UE23CS352A: Machine Learning
Hackathon Report – Hangman

Name: Tanushri Mohan Chougale
SRN: PES2UG24CS826
Section: F

Key Observations

The most challenging part was correctly implementing the Hidden Markov Model (HMM) and ensuring smooth integration with the reinforcement learning environment. Understanding how the model learns letter transition probabilities and tuning hyperparameters such as smoothing factors were key learning points. A major insight was how the RL agent gradually improved its guessing ability as the HMM captured stronger letter-sequence dependencies from the training corpus.

Strategies

The HMM was designed using bigram probabilities, where each letter depends on the preceding one. The transition and initial probabilities were computed from the corpus text file with Laplace smoothing to handle unseen letter pairs. The reinforcement learning component used a Q-learning approach, defining states as the partially guessed word patterns and actions as the next possible letter guesses. Rewards were structured as +1 for correct guesses, -1 for wrong guesses, and -0.5 for repeated guesses to encourage exploration and penalize redundancy.

Exploration

An ε-greedy exploration strategy was implemented to balance exploration and exploitation. Initially, ε was higher (e.g., 0.3) to allow broader exploration of letter predictions, which helped the agent learn transition probabilities effectively. Gradually, ε was reduced, allowing the agent to rely on the learned HMM-based Q-values. This improved both success rate and average score stability across episodes.

Future Improvements

In future iterations, the model could be enhanced using trigram-based HMMs or even neural architectures like Recurrent Neural Networks (RNNs) or LSTMs for capturing long-term dependencies. Additionally, integrating a Deep Q-Network (DQN) could improve learning efficiency in complex environments. More detailed evaluation metrics such as precision, recall, and letter entropy could also provide deeper insight into agent performance. Further tuning of reward structures might help balance between learning efficiency and exploration behavior.