《資訊管理與資通安全概要》

一、機器學習（Machine Learning）是人工智慧（Artificial Intelligence）的一個分支；深度學習（Deep Learning）則是機器學習的一個分支。

（一）請說明深度學習與機器學習的差異性。（10分）

（二）請說明深度學習模型之基本概念。（10分）

（三）深度學習的辨識效能仰賴大量且有效的訓練資料集。請說明為何深度學習的辨識效能效佳。（10分）

試題評析	近年來由於人工智慧的崛起，與人工智慧有關的名詞和領域越來越常出現在考題中，此題的機器學習、深度學習便是。故近幾年在準備考試時，需要多關注與此有關的議題與名詞，當有常出現名詞時，要能夠理解其基本意義並熟記，依照此種準備方式即可應付大部分相關的時事考題，並且在答題上要能夠掌握機器學習的基本理論與架構，和它相關的其他理論。
考點命中	《高點・高上資訊管理與資通安全講義》第二回，蕭老師編撰，頁24-27。

答：

(一)機器學習為透過從過往的資料和經驗中學習並找到其運行規則，最後達到人工智慧的方法，透過樣本訓練機器辨識出運作模式，而不是用特定的規則來編程。換句話說，機器學習是一種弱人工智慧，它從資料中得到複雜的函數(或樣本)來學習以創造演算法(或一組規則)，並利用它來做預測。其也是一門多領域交叉學科，涉及機率論、統計學、逼近論、凸分析、計算複雜性理論等多門學科。其理論主要是設計和分析一些讓電腦可以自動「學習」的演算法機器學習是可以尋找適合讓電腦做預測或數學模型分類的一種演算方法。這種演算方法主要透過蒐集大量原始數據與標準答案，以訓練資料調整且選擇相應的數學模型，同時並藉由驗證資料比對計算分類結果，來判定模型是否適合用來預測或分類。

深度學習為機器學習的分支，為一種實現機器學習的技術。其試圖使用包含複雜結構或由多重非線性變換構成的多個處理層對資料進行高層抽象的演算法，並將其中函數作多重非線性轉換，使之增加高度抽象化資料、記憶資料影響能力。又可分為監督式、非監督式或半監督式的特徵學習方法和特徵提取高效演算法來替代手工取得特徵。

(二)深度學習的基本概念為機器學習中的分散表示（distributed representation），其假定觀測值是由不同因子相互作用生成。而深度學習更進一步假定這一相互作用的過程可分為多個層次，代表對觀測值的多層抽象。不同的層數和層的規模可用於不同程度的抽象。深度學習運用了這分層次抽象的思想，更高層次的概念從低層次的概念學習得到。這一分層結構常常使用貪婪演算法逐層構建而成，並從中選取有助於機器學習的更有效的特徵。不少深度學習演算法都以無監督學習的形式出現，因而這些演算法能被應用於其他演算法無法企及的無標籤資料，這一類資料比有標籤資料更豐富，也更容易獲得，這也是深度學習重要的優勢。

(三)深度學習透過貪婪演算法並分層建構而成，從中選擇有助於機器學習的有效特徵，而深度學習中所採用的分層次的抽象方法，會讓較高層次的概念從較低層次的概念中學習。之後再採用無監督式訓練來構建神經網路，用以發現有效的特徵，此後再採用監督式的反向傳播以區分有標籤資料。當足夠多的層數被學習完畢，這一深層結構生成為一個模型，可以通過自上而下的採樣重構整個資料集，此模型在高維度的結構化資料上能夠有效地提取特徵。

二、因應巨量資料的來臨，資料不僅在數量上變多，而且日益複雜，資料倉儲系統面臨許多挑戰與改變。

（一）請說明料倉儲系統的主要功能。（10分）

（二）面對巨量資料收集與分析，請說明巨量資料處理所需之主要技術與硬體架構。（10分）

試題評析	巨量資料為一個可深可淺的主題，大部分市售書與講義都有很完整的說明其概念與大致功能，所以第一題對同學來說不會有太大的困難。 第二題就是較深入的考題，需要對巨量資料的處理與架構有基礎的認識，並且能夠說明。而技術與架構現在大多使用Hadoop此方式，故除了一般的介紹要理解外，也要多熟悉上課之補充。

