



Explainable Prediction Of Digital Brain Function Screen Score From Self-reported Lifestyle And Clinical Context

Jermyn See¹, See Ann Soo¹, Nav Vij¹, Prem Pillay²

¹Neurowyze Pte Ltd, Singapore, ²Singapore Brain Spine Nerves Center, Singapore

A 500x bootstrap-bagged Explainable Boosting Machine (EBM) trained on N=4,345 routine screenings (Singapore & Malaysia) predicts Digital Brain Function Screen (DBFS) cognitive score from self-reported lifestyle and demographic context ($R^2=0.554$, $r=0.744$, $RMSE=21.6$). Cognitively stimulating activity (~14 pts) and education (~7 pts) show the largest feature-response spans, with transparent per-patient explanations supporting structured clinician-led lifestyle counselling.

INTRODUCTION

Background

- Lifestyle factors are **modifiable determinants of cognitive decline**.
- Most screening pathways lack a practical way to **translate self-report into individualised guidance** alongside the test result.
- Black-box models don't fit clinic workflows**; clinicians need to see *why* a score is what it is before acting on it.

Objective

- Build an **interpretable model** paired with DBFS that estimates expected screening performance from routine self-report and **produces transparent, patient-specific explanations** for clinician-led follow-up.

About DBFS

- 15-minute digital cognitive assessment**: memory, executive function, attention, validated for use in routine health screening.
- Returns an **overall DBFS score** that is a readout of cognitive health and **indicator of MCI risk**.
- Includes a **self-reported lifestyle questionnaire** (Table 1).

METHODS

Data

- N = 4345**, routine health screenings, Singapore & Malaysia
- Exclusion criteria**: individuals with diagnosed dementia or major neurologic/psychiatric illness.

Category	Variables Assessed
Modifiable	
Physical Activity	Exercise frequency
Nutrition	Olive oil, dairy, vegetable, etc. intake
Sleep	Average nightly sleep duration
Substance Use	Smoking status, alcohol use
Medical Conditions	Diabetes, hypertension, high cholesterol, etc.
Psychosocial Wellbeing	Mood symptoms, depression, etc.
Cognitive Engagement	Frequency of work/leisure cognitively stimulating activity
Non-Modifiable	
Age	Continuous (in years)
Sex	Male/female
Education	Degree attainment

Table 1. Details of self-reported demographics, lifestyle and clinical history questionnaire deployed in DBFS.

Model: Explainable Boosting Machine (EBM)

- Glassbox additive model**: each feature's contribution is directly readable as a term-contribution shape function.
- 500x bootstrap-bagging**: term contributions averaged across resamples for stability; ± 1 SD bands shown per term.

Evaluation

- Held-out 80/20 split** (single constrained EBM) as the generalisation estimate (Fig 1); deployed model is the **full 500x bag**.

RESULTS

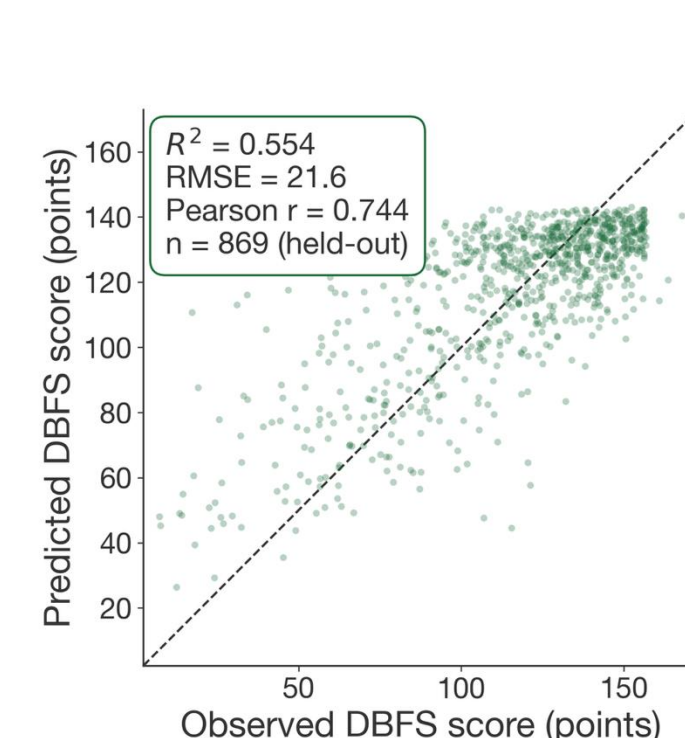


Fig. 1. Predicted vs. observed DBFS scores on held-out test set ($n=869$, 80/20 split). Dashed line = perfect agreement. $R^2=0.554$, $RMSE=21.6$, $Pearson\ r=0.744$.

