ASSIGNMENT 5

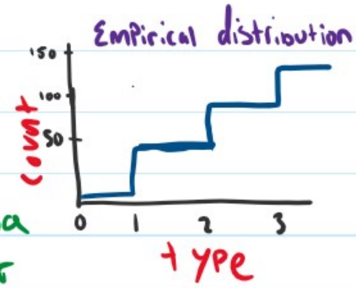
Monday, April 30, 2018 1:36 PM

subset of data collected

1. Compute the empirical distribution of the three types of irises in the iris dataset (see part II). Using the empirical distribution, compute the entropy of the iris species.

Empirical distribution:

In statistics, an empirical distribution function is the distribution function associated with the empirical measure of a sample. This cumulative distribution function is a step function that jumps up by 1/n at each of the n data points. Its value at any specified value of the measured variable is the fraction of observations of the measured variable that are less than or equal to the specified value



number_of_samples = 150

number_of_classes = 3 types

number_of_class_one_samples = 50 Setosa

number_of_class_two_samples = 50 Versicolor

number_of_class_three_samples = 50 Virginica

$P(V_k) = \left(\frac{1}{\text{total samples}}\right) * \sum$ # of class K samples as labels in data set

$P(\text{Setosa}) = \frac{1}{150} * 50$

$P(\text{Versicolor}) = \frac{1}{150} * 50$

$P(\text{Virginica}) = \frac{1}{150} * 50$

Note:

Log Rules:

$$(-1)\cdot\left(\frac{1}{n}\right)\log_2\left(\frac{1}{n}\right) == (+1)\cdot\left(\frac{1}{n}\right)\log_2\left(\frac{n}{1}\right)$$

Entropy is the measure of uncertainty of a random variable

Acquisition of information corresponds to a reduction in entropy.

$$H(v) = \sum_{K=3} P(v_k)\log_2\frac{1}{P(v_k)} = -\sum_K P(v_k)\log_2(P(v_k)) = +\sum_K P(v_k)\log_2\left(\frac{1}{P(v_k)}\right)$$

if # of classes in data set = 3

take (-) away with log rules

$$\underbrace{H(x) = -\frac{p}{p+n}\log_2\frac{p}{p+n} + -\frac{n}{p+n}\log_2\frac{n}{p+n}}_{1 \text{ cass}} = 3*\left(\left(\frac{1}{3}\right)\log_2(3) + \left(\frac{2}{3}\right)\log_2\left(\frac{3}{2}\right)\right)$$

2. Regularization is a model selection technique.
This approach searches for a hypothesis that directly minimizes the weighted sum of emirical loss and the complexity of the hypothesis, which can be called the total cost:

$$\text{Cost}(h) = \text{EmpiricalLoss}(h) + \lambda\,\text{complexity}(h)$$

$$\hat{h}^* = \underset{h \in H}{\text{argmin}}\ \text{Cost}(h)$$

λ is a hyper-parameter; it serves as a conversion rate between loss and hypothesis complexity, because each hypothesis are not measued on the same scale.

This approach combined loss and complexity into one metric, allowing the computer to find the best hypothsis all at once.

We still have to compute the cross validation search to find the hypothesis that generalizes best, but we swap the size attribute with λ attribute. We select the value of λ that gives us the best validation set score.

The process of explicitly penalizing complex hypothesis is the idea of Regularization because it looks for a function that is more regular or less complex.

Note: The choice of regularization function depends on the hypothesis space.

This is important because it makes sure you do not overfit your model, so it can be applied to the real world.

3. Steps: Reasoning:

Steps	Reasoning
irrigation	Fact
Water	irrigation => Water
Sun	Fact
nutrients	Fact
Flower	nutrients ^ Water ^ sun => flower
bat	Fact
Pollinator	bat v bird v bee => pollinator
∴ Fruit	Flower ^ Pollinator => Fruit

4. a. Percepts:

- Breeze ; directly adjacent from Pit
- Stench ; directly adjacent from Wumpus
- glitter ; in square w/ gold

b. Propositional logic → the way in which the truth of sentence(s) is determined

Part one: Syntax

Backus-Naur form (BNF) four components:

- terminal symbols ;
- nonterminal symbols ;
- Start symbol
- rewrite rules ;

Of form LHS ⟶ RHS

- LHS - nonterminal
- RHS - sequence of zero or more symbols

Note: Symbols either terminal or nonterminal or ϵ (empty string)

Rewrite rule of the form: Sentence → Noun Phrase Verb Phrase

subject, object

Part of sentence containing verb & any indirect sentence object.

rewrite rule means whenever we have two strings categorized as a NounPhrase and VerbPhrase, we can append them together and categorize the results as a Sentence.

