Smart Sensor Calibration using Auto-Rotating Perceptrons

LatinX in AI Research Workshop - ICML 2020

Daniel Saromo, Leonardo Bravo and Elizabeth Villota

Biomechanics and Applied Robotics Research Laboratory (GIRAB)
Pontifical Catholic University of Peru (PUCP)

We used the Auto-Rotating Perceptron (ARP) neural unit to calibrate a wearable sensor.

Our results show that **when changing** classic perceptrons **to ARP**, the **test loss** of the sigmoid networks was **reduced by a factor of 15**.

Introduction — Motivation

- Sports analysis techniques help athletes to increase their performance and to avoid incorrect practices that could lead to injuries.
- At sports court: Qualitatively evaluation is done by the coach.

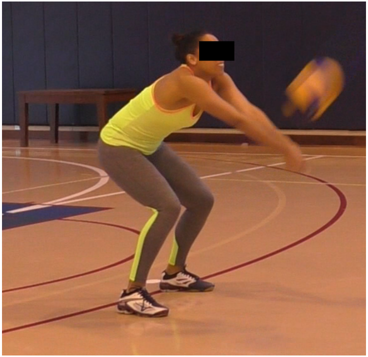


Figure 1: Volleyball athlete executing the technical reception gesture.

Introduction — Motivation

- Sports analysis techniques help athletes to increase their performance and to avoid incorrect practices that could lead to injuries.
- At sports court: Qualitatively evaluation is done by the coach.

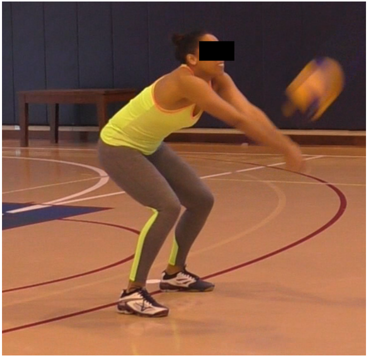


Figure 1: Volleyball athlete executing the technical reception gesture.

- **Main problems:** Lack of: objective assessment and wearable sensing.

Introduction — Motivation

- Research project:
  - Deals with biomechanical characterization of the service reception in volleyball.
  - A key cuantitative measurement is the Ground Reaction Force (GRF).

Introduction — Motivation

- Research project:
  - Deals with biomechanical characterization of the service reception in volleyball.
  - A key cuantitative measurement is the Ground Reaction Force (GRF).
  - Reliable sensors available are expensive, not portable and of limited area (such as the AMTI force plate).

Figure 2: AMTI sensor with its wooden pallet (left) and volleyball athlete executing reception service on the AMTI sensor at the sports court (right).

Introduction — Motivation

- Research project:
  - Deals with biomechanical characterization of the service reception in volleyball.
  - A key cuantitative measurement is the Ground Reaction Force (GRF).
  - Reliable sensors available are expensive, not portable and of limited area (such as the AMTI force plate).

Figure 2: AMTI sensor with its wooden pallet (left) and volleyball athlete executing reception service on the AMTI sensor at the sports court (right).

- More information about the project: `https://youtu.be/z8aMblOUp_I`.

Problem definition — WEVES

- An insole-type WEarable VErtical Sensor system (WEVES) for GRF measurement was developed at GIRAB laboratory, see Figure 3.

(a) (b)

Figure 3: **(a)** WEVES with insole detail. **(b)** Stand-up straight position test with AMTI (red) and WEVES (green) sensors.

Problem definition — WEVES

- An insole-type WEarable VErtical Sensor system (WEVES) for GRF measurement was developed at GIRAB laboratory, see Figure 3.

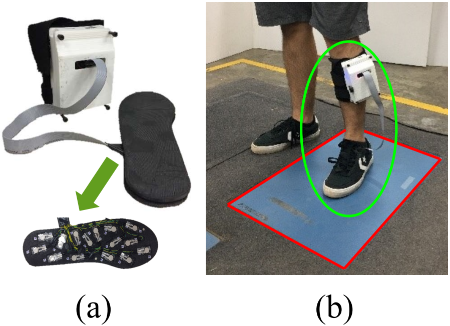


Figure 3: **(a)** WEVES with insole detail. **(b)** Stand-up straight position test with AMTI (red) and WEVES (green) sensors.

