Transfer learning for neuroimaging data

Esten H. Leonardsen

April 8, 2025

UNIVERSITETET
I OSLO

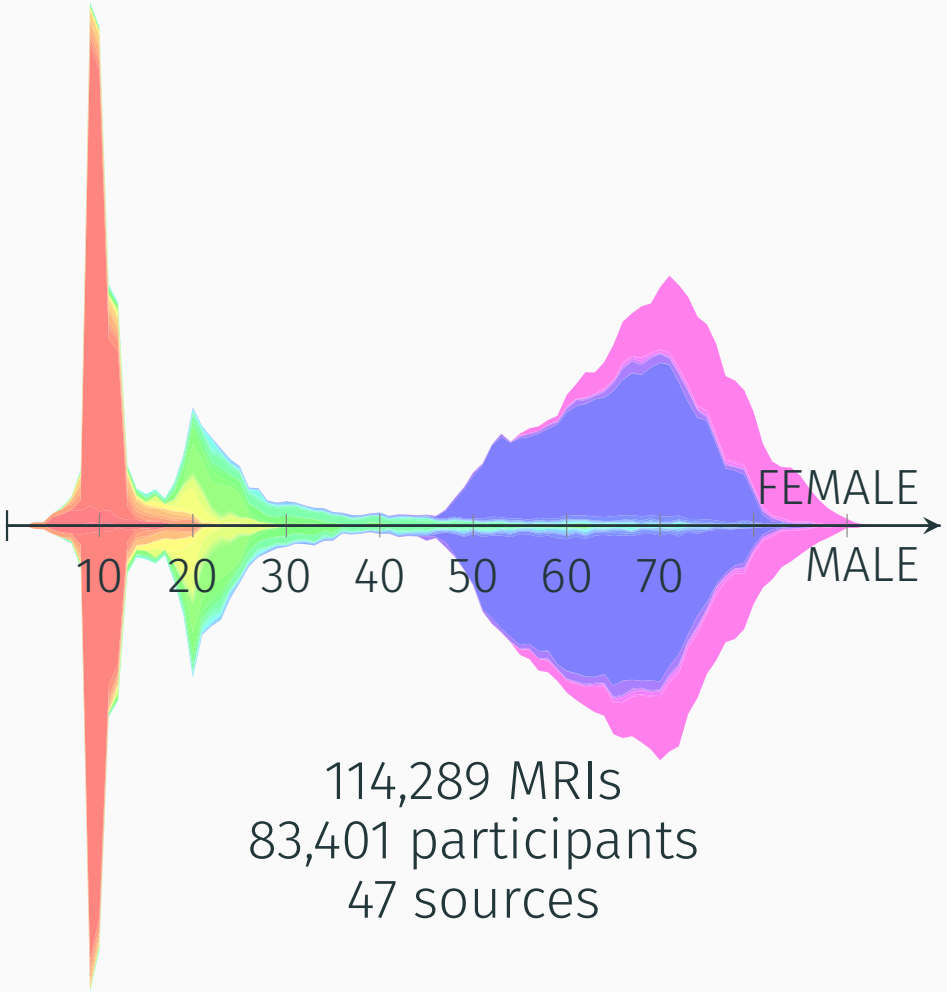
Background

- Pretrained models for computer vision have facilitated the democratization of cutting-edge technology for processing natural images, allowing practitioners to surpass the state-of-the-art in a wide range of tasks even when resources are limited

→ Why not try to achieve this for neuroimaging data as well?

