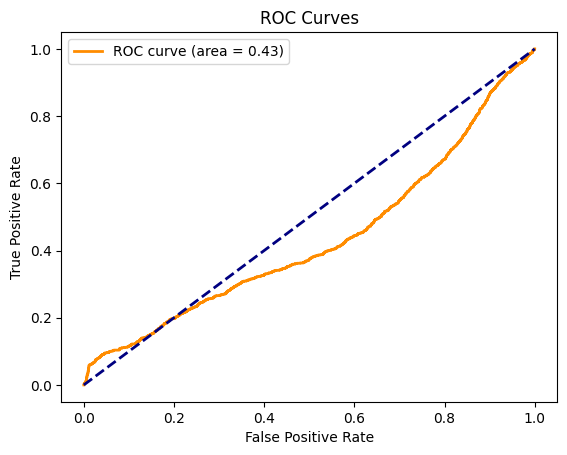
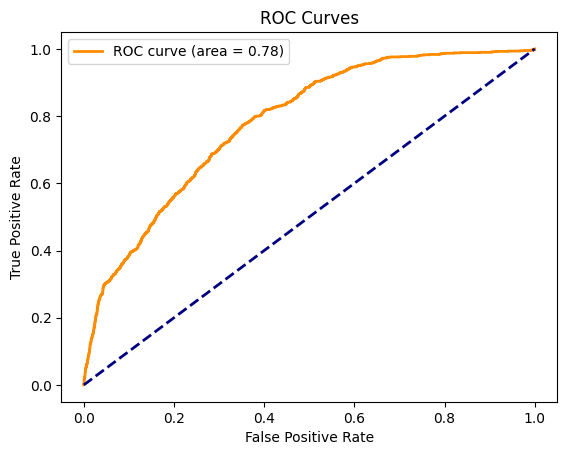
**SEMI-SUPERVISED ENSEMBLE**

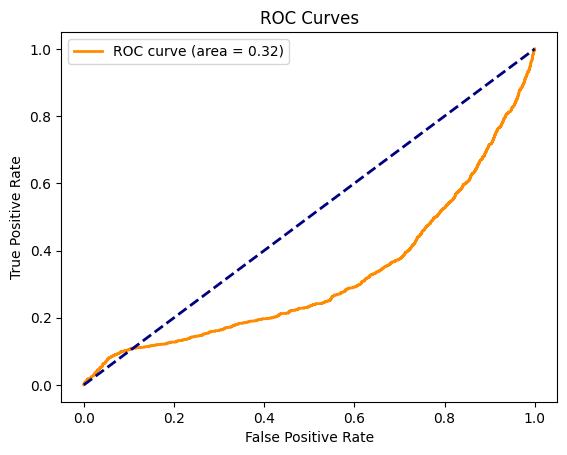
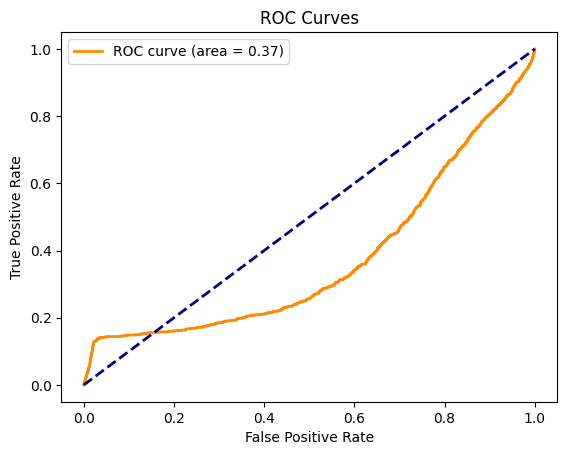
For the semi-supervised ensemble, we used Decision tree classifier and the Support Vector Machine models that we trained with the labelled data. After that we used theses models to generate pseudo labels on the unlabeled data. We used then an ensemble classifier in our case the voting classifier to make prediction on the combined dataset generated with labelled data and pseudo labelled data.

Here are the results we got from it

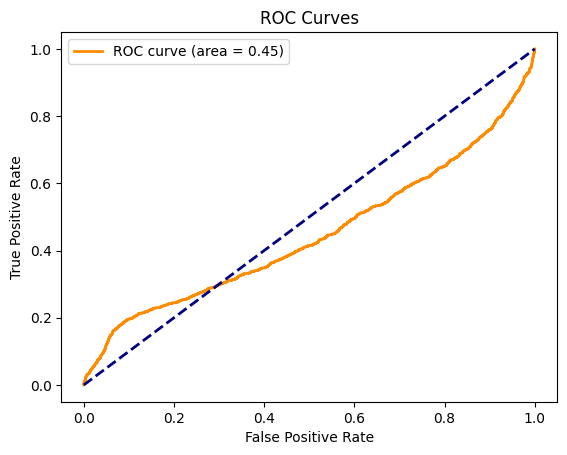
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0.5 | 0.75 | 0.9 | 0.95 | 0.99 |
| Accuracy | 0.8759 | 0.8752 | 0.8762 | 0.878 | 0.8786 |
| F1-score | 0.3461 | 0.1999 | 0.1012 | 0.0433 | 0.01437 |
| Runtime | 98.367 | 39.7415 | 16.596 | 8.9232 | 2.8842 |



