COMP3308/3608, Lecture 13b

ARTIFICIAL INTELLIGENCE

Preparation for the Exam

Sample exam/revision questions are available on Canvas

3 documents:

1) Sample exam questions

2) Search: Quiz practice with quokkas

3) Bayesian networks: BN practice questions

Marking - Reminder

- **No more homeworks and assignments - well done for completing all assessments!**
- **Homeworks – the 3 selected homeworks were announced on Piazza, marking is completed and the marks are in Canvas**
- **Assignment 2**
  - **Code (/12) - marked by PASTA - your mark is already available in PASTA; it is determined by the number of tests you have passed**
  - **Report (/12) – we aim to complete the marking by Thursday week 13**
  - **Check PASTA – your mark and feedback may be available already**
  - **We mark progressively and as soon as your report is marked, you can see your mark in PASTA, together with the marking criteria, the breakup of your marks and additional feedback on your assignment**
  - **To see the marking sheet and feedback you may need to reload the page in PASTA**
  - **Marks will be uploaded in Canvas after all marking is completed and the late penalties will be applied in Canvas**
  - **You will be asked to check your mark in Canvas**

Last week of the semester - we are approaching the finish line!

What a semester!
So different and difficult.

Eliud Kipchoge from Kenya wining the men's marathon at the Rio Olympic Games in 2016

- **How do you feel?**
- **happy, excited, tired, exhausted, relieved, scared ...**

- **Do you think you have learned useful things in this course?**

Images from: https://www.iaaf.org/news/report/rio-2016-men-marathon, Getty images

Outline

- **Complete the Unit of Study Survey**
  https://student-surveys.sydney.edu.au/students/

- **Talk about the exam**

Unit of study surveys – why you should do them?

- **We really need you do complete the survey. Please do it now!**
  https://student-surveys.sydney.edu.au/students/
- **These surveys are required for all courses, not only for AI**

- **Please take them seriously and complete them properly – so many important decisions are based on them, e.g.**
  - **Courses with low student satisfaction are investigated by the Head of School and Associate Dean Education**
  - **Courses with high student satisfaction are commended**
- **Courses with low response rate are also investigated – the required minimum response rate is 50%**

Why you should do them? (2)

- **We need to know what went well and what didn't**
- **We put a lot of effort in this course – we need your feedback**
- **We make changes every year based on the student surveys**
- **If you had any difficulties or concerns, please be specific and explain them**
  - **You don't want your comments to be dismissed, e.g. "It is natural students to complain when the material is difficult/not interesting/etc."**
  - **You don't want someone who doesn't teach well to say: "I don't know what the student meant; I did a good job..."**
- **I can assure you that I read every single comment and make changes based on the surveys**

Why you should do them? (3)

- **There are prizes! ☺**
- **Each semester there is a prize draw for a:**
  - **1st prize - Apple iPad Air 64GB**
  - **2nd prize – Apple Watch**
  - **3rd prize - 2 x $200 JB HiFi Giftcards**
  - **4th prize - 6 x $100 JB HiFi Giftcards**
- **The draw will be on 30 June 2020**

- **Last semester's winners (semester 2, 2019):**
- **https://student-surveys.sydney.edu.au/students/complete/prizes.cfm**

Exam question: What is the probability to win these prizes? ☺

Unit of study surveys (4)

- **The surveys are *anonymous***
  - **We receive the results after all marks are finalised and don't know which student gave which rating or comment**
- **After the marks are finalised, you will receive a summary of the numeric ratings provided by all students in the course, together with my reply to your ratings and comments**
- **Q6 is about *feedback on your learning***
  - **Was your learning supported well, were your questions answered well on Piazza and during the tutorials; feedback on the assessments**
- **Q10 is about the *tutorials*. Write the name of your tutor – you know them from the Zoom sessions:**
  - **COMP3308:**
    - **Tue 4pm and 5pm: David**
    - **Wed 10am and 11am: Tristan**
    - **Wed 1pm: Nick and David**
    - **Wed 2pm: Nick and Stephen**
    - **Wed 3pm: James**
    - **Wed 4pm: Stephen and James**
    - **Wed 5pm: Stephen**
  - **COMP3608: Jessica (all tutorials)**
- **Complete the survey now: https://student-surveys.sydney.edu.au/students/**

I will pass all your comments to your tutors!

