

MSc in Software Engineering and Database Technologies

CT621 Artificial Intelligence

**Student Name:** Cristina Borges

**NUIG ID Number:** 20234955

**Course:** CT621 Artificial Intelligence

**Assignments:** Final Exam

**Date of Submission:** 11/07/2021

Contents

1. Question 1 3

(a) (i) Parental Tree Diagram 3

(a) (ii) 3

(b) 3

(c) 4

2. Question 2 8

(a) (i) 8

(a) (ii) 8

(a) (iii) 9

(a) (iv) 9

(a) (v) 10

(b) (i) 11

(b) (ii) 12

(c) (i) 12

(c) (ii) 12

(c) (iii) 12

(c) (iv) 13

Question 3 14

(a) 14

(b) 16

(c) 19

Question 4 20

(a) 20

(b) 22

Appendix 1 23

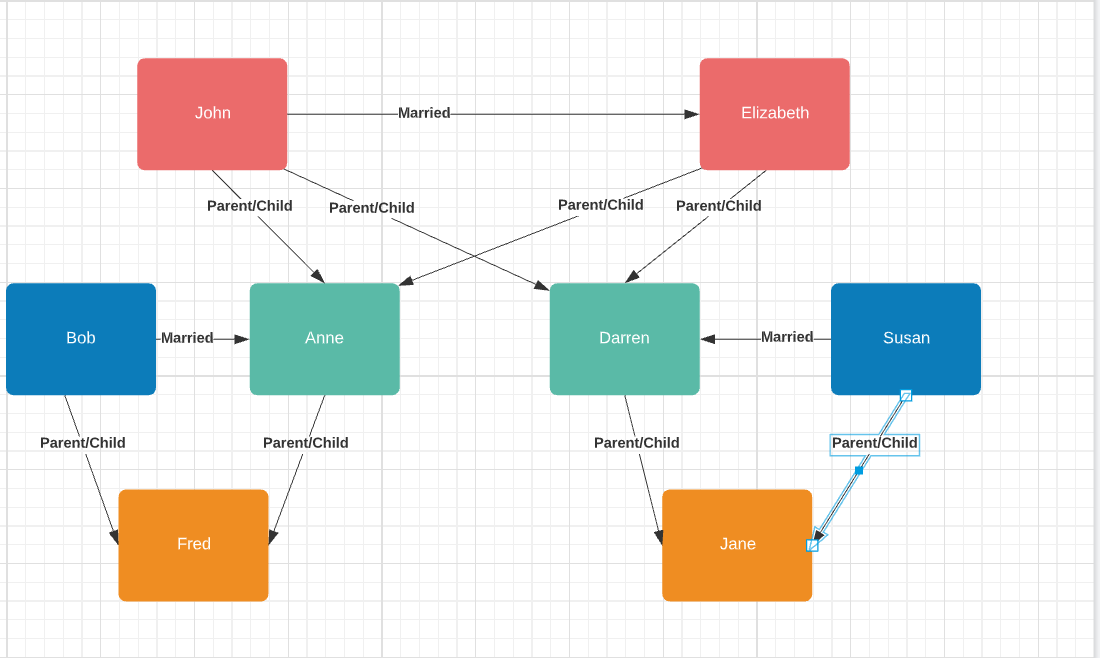
Appendix 2 25

Appendix 3 30

Appendix 4 33

## 1. Question 1

## (i) Parental Tree Diagram



Lucid.app

## (ii)

Male(john)

Male(bob)

Male(fred)

Male(darren)

Female(elizabeth)

Female(anne)

Female(susan)

Female(jane)

Married(elizabeth, john)

Married(anne, bob)

Married(susan, darren)

Parent(bob, fred)

Parent(darren, jane)

Parent(john, anne)

Parent(elizabeth, darren)

## 

is\_married(X, Y) :- married(X, Y), married(Y, X).

is\_married(Y, X) :- married(X, Y), married(Y, X).

father(X, Y)

mother(X, Y)

## 

A .pl file has been created and tested in SWI prolog – a copy of the file has also been shared by email and the contents are shared in Appendix A

I am summarising the expressions set out in the code:

The following from the code are facts:

*male(john).*

*male(bob).*

*male(fred).*

*male(darren).*

*female(elizabeth).*

*female(susan).*

*female(jane).*

*female(anne).*

*parent(john, anne).*

*parent(elizabeth, anne).*

*parent(john, darren).*

*parent(elizabeth, darren).*

*parent(bob, fred).*

*parent(anne, fred).*

*parent(darren, jane).*

*parent(susan, jane).*

*married(elizabeth, john).*

*married(bob, anne).*

*married(susan, darren).*

*married(john, elizabeth).*

*married(anne, anne).*

*married(darren, susan).*

*father(john, darren).*

*father(john, anne).*

*father(darren, jane).*

*father(bob, fred).*

*mother(elizabeth, darren).*

*mother(elizabeth, anne).*

*mother(susan, jane).*

*mother(anne, fred).*

The facts summarise the following:

* Gender of all individuals
* All parent, child relationships
* All mother, child relationships
* All father, child relationships
* All marriages

The following from the codes are rules:

*father(X,Y):- male(X),*

*parent(X,Y).*

*mother(X,Y):- female(X),*

*parent(X,Y).*

*grandfather(X,Y):- father(X,Z),*

*parent(Z,Y).*

*is\_married(X,Y):- father(X,Z), mother(Y,Z).*

*is\_married(X,Y):- mother(X,Z), father(Y,Z).*

The rules set out that:

* A father (x) to a child (Y) means that the father is male and is a parent of child Y
* A mother (x) to a child (Y) means that the mother is female and is a parent of child Y
* A grandfather(x) to a grandchild (y) means that the child (Y) has a parent (Z). Parent (Z) is then in turn the child of a male (X) i.e. X is the father of Z
* A father (X) to child Z is married to the mother (Y) of child Z

The following screenshots from the SWI Prolog demonstrate the following relationships:

* john is the father of anne – shown as true
* elizabeth is the mother of anne – shown as true
* bob is the father of fred – shown as true
* elizabeth is the mother of Darren – shown as true

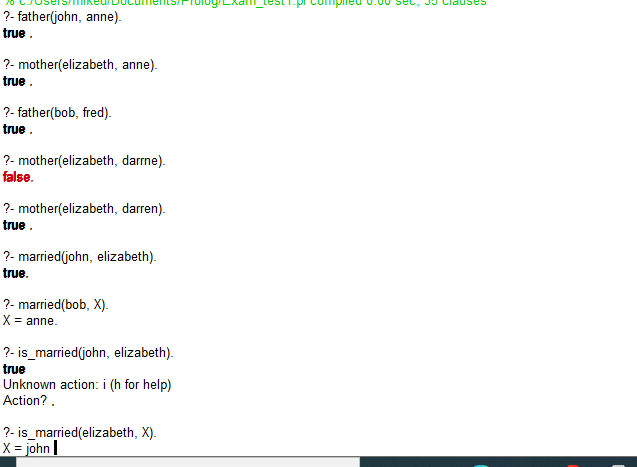
These 4 relationships are true which is consistent with the expected relationships as outlined in the question.

Other scenarios tested were:

* john is married to elizabeth – shown as true
* solve for who elizabeth is married to – answer returned of john which is correct
* solve for who is parent of fred – answer returned of bob
* solve for who is child of anne – answer returned of fred

I also tested the following relationships, all of which returned false as expected:

* test if anne is parent of jane
* test if fred is father of fred
* test if john is mother of anne
* test if john is married to bob
* test if susan is married to fred





I chose to describe the predicate grandfather(X,Y) in my code. When testing this, I tried both:

grandfather(john, fred) – which returned true

grandfather(john, jane) – which returned true

grandfather(X, fred) – which returned John

These results were all as expected.

