INDEPENDENT STUDY [6V81 Spring 2018]

Computational Ethics : Search and Rescue Robots

MEHAK BERI

[mxb166430@utdallas.edu]

MENTOR: PROFESSOR CHRIS IRWIN DAVIS

[chrisirwindavis@utdalla.edu]

INDEX

Sr. No.	Topic	Page Number
1	Abstract	3
2	Introduction	3
3	Approach	4
4	Experiment description	6
5	Analysis	9
6	Conclusion	14
7	Future Work	15
8	References	16

1. Abstract

In this research, artificial intelligence and machine learning techniques were used to model human decisions in search and rescue robots. The robots can make human-like decisions to prioritize treatment of injured victims with 63% accuracy. To model human decisions, 328 responses to a questionnaire emulating different scenarios, including TRIAGE injury levels of the victims, was used. Multiple machine learning techniques were used to learn from the data and corresponding models were evaluated on the data set. Out of the used algorithms, Support vector machines and Fast decision trees perform better than their counterparts and their results have been reported.

2. Introduction

If there were two injured people on a battlefield, one a female soldier aged 25, and another a male soldier aged 35, and you had the resources to attend to one person at a time, who would you attend to first?

Ethical dilemmas are hard to solve for humans and more so, for machines. The concept of ethics in machines is still a grey area of research with no right answers. Essentially, we want an intelligent machine to make decisions which are as "human" as possible, including the human perception of an ethical situation.

Wikipedia defines computational ethics as: “Machine ethics (or machine morality, computational morality, or computational ethics) is a part of the ethics of artificial intelligence concerned with the moral behavior of artificially intelligent beings.”

In this research, the problem of computational ethics for Search and Rescue robots has been tackled using Machine learning techniques and human annotated data. Nearly 350 people answered a questionnaire posing them ethical dilemmas in situations where a rescue robot could prioritize injured victims, given their injury level, gender, age and status (civilian/ doctor/ soldier). The research brought out not only solutions as accurate as 63%, but also, the prejudices we carry with us as a society.

3. Approach

In order to combat the problem of computational ethics in search and rescue robot domain, we imagined the following scenarios where ethical dilemmas might arise:

- Natural Disaster in the city
- Post battle combat zone
- Hospital during natural disaster

Situation:

There is a rescue robot, whose task is to save as many lives as possible in case of a mass casualty incident. Unfortunately, the robot can only attend to one person at a time. You need to provide judgement on who a rescue robot should attend to first, provided the injury level and the type/ condition of person A and person B.

Injury level 1 implies that the patient needs immediate attention. It is the most severe degree of injury. It is followed by Injury level 2, 3 and 4, where a person with Injury level 4 can survive even with delayed medical care.

Each situation was modeled to incorporate specific features, which were later converted into data points.

Features

- Age: Continuous valued numbers representing ages of victim A and victim B
- Gender: Discrete values representing the gender of the victim. For the purpose of data refinement, male was taken as '0' and female as '1'. Pregnant female was represented as '2'
- Status/features: Special features for each of the victim. The feature was situation specific. In the post battle combat zone, a soldier was given feature value of 1, whereas a civilian was given a feature value of 0. Similarly, a doctor was given a feature value of 1 and rest of the people a feature value of 0. In the case of natural disaster in a city, every body was given a feature value of 0
- Level: The injury level represented the severity of injury for a given victim. These injury levels were inspired from TRIAGE injury levels.

TRIAGE

As per Wikipedia, Triage is the process of determining the priority of patients' treatments based on the severity of their condition. This rations patient treatment efficiently when resources are insufficient for all to be treated immediately. Triage may result in determining the order and priority of emergency treatment, the order and priority of emergency transport, or the transport destination for the patient.

Triage may also be used for patients arriving at the emergency department, or telephoning medical advice systems,[2] among others. This article deals with the concept of triage as it occurs in medical emergencies, including the prehospital setting, disasters, and emergency department treatment.

