

CURSO: CC471

Proyecto del curso

Sistema de Ayuda para el análisis de información Genética de 4 tipos de virus y sus variantes (puede ser: Influenza virus, Dengue virus, West Nile virus, MERS coronavirus, Ebolavirus, Rotavirus, Zika virus, SARS-CoV-2)

1. Los objetivos del proyecto son:

1 .Desarrollar una aplicación: -- software / plug-in/ Librería – utilizando Bio Python, y otras librerías disponibles, que sirva para los siguientes propósitos:

2. Realizar el análisis bioinformático de información genética de genes específicos y proteínas pertenecientes a especies de virus que provienen de diferentes regiones del mundo.

3. A partir de los árboles filogenéticos generados descubrir / Mostrar / Analizar las relaciones evolutivas de 10 o más las variantes de cada uno de los virus escogidos.

3. El Software **debe** incluir como mínimo:

- La posibilidad de Escoger /VIRUS/GENES/ PROTEINAS relacionados para hacer el análisis posterior.
- Alineamiento múltiple de Genes homólogos.
- Alineamiento de Proteínas.
- Generación de Árboles Filogenéticos

Es deseable que el análisis **también** incluya :

- Análisis de la expresión de genes en diferentes circunstancias y/o organismos.
- Modelamiento de la Estructura de Proteínas.

El proyecto deberá cumplir con los siguientes requerimientos obligatorios:

2. Requerimientos Obligatorios:

El proyecto se realizará por un grupo de 2 personas.

El software generado se probará en un ejemplo exhaustivo que muestre todas las características solicitadas. El software deberá generar resultados y archivos en formatos estándar. Los resultados se presentarán en el informe final.

2.1.El Gene o conjunto de genes a analizarse como ejemplo, puede ser escogido por el grupo utilizando el software desarrollado, según cualquiera de los siguientes criterios:

- Debe ser parte importante en la manifestación de alguna característica específica del conjunto de individuos de los virus escogida(s)
- Puede ser parte importante en la codificación de proteínas esenciales para una función específica del organismo.
- Debe contar con información pública disponible de Bases de Datos de experimentos reales. p.ej. GEO -Datasets Database (GEO- Gene Expression Omnibus) , NCBI, etc.

- El resultado debe incluir el análisis filogenético de las variantes escogidas de los virus y de las regiones de las que provienen.

3. Aclaratorias Importantes:

3.1. La tarea puede ser realizada utilizando e integrando recursos web o herramientas disponibles públicamente. **No debe incluir código de otros grupos.**

3.2. Para cada proyecto se debe realizar todas las características obligatorias. **El código debe estar profusamente comentado. Indicando en cada módulo o funcion, El autor, El grupo, El Curso, El periodo académico.**

3.3. El primer día de retraso será penalizado con dos (2) puntos y luego cada día de retraso adicional será penalizado con un (1) punto.

3.4 La Nota Final del Proyecto tendrá un peso de (40%) en el examen Final. Y será calculada como el promedio de los entregables 2,3,y 4.

4. Entregables:

4.1. Entregable 1. Práctica Calif N.2 Propuesta del Proyecto. Incluye cronograma y definicion de entregables .

4.2. Entregable 2. Práctica Calif N.3. Version final de la propuesta en PDF y Demostracion al 30% (Definicion de: Especies, Genes, y Proteinas) Razones por las que se escogieron

4.3. Entregable 3. Practica Calif. N.4 Demostracion al 70% (Realización y análisis de Alieneamientos de Genes y proteínas escogidos con el código Biopython)

4.4. Entregable 4. Lab. Calif. N.4 – Análisis y Generación de Árboles Filogeneticos a partir de la información generada. Demostracion al 100% incluyendo Codigo fuente, Librerias, datos de prueba, reporte final. En formato digital (DVD)

Los proyectos y documentos entregados deberan ser entregados con licencia Open Source y Creative Commons.

5. Secciones que se deben incluir en la Propuesta de Proyecto y en el Reporte Final:

5.1 **Resumen Ejecutivo.**[Lo que Ud. Esta Proponiendo Hacer, en caso de la propuesta] Una vision general de lo que ha hecho en su proyecto.

5.2. **Descripción del Proyecto. Debe incluir los alcances y el cronograma del proyecto.**

Las fechas en el cronograma deben coincidir con las fechas propuestas para evaluaciones de práctica del proyecto. En dichas ocasiones se presentarán y evaluarán los avances del proyecto.

5.3. **Algoritmos e implementación computacional** - Una descripcion de los algoritmos y herramientas que se [planean utilizar en caso de la propuesta] utilizados incluyendo pseudo código y codigo fuente.

5.5. **Resultados** – Una descripción de los resultados [esperados en el caso de la propuesta]. Un reporte integrando los resultados proporcionados por la herramienta.

5.6. **Conclusiones.** - Incluye las ventajas y desventajas del enfoque utilizado, aspectos inesperados del proyecto, trabajo futuro, etc.

5.7 **Apéndice** – Incluye información suplementaria

5.8. **Bibliografia.**