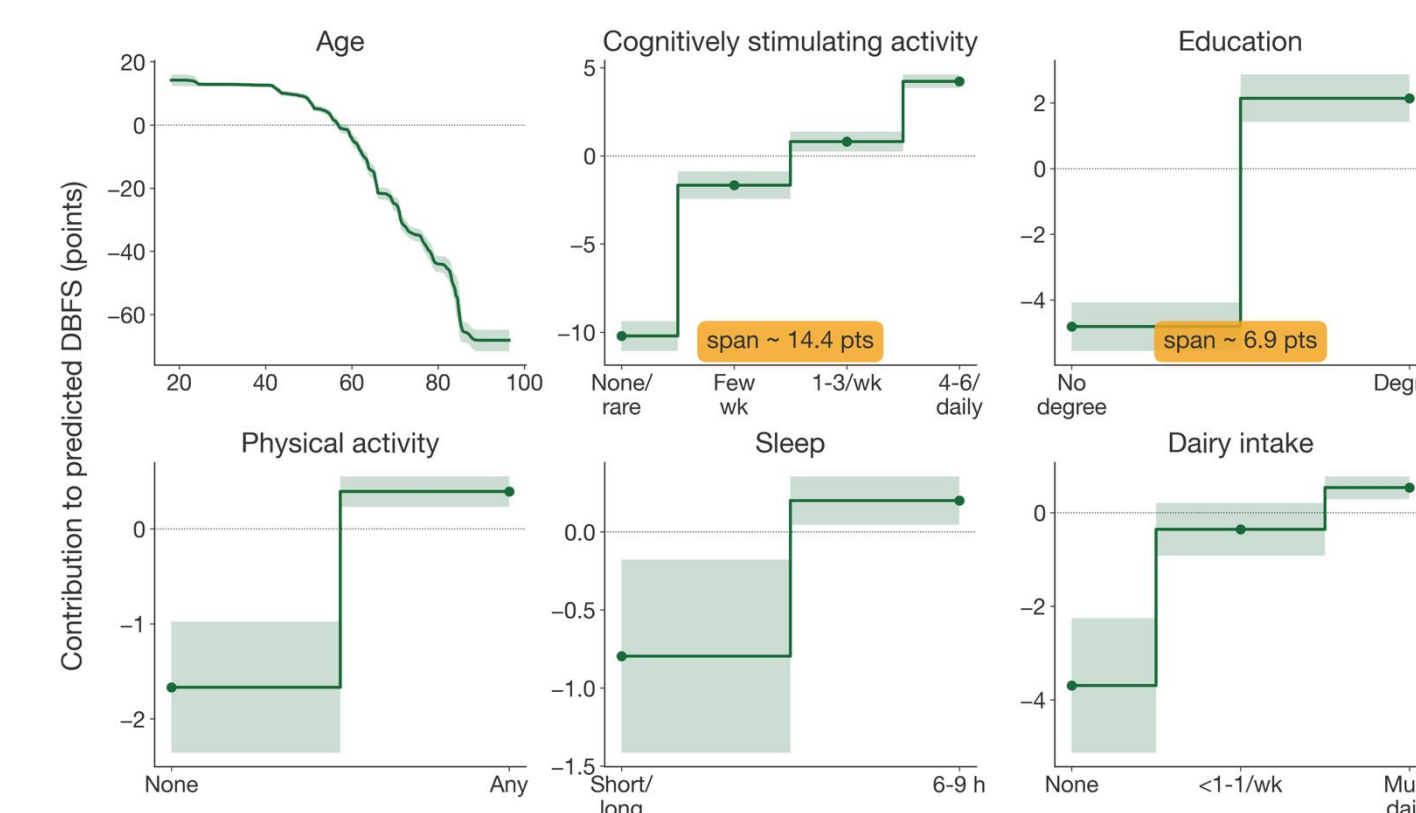


Fig. 2. Term-contribution shape functions for six model features, with ± 1 SD bands across 500 bootstrap resamples (bagging stability). Y-axis: contribution to predicted DBFS score (points), mean-centred. Annotated spans reflect low-to-high category difference for cognitively stimulating activity (~14.4 pts) and education (~6.9 pts).

