On potential limitations of differential expression analysis with non-linear machine learning models – Supplementary materials

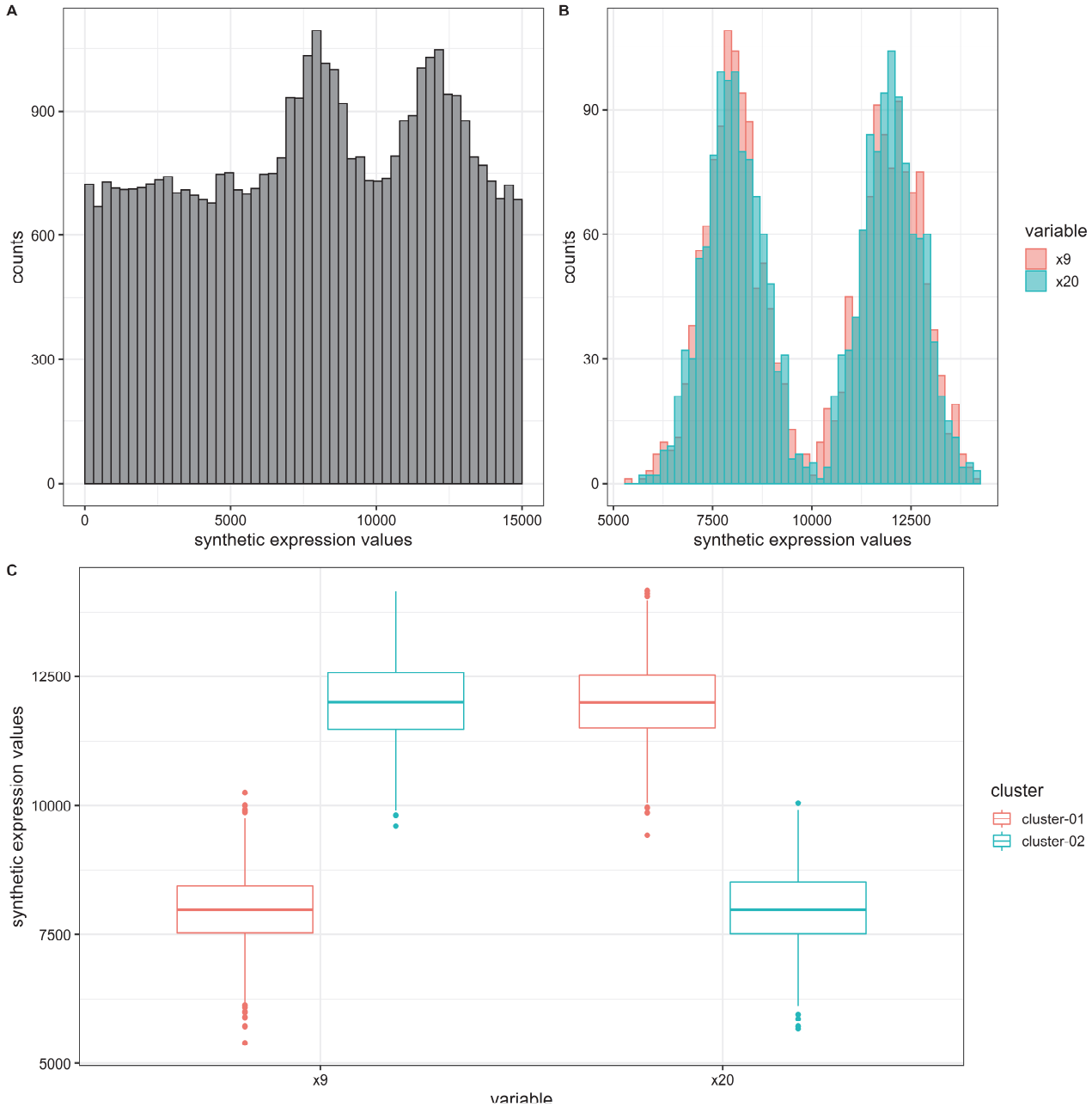
Supplementary data

We built a supplementary dataset, named *(D) big circles (negative binomial)*, similar to dataset *(C) circles (big)* but with a few major differences:

