BIO543: Big Data Mining Healthcare

(6th February 2025, Quiz1)

Maximum Marks: 20   **Duration: 20 Minutes**

Name:_____   **Admission No:** _____

Instructions: Attempt all questions, each question carry one mark. Please tick only one best option.

1. Which of the following contain genetic information in humans?

    a) Proteins   **c) Chromosomes**

    b) Lipids   d) Carbohydrates

2. What is the "start" codon in translation?

    a) UGA   c) UAA

    b) AUG   d) UAG

3. Proteins are made of __ types of natural amino acids.

    a) 4   c) 16

    b) 20   d) 64

4. Which of the following is used to cut DNA at specific sites?

    a) DNA polymerase   c) RNA polymerase

    b) Restriction enzyme   d) Telomerase

5. How many PCR cycles are required to produce 32 duplexes starting from 2 duplexes?

    a) 3   c) 5

    b) 4   d) 6

6. Which of the following atoms is NOT a part of proteins?

    a) Carbon   c) Hydrogen

    b) Oxygen   **d) Phosphorus**

7. Which of the following contain Sulfur?

    a) Protein   c) DNA

    b) RNA   d) Lipid

8. Which Docker command is used to list all running containers?

    a) docker images   c) docker list

    b) docker ps   d) docker run

9. Which of the following is not a Unix shell?

    a) bash   c) ksh

    b) tcsh   **d) ush**

10. Which of the following is NOT a valid Python data type?

    a) List   c) Dictionary

    b) Tuple   **d) Array**

11. Which pandas function is used to read a CSV file?

    a) pd.open_csv()   **c) pd.read_csv()**

    b) pd.load_csv()   d) pd.import_csv()

12. Which of the following is not a function of NumPy?

    a) multiply()   **c) upper()**

    b) binomial()   d) randint()

13. Which Python library is commonly used for data visualization?
    - a) Scikit-learn
    - **b) Seaborn**
    - c) NumPy
    - d) Pandas

14. Which of the following is NOT a major repository for healthcare?
    - a) IEDB
    - b) PRRDB
    - c) GEO
    - **d) SCOPUS**

15. Which repository is used for medical imaging data?
    - **a) TCIA**
    - b) PubMed
    - c) UniProt
    - d) PDB

16. The FDA and PubChem databases are mainly used for maintaining information on:
    - a) Vaccine candidates
    - b) Genetic sequences
    - **c) Chemical drugs**
    - d) Microbial genomes

17. COVAXIN is a _____ based vaccine
    - a) Recombinant protein
    - b) DNA
    - c) Epitope
    - **d) Whole organism**

18. Which database maintains human GAIT data?
    - **a) PhysioNet**
    - b) ThpDB
    - c) TCGA
    - d) OMIM

19. Which of the following web server is used in computer-aided vaccine design?
    - **a) IL6pred**
    - b) b3pred
    - c) 1mg
    - d) CancerEnd

20. Which of the following database contain drug like molecules?
    - a) B3pdb
    - b) CancerCSP
    - **c) ChEMBL**
    - d) IEDB