Table 2-2 Five-level Triage Systems

System	Countries	Levels	Patient should be seen by provider within
Australasian Triage Scale (ATS) (formerly National Triage Scale of Australia)	Australia New Zealand	1 - Resuscitation 2 - Emergency 3 - Urgent 4 - Semi-urgent 5 - Nonurgent	Level 1 - 0 minutes Level 2 - 10 minutes Level 3 - 30 minutes Level 4 - 60 minutes Level 5 - 120 minutes
Manchester	England Scotland	1 - Immediate (red) 2 - Very urgent (orange) 3 - Urgent (yellow) 4 - Standard (green) 5 - Nonurgent (blue)	Level 1 - 0 minutes Level 2 - 10 minutes Level 3 - 60 minutes Level 4 - 120 minutes Level 5 - 240 minutes
Canadian Triage and Acuity Scale (CTAS)	Canada	1 - Resuscitation 2 - Emergent 3 - Urgent 4 - Less urgent 5 - Nonurgent	Level 1 - 0 minutes Level 2 - 15 minute Level 3 - 30 minutes Level 4 - 60 minutes Level 5 - 120 minutes

(Australasian College for Emergency Medicine, 2002; Canadian Association of Emergency Physicians, 2002; Manchester Triage Group, 1997)

ASSUMPTIONS

In real life, a lot of other factors come into the picture including the demographics/ origin of the victims, hidden health conditions of the victim, obstructions on the path of victim retrieval, distance from each victim etc.

In this research, we have discarded the above assumptions to keep the focus of the research on the ethical dilemma of the situation at hand. Hence, the other factors which have not been extracted as features in the data have been discarded.

4. Experiment description

The experiment comprised of modeling ethical dilemma situations into questions and compiling them into a quiz. The questionnaire was answered by 328 people over a span of two months.

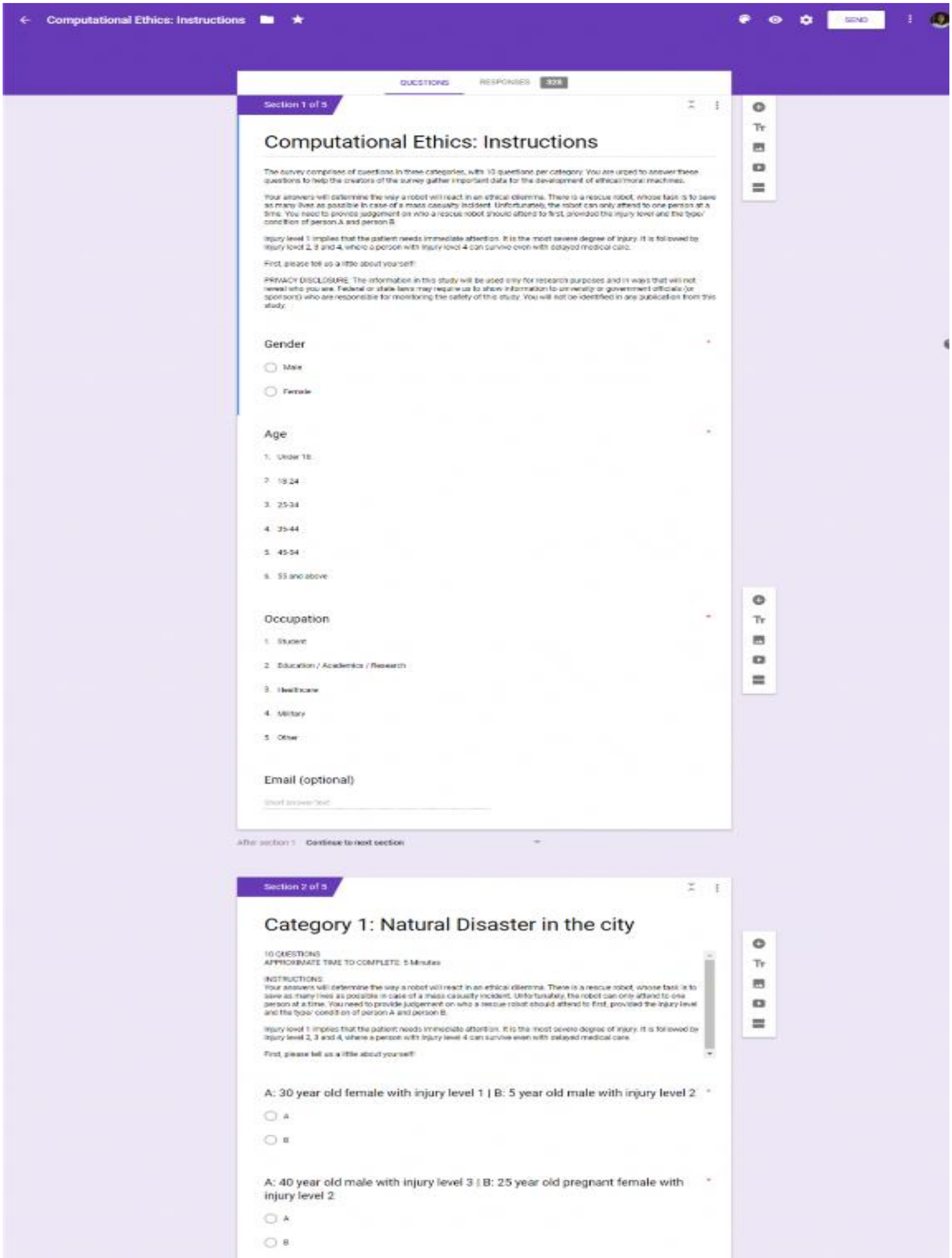
The number of questions were 30, which were equally distributed in three sections comprising of 10 questions each for each of the three situations considered.

