

Tercer examen parcial

MSc. Vicente Machaca Arceda

12 de julio de 2021

DOCENTE	CARRERA	CURSO
MSc. Vicente Machaca Arceda	Escuela Profesional de Ciencia de la Computación	Bioinformática

1. Competencias del curso

- Aplica las bases matemáticas y la teoría de la informática en algoritmos de Bioinformática.
- Analiza, diseña y propone soluciones frente a problemas bioinformáticos.
- Sabe cómo utilizar y conoce las bases computacionales de herramientas modernas de secuenciamiento, alineamiento, árboles filogenéticos y mapeo de genomas.

2. Competencias del trabajo

- Implementar un paper de investigación en Bioinformática.

3. Equipos y materiales

- Editor de texto Latex

4. Entregables

- Se debe elaborar un informe en Latex donde se desarrolle el trabajo solicitado.
- El informe se desarrollará en grupos de 4.
- El informe deberá estar correctamente citado utilizando las normas APA o IEEE.

5. Descripción del trabajo

Implementar un artículo académico de Bioinformática. Estos trabajos han sido escogidos, de manera tal que tengan una complejidad de nivel medio, se cuente con todos los recursos (bases de datos, librerías) y pueden ser llevados a cabo en una PC básica. Los trabajos a implementar son:

- *DNA sequence similarity analysis using image texture analysis based on first-order statistics* [1].
- *Use of image texture analysis to find DNA sequence similarities* [2].
- A New Local Search Algorithm for the DNA Fragment Assembly Problem [3].
- *Toward an Alignment-Free Method for Feature Extraction and Accurate Classification of Viral Sequences* [4].
- *SplitThreader: Exploration and analysis of rearrangements in cancer genomes* [5].
- *Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13)* [6].
- *OmicsNet: a web-based tool for creation and visual analysis of biological networks in 3D space* [7].
- *Ribbon: intuitive visualization for complex genomic variation* [8].

6. Rúbricas

Rúbrica	Cumple	Cumple con obs.	No cumple
Informe: El informe debe estar en Latex, con un formato limpio y fácil de leer. Además, debe contener: descripción de los algoritmos utilizados, código fuente y resultados obtenidos.	3	1.5	0
Implementación: Los alumnos han logrado implementar el algoritmo del paper.	7	3.5	0
Resultados: El grupo ha logrado replicar los resultados en un 50%. Es decir, no es necesario utilizar todas las bases de datos utilizadas en el paper.	5	2.5	0
Presentación: El alumno demuestra dominio del tema y conoce con exactitud cada parte de su código. Además, demuestra conocer la base matemática de su implementación. Se realizará preguntas a cada integrante del grupo, si un alumno no logra responder, tendrá cero en esta rúbrica.	5	2.5	0

Referencias

- [1] E. Delibaş and A. Arslan, “Dna sequence similarity analysis using image texture analysis based on first-order statistics,” *Journal of Molecular Graphics and Modelling*, p. 107603, 2020.
- [2] W. Chen, B. Liao, and W. Li, “Use of image texture analysis to find dna sequence similarities,” *Journal of theoretical biology*, vol. 455, pp. 1–6, 2018.
- [3] E. Alba and G. Luque, “A new local search algorithm for the dna fragment assembly problem,” in *European Conference on Evolutionary Computation in Combinatorial Optimization*, pp. 1–12, Springer, 2007.
- [4] D. Lebatteux, A. M. Remita, and A. B. Diallo, “Toward an alignment-free method for feature extraction and accurate classification of viral sequences,” *Journal of Computational Biology*, vol. 26, no. 6, pp. 519–535, 2019.
- [5] M. Nattestad, M. C. Alford, F. J. Sedlazeck, and M. C. Schatz, “Splitthreadder: Exploration and analysis of rearrangements in cancer genomes,” *bioRxiv*, p. 087981, 2016.
- [6] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. Nelson, A. Bridgland, *et al.*, “Protein structure prediction using multiple deep neural networks in the 13th critical assessment of protein structure prediction (caspl3),” *Proteins: Structure, Function, and Bioinformatics*, vol. 87, no. 12, pp. 1141–1148, 2019.
- [7] G. Zhou and J. Xia, “Omicsnet: a web-based tool for creation and visual analysis of biological networks in 3d space,” *Nucleic acids research*, vol. 46, no. W1, pp. W514–W522, 2018.
- [8] M. Nattestad, R. Aboukhalil, C.-S. Chin, and M. C. Schatz, “Ribbon: intuitive visualization for complex genomic variation,” *Bioinformatics*, vol. 37, no. 3, pp. 413–415, 2021.