Most existing ensemble methods aim to train the underlying embedded models independently and simply aggregate their final outputs via averaging or weighted