COMP5318 Machine Learning and Data Mining

Week 13 Revision and Preparation for the Exam

I. Unit of study survey

Please complete the unit of study evaluation survey.
We **really need you to do this:**

- We would like to hear your feedback – what worked well and what didn't. We make changes every year based on the student surveys to improve the course.
- The survey is part of the faculty teaching quality process and courses with low scores are investigated, as well as courses with a low response rate. **Important decisions are made based on the surveys, so please take the survey seriously.**

The survey is anonymous. To complete it:
1) click on: https://student-surveys.sydney.edu.au/students/
2) fill in your Unikey details
3) complete the survey and remember to press "Submit"

You can also use this QR code:

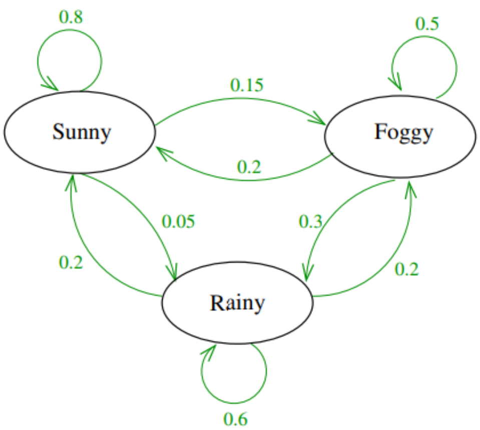
Thank you for completing the survey!

II. Tutorial questions

See the document "Sample exam questions" on Canvas. We will do some of the questions during the tutorial today.

<u>Exercise 1.</u> Markov models

Given is the following Markov model for the weather in Sydney:



If the weather yesterday was *Rainy*, and today is *Foggy*, what is the probability that tomorrow it will be *Sunny*?

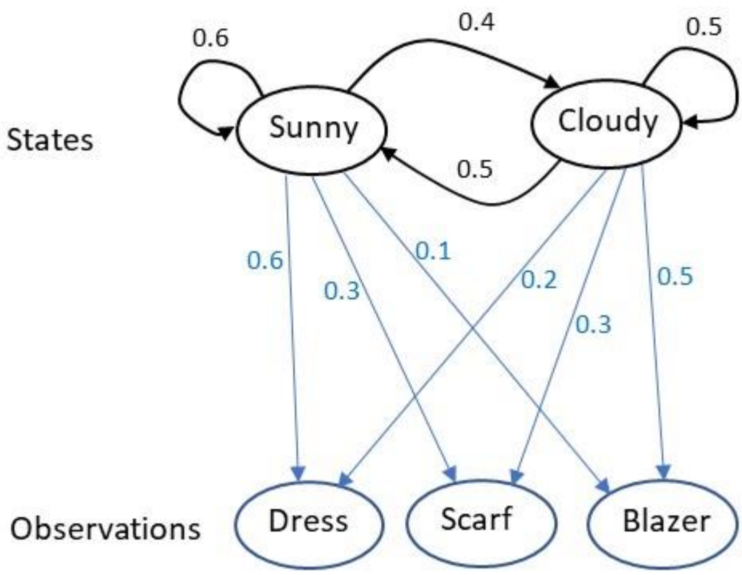
Solution:

$P(q_3 = Sunny \mid q_2 = Foggy, q_1 = Rainy) =$
$= P(q_3 = Sunny \mid q_2 = Foggy) =$ (Markov assumption) $= 0.2$

<u>Exercise 2.</u> Hidden Markov models

Julia tested positive to COVID and had to quarantine at home for several days. Her friend Nicole came to bring her food every day. We don't know what the weather was on the quarantine days but we know the type of clothing Nicole wore and it provides evidence about the weather.

The following Hidden Markov Model models the situation. The initial state probabilities are: A_0(Sunny)=0.5 and A_0(Cloudy)=0.5.



Suppose that on the first quarantine day Nicole wore a dress and on the second she wore a blazer.

a) What is the probability of the observation sequence?
b) What is the most likely sequence of hidden states?

Briefly show your calculations.

