people get stuck in loops and find it difficult to change their ways. However the ability to adapt to the project requirements are a high priority as the world is constantly changing and therefore so are the circumstances.(37–39)

As the project progressed since I implemented a scrum chart, I was able to adapt to changes in which the project took so that I could easily and smoothly change paths.

# Design

As the project is mainly research orientated there section of design will not incorporate as much as a development orientated project would. However it is still possible to give a low level design of some key applications I will create.

In order to model these low level designs it is important to pick a frame work in order to design them effectively. Therefore I have adapted the UML standard for a majority of the design.

Use Case:

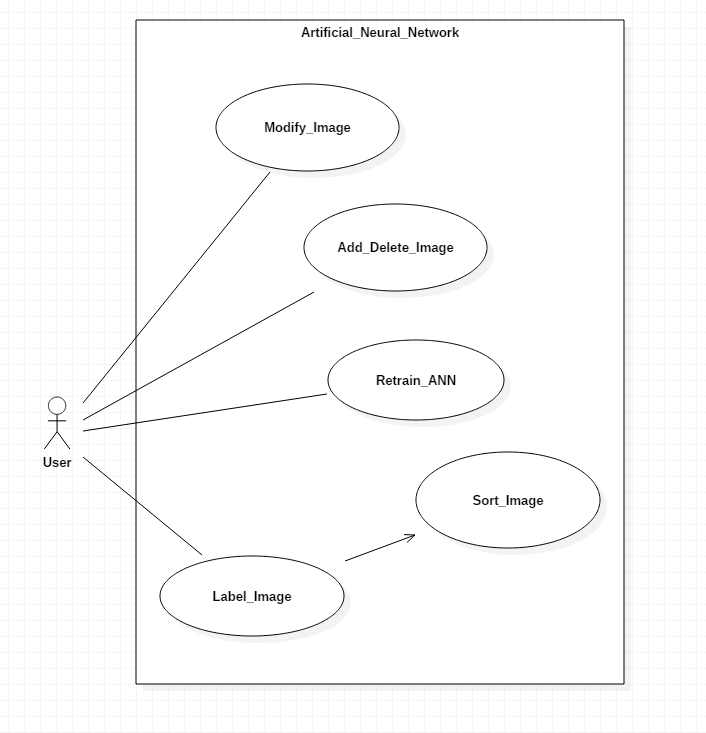


Figure 20 - ANN Use Case

Here we can identify a common user of the **Artificial Neural Network** where the user has access to a few functions:

* The modify image gives the user access to a function which may transform the images currently in the training dataset.
* Delete or add an image gives the user access to a function which will manually add or delete an image from a dataset.
* Retrain ANN gives the user access to the function that retrains the ANN using the training dataset.
* Label image gives the user access to a function which takes in an image or a set of images and classifies them based on the calculation done in the neural network.

On a higher level the design of the neural network can be shown in the diagram below:

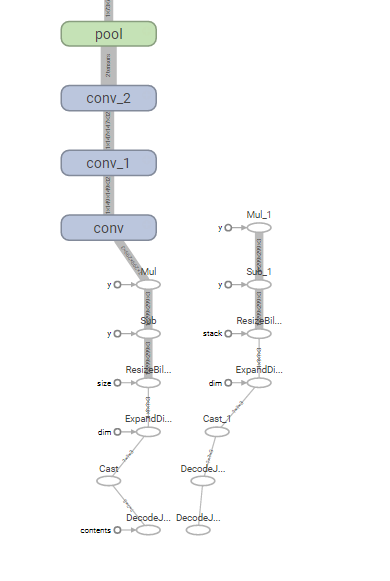


Figure 21 - Head Section of ANN

This section of the neural network includes the breaking down of the images and the beginning stages of the learning section of the ANN.

