Algorithm 1: Co-Training

Input: Labelled data $\boldsymbol{L}$, unlabelled data $\boldsymbol{U}$, classifiers $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, p (positives), n (negatives), u (number of initial data), k (iterations)

Output: Trained classifiers

Create a subset $\boldsymbol{U'}$ selecting u random instances of $\boldsymbol{U}$

for *k iteraciones*

    Train $\boldsymbol{H}_1$ with $\boldsymbol{L}$ only considering a subset ($\boldsymbol{x}_1$) of the features of each instance (x)

    Train $\boldsymbol{H}_2$ with $\boldsymbol{L}$ only considering a subset ($\boldsymbol{x}_2$) of the features of each instance (x)

    Make $\boldsymbol{H}_1$ predict p positive instances and n negative instances from $\boldsymbol{U'}$ that have the highest confidence

    Make $\boldsymbol{H}_2$ predict p positive instances and n negative instances from $\boldsymbol{U'}$ that have the highest confidence

    Add those selected instances to $\boldsymbol{L}$

    Replenish $\boldsymbol{U'}$ by adding $2p + 2n$ instances from $\boldsymbol{U}$.

endfor

return $\boldsymbol{H}_1, \boldsymbol{H}_2$