No personal information was gathered in order to provide freedom of answering without being judged as well as to protect the sanctity of the data.

Following is the link to the survey which is powered by Google forms:

https://goo.gl/forms/ABukLIt0QaIujrIh2

Following is the screenshot of the survey:



Computational Ethics: Instructions

QUESTIONS RESPONSES

Section 1 of 5

Computational Ethics: Instructions

The survey comprises of questions in three categories, with 10 questions per category. You are urged to answer these questions to help the creators of the survey gather important data for the development of ethical/moral machines.

Your answers will determine the way a robot will react in an ethical dilemma. There is a rescue robot, whose task is to save as many lives as possible in case of a mass casualty incident. Unfortunately, the robot can only attend to one person at a time. You need to provide judgement on who a rescue robot should attend to first, provided the injury level and the type/condition of person A and person B.

Injury level 1 implies that the patient needs immediate attention. It is the most severe degree of injury. It is followed by injury level 2, 3 and 4, where a person with injury level 4 can survive even with delayed medical care.

First, please tell us a little about yourself!

PRIVACY DISCLOSURE: The information in this study will be used only for research purposes and in ways that will not reveal who you are. Federal or state laws may require us to show information to university or government officials (or sponsors) who are responsible for monitoring the safety of this study. You will not be identified in any publication from this study.

Gender *

- Male
- Female

Age *

1. Under 18
2. 18-24
3. 25-34
4. 35-44
5. 45-54
6. 55 and above

Occupation *

1. Student
2. Education / Academics / Research
3. Healthcare
4. Military
5. Other

Email (optional)

Short answer text

After section 1 Continue to next section

Section 2 of 5

Category 1: Natural Disaster in the city

10 QUESTIONS
APPROXIMATE TIME TO COMPLETE: 5 Minutes

INSTRUCTIONS:
Your answers will determine the way a robot will react in an ethical dilemma. There is a rescue robot, whose task is to save as many lives as possible in case of a mass casualty incident. Unfortunately, the robot can only attend to one person at a time. You need to provide judgement on who a rescue robot should attend to first, provided the injury level and the type/condition of person A and person B.

Injury level 1 implies that the patient needs immediate attention. It is the most severe degree of injury. It is followed by injury level 2, 3 and 4, where a person with injury level 4 can survive even with delayed medical care.

First, please tell us a little about yourself!

A: 30 year old female with injury level 1 | B: 5 year old male with injury level 2 *

- A
- B

A: 40 year old male with injury level 3 | B: 25 year old pregnant female with injury level 2 *

- A
- B

The responses were downloaded in excel format.

