**Unit 4 Algorithmics**

**SAC 3 – Outcome 3**

Answer all questions on lined paper.

1. “The barber in Seville shaves all men who do not shave themselves, and only those men. Does the barber shave himself?” Explain the problem behind this question, and how it was part of the crisis facing mathematics in the early 20th century. (2 marks)

1m for relating it to (Russell’s) Paradox or the set of all sets that do not contain themselves.

1m challenged the idea that mathematics was a self-contained and logically consistent system

Most students got the first mark but not the second.

1. Describe key features of a Turing machine. (3 marks)

3 of:

Infinite tape

Moves forwards and backwards

State register

Finite symbol set

Max 2m if “can only write 0/1/blank” or similar – although this is commonly how they are used, it is not a requirement.

1. State what the Halting Problem sought to find. (1 mark)

An algorithm which can tell for any given program and input, whether or not the program halts.

Many students confused the Halting Problem with Turing’s conclusion that the Halting Problem was undecidable.

1. Explain Turing’s argument which shows that the Halting Problem is undecidable. (3 marks)

For example:

Imagine a program (P) which can determine, for a given algorithm and input, if it will halt or run forever. If the input runs forever then P will halt. If the input halts then P will run forever, i.e. not halt. Now feed that program into itself, P+. Note that P and P+ are the same program. If P halts, then P+ doesn’t halt. If P doesn’t halt then P+ does. Because P and P+ are the same program, we have a contradiction. Hence it is impossible to design a machine which will tell us, for any machine, whether or not it will halt on a given input.

Some students lost a mark for saying that the output was inconsistent/changed each time.

1. What is the significance of Turing’s conclusion for the entscheidungsproblem? (2 marks)

The ESP sought an algorithmic method to determine the truth or falsity of mathematical/logical statements. Turing shows that no general algorithm can do this for all statements.

1. Rajesh argues that we could change the design of our hypothetical Halting machine so that the paradox does not arise. Explain why this does not affect the conclusion of the Halting Problem. (1 mark)

The Halting Problem requires a machine which can tell, for any program, whether or not it will halt. We only have to design one program for which this is impossible in order to demonstrate that no such machine can exist.

Had to get the idea that we only need one counterexample to make the point. Many students talked generally about the HP without addressing Rajesh’s claim.

1. With reference to our knowledge of other minds, discuss whether or not it is reasonable to believe that a computer is sentient (conscious) in light of the Chinese Room Argument. Make sure that you present a balanced discussion. (3 marks)

The Chinese Room Argument states that a man in a room manipulating Chinese characters according to rules can present the appearance of understanding Chinese without actually understanding, and that this is what is happening with a computer.

The other minds response says that we only know that other people exist based on their behaviour, and so we should apply the same burden of proof to computers.

However, we also assume other minds are conscious because of their biological origins, not just their behaviour.

Still, what if we created a computer using similar biological processes? Would this then be judged as conscious on the same basis as people?

Many students scored 2m for the middle two points above. A third mark could be earned for either a brief summary of CRA or further depth of discussion such as the final point.

Question 8 to 11 relate to the following case study.

The government has decided to use artificial intelligence in order to predict the probability of criminals reoffending, and hence whether or not to release them from prison early. They have gathered data on 50,000 offenders, including: age, sex, ethnicity, nature of offence committed, history of offending and religious belief.

The outcomes we aim to predict (and are known for our training data) are:

1. whether or not the person reoffended during the two years following their release
2. if so, what was the seriousness of that offence, on a binary scale – serious or minor.
3. Explain how a binary classification SVM could be best constructed to solve this problem. (2 marks)

The data would be fed into a SVM and separated using a multidimensional hyperplane into “offends” or “doesn’t re-offend”.

One approach would then take the “offends” data and put it into another SVM to separate into “serious” or “minor” reoffence. This was quite common.

Alternatively one could ignore the second outcome and just focus on whether re-offence occurred.

Some students did not show an understanding that the output must be binary, in which case only scored 1m.

1. We find that the margin generated is very small. What problem does this indicate? Explain how this problem could cause our SVM to make an incorrect prediction. (2 marks)

Multiple options here. Both overfitting or underfitting were possible. Also it could be that the data was simply hard to separate. (We needed to know how accurate the model was against the training data to be more sure). Depending on the problem chosen, the incorrect prediction could be caused by the model failing to accurately separate the data in some way.

1. How would a neural network allow us to generate a more nuanced model? Answer specifically to the scenario as stated. (1 mark)

A NN allows non-binary outputs, e.g. the probability of reoffending and the seriousness of the offence.

Some students said that a NN was needed for non-linear relationships, however the SVM can handle these too via transformations. Also many talked about hidden layers and nuance, but this was not convincing and generally not related to the context.

1. Discuss the ethical issues that are raised by using artificial intelligence in this way when compared to using human judgement (i.e. a parole board, usually comprised of three people). Explain whether the issues you discuss are more likely to arise for a neural network or a SVM. You should consider transparency, bias, fairness and human dignity. (5 marks)

Points include:

**Transparency**

More likely a problem with a NN

If we are denying parole, we need to be able to explain why, even if we can predict reoffending with 100% accuracy.

On the other hand, humans are no better at explaining such things – we often make a decision and then look for reasons to justify it.

Also this might make the system gameable – if e.g. Pastafarians are less likely to reoffend, then people could change religion.

**Bias**

Could arise with either SVM/NN

If historically there are certain ethnic groups which have been unfairly targeted by the justice system,i.e. more likely to be rearrested than another person for the same offence, then they will be given a higher probability of reoffending.

Human judges would still have this bias. In either case it could be manually corrected – both judges and our AI might be told to treat some groups more leniently.

Also there may be good reasons for bias – violent men may be more likely to reoffend than violent women.

**Fairness**

Both AIs would account for all information in a systematic way, so that it should be fairer than humans who are subject to moods, pre-conceptions etc.

However we have to make sure we choose an adequate number of attributes to measure to make sure we are including all relevant information in our decisions.

**Human Dignity**

Humans should not have life-changing decisions made about them by a machine. They should be able to present their case to a panel.

However plenty of other such decisions are made by machine, e.g. credit decisions. If it’s fairer and safer for society then this overrides other concerns.

A panel favours people who are articulate and convincing, but this does not predict criminality. (This point could be made under fairness or bias too.)

For 5 marks students needed some depth rather than lists of points. Many students scored 3 marks for presenting one side of an argument, or for omitting any distinction between NN/SVM.