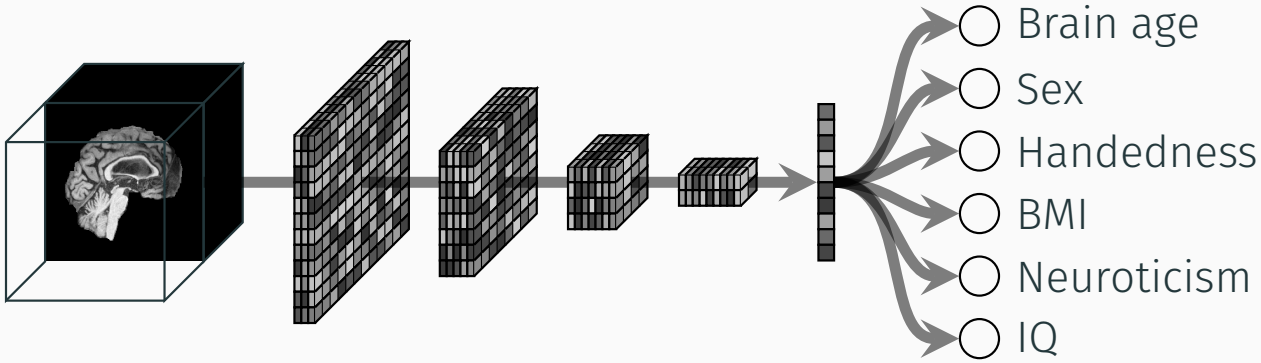
- Bonus: Large, complex neural networks trained on vast amounts of heterogeneous data can learn non-linear and abstract representations that allow us to further understand brain variability.

Methods



- ds002424
- HBN
- ABCD
- QTAB
- PING
- ADHD200
- PNC
- ABIDE-II
- ds000119
- ABIDE-I
- BRAINMINT
- SLIM
- QTIM
- Beijing
- AOMIC-PIOP2
- ds000202
- AOMIC-PIOP1
- AOMIC-ID1000
- CoRR
- HCP
- FCON1000
- ds000171
- TOP
- SCZ-Z
- NIMH
- NKI-RS
- MPI-LEMON
- ds003592
- ds004302
- ds000222
- SALD
- IXI
- DLBS
- Cam-CAN
- StrokeMRI
- PPMI
- UKBB
- Tao-Wu
- ds000245
- OASIS3
- Demgen
- NEUROCON
- MIRIAD
- ds004392
- AIBL
- ANM
- ADNI