Solution:

States: Sunny (S) and Cloudy (C)
Observations: Dress (D), Scarf (Sc) and Blazer (B)
Observation sequence: X=Dress (D), Blazer (B)

a) This is HMM problem 1. We can use the Forward algorithm to solve it.

<u>Step 1: Initialization</u>

Day 1: D observed

$V_S(1) = A_0(S)e_S(D)$= *0.5*0.6=0.3*
$V_C(1) = A_0(C)e_C(D)$= *0.5*0.2=0.1*

Step 2: Iteration

Day 2: B observed

$V_S(2) = e_S(B)*(V_S(1)a_{SS} + V_C(1)a_{CS})=$
$= 0.1*(0.3*0.6+ 0.1*0.5)=0.1*(0.18+0.05)=0.023$

$V_C(2) = e_C(B)*(V_S(1)a_{SC} + V_C(1)a_{CC})=$
$= 0.5*(0.3*0.4 + 0.1*0.5)=0.5*(0.12+0.05)=0.085$

Step 3: Termination

P(X)=P(D,B)=0.023+0.085=0.108

The probability of the observation sequence Dress, Blazer is 0.108.

b) This is HMM problem 2. We can use the Viterbi algorithm to solve it.

Step 1: Initialization

Day 1: D observed

$V_S(1) = A_0(S)e_S(D)= 0.5*0.6=0.3$
$V_C(1) = A_0(C)e_C(D)= 0.5*0.2=0.1$

Step 2: Iteration

Day 2: B observed

$V_S(2) = e_S(B)*\max(V_S(1)a_{SS},\ V_C(1)a_{CS})=$
$= 0.1*\max(\mathbf{0.3*0.6},\ 0.1*0.5)=0.1*0.3*0.6=0.018$
$Ptr_S(2) = \mathrm{argmax}(\mathbf{0.3*0.6}, 0.1*0.5) = 1$, i.e. S

$V_C(2) = e_C(B)*\max(V_S(1)a_{SC},\ V_C(1)a_{CC})=$
$= 0.5*\max(\mathbf{0.3*0.4},\ 0.1*0.5)=0.5*0.3*0.4=0.06$
$Ptr_C(2) = \mathrm{argmax}(\mathbf{0.3*0.4}, 0.1*0.5) = 1$, i.e. S

Step 3: Termination and trace-back

Final state = argmax($V_S(2), V_C(2)$)=argmax(0.018, **0.06**)=2, i.e. C

Trace-back through pointers:
$Ptr_C(2)$ =S

Hence, the most likely sequence of hidden states is S, C (Sunny, Cloudy)

<u>Exercise 3.</u> Perceptron

Given is the following training set:

	input	output
ex. 1:	1 0 0	1
ex. 2:	0 1 1	0
ex. 3:	1 1 0	1
ex. 4:	1 1 1	0
ex. 5:	0 0 1	0

a) Train a perceptron **with a bias** on this training set. Assume that all initial weights (including the bias of the neuron) are 0. Show the set of weights (including the bias) at the end of the first epoch. Apply the examples in the given order.

Recall that the perceptron uses a step function defined as:

step(n) = 1, if n >= 0
= 0, otherwise.

Solution:

a) starting point: w=[0 0 0], b=0

applying ex.1: p_1= [1 0 0], t_1=1
a=step([0 0 0][1 0 0]+0)=step(0)=1
e=t_1-a=1-1=0
w^{new}=[0 0 0]+0[1 0 0] =[0 0 0]
b^{new}= 0+0=0
i.e. no change in w and b as e=0

applying ex.2: p_2= [0 1 1], t_2=0
a=step([0 0 0][0 1 1]+0)=step(0)=1
e=t_2-a=0-1=-1
w^{new}=[0 0 0]+(-1)[0 1 1] =[0 -1 -1]
b^{new}= 0+(-1)=-1
i.e. w an b have been updated

Similarly:
After applying ex. 3: w^{new}=[1 0 –1], b^{new}=0
After applying ex. 4: w^{new} =[0 -1 -2], b^{new} =-1
After applying ex. 5: no change
end of epoch 1

Weight vector and bias at the end of epoch 1: w =[0 -1 –2], b =-1

Thank you for doing this course, we hope you found it useful. Good luck with the exam!