The Exam

- **Online, on Canvas, set as a Quiz**
- **Unproctored**
- **Duration: 2 hours + 10 minutes reading time = 130 minutes**
  - **The Canvas Quiz will open at the scheduled time (as per your exam timetable) and close after 130 minutes**
- **Marks total: 60 (=60% of your final mark)**
- **A minimum of 24 marks (40%) on the exam is required to pass the course**

- **Have a calculator ready as there are calculation questions**
- **The recommended browsers for Canvas are Chrome and Firefox, not Internet Explorer or others**

Canvas Exam Sites

- **The Canvas site for the exam is different that the Canvas site we use during the semester for teaching – 2 exam sites:**
  - **Final exam for COMP3308**
  - **Final exam for COMP3608**
- **The Exams Office will give you access to your exam site**

Students in Different Time Zones

- **Important for international students who could not come to Australia and exchange students who had to return to their country (we have >30 exchange students from all continents):**
- **If you are in a different time zone and the exam is too early/too late for you, you can apply for Special Consideration to have special arrangements for the exam:**

https://www.sydney.edu.au/students/special-consideration.html#time-zone

Examinable Content

- **All material is examinable <u>except</u>:**
  - **week 1 - introduction**
  - **week 13a (applications of AI – recommender systems)**
  - **historical context, Matlab and Weka**

Other Advice

- **Check and double-check the day and time of your exam**
- **Be calm and well organized**
- **Allocate time according to marks for each question**

- **You should find the exam easy if you:**
  - **Have attended lectures and followed up**
  - **Have attended tutorials, have done all tutorial exercises and followed up**
  - **Have read the recommended reading from the text books**
  - **Have done well in the assignments**

Types of Questions

- **The exam contains 3 types of questions:**
  - **problem solving questions**
  - **questions requiring short answers - test your understanding and ability to relate concepts; be clear and concise**
  - **multiple-choice questions (small number)**

- **The exam questions are similar to the tutorial questions**

Don’t Waste Your Time!

- **During the exam:**
  - **You need to be quick!**
  - **The questions are not easily googalable**
  - **There are different versions of questions to prevent cheating**

- **Do what the question asks for:**
  - **Some questions ask you to show your calculations/working, some others - just the final answer**

Short-Answer Questions

a) *Gain ratio* is a modification of *Gain* used in decision trees. What is its advantage?

It penalizes highly-branching attributes by taking into account the number and the size of branches.

b) Why do we need to normalize the attribute values in the k-nearest-neighbor algorithm?

As different attributes are measured on different scales, without normalization the effect of the attributes with smaller scale of measurement will be less significant than those with larger.

c) What is the main limitation of the perceptrons?

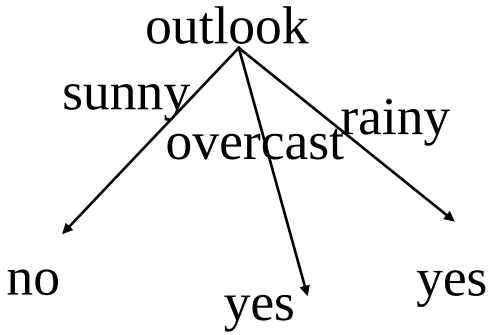
Can separate only linearly separable data.

Short-Answer Questions (2)

d) The 1R algorithm generates a set of rules. What do they test?

