**Desafio 02 – Resumo de Artigo Considerado Essencial no Âmbito da Computação