CANDIDATE NUMBER………………………………………………….

“Investing in Africa’s Future”

COLLEGE OF HEALTH, AGRICULTURE AND NATURAL SCIENCES

DEPARTMENT OF BIOMEDICAL AND LABORATORY SCIENCES

BACHELOR OF MEDICAL LABORATORY SCIENCES HONOURS DEGREE

MLS3103: BIOINFORMATICS AND AI

END OF SEMESTER FINAL EXAMINATIONS

NOVEMBER 2025

LECTURER: Mr Z CHIWODZA

DURATION: 3 HOURS

INSTRUCTIONS

1. Write your candidate number on the space provided on top of each page
2. Answer **all** questions in sections A on the question paper.
3. Answer **all** questions in section B on separate answer sheets provided.
4. Answer any **3** questions in section C on separate answer sheets provided
5. The mark allocation for each question is indicated at the end of the question
6. Credit will be given for logical, systematic and neat presentations in sections B and C

CANDIDATE NUMBER……………………………………………….

SECTION A: MULTIPLE CHOICE [40 MARKS]

- **Answer all questions by encircling the correct response T for TRUE or F for FALSE for each statement in all the questions**
- **Each correct response is allocated half a mark**

---

1. **Regarding biological databases:**
   - a) T F FASTA format includes both sequence and quality scores.
   - b) T F GenBank is a primary database containing annotated DNA sequences.
   - c) T F PDB is mainly used for protein 3D structure data.
   - d) T F EMBL is a European-based nucleotide sequence database.

2. **Concerning on sequence alignment:**
   - a) T F BLAST can align both nucleotide and protein sequences.
   - b) T F A lower E-value indicates a more significant match.
   - c) T F Gaps in alignments may represent insertions or deletions.
   - d) T F A high bit score indicates weak similarity.

3. **Regarding Next Generation Sequencing (NGS) technologies:**
   - a) T F Illumina sequencing produces long reads (>10 kb).
   - b) T F Oxford Nanopore sequencing detects changes in ionic current.
   - c) T F PacBio uses single-molecule real-time (SMRT) sequencing.
   - d) T F Coverage depth affects the confidence in mutation calling.

4. **Artificial Intelligence in medical diagnostics:**
   - a) T F Machine learning can be used for automated image analysis in histology.
   - b) T F AI models can eliminate the need for medical laboratory scientists.
   - c) T F Deep learning is useful in pattern recognition tasks such as radiology.
   - d) T F AI systems are always free from bias if trained on large datasets.

5. **Applications of bioinformatics in infectious disease:**
   - a) T F Whole-genome sequencing helps track antimicrobial resistance.
   - b) T F Phylogenetics can be applied to study viral outbreaks.
   - c) T F Consensus sequences are created from raw reads without alignment.
   - d) T F Databases like GISAID store influenza and coronavirus sequences.

6. **With regards to Health informatics:**
   - a) T F Laboratory Information Systems (LIS) are used for patient result reporting.
   - b) T F Interoperability allows different hospital systems to share data.
   - c) T F Data privacy is less important in AI systems.
   - d) T F HL7 and FHIR are standards for health data exchange.

7. **Concerning Phred quality scores:**
   - a) T F A Phred score of 20 means 1% error probability.
   - b) T F Quality scores are irrelevant once reads are aligned.

c) T F High-quality reads reduce false variant calls.
d) T F FASTQ format includes both sequence and quality scores.

8. **Regarding ethical issues in AI:**
a) T F AI systems can raise concerns about patient consent and data use.
b) T F Algorithmic bias may worsen health disparities.
c) T F AI models must always be transparent (explainable AI).
d) T F Patients should never be informed if AI contributes to diagnosis.

9. **A reference lab receives a viral sequence from a patient with suspected HIV. They run a BLAST search to compare it against known sequences.**
a) T F BLAST can be used to confirm whether a patient's viral sequence belongs to HIV-1 or HIV-2.
b) T F A BLAST search can identify antimicrobial resistance genes in bacterial isolates.
c) T F BLAST always gives only one correct match.
d) T F A higher bit score indicates a more reliable alignment.

10. **A sequencing lab is validating its pipeline for variant calling in clinical samples.**
a) T F Coverage depth of >30× is typically considered reliable for variant calling.
b) T F A FASTQ file includes sequence quality scores.
c) T F The consensus FASTA is identical to the reference genome.
d) T F Misalignments can result from low-quality reads or repetitive sequences.

11. **Regarding Genomics, Bioinformatics and AI in outbreak management:**
a) T F Sequencing data can be shared globally for epidemic preparedness.
b) T F AI can predict future outbreaks by analyzing epidemiological data.
c) T F Genomic surveillance has no role in public health decision-making.
d) T F Real-time dashboards integrating AI and bioinformatics can aid health ministries.