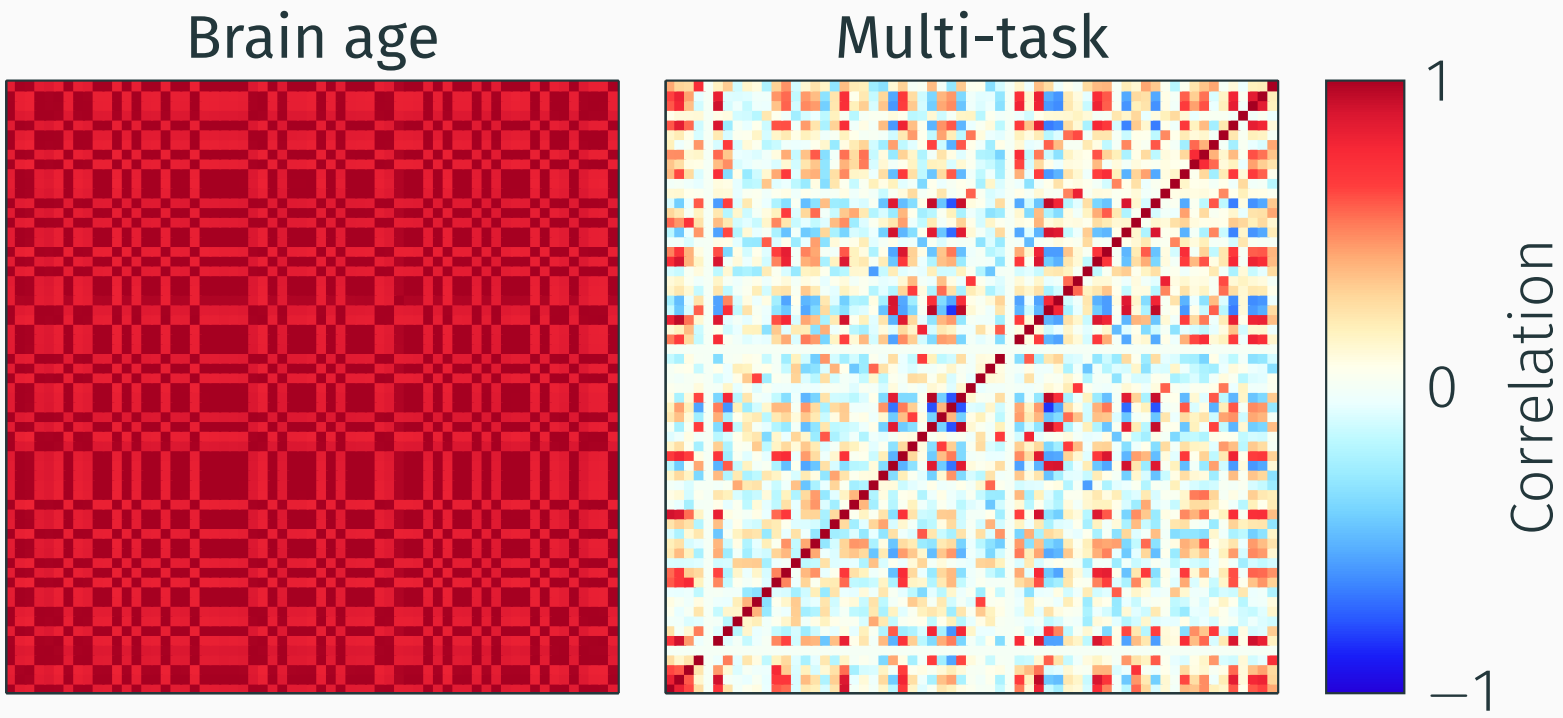
Methods



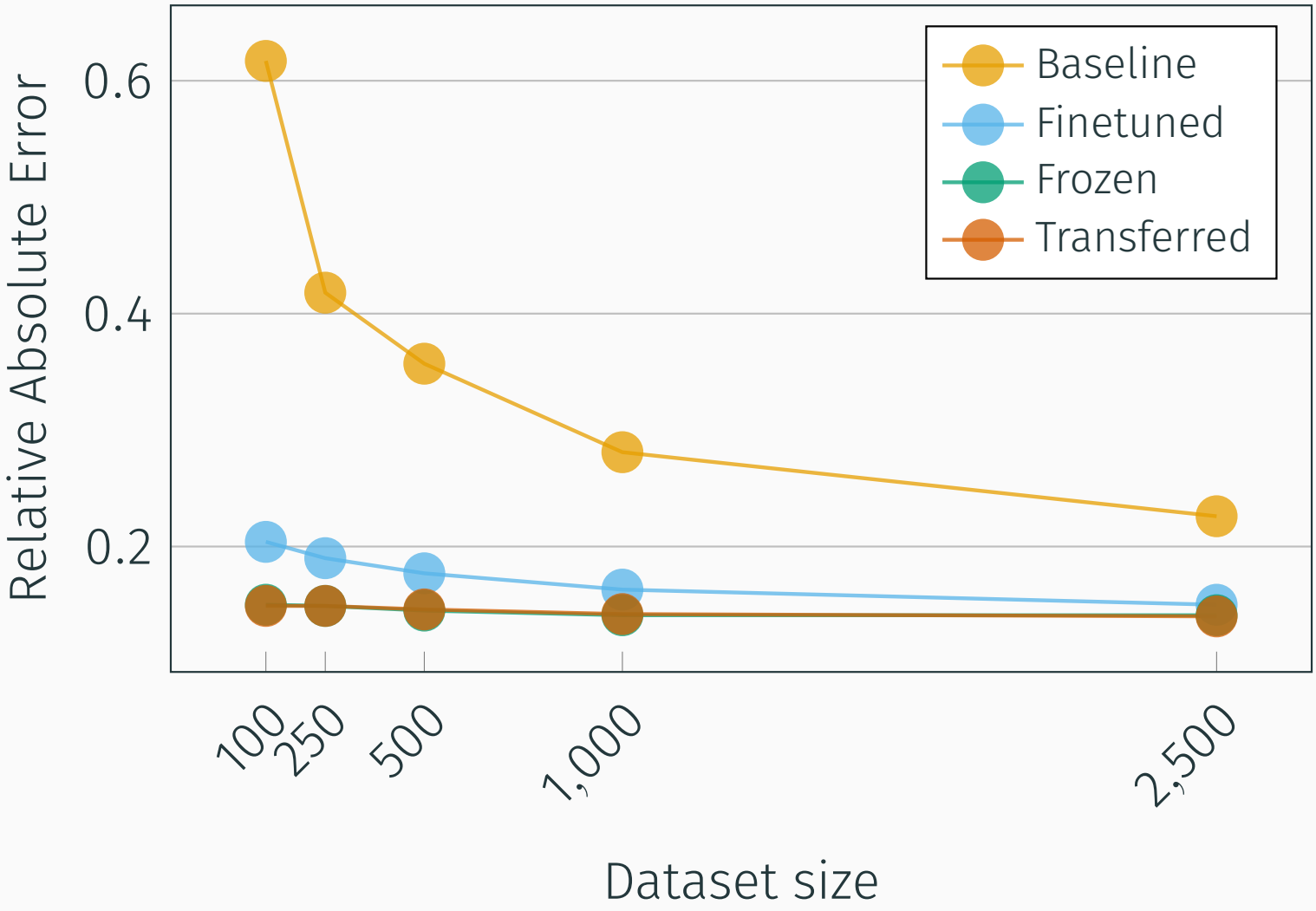
Results


Age
Sex
Handedness
BMI
Intelligence
Neuroticism
Multi-task (validation)
Multi-task (test)
Brain age (test)
Predicted
Observed
MAE 2.21
AUC 0.99
AUC 0.61
MAE 2.38
R 0.27
R 0.17
MAE 3.41
AUC 0.98
AUC 0.48
MAE 1.95
MAE 3.68

