GLM Boost 🦠🙌📈

Our Question:

- Can Low/High h1 HAI responders be predicted from baseline - day 7 data?

Our Process:

- Try various models predicting mucosal, humoral, or cellular responders, and comparing AUC metrics.
- GLM Boost performed well, so we examined theory & verified it as a strong model for our data & question.

Our Model:

- GLM Boost, is a Generalized linear model (GLM), a regression model that can be fit to non gaussian data & optimized via boosting, which is a process that combines a bunch of weak learners to create a strong learner.

Our Results :

- Train AUC: *0.958*
- Predict AUC: *0.825*
- Balanced dataset

Biological Significance:

- Variables of Importance:

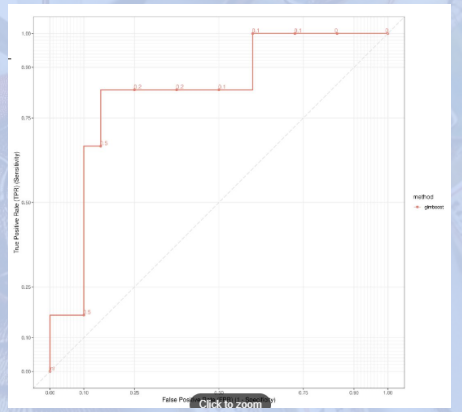
feature	score_percentage
blood_baseline_go.0000187	100
blood_baseline_go.0000240	36
blood_baseline_go.0000211	29
blood_baseline_go.0000221	25
h1_v0_seropositive	21
blood_baseline_go.0000244	21
blood_baseline_go.0000253	3
blood_baseline_go.0000225	0

Genes involved in

- MAPK pathway
- Cell division
- Mitochondrial activity
- Protein Tagging

Seropositivity: baseline antibody levels

- Immunity before vaccination



Predicted AUC	Training AUC	Training Specificity	Training Sensitivity	Training Precision	Training Recall	Training Precisio...	Precision-Recall ...	Details
0.825	0.9583	0.3333	0.975	0.92	0.975	0.3911	0.9676	93 - 47 - 46 | 1 - 8