- WEVES measurement signal $\mathbf{w}$ must be as close as possible to the AMTI reference signal $\mathbf{p}$.

Problem definition — WEVES calibration

- Results from the WEVES output $\mathbf{w}$ and the AMTI platform signal $\mathbf{p}$ show the same shape with differences in amplitude, for all movements tested.

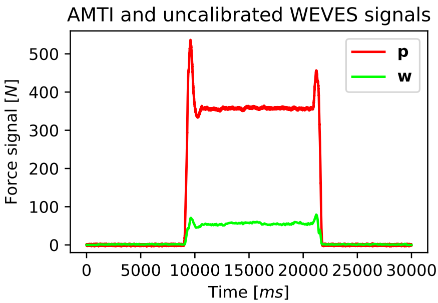


Figure 4: Force measurements (N): AMTI ($\mathbf{p}$, red) and uncalibrated WEVES ($\mathbf{w}$, green).

Problem definition — WEVES calibration

- Results from the WEVES output $\mathbf{w}$ and the AMTI platform signal $\mathbf{p}$ show the same shape with differences in amplitude, for all movements tested.

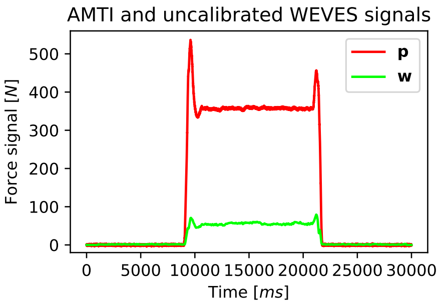


Figure 4: Force measurements (N): AMTI ($\mathbf{p}$, red) and uncalibrated WEVES ($\mathbf{w}$, green).

- Same shape

Problem definition — WEVES calibration

- Results from the WEVES output $\mathbf{w}$ and the AMTI platform signal $\mathbf{p}$ show the same shape with differences in amplitude, for all movements tested.

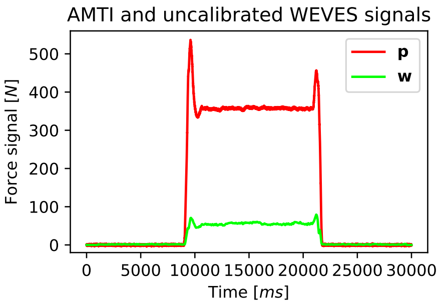


Figure 4: Force measurements (N): AMTI ($\mathbf{p}$, red) and uncalibrated WEVES ($\mathbf{w}$, green).

- Same shape $\rightarrow$ What happens if we scale the signal $\mathbf{w}$ until we reach $\mathbf{p}$?

Problem definition — WEVES calibration requires scaling

- What happens if we scale the signal $\mathbf{w}$ until we reach $\mathbf{p}$?

- What happens if we scale the signal $\mathbf{w}$ until we reach $\mathbf{p}$?

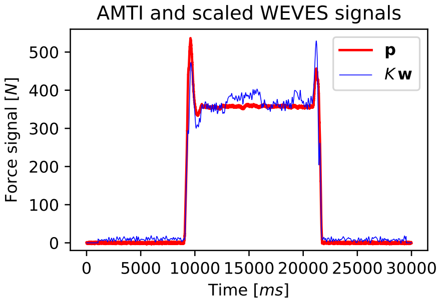


Figure 5: Force measurements (N): AMTI ($\mathbf{p}$, red) and scaled WEVES ($K\mathbf{w}$, blue).

Problem definition — WEVES calibration requires scaling

- What happens if we scale the signal $\mathbf{w}$ until we reach $\mathbf{p}$?

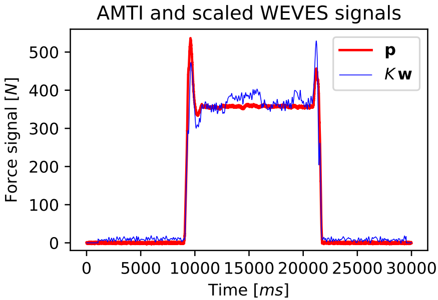


Figure 5: Force measurements (N): AMTI ($\mathbf{p}$, red) and scaled WEVES ($K\mathbf{w}$, blue).

