1. *Warm Up:*

# How many distinct weights must be learned for the connections to the Convolution layer?

A. 16 B. 169 C. 2,704 D. 16,900 E. None of the above

Answer: A because 4 x 4 = 16 weights because weights are shared.

# How many units are in the Pooling layer?

A. 4 B. 25 C. 81 D. 100 E. 169

Answer: B because 5 x 5 = 25 units

# How many distinct weights must be learned for the connections to the Output layer?

A. 4 B. 5 C. 16 D. 100 E. 104

Answer: E because (25 x 4) + 4 = 104 (including the bias weights)

# Which one of the following is equal to P(A, B, C) given Boolean random variables A, B and C, and no independence or conditional independence assumptionsbetween any of them?

* + 1. *P(A | B) P(B | C) P(C | A) B. P(C | A, B) P(A) P(B) C. P(A, B | C) P(C)*

D. P(A | B, C) P(B | A, C) P(C | A, B) E. P(A | B) P(B | C) P(C)

Answer: C

# A 6-sided die is rolled 15 times and the results are: side 1 comes up 0 times; side 2: 1 time; side 3: 2 times; side 4: 3 times; side 5: 4 times; side 6: 5 times. Based on these results, what is the probability of side 3 coming up when using Add-1 Smoothing?

A. 2/15 B. 1/7 C. 3/16 D. 1/5 E. None of the above Answer: B

Because (2+1)/(1+2+3+4+5+6) = 1/7

# Compute P (D | TP), the posterior probability that you have disease D when the test is positive.

A. 0.0495 B. 0.078 C. 0.635 D. 0.97 E. None of the above

Answer: C

Because P(D | TP) = (P(TP | D) P(D)) / P(TP) = ((.99)(.05)) / .078 = 0.635

# The AdaBoost algorithm creates an ensemble of weak classifiers by doing which one of the following before determining the next weak classifier:

* + 1. *Chooses a new random subset of the training examples to use*
    2. *Decreases the weights of the training examples that were misclassified by the previous weak classifier*
    3. *Increases the weights of the training examples that were misclassified by the previous weak classifier*
    4. *Removes the training examples that were classified correctly by the previous weak classifier*
    5. *None of the above Answer: C*

# True or False:

* 1. The Perceptron Learning Rule is a sound and complete method for a Perceptron to learn to correctly classify any 2-class classification problem.
     1. *True B. False*

# The back-propagation algorithm, when run until a minimum is achieved, always finds the same solution (i.e., weights) no matter what the initial set

of weights are.

* + 1. *True B. False*

# Convolutional Neural Networks (CNNs) can learn to recognize an object in an image

no matter how the object is translated (i.e., shifted horizontally and/or vertically) even if the training set only includes that object in one position.

* + 1. *True B. False*

# True or False: Given a linearly-separable dataset for a 2-class classification problem,

a Linear SVM is better to use than a Perceptron because the SVM will often be able to achieve better classification accuracy on the testing set.

* + 1. *True B. False*

# True or False: For a 2-class classification problem where AdaBoost has selected five

weak classifiers, the classification of a test example is determined by the majority class of the classes predicted by the five weak classifiers.

* + 1. *True B. False*

# Technical Questions:

* 1. If f (s), g(s) and h(s) are all admissible heuristics then which of the following are also guaranteed to be admissible heuristics:

f (s) + g(s) + h(s)  f (s)/3 + g(s)/3 + h(s)/3



f (s)/6 + g(s)/3 + h(s)/2  f (s) ∗ g(s) ∗ h(s)

min(f (s), g(s), h(s))  min(f (s), g(s) + h(s))

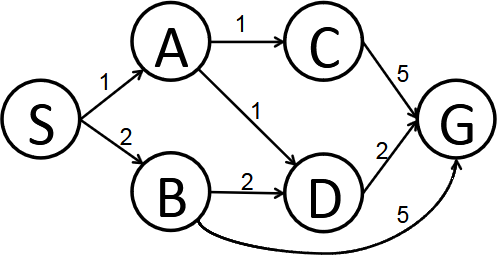
max(f (s), g(s), h(s))  max(f (s), g(s) + h(s))

# Answer the following questions about the search problem shown above. S is the start- state, G is the (only) goal-state. Break any ties alphabetically. For the questions that ask for a path, please give your answers in the form ‘S − A − D − G.’

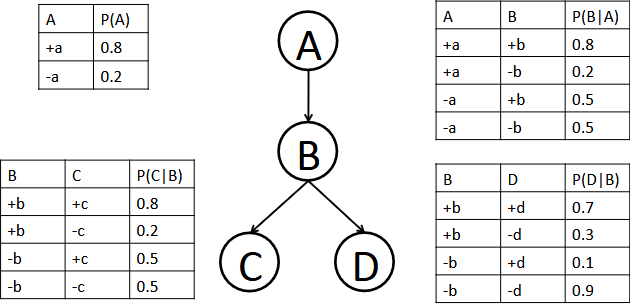
* + *What path would breadth-first graph search return for this search problem?*

Answer: S − B − G

* + *What path would uniform cost graph search return for this search problem? Answer: S − A − D − G*
  + *What path would depth-first graph search return for this search problem? Answer: S − A − C − G*
  + *What path would A\* graph search, using a consistent heuristic, return for this search problem? Answer: S − A − D – G*