* **Conv**: Convolution Layer
* **Pool**: Down Sizing of Data

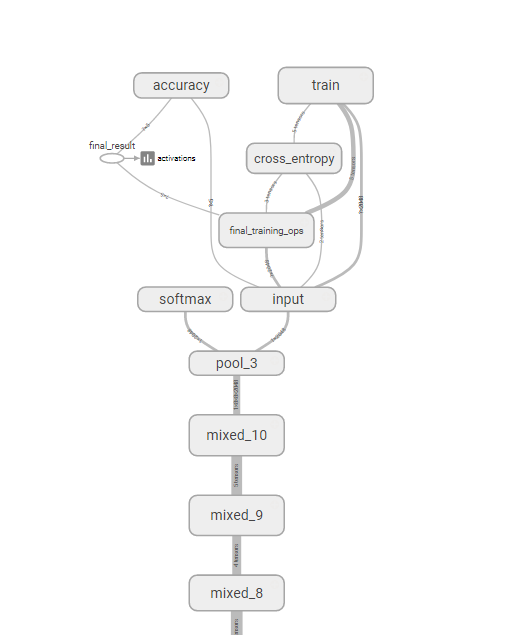


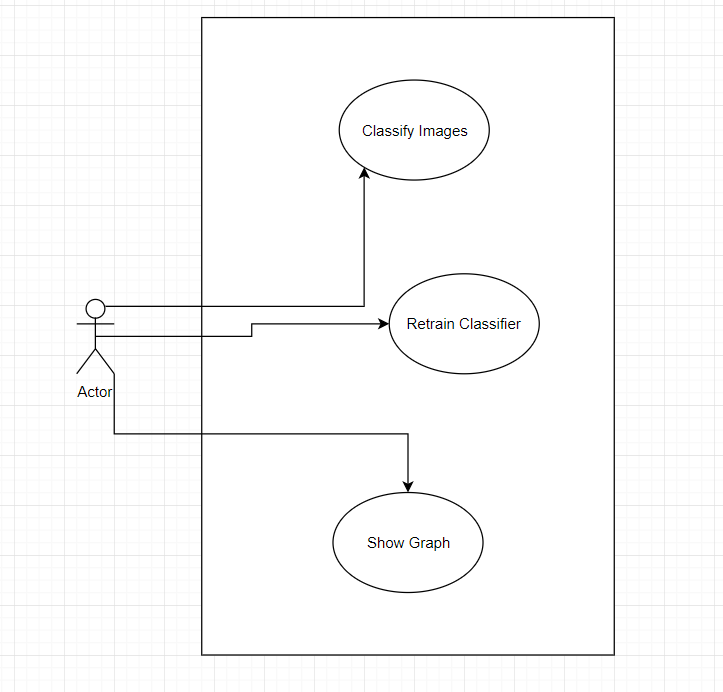
Figure 22 - Tail Section of ANN

This section is the end section of the neural network where the final stages of training are performed and then prepared to be mapped and outputted.

* **Mixed**: adjusts data using such techniques as back propagation
* **Softmax**: Algorithm for Regression
* **Cross Entropy**: Calculation of Weights
* **Accuracy**: Takes the weights and outputs

(3,40)

KNN Classifier:



An example use case of the KNN classifier in which I intend to implement. It has 3 main functions:

1. Classify using supervision on images from a dataset.
2. Retrain the Model using new images.
3. Show the Graphs for the Classification and accuracies.

# Prototyping and Development

During the prototyping and development phase I conducted research into the area of different methods of design and also different methods for development. Some methods in which I researched are listed below:

The first is **Moqups**. It is a **wire framing** type of software. It is a Rich Web application that a user can utilise in order to create UI. Not only is the user interface of the application very good, it also is very easy to use. This is due to the well layered out areas for adding different modules. However the weakness for this software is that you are limited to only one project at a time. While this may be a problem for great size projects, since I only need one project then it is still quite suitable. The fact it is also a software gives many benefits such as file storage. Unfortunately it also lacks in giving functionality depth as it is solely for UI design. (41,42)

Unfortunately because it is an application, it has its limits in the area of creativity but it still is good for designing the foundation sketches.

The second technique I will consider is **POP** which means **Prototyping on Paper**. This technique is different for each designer as it relies heavily in the skills of the prototype. This technique is easily the best for creativity as it has no constraints apart from the pen and paper. The main downside to this is that the longevity of the designs are short as newly created designs must be redrawn. This may take up a lot of time however for no-complex applications this wouldn’t be a huge problem. One Weakness it also has over the Moqups is that the file handing system must be carefully implemented which usually would include the scanning of designs and the sorting into folders of the designs.

(43)

This technique is quite nonreciprocal however it can also provide a slight depth into functionality which may serve to increase its effectiveness. However the general effectiveness is low.

A software which I thought was great but I cannot use as it is only Mac Specific is **Origami Studio** which is made by Facebook to help the teams to design and build products. It is a really easy to use prototyping software that can create medium to high level designs. It includes file storage and makes design longevity last.

You can add basic levels of interaction into the designs which are great for showing the foundation meanings to your design. Overall it is a really useful prototyping software for MacOS designers.(44)

I wish to implement a Low to Medium Level of design as I feel that there is no need for the other two types as the complexity in the User Interface is quite low as most of the complexity is in the operations.

Three levels of design:

* **Wireframe:**  Low to Medium Level
* **Mocup:**  Medium Level
* **Prototype:**  Medium to High Level

Therefore I would like to implement a Wireframe Design using the software Moqups.