I wrote a python script to clean the downloaded data and convert it into features for victim A and for B, so that the data is prepared to be entered into the WEKA software for experimenting.

The following is a screenshot of how the data looks after cleaning:

A_AGE	A_GENDER	A_STATUS	A_LEVEL	B_AGE	B_GENDER	B_STATUS	B_LEVEL	OUTPUT
60	1	1	3	15	0	0	1	A
22	0	0	2	50	0	1	1	B
23	1	1	1	45	2	0	2	B

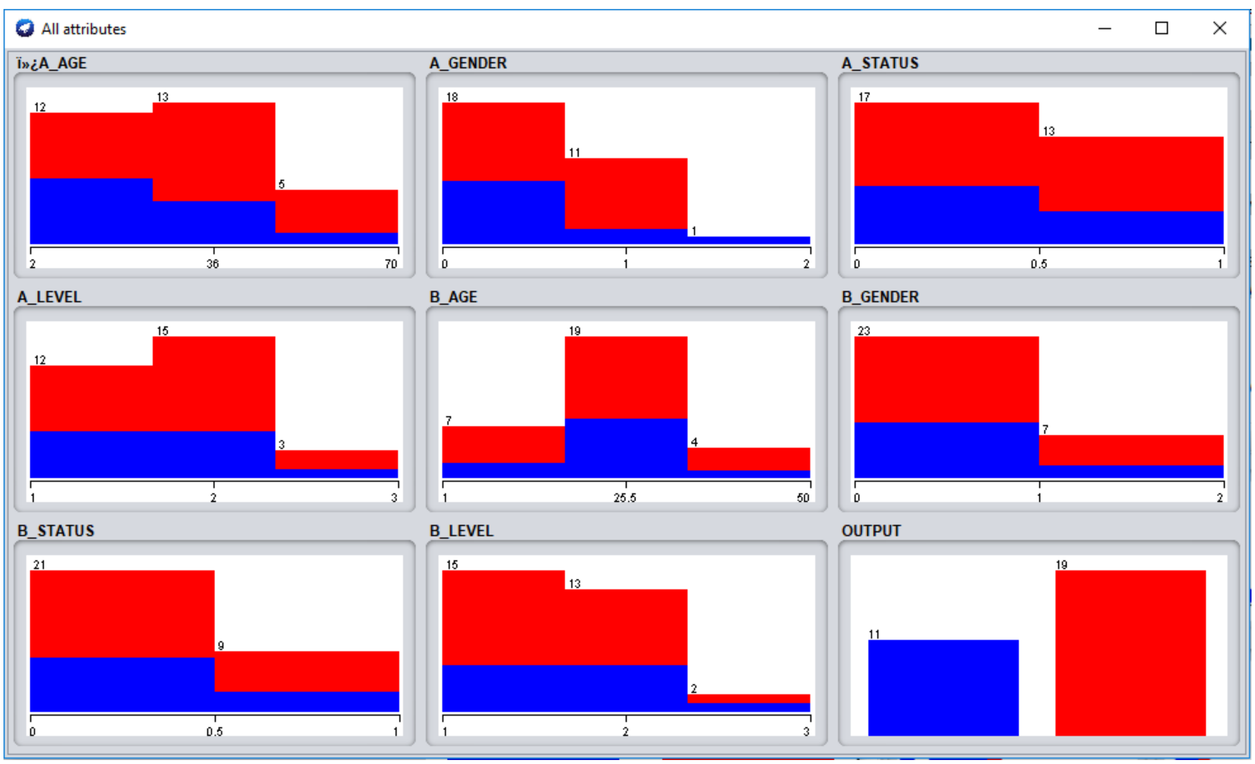
The data was fed into the WEKA software and multiple machine learning algorithms were run on it and their accuracy was noted.

