**PRUEBA ESTANDARIZADA DE “MACHINE LEARNING FOR GENOMIC”**

**Nombres y Apellidos: …………...............................................................**

**Expositor:** Student PhD & MSc. Ronald Eleazar Huarachi Olivera

**Organizador y Coordinador:** Dr. Antonio Mateo Lazarte Rivera

**LABORATORIO DE BIOTECNOLOGÍA CELULAR Y MOLECULAR AVANZADA (LAB-BIOTCEMA); ESCUELA PROFESIONAL DE BIOLOGÍA; UNIVERSIDAD NACIONAL DE SAN AGUSTÍN; AREQUIPA-PERÚ**

**INDICACIONES:**

* El estudiante deberá responder eficientemente el control, el cuál se asignará una nota de 0 a 20, donde la mínima nota aprobatoria corresponde a 10.5 y la máxima nota aprobatoria corresponde a 20.
* La nota aprobatoria de 10,5 a 20.
* Durante la búsqueda de información en inglés, se recomienda de que el participante interprete correctamente sus respuestes, no solamente traducir de inglés al español, sino que el participante deberá interpretar su respuesta de manera correcta.
* El estudiante deberá enviar el control (en Word) al correo electrónico: workshop.biol.molecular@gmail.com con el siguiente título o código **“2preworkSecMasiva\_NOMBRE\_APELLIDO”** a más tardar el lunes 24 de Febrero hasta las 20:00 hrs.

RESPONDA CLARAMENTE LAS SIGUIENTES PREGUNTAS CITANDO REFERENCIAS BIBLIOGRAFICAS (puede basarse en las diapositivas del curso). OJO: NO SERÁN EVALUADAS LAS PRUEBAS CLONADAS. RESUELVA PERSONALMENTE. SI DESEA TAMBIEN PUEDE ENVIAR CAPTURA DE PANTALLA, CON LA FINALIDAD DE SABER QUE UD. RESOLVIO LA PRUEBA ESTANDARIZADA.

**1.- Con buenas fuentes bibliográficas mencione cual es la diferencia entre aprendizaje No Supervisado y Aprendizaje supervisado. COLOQUE 1 EJEMPLO DE CADA UNO. (1 PUNTO)**

**2.- Mencione cuales son los tipos de clustering y a que se refiere cada tipo? (1 PUNTO)**

**3.- En Clustering jerarquico aglomerativo, cree su propia matriz y arme su dendrograma usando lo aprendido en práctica de “Machine Learning for Genomic” (3 PUNTOS)**



**4.- En cuanto al K-means, en base a que criterios elijo el “K”? Y como los reubico a los centroides (1 PUNTO)**

**5.- Con la lista de genes proporcionados durante la practica de MACHINE LEARNING (Análisis de aprendizaje no supervisado), el participante puede manipular los datos a su manera para obtener un CLUSTERING BICLUSTERING para obtener supuestos genes rítmicos o circadianos. Así como en el siguiente ejemplo:**

****

**Deberán mostrar la tabla manipulada con genes de ritmicidad como dato adjunto de la prueba estandarizada (3 PUNTOS)**

**6.- En base a que criterios elijo el número de componentes principales en PCA (1 PUNTO)**

**7.-Que es el Single Cell, ¿Qué tecnología uso para secuenciar?** **(2 PUNTOS)**

**8.- Cual es la diferencia entre PCA y t-SNE (cite fuentes bibliográficas indexada en SCOPUS) (1 PUNTO)**

**9.- Mencione un ejemplo sobre la sobreexpresión y subexpresiòn usando el log2(fold change) asigne los valores usted desee.(Cite fuentes bibliográficas indexadas en SCOPUS para mencionar su teoría)** **(2 PUNTOS)**

**10.- ¿Que diferencias existe entre el KNN y Linear (and Quadratic) discriminant analysis (L/QDA) para separar cluster? (1 PUNTO)**

**11.- Usando el Suppert Vector Machine como realizo la separación entre diferentes cluster(1 PUNTO)**

**12.- Siendo usted un(a) profesional de las Ciencias Biológicas o Biomédicas ya sea Biólogo, Biotecnólogo, Médico, Nutricionista ,etc. Usando MACHINE LEARNING cual sería su propuesta para formar su propia empresa en el Perú en base a la genómica y/o biología molecular. SI desea elabore esquemas (3 PUNTOS)**