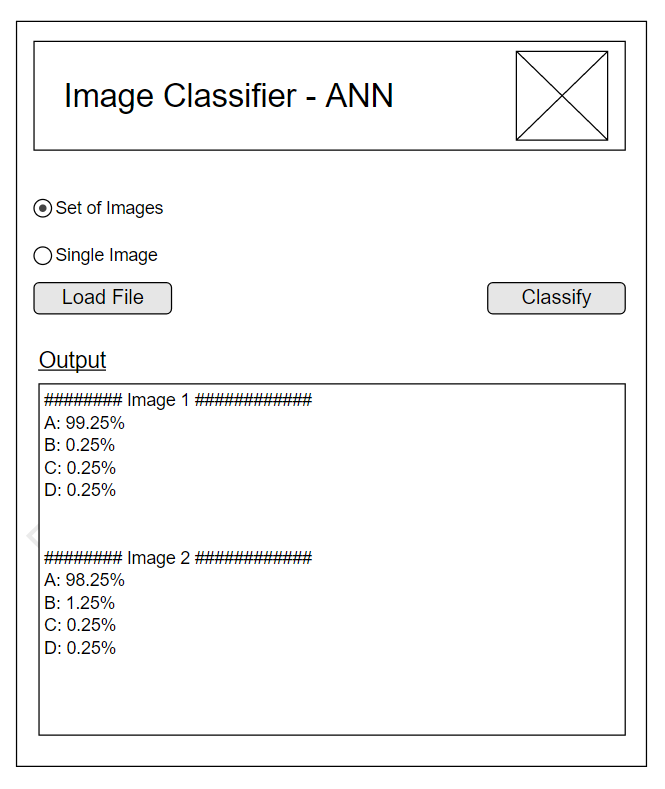


Figure 23 – Low to Medium Level Prototype

As the complexity is in the operation rather than the user interface I decided on this design as it clearly marks out the functions. It gives the option to load an image to classify by set of images or by a single image. A simple classify button that will run the operations on the intended images. It also includes a convenient textbox where the results will be outputted to.

# Testing

In the testing stage I conducted research into the area of development strategies that would assist and benefit testing. Below are a 2 strategies:

The first technique is called Agile Scrum Testing Process. It involves separating different features into modules that will be the aim for every release. Builds are incremental and therefore the final build should incorporate all features that we promised to be delivered. (31,32)

Some key features of this process are:

* Testing is done throughout the project lifecycle.
* The project team are responsible for decisions.
* Iterative Development Technique implemented.

One other key feature which is the implementation of User Stories. This is where requirements for the project are not written in one huge block but are built up. In a case, a user story which needs to be delivered would be:

A User

Wants to achieve a target

For a certain Reason

Iteratively these stories are built up. Therefore instead of creating one huge requirements document they’re iteratively compiled. One of the main strengths of scrum are that it’s very simple to understand and straightforward. (31,32)



Figure 24 - Scrum Operation Flow

(31)

The second technique is called the RUP. Originally designed as a software process product by Rational Software. This technique breaks testing down into 4 stages.

* **Inception**
  + The idea for the project is defined and the project resources are checked to see if it is possible to complete.
* **Elaboration**
  + The project architecture and required resources are checked. Time lines and such things as costs are considered.
* **Construction**
  + The project is developed and completed. Software has been Designed, Written and Tested.
* **Transition**
  + Software is released to the public however some testing continues based on user-feedback.

(45)

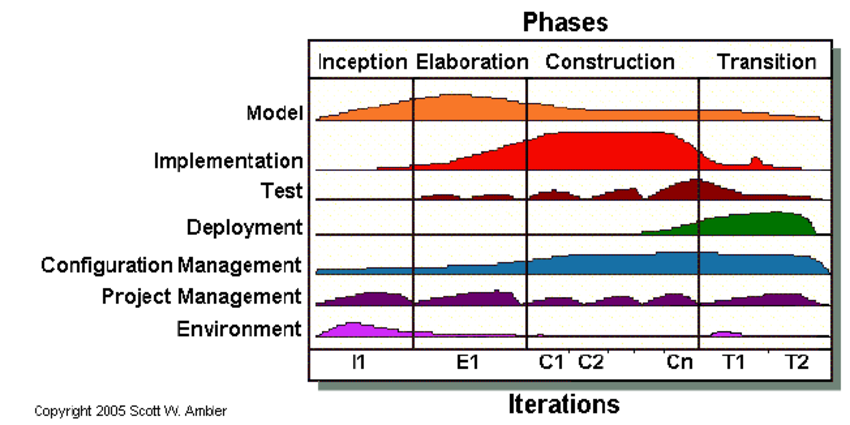


Figure 25 - Rational Unified Process Flow

The strength of RUP is that it effectively reduces the costs of unexpected development costs. However I believe for this project the lifecycle of RUP is too long and scrum is more suitable.

