Supplementary material/Statistical modelling
F#ck in digital networks, Laitinen et al.
Contact: paula.rautionaho@uef.fi

Section 1 provides a summary of the prediction and classification accuracy measures, and Section 2 provides the summary of the GLMM tree results.

1. Prediction and classification accuracy

C-index	Somers dxy	Confusion matrix accuracy
0.77	0.53	62.8%

Confusion Matrix

```
Confusion Matrix and Statistics

          Reference
Prediction  Low High
      Low  1941 1232
      High  876 1611

               Accuracy : 0.6276
                 95% CI : (0.6148, 0.6402)
    No Information Rate : 0.5023
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.2555

 Mcnemar's Test P-Value : 1.059e-14

            Sensitivity : 0.5667
            Specificity : 0.6890
         Pos Pred Value : 0.6478
         Neg Pred Value : 0.6117
             Prevalence : 0.5023
         Detection Rate : 0.2846
   Detection Prevalence : 0.4394
      Balanced Accuracy : 0.6278

       'Positive' Class : High
```

The GLMM tree performs significantly better than a baseline model ($p < 2.2e-16$). 62.8% of all instances are predicted correctly. 68.9% of the Lows and 56.67% of the Highs are predicted correctly. The accuracies for variety-specific predictions are AU 65.5%, UK 55.8 % and US 63.2%.

2. Summary of the GLMM trees

```
> print(gtmedian)

Model formula:
freqmedian ~ 1 | variety + network_size + NSI


Fitted party:
[1] root
|   [2] variety in AU, UK
|   |   [3] variety in AU
|   |   |   [4] NSI <= 0.46
|   |   |   |   [5] network_size <= 13: n = 82
|   |   |   |       (Intercept)
|   |   |   |         -1.417066
|   |   |   |   [6] network_size > 13: n = 1347
|   |   |   |       (Intercept)
|   |   |   |        -0.4483421
|   |   |   [7] NSI > 0.46: n = 411
|   |   |       (Intercept)
|   |   |         -1.215607
|   |   [8] variety in UK: n = 897
|   |       (Intercept)
|   |        -0.2351919
|   [9] variety in US
|   |   [10] NSI <= 0.46
|   |   |   [11] network_size <= 18: n = 387
|   |   |       (Intercept)
|   |   |        0.02584123
|   |   |   [12] network_size > 18: n = 2100
|   |   |       (Intercept)
|   |   |         0.725466
```

| | [13] NSI > 0.46: n = 436

| | (Intercept)

| | -0.1562806

Number of inner nodes: 6

Number of terminal nodes: 7

Number of parameters per node: 1

Objective function (negative log-likelihood): 3672.863