Eco-RETINA: A Green, Flexible Algorithm for Model Building

Extended Abstract

Javier Capilla, Alba Alcaraz, UCM,

Ángel Valarezo, Alfredo García-Hiernaux, Teodosio Pérez-Amaral, ICAE, UCM

Eco-RETINA is an innovative and eco-friendly algorithm explicitly designed for out-of-sample prediction. Functioning as a regression-based flexible approximator, it is linear in parameters but nonlinear in inputs, employing a selective model search to optimize performance. The algorithm adeptly manages multicollinearity while emphasizing speed, accuracy, and environmental sustainability. Its modular and transparent structure facilitates easy interpretation and modification, making it an invaluable tool for researchers in developing explicit models for out-of-sample forecasting. The algorithm generates outputs such as a list of relevant transformed inputs, coefficients, standard deviations, and confidence intervals, enhancing its interpretability.

Now implemented in Python and available on GitHub, Eco-RETINA introduces several new features, including measuring CO_2 emissions and energy consumption. These enhancements, alongside improved data transformations, bottleneck elimination, and a user-friendly interface, significantly boost its performance. The algorithm achieves remarkable reductions in carbon footprint and power consumption (ranging from 50% to 90%) while significantly reducing computational time. Empirical results indicate that Eco-RETINA is not only a sustainable alternative to conventional neural networks but also surpasses them in certain aspects, offering a competitive edge in accuracy, interpretability, and environmental impact.

Motivation and Characteristics

The power consumption and carbon footprint associated with AI-related algorithms have become pressing concerns for researchers, industry, and policymakers. Lynn et al. (2021), OECD (2022), Barbierato et al. (2024), among others, have highlighted the critical importance of addressing these issues, with some researchers predicting the potential depletion of available energy resources by 2030. To tackle these challenges, major technological firms and initiatives such as the European Commission's Enfield project are investing heavily in sustainable AI solutions.

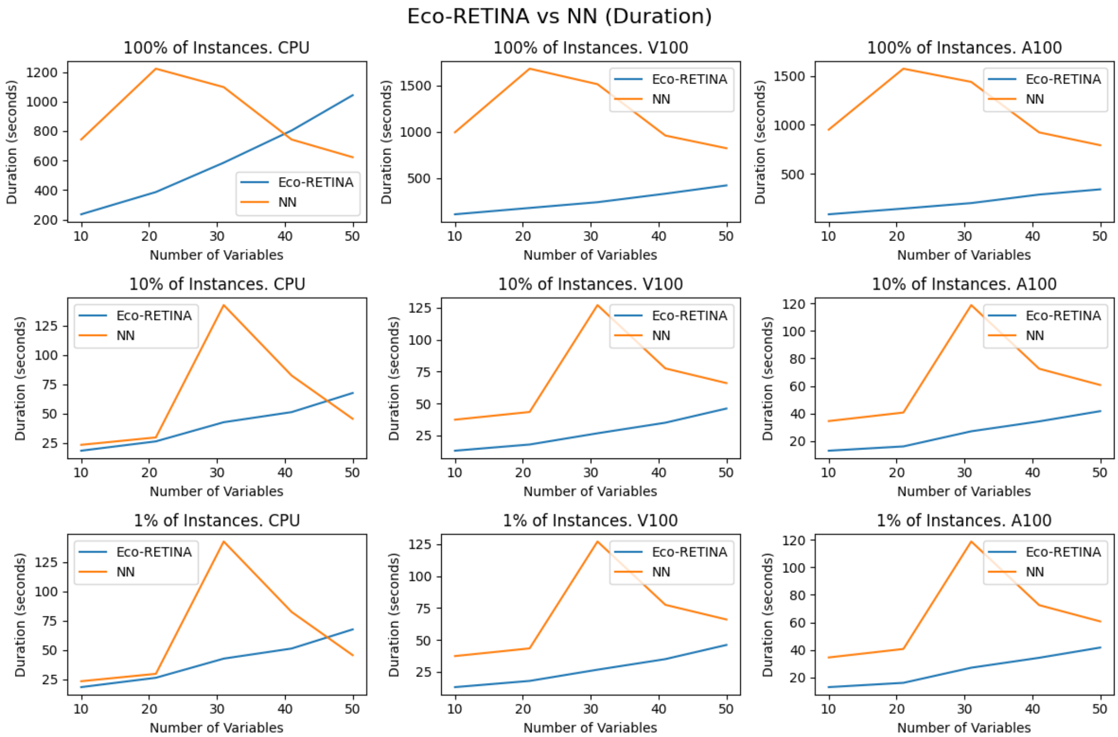
The energy costs of developing, training, and deploying deep learning models have surged in recent years, driven by the increasing complexity of neural networks and their reliance on computationally intensive resources like GPUs and TPUs. As a result, the development of efficient and sustainable algorithms has become imperative. However, many of these innovations come at the cost of reduced performance compared to their original counterparts.

Eco-RETINA addresses this gap by providing a green algorithm specifically designed for out-of-sample prediction. Building on the foundational work of Pérez-Amaral et al. (2003), Eco-RETINA (Capilla, 2024) integrates substantial improvements over previous versions. It operates as a regression-based flexible approximator, linear in parameters for improved convergence and nonlinear in inputs for enhanced flexibility. The algorithm's selective search process allows faster execution without compromising accuracy. It effectively handles multicollinearity through a threshold-based approach and incorporates robust mechanisms for outlier detection and management.