The reasons for this decision will be brought into more depth below.

* Scrum teams work very closely with the users and therefore can respond very quickly to user needs as a product backlog that is updated and prioritised aids this.
* Those who decide the stories are the ones going to be executing those stories and therefore they should be more manageable as they are specific to each member’s capabilities and performance.
* It has a high focus on communication among departments or sectors.
* The Progress is very visible as daily scrums shed light upon it.

(31,32,46)

In conclusion the scrum technique is very effective with small teams or in my case it is most relevant to my project as I am the sole stakeholder in this project and I will be the sole decision maker.

I have also implemented a spreadsheet for tracking Bugs during testing procedures so that they would not cause issues in the later stages.

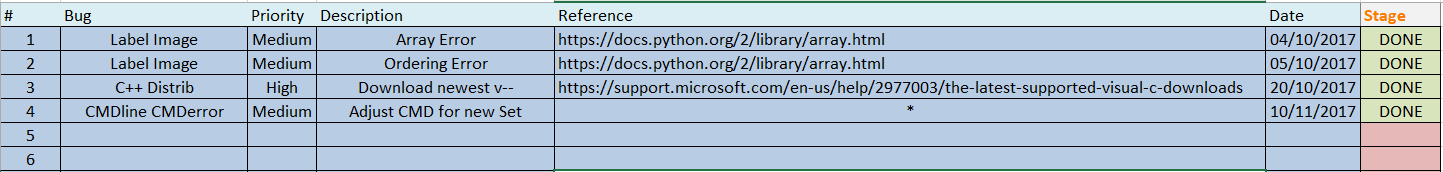


Figure 26 - Bug Tracking Spreadsheet

* **#**: Is relative to the id number of the bug.
* **Bug**: Is relative to the section where the bug applies.
* **Priority**: Is relative to the level of urgency in which it needs to be completed.
* **Description**: Gives a more in-depth description of the bug.
* **Reference**: Provides a link from research to help fix the bug.
* **Date**: When the bug was found to give an insight to how long it hasn’t been fixed.
* **Stage**: Green represents bug fixed, yellow represents bug currently fixing, red represents not yet accepted.

(47)

This spread sheet has been proved useful for keeping track of bugs across sprints so that if an error of the same type re-occurs it can be tracked down and a similar procedure for fixing can be applied.

# Issues and Risks

At this stage of the project there has been many objectives that have been overcome but there also have been some issues in which have been more difficult to complete. Some of the main areas in which there has been problems are:

**Finding good comparison articles**

The meaning of this is that when search for articles about image classification with deep learning or without while they would discuss the underlying technology they would not often doing comparisons of other methods.

This made it difficult to draw conclusions of which methods are better in different situations as many of the solutions were general.

However I would still be able to draw a very general conclusion such as which algorithm is useful for image classification and which are not.

An example of this would be the difference between Softmax regression and logistic regression. Softmax regression is good for many different classes while logistic is only good for determining if it is a positive class or a negative class hence the name logistic. Either a Positive or a Negative.

**Too broad research**

In the beginning I didn’t really know exactly what I wanted to do and therefore that would cause problems for later on.

However I managed to fix this very broad sense of research by adjusting my research spectrum.

I started with Machine Learning and Image Classification which was too broad and then changed to neural networks and image classification, but now I am focused on the applications of image classification in deep learning.

This spectrum allows me to do research into machine learning while also giving some room to non-machine learning tactics in order to compare effectiveness among a few descriptors.

**Technology**

After about a month into the project my main laptop broke with a lot of important software that could not have been backed up.

Although I managed to save all my data due to constant backups of my crucial data into the cloud, some of the software lost was used in articles I had wrote.

Therefore the timeframe in which I originally thought out had to be adapted to suit the new requirements in that specific iteration.

Now that the main areas of high priority risk have been assessed it is important to implement a type of Risk Strategy or create your own by adapting a known strategy to suit the requirements of your system. I chose to implement my own strategy of Risk Management – RM.

This techniques implements the three main areas that risk management may fall into. Such categories are listed below for reference:

**Avoidance – Eliminate or Disassociate.**

Avoidance of a Risk may seem harmless but it is also a bad practise as if nothing is gained and nothing is lost then there has not been any improvements nor any negative areas that feedback can be derived from in order to improve the system as a whole.

Avoidance of a risk would be not research into an area of neural networks as you may think it’s too complex. Simply avoiding it would mean that you would lose beneficial information due to the unknown.(48–50)

I made sure to research into all areas in which were open to me in order to avoid this risk and if I became stuck on an issue I would schedule to go through it in-depth at another time.