- We need to find the calibration factor K that makes $\mathbf{p}$ and $K\mathbf{w}$ similar.

Problem definition — WEVES calibration requires scaling

- How much do we have to *lift* the WEVES signal in order to reach $\mathbf{p}$?

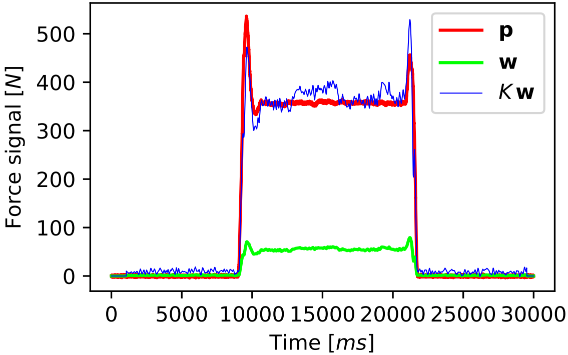


Figure 6: Force measurements (N) overlapped: AMTI ($\mathbf{p}$, red), uncalibrated WEVES ($\mathbf{w}$, green), and scaled WEVES ($K\mathbf{w}$, blue).

Problem definition — WEVES calibration requires scaling

- How much do we have to *lift* the WEVES signal in order to reach $\mathbf{p}$?

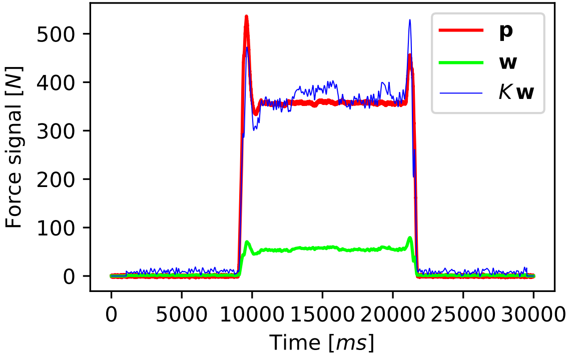


Figure 6: Force measurements (N) overlapped: AMTI ($\mathbf{p}$, red), uncalibrated WEVES ($\mathbf{w}$, green), and scaled WEVES ($K\mathbf{w}$, blue).

- In real applications, only the WEVES signal $\mathbf{w}$ will be available to find the corresponding calibration factor K.

Methodology — Calculating the calibration factor K

- WEVES calibration posed as a supervised regression problem.
- Input: $\mathbf{w}$. Output: K.

Methodology — Calculating the calibration factor K

- WEVES calibration posed as a supervised regression problem.
- Input: $\mathbf{w}$. Output: K.

1. **Dataset generation:** We measured the difference between $\mathbf{p}$ and scaled $\mathbf{w}$ using the Root Mean Square Error (RMSE). Then, finding K is posed as an optimization problem:

$$K = \arg\min_{K} \left\{ \mathrm{RMSE}\left(\mathbf{p},\, K\mathbf{w}\right) \right\},\ K > 0.$$

Figure 7: Searching the K values that make $\mathbf{p}$ and $K\mathbf{w}$ similar.

- Optimizers tested to calculate K: Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO).

Methodology — Calculating the calibration factor K

2. **Factor K prediction:** We trained four neural regression model types to predict the target K using only $\mathbf{w}$.

$$\mathbf{w} \xrightarrow{\text{Deep Learning}} K$$

Figure 8: Inferring K from the WEVES signal $\mathbf{w}$.

2. **Factor K prediction:** We trained four neural regression model types to predict the target K using only $\mathbf{w}$.

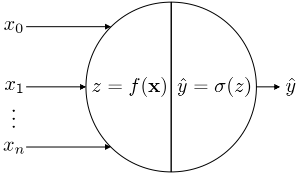
$$\mathbf{w} \xrightarrow{\text{Deep Learning}} K$$

Figure 8: Inferring K from the WEVES signal $\mathbf{w}$.

- We looked for reducing the loss prediction error of the regression. For this problem, we tested the effect of changing classic perceptrons to Auto-Rotating Perceptrons (ARP), with two activation functions (ReLU and sigmoid).

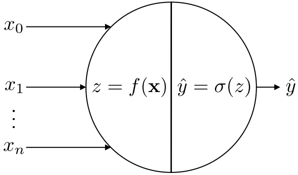
Methodology — What is an Auto-Rotating Perceptron?