5. Analysis

On first look, the following facts were noticed about the data:

- The input attributes are all numerical and have differing scales. We may see some benefit from either normalizing or standardizing the data.
- There are no missing values marked.
- The class attribute is nominal and has two output values meaning that this is a two-class or binary classification problem.
- The class attribute is unbalanced, 1 “A” outcome to 1.7 (=19/11) “B” outcomes, nearly double the number of negative cases. We may benefit from balancing the class values

<u>Attribute Distributions:</u>



We can notice a few things about the shape of the data:

- Some attributes have a Gaussian-like distribution such as A_Gender and B_Age, suggesting methods that make this assumption could achieve good results, like Logistic Regression and Naive Bayes.
- A lot of overlap can be seen between the classes across the attribute values. The classes do not seem easily separable.
- The class imbalance graphically depicted can be seen clearly.

Normalized View:

The first view created is of all the input attributes normalized to the range 0 to 1. This may benefit multiple algorithms that can be influenced by the scale of the attributes, like regression and instance based methods.

Standardized Views**:**

It was noted in the previous section that some of the attribute have a Gaussian-like distribution. We can rescale the data and take this distribution into account by using a standardizing filter.

This will create a copy of the dataset where each attribute has a mean value of 0 and a standard deviation (mean variance) of 1. This may benefit algorithms in the next section that assume a Gaussian distribution in the input attributes, like Logistic Regression and Naive Bayes

Baseline Algorithm:

You cannot know which algorithm will perform the best for your problem before hand so you must try a suite of algorithms and see what works best, then double down on it.

As such, it is critically important to develop a baseline of performance when working on a machine learning problem. A baseline provides a point of reference from which to compare other machine learning algorithms.

It is possible to get an idea of both the absolute performance increases you can achieve over the baseline as well as lift ratios that show you relatively how much better you are doing. Without a baseline you do not know how well you are doing on your problem. You have no point of reference to consider whether or not you have or are continuing to add value. The baseline defines the hurdle that all other machine learning algorithms must cross to demonstrate “skill” on the problem.

I chose the following as baseline algorithm:

ZeroR**:** ZeroR is the simplest classification method which relies on the target and ignores all predictors. ZeroR classifier simply predicts the majority category (class). Although there is no predictability power in ZeroR, it is useful for determining a baseline performance as a benchmark for other classification methods.

<u>Algorithm Evaluation (Unbalanced classes; A=11; B=19):</u>

I applied the following machine learning algorithms:

1) rules.ZeroR [simply predicts the majority category]
2) bayes.NaiveBayes [Naïve bayes algorithm]
3) functions.Logistic [logistic regression]
4) functions.SMO [Support vector machine] (SMO Implements John Platt's sequential minimal optimization algorithm for training a support vector classifier]
5) lazy.IBk [3-nearest neighbors]
6) rules.PART [Class for generating a PART decision list]
7) trees.REPTree [Fast decision tree learner]
8) trees.J48 [decision tree -> Class for generating a pruned or unpruned C4.5 decision tree]

Analysis:

Analysing: Percent_correct

Datasets: 3

Resultsets: 8

Confidence: 0.05 (two tailed)

Dataset	(1) rules.Ze |	(2) bayes	(3) funct	(4) funct	(5) lazy.	(6) rules	(7) trees	(8) trees
overall_1(100)	63.33 |	45.00 *	39.00 *	62.00	50.00	42.33	62.33	49.33 *
overall_1-weka.filters.un(100)	63.33 |	44.67 *	39.00 *	62.33	50.00	42.33	62.33	49.33 *
overall_1-weka.filters.un(100)	63.33 |	42.33 *	39.00 *	62.33	50.00	42.33	62.33	49.33 *

We can see that relatively, Support vector machines and Fast decision trees perform better than their counterparts.

Here the first column in the table represents the three views of the data as described above.