**Reduction – Optimize through Mitigation.**

Reduction of a Risk or better known as risk optimisation. While reduction of risks as a concept is acceptable there are also other factors at play that may limit or hinder the paradigms implementation. This area of risk management takes a reduction strategy and weights the positive and negative outcomes and therefore try’s to find either a balance of negative and positive or attempts to find a strategy that positive outlies the negative.(48–50)

In the beginning I had a bad working example of a complex ANN. It was running off Docker which caused many complications and therefore to reduce the risk , I found a way to implement it without the use of docker.

**Retention – Accept and Budget.**

Risk Retention is the acceptance of the risk when the cost of insuring against the risk would be greater over time than the total losses sustained.(48–50)

Originally I didn’t have some functionalities working in my ANN example however I didn’t want to change my file system in case the ANN stopped working. However the I deemed that accepting the risk was not of value and that it was more necessary to apply the fixes.

**Risk analysis** involves in the assessment of risks through the usage of qualitative and quantitative strategies to further define a situation and therefore recognise threats.

Anything in moderation is acceptable although when the extremes of an entity are taken then they are to cause issues. The same applies to risk management in the form of hindrance. (50)

Factors that may hinder this process are:

**Over Complexity**

**Uncertainty**

**Resource Wastage**

If the system which aids in risk management is too complex for the requirements then the overall process is slowed beyond the point of reason. (51)

For example if I am conducting research but I require second and third authenticators and letters of authenticity then this levels of risk management would be too much for checking information from a supposed legitimate source.

However if it were from a source that isn’t as legitimate then perhaps an email requesting their references surely would be sufficient.

**Uncertainty** is one the main factors that degrade the risk management system. This involves in the moderators of the system being uncertain of how to conduct a risk assessment and therefore much time would be wasted. (52)

This kind of factor can usually be overcome with the implementation of experienced analysists who using logic and experience can make informed solutions to Risks.

During the process of handling Issues and Risks I implemented a spreadsheet in order to manage issues and risks throughout the week so that I can accordingly keep track of them. The headers used are:

* **Issue Number**
  + So that Issues can be easily referenced and if multiple people are working together that the spreadsheet can be easily managed.
* **Issues and Risks ( Title )**
  + To give a quick and simple name to the area in which the risk or issues affects.
* **Priority**
  + The level of priority relates to the urgency of the issue or risk. In this case putting together the interim report is of high importance all round.
* **Description**
  + Gives a more in detail description of what the issues and risks are as it helps to describe the problem.
* **Date**
  + To keep track of how long the issue has been left unresolved so that its priority may be raised or lowered.
* **Stage**
  + To keep track of the issue or risk to make sure none slip through the cracks and cause a problem at a later sprint.

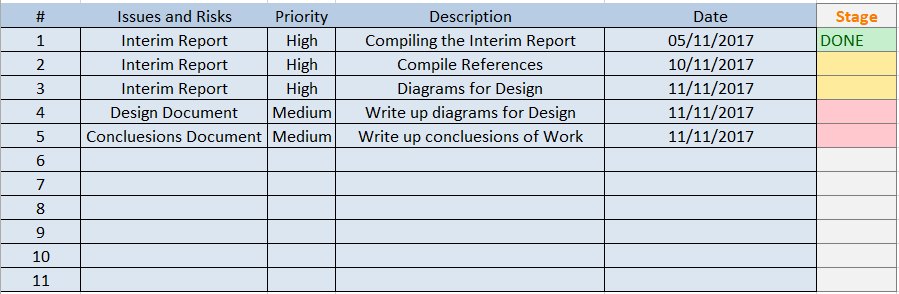


Figure 27 - Issues and Risks Tracking Spreadsheet

The Colour of stage is relative to whether or not the issue or risk has been accepted and therefore we clearly see what is left to work on and should still be worked on.

* Green: Complete.
* Yellow: Currently in Progress.
* Red: Not yet Accepted.
* Grey: No Risk Found yet.

# Plan and Future Work

As of the submission date I have accomplished nearly all objectives that were set out both in the Gant Chart and in the Scrum Iterations.

However there is still much work that is needed to be done as the project has only passed the halfway mark. Some areas in which I will soon begin work on are listed below:

**Further Analysis of Deep learning Techniques**

Although I have already conducted a thorough investigation into deep learning as set out by deadlines I still believe there is much I can learn from the area. When going forward to do more research into Image Processing and Image Classification I believe there will be topics that are relatable and therefore useful in the project.