The values of a single attribute.

if outlook=sunny then play=no
elseif outlook=overcast then play=yes
elseif outlook=rainy then play=yes



e) The problem of finding a decision boundary in SVM can be formulated as an optimisation problem using Lagrange multipliers. What are we maximizing? The margin of the hyperplane.

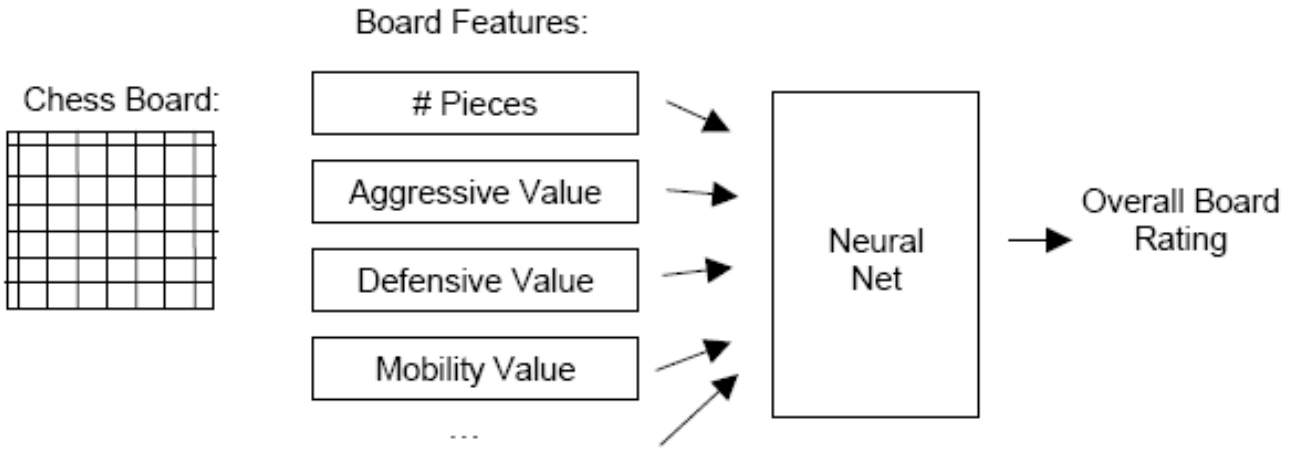
f) In linear SVM, we use dot products both during training and during classification of a new example. What vectors are these products of?

During training: Pairs of training examples.

During classification of a new example: The new example and the support vectors.

Multiple Choice Questions - Backpropagation

- **Your task is to develop a computer program to rate chess board positions**
- **You got an expert chess player to rate 100 different chessboard positions and then use this data to train a backpropagation neural network, using board features as the ones shown in the figure below:**



Multiple Choice Questions – Backpropagation (2)

Answer with yes or no:

1. The backpropagation network may be susceptible to overfitting of the training data, since you tested its performance on the training data instead of using cross validation.

Yes

2. The backpropagation neural network can only distinguish between board positions that are completely good or completely bad.

No

The backpropagation neural network can be used for both classification and regression tasks. In regression tasks the output is numeric and can represents different levels of good and bad.

Multiple Choice Questions - Backpropagation

3. The backpropagation network implements gradient descent, so your network may converge to a set of weights that is only a local minimum rather than the global minimum.

Yes

4. You should have used higher learning rate and momentum to guarantee convergence to the global minimum.

No

The use of momentum reduces the oscillations when using a higher learning rate but it doesn't guarantee convergence to the global minimum.