## 2. Question 2

## (i)

**Markov Decision processes**

A Markov chain has the objective of making predictions about transitions to future states based solely on the current one (Tanwar S, 2019). An extension of the Markov chain is the Markov decision process which sets out a framework for a multi-state model with transition probabilities between each potential state.

Some of the key features of a Markov Decision process are that:

1. all future transitions are dependant on the current state
2. the probability of moving to a different state depends on the current state only
3. there will be some states which are not reachable from other
4. There can be absorbing states from which the process will not move

Markov decisions processes are typically what is know as a stochastic process, where the probability of a transition from one state to another is modelled by a random variable rather than a single deterministic probability. The use of stochastic modelling in the Markov process means that the uncertainty within the process can be accounted for and allows for a range of outcomes to be produced.

A real world example of where a Markov decision process could be used is when modelling the health of an individual. In a set up like that, the key features might be:

1. States of healthy, sick, terminal and dead
2. The “dead” state would be what is called an absorbing state where the probability of leaving the “dead” state once entered is zero.
3. The healthy state might only be able to reach the sick or the dead state
4. Sick state could reach healthy, terminal or dead states
5. Terminal state could only reach sick or dead states
6. The probability of transitioning from 1 state to another may be modelled using a stochastic random variable
7. Could have a time element i.e. the longer that person is in terminal state, the more likely they move to dead state rather than recovering to healthy

In the context of AI, the Markov Decision Processes are often used to determine the appropriate balance to take between exploration and exploitation (DeepAI)

## (ii)

**Partial Observability in MDPs**

A Partially Observable Markov Decision Process (POMDP) is a form of a Markov Decision Process (MDP). The key difference is that for a POMDP, the current state may not be directly observable (Background on POMDP’s). A POMDP will allow for an observation to be made that will depend on the state being occupied at set points in time to determine the state it is currently believed to be in (H Kamalzadeh, M Hahsler 2019). A probability distribution function (pdf) can be assigned to the POMDPs view on their current state. By using a pdf, we can get results giving a range of outcomes with different probabilities assigned to intervals of the range.

An interesting example of the real life application is in the area of navigation of robotics where the POMDP can allow for the uncertainty of movement between states (locations), Pineau and Thurns paper shares the results of an experimental example in a nursing home. The findings of their study was that the use of robot using POMDP to manage uncertainty in decision making led to positive outcomes following interactions and the key problems identified were largely design/hardware issues i.e. inappropriate speech recognition system and initial issue relating to the walking pace of patients – these were not issues relating to the use of the POMDP.

A key issue identified for POMDP’s is the computational limitations imposed on them. Real world systems can have a near infinite number of possible states, interactions and observations, however, most POMDP’s can only handle a small number (N Roy, G Gordon, S Thrun 2003). For the use of POMDP’s to become more widespread this technical limitation will need to be overcome or the systems being modelled will need to be simplified significantly – potentially at the cost of usefulness of the system.

## (iii)

**Lisp programming language**

Lisp is one of the oldest and most widely used programming languages for AI – even now it remains one of the most widely used programming languages for AI. It is a form of object oriented programming and objects can take the form of functions, data items or data structures. Initially Lisp was the preferred language to use for AI as it had sufficient flexibility to allow for easy experimentation and first introduced constructs like “if-then-else” (S Valencia 2017). These were the sort of constructs required for conditional decisions being taken by AI systems.

Now when looking at Lisp for AI, the key benefits of its application in AI is that (What is Lisp Good For?):

1. Remains highly customisable
2. It would be impossible to develop certain types of applications in other programming languages
3. Suitable for applications that will require continued availability even if changes are required after deployment

Another reason that Lisp is favoured for use in AI programming is that it allows for the programmer to adapt to their target solution and for inductive logic problems which can form part of AI decision making (A Melnichuk).

Despite these benefits there are a few key disadvantages to the Lisp programming language, namely (A Stephens, C Owens):

1. Difficult to read – doesn’t follow same conventions as other languages on semicolons for example
2. Not very secure as functions can change themselves
3. Commercial versions are expensive
4. No longer widely used

Despite its obvious strengths, it seems likely that the use of Lisp will continue to decline as alternatives like Java, Python and C++ are also widely used for the purpose of programming in AI.

## (iv)

**Memory Augmented Neural Networks**

Firstly, a neural network is a model of connected neurons or nodes (processing units) that aim to work together to solve a problem. A neural network learns by example so it is able to change its structure based on the training phase where training data is provided to the system (NUIG - Machine Learning (Part 3) Workshop 6 Section 1).

Memory Augmented Neural Networks (MANN) are a form of neural network where the processing is separated from the memory (mcollier). It will typically consist of three main parts (Afham M 2020):

1. Controller network – this is the standard neural network that will learn and make decisions/predictions based on the training data provided
2. External memory module – this will contain the memory of key parameters for tasks
3. Read – write heads – this is used to drive to communication between the controller network and the external memory module.

MANN is generally used for one-shot tasks where there are very few iterations or limited data available for training as it has shown to be more successful in doing so than other deep learning algorithms.

## (v)

**Ensemble Learning**

Ensemble learning is a machine learning method whereby you take a combination of multiple machine learning models with the objective of improving the overall accuracy. When using ensemble learning, you will need to have trained each of the models being combined on a subset of the training data (B Dickson 2020). When constructing the data for the training set this can be done by what is known as bootstrapping, i.e. taking samples from the dataset with replacement or by pasting i.e. taking samples from the data set without replacement.

Another ensemble learning method raised by Dickson is Boosting – that is where each model is built on the previous one with the aim of resolving any flaws or inefficiencies identified in the previous phase. This would mean that each iteration of an ensemble learning model would perform at least as well or better than the last.

The main advantages of Ensemble Learning are that it allows us to get to the best model for a particular scenario and improve predictive accuracy. However, ensemble learning can lead to models that are difficult to interpret and can be significantly more expensive and time consuming to construct (Kapalko).

Another interesting benefit of ensemble learning is the green benefits (K Gidney 2020). As ensemble learning can use smaller amounts of data and computing networks, the cost of running ensemble networks (both in terms of finance and emissions) will become more attractive to users.

## (i)

Probability is the mathematical method typically used to demonstrate uncertainty which is used in decision making, maths, economics and across industries.

Firstly, the use of probability allows us to assess and quantify uncertainty in a consistent and repeatable way – it is a quantitative rather than qualitative measure so it is less open to incorrect interpretations.

There are two main types of uncertainty in AI, namely epistemic uncertainty and aleatoric uncertainty. Epistemic uncertainty refers to uncertainty because of limitations of the model which could relate to date or other model knowledge – this can be reduced and managed by finding sufficient training data. Aleatoric uncertainty arises due to the uncertainty in observations and cannot be reduced by adding data (M Kana 2020). So the use of adequate data allows us to remove all unnecessary uncertainty.

Most AI models are predictive in nature and are designed, trained and tested using probabilistic methods (Brownlee) 2019. When building each step of a model, the level of uncertainty can be controlled by the designer, for example how will a system respond to certain prompts.

In the case of AI, it can be used to help illustrate the uncertainty at the three stages of an AI model:

1. Uncertainty of inputs – for example driving conditions for AI car
2. Uncertainty of decision – for example an AI car encountering a pedestrian and deciding whether to swerve, brake or do nothing (simplified set of choices)
3. Uncertainty of outputs – for example after decision taken are passengers uninjured/injured/dead is pedestrian uninjured/injured/dead

Probability allows for deterministic or stochastic models to be used to assess each stage. The use of stochastic models allow for uncertainty as a probability distribution can be used as an input (Brownlee 2019). As a probability distribution function is being used, its parameters can be flexed based on the views of the person building the AI model (or simply for testing).

With AI (as with people), their actions and decisions will depend entirely upon what they encounter – for example with an AI car, if it starts raining, the car will slow down – we could therefore use a probability model to determine the likely speed of a car based on our assessment of potential speed. By using probability theory, we can model complex relational interactions and assess the likelihood of each path being taken.

