Machine Learning HW #1

R01942135 電信所 周伯威

1. (a)
   - (i) Not suitable. There’s an easy definition about whether a number is a prime. However we all hope that one day in the future, machine will be intelligent enough to tell whether a number is a prime.
   - (ii) Yes. Pattern: custom behavior. Definition: not easily programmable. Data: history of bank operation and abnormal ones.
   - (iii) No. There’s a programmable definition and we all already know the underlying rules, i.e. Newton's law of universal gravitation.
   - (iv) Yes. Pattern: from all driver’s behavior. No easy definition. Data: all the history of traffics
   - (v) Yes. There’s some unknown underlying pattern, no easy definition, but it can be learned from medical records.
2. (b)
   The recommendation can be learned from partial / implicit information, sequentially. Most importantly, user feedback only serve as “goodness” rather than a definitive label or answer.
3. (c)
   Every credit card owner has their own purchasing behavior and each one of that is quite different from one another, and there’re generally no labels.
4. (a)
   Teacher will tell you which is wrong and which is right.
5. (a)
   Training labels come from those who cannot pay off the debt. Therefore an supervised learning algorithm is best suitable for it.

6. (d)
   (1) Error only comes from even points.
   (2) The floor of x/2, is the number of odd number below x.
   Therefore, by subtracting $\left\lfloor\frac{N+L}{2}\right\rfloor - \left\lfloor\frac{N}{2}\right\rfloor$, we get the total number of errors.

N	N+L	O X O X O X O
odd	odd	↑ ↑
odd	even	↑ ↑
even	odd	↑ ↑
even	even	↑ ↑

7. (c)

   From lecture slides p.5 in 04_handout.pdf, we can choose +1/-1 for the additional L points as long as the N points are satisfied $\rightarrow 2^{\mathrm{L}}$.

8. (a) (c)

   For (a), (b), the number of combinations of picking k points out of L is $\binom{L}{k}$.

   For (c), (d), the probabilities are all the same, and therefore the expected value of OTS is an OTS of one of them.

9. (a) It's a binomial distribution.

$$\left(\frac{1}{2}\right)^{10}\binom{10}{5} \cong 0.246$$

10. (b)

$$0.1^{1}0.9^{9}\binom{10}{1} \cong 0.39$$

11. (d) 1 orange + 0 orange

$$0.1^{9}0.9^{1}\binom{10}{1} + 0.1^{10}\binom{10}{0} = 9.1 \cdot 10^{-9}$$

12. (b) From Hoeffding's inequality,

$$\frac{1}{2}P(|v - 0.9| > 0.8) \leq e^{-2\cdot 0.8^{2}\cdot 10} = 5.52 \cdot 10^{-6}$$

13. (b) Only B-type and C-type have orange 1 on dice. Therefore

$$\mathrm{P} = \left(\frac{2}{4}\right)^{5} = \frac{8}{256}$$

14. (c)

Orange 1	Orange 2	Orange 3	Orange 4	Orange 5	Orange 6
BC	AC	BC	AD	BD	AD

   Consider the cases, 4X1Y, 3X2Y, 2X3Y, 1X4Y, for (X, Y) = (B, C), (A, C), (A, D) and (B,D). There're $4 \times (5 + 10 + 10 + 5) = 120$.

   There're 4 left to be considered, i.e. 5A, 5B, 5C, 5D. => 124 in total.

For Problem 15-20, please download my MATLAB codes from CEIBA. Just run "go.m" and you'll get all the computed outputs for each problem.