Eco-RETINA is modular and transparent, enabling users to replace individual components with alternative modules tailored to specific problems. Unlike black-box models, its transparent nature allows for modification and understanding by the user. Its outputs, akin to those of traditional econometric packages, include explicit forecasting models, point estimates, parameter standard deviations, and confidence intervals, all of which are interpretable.

Written in Python and available on GitHub, Eco-RETINA incorporates functionalities for measuring CO_2 emissions and energy expenditure. The algorithm also allows for a range of data transformations, including ratios, cross-products, inverses, and logarithmic transformations of inputs. Bottleneck elimination during programming ensures optimal performance by avoiding repetitive, time- and energy-intensive operations. A newly developed user interface enhances accessibility and usability.

Figure 1 The duration of Eco-RETINA vs a relevant NN



Eco-RETINA is faster over much of the relevant range of the number of variables.

Figure 2. Eco-RETINA output summary

```
eco_retina.model.summary()
```

✓ 0.0s

OLS Regression Results

Dep. Variable:	obj_variable	R-squared:	0.839
Model:	OLS	Adj. R-squared:	0.839
Method:	Least Squares	F-statistic:	5.631e+04
Date:	Sun, 14 Apr 2024	Prob (F-statistic):	0.00
Time:	18:39:16	Log-Likelihood:	-7.0650e+05
No. Observations:	86553	AIC:	1.413e+06
Df Residuals:	86544	BIC:	1.413e+06
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
constant	2406.9761	1093.325	2.202	0.028	264.068	4549.884
carat_weight^1*meas_width^1	1012.5979	1.530	662.004	0.000	1009.600	1015.596
meas_width^-1*meas_length^1	265.2951	18.546	14.304	0.000	228.944	301.646
polish_Excellent*depth_percent	110.3873	25.818	4.276	0.000	59.784	160.991
polish_Excellent*table_percent	-75.9719	27.756	-2.737	0.006	-130.373	-21.571
polish_Very Good	2149.2133	603.394	3.562	0.000	966.566	3331.860
polish_Excellent	257.7628	584.608	0.441	0.659	-888.064	1403.590
polish_Very Good*depth_percent	91.2128	26.122	3.492	0.000	40.013	142.412
polish_Very Good*table_percent	-91.0710	27.773	-3.279	0.001	-145.506	-36.636
table_percent^-1*depth_percent^1	-5306.3593	1510.614	-3.513	0.000	-8267.151	-2345.568

Omnibus:	13326.421	Durbin-Watson:	2.006
Prob(Omnibus):	0.000	Jarque-Bera (JB):	166250.786
Skew:	0.336	Prob(JB):	0.00
Kurtosis:	9.756	Cond. No.	2.34e+16

Model of the prices of diamonds as a function of their characteristics

- constant
- carat_weight*meas_width: carat weight times measured width
- meas_length/meas_width: measured length over measured width
- polish_Excellent*depth_percent: Excellent polish times percent depth
- polish_Excellent*table_percent: Excellent polish times percent table
- polish_very good: very good polish
- polish-Excellent: Excellent polish
- polish_Very Good*table_percent: very good polish times percent table
- polish_very good*depth_percent: very good polish times percent depth
- depth_percent/table_percent: percent depth over percent table

The algorithm gives us a parsimonious model using 9 (possibly transformed) inputs plus the constant. Eco-RETINA also provides the usual statistics for regression.

Conclusions

Eco-RETINA represents a significant advancement over its predecessors by integrating additional capabilities such as outlier management, enhanced multicollinearity handling, expanded input transformations, bottleneck elimination, and the ability to measure power consumption, carbon footprint, and execution time. Experiments demonstrate that Eco-RETINA achieves prediction errors comparable to earlier versions while significantly reducing power consumption and computational time (up to 4.25 times faster).

When compared to neural networks, Eco-RETINA shows several advantages. While neural networks may achieve slightly lower prediction errors, Eco-RETINA's training emissions are substantially smaller in most experiments. Additionally, it outperforms neural networks in accuracy when used with a moderate number of inputs. Its openness, interpretability, and speed position it as a strong contender among Green AI algorithms.

Eco-RETINA provides an explicit forecasting model based on transformed inputs, along with comprehensive diagnostics such as estimated coefficients, t and F statistics, and coefficient of determination. As a transparent and efficient tool, it complements neural networks and serves as a cost-effective exploratory data tool for identifying promising subsets of transformed inputs.

In an era where sustainable AI is increasingly essential, Eco-RETINA's emphasis on speed, accuracy, and environmental sustainability underscores its critical role in the future of algorithm development. By offering a green, interpretable, and high-performing alternative, Eco-RETINA sets a new standard for sustainable machine learning and forecasting tools.

References

- Barbierato, E., & Gatti, A. (2024). Towards Green AI. A methodological survey of the scientific literature. IEEE Access.
- Capilla, J. (2024). *Modelado automático con RETINA desde un enfoque de Green AI*. Facultad de Económicas y Empresariales, Universidad Complutense de Madrid.
- Lynn, H. K., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., et al. (2021). *Aligning artificial intelligence with climate change mitigation*. HAL Archives.
- OECD (2022). Measuring the environmental impacts of AI compute and applications: the AI footprint: digital economy papers, November 2022 No. 341.
- Pérez-Amaral, T., Gallo, G., & White, H. (2003). *A Flexible Tool for Model Building: The Relevant Transformations of the Inputs Network Approach*

(RETINA). Oxford Bulletin of Economics and Statistics, 65(Supplement), 305–323.

- White, H. (1998). *Artificial Neural Networks and Alternative Methods for Assessing Naval Readiness*. Technical Report, NRDA, San Diego.