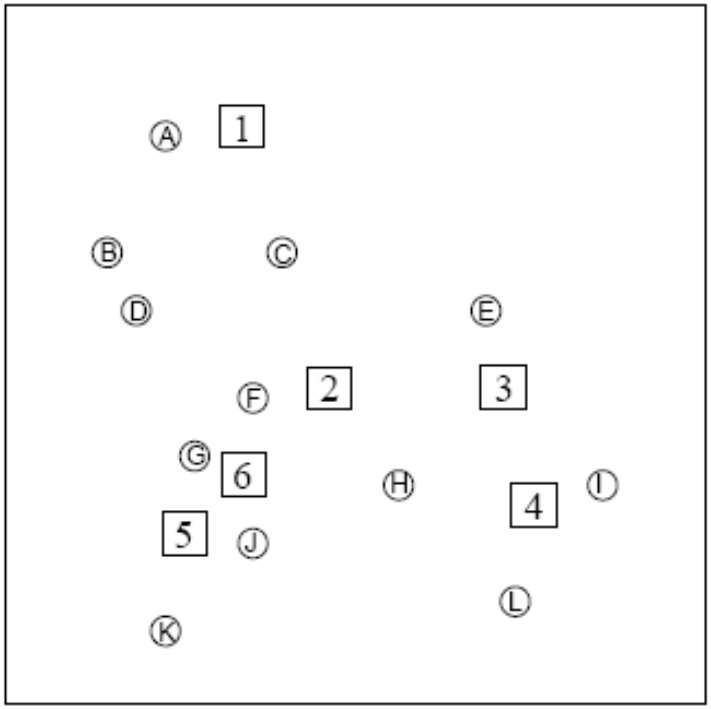
5. The topology of your neural net might not be adequate to capture the expertise of the human expert.

Yes

Too few neurons – underfitting; too many – overfitting.

Problem Solving Question - Nearest Neighbor

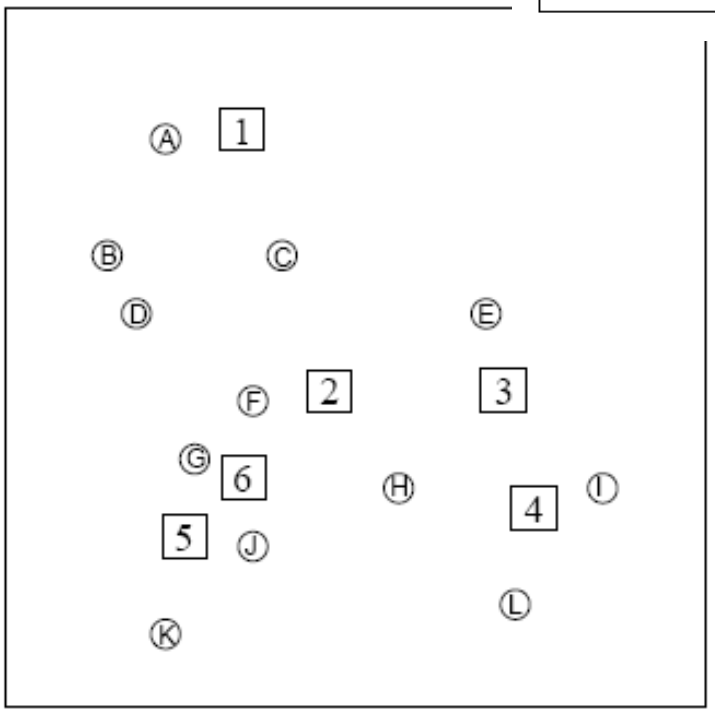
- **In the figure below, the circles are training examples and the squares are test examples, i.e. we are using the circles to predict the squares**
- **Two algorithms are used: 1-Nearest Neighbor and 3-Nearest Neighbor**



Problem Solving Question - Nearest Neighbor

We know that:

Square Point	Using 1 nearest neighbors	Using 3 nearest neighbors
1	-	+
2	-	
3		+
4	+	-
5		-