# Representation Consider the joint distribution P (A, B, C, D) defined by the Bayes’ net below.



Compute the following quantities:

P (A = +a) =

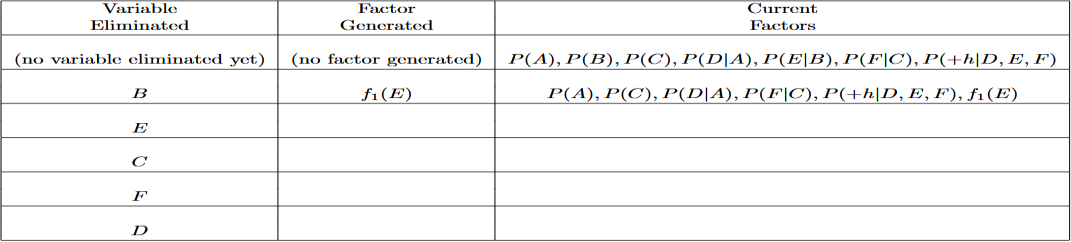
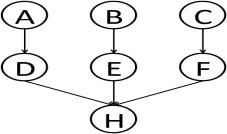
Answer: 0.8

P (A = +a, B = −b, C = −c, D = +d) = Answer: 0.8 ∗ 0.2 ∗ 0.1 ∗ 0.5 = 0.008

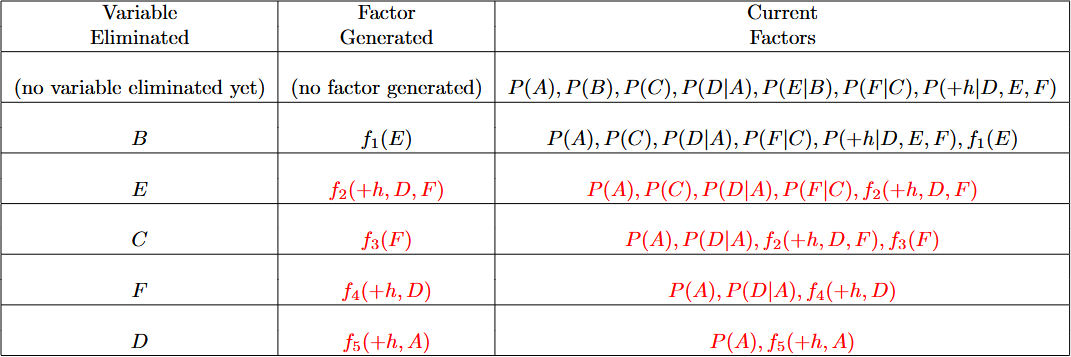
P (A = +a|B = −b, C = −c, D = +d) =

Answer: ((0.8∗0.2∗0.1∗0.5)/(0.8∗0.2∗0.1∗0.5+0.2∗0.5∗0.1∗0.5)) = 0.615

1. For the Bayes’ net shown below, consider the query P (A|H = +h), and the variable elimination ordering B, E, C, F, D.
2. In the table below fill in the factor generated at each step — we did the first row for you.



1. Which is the largest factor generated? Assuming all variables have binary-valued domains, how many entries does the corresponding table have?

Answer: f2(+h, D, F ), its table has 22 = 4 entries

# Consider the geometric distribution, which has P (X = k) = (1 − θ)k−1θ. Assume in our training data X took on the values 4, 2, 7, and 9.

Write an expression for the log-likelihood of the data as a function of the parameter θ.

Answer: L(θ) = P (X = 4)P (X = 2)P (X = 7)P (X = 9) = (1 − θ)3θ(1 − θ)1θ(1 − θ)6θ(1 − θ)8θ = (1 − θ)18θ4

log L(θ) = 18 log(1 − θ) + 4 log θ

# Definitions:

1.Answer: Supervised learning involves the use of labeled data to teach a machine to recognize patterns and make predictions. Unsupervised learning is the use of unlabeled data to allow the machine to explore and find patterns in the data.

2.Answer:A neural network is an AI system that is made up of interconnected layers of neurons that are used to recognize patterns and make predictions. A deep learning system is a type of neural network that uses multiple layers of neurons to process large amounts of data and make more accurate predictions.

3.Answer: Supervised learning, unsupervised learning, semi-supervised learning.

4.Answer:As a student in computer engineering or computer science, incorporating artificial intelligence (AI) into my studies can significantly enhance my skills and expand my career opportunities. Start by understanding the fundamentals of AI, including machine learning, neural networks, natural language processing, and computer vision, while also strengthening my knowledge in mathematics and statistics, which are essential for many AI algorithms. Engage in hands-on projects by implementing basic AI algorithms using programming languages like Python and libraries such as TensorFlow, Keras, or PyTorch. Participating in hackathons can also provide valuable practical experience. Familiarize myself with popular AI frameworks and tools, and learn to use version control systems like Git for collaboration. Seek out research opportunities with professors or researchers to gain insights into real-world AI applications, and contribute to open-source projects on platforms like GitHub to build my portfolio. Explore how AI is applied in various fields such as healthcare, finance, and robotics, and analyze case studies of successful implementations to understand best practices. Additionally, join AI-focused communities, attend workshops and conferences to network with professionals, and stay updated with the latest research and trends by following AI journals and thought leaders in the field. By building a solid foundation, gaining hands-on experience, collaborating with others, and engaging with the AI community, you can develop the necessary skills to thrive in this rapidly evolving field.