

**TP555 – Inteligência Artificial e Machine Learning**  
**Josino Villela da Silva Neto – Matrícula: 854 – Mestrado em Engenharia de Telecomunicações**

**Lista de Exercícios #8**

**1.** Sim. Os cinco diferentes modelos podem ser combinados com a finalidade de se obter melhores resultados. A técnica que faz uso dessa combinação de modelos/preditores, é chamada de Ensemble Learning. Dentre alguns métodos Ensemble, pode-se mencionar: bagging, boosting, stacking e etc.

**2.** Diferentes modelos de classificadores realizam diferentes predições. Dadas estas predições, a classe que foi predita pelo maior número de classificadores é dada como a que obteve a maioria dos votos. Sendo assim, um classificador de votação rígida, é aquele que considera a classe que obteve a maioria dos votos para realizar sua predição. Em outras palavras, um classificador de votação rígida é aquele que escolhe a classe que obteve a maioria dos votos.

Quando os modelos utilizados são capazes de estimarem as probabilidades de cada classe, estas podem ser utilizadas para realizar a predição, com base na classe de maior chance de ocorrência. O classificador que utiliza estes dados (probabilidades de cada classe), é chamado de classificador de votação suave.

Resumidamente, os classificadores de votação rígida utilizam as predições realizadas pelos diferentes modelos, ao passo que os classificadores de votação suave utilizam as probabilidades de ocorrência de cada classe fornecidas pelos diferentes modelos.

**3.** Bagging Ensemble, Pasting Ensemble e Florestas Aleatórias (por serem treinadas através de métodos Bagging e as vezes Pasting), podem ser implementados em diferentes servidores e treinados em paralelo, objetivando acelerar o treinamento. São modelos populares pela sua escalabilidade.

**4.** O benefício da avaliação out-of-bag está no fato de que com ela, não são necessários conjuntos de validação ou validação cruzada para a avaliação do Ensemble (conjunto/combinação de modelos).

**5.** As árvores extras são mais aleatórias que as florestas aleatórias, pois utilizam limiares aleatórios para cada feature, ao invés de buscarem pelos melhores valores, como ocorre com as florestas aleatórias. Esta aleatoriedade pode ajudar, pois as árvores extras são mais rápidas de serem treinadas.

**6.** O segundo modelo (ensemble com um conjunto de árvores de decisão), generalizará melhor. É possível observar isto através das fronteiras de decisão (contornos mais adequados, sem indícios de sobreajuste) e também através da área abaixo da curva ROC (ensemble possui maior área e consequentemente maior precisão).