## (ii)

P(catch) = P(catchIcavity) + P(catch**¬cavity) = 0.108 + 0.072 + 0.016 + 0.144 = 0.34**

**P(catch) = (0.34, 0.66)**

P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = **0.20**

P(cavityIcatch) = P(cavity) / P(catch)

P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2

P(cavity) = (0.2, 0.8)

P(cavityIcatch) = P(cavity) / P(catch) = 0.2 / 0.34 = 0.5294

## (i)

**Depth First Search Algorithm**

**Depth first search (DFS) is an algorithm used for finding a path between vertices in a decision tree structure. Using DFS involves starting at the top level of the graph and moving down each level as far as possible before backtracking to find the next unchecked path. The algorithm repeats itself until all branches have been checked.**

**The main benefits of DFS are that it uses less memory and reaches the furthest layer faster, however it may check a number of unnecessary paths and will rarely find the optimal solution for a problem (java9pro).**

**The main uses of DFS are to solve puzzles with a single possible solution, job scheduling processes when there are dependencies between jobs and path finding processes (Geeksforgeeks).**

**The path taken to reach the stage Q would be**

**A, B, E, K, L, F, M, N, S, T, C, G, H, O, P, D, I, J, Q (19 nodes ignoring track backs, over 30 with trackbacks)**

## (ii)

**Breadth First Search Algorithm**

**Breadth first search (BFS) is an algorithm used for reaching a decision tree structure with the objective of finding an node that satisfies certain conditions. It is a form of graph traversal which follows a path where it checks all nodes on the same level as the current selected nodes before moving down to the next level (guru99). After checking a node it is marked as visited in the queue and stored to the memory. It is also a first (hackerearth). The BFS is best suited for solving puzzles – examples I found reference to include rubix cubes. The main benefits of BFS are that it will always find the optimal solution as it gets to the closes goal in the shortest time possible – there are no redundant paths followed. The main disadvantage of the BFS is that it uses a lot of data due to the number of nodes stored (java9pro).**

**The path taken to reach the stage Q would be**

**A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q (16 nodes visited)**

## (iii)

**Iterative Deepening Search Algorithm**

**Iterative deepening search (IDS) is a combination of the DFS and BFS algorithms. When running and IDS, it will go to a set depth only before returning to the original level and moving to an adjacent node before starting another depth first search to an agreed level. By setting a limit on the depth that is searched to it reduces the risk of following a very long path to a dead end (educative). As most of the nodes are at the lower levels, the overhead of running an IDS are relatively low. IDS is generally suitable to use when we have a large number of nodes to search but we don’t know how many levels deep it goes. The main benefits of IDS are that it will have a quick responsive time and greater efficiency when solutions are at lower levels. However, there is a high time costs and a large number of unnecessary calculations. It is also a better choice of algorithm when there is a high branching factor (educba)**

**Number of nodes visited depends on the assumed starting depth, if we say from level 0 to level 3 permitted, path would be:**

**A, B, E, K, L, F, M, N, C, G, H, O, P, D, I, Q (16 unique nodes visited)**

## (iv)

**Bi-directional Process**

**A bi-directional search process is one where the algorithm begins searching from the starting and target end node simultaneously to find the shortest path until they meet. This approach allows for significant improvements in search times and reduces the overheads in terms of resources used. However, it does require that the user knows the end state, it may also not be possible to search back through all states (educba). Bi-directional processes also require more significantly more processing power which can have a negative impact on performance. They most common use of bi-directional processes is in neural networks which are used for predictive forecasting models.**

**Number of nodes visited will depend on the process used.**

## Question 3

## (a)

The information gain of an attribute is the reduction in entropy from partitioning the data according to that attribute. The provide a quantitative measure of the quality of a split in the data.

The following table summarises the data provided and I have included a row identifier:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Sensor-XT | Motor-T | Level | Arrow | Rocket Connected |
| 1 | T | T-45 | one | A-North | Connected |
| 2 | F | T-45 | one | A-North | Fail |
| 3 | T | T-30 | two | A-North | Connected |
| 4 | T | T-40 | three | A-North | Fail |
| 5 | T | T-35 | three | A-West | Connected |
| 6 | F | T-30 | three | A-West | Fail |
| 7 | F | T-35 | two | A-South | Connected |
| 8 | T | T-35 | one | A-North | Connected |
| 9 | T | T-30 | one | A-West | Connected |
| 10 | T | T-40 | three | A-West | Fail |
| 11 | F | T-40 | one | A-West | Connected |
| 12 | F | T-40 | two | A-North | Connected |
| 13 | T | T-45 | two | A-West | Connected |
| 14 | F | T-40 | three | A-North | Fail |
| 15 | T | T-45 | one | A-West | Connected |
| 16 | F | T-30 | three | A-South | Fail |
| 17 | T | T-30 | three | A-South | Connected |
| 18 | T | T-45 | one | A-North | Fail |
| 19 | F | T-35 | two | A-West | Fail |

When reviewing the data, I noted that the rows highlighted in yellow above (row ID 1 and 18) have the same results for each of the variables but have conflicting results for the outcome Rocket Connected – 1 connected & 1 failed. These will need to be removed from the dataset.

Revised dataset is therefore:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Sensor-XT | Motor-T | Level | Arrow | Rocket Connected |
|  |  |  |  |  |  |
| 2 | F | T-45 | one | A-North | Fail |
| 3 | T | T-30 | two | A-North | Connected |
| 4 | T | T-40 | three | A-North | Fail |
| 5 | T | T-35 | three | A-West | Connected |
| 6 | F | T-30 | three | A-West | Fail |
| 7 | F | T-35 | two | A-South | Connected |
| 8 | T | T-35 | one | A-North | Connected |
| 9 | T | T-30 | one | A-West | Connected |
| 10 | T | T-40 | three | A-West | Fail |
| 11 | F | T-40 | one | A-West | Connected |
| 12 | F | T-40 | two | A-North | Connected |
| 13 | T | T-45 | two | A-West | Connected |
| 14 | F | T-40 | three | A-North | Fail |
| 15 | T | T-45 | one | A-West | Connected |
| 16 | F | T-30 | three | A-South | Fail |
| 17 | T | T-30 | three | A-South | Connected |
|  |  |  |  |  |  |
| 19 | F | T-35 | two | A-West | Fail |

Information Gain for Sensor-XT attribute

Sensor-XT has 11 results marked True, 9 after the conflicting data removed and 8 results marked false.

The formula to calculate information gain is:

Gain(S,A) = Ent(S) -

All logs are base 2 and expressed as log(number, base) in formulae.

Ent(S) = Ent([9T, 8F])

=-9/17 x log(9/17,2) – 8/17 x log (8/17,2)

= 0.4858 + 0.5117

= 0.9975

For Sensor-XT = T, Rocket Connected for 8 – this was reduced to 7 after data correction and failed for 3 – this was reduced to 2 after data correction

For Sensor-XT = F Rocket Connected for 3 and failed for 5

Ent(Sensor-XT = T) = Ent([7T, 2F])

=-7/9 x log(7/9, 2) – 2/9 x log(2/9, 2)

= 0.2820 + 0.4822

= 0.7642

Ent(Sensor-XT = F) = Ent([3T, 5F])

=-3/8 x log(3/8, 2) – 5/8 x log(5/8, 2)

=0.5306 + 0.4238

= 0.9544

**Information Gain**

0.9975 – 9/17 x 0.7642 – 8/17 x 0.9544

= 0.1438

Value for information gain is relatively low - which implies that the Sensor-TX attribute does not contribute significant information to our result

If I had not removed the inconsistent data, I would have gotten an information gain of 0.1013 – this makes sense intuitively as we would expect garbled data to be less informative.

