INSTRUCTIONS:

- This question paper has **FOUR** questions.
- Answer any **THREE** questions.

1. (a) Using examples, explain the differences between voiced and unvoiced sounds? (2 marks)
   (b) The English word 'see' contains both voiced and unvoiced sounds.
      i. Identify these voiced and unvoiced sounds. (2 marks)
      ii. Give the expression used to compute the zero crossing rate of a signal. (2 marks)
      iii. Sketch a plausible plot of the zero-crossing rate of the recording of the word 'see' and explain your sketch. (4 marks)
   (c) Describe the speech chain? (2 marks)
   (d) Describe the main parts of the speech production apparatus. Include a diagram. (4 marks)
   (e) Using an appropriate block diagram, explain the operation of an adaptive differential PCM system. (4 marks)

2. (a) The sampling rate of a given speech recording is 44.1KHz. You are required to perform short time analysis of the signal using a window size of 40ms and an overlap between windows of 75%.
      i. How many samples of the speech signal are contained in a window? (2 marks)
      ii. What is the appropriate size of the Fast Fourier Transform? (2 marks)
      iii. If the signal is 10 seconds long, how many frames will be generated? (3 marks)
      iv. Write pseudo-code that explains the generation of a spectrogram of the signal (3 marks)
   (b) Give the expression used to compute the short time energy of a signal. (2 marks)
   (c) Using the short time energy of a speech signal as a feature, design a voice activity detection system using Gaussian Mixture Models. Comment on (8 marks)
      i. The suitability of GMMs
      ii. The parameters of the GMM
      iii. How to estimate the parameters of the model

3. (a) Describe the source filter model of speech using relevant equations and diagrams? (4 marks)
   (b) Derive the Yule-Walker equations for estimation of linear prediction coefficients. (8 marks)
   (c) Describe the Levinson-Durbin Algorithm for solution of the Yule-Walker equations (8 marks)

4. (a) You are to design a speech recognition system to recognise two words 'yes' and 'no' based on Mel - frequency cepstral coefficients (MFCCs) and using a logistic regression classifier (8 marks)
      i. Describe the process of extraction of MFCCs from the speech signal.
      ii. Describe the operation of the logistic regression classifier.

iii. Describe the loss function of the logistic regression classifier.

iv. Describe the training procedure of the logistic regression classifier.

(b) Let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ be a sequence of random variables.

i. Show that

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = p(\mathbf{x}_1) \prod_{i=2}^{N} p(\mathbf{x}_i | \mathbf{x}_{i-1}, \ldots, \mathbf{x}_1)$$

(2 marks)

ii. If the sequence forms a first order Markov chain, write the expression for $p(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ (3 marks)

(c) Give the mathematical description of a Hidden Markov model (HMM) clearly stating the parameters used to describe it. (3 marks)

(d) Describe the operation of a code excited linear prediction system. (4 marks)