Total No. of Questions :12] [Total No. of Pages : 3

P1570

[3764]-434
B.E. (I.T.)
BIO - INFORMATICS
(Elective - I)

Time :3 Hours] *[Max. Marks : 100*

Instructions to the candidates:

1) *Answer any three questions from each section.*
2) *Answers to the two sections should be written in separate books.*
3) *Neat diagrams must be drawn wherever necessary.*
4) *Assume suitable data, if necessary.*

SECTION - I

Q1) a) Explain the Central Dogma of Molecular Biology. [8]

b) Explain the Gene Mapping Process in detail. [8]

OR

Q2) a) Explain the odds-likelihood form of Bayes' Theorem and explain any two limitations of Bayes' Theorem? [8]

b) The probability of a patient having a particular genetic disease is 0.5. calculate the pretest odds? If the Likelihood ratio is given as 2.75, calculate the post-test odds? Find the probability of the patient suffering from the genetic disease? [8]

Q3) a) Explain Microarray Spotting Process Flow? [8]

b) What is Clustering? Explain Hierarchical Clustering. Explain K-means clustering? [8]

OR

Q4) a) For the given fluorescence data as x[n] in the table below, calculate mean, standard deviation and variance? [8]

n	1	2	3	4	5	6	7
x [n]	3.2	1.6	4.4	3.3	2.7	1.7	6.8

P.T.O.

b) Explain the concept of True Positives, True Negatives, False Positives and False Negatives. **[4]**

c) Explain the concept of Sensitivity and Specificity along with the formulae. **[2]**

d) Explain the concept of Receiver Operating Characteristics. **[2]**

Q5) a) Discuss following data mining methods in detail : **[10]**

i) Classification.
ii) Regression.
iii) Link Analysis.
iv) Deviation Detection.
v) Segmentation.

b) Explain and differentiate Pattern Recognition from Pattern Discovery. **[8]**

OR

Q6) a) Explain following terms : **[16]**

i) Genetic Programing.
ii) Neural Networks.
iii) Hidden Markov Models.
iv) Decision Trees.

b) Explain Inductive Logic Programming? **[2]**

SECTION - II

Q7) a) For the given two nucleotide sequences calculate the alignment score. Use gap penalty of (- 0.5) per gap. Assuming opening cost and extension cost of (- 0.5) each calculate the penalty gap, using this also calculate expanded gap penalty. **[12]**

Sequence 1 : ATTCGGCATTCAGAGCTAGA
Sequence 2 : ATTCGACATT----GCTAGTGGTA

b) Given A =[2 3 8 4 1] and B=[9 11 1 0 2 4 5 6 7 3 2], calculate **[6]**
Max Value = $f(A_1, B_i)$, where, i=1, 2,....., 11.

OR

Q8) Explain following in brief : **[18]**

- a) DNA Probes.
- b) Genetic Markers.
- c) Applications of genetic engineering.
- d) Polymerase chain reaction.

Q9) BLAST and FASTA are two widely used tools for sequence alignment. Explain the similarities and differences in their approach. **[16]**

OR

Q10) a) What is an E-value? You do a databank search using FASTA with an amino acid sequence as a query. The only reported match has an E-value of 10. What does this mean for the similarity of the two sequences? **[8]**

b) Explain PSI-BLAST with the schematic. List any four applications for which PSI-BLAST can be used. **[8]**

Q11) Explain the concept behind Collaboration and Communication. Explain clearly the hierarchy. **[16]**

OR

Q12) a) What is the significance of Biotechnology? **[8]**

b) Discuss various factors that are responsible for degradation of ecosystem. **[8]**

# # # #