

**TEST DE EVALUACIÓN PARA EL PUESTO DE PRACTICANTE DE INFORMÁTICA
BOLSA DE TRABAJO PUCP**

Febrero 2018

** La presentación de los resultados tendrá una consideración especial.*

Escoja una de las siguientes opciones :

- a) Usando Python, realice la clasificación de un dataset de convocatorias de trabajo. El objetivo es predecir si una convocatoria debe ser aceptada o no en función del título, descripción y requerimientos de la misma. Es libre de escoger las librerías y algoritmos que usted considere necesario.
- b) Usando Python realice la clasificación de un dataset basado en textos. Para esto, debe realizar una implementación básica del TF-IDF y utilice un método de clasificación bayesiano. Debe utilizar la librería Sklearn. Recuerde que el dataset escogido debe incluirse en su repositorio.

Pregunta Obligatoria :

- a) Investigue sobre Apache Cassandra, para qué sirve y cómo interactúa con Python. La respuesta se debe colocar en un archivo .txt dentro del repositorio.

Nota importante:

La implementación de los algoritmos pedidos deben estar en un repositorio de GitHub, para ello se le pide que envíe el link del repositorio. El envío de archivos por otros medios no será revisado. Serán evaluadas las buenas prácticas de programación.