

EXTRA 3 — post_mlp: $\text{logit(' algorithm')} - \text{logit(' wizard')}$