考點命中 《高點・高上資訊管理與資通安全講義》第二回，蕭老師編撰，頁19-23。

答：

(一)資料倉儲能將使用者所需要的資料做有效的管理，讓資料的存取更為有效率並且利用得到資料做進一步的分析並將資料轉換成對企業有利的資訊，其具有主題導向（Subject-Oriented）、整合性（Integrated）、不會變動（Non-Volatile）和隨時間變異（Time-Variant）之特性，其具體功能為：

1.讓組織可以更容易取得最新與最廣泛的資料，免於傳統文件的往返時間。

2.對於資料模型(Model)重新模式化的能力。

3.可以滿足即時決策之需求。

4.存取資料倉儲的資料不會影響正在執行的作業與效率。

(二)目前大多巨量資料的處理是利用Hadoop，其核心技術原理是源自Google打造搜尋引擎的關鍵技術，在Hadoop平臺中，核心用途是儲存空間的資源管理，以及記憶體空間和程式排程的安排。透過分散式架構的HDFS檔案系統、搭配可分散運算的MapReduce程式演算方法，可以將多臺一般商用等級的伺服器組合成分散式的運算和儲存叢集，來提供巨量資料的儲存和處理能力。其架構包含：

1.HDFS (Hadoop Distributed File System)

是Yahoo提出的一套分散式檔案系統，可將一整個叢集(Cluster)視為一台電腦，進行檔案存取的操作。其組成元件為：

(1)Name Node

Name Node(NN)為HDFS的核心元件，管理整個HDFS(檔案讀取寫入...等操作)，整個叢集裏面只會有一個NN，所以當其當掉時，會讓整個HDFS無法運作。

(2)Secondary NameNode

HDFS在系統一開始的時候會讓NN讀取相關的狀態進到記憶體中，並且編輯Edit log。SN會定期從NN那下載edit log，然後合併，再傳回去給NN覆蓋，自己也保留一份備份，這樣就可以減少NN的負擔。

(3)Backup NameNode

功能就和NN完全一樣了，另外他也進行SN的工作，一來可以替NN建立還原點，二來當NN掛掉的時候BN可以馬上起來接替NN進行動作，角色轉為NN，而原來的NN重新啟動後就會轉成BN。

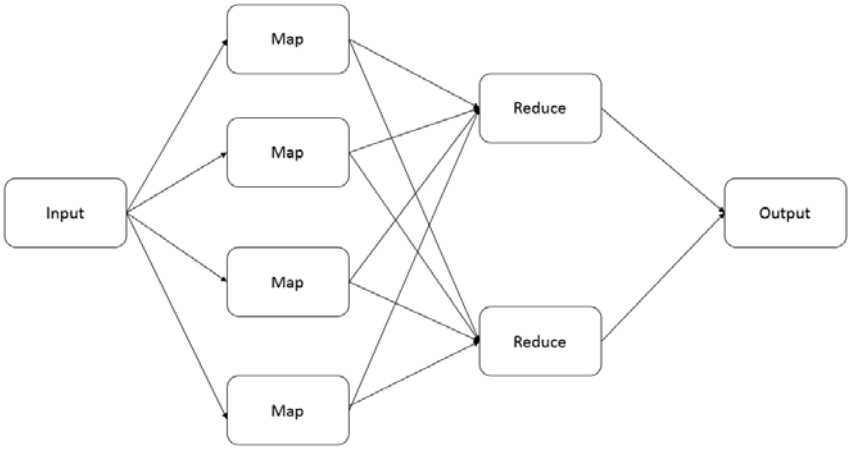
(4)Data Node

主要就是資料儲存的節點，儲存方式以Block為單位。

2.MapReduce

為一種解決問題的程式開發模式，開發人員需要先分析待處理問題的解決流程，找出資料可以平行處理的部分，也就是那些能夠被切成小段分開來處理的資料，再將這些能夠採用平行處理的需求寫成Map程式。

然後就可以使用大量伺服器來執行Map程式，並將待處理的龐大資料切割成很多的小份資料，由每臺伺服器分別執行Map程式來處理分配到的那一小段資料，接著再將每一個Map程式分析出來的結果，透過Reduce程式進行合併，最後則彙整出完整的結果。