As an abbreviation two rules $(S \to A)$ and $(S \to B)$ can be written as $(S \to A|B)$ and

Expr $\to$ expr operator expr | (expr) | Number

Number $\to$ Digit | Number Digit (aka string of digits)

Digit $\to$ 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

Operator $\to$ + | - | ÷ | * | ^ | v |

- $\Rightarrow$ implies or called an implication

  i.e. $\underbrace{(W_{13} \wedge P_{31})}_{\text{Premise}} \Rightarrow \underbrace{\neg W_{22}}_{\text{Conclusion}}$

  implication also known as if-then & rules statements

- $\Leftrightarrow$ biconditional (if and only if)

  i.e $W_{1,3} \Leftrightarrow \neg W_{2,2}$

Sentence $\to$ Atomic Sentence | Complex Sentence

Atomic Sentence $\to$ True | False | P | Q | R | (single symbol w/ no operator tagged on)

Complex Sentence $\to$ *(Sentence == S); (S) | [S] | ¬S | S ∧ S | S ∨ S | S → S | S ⇔ S

The BNF Grammar by itself is Ambiguous;

a sentence with several operators can be parsed by grammar in multiple ways

To eliminate ambiguity we define a Precedence for each operator

Operator Precedence: $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$

$\neg$ binds most tightly in example $\neg A \wedge B == (\neg A) \wedge B$

rather than $== \neg(A \wedge B)$

When in doubt use parentheses!

rather than -- $\neg(P \wedge Q)$

When in doubt use Parentheses

Part two: Semantics

The Semantics (Note: Plural) defines the rules for determining the truth of a Sentence with respect to a Particular model.

In Propositional logic, a ==model== Simply fixes the truth value - **true** or **false** - for every Propositional Symbol.

$$m_1 = \{ P_{1,2} = \text{false},\ P_{2,2} = \text{false},\ P_{3,1} = \text{true} \}$$

With three Propositional symbols there are $2^3 = 8$ Possible models.

The Semantics for Propositional logic Must Specify how to compute the truth of any sentence, given a model.
This is done recursively. All sentences are constructed from atomic sentences.

All Sentences are Constructed from atomic Sentences & five Connections; therefore,

we need to Specify how to:

- Compute the truth of atomic Sentences
- Compute truth of sentences formed with each of 5 connectives

atomic Sentences:

- True is true in every model & False is false in every model
- The truth value of every other Proposition Symbol Must be Specified directly in model. i.e in m_1 $P_{1,2}$ is False.

- The truth value of every other Proposition Symbol must be specified directly in model. i.e in M_1, $P_{1,2}$ is False

Complex Sentences:

We have 5 rules, which hold for any subsentences P and Q in any model m (iff :: if & only if)

- $\neg P$ is true iff P is false in m
- $P \wedge Q$ is true iff P and Q are true in m
- $P \wedge Q$ is true iff either P or Q is true in m
- $P \Rightarrow Q$ is true ~~iff~~ unless P is true and Q is false in m

read: "If P is true then I am claiming that Q is true. Otherwise I am making no claim."

The only way for this sentence to be false is if P is true & Q is false

Propositional logic does Not require any relation of causation or relevance between P and Q

- $P \Leftrightarrow Q$ is true iff P and Q are both true or both false in m

$P \Leftrightarrow$ is true whenever both $P \Rightarrow Q$ & $Q \Rightarrow P$ are true.

i.e. 1.

sentence $\neg P_{1,2} \wedge (P_{2,2} \vee P_{3,1})$ evaluated in M_1

gives true $\wedge$ (false $\vee$ true) = true $\wedge$ true = true

i.e. 2.

Many of the rules of the wumpus world are best written using $\Leftrightarrow$.

a square is breezy if a neighboring square has a Pit, and a square is breezy only if a neighboring square has a Pi

$$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$$

$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$.

Subjects: noun, or Pronoun Performing a Verb in a Sentence.