## (b)

The dataset chosen for analysis is the Fisher Iris Data set which has been taken from:

<https://storm.cis.fordham.edu/~gweiss/data-mining/weka-data/iris.arff>

**About the dataset**

Dataset consists of 3 classes, each with 50 instances giving 150 instances in total.

There are 5 attributes in total, the length and width of the sepal and petal and the class.

Each class refers to a different type of iris plant (Iris-setosa, Iris-versicolor and Iris-virginica)

The complete dataset is included in appendix 3 for reference.

**About the algorithm**

The algorithm chosen is a classification via regression algorithm. The objective is to find a relationship between the class (iris type) and the other four attributes in the data set so that the model can predict the class that a particular instance belongs to. .

As we are looking for a linear relationship between attributes for each classification, it is reasonable to use a linear regression model for classification, however, if the dataset was looking at binary attributes, it may not be appropriate (jinglecode).

**Analysis of accuracy**

I have included a copy of the model output in Appendix 4 for references.

The classification via regression algorithm was able to correctly classify 98% of instances which would be considered an extremely high level of accuracy.

As explained by Parson, the Kappa Statistic tells us how accurate the algorthim is vs what was expected i.e compared to random chance – our value of 0.97 suggests it is more accurate than random chance. Generally we would need to look at other measures before we can say the Kappa statistic suggests the algorithm is accurate, however, given how close it is to 1, my view is unlikely to change.

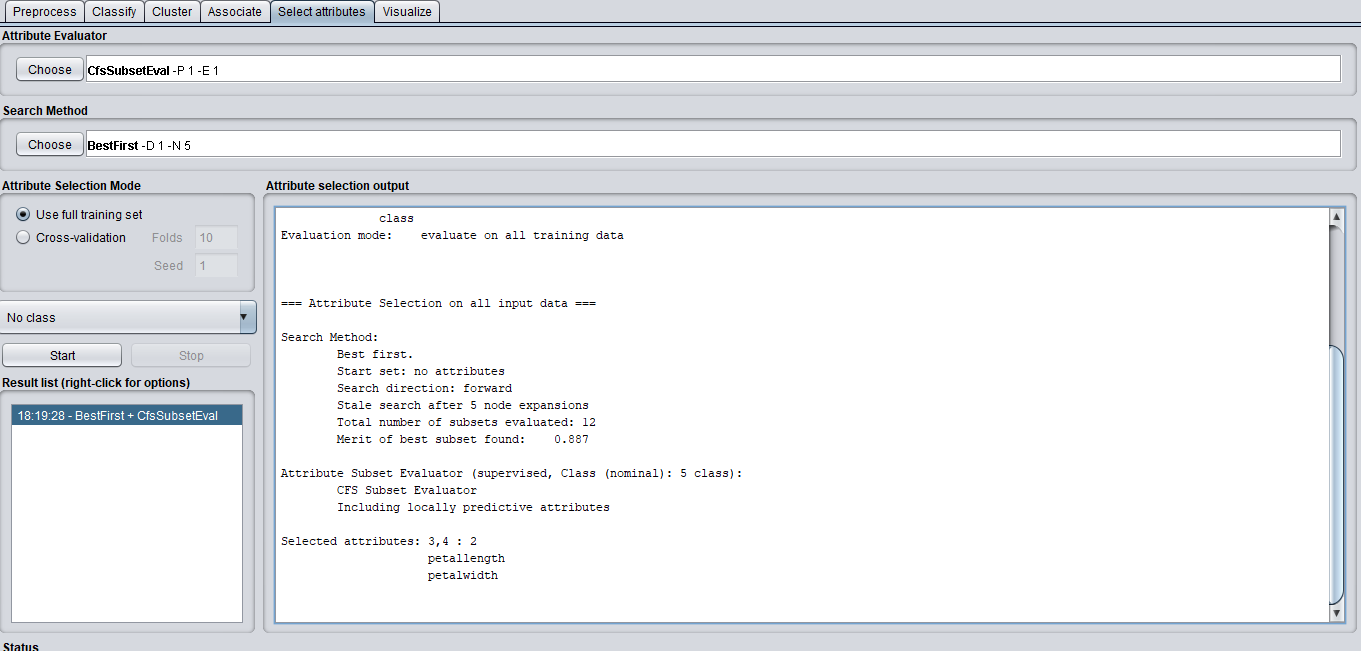
Mean absolute error being quite low, again demonstrates that predictions are largely correct. I also have a low root mean squared error of 0.1298 which suggests high accuracy

The Relative absolute error and root relative squared error have very low values as well which further reiterates that the algorithm has been accurate in classifying the data.

The confusion matrix summarises how good the model at correctly predicting each class.

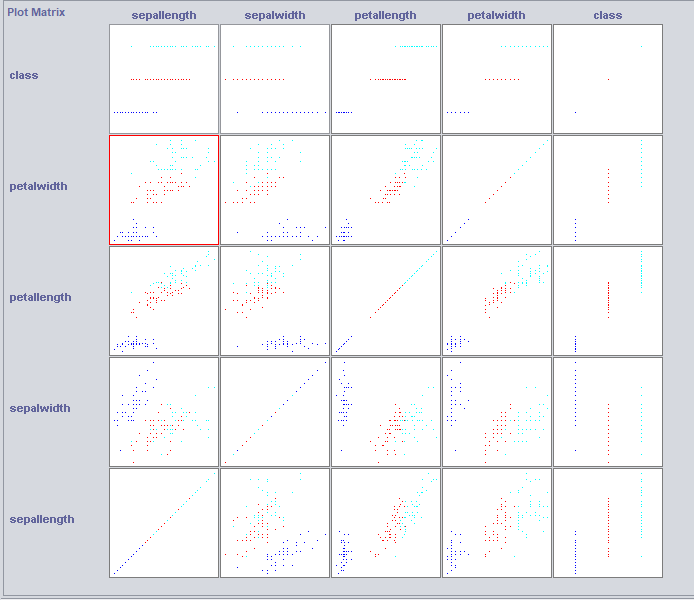
As per the outputs, the confusion matrix confirms the high accuracy of the algorithm with 50/50 classified correctly for the Iris-setosa, 48/50 classified correctly for the Iris-versicolor and 49/50 classified correctly for the Iris-virginica.

When we navigate to the select attribute tab and evaluate the attributes that give the best fit, it returns petal length and petal width. I tried removing 1 of the other attributes and re-running the classifier resulted in no change to the confusion matrix for the scenario when the sepal width is removed, while there is a slight drop when sepal length is removed.



When we look at the plot matrix of each attribute, we can see for class v petal length and class v petal width shows there are relatively clear divisions between the plot points for each iris type.

By contract, if you plot class against sepal length, you can see there are significant overlaps between the sepal length recorded for the different classes.



**Summary of learnings**

**The dataset is fairly simple and by visual inspection and sorting using excel, you can see the attributes that are best to identify the class of iris.**

**The choice of algorithm wasn’t particularly important in this case as the results are all largely similar.**

**A summary of the % correctly classified using different algorithms were:**

**Iterative Classifier Optimizer 96.6% correctly classified**

**Multi Class Classifier 96.6% correctly classified**

**J48 98% correctly classified**

**Random Forest 100% correctly classified**

**KStar 100% correctly classified**

**Logistic 98% correctly classified**

**Naïve Bayes 96% correctly classified**

## (c)

The main cloud machine learning services currently on offer in the market are:

* Amazon ML which is available on Amazon Web Services (AWS)
* Azure AI platform with Microsoft
* Google AI platform
* Watson Machine Learning from IBM

Having considered the options, I would recommend Amazon ML from Amazon Web Services. Firstly when comparing them with the main alternatives, it is the most versatile as it offers more services than the Google and IBM offerings (altexsoft).