**Analysis into different techniques for image classification**

So far I have done research into deep learning techniques for classification and different kinds of learning paradigms and analysis however I still believe there is a decent amount which can be learned about specific image classification in relation to Deep Learning. Such headings as:

* Image Classification Algorithms
* Comparisons of different algorithms
* Preparation of the Datasets for Image Classification
* Procedure of Image Classification

**Image Processing Techniques**

Image processing is a topic which often is relatable to the area of image classification as it often involves the modifying of images through the usage of libraries like OPENCV to better suit the system’s needs.

Such modifications could be:

* Changing the Colour Scale
* Colour Segmentation
* Morphing of Images
* Improving of images through noise reduction techniques

**Background research**

While I believe I need to do research into the other entities in which I want to research and implement, I still need to do the underlying technology research for them.

This could be what kind of languages or environments are needed to effectively carry out the investigation

The plan in which I will follow will be the same as specified in the gate chart except it will continue on from the current iteration and date.

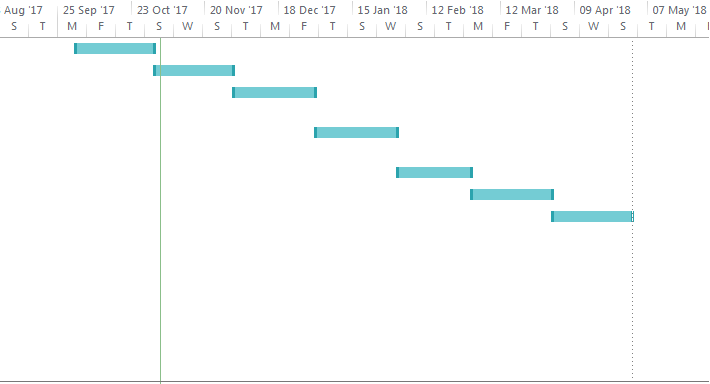


Figure 28 - Current Gant Timeline

In the Gant chart above it illustrates the date timelines in which will follow the submission however since these are only done by month and the scrum chart follows this , the scrum chart will be more in detail for the weekly objectives as it is designed to be dynamic and react well to changes and deliverables.

Now that we have reached month 3 in the Gant charts progression we can see the kind of work that is planned to take place.

1. Implementation of Deep Learning / Continue Research.
2. Image Classification Research and continue a little Deep Learning.
3. Image Classification Research and Implementation.
4. Implement Testing Phase and Plan Scrum for final Iterations.
5. Plan on Adding the Finishing touches to the Main Document.



Figure 29 - Current Gant Workload Breakdown

An example Scrum Chart Iteration within the month #5 could be:

* Finding a comparative paper.
* Finding an Insightful book.
* Conduct Project Management Design.
* Check Past Written Documents for Error.

# Conclusions

During the course of the investigation I have come to draw conclusions through the comparison and clarification of topics. Some of the main conclusions I have found thus far are:

**Supervised Unsupervised**

I found out that supervised learning is very efficient for non-linear datasets while unsupervised learning is efficient for linear datasets. This is according to the article written by Sathya and research collected from numerous books I had studied.

**Softmax Logistical Regression**

I found out the key differences between a basic Logistic Regression and a Softmax Logistical Regression and that in the case in which I would like to implement that a Softmax Logistical regression would be more beneficial as it works well with many classes.

**Layering of Neural Networks**

I discovered that there are many different kinds of neural networks and then again within those categories exist further deviations of those kinds. For example a forward propagation neural network with MLP implemented.

**Regression**

I had to do a lot of research into the area of what regression actually is and through the implementation of a linear regression model and heavy research into more complex structures such as Softmax Logistical Regression and Logistical Regression I found out that regression models are often used for the prediction of data.

**Data Mining**

I gained a greater understanding in the area of the data mining process that would be incorporated into a Machine Learning system. Such things as its general structure and different methods for finding patterns and therefore learning that it’s possible to extract further data embedded into datasets once thought to have been exhausted of information.

I have also made vast steps toward understanding the flow of work in an efficient development cycle as I have deeply studied the effects and advantages of different management techniques including where their disadvantages are most.

**Scrum Development Technique**

After researching the Scrum Methodology using their official website and a few books I found that scrum is one of the top development structures currently and that its benefits are huge in the communication side of development where teams communicate with teams and then with the major stakeholders. It also has strong benefits in the area of project planning as it’s easily able to adapt to new situations.

**Issues, bug and risk management**

During the evaluation of different risk, bug and issue managing techniques I concluded that Scrum was highly effective at tracking weekly progress to insure requirements are met. I also have concluded that the gant chart is useful for setting monthly goals.

In conclusion although I still have much to learn I believe I have come far since I began this project gaining great knowledge in the area of project management and the topics surrounding deep learning and image classification.

