# MECON6102-
# 2019 exam

Dr. Ye Luo

March 24, 2022

## 1 Multiple Choice

[50pts]

1. Suppose you would like to estimate a model that predicts $y$ from $x$, and you think there may be non-linearity. What can you do to test non-linearity?

(a) Run a regression of $y$ on $x$.

(b) Run a neural networks.

(c) Run an AIC/BIC using non-linear transformations.

(d) Run a logistic regression.

Answer: (c) - you can test non-linearity by adding non-linear transforms of the original variables.

2. Which models is most powerful for risk prediction in modern finance?

(a) Linear regression.

(b) Logistic regression.

(c) Support vector machine.

(d) Kernel support vector machine.

Answer: (d). (b), (c) are linear.

3. What does support vector machine do exactly?

(a) Minimize classification error.

(b) Minimize Mean Squared Error.

(c) Maximize Separating gap.

(d) None of above.

Answer: (c).

4. In the decision theory, what does temporal difference algorithm do?

(a) Learning the Q function based on reward.

(b) Learning the value function V based on reward

(c) Learning the Q function based on the difference between reward and expected reward.

(d) Learning the value function V based on the difference between reward and expected reward.

Answer: (c) and (d).

5. What is Not right for decision tree?

(a) Good for interpretation.

(b) Robust decision.
(c) More flexible than rule based systems.
(d) Can learn from data.
Answer: (b). RF is robust, DT is not.
6. What methods are Not often used in Fraud detection?
(a) Linear regression.
(b) Logistic regression.
(c) Neural Networks.
(d) Random Forest.
Answer: (a) - linear regression is for prediction rather than classification.
7. In random forest algorithm, which is Not ture:
(a) You need to learn many decision trees.
(b) You learn the decision tree from randomize data.
(c) The final decision depends on majority votes from the trees.
(d) You learn as many trees are the sample size.
Answer: (d).
8. Which of the following is true for clustering:
(a) Clustering can be used to directly predict returns.
(b) Clustering needs labels to be predicted.
(c) K-means algorithm is one way to do clustering.
(d) Support vector machine is one way to do clustering.
Answer: (c). (Note: we haven't discussed clustering this semester, so just ignore this question)
9. In high dimensional data, which is true:
(a) You can not use linear regression.
(b) You can not use LASSO.
(c) You can not use AIC/BIC.
(d) You can not use Deep Learning.
Answer: (a). In high dimensional data, you must perform model selection.
10. What is Not the feature of modern natural language processing?
(a) The words are vectorized by clustering.
(b) You can use modern nlp to analyze alternative financial data.
(c) The modern NLP is rule based.
(d) The modern NLP is data based.
Answer: (c) (Note: we haven't discussed nlp this semester, so just ignore this question) .

## 2 Short Essay

[50pts]
1. Briefly discuss the difference between modern AI and human intelligence.
Answer:

- Modern AI is data-driven, based on big data/simulation data and a great amount of computing power. Human intelligence relies on less data.
- Modern AI is single tasking, human intelligence is multi-tasking.
- AI is experience driven, Human intelligence relies both on experience and models.
- Modern AI learns from scratch, human learns from each other (knowledge).

2. Write down the formulation of random forest, and briefly discuss the advantage of it.

- For $B = 1, 2, ..., M$, random draw sample $(x_1^B, y_1^B)..., (x_n^B, y_n^B)$ from original sample $(x_1, y_1), ..., (x_n, y_n)$.
- Train decision tree based on the sample $(x_1^B, y_1^B)..., (x_n^B, y_n^B)$, denote as $h^B(x)$.
- Make decision based on the voting rule $h(x) := 1(\frac{1}{M}\sum_{B=1}^{M} h^B(x) > c)$ for some cut-off $c$, e.g., $c = 0.5$.

Advantage:

- More robust than decision trees, better out-of-sample performance.
- Great model explainability.
- Non-linear modeling.