**Established Platform**

One of the criteria for assessment that was referred to was that the platform be established and this is something that AWS satisfies as demonstrated by their current customer list (AWS Machine Learning Customers) – they currently provide services to some of the biggest names in entertainment, sport, finance and technology hardware with over 100k enterprise customers. As the AWS ML service is so ubiquitous across our industry, there is a broad depth of resources and knowledge already available for learning and training.

**Number of languages supported**

The Amazon Web Services umbrella provides support to the most widely used languages such as Java, Python, Ruby meaning it will be versatile enough to work for most projects.

**Integration**

The AWS AI service can be integrated to business applications and can include human oversight of predictions (Augmented AI). The Machine Learning services are also integrated with the rest of the AWS platform

**Flexibility & Accessibility**

The Amazon Machine Learning offering is highly scalable and allows for the customers to only pay for what they use and can allow developers with little or no data science knowledge (H Reese 2016). It also includes various visual aids and analytics to help users better understand what the machine learning algorithm is doing.

**Security**

Amazon’s AI offering now includes a rapid threat detection system which protects against threats to the AWS Infrastructure, Platform and cloud services (Malenfant 2021) which reduce the risk related to any updates released from the cloud.

## Question 4

## (a)

I have read a summary of the EU press release and think the first thing that needs to be said is that their announcement to create a legal framework to govern AI is to be welcomed. To date, the United States has not yet implemented any comprehensive national legislation on AI (Zhu) meaning the proposed rules being announced by the EU will likely become the new de-facto industry standards.

There are a few (occasionally overlapping) areas that I would like to focus in on which I have broken into some broad headings for my discussion

**EU Proposed Rules for AI**

The EU press release confirms there will be a co-ordinated plan across member states in terms of investment and policy. By having uniformity across so many countries, should, I would expect make investment in AI innovation easier to justify for businesses. It also ensures there is a critical mass of potential customers for them to sell in to (assuming compliance. All of this largely speaks to the first principal referred to in their co-ordinated plan, namely the creation of conditions suitable for AI’s development and the investment it will attract will help them with their second objective of fostering excellence in AI.

The third principal referred to is ensuring that AI works for the people that use it – this is where we need to start looking in a more granular detail.

**Trust in AI**

A 2018 survey by statista showed that trust in AI for EU nations was relatively low, for example just 21% of Germans stating that they do trust AI so the EU press release feels like they have measured the temperature across national capitals and realised that proactive action is needed. When looking for how to improve this, I felt that the factors proposed by Naveen Joshi (Forbes 2019) seem like sensible building blocks that warrant some discussion.

The first is explainability – for many, AI is like a genie in a lamp that performs magic for us. By making it more explainable, it will mean that users with a better understanding of how their decisions may impact a result can have greater confidence and trust in the system being used. An example of where AI is used in an explainable manner is to generate online insurance quotes for customers (G Clarke 2019). As users are already familiar with how an insurance policy pricing works, they can feel confident they can trust the algorithm just by testing how certain changes impact the results (for example, check what happens when you get test quotes with 0 up to 10 penalty points). Improvements in explainability will require increased education of users by both the businesses that use AI and regulators. Explainability can also be improved by trying to use existing well understood concepts when developing an AI system, however, this would significantly stifle innovation and creativity in a very young and developing industry – this would also go against the suggested EU principal of fostering excellence.

Another area Joshi raised that can help increase trust in AI systems is the idea of machine learning integrity – this would mean that it does exactly as the developer intended. I have difficulty coming on board with this one as it totally discounts the notion of malicious actors who could create an AI system with less than noble intentions. A relatively high-profile example of this is Cambridge Analytica and its use of AI to influence voters by using specifically tailored advertisements (J Deckler 2020). This also ties in neatly with the concept of conscious development – by ensuring this principle is followed, any concerns relating to integrity can fall away.

Joshi also noted that results must be reproducible, to me this would mean that the AI model has been well tested and validated to ensure it is accurate.

The final factor raised by Joshi is regulation which is precisely what the EU have now brought into being so this box is now being ticked off. It will be interesting to see if they have any meaningful impact on the next statista survey.

**Risk levels**

The proposed framework of following a risk based approach with risk posed by AI being classified on a scale ranging from unacceptable down to high risk, limited risk then minimal risk does appear to be a sensible approach as a starting point, however, I think there is significant further thought that needs to go into the classification system.

For example, an unacceptable risk referred to is social scoring, however, this may be something that financial institutions use for risk management (M Gromek 2018) meaning a lot of these profiles are being built anyway. I am not sure that it is sensible to allow social scoring AI at all unless the responsibility for its use resides with the customer rather than the developer. I also found it noteworthy that national militaries (which are under direct control of governments) will be exempt from these rules (N Lomas 2021).

There is also recognition within the rules for when an AI system is changed substantially by the user instead of the developer, in the case of high-risk systems, responsibility for any regulatory compliance would move to the user in those situations (M MacCarthy, K Propp 2021). I haven’t had an opportunity to read the complete regulations and related recitals, however, there would need to be an unambiguous definition given for what is considered substantial change.

**Future of AI**

While AI is continuing to evolve, there are a number of areas where the EU has remained silent but will need to consider very soon.

Firstly, the economic impact of AI. A paper from the European Parliament Research Service suggested that global GDP could increase by the order of 14% as a result of AI advancements. With this level of advancement, there is also likely to be a material impact on the labour markets as advancements in automation could incentivise companies to significantly reduce their workforce moving large segments of the labour market to unemployment.

Any developments in AI will also have a spill over into other longer established industries. One such example is insurance where there will now be a demand for a new type of insurance to protect against AI risks (RSS Kumar, K Nagle 2020). Perhaps one of the biggest emerging lines will be AI car insurance.

While there is still some time to go before autonomous vehicles become commonplace, it is vital that the infrastructure is in place from industries like insurance when they are ready to go live.

Another area where EU regulation may need to reach would be in the creation of a regulatory body that could help the AI industry to develop codes of practice or ethical standards expected from participants. This would ensure that those working in the industry are properly versed in what is required from them and could help contribute to improved trust from users.

**Conclusion**

The regulation and legislation around the use of AI is still in its infancy and the areas that will be most impacted are still not fully known or identified. As we learn more about the development of the AI industry and what it can possibly achieve, it will be vital that any regulations put in place are continuously evolved and developed at the same pace as the industry itself.

## (b)

Prolog code to be completed

## Appendix 1

**Question 1**

**Prolog code**

male(john).

male(bob).

male(fred).

male(darren).

female(elizabeth).

female(susan).

female(jane).

female(anne).

parent(john, anne).

parent(elizabeth, anne).

parent(john, darren).

parent(elizabeth, darren).

parent(bob, fred).

parent(anne, fred).

parent(darren, jane).

parent(susan, jane).

married(elizabeth, john).

married(bob, anne).

married(susan, darren).

married(john, elizabeth).

married(anne, anne).

married(darren, susan).

father(john, darren).

father(john, anne).

father(darren, jane).

father(bob, fred).

mother(elizabeth, darren).

mother(elizabeth, anne).

mother(susan, jane).

mother(anne, fred).

father(X,Y):- male(X),

parent(X,Y).

mother(X,Y):- female(X),

parent(X,Y).

grandfather(X,Y):- father(X,Z),

parent(Z,Y).

is\_married(X,Y):- father(X,Z), mother(Y,Z).

is\_married(X,Y):- mother(X,Z), father(Y,Z).