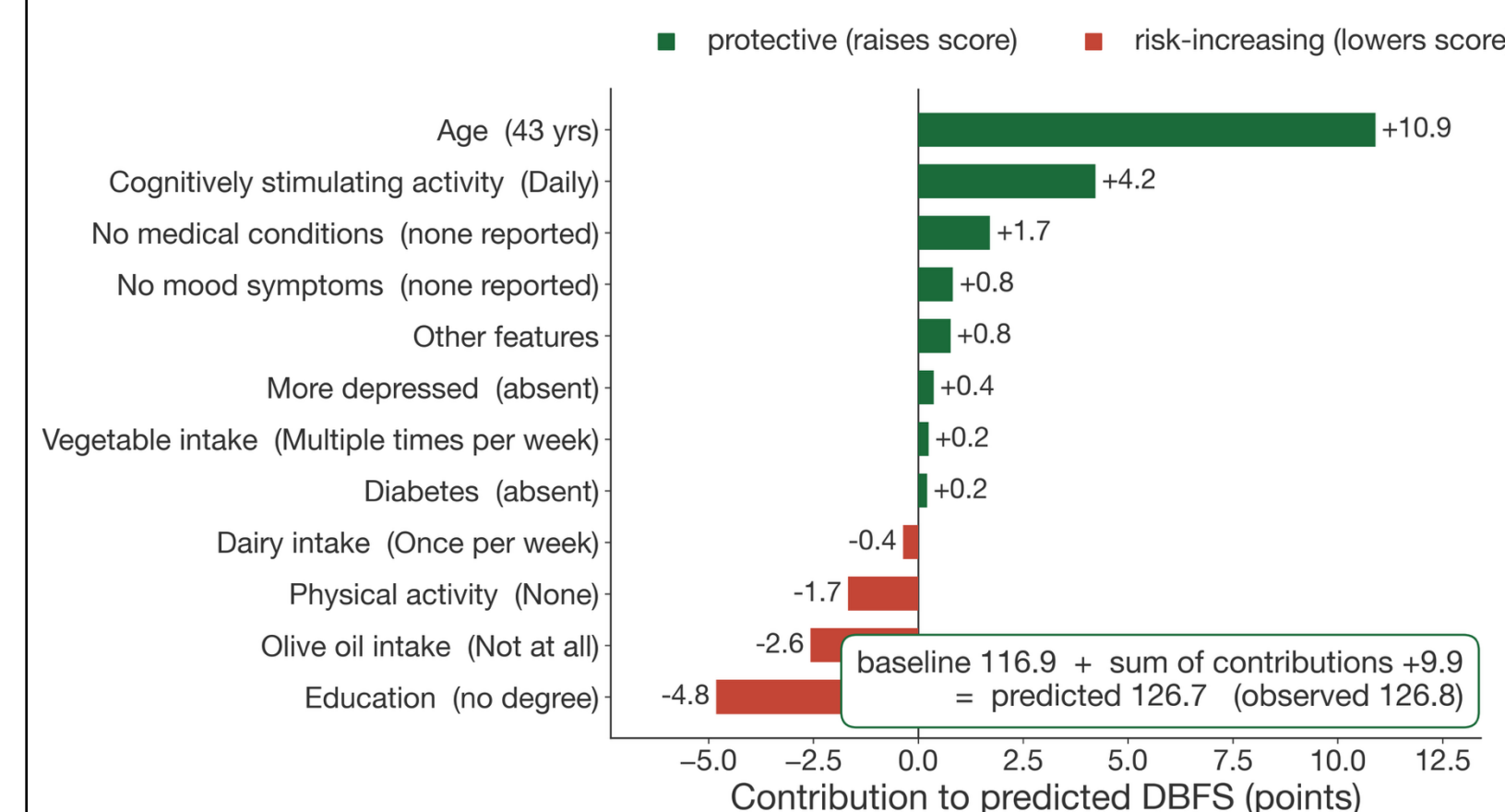


Fig. 3. Patient-specific explanation for a representative case (observed 126.8, predicted 126.7). Bars show each feature's signed contribution to the predicted score relative to the model baseline (116.9). Green = protective (raises score); red = risk-increasing (lowers score).

- Held-out performance**: $R^2=0.554$, $RMSE=21.6$, $r=0.744$ ($n=869$); **ceiling compression** above ~130 pts
- Largest modifiable effects**: cognitively stimulating activity (~14.4 pts) and education (~6.9 pts); wide stability bands on rare categories (physical activity, sleep) reflect data sparsity, not instability (Fig. 2)
- Per-patient explanations are fully auditable**: baseline + signed feature contributions sum to the predicted score (Fig. 3)

CONCLUSIONS

Conclusions

- Bagged EBM **predicts DBFS performance** from routine self-report with transparent, per-patient explanations, **no extra data collection needed**.
- Largest modifiable effects** (cognitive activity, education) are consistent with cognitive reserve; appropriate for contextualising scores, not causal targets.
- Supports **structured lifestyle counselling** as a direct add-on to DBFS screening
- Limitations**: internal evaluation only (external/temporal validation pending); cross-sectional. Figures reflect updated model ($N=4,345$); metrics differ slightly from submitted abstract.

REFERENCES

- Livingston G, et al. *Lancet*. 2024;404:572–628.
- Lou Y, Caruana R, Gehrke J. *Proc. ACM SIGKDD*. 2012:150–158.
- Nori H, et al. InterpretML. arXiv:1909.09223. 2019.
- Liu Y, et al. *Front Aging Neurosci*. 2024;16:1358992.
- Chew EY, et al. *Bioinformatics*. 2023;19(5):522–524.

For further inquiries about the study, please contact Jermyn See at jermyn.see@neurowyze.com.