Results



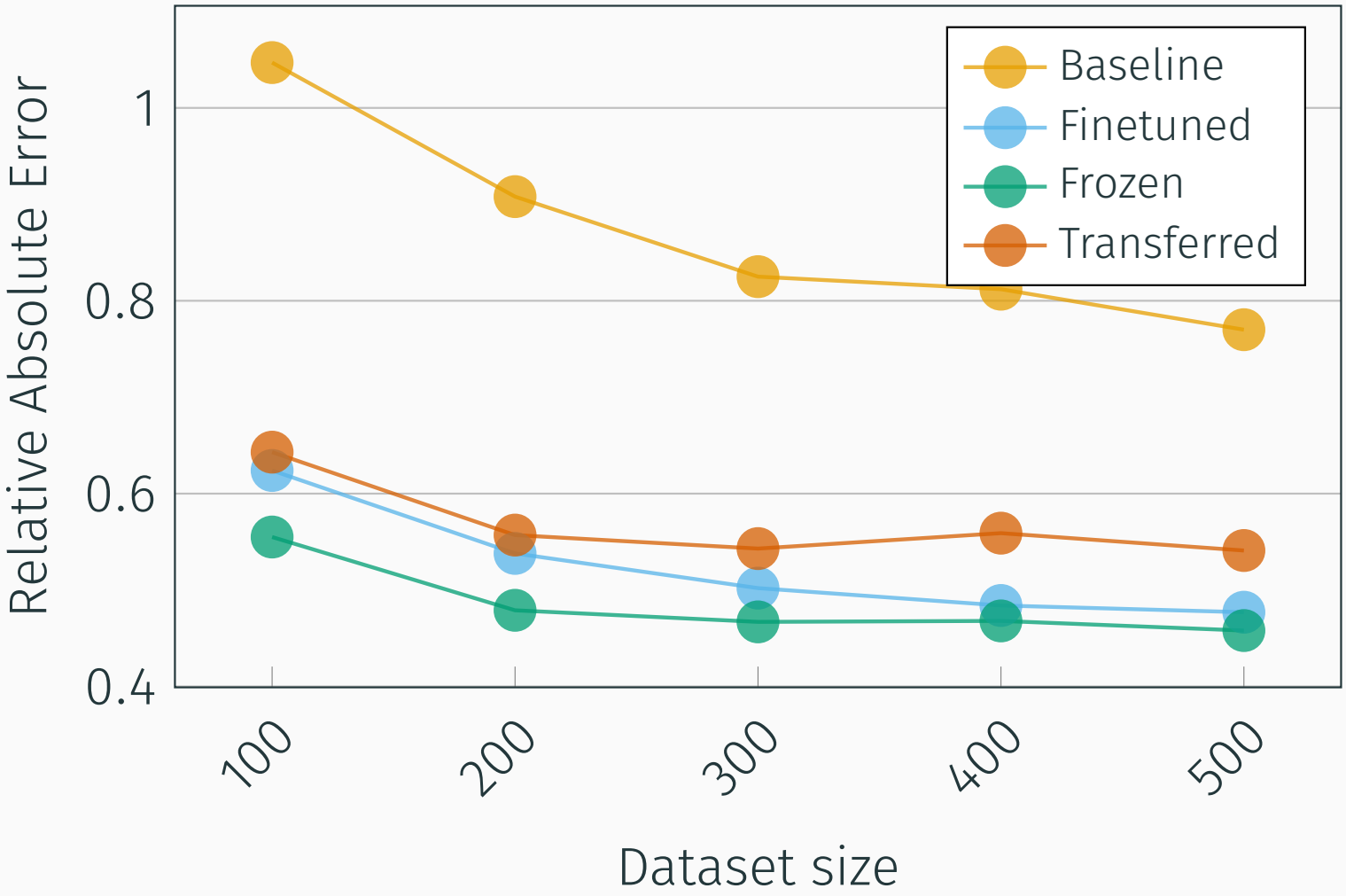
Results


Relative Absolute Error
0.6
0.4
0.2
Baseline
Finetuned
Frozen
Transferred
100
250
500
1,000
2,500
Dataset size