Methodology — What is an Auto-Rotating Perceptron?

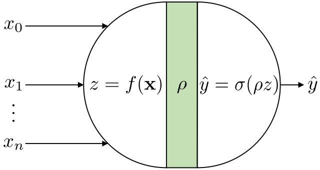


$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + w_0\, x_0 \qquad x_0 = 1$

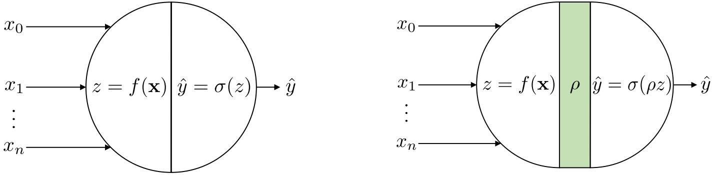
Methodology — What is an Auto-Rotating Perceptron?



$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + w_0\, x_0 \qquad x_0 = 1$



$g(\mathbf{x}) = \rho\, f(\mathbf{x}) \qquad \rho = \frac{L}{|f(\mathbf{x}_Q)|}$



$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + w_0\, x_0 \qquad x_0 = 1$

$g(\mathbf{x}) = \rho\, f(\mathbf{x}) \qquad \rho = \frac{L}{|f(\mathbf{x}_Q)|}$

Figure 9: Classic perceptron (left) and ARP (right).

Methodology — What is an Auto-Rotating Perceptron?

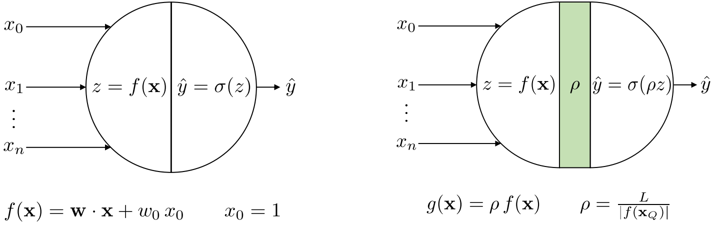


Figure 9: Classic perceptron (left) and ARP (right).

- The ARP, proposed by Saromo et al. [1], is an innovative neural unit that aims to avoid the vanishing gradient problem by making the z inputs of the perceptron activation $\sigma(z)$ near zero with no learning alteration.
- The modification is achieved by multiplying the linear transformation $f(\mathbf{x})$ with an scalar coefficient ρ before the activation function $\sigma(\cdot)$.
- ARP has two hyperparameters: $\mathbf{x}_Q = \langle x_Q, \cdots, x_Q \rangle \in \mathbb{R}^n$ and $L \in \mathbb{R}$.

Methodology — Dynamic region of the neurons

- e.g., $\sigma(z)$: Unipolar sigmoid. If $\sigma'(z) \approx 0 \rightarrow$ Unwanted node saturation.
- Nodes need to be in their **dynamic region $\mathcal{L}$**. ARP let us control that.

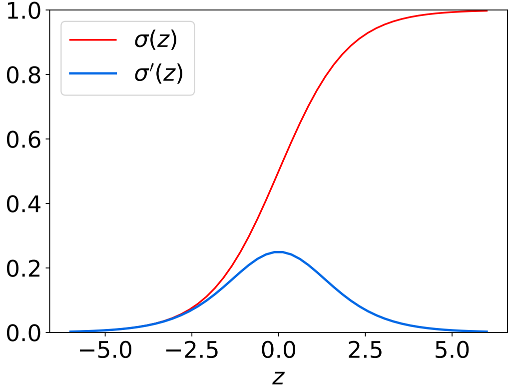


Figure 10: Sigmoid activation function $\sigma(z)$ and its derivative.

Methodology — Dynamic region of the neurons

- e.g., $\sigma(z)$: Unipolar sigmoid. If $\sigma'(z) \approx 0 \rightarrow$ Unwanted node saturation.
- Nodes need to be in their **dynamic region $\mathcal{L}$**. ARP let us control that.

