UNIVERSIDAD COMPLUTENSE MADRID

Genetic Programming based on Grammars Challenge

Iñaki Hidalgo

Adaptive & Bioinspired Systems (ABSys) Research Group

J. Ignacio Hidalgo (Iñaki)

Universidad Complutense de Madrid

COMPLEJO HOSPITALARIO DE TOLEDO

Hospital Virgen de la Salud de Toledo

Hospital Universitario Príncipe de Asturias
SaludMadrid
Comunidad de Madrid

Hospital U. Príncipe de Asturias

Instituto de Matemática Multidisciplinar (UPV)

im^2

Universidad de Extremadura

UEX
UNIVERSIDAD DE EXTREMADURA

People

- Iñaki Hidalgo
- Juan Lanchares
- Oscar Garnica
- Manel Velasco
- Almudena Sánchez
- Carlos Cervigón

- Esther Maqueda
- Marta Botella

- Rafael Villanueva
- Juan Carlos Cortés

- Francisco Fernández de Vega
- Francisco Chávez

Ph. D. Students

- Alberto Gutierrez
- Jorge Alvarado
- Daniel Parra
- Javier Villanueva
- Jorge Koronis
- Laura Millán

M Sc.

- Marina de la Cruz
- Jorge Cano

3

My University

- UCM: Universidad Complutense de Madrid
  - Originally from Alcalá de Henares (Complutum for Ancient Romans)
  - Funded in 1499 by Cardinal Cisneros
- 75.000 undergraduate students
- 15.000 + graduates and others
- Two different campuses
  - Main: Moncloa
  - Somosaguas 7km
- Facultad de Informática (Computer Science)
  - Computer Architecture Department

Ciudad Universitaria

UNIVERSIDAD
COMPLUTENSE
MADRID

Projects

- Artificial intelligence on specialised hardware accelerators and embedded systems for precision personalised treatment of diabetes (PID2021-125549OB-I00)

- Artificial intelligence wearable decision support system for people with diabetes (PDC2022-133429-I00)

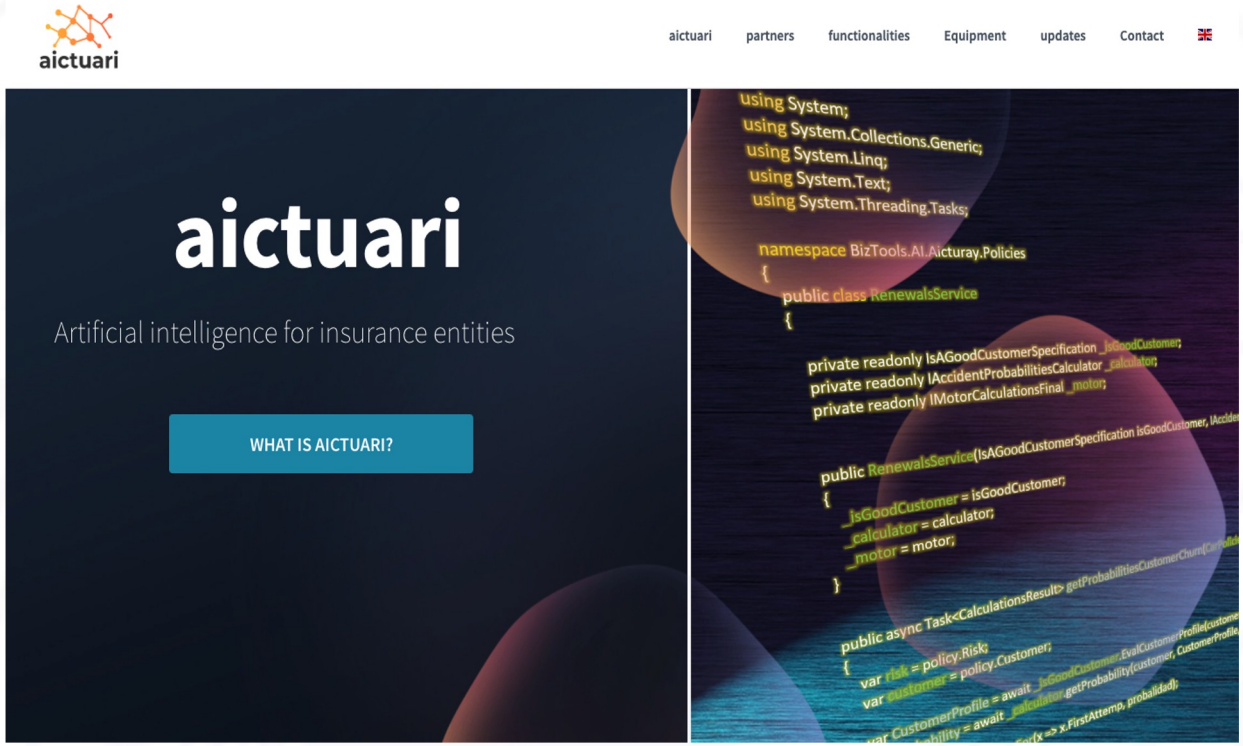
- Design, using artificial intelligence, of predictive algorithms for the identification of individuals at risk of developing overweight/obesity and its associated pathologies: Contribution of genetic analysis (GenObIA-CM) S2017/BMD-3773