三、勒索軟體（Ransomware）是惡意軟體的一種，卻造成組織非常大的威脅。

（一）請說明勒索軟體之惡意行為，並舉例說明遭受勒索軟體攻擊之原因。（10分）

（二）請說明勒索軟體運用的密碼學技術。（10分）

（三）請說明如何預防此類攻擊。（10分）

試題評析	本題第一與第三小題屬於基本概念，只要對於攻擊方式有基本認知，便可以輕鬆應付勒索軟體的基本定義與防範，對同學不構成威脅。唯獨第二題看似較為困難，但是其背後的運作理論和加解密的基本概念相同，只要套用非對稱式加密的概念，此題也是可以迎刃而解。
考點命中	第二小題：《高點‧高上資訊管理與資通安全講義》第三回，蕭老師編撰，頁58-59。

答：

(一)是一種特殊的惡意軟體，被歸類於「阻斷存取式攻擊」（denial-of-access attack）。其攻擊方式大概可以分為「僅是單純地將受害者的電腦鎖起來」與「系統性地加密受害者硬碟上的檔案」，這個攻擊方式不會立即讓使用者感受到，惡意軟體會在背地裡悄然運作，直到系統或資料鎖定機制部署完成。部署完成後，才會告訴使用者資料遭上鎖並藉此勒索贖金以取回對電腦的控制權，或是取回受害者根本無從自行取得的解密金鑰以便解密檔案。

其通常透過木馬病毒的形式傳播，例如透過假冒成普通的電子郵件等社會工程學方法欺騙受害者點擊連結下載，主要會將自身掩蓋為看似無害的檔案。或是透過路過式攻擊，也就是使用者瀏覽網頁或惡意廣告就會中毒。

(二)其大多採用採高安全性非對稱加密，即攻擊者的電腦會產生1組加密與解密金鑰，透過網路將加密金鑰傳送至受害者的電腦，並將檔案加密，由於受害者的電腦中並沒有解密金鑰，所以無法將檔案解密，因此將無法讀取被影響的檔案。由於解密金鑰不曾於網路上傳輸，除非受害者反過來主動入侵攻擊者的電腦，否則無法取得解密金鑰，再加上CryptoLocker採用的是金鑰長度為1024bit甚至是2048bit的RSA加密演算法，在實作上仍無法直接破解密碼、救回檔案。

(三)預防方式可為：

1.有感染跡象時立即斷網、關機並暫時停止該帳號的網路存取登入權限
2.使用防毒軟體
3.定期備份重要檔案
4.定期更新軟體
5.避免點選不明連結
6.組織提供不間斷的員工警覺性訓練
7.控管共用資料夾的存取權限，例如：不提供寫入權限

四、安全資訊與事件管理（Security Information and Event Management，簡稱SIEM）乃結合安全資訊管理（Security Information Management，簡稱SIM）與資安事件管理（Security Event Management，SEM）之整合系統，提供管理者整合組織企業之資訊安全管理所需資訊與管理功能。

（一）請說明安全資訊與事件管理系統（SIEM）之基本目標與功能。（10分）

（二）請說明組織採用SIEM系統時，應考量那些成本？（10分）

試題評析	在建置資安防護中心時有許多不同的標準，其中又以SIEM最為常見與有效，只要掌握基本概念，此題即可輕鬆回答。

答：

(一)SIEM可提供從多種資料來源中即時收集或從歷史資安事件分析而產生的威脅偵測及資安事故應變。同時也提供合適的報表以及歷史資安事故的分析。SIEM的核心能力就是從各種不同的來源收集、關聯、以及分析資安事件。良好的SIEM可以讓資安防護中心事半功倍，故在建制資安防護中心時，勢必要將SIEM列入評估的考量。其對於安全事件、使用的記錄都有使用AES加密保存，並且透過系統的通報可以即時通知使用者安全事件的發生。因此對於實踐個資法的安全措施，十分實用。

(二)必須考慮的成本大概可分為：

1.儲存成本

DB/ Storage/ NAS等儲存設備的成本。其中又再包含「確認SIEM的內建儲存庫型別以及儲存空間」與「確認其與外部儲存的相容性，以防碰到龐大儲存需求的情況」。

2.許可標準

此項為決定和影響SIEM總成本的主要關鍵，取決於現成裝置需要整合的數量或需要合併的資產數量、使用者數量或其他標準。

3.專業服務

包含開發解析機器、SIEM管理和事件處理培訓專案、SIEM裝置高優先順序問題的相關解決方案與SIEM的整合方案。

4.其他成本

(1)硬體解決方案：包含所需裝置的數量和型別與所需的支援性軟體。

(2)軟體解決方案：符合SIEM的最低系統軟體要求。