# Bibliography

1. The Best Machine Learning Libraries in Python [Internet]. Stack Abuse. 2015 [cited 2017 Nov 23]. Available from: http://stackabuse.com/the-best-machine-learning-libraries-in-python/

2. scikit-learn: machine learning in Python — scikit-learn 0.19.1 documentation [Internet]. 2017 [cited 2017 Nov 13]. Available from: http://scikit-learn.org/stable/

3. TensorFlow [Internet]. TensorFlow. 2017 [cited 2017 Nov 13]. Available from: https://www.tensorflow.org/

4. OpenCV library [Internet]. 2017 [cited 2017 Nov 13]. Available from: https://opencv.org/

5. Java Machine Learning Library (Java-ML) [Internet]. [cited 2017 Nov 16]. Available from: http://java-ml.sourceforge.net/

6. Welcome to OpenCV Java Tutorials documentation! — OpenCV Java Tutorials 1.0 documentation [Internet]. [cited 2017 Nov 16]. Available from: http://opencv-java-tutorials.readthedocs.io/en/latest/

7. Bobonich C. An Introduction to Philosophy [Internet]. Stanford University; [cited 2017 Nov 23]. Available from: https://web.stanford.edu/~bobonich/glances%20ahead/III.logic.language.html

8. Smith Robin. Aristotle’s Logic. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy [Internet]. Spring 2017. Metaphysics Research Lab, Stanford University; 2017 [cited 2017 Oct 9]. p. 4. Available from: https://plato.stanford.edu/archives/spr2017/entries/aristotle-logic/

9. 20th WCP: Computational Complexity and the Origin of Universals [Internet]. [cited 2017 Nov 16]. Available from: https://www.bu.edu/wcp/Papers/Mind/MindPerl.htm

10. What is Data Mining, Predictive Analytics, Big Data [Internet]. [cited 2017 Nov 16]. Available from: http://www.statsoft.com/Textbook/Data-Mining-Techniques

11. Jiawei Han, Micheline Kamber. Data Mining - Concepts and Techniques. 2nd ed. Morgan Kaufmann; 2006. 702 p.

12. Ethem Alpaydin. Introduction to Machine Learning. 2nd ed. Massachusetts Institute of Technology; 2010.

13. Ullrich C. Descriptive and Prescriptive Learning Theories. In: Pedagogically Founded Courseware Generation for Web-Based Learning [Internet]. Springer, Berlin, Heidelberg; 2008 [cited 2017 Nov 16]. p. 37–42. (Lecture Notes in Computer Science). Available from: https://link.springer.com/chapter/10.1007/978-3-540-88215-2\_3

14. Rokach L. Data mining with decision trees : theory and applications : Series in machine perception and artificial intelligence. World Scientific; 2010.

15. Jared Dead. Big Data, Data Mining and Machine Learning. Hoboken, N.J.: John Wiley & Sons; 2014.

16. Hazewinkel M. Encyclopaedia of mathematics. Dordrecht: Kluwer Academic Publishers; 2000.

17. Cascade Classification — OpenCV 2.4.13.4 documentation [Internet]. [cited 2017 Nov 16]. Available from: https://docs.opencv.org/2.4/modules/objdetect/doc/cascade\_classification.html

18. Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann; 2016. 655 p.

19. Stewart J. Calculus. 8th ed. McMaster University: Cengage Learning; 2016.

20. Aldrich J. Correlations Genuine and Spurious in Pearson and Yule [Internet]. Vol. 4. 1995 [cited 2017 Nov 16]. 364-367 p. Available from: https://web.archive.org/web/20060219042545/http://www.economics.soton.ac.uk/staff/aldrich/spurious.pdf

21. Unsupervised Feature Learning and Deep Learning Tutorial [Internet]. [cited 2017 Nov 16]. Available from: http://ufldl.stanford.edu/tutorial/supervised/SoftmaxRegression/

22. Chang J. Linear Regression [Internet]. Yale University; 1997 [cited 2017 Nov 16]. Available from: http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm

23. Raschka S. Python Machine Learning. Birmingham, Uk: Packt Publishing; 2015.

24. Kühnberger K-U, Rudolph S, Wang P. Artificial General Intelligence: 6th International Conference, AGI 2013, Beijing, China, July 31 -- August 3, 2013, Proceedings. Springer; 2013. 217 p.

