Supporting Information (Q.-Y. Zhang and J. Aires-de-Sousa)
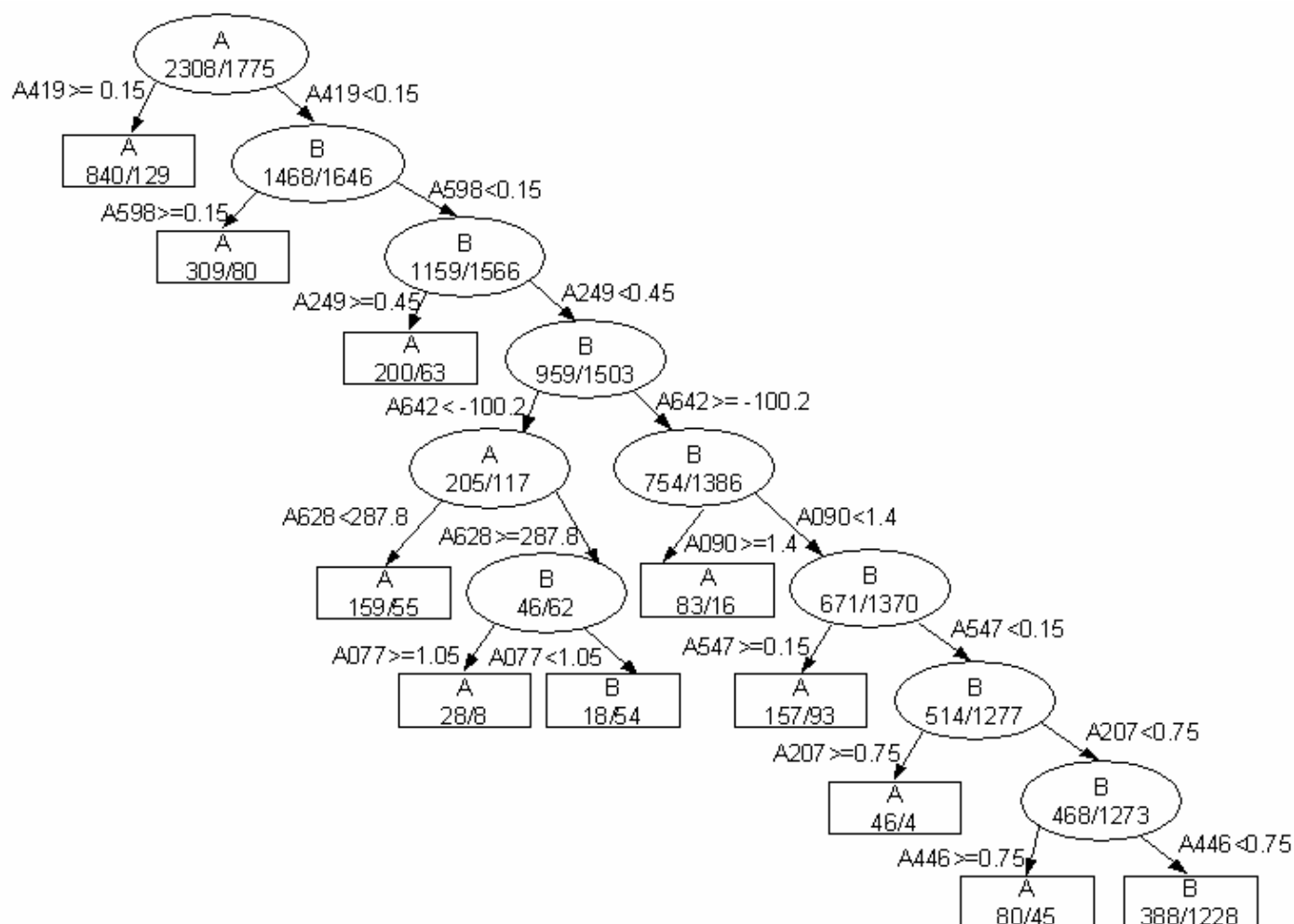
Graphical representation of the classification tree obtained for mutagenicity prediction.

Ovals represent intermediate nodes, and rectangles represent terminal nodes. Inside each node, a classification is given for the node (A: mutagen; B: nonmutagen), as well as the number of mutagens/nonmutagens of the training set falling into it. Descriptor A628 is the molecular weight. Descriptor A642 is the ring strain energy calculated by PETRA. The other descriptors used by the tree are components of the MOLMAP descriptor. The first (A419) and the third (A249) correspond mostly to C-H bonds of polyaromatic compounds, while the second (A598) corresponds to C-Cl bonds, and to C-N bonds in azides.