

Fundamentals of Artificial Intelligence

COSC 4368

Some Solution Sketches

Final Exam

Monday, May 6, 2p

*Name:*

*SSN:*

1. Naïve Bayes and Belief Networks (17 points)
2. FOPL as a Language (10 points)
3. Support Vector Machines (7 points)
4. FOPL Theorem Proving Using Resolution (13 points)
5. Multi-Agent Systems (5 points)
6. Ethics for AI (8 points)
7. Supervised Learning and Neural Networks (13 points)
8. AI in General (7 points)

Point Total (out of 73):

Number Grade:

The exam is “open books” and you have 120 minutes to complete the exam.

1) Belief Networks and Naïve Bayes [17]

a) Consider the following belief network that consists of variables A, B, C, D, E all of which have two states {true, false} and whose structure is depicted below is given.

 B

A D E

 C

1. Is A|∅ d-separable from E|∅[[1]](#footnote-1). Given reasons for your answer! [4]

There are two paths:

A-C-D-E and A-B-D-E

The first path is blocked in node C (pattern3 as C in not in evidence) and

The second path is blocked in B (pattern3 as B is not in evidence)

As both paths are block, A and E are independent

1. Is B|A d-separable from C|A. Give reasons for your answer! [4]

There are two paths: B-D-C and B-A-C

The first path is not blocked in node D, as D is not in evidence

The second path is blocked in A (pattern2 as A is not in evidence) if they do not write that line they still deserve full credit

As the first path is not blocked B|A is not d-separable (independent) from C|A

1. What advantage you see in using Belief Networks instead of using a naïve Bayesian approach? [4]

Using Belief-networks you can express dependencies between random variables that are not independent of each other, using domain specific knowledge[2]. By doing that BBNs will obtain “better”, more accurate predictions than the naïve Bayesian approach, as making conditional independence assumptions that are violated by the observed data will increase prediction errors.

1. Assume P(D)=0.02, P(S1)=0.2 P(S1|D)=0.4 P(S2)=0.1 P(S2|D)=0.3. Compute P(D|S1,S2) using a Naïve Bayesian approach! [3]

P(D|S1,S2)=0.02\*2\*3=0.12 No partial credit!

1. What specific conditional independence assumptions are made in your computations for problem iv)? [2]

S1 is independent of S2[1], and S1|D is independent from S2|D[1].

2) FOPL as a Language [10]

Express the following natural language statements using first order predicate calculus formulas:

1. Flipper is an intelligent dolphin. [2]
2. There is a dog-catcher in Texas that got bitten by a dog in each county in Texas. [4]
3. No student who took the final exam of COSC 4368 got a grade of ‘C’ in the final exam. [4]
4. dolpin(Flipper) ∧ intelligent(Flipper) no partial credit; using ∃ is wrong!
5. ∃x (dogcatcher(x) ∧ lives(x,Texas) ∧ ∀c (county(c,Texas)🡪 ∃d bitten(x,d) ∧ located(d,c)))
6. ~∃s (took-final(COSC4368,s) ∧ grade(s, COSC4368,Final,C)); might also give full credit for ~∃s grade(s, COSC4368,Final,C))

Can give partial credit up to 1 point for c and up to 2.5 points for b!

3) Support Vector Machines [7]

Assume a support vector machine hyperplane that has been learnt for a dataset having attribute A, B, C is given:

A2+B3-C4-2

e.g. the support vector machine classifies an example ex1 (0,1,1) where 0, 1, 1 are the values for attributes A, B, C as belonging to the negative class as 0\*2+1\*3-4\*1-2=3.

Assume we have 3 more examples ex2, ex3, ex4 are given for which the hyperplane equation returns -20, 0, +4. What does this tell you about the examples [3].

ex1and ex2 are on the side of the negative class of the hyperplane[1], but ex2 is much further away from the hyperplane than ex1[0.5]; ex3 is located on the hyperplane[1], and ex4 is one the other (positive class) side if the hyperplane [0.5]

Can the fact that the support vector machine equation returns a negative or positive number—and not just a class label of the predicted class—be used for anything useful? [4]

If the values returned for the hyperplane equation for a testing example e is very close to 0 we are much less confident about the correctness of the SVM’s prediction compared to the case where plugging the attribute values of e into the hyperplane equation returns large negative or positive values.