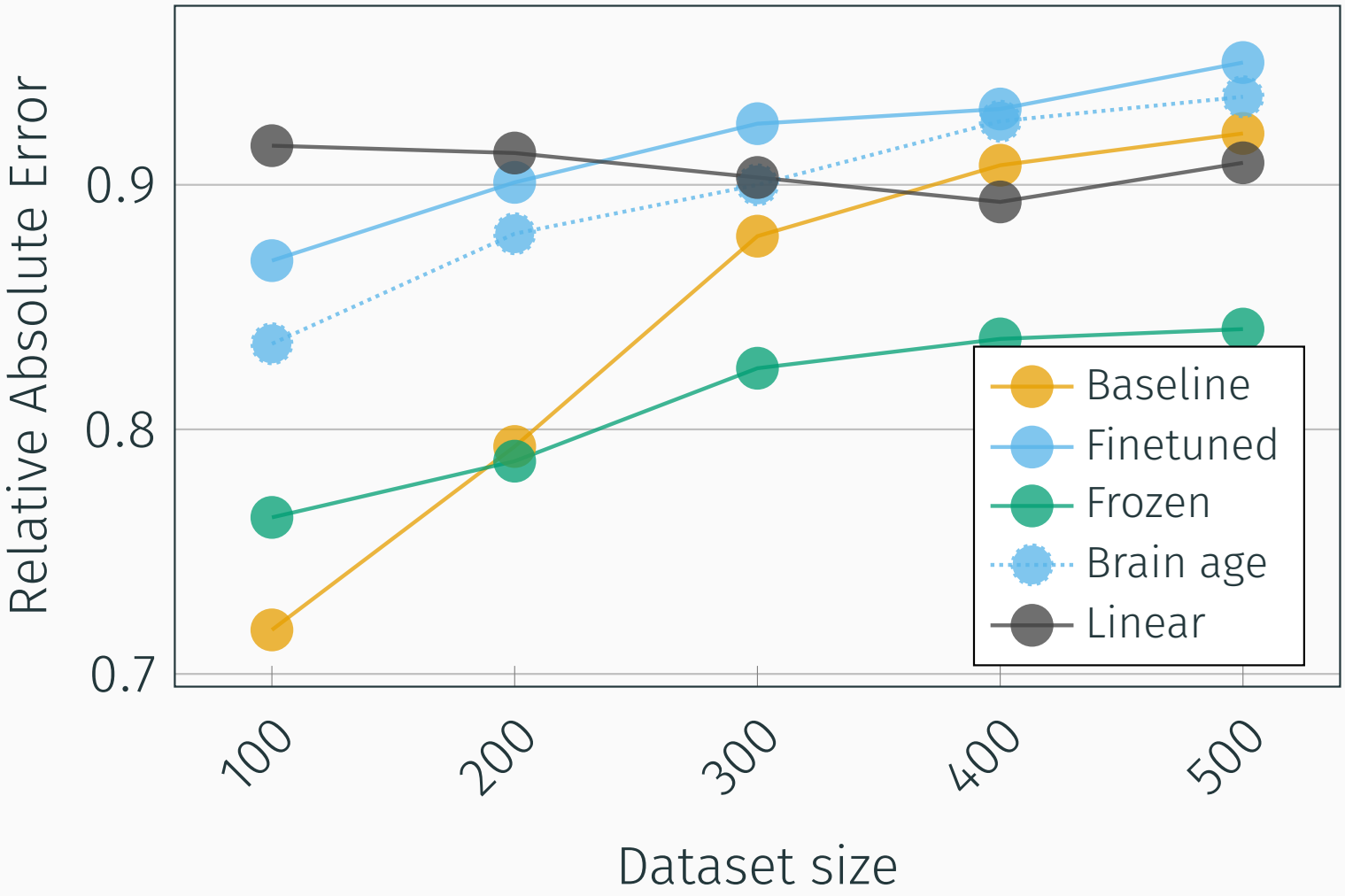
Results

- Each entry is the average over 100 bootstrapped training samples of size n
  - Additional 20% for validation and 20% for testing
  - 9 models with different hyperparameters are fit, the best is selected based on validation performance
- A model trained from scratch (baseline), and three using different transfer learning strategies
  - Transferred: The pretrained model is directly applied
  - Frozen: The pretrained model is retrained while freezing everything but the final layer
  - Finetuned: The pretrained model is used for initialization, then fully retrained