## Appendix 2

**Question 2**

**References**

CT621 Artificial Intelligence (2021) Machine Learning (Part 2) Workshop 6 Section 2

Available at: NUIG Blackboard

Accessed 10th July 2021

Tanwar S (2019) Markov chains and Markov Decision process

Available at: <https://medium.com/@sanchittanwar75/markov-chains-and-markov-decision-process-e91cda7fa8f2>

Accessed 10th July 2021

DeepAI (undated) Markov Decision Process

Available at: <https://deepai.org/machine-learning-glossary-and-terms/markov-decision-process>

Accessed 10th July 2021

Geeksforgeeks (2018) Markov Decision Process

Available at: <https://www.geeksforgeeks.org/markov-decision-process/>

Accessed 10th July 2021

Brown University (undated) Background on POMDPs

Available at: https://cs.brown.edu/research/ai/pomdp/tutorial/pomdp-background.html

Accessed 10th July 2021

Kamalzadeh H, Hahsler M (2019) POMDP: Introduction to Partially Observable Markov Decision Processes

Available at: <https://cran.r-project.org/web/packages/pomdp/vignettes/POMDP.html>

Accessed 10th July 2021

Pineau J, Thrun S (2002) High-level robot behavior control using POMDPs

Available at: <https://www.aaai.org/Papers/Workshops/2002/WS-02-05/WS02-05-012.pdf>

Accessed 10th July 2021

Roy N, Gordon G, Thrun S (2003) Finding Approximate POMDP Solutions Through Belief Compression

Available at: <http://robots.stanford.edu/papers/Roy04a.pdf>

Accessed 10th July 2021

Valencia S (2017) The Lisp approach to AI (Part 1)

Available at: <https://medium.com/ai-society/the-lisp-approach-to-ai-part-1-a48c7385a913>

Accessed 10th July 2021

Melnichhuk A (2019) AI Programming: 5 Most Popular AI Programming Languages

Available at: <https://ncube.com/blog/ai-programming-languages>

Accessed 10th July 2021

Stephen A, Owen C (undated) Lisp

Available at: <http://www.ai.sri.com/~delacaze/alu-site/alu/table/good.htm>

Accessed 10th July 2021

Available at: <https://www2.cs.siu.edu/~rahimi/cs311/langpresentations/LISP.pptm>

Accessed 10th July 2021

CT621 Artificial Intelligence (2021) Machine Learning (Part 3) Workshop 6 Section 1

Available at: NUIG Blackboard

Accessed 10th July 2021

Afham M (2020) Memory Augmented Neural Network for Meta Learning — Case Study

Available at: <https://medium.com/the-ai-team/memory-augmented-neural-network-for-meta-learning-case-study-56af9cc81ae2>

Accessed 10th July 2021

Mcollier (2019) Memory-Augmented Neural Networks for Machine Translation (Pre-Print)

Available at: <https://isg.beel.org/blog/2019/06/24/memory-augmented-neural-networks-for-machine-translation-pre-print/>

Accessed 10th July 2021

Santoro (2016) Explanation of

One-shot Learning with Memory-Augmented Neural Networks

Available at: <https://rylanschaeffer.github.io/content/research/one_shot_learning_with_memory_augmented_nn/main.html>

Accessed 10th July 2021

Dickson B (2020) What is ensemble learning?

Available at: <https://bdtechtalks.com/2020/11/12/what-is-ensemble-learning/>

Accessed 10th July 2021

Kapalko R (2019) Predictive Model Ensembles: Pros and Cons

Available at: <https://blogs.perficient.com/2019/11/07/predictive-model-ensembles-pros-and-cons/>

Accessed 10th July 2021

Gidney K (2020) HOW ENSEMBLES CAN REDUCE MACHINE LEARNING’S CARBON FOOTPRINT

Available at: <https://dataconomy.com/2020/03/how-ensembles-can-reduce-machine-learnings-carbon-footprint/>

Accessed 10th July 2021

Kana M (20020 Uncertainty in Deep Learning. How to Measure

Available at: <https://towardsdatascience.com/my-deep-learning-model-says-sorry-i-dont-know-the-answer-that-s-absolutely-ok-50ffa562cb0b>

Accessed 10th July 2021

Brownlee J (2019) 5 Reasons to Learn Probability for Machine Learning

Available at: https://machinelearningmastery.com/why-learn-probability-for-machine-learning/

Accessed 10th July 2021

Brownlee J (2019) What Does Stochastic Mean in Machine Learning

Available at: https://machinelearningmastery.com/stochastic-in-machine-learning/

Accessed 10th July 2021

Guru99 (undated) Breadth First Search (BFS) Algorithm with Example

Available at: https://www.guru99.com/breadth-first-search-bfs-graph-example.html

Accessed 10th July 2021

hackerearth (undated) Breadth First Search

Available at: https://www.hackerearth.com/practice/algorithms/graphs/breadth-first-search/tutorial/

Accessed 10th July 2021

Jva9pro (2017) Advantages and disadvantages of DFS and BFS

Available at: https://java9pro.blogspot.com/2017/08/advantages-and-disadvantages-of-dfs-and.html

Accessed 10th July 2021

Moore K, Jennison K, Khim J (undated) Depth First Search (DFS)

Available at: https://brilliant.org/wiki/depth-first-search-dfs/

Accessed 10th July 2021

Geeksforgeeks (2021) Applications of Depth First Search

Available at: https://www.geeksforgeeks.org/applications-of-depth-first-search/  
Accessed 10th July 2021

Educative (undated) What is Iterative Deepening Search

Available at: <https://www.educative.io/edpresso/what-is-iterative-deepening-search>

Accessed 10th July 2021

Pedamkar P (undated) Iterative Deepening Depth-First Search

Available at: https://www.educba.com/iterative-deepening-depth-first-search/

Accessed 10th July 2021

Pedamkar P (undated) Bidirectional search

Available at: https://www.educba.com/bidirectional-search/

Accessed 10th July 2021

jinglescode (2020) Why Linear Regression is not suitable for Classification

Available at: <https://jinglescode.github.io/2019/05/07/why-linear-regression-is-not-suitable-for-classification/>

Accessed 10th July 2021

Mlr (undated) Classification via regression wrapper.

Available at: <https://mlr.mlr-org.com/reference/makeClassificationViaRegressionWrapper.html>

Accessed 10th July 2021

Parson Dr D (undated) CSC 558 – Data Mining and Predictive Analysis

Available at: <http://faculty.kutztown.edu/parson/fall2019/Fall2019Kappa.html#:~:text=%E2%80%9CThe%20Kappa%20statistic%20(or%20value,to%20evaluate%20classifiers%20amongst%20themselves>.

Accessed 10th July 2021

Altexsoft (2021) Comparing Machine Learning as a Service: Amazon, Microsoft Azure, Google Cloud AI, IBM Watson

Available at: <https://www.altexsoft.com/blog/datascience/comparing-machine-learning-as-a-service-amazon-microsoft-azure-google-cloud-ai-ibm-watson/>

Accessed 10th July 2021

AWS Machine Learning Customers

Available at: [https://aws.amazon.com/machine-learning/customers/?c=ml&sec=cst&customer-references-cards.sort-by=item.additionalFields.publishedDate&customer-references-cards.sort-order=desc&awsf.customer-references-location=\*all&awsf.customer-references-segment=\*all&awsf.customer-references-industry=\*all](https://aws.amazon.com/machine-learning/customers/?c=ml&sec=cst&customer-references-cards.sort-by=item.additionalFields.publishedDate&customer-references-cards.sort-order=desc&awsf.customer-references-location=*all&awsf.customer-references-segment=*all&awsf.customer-references-industry=*all)

Accessed 10th July 2021

Ziegler JK (2014) Choosing the Right Programming Language for your Startup

Available at: <https://aws.amazon.com/blogs/startups/choosing-the-right-programming-language-for-your-startup/>

Accessed 10th July 2021

Explore AWS AI Services

Available at: <https://aws.amazon.com/machine-learning/ai-services/>

Accessed 10th July 2021

Augmented AI

Available at: <https://aws.amazon.com/augmented-ai/>

Accessed 10th July 2021

Reese H (2016) Should Amazon be your AI and machine learning platform?