Predicates: Verb and its Pals

i.e. $\boxed{\text{I}}$ $\boxed{\text{bought a } \boxed{\text{Can of Peaches}}}$.

Subject
Predicate

Note: can of peaches is a noun but its not the noun Performing a verb. i.e. its a Predicate.

b

- write a set of logical sentences that describe the Predicates that can be infered by Percepts at 2,3.

~~Model$_1$ (B_{23} = True S_{23} = True, G_{23} = True, P_{23} = False~~ not what was Asked

$$B_{2,3} \Leftrightarrow (\underbrace{P_{1,3} \vee P_{2,4}} \vee P_{3,3} \vee P_{2,2})$$

Note: $P_{1,3}$ & $P_{2,4}$ can both be true but remember!

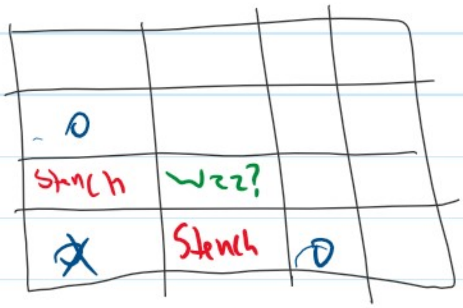
$$G_{2,3} \Leftrightarrow G \wedge \neg(G_{1,1} \wedge G_{2,1} \wedge G_{3,1} \wedge G_{4,1} \wedge G_{1,2} \wedge G_{2,2} \wedge G_{3,2} \wedge G_{4,2} \wedge G_{1,3} \wedge G_{3,3} \wedge G_{4,3} \wedge G_{1,4} \wedge G_{2,4} \wedge G_{3,4} \wedge G_{4,4})$$

$$S_{2,3} \Leftrightarrow (W_{1,3} \vee W_{2,4} \vee W_{3,3} \vee W_{2,2})$$

5. Starting Sentences: $\neg P_{1,1}$ and $\wedge$ arrow

5. Starting Sentence: $\neg P_{1,1}$ and $\wedge$ arrow
additional info: $S_{2,1}$ and $\wedge$ $S_{1,2}$
Don't forget What were trying to Prove is a Sentence that will be negated
Write Knowledge base in Conjunctive normal form and Show whether or not question $W_{2,2}$ is entailed by the knowledge base using the resolution rule.

goal decide $Kb \models \alpha$ for some Sentence α.



R_1: $\neg P_{1,1}$
R_2: arrow
R_3: $S_{2,1}$
R_4: $S_{1,2}$

$Kb = \neg P_{1,1} \wedge arrow \wedge S_{2,1} \wedge S_{1,2}$
Show $Kb \models W_{22}$ through

Write kb in conjunctive normal resolution rule

Complex sentences to use only $\vee$ operator
Logical equivalence:

Logical equivalence:

two sentences α & β are equivalent if they are true only if each of them entails the other

$\alpha \equiv \beta$ iff $\alpha \models \beta$ and $\beta \models \alpha$

Validity: a sentences is valid if it is true in all models
i.e. $P \vee \neg P$ is valid

Valid sentences are also known as tautologies, they are necessarily true.

Because the sentence True is true in all models, every valid sentence is logically equivalent to True.

Rule: For any sentences α and β, $\alpha \models \beta$

Proof By Contradiction: Iff the sentence $(\alpha \Rightarrow \neg\beta)$ is Valid

$$\overbrace{(\neg W_{11} \wedge \neg P_{1,1} \wedge S_{2,1} \wedge S_{1,2})}^{\alpha} \Rightarrow \overbrace{(\neg W_{22})}^{\beta}$$

Put in conjunctive normal, solve w/ resolution rule.
Hint: Break up sentences to make more inferences that get added to knowledge base

$\neg(\neg W_{11} \wedge \neg P_{1,1} \wedge S_{2,1} \wedge S_{1,2}) \vee (\neg W_{22})$: $\neg\alpha \vee \beta$ implication elimination

$(W_{11} \vee P_{1,1} \vee \neg S_{2,1} \vee \neg S_{1,2}) \vee (\neg W_{22})$: Demorgans

$W_{11} \vee P_{1,1} \vee \neg S_{2,1} \vee \neg S_{1,2} \vee \neg W_{22}$