

經濟部所屬台灣電力股份有限公司 109 年新進博士級人員甄試

類別：資訊工程

科目：資訊工程

| | |
|------|---|
| 注意事項 | <ol style="list-style-type: none">1. 本試題共 2 頁(A4 紙 1 張)。2. 可使用本甄試簡章規定之電子計算器。3. 本試題分 7 大題，每題配分於題目後標明，共 100 分。須用藍、黑色鋼筆或原子筆在答案卷指定範圍內作答，不提供額外之答案卷，作答時須詳列解答過程，於本試題或其他紙張作答者不予計分。4. 本試題採雙面印刷，請注意正、背面試題。5. 考試結束前離場者，試題須隨答案卷繳回，俟本節考試結束後，始得至原試場或適當處所索取。6. 考試時間：150 分鐘。 |
|------|---|

一、新冠肺炎測試中，假設偽陽率是 0.1%。假設 99% 生病的人都抓得到，也假設地球 60 億人有 3000 萬人染病，所以每 200 人有 1 人染病。今天發布 1 人機場檢驗呈陽性反應，真的嗎？請回答是真的機率為何？（8 分）

二、在機器學習中，模型通常會遇到 Bias-Variance tradeoff 問題：（每題 5 分，合計 20 分）

- （一）請闡述 Bias error 如何產生？
- （二）請闡述 Variance error 如何產生？
- （三）請問一個模型的 total error 由哪 3 種 error 組成？
- （四）請在模型複雜度(x-axis)對上模型錯誤率(y-axis)座標上畫出 bias、variance、以及 total error 之趨勢圖。

三、請回答以下關於 ensemble learning 之問題：（每題 5 分，合計 10 分）

- （一）請解釋什麼是 ensemble learning？
- （二）請解釋 Random forest 所使用的技巧？

四、自然語言處理中，為了使字元之初始 one-hot 向量能夠找出相關性，科學家使用 word embedding 技法，請解釋以下幾種常見方法之運作原理：（每題 5 分，合計 10 分）

- （一）Continuous Bag-of-Words Model (CBOW)
- （二）Continuous Skip-gram Model (Skip-gram)

五、比特幣(Bitcoin)以及以太坊(Ethereum)皆使用 Proof of Work 共識機制來避免 double spending 問題，但其交易模型不同，比特幣使用 UTXO，而以太坊使用 Account based 交易模型。請回答以下問題：（每題 5 分，合計 20 分）

- （一）請闡述 Proof of Work 機制。
- （二）請說明 Selfish mining attack 如何運作？
- （三）請闡述 Unspent Transaction Output (UTXO)交易模型。
- （四）請闡述 Account-based 交易模型。

六、區塊鏈系統之背後原理為分散式系統之理論，請回答以下問題：

- (一) 請解釋 CAP 定理，以其闡述其中 C、A、P 代表之意。(5 分)
- (二) 在分散式系統理論中，Fischer 等人證明 FLP impossibility，但後人使用其他方法來規避完全非同步共識演算法之限制，請舉出 2 種方法並各舉 1 共識演算法為例。(8 分)
- (三) 分散式系統中共識演算法會定義節點行為，其一為故障容錯，另一為拜占庭容錯，請說明此 2 種節點之行為。(5 分)
- (四) 請舉 1 例來說明 Paxos 共識演算法無法容忍任一個拜占庭容錯節點。(6 分)

七、給定以下字元以及出現之頻率，請畫出其 optimal Huffman Codes 之樹狀表示結構。(A : 3、B : 2、C : 5、D : 4、E : 1、F : 6) (8 分)