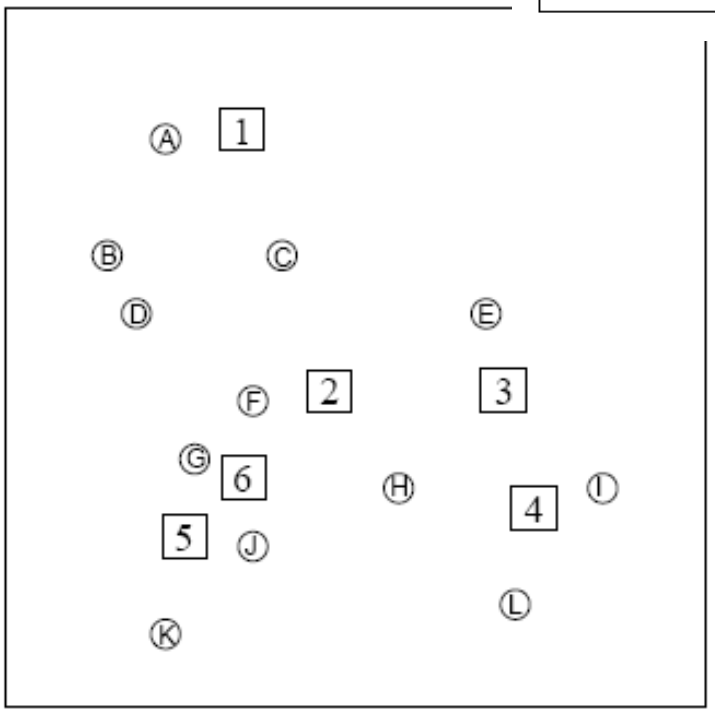
What will be the class of the following examples?
Write +, - or U for cannot be determined.

Circle L:
Circle I:
Circle H:
Circle E:
Circle K:
Circle C:
Square 6 using 1-Nearest Neighbour:
Square 6 using 3-Nearest Neighbour:
Square 3 using 1-Nearest Neighbour:
Square 5 using 1-Nearest Neighbour:

Answer

We know that:

Square Point	Using 1 nearest neighbors	Using 3 nearest neighbors
1	-	+
2	-	
3		+
4	+	-
5		-

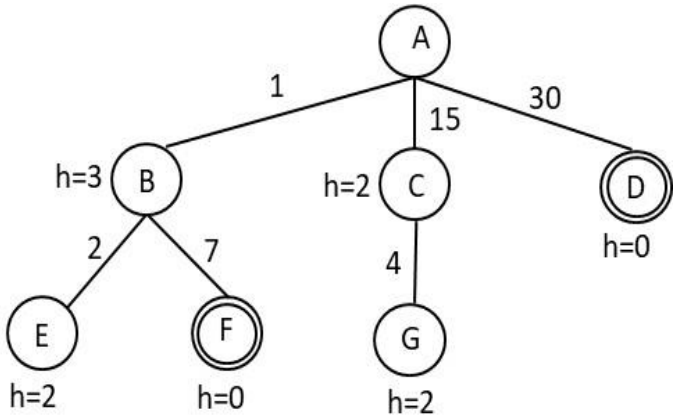


What will be the class of the following examples? Write +, - or U for cannot be determined.

Circle L: -
Circle I: +
Circle H: -
Circle E: +
Circle K: U
Circle C: +
Square 6 using 1-Nearest Neighbour: U
Square 6 using 3-Nearest Neighbour: -
Square 3 using 1-Nearest Neighbour? +
Square 5 using 1-Nearest Neighbour? U

Problem Solving Question - Search

- **In the tree below the step costs are shown along the edges and the h values are shown next to each node. The goal nodes are double-circled: F and D.**