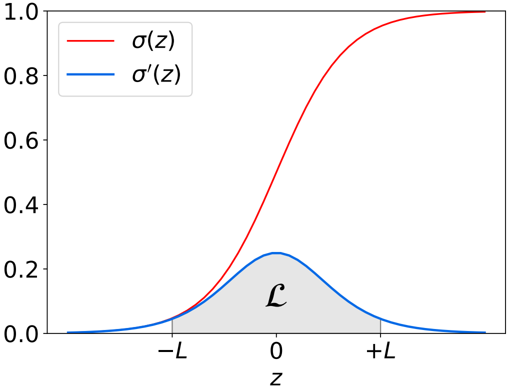


Figure 10: Sigmoid activation function $\sigma(z)$ and its derivative.

Experimentation — Learning results

- ARP hyperparameters: $x_Q = 2.6$, since the maximum input value is $1 < x_Q$; and $L = 3$, because the sigmoid derivative isn't very small for $|z| \leq 3$.

- ARP hyperparameters: $x_Q = 2.6$, since the maximum input value is $1 < x_Q$; and $L = 3$, because the sigmoid derivative isn't very small for $|z| \leq 3$.

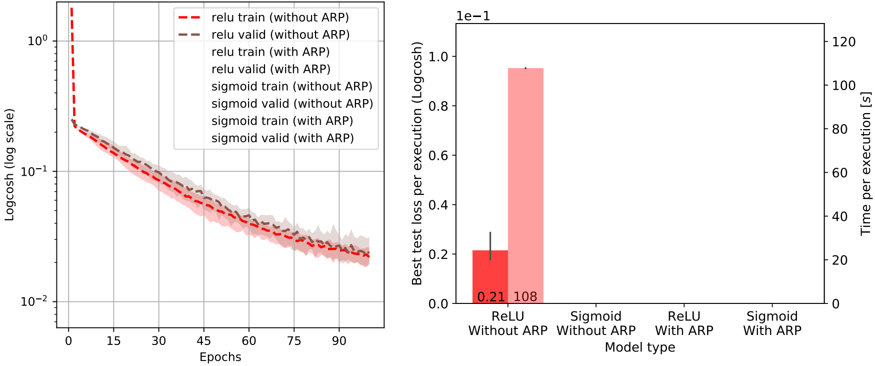


Figure 11: Comparison of the four model types tested. For each model family: 50 executions with 100 epochs per execution.

Experimentation — Learning results

- ARP hyperparameters: $x_Q = 2.6$, since the maximum input value is $1 < x_Q$; and $L = 3$, because the sigmoid derivative isn't very small for $|z| \leq 3$.

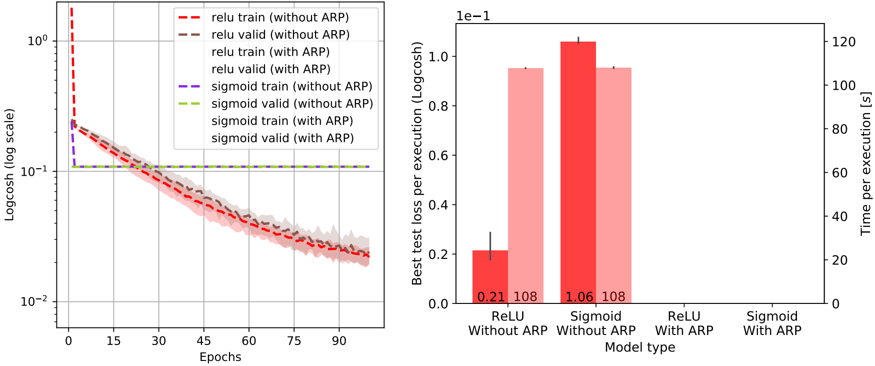
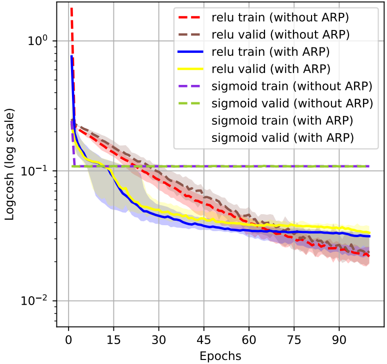


Figure 11: Comparison of the four model types tested. For each model family: 50 executions with 100 epochs per execution.