Rerunning analysis using Support vector machine (SMO) as test base:

Analysing: Percent_correct

Datasets: 3

Resultsets: 8

Confidence: 0.05 (two tailed)

Dataset	(4) function		(1) rules	(2) bayes	(3) funct	(5) lazy.	(6) rules	(7) trees	(8) trees
overall_1 (100)	62.00	|	63.33	45.00 *	39.00 *	50.00	42.33	62.33	49.33
overall_1-weka.filters.un(100)	62.33	|	63.33	44.67 *	39.00 *	50.00	42.33	62.33	49.33
overall_1-weka.filters.un(100)	62.33	|	63.33	42.33 *	39.00 *	50.00	42.33	62.33	49.33

Final model described in terms of Support vector machines with standard deviation in brackets:

Dataset(4)	functions.SMO '-C
overall_1 (100)	62.00(12.55) |
overall_1-weka.filters.un(100)	62.33(12.22) |
overall_1-weka.filters.un(100)	62.33(12.22) |

Final model described in terms of RepTrees with standard deviation in brackets:

Dataset	trees.REPTree '-M
overall_	62.33(12.22) |
overall_1-weka.filters.un(100)	62.33(12.22) |
overall_1-weka.filters.un(100)	62.33(12.22) |

Algorithm Evaluation(Balanced classes; A=15; B=15):

I balanced the classes by flipping the A and B attributes of 4 questions, so that the data now has balanced classes : A=15 and B=15

Dataset		(1) rules.Ze |	(2) bayes	(3) funct	(4) funct	(5) lazy.	(6) rules	(7) trees	(8) trees
overall_1	(100)	63.33 |	45.00 *	39.00 *	62.00	50.00	42.33	62.33	49.33 *
overall_1-weka.filters.un	(100)	63.33 |	44.67 *	39.00 *	62.33	50.00	42.33	62.33	49.33 *
overall_1-weka.filters.un	(100)	63.33 |	42.33 *	39.00 *	62.33	50.00	42.33	62.33	49.33 *

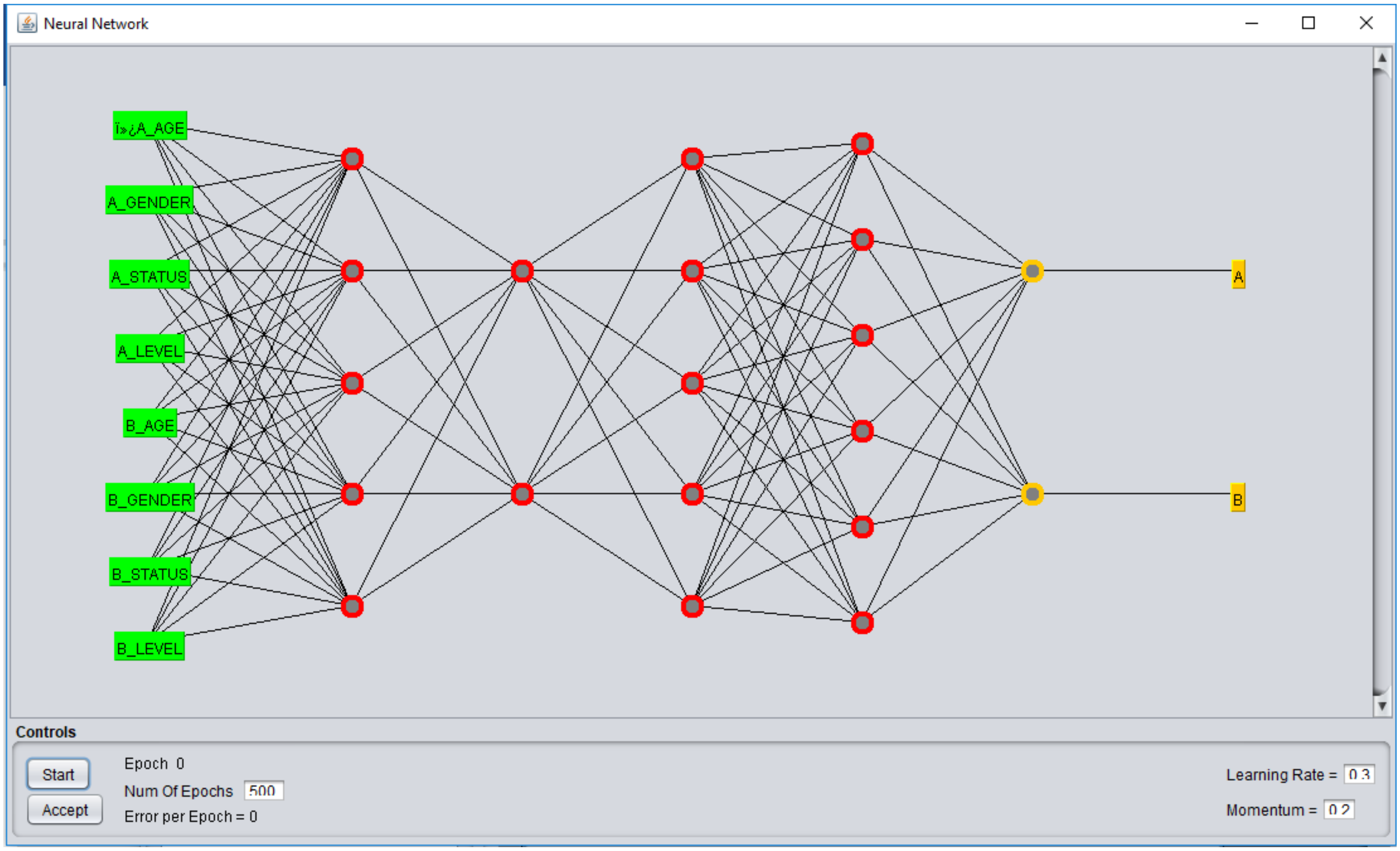
Surprisingly, there was not much improvement over balanced classes!

