1.Write the expression for 2nd order Markov Chain. If we have large corpus available then what can be a good advantage of 4 order markov chain over second order Markov chain?   [.5+.5]

2. Mention the limitation of nGram model . What is the solution to handle that limitation.   [.5+.5]

3. Give the intuition behind the evaluation technique called perplexity? why it is called bad approximation?   [1+1]

4.There are some numbers with repetition, and total numbers are 10. one number comes 2 times and another number comes 3 time and remaining number comes single time.
Calculate the perplexity of the system.   [2]

5. **A model is trained based on following corpus** **[2]**

Dr. Saikat Singh is very popular in rural areas of Delhi due to low fees. Dr. charging such low fees are a big relief for the common people. Prof. Saikat is also very popular among students for his expertise and contribution in Math. His wife is also a Prof. and his son also wants to be a Prof. in Nuclear Physics.

Question: Predict the most probable profession of **Saikat Singh** using a simple Interpolation technique. (Show the overall calculation steps) $\lambda1$= 0.5 ,$\lambda2$=0.3, $\lambda3$=0.2 , total words =60

6. From a pack of cards, three cards are drawn at random. Find the probability that each card is from different suit. **[2]**