Experimentation — Learning results

- ARP hyperparameters: $x_Q = 2.6$, since the maximum input value is $1 < x_Q$; and $L = 3$, because the sigmoid derivative isn't very small for $|z| \leq 3$.



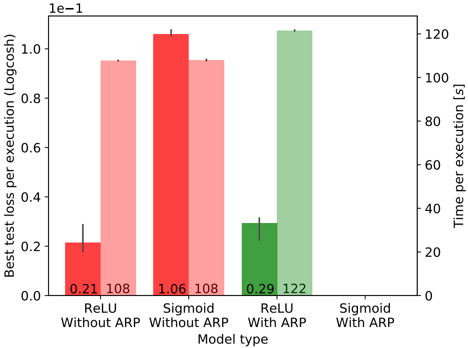


Figure 11: Comparison of the four model types tested. For each model family: 50 executions with 100 epochs per execution.

- ARP hyperparameters: $x_Q = 2.6$, since the maximum input value is $1 < x_Q$; and $L = 3$, because the sigmoid derivative isn't very small for $|z| \leq 3$.

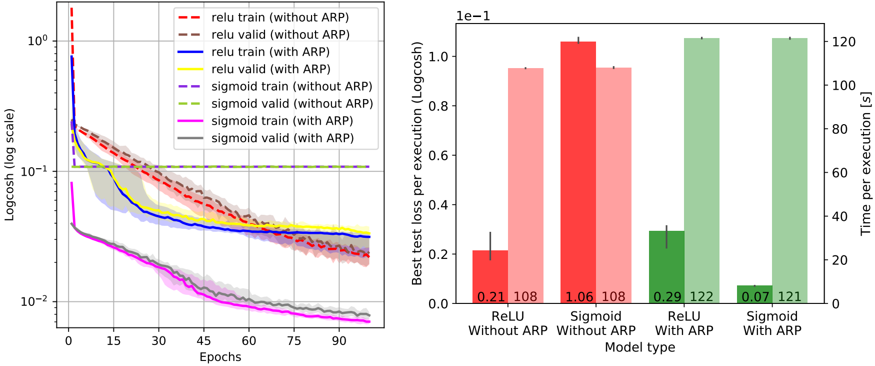


Figure 11: Comparison of the four model types tested. For each model family: 50 executions with 100 epochs per execution.

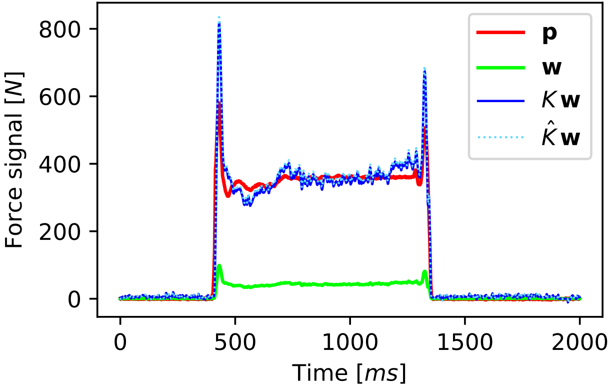


Figure 12: Force measurements (N) overlapped: AMTI ($\mathbf{p}$, red); uncalibrated WEVES ($\mathbf{w}$, green); scaled WEVES ($K\mathbf{w}$, blue) with K calculated using $\mathbf{p}$ and $\mathbf{w}$; and calibrated WEVES ($\hat{K}\mathbf{w}$, blue) with $\hat{K}$ calculated using only $\mathbf{w}$.

Conclusion — Key Takeaways

- We employed the ARP neural unit aiming to calibrate a wearable GRF sensor.
- ARP-sigmoid networks can have a better performance than ReLU networks with classic neurons **without altering the inference structure** learned by the perceptron.
- Compared with classic perceptrons that use sigmoid, the **test loss** of the sigmoid-ARP networks was **reduced by a factor of 15** at the cost of increasing the execution time by $\sim 12\%$.

Acknowledgments

References

SAROMO, Daniel; VILLOTA, Elizabeth; VILLANUEVA, Edwin.
Auto-Rotating Perceptrons.
LXAI Workshop at NeurIPS. Vancouver, 2019.
https://arxiv.org/pdf/1910.02483.pdf

More information about the ARP neural unit available at:
`https://danielsaromo.xyz/ARP`