12. **A research team builds a predictive AI model trained on genomic datasets.**
a) T F AI models can also be trained on bioinformatics datasets to predict resistance patterns.
b) T F Bioinformatics is limited to only bacterial genome analysis.
c) T F AI can assist in predicting protein structure from sequences.
d) T F Integration of AI and bioinformatics supports precision medicine.

13. **During an Ebola outbreak, genomic epidemiology was used to trace chains of transmission.**
a) T F Phylogenetic trees can reveal transmission chains during an outbreak.
b) T F Ebola outbreaks have been traced using genomic sequencing.
c) T F Phylogenetic analysis can only be done with protein sequences.
d) T F Multiple sequence alignment is required before phylogenetic tree building.

14. **A regional hospital adopts AI-based digital pathology for routine lab tests.**
a) T F AI can count blood cells in digital smears.
b) T F Algorithms can flag abnormal haemoglobin patterns.
c) T F AI can assist in interpreting liver enzyme trends in chemical pathology.
d) T F Machine learning can predict patient prognosis from laboratory data.

15. **A genomic research team plans to deposit anonymized patient sequences into GenBank.**
    a) T F Patient consent is not needed when depositing clinical sequences in public databases.
    b) T F Data anonymization is important when sharing patient genomic data.
    c) T F Algorithmic bias can affect how AI performs across different populations.
    d) T F AI regulation in healthcare is still developing globally.

16. **An AI tool is piloted in a malaria-endemic area to assist with rapid screening of blood smears.**
    a) T F AI can be trained to detect malaria parasites on blood smears.
    b) T F Deep learning algorithms are commonly used in medical image recognition.
    c) T F AI models cannot be applied to haematology slides.
    d) T F Machine learning can assist in interpreting Full Blood Count results.

17. **Which of the following are examples of Secondary Databases in bioinformatics**
    a) T F BLAST
    b) T F UniProt
    c) T F Stanford's HIVdb
    d) T F ENSMBL

18. **A physician uses a CDSS that alerts when a prescribed antibiotic may interact with another medication.**
    a) T F CDSS can help reduce prescription errors.
    b) T F CDSS rely only on manual data entry and not on EHR integration.
    c) T F AI-enhanced CDSS can provide personalized treatment recommendations.
    d) T F Over-alerting is a disadvantage of CDSS.

19. **A rural clinic adopts telemedicine services with mobile health (mHealth) applications.**
    a) T F Telemedicine allows remote consultations between patients and doctors.
    b) T F mHealth apps can track chronic disease parameters like blood glucose.
    c) T F Telemedicine cannot be integrated with laboratory information systems.
    d) T F Limited internet connectivity can hinder telemedicine effectiveness.

20. **A national health ministry deploys an AI-driven dashboard integrating hospital, laboratory, and surveillance data**.
    a) T F Big data analytics can identify disease trends at a population level.
    b) T F AI can predict hospital bed occupancy during outbreaks.
    c) T F Health informatics systems cannot integrate genomic data with clinical data.
    d) T F Data quality and completeness are critical for reliable AI predictions.

CANDIDATE NUMBER……………………………………………….

SECTION B: [20 MARKS]

Answer all questions on separate answer sheets provided

1. List five (5) types of biological databases used in bioinformatics and give one use for each. **[5 marks]**
2. State five (5) advantages of using AI in medical laboratory diagnosis. **[5 marks]**
3. Briefly describe five (5) steps in a standard NGS workflow from sample to consensus sequence. **[5 marks]**
4. List five (5) legal challenges associated with using AI in healthcare. **[5 marks]**

SECTION C: [75 marks]

Answer any 3 questions from this section on separate answer sheets provided

1. a) Discuss the role of NGS and bioinformatics in detecting emerging infectious diseases **[15 marks]**
   b) Explain how sequence alignment and BLAST help in pathogen identification. **[10 marks]**
2. Evaluate the benefits and limitations of using AI in laboratory results interpretation compared to human experts **[25 marks]**
3. a) Outline the key functions of a Laboratory Information System (LIS). **[10 marks]**
   b) Discuss the importance of interoperability and standards in integrating bioinformatics and AI outputs into healthcare systems. **[15 marks]**
4. a) Compare Illumina, Oxford Nanopore, and PacBio sequencing technologies. **[15 marks]**
   b) Discuss the significance of coverage depth and Phred quality scores in ensuring accurate variant calling. **[10 marks]**
5. Discuss major ethical challenges and solutions in applying AI in healthcare **[25 marks]**