50 % of data unlabeled 75 % of data unlabeled



90 % of data unlabeled 95 % of data unlabeled

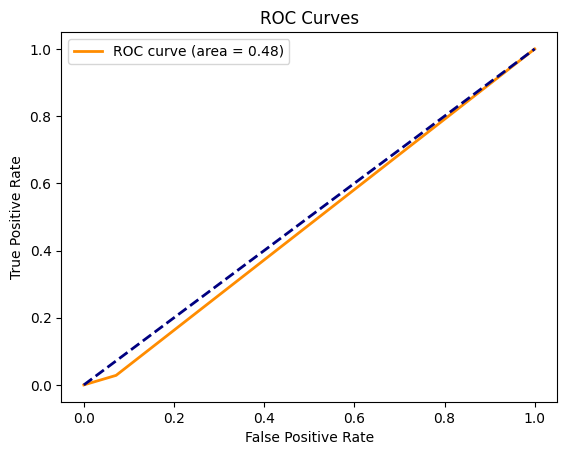
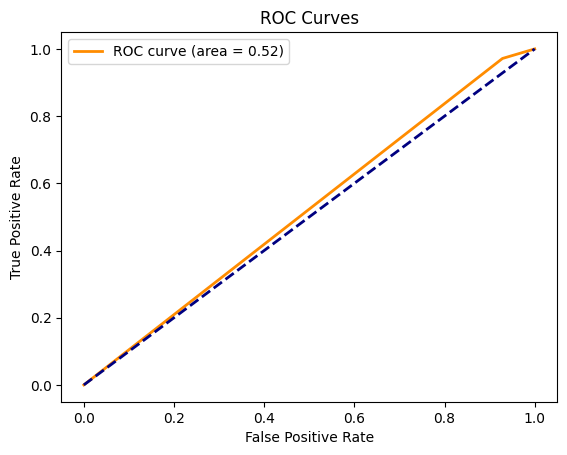


99 % of data unlabeled

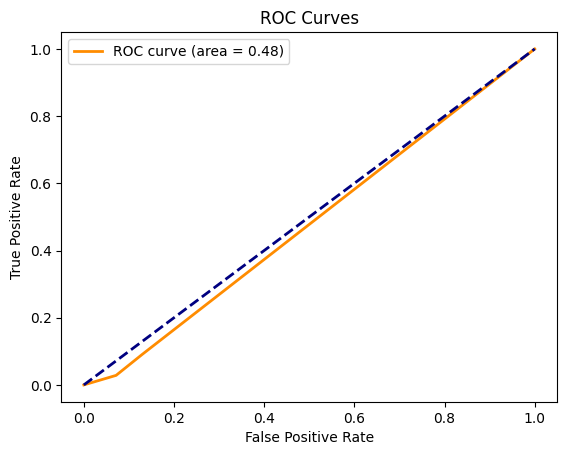
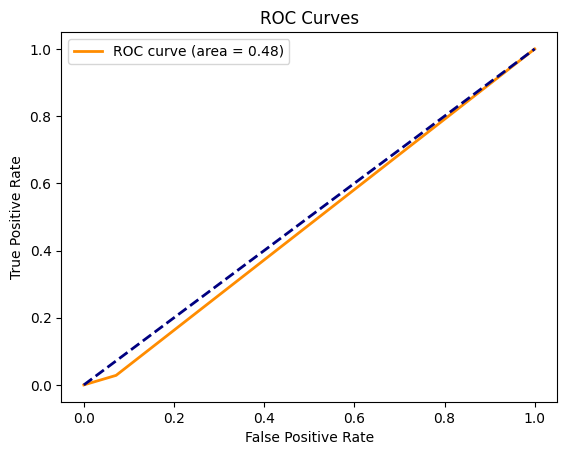
**UNSUPERVISED PRE-TRAINED**

For this question used the unsupervised pre-trained method with the restricted Boltzmann machines (RBMs). We trained it layer by layer using the unlabeled dataset. After that we used the parameter got from this training to create a 2-layers neural network for the classification task with the label data.

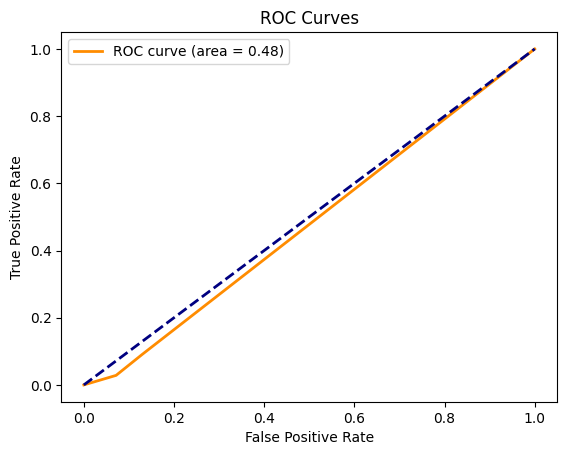
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0.5 | 0.75 | 0.9 | 0.95 | 0.99 |
| Accuracy | 0.8793 | 0.8793 | 0.8793 | 0.8793 | 0.8793 |
| F1-score | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Runtime | 7.281 | 6.40 | 5.844 | 5.657 | 5.523 |



50 % of data unlabeled 75 % of data unlabeled



90 % of data unlabeled 95 % of data unlabeled



99 % of data unlabeled