Available at: <https://www.zdnet.com/article/should-amazon-be-your-ai-and-machine-learning-platform/>

Accessed 10th July 2021

Malenfant J (2021) Vectra Introduces Detect for AWS:Threat Detection and Response for IaaS and PaaS

Available at: <https://www.vectra.ai/blogpost/detect-for-aws>

Accessed 10th July 2021

Zhu C (2021) United States: Artificial Intelligence Comparative Guide

Available at: https://www.mondaq.com/unitedstates/technology/1059776/artificial-intelligence-comparative-guide

Accessed 10th July 2021

European Commission (2021) Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence

Available at: https://ec.europa.eu/commission/presscorner/detail/en/ip\_21\_1682

Accessed 10th July 2021

statista (2018) Trust in artificial intelligence by country 2018

Available at: https://www.statista.com/statistics/948531/trust-artificial-intelligence-country/

Accessed 10th July 2021

Joshi N (2019) How We Can Build Trustworthy AI

Available at: https://www.forbes.com/sites/cognitiveworld/2019/07/30/how-we-can-build-trustworthy-ai/?sh=3e977fd3d5e3

Accessed 10th July 2021

Clarke G (2019) How Your Insurance Quote Is Powered By Artificial Intelligence

Available at: https://www.forbes.com/sites/ginaclarke/2019/01/21/how-your-insurance-quote-is-powered-by-artificial-intelligence/?sh=62b2a9942811

Accessed 10th July 2021

Deckler J (2020) POLITICO AI: Decoded: How Cambridge Analytica used AI

Available at: https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-how-cambridge-analytica-used-ai-no-google-didnt-call-for-a-ban-on-face-recognition-restricting-ai-exports/

Accessed 10th July 2021

Lomas N (2021) Europe lays out plan for risk-based AI rules to boost trust and uptake

Available at: https://techcrunch.com/2021/04/21/europe-lays-out-plan-for-risk-based-ai-rules-to-boost-trust-and-uptake/?guccounter=1&guce\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce\_referrer\_sig=AQAAAJcMcNfm\_k1YeNMn-iwU1Lr5MQarF5oo8FVF4-pivD7HjQtPx36poWcKACQpRCkst8hkGLEoXBR7zERSFIuzqOhKcdgPhUEEdrOpukKojASn9X1atdk2TuIqey-Oy4Y-JS-TrFCWkh6DKliPBMXWKXsFArRGYy\_IGooKZ4GgSo\_j

Accessed 10th July 2021

Gromek M (2018) Social Scoring In Finance: We Are Partly Already There

Available at: https://www.forbes.com/sites/michalgromek/2018/07/25/social-scoring-in-finance-we-are-partly-already-there/?sh=3b8e8e3360bd

Accessed 10th July 2021

MacCarthy M, Propp K (2021) Machines learn that Brussels writes the rules: The EU’s new AI regulation

Available at: https://www.brookings.edu/blog/techtank/2021/05/04/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/

Accessed 10th July 2021

European Parliment (2019) Economic impacts of artificial intelligence (AI)

Available at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS\_BRI(2019)637967\_EN.pdf

