



PRÁCTICA: ANÁLISIS DE DATOS GENÉTICOS MASIVOS MEDIANTE PARALELISMO Y REDES NEURONALES

INTRODUCCIÓN

En la era de la información, la bioinformática se ha convertido en un campo esencial para procesar y analizar grandes volúmenes de datos genéticos. La secuenciación masiva del ADN genera una cantidad inmensa de datos que requieren tecnologías avanzadas para su procesamiento eficiente. En este contexto, la computación paralela y las redes neuronales juegan un papel fundamental al permitir analizar millones de fragmentos de ADN simultáneamente, reduciendo significativamente el tiempo necesario para obtener resultados.

Esta práctica se centra en la aplicación de estrategias de paralelismo y algoritmos basados en redes neuronales para resolver problemas complejos relacionados con la bioinformática y explorar cómo este enfoque puede ser adaptado a otros campos de la ciencia y la tecnología.

OBJETIVOS

- Comprender cómo la computación paralela mejora el rendimiento en el análisis de datos genéticos masivos.
- Explorar el uso de redes neuronales para la identificación de patrones y ensamblaje de secuencias de ADN.
- Reflexionar sobre las implicaciones técnicas y éticas del uso de estas tecnologías en la bioinformática y otros campos.

DESARROLLO DE LA PRÁCTICA

Pregunta de investigación

Investiga y compara el tiempo estimado que llevaría secuenciar el genoma humano hoy en día utilizando tecnologías de computación paralela y redes neuronales, en comparación con el método original utilizado en el Proyecto Genoma Humano.

Guía para la investigación:

1. **Búsqueda de información:**
 - Investiga cuánto tiempo tomó completar la secuenciación del Proyecto Genoma Humano utilizando tecnologías secuenciales.
 - Busca estimaciones actuales sobre el tiempo de secuenciación del genoma utilizando computación paralela y redes neuronales modernas.
2. **Factores clave en la reducción del tiempo:**



- Identifica y explica qué tecnologías específicas han permitido reducir el tiempo de secuenciación (por ejemplo, arquitecturas paralelas como GPUs y FPGAs).
- Describe cómo estas tecnologías dividen el trabajo de secuenciación y mejoran su eficiencia.

3. **Beneficios de las arquitecturas paralelas:**

- Investiga cómo las GPUs y FPGAs han sido utilizadas en la práctica para acelerar el procesamiento de datos genéticos.
- Explica cómo las redes neuronales pueden detectar errores y patrones para mejorar la precisión del ensamblaje de ADN.

ESTRUCTURA RECOMENDADA PARA EL INFORME:

1. **Introducción:** Presenta el objetivo de la investigación y la relevancia de la computación paralela.
2. **Desarrollo:** Expone los resultados de la investigación sobre el tiempo de secuenciación y los factores que permiten su reducción.
3. **Conclusiones:** Reflexiona sobre los avances tecnológicos y su impacto en la ciencia y la tecnología.

EVALUACIÓN

- **60%** Informe individual entregado, evaluado en función de su claridad, reflexión crítica y comprensión de los conceptos.
- **20%** Precisión en la comparación de datos y análisis de los factores clave que explican la mejora en el tiempo de secuenciación.
- **20%** Propuestas innovadoras o enfoques creativos en la forma de plantear o explicar la información obtenida.

NOTA: Los datos necesarios para responder estas preguntas deben ser investigados por los estudiantes utilizando fuentes confiables de internet, artículos científicos o bases de datos académicas.