Other way to say it: Closeness to 0 can be used to assess a degree of confidence the SVM has with respect to the correctness of the prediction it made for e!

4) Resolution for FOPL [13]

a) Show using Resolution (and not by using other methods!):

* + 1. ∀y (R(y,y) 🡪 P(2,y,y))
		2. ∀x∀y∀z (P(x,y,z) 🡪 R(y,z) )
		3. ∀x∀y (R(x,y) 🡪 R(y,y) )
		4. ∀x ∃z (P(x,x,z) 🡪 R(x,z))
		5. ∃z ∀x∀y (P(z,x,y) 🡪 ~R(x,y))
		6. P(4,5,6)

|-

(X) P(2,6,6)

i) First transform the FOPL formulas into clauses! [5]

ii)Next, using the resolution method try to infer the empty clause! [5]

1. ~R($y,$y) v P(2,$y,$y)) [1]
2. ~P($x,$y,$z) v R($y,$z) [1]
3. ~R($x,$y) v R($y,$y) [1]
4. ~P($x,$x,Z($x) v R($x,Z($x)) [1]
5. ~P(Z,$x,$y) v ~R($x,$y)) [1]
6. P(4,5,6)
7. ~ P(2,6,6)
8. ~R(6,6) using (7) and (1)
9. R(5,6) using (6) and (2)
10. R(6,6) using (9) and (3)
11. empty clause using (10) and (8)

Can give partial credit of up to 2.5 points for incorrect proofs!

b) What role does unification play in proving FOPL theorems using resolution? [3]

Unification computes the most general substitution that makes two clauses with match variables equal [1.5]; the resolution proof methods finds a clause and a negated clause that unify, and obtains a new clause by applying the obtained substitution to the remaining clauses.[2.5]

At most 3 points! Other answers might deserve partial credit!

5) Multi-Agent Systems [5]

Describe an application that might benefit from using Multi-Agent System technology. Explain briefly how the application benefits in particular from using multi-agent technology! Limit your answer to 5-6 sentences! [5]

No Answer given!

6) Ethics for AI [8; up to 2 extra points]

Pick an AI Ethics Problem of you own preference. Briefly describe the problem and then propose what should be done in the future to alleviate the problem. Limit you answer to 8-10 sentences (8 points, and up to 2 extra points).

No Answer given!

7) Supervised Learning and (Deep) Neural Networks [13]

a) What role do validation sets play in Supervised Learning? [3]

It is used during training to determine the optimal setting for the employed machine learning algorithm’s (hyper) parameters.

If the do not mention ‘during training’ only 2 points!

b) Give one reason[[2]](#footnote-2) why deep neural network technology has been quite successful in the last 5 years in learning classification models for challenging datasets that exhibit quite high accuracies, Explain the reason you give in 2-3 sentences! [4]

no answer given; look at slideshow on deep learning!

c) 10-fold cross-validation is usually repeated 10 times; assume you use leave-one-out cross-validation instead: should it also be repeated 10 times? Give a reason for your answer! [3]

It does not make any sense to repeat it [1.5] as the training-test-set pairs used in leave-one-out cross validation are always the same[1.5]; that is if we repeat leave-one out, we get 10 times the same result!

d) Neural networks that use at least one (or a lot) intermediate layers have been much more successful than 2-layer neural networks which directly link inputs with outputs. Why do you believe this is the case? [3]

The intermediate layer allows to create new features that facilitate getting high accuracies [3].

Other answers might deserve partial credit to up to 2 points.

8) AI in General [7]

Recently, foreign governments have assigned a lot of resources to AI education, research and development. For example, the Chinese Government wants “*AI to be Made in China*” by 2030. What do you believe are the reasons for this development? Limit your discussion to 5-8 sentences!

No Answer given!

1. ∅ represents “no evidence”; question i basically asks if A and E are independent; that is, if P(A∧E)=P(A)\*P(E) [↑](#footnote-ref-1)
2. There might be multiple reasons; however, you will do not get more points by descibing a second, third,… reason! [↑](#footnote-ref-2)