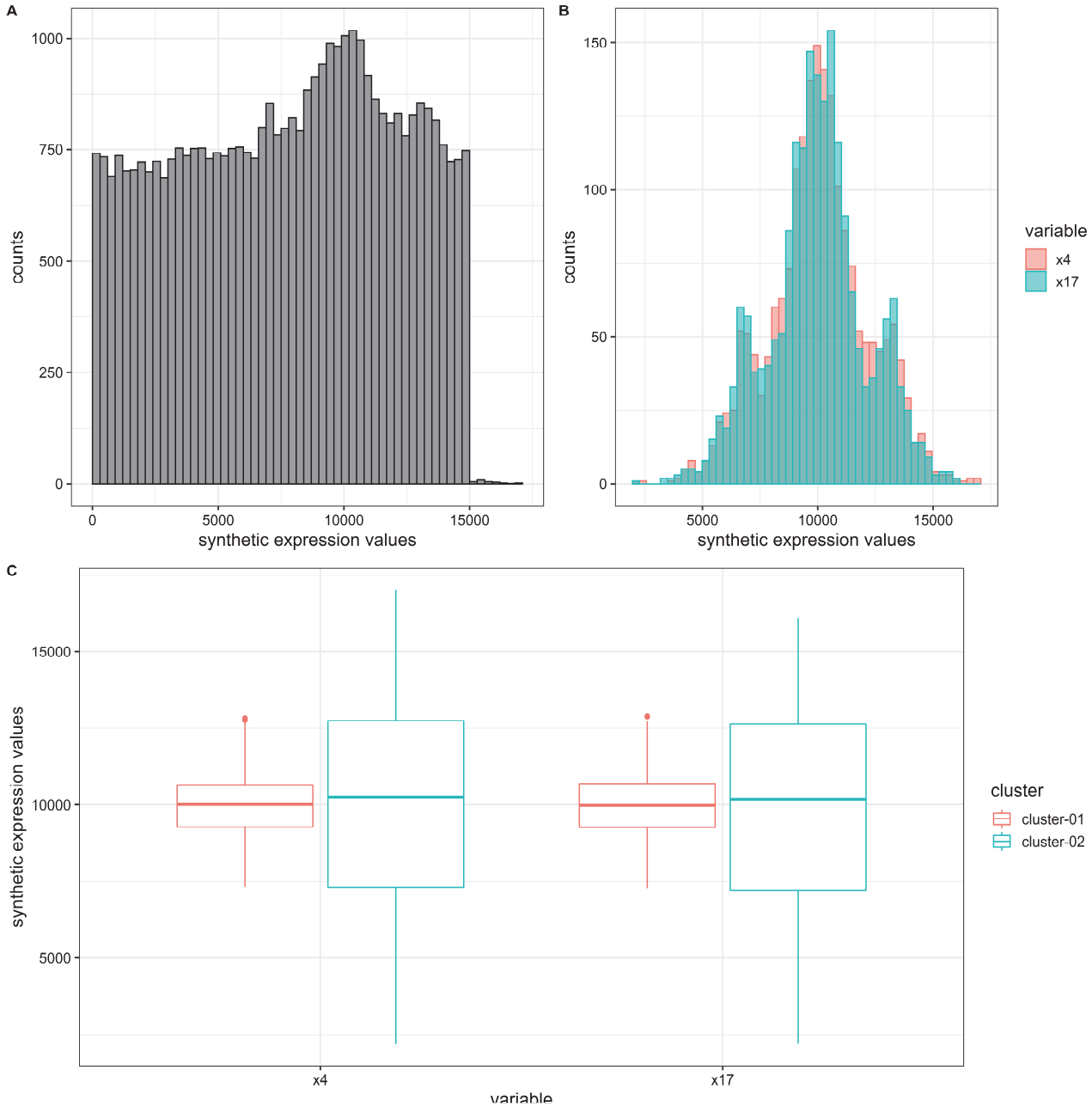
1. The dataset is characterised by 200 significant variables (instead of just 2); the purpose is to show that the XAI solution is not affected by the number of significant variables
2. The values are sampled from negative binomial distributions, mimicking real gene expression data; the purpose is to show that the underlying distribution is not affecting the XAI solution either

The decision boundary is a 200-dimensional hyper-sphere, which generalises the 2-dimensional circle of dataset (C).