- Determination of microscopic residual stresses using diffraction methods, EBSD maps, and evolutionary algorithms Y2018/NMT-4668 (Micro-Stress- MAP-CM)

- Development of advanced artificial intelligence techniques for cost optimization and risk minimization innsurance companies. (Industrial Doctorate IND2020/TIC-17435)

AICTUARI

- Development of advanced artificial intelligence techniques for cost optimization and risk minimization innsurance companies. (Industrial Doctorate IND2020/TIC-17435)



- Perform risk prediction modelling and design an automatic premium calculation system.
- Design implementations on state-of-the-art graphics hardware and FPGA-based systems. Study the feasibility and suitability of other types of hardware implementation.

MICRO-STRESS

Determination of microscopic residual stresses using diffraction methods, EBSD maps, and evolutionary algorithms

Y2018/NMT-4668 (Micro-Stress- MAP-CM)

Objetivos

Desarrollar una metodología basada en técnicas computacionales (Algoritmos

Challenge

- Estimation of total body fat using symbolic regression and evolutionary algorithms.
  - Body mass index (BMI) is commonly used to determine whether a person is overweight or obese.

$$BMI = \frac{Weight(kg)}{Height(m) \times Height(m)}$$

  - However, this value is inaccurate and is usually calculated using only the person's mass and height.
  - This challenge proposes the application of evolutionary computation to obtain a reliable and interpretable expression of body fat percentage
    - from public anthropometric data
    - National Health and Nutrition Examination Survey (NHANES)
      - US Centers for Disease Control and Prevention (CDC).
  - Collaboration with MIT (Prof. Omar Costilla)

- Symbolic Regression Problem
  - Obtain an expresión that best represent the data
  - 2404 samples
  - 11 variables
    - You can use more variables or less

NHANES Variable	Variable Description	Mean	Std.	Min.	25%tile	50%tile	75%tile	Max.
SEQN	Anonymous ID Number	–	–	–	–	–	–	–
RIAGENDR	Gender (1='M', 0='F')	–	–	–	–	–	–	–
RIDAGEYR	Age (years)	38.1	12.6	18.0	27.0	38.0	49.0	59.0
BMXWT	Weight (kg)	79.7	20.4	36.2	64.9	76.9	91.9	176.5
BMXHT	Height (cm)	166.6	9.3	138.3	159.4	166.5	173.8	190.2
BMXLEG	Upper Leg Length (cm)	39.5	3.6	26.0	37.0	39.5	42.0	50.0
BMXARML	Upper Arm Length (cm)	37.0	2.7	29.6	35.0	37.0	39.0	45.5
BMXARMC	Arm Circumference (cm)	33.1	5.1	20.7	29.4	32.9	36.4	52.7
BMXWAIST	Waist Circumference (cm)	96.0	16.3	56.4	83.8	94.7	106.4	154.9
BMXHIP	Hip Circumference (cm)	104.6	12.8	77.8	95.5	102.7	111.6	168.5
DXDTOPF	**Total Body Fat %**	**33.1**	**8.6**	**12.1**	**27.1**	**32.9**	**40.2**	**56.1**

Challenge

- Use Genetic Programming based on grammars
- Select the technique
- Design the grammar
- Tune the parameters
- Improve the results of
  - Information Fusion via Symbolic Regression: A Tutorial in the Context of Human Health. Jennifer J. Schnura, Nitesh V. Chawlaa, Lucy Family Institute for Data and Society, Department of Computer Science and Engineering, University of Notre Dame, 46556, IN., USA
  - https://doi.org/10.1016/j.inffus.2022.11.030

Software

- JECO: Java Evolutionary COmputation library.
  - GE and more EA
  - https://github.com/ABSysGroup/jeco
- PonyGE2
  - https://github.com/PonyGE/PonyGE2/
- Structured GE
  - https://github.com/nunolourenco/sge3
- Genetic Engine
  - A hybrid between strongly-typed (STGP) and grammar-guided genetic programming (GGGP).
  - https://pypi.org/project/GeneticEngine/

Why select this challenge?

- Real World Application
- Useful in Medicine
- Other challenges are very interesting
  - Some of you are going to select them for sure
- All my students are now finishing their PhDs
  - Nobody cares my problems in the Lab till december
- Anna Says that I am very “majo”