Accessed 10th July 2021

Kumar RSS, Nagle F (2020) The Case for AI Insurance

Available at: https://hbr.org/2020/04/the-case-for-ai-insurance

Accessed 10th July 2021

## Appendix 3

**Question 3**

**Iris dataset**

@RELATION iris

@ATTRIBUTE sepallength REAL

@ATTRIBUTE sepalwidth REAL

@ATTRIBUTE petallength REAL

@ATTRIBUTE petalwidth REAL

@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}

@DATA

5.1,3.5,1.4,0.2,Iris-setosa

4.9,3.0,1.4,0.2,Iris-setosa

4.7,3.2,1.3,0.2,Iris-setosa

4.6,3.1,1.5,0.2,Iris-setosa

5.0,3.6,1.4,0.2,Iris-setosa

5.4,3.9,1.7,0.4,Iris-setosa

4.6,3.4,1.4,0.3,Iris-setosa

5.0,3.4,1.5,0.2,Iris-setosa

4.4,2.9,1.4,0.2,Iris-setosa

4.9,3.1,1.5,0.1,Iris-setosa

5.4,3.7,1.5,0.2,Iris-setosa

4.8,3.4,1.6,0.2,Iris-setosa

4.8,3.0,1.4,0.1,Iris-setosa

4.3,3.0,1.1,0.1,Iris-setosa

5.8,4.0,1.2,0.2,Iris-setosa

5.7,4.4,1.5,0.4,Iris-setosa

5.4,3.9,1.3,0.4,Iris-setosa

5.1,3.5,1.4,0.3,Iris-setosa

5.7,3.8,1.7,0.3,Iris-setosa

5.1,3.8,1.5,0.3,Iris-setosa

5.4,3.4,1.7,0.2,Iris-setosa

5.1,3.7,1.5,0.4,Iris-setosa

4.6,3.6,1.0,0.2,Iris-setosa

5.1,3.3,1.7,0.5,Iris-setosa

4.8,3.4,1.9,0.2,Iris-setosa

5.0,3.0,1.6,0.2,Iris-setosa

5.0,3.4,1.6,0.4,Iris-setosa

5.2,3.5,1.5,0.2,Iris-setosa

5.2,3.4,1.4,0.2,Iris-setosa

4.7,3.2,1.6,0.2,Iris-setosa

4.8,3.1,1.6,0.2,Iris-setosa

5.4,3.4,1.5,0.4,Iris-setosa

5.2,4.1,1.5,0.1,Iris-setosa

5.5,4.2,1.4,0.2,Iris-setosa

4.9,3.1,1.5,0.1,Iris-setosa

5.0,3.2,1.2,0.2,Iris-setosa

5.5,3.5,1.3,0.2,Iris-setosa

4.9,3.1,1.5,0.1,Iris-setosa

4.4,3.0,1.3,0.2,Iris-setosa

5.1,3.4,1.5,0.2,Iris-setosa

5.0,3.5,1.3,0.3,Iris-setosa

4.5,2.3,1.3,0.3,Iris-setosa

4.4,3.2,1.3,0.2,Iris-setosa

5.0,3.5,1.6,0.6,Iris-setosa

5.1,3.8,1.9,0.4,Iris-setosa

4.8,3.0,1.4,0.3,Iris-setosa

5.1,3.8,1.6,0.2,Iris-setosa

4.6,3.2,1.4,0.2,Iris-setosa

5.3,3.7,1.5,0.2,Iris-setosa

5.0,3.3,1.4,0.2,Iris-setosa

7.0,3.2,4.7,1.4,Iris-versicolor

6.4,3.2,4.5,1.5,Iris-versicolor

6.9,3.1,4.9,1.5,Iris-versicolor

5.5,2.3,4.0,1.3,Iris-versicolor

6.5,2.8,4.6,1.5,Iris-versicolor

5.7,2.8,4.5,1.3,Iris-versicolor

6.3,3.3,4.7,1.6,Iris-versicolor

4.9,2.4,3.3,1.0,Iris-versicolor

6.6,2.9,4.6,1.3,Iris-versicolor

5.2,2.7,3.9,1.4,Iris-versicolor

5.0,2.0,3.5,1.0,Iris-versicolor

5.9,3.0,4.2,1.5,Iris-versicolor

6.0,2.2,4.0,1.0,Iris-versicolor

6.1,2.9,4.7,1.4,Iris-versicolor

5.6,2.9,3.6,1.3,Iris-versicolor

6.7,3.1,4.4,1.4,Iris-versicolor

5.6,3.0,4.5,1.5,Iris-versicolor

5.8,2.7,4.1,1.0,Iris-versicolor

6.2,2.2,4.5,1.5,Iris-versicolor

5.6,2.5,3.9,1.1,Iris-versicolor

5.9,3.2,4.8,1.8,Iris-versicolor

6.1,2.8,4.0,1.3,Iris-versicolor

6.3,2.5,4.9,1.5,Iris-versicolor

6.1,2.8,4.7,1.2,Iris-versicolor

6.4,2.9,4.3,1.3,Iris-versicolor

6.6,3.0,4.4,1.4,Iris-versicolor

6.8,2.8,4.8,1.4,Iris-versicolor

6.7,3.0,5.0,1.7,Iris-versicolor

6.0,2.9,4.5,1.5,Iris-versicolor

5.7,2.6,3.5,1.0,Iris-versicolor

5.5,2.4,3.8,1.1,Iris-versicolor

5.5,2.4,3.7,1.0,Iris-versicolor

5.8,2.7,3.9,1.2,Iris-versicolor

6.0,2.7,5.1,1.6,Iris-versicolor

5.4,3.0,4.5,1.5,Iris-versicolor

6.0,3.4,4.5,1.6,Iris-versicolor

6.7,3.1,4.7,1.5,Iris-versicolor

6.3,2.3,4.4,1.3,Iris-versicolor

5.6,3.0,4.1,1.3,Iris-versicolor

5.5,2.5,4.0,1.3,Iris-versicolor

5.5,2.6,4.4,1.2,Iris-versicolor

6.1,3.0,4.6,1.4,Iris-versicolor

5.8,2.6,4.0,1.2,Iris-versicolor

5.0,2.3,3.3,1.0,Iris-versicolor

5.6,2.7,4.2,1.3,Iris-versicolor

5.7,3.0,4.2,1.2,Iris-versicolor

5.7,2.9,4.2,1.3,Iris-versicolor

6.2,2.9,4.3,1.3,Iris-versicolor

5.1,2.5,3.0,1.1,Iris-versicolor

5.7,2.8,4.1,1.3,Iris-versicolor

6.3,3.3,6.0,2.5,Iris-virginica

5.8,2.7,5.1,1.9,Iris-virginica

7.1,3.0,5.9,2.1,Iris-virginica

6.3,2.9,5.6,1.8,Iris-virginica

6.5,3.0,5.8,2.2,Iris-virginica

7.6,3.0,6.6,2.1,Iris-virginica

4.9,2.5,4.5,1.7,Iris-virginica

7.3,2.9,6.3,1.8,Iris-virginica

6.7,2.5,5.8,1.8,Iris-virginica

7.2,3.6,6.1,2.5,Iris-virginica

6.5,3.2,5.1,2.0,Iris-virginica

6.4,2.7,5.3,1.9,Iris-virginica

6.8,3.0,5.5,2.1,Iris-virginica

5.7,2.5,5.0,2.0,Iris-virginica

5.8,2.8,5.1,2.4,Iris-virginica

6.4,3.2,5.3,2.3,Iris-virginica

6.5,3.0,5.5,1.8,Iris-virginica

7.7,3.8,6.7,2.2,Iris-virginica

7.7,2.6,6.9,2.3,Iris-virginica

6.0,2.2,5.0,1.5,Iris-virginica

6.9,3.2,5.7,2.3,Iris-virginica

5.6,2.8,4.9,2.0,Iris-virginica

7.7,2.8,6.7,2.0,Iris-virginica

6.3,2.7,4.9,1.8,Iris-virginica

6.7,3.3,5.7,2.1,Iris-virginica

7.2,3.2,6.0,1.8,Iris-virginica

6.2,2.8,4.8,1.8,Iris-virginica

6.1,3.0,4.9,1.8,Iris-virginica

6.4,2.8,5.6,2.1,Iris-virginica

7.2,3.0,5.8,1.6,Iris-virginica

7.4,2.8,6.1,1.9,Iris-virginica

7.9,3.8,6.4,2.0,Iris-virginica

6.4,2.8,5.6,2.2,Iris-virginica

6.3,2.8,5.1,1.5,Iris-virginica

6.1,2.6,5.6,1.4,Iris-virginica

7.7,3.0,6.1,2.3,Iris-virginica

6.3,3.4,5.6,2.4,Iris-virginica

6.4,3.1,5.5,1.8,Iris-virginica

6.0,3.0,4.8,1.8,Iris-virginica

6.9,3.1,5.4,2.1,Iris-virginica

6.7,3.1,5.6,2.4,Iris-virginica

6.9,3.1,5.1,2.3,Iris-virginica

5.8,2.7,5.1,1.9,Iris-virginica

6.8,3.2,5.9,2.3,Iris-virginica

6.7,3.3,5.7,2.5,Iris-virginica

6.7,3.0,5.2,2.3,Iris-virginica

6.3,2.5,5.0,1.9,Iris-virginica

6.5,3.0,5.2,2.0,Iris-virginica

6.2,3.4,5.4,2.3,Iris-virginica

5.9,3.0,5.1,1.8,Iris-virginica

%

%

%

## Appendix 4

**Question 3**

**Weka Output**

=== Run information ===

Scheme: weka.classifiers.meta.ClassificationViaRegression -W weka.classifiers.trees.M5P -- -M 4.0 -num-decimal-places 4

Relation: iris

Instances: 150

Attributes: 5

sepallength

sepalwidth

petallength

petalwidth

class

Test mode: evaluate on training data

=== Classifier model (full training set) ===

Classification via Regression

Classifier for class with index 0:

M5 pruned model tree:

(using smoothed linear models)

petallength <= 2.45 : LM1 (50/0%)

petallength > 2.45 : LM2 (100/0%)

LM num: 1

class =

-0.0571 \* petallength

+ 1.0607

LM num: 2

class =

-0.0323 \* petallength

+ 0.1647

Number of Rules : 2

Classifier for class with index 1:

M5 pruned model tree:

(using smoothed linear models)

petallength <= 2.45 : LM1 (50/0%)

petallength > 2.45 :

| petallength <= 5.15 : LM2 (66/51.632%)

| petallength > 5.15 : LM3 (34/0%)

LM num: 1

class =

-0.105 \* sepalwidth

+ 0.0469 \* petallength

- 0.1087 \* petalwidth

+ 0.3515

LM num: 2

class =

0.0316 \* sepallength

+ 0.3442 \* sepalwidth

- 0.0354 \* petallength

- 1.0351 \* petalwidth

+ 1.2448

LM num: 3

class =

0.0522 \* sepallength

+ 0.0225 \* sepalwidth

- 0.0758 \* petallength

- 0.2432 \* petalwidth

+ 0.5763

Number of Rules : 3

Classifier for class with index 2:

M5 pruned model tree:

(using smoothed linear models)

petallength <= 4.45 : LM1 (79/0%)

petallength > 4.45 :

| petallength <= 5.15 : LM2 (37/61.258%)

| petallength > 5.15 : LM3 (34/0%)

LM num: 1

class =

0.0761 \* petalwidth

- 0.038

LM num: 2

class =

-0.2629 \* sepallength

+ 0.6777 \* petallength

+ 0.678 \* petalwidth

- 2.3084

LM num: 3

class =

-0.0827 \* sepallength

+ 0.1333 \* petallength

+ 0.2223 \* petalwidth

+ 0.2808

Number of Rules : 3

Time taken to build model: 0.56 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.01 seconds

=== Summary ===

Correctly Classified Instances 147 98 %

Incorrectly Classified Instances 3 2 %

Kappa statistic 0.97

Mean absolute error 0.0712

Root mean squared error 0.1298

Relative absolute error 16.0108 %

Root relative squared error 27.5407 %

Total Number of Instances 150

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

1,000 0,000 1,000 1,000 1,000 1,000 1,000 1,000 Iris-setosa

0,960 0,010 0,980 0,960 0,970 0,955 0,999 0,998 Iris-versicolor

0,980 0,020 0,961 0,980 0,970 0,955 0,999 0,998 Iris-virginica

Weighted Avg. 0,980 0,010 0,980 0,980 0,980 0,970 0,999 0,999

=== Confusion Matrix ===

a b c <-- classified as

50 0 0 | a = Iris-setosa

0 48 2 | b = Iris-versicolor

0 1 49 | c = Iris-virginica