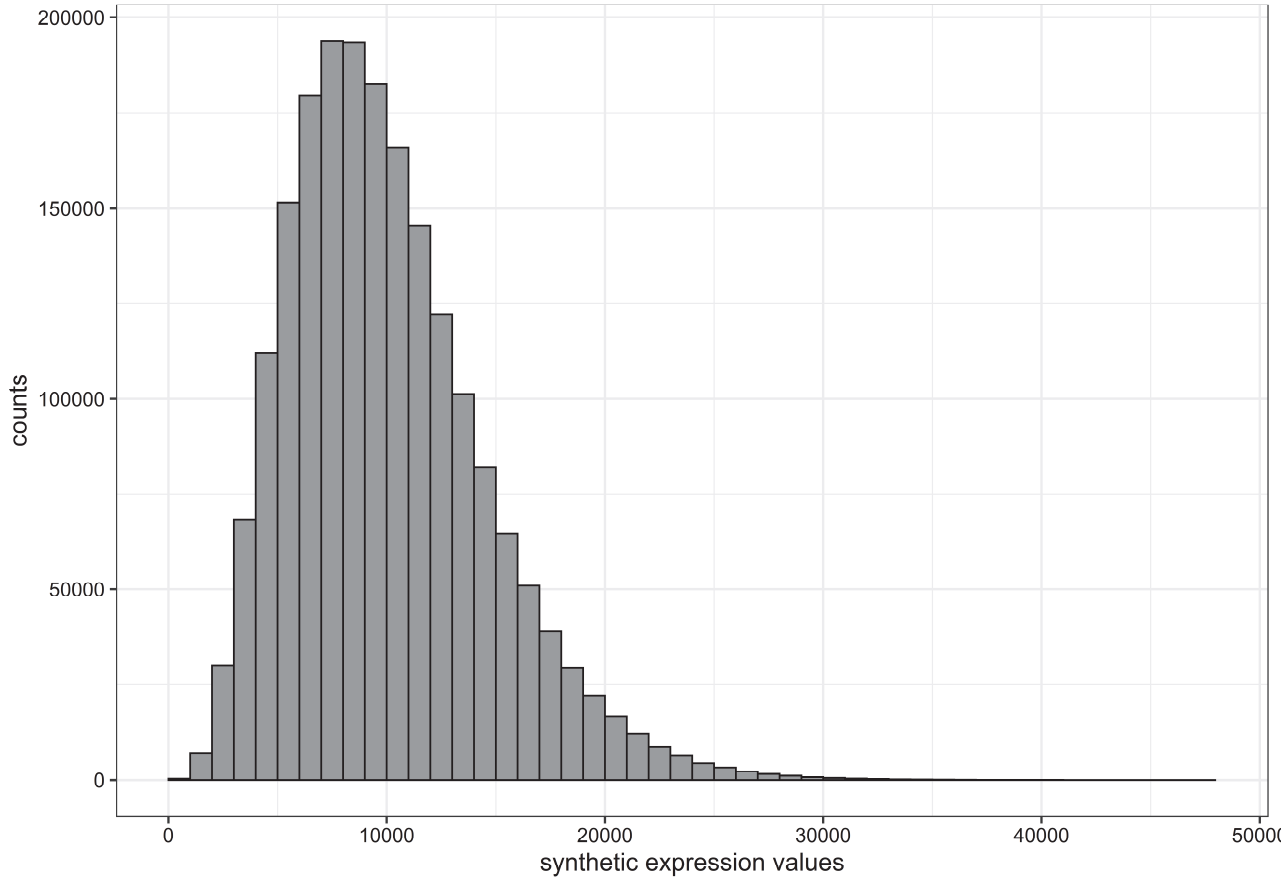
Supplementary figures



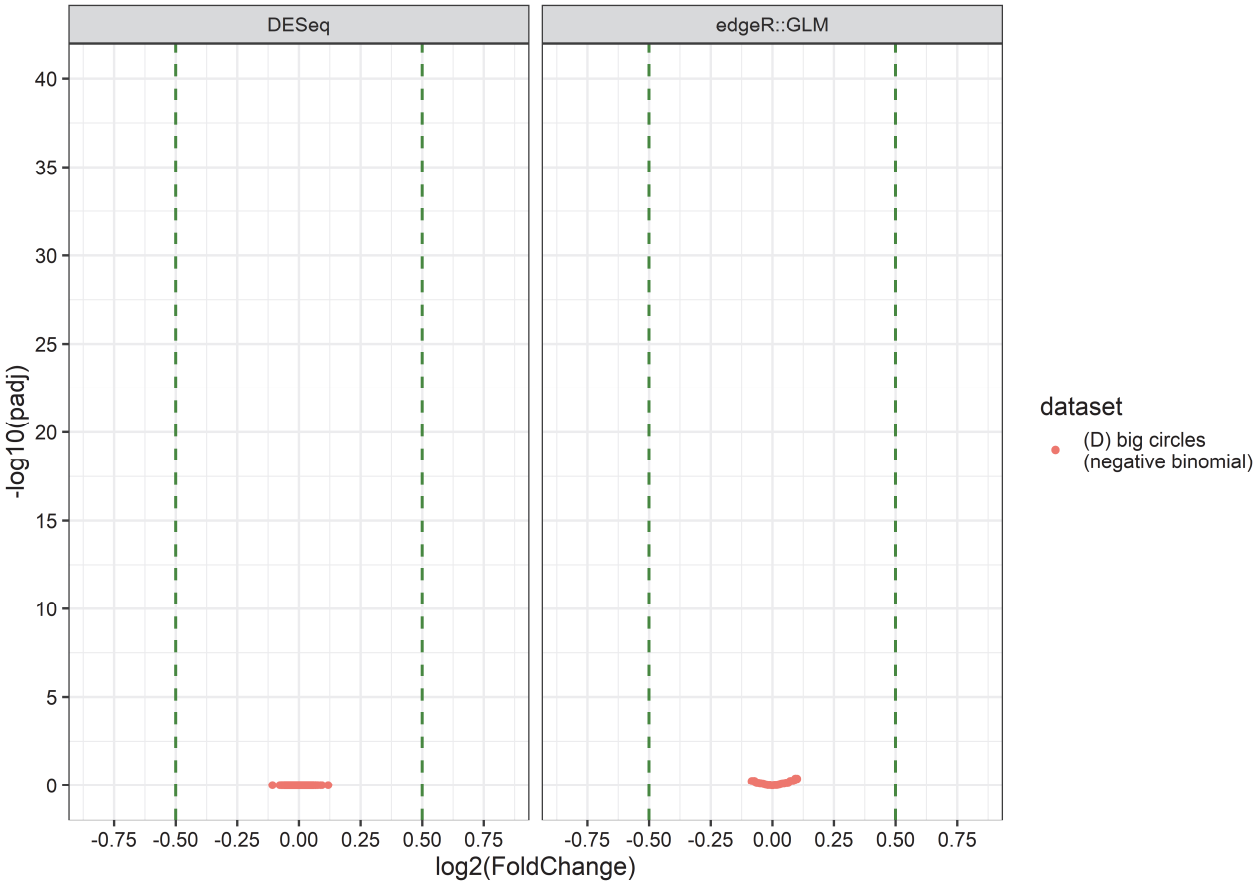
Supplementary Figure 1. Distributions of the synthetic expression values in dataset *(A) clouds*. A) Histogram of all the n=2000 values for all the n=20 features. B) Histogram of the synthetic expression values of the two relevant features x9 (red) and x20 (blue). C) Boxplots of the synthetic expression values of the two relevant features (x9 and x20), in cluster-01 (red) and cluster-02 (blue) respectively.