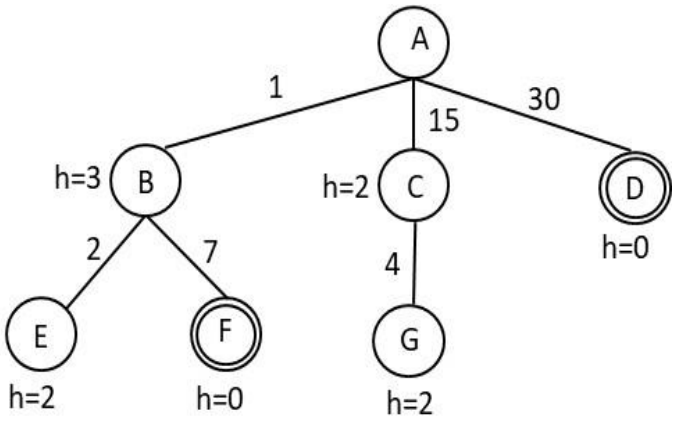


- **Write down the order in which nodes are expanded using:**
  - **a) BFS**
  - **b) DFS**
  - **c) UCS**
  - **d) IDS**
  - **e) Greedy search**
  - **f) A***
- **In case of ties, expand the nodes in alphabetical order.**

Search - Review

- **BFS: Expands the shallowest unexpanded node**
- **DFS: Expands the deepest unexpanded node**
- **UCS: Expands the node with the smallest path cost g(n) (from the root!)**
- **IDS: DFS at levels l = 0, 1, 2, etc.; expands the deepest unexpanded node within level l**
- **Greedy: Expands the node with the smallest heuristic value h(n)**
- **A*: Expands the node with the smallest f(n)=g(n)+h(n)**

Solution



- **Write down the order in which nodes are expanded using:**
  - **a) BFS: ABCD**
  - **b) DFS: ABEF**
  - **c) UCS: ABEF**
  - **d) IDS : AABCD**
  - **e) Greedy search: AD**
  - **f) A*: ABEF**

Short Answers - Search

- **A* uses admissible heuristics. What happens if we use a non-admissible one? Is it still useful to use A* with a non-admissible heuristic?**

Not optimal anymore. But it could still find a reasonably good solution in acceptable time, depending on how good the heuristic is.

- **What is the advantage of choosing a dominant heuristic in A* search?**

Expands a fewer number of nodes. As a result, the optimal solution will be found quicker.

- **What is the main advantage of hill climbing search over A* search?**

Space complexity – keeps only the current node in memory

Some ML Competitions

Competition sites:

- **Kaggle:** http://www.kaggle.com/
- **KD nuggets:** http://www.kdnuggets.com/datasets/competitions.html

Where to Apply AI and ML?

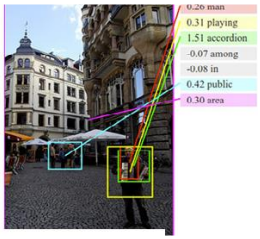
- **IBM's big 5 ML application areas:**

http://research.ibm.com/cognitive-computing/machine-learning-applications

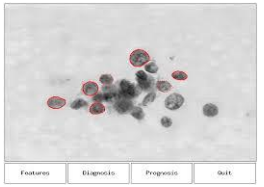
1. ***Education*** **- the classroom of the future will "learn" the student**
2. ***Shopping*** **- personal shopping, predicting what customers will like**
3. ***Medicine*** **– personalised disease treatment based on your DNA**
4. ***Security*** **– fraud detection and protection (bank accounts, passwords on social media, etc.)**
5. ***Smart cities*** **– understanding and responding to citizen's needs (transport, emergencies, housing, electricity consumption; data from crowdsourcing, mobile applications and sensors)**

- **Application areas aligned with UN development goals - e.g. quality education, good health, 0 hunger and homelessness, clean water, etc. http://www.un.org/sustainabledevelopment/sustainable-development-goals/**
- **Workshop on Data Science for Social Good at the European Conference on ML & DM: for https://sites.google.com/view/ecmlpkddsogood2020/**

The Last Slide



- **We hope that you found this course useful**
- **Use what you have learned in this course!**
- **Apply AI and ML – there are so many opportunities!**
- **There are already some great AI and ML success stories, but there will be even greater ones in the future and you can be part of them! You can contribute to them or lead them!**



All the best!

Jessica, Givanna, Stephen, David, Tristan, Nicholas, James and Irena