First article:

-Ensemble learning in machine learning uses finite concrete models to obtain better predictive performance unlike statistical ensemble that use infinite methods.

-In supervised learning the model tries many hypothesis spaces until it find good one with good predictions but it is hard to find good one for every problem so we use ensemble learning to combine many hypotheses to find hopefully better one.

-ensemble learning can be used to increase the performance of weak learning models by combining them with other models and it can be also more efficient than giving more data to one model.

-random first is one of the examples of ensemble learning which use many decision trees to get better predictions.

-it can be used in unsupervised learning .

-Theory : Ensemble learning gives better performance when there is a diversity between the models.

-Ensemble size: even though the ensemble size and the number of models is important and have great impact on the final predictions but there is limited studies about this problem but there is a concept called “the law of diminishing returns in ensemble construction” states that as you add more models at some point this will not be efficient but also may lead to reducing the performance of the predictions.

– Types of ensemble learning:

-Baye’s optimal classifier : it ensembles all the hypothesis in the hypothesis space and

It’s formula is : 

-Bagging : it ensembles on bootstrapped data (subsets) it done by voting of the ensemble members to get the prediction and random forest is example of this method.

-Boosting: training successively models on the data and as every model train on the data the wrong predictions get greater weight than the correct prediction so the model after it learn from it and so on it sometimes gives better accuracy than bagging but it leads to overfitting.

-Bayesian model averaging: it get the prediction by averaging the models predictions weighted by their posterior probabilities that the efficient ones get greater weight than the inefficient and weak models.

-Bayesian model combination : it i the algorithmic correction of the averaging as it distribute the weight on the models unlike BMA that tends to converge the weight towards the closest model and it tends to choose the best combination of the models but it is more expensive than BMA.

-Bucket of models : it makes all the models train on all the problems and choose for each problem the best model so it can give better performance by the combination of the models and giving each one the task that he can do the best.

-stacking : it train models and then use there predictions to help the final model to predict more efficient results and avoid overfitting.

Second article :

Ensemble techniques:

-Max voting: it is done by training each model on the data and then get the most occurred predict as the final one , important code part:

from sklearn.ensemble import VotingClassifier

VotingClassifier(m1,m2, voting=”hard”)

-Averaging: it sums all the data an dividing them by the num of models you are using

Coding part :

Finalprediction = (p1+p2+p3+...+pn)/n

-Weighted average: it is extension of averaging but it gives each model weight according to their priority and efficiency , coding part:

Finalpred = (p1\*w1 + p2\*w2 + … + pn\*wn) where w is between 0 and 1

-Stacking : is to train models on the data that splitied into small subsets and then take their predictions for another mother model so it can give us final prediction.

-Blending: the train data is splitted into two parts the train and validation the models are trained and then predict the validation and test and then takes their predictions and ude the to make the final model get more efficient predictions.

-Bagging : is to split the training data into subsets and train models on the subsets and the combine all their predictions to get final predictions.

-Boosting : it is sequential where each model learn from the other model error and try to make it better.