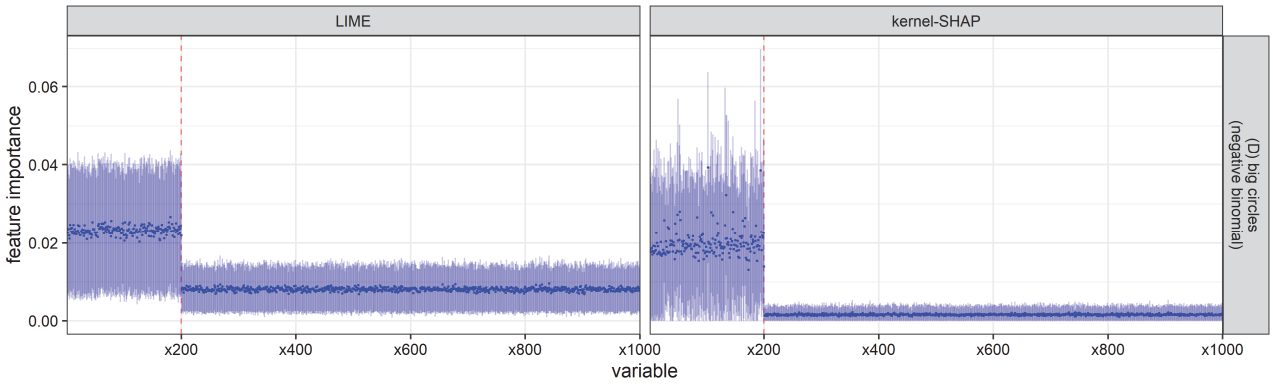
Supplementary Figure 2. Distributions of the synthetic expression values in dataset *(B) circles*. A) Histogram of all the n=2000 values for all the n=20 features. B) Histogram of the synthetic expression values of the two relevant features x4 (red) and x17 (blue). C) Boxplots of the synthetic expression values of the two relevant features (x4 and x17), in cluster-01 (red) and cluster-02 (blue) respectively.



Supplementary Figure 3. Distribution of the synthetic expression values in dataset *(D) big circles (negative binomial)*: histogram of all the n=2000 values for all the n=1000 features.



Supplementary Figure 4. Volcano plot of the DE results on dataset (D) big circles (negative binomial) obtained with DESeq (left) and edgeR-GLM (right) respectively. Adjusted *p*-values have been computed using the Benjamini-Hochberg correction.



Supplementary Figure 5. Feature importance of dataset *(D) big circles (negative binomial)* computed using two different explainers, namely LIME (left column) and kernel-SHAP (right column). Barplots report feature importance (mean and standard deviation) based on local explanations of each sample. For a better visualisation of the emerging pattern, the features have been ordered so that the first 200 are the relevant ones, whereas the remaining 800 are noisy.