25. Plunket, Elman. Exercises in Rethinking Innateness [Internet]. MIT Press; 1997 [cited 2017 Nov 16]. 166 p. Available from: http://www.cse.unsw.edu.au/~billw/cs9444/crossentropy.html

26. Grosse R. Training a Classifier [Internet]. Computer Science University of Toronto; [cited 2017 Nov 16]. Available from: http://www.cs.toronto.edu/~rgrosse/courses/csc321\_2017/readings/L04%20Training%20a%20Classifier.pdf

27. Lipmann RP. An Introduction to Computing with Neural Nets. IEEE ASSP Magazine; 1987.

28. Ng A, Ngiam J. UFLDL Tutorial [Internet]. 2013 [cited 2017 Nov 5]. Available from: http://ufldl.stanford.edu/tutorial/supervised/SoftmaxRegression/

29. Sathya R, Abraham A. Comparison of supervised and unsupervised learning algorithms for pattern classification. Int J Adv Res Artif Intell. 2013;2(2):34–38.

30. What is a Gantt Chart? Gantt Chart Software, Information, and History [Internet]. [cited 2017 Nov 16]. Available from: http://www.gantt.com/

31. What is Scrum? [Internet]. Scrum.org. 2017 [cited 2017 Nov 12]. Available from: http://www.scrum.org/resources/what-is-scrum

32. Getting Started with Agile Scrum Methodology: The Complete Guide for Software Developers and Testers — Software Testing Help [Internet]. 2017 [cited 2017 Nov 5]. Available from: http://www.softwaretestinghelp.com/agile-scrum-methodology-for-development-and-testing/

33. Craiglow H. Research Guides: MSF Orientation : Vetting & Citing Your Sources [Internet]. 2017 [cited 2017 Nov 21]. Available from: http://researchguides.library.vanderbilt.edu/c.php?g=68930&p=445413

34. Help finding information online [Internet]. The University of Edinburgh. [cited 2017 Nov 16]. Available from: https://www.ed.ac.uk/information-services/library-museum-gallery/finding-resources/library-databases/databases-overview

35. Save Time and Improve your Marks with CiteThisForMe, The No. 1 Citation Tool [Internet]. Cite This For Me. [cited 2017 Nov 21]. Available from: http://www.citethisforme.com

36. Zotero | Home [Internet]. [cited 2017 Nov 21]. Available from: https://www.zotero.org/

37. Boeing B, Shannon D. Systems and Methods for Analyzing Requirements. 2016.

38. Michael Christel, Kyo C. Kang. Issues in Requirements Elicitation. Technical Report CMU/SEI-92-TR-012. CMU / SEI.; 1992.

39. Tran E. Requirements & Specifications [Internet]. 18-849b Dependable Embedded Systems Carnegie Mellon University; 1999 [cited 2017 Nov 16]. Available from: https://users.ece.cmu.edu/~koopman/des\_s99/requirements\_specs/

40. TensorFlow For Poets [Internet]. [cited 2017 Nov 16]. Available from: https://codelabs.developers.google.com/codelabs/tensorflow-for-poets/#0

41. What is wireframing [Internet]. Experience UX. [cited 2017 Nov 16]. Available from: http://www.experienceux.co.uk/faqs/what-is-wireframing/

42. Online Mockup, Wireframe & UI Prototyping Tool · Moqups [Internet]. [cited 2017 Nov 16]. Available from: https://moqups.com/

43. Paper Prototyping: The 10-Minute Practical Guide [Internet]. Studio by UXPin. 2016 [cited 2017 Nov 16]. Available from: https://www.uxpin.com/studio/blog/paper-prototyping-the-practical-beginners-guide/

44. Origami Studio — Design Prototyping [Internet]. [cited 2017 Nov 16]. Available from: https://origami.design/

45. Rational Unified Process. Ration Softw. 1998;18.

46. An Empirical Framework For Learning (Not a Methodology) [Internet]. [cited 2017 Nov 17]. Available from: http://scrummethodology.com/

47. CDC. CDC Works 24/7 [Internet]. Centers for Disease Control and Prevention. 2017 [cited 2017 Nov 17]. Available from: https://www.cdc.gov/index.htm

48. Dorfman MS. Introduction to Risk Management and Insurance. Englewood Cliffs: Prentice Hall; 2007.

49. Verne U of LV 1950 TSL, Verne C 91750593-3511 © U of L. Risk Reduction and Avoidance [Internet]. Risk Management. 2011 [cited 2017 Nov 17]. Available from: https://sites.laverne.edu/risk-management/risk-reduction-avoidance/

50. Hubbard D. The Failure of Risk Management: Why It’s Broken and How to Fix It. John Wiley & Sons; 2009. 46 p.

51. Technology C © MI of, reserved 1977-2017 All rights. Understanding and Managing Complexity Risk [Internet]. MIT Sloan Management Review. [cited 2017 Nov 17]. Available from: http://sloanreview.mit.edu/article/understanding-and-managing-complexity-risk/

52. Kaan D. Defining risk and a framework for moving towards resilience in agriculture. Colo State Univ. 2002;