Modelling data using Neural Networks:

No. of hidden layers = 2 ; learning rate=0.3; momentum=0.2 ; cross validation=2; gives 50% accuracy

No. of hidden layers = 5 with 2,5,6 nodes in respective layers, learning rate=0.3; momentum=0.2, cross validation=2; gives

Correctly Classified Instances	19	63.3333 %
Incorrectly Classified Instances	11	36.6667 %



SCREENSHOT FROM WEKA SOFTWARE

6. Conclusion

Finalizing model and Presenting Results:

We can use the mean and standard deviation of the model accuracy collected in the last section to help quantify the expected variability in the estimated accuracy of the model on unseen data.

*We can generally expect that the performance of the model on unseen data will be 62.33% plus or minus (2 * 12.22)% or 24.44%. We can restate this as between 37.89% and 86.77% accurate.*

Cross validation results for same model (with c values<= 1; for c values > 1 the correctly classified percentage decreases)

Correctly Classified Instances	19	63.3333 %
Incorrectly Classified Instances	11	36.6667 %

7. Future Work

I plan to publish the study in the form of a research paper under the guidance of Professor Chris Davis. The following is the future work which can be done on the data and which can yield improved results:

- Separating the complete data into its specific 3 situations or domains and running the algorithms on domain specific/ only situational data.
- Tinkering with default neural network building scheme in WEKA and trying permutations of number of layers/nodes.
- Utilizing the percentage of preference for victim A and percentage of preference for victim B as features.
- Looking at the demographics of the respondents to get an insight into why some people answered the way they did
- Introduce a feature to represent the strength of classification based on data points.

8. References

I referred to the following articles and links to help me in my efforts for this course:

[1] https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-262

[2] http://www.firstaidforfree.com/a-guide-to-triage-for-first-aiders-and-first-responders/

[3] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839472/

[4]http://www.fireengineering.com/articles/print/volume-161/issue-3/features/ems-triage-sorting-through-the-maze.html

[5] http://www.jems.com/articles/2012/06/triage-distance.html?c=1

[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3148628/

[7] http://www.mariposacounty.org/DocumentCenter/Home/View/3336

[8] http://citmt.org/Start/answers.htm

[9] https://www.doomandbloom.net/the-mass-casualty-incident-triage-part-3/

[10] Emergency Severity Index, version 4 – Agency for healthcare research and quality

[11] https://machinelearningmastery.com/binary-classification-tutorial-weka/

[12]https://machinelearningmastery.com/estimate-baseline-performance-machine-learning-models-weka/