# Additional file 4: Formulae of assessment metrics

This additional file describe the formulae of assessment metrics to evaluate the prediction performance

$$Sen=1-\frac{N\_{-}^{+}}{N^{+}} , 0\leq Sen\leq 1 (1)$$

$$Spec=1-\frac{N\_{+}^{-}}{N^{-}} , 0\leq Spec\leq 1 (2)$$

$$Acc=1-\frac{N\_{-}^{+}+N\_{+}^{-}}{N^{+}+N^{-}} , 0\leq Acc\leq 1 (3) $$

$MCC= \frac{1-\left(\frac{N\_{-}^{+}}{N^{+}}+\frac{N\_{+}^{-}}{N^{-}}\right)}{\sqrt{\left(1+\frac{N\_{+}^{-}-N\_{-}^{+}}{N^{+}}\right)\left(1+\frac{N\_{-}^{+}-N\_{+}^{-}}{N^{-}}\right)}} , -1\leq MCC\leq 1 (4)$

TP, FP, TN, FN stand for true positive, false positive, true negative, and false negative respectively.

The relations between these symbols and the symbols in Eqs. (1, 2, 3 and 4) are given by:

$$\left\{\begin{matrix}N\_{+}^{-}=FP\\N\_{-}^{+}=FN\\\begin{matrix}N^{+}=TP+N\_{-}^{+}\\N^{-}=TN+N\_{+}^{-}\end{matrix}\end{matrix}\right. \left(5\right)$$

where, $N\_{-}^{+}$ represents the number of positive samples (TNFs) incorrectly predicted to be negative samples (non-TNFs), $N\_{+}^{-}$ represents the number of negative samples incorrectly predicted to be positive sample, $N^{+} $represents the number of positive sample surveyed and $N^{-} $represents the number of negative sample surveyed.