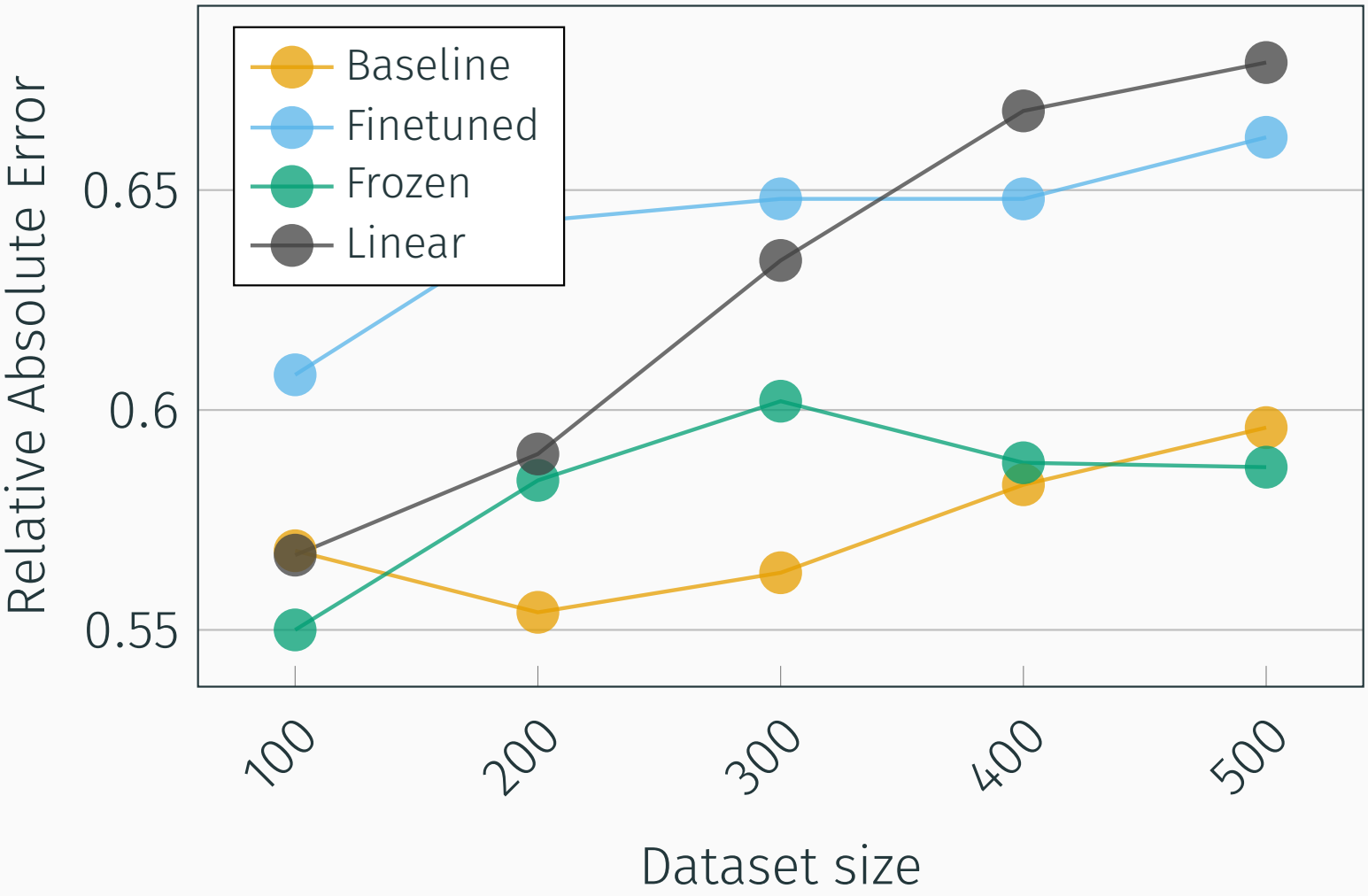
Results


Relative Absolute Error
1
0.8
0.6
0.4
100
200
300
400
500
Dataset size
Baseline
Finetuned
Frozen
Transferred

