

Desarrollo de Software Estadístico Descriptivo

Objetivo del examen

Por medio de este examen se evalúan los conocimientos y habilidades del alumno en el desarrollo de algoritmos de programación en lenguaje C, así como buenas prácticas de programación. Un elemento importante para este examen es la evaluación por medio de un documento.

Desarrollo

Se desea contar con un documento que muestre la solución a un problema específico de programación. Todo el desarrollo debe ser realizado en forma individual.

Problema: Dentro del proceso de Desarrollo de algoritmos de inteligencia artificial, uno de los pasos más importantes es el resumen de datos en simples medidas estadísticas. Esto, no nada mas permite conocer la estructura de la información a utilizar, si no también es paso importante para la elección de tipo de algoritmo de intelifencia a utilizar. Es por esto, que se desea generar un software capaz de generar medidas estadísticas básicas, además de generar un file con el reporte de estas.

El programa de software **TDA_Statistics** deberá de ser capaz de calcular las siguientes cantidades:

1) **Medidas de valor central:**

- a. Media. $\mu = \frac{1}{N} \sum_{n=0}^{N-1} X_n$
- b. Mediana. $med(sort(X_n))$
- c. Moda.

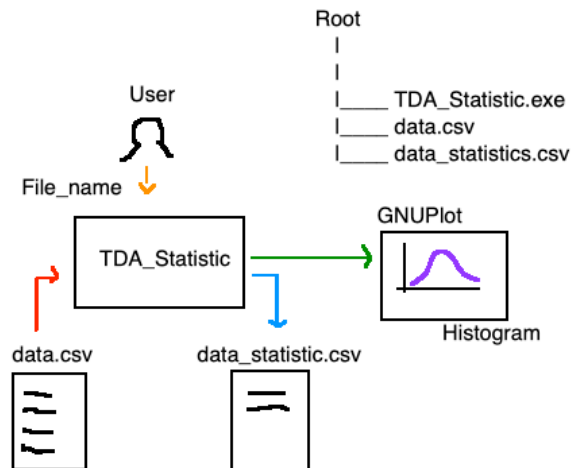
2) **Medidas de dispersion:**

- a. Desviación estandar $\sigma = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (X_n - \mu)^2}$
- b. Rango. $\min(X), \max(x)$
- c. Distancia intercuartíl. $Sort(X[95\%]) - sort(X[25\%])$

El software debe ser capaz de leer los datos desde un file de entrada de tipo **.csv**. Para poder generar el histograma, de deberá utilizar GNUPLOT.

P.S Bibliotecas de Sorting de C pueden ser utilizadas!

El diagrama de interacción de propone ser algo similar a:



El reporte generado con las medidas estadísticas deberá ser almacenado en el mismo directorio del cual los datos fueron obtenidos.

Recuerda antes de programar, diseñar y documentar tu software!! Es parte de la entrega!

Entregas:

- File con diseño, documentación y capturas de pantalla.
- Código Fuente y Header files de bibliotecas modularizadas.

El examen tendrá duración de 2 horas (durante la clase) y deberá ser entregado a través de un pull-request (5 decimas extra) de GitHub y/o a través de las entregas de Teams!

Mucha suerte!

Data_Statistic.csv

Media:

12

Moda:

4

Data.csv

12

11

23

12

0

23