Assignment11

Part1

1. **Bengio, Yoshua, et al.**

   "Learning long-term dependencies with gradient descent is difficult."

   IEEE Transactions on Neural Networks (1994).

   🔗 https://doi.org/10.1109/72.279181

2.

   Rumelhart, David E., Hinton, Geoffrey E., and Williams, Ronald J.

   "Learning representations by back-propagating errors."

   Nature, 1986.

   🔗 https://doi.org/10.1038/323533a0

3. Yes. There are many existing results showing that the approximation lemmas for tanh networks extend to other activation functions such as ReLU and sigmoid.

   Hornik, Kurt.

   "Approximation capabilities of multilayer feedforward networks."

   Neural Networks, 1991.

   🔗 https://doi.org/10.1016/0893-6080(91)90009-T

4. Yes. Handling missing values is a well-studied topic in meteorology.

   Kalnay, Eugenia.

   "Atmospheric Modeling, Data Assimilation and Predictability."

   Cambridge University Press, 2003.

   🔗 https://doi.org/10.1017/CBO9780511802270

5. The $\log\|\Sigma\|$ term appears in the log-likelihood of a multivariate normal distribution due to the normalization factor of the density function.

   Bishop, Christopher M.
   "Pattern Recognition and Machine Learning."
   Springer, 2006.
   (第 2 章：Multivariate Gaussian distributions；第 10 章：MLE)

🔗 https://link.springer.com/book/10.1007/978-0-387-45528-0

6. Yes. The limitations of GDA under non-Gaussian data distributions are well-studied in both statistics and machine learning.

   If the true class-conditional distributions are skewed, heavy-tailed, or multimodal, the Gaussian assumption becomes severely mis-specified.

   This leads to biased parameter estimates, incorrect covariance structure, and suboptimal decision boundaries.

   Hastie, Tibshirani, Friedman

   "The Elements of Statistical Learning."

   Springer, 2009.

7. Yes. There is established research comparing standard score matching with sliced score matching.

   Mathematically, standard score matching involves second-order derivatives (Hessians), making it computationally expensive in high dimensions.

   Sliced score matching replaces full gradients with random one-dimensional projections, avoiding Hessians and greatly reducing computational cost.

   Hyvärinen, Aapo.

   "Score matching: A new method for estimating parameters of unnormalized statistical models."

   ICML, 2005.

   🔗 https://proceedings.mlr.press/r2/hyv02a/hyv02a.pdf

8. Yes, the difference between the Itô integral and the standard Riemann integral is well studied.

   Brownian motion has infinite variation and is nowhere differentiable, so the Riemann integral cannot be used.

   Instead, the Itô integral is defined via mean-square limits, not classical Riemann sums.

   Øksendal, Bernt.
   "Stochastic Differential Equations: An Introduction with Applications."
   Springer, 2003.

   🔗 https://doi.org/10.1007/978-3-642-14394-6

9. Yes, this topic is well studied in labor economics, AI policy, and education research.

   Studies consistently show that the first jobs to disappear are routine, repetitive, or easily codifiable tasks.

   AI struggles to replace roles requiring creativity, empathy, persuasion, or complex social interactions.

   Frey, Carl Benedikt & Osborne, Michael.

   "The future of employment: How susceptible are jobs to computerisation?"

   Oxford University, 2013.

   🔗 https://www.oxfordmartin.ox.ac.uk/publications/the-future-of-employment/

Part2

在最終的 Final Project 願景中，我期望未來 20 年的人工智慧能真正理解並生成完整的互動式遊戲世界，其中包含動態地形建構、角色行為決策、敘事推進、藝術風格一致的視覺內容，以及對玩家行為的即時回應。具體而言，AI 應能接收玩家輸入的自然語言指令，並以此即時調整遊戲場景、任務與角色互動，使遊戲世界具備高度開放性與可延展性。如此的願景代表著遊戲製作流程將被根本改變：內容生成將從「人類大量手工製作」轉變為「AI 理解玩家意圖後自動生成世界」，這是一個需要跨領域能力（語言理解、視覺生成、規劃與控制）的長期研究方向。

然而，若要讓 AI 逐步邁向上述願景，必須從一個目前可行且技術上「可完全實作」的簡化模型開始。基於此，我提出以下 Toy Model，作為通往完整 AI 遊戲生成系統的起點。

Toy Model：AI 從一段「世界狀態」中生成一個合理的遊戲事件

Input（世界狀態）：

Location: forest

Time: night

Player status: injured

NPC nearby: hunter

Weather: foggy

Output（**AI** 生成事件）：

A hunter notices the injured player and offers herbal medicine.