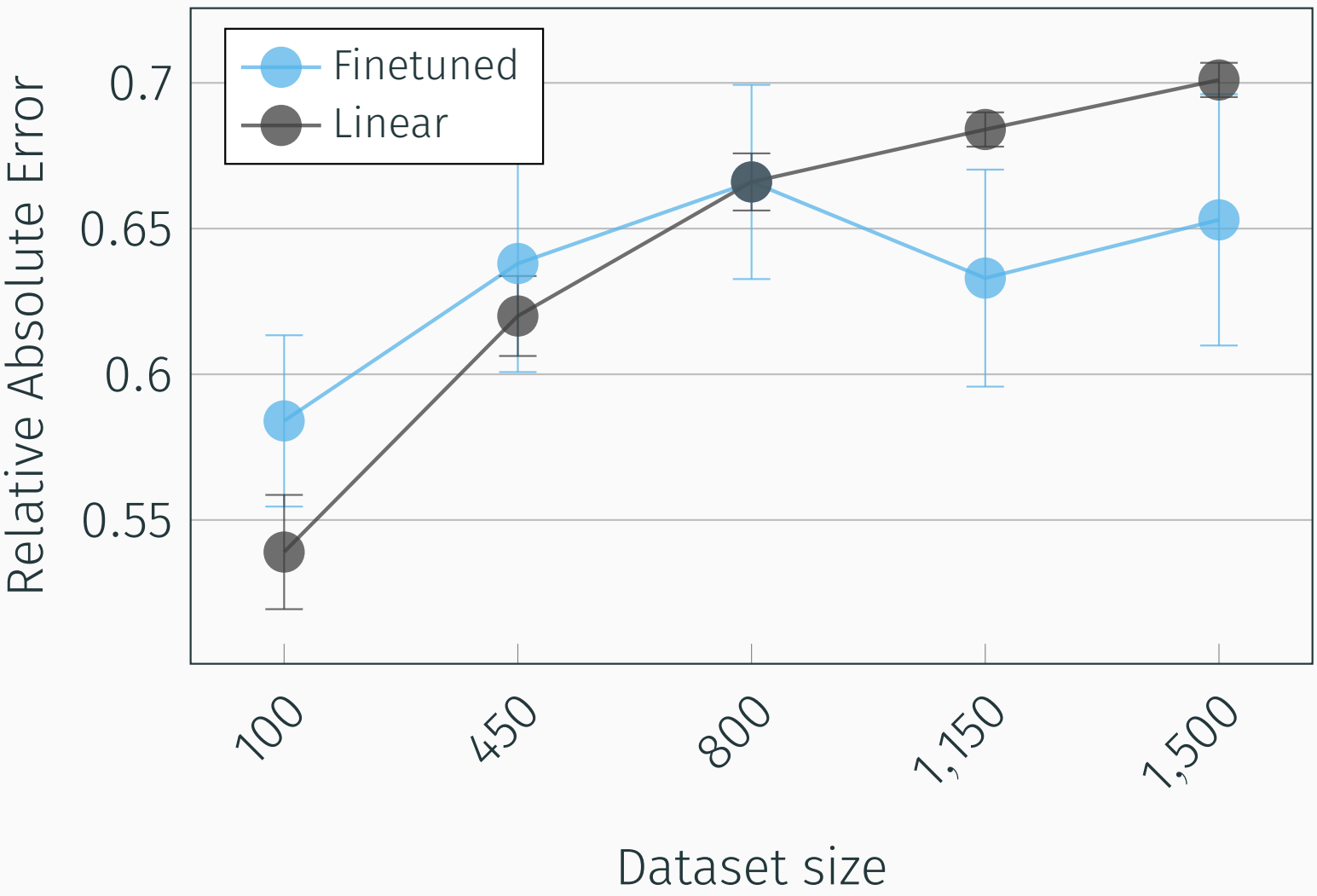
Results


Relative Absolute Error
0.9
0.8
0.7
100
200
300
400
500
Dataset size
Baseline
Finetuned
Frozen
Brain age
Linear

Results


Baseline
Finetuned
Frozen
Linear
Relative Absolute Error
0.65
0.6
0.55
100
200
300
400
500
Dataset size

Results


Finetuned
Linear
Relative Absolute Error
0.7
0.65
0.6
0.55
100
450
800
1,150
1,500
Dataset size

Results

Why does the deep learning model not behave asymptotically?

- It shouldn't (bad luck?)
- Double descent? (🤞)
- Some property of the data? (heterogeneity?)
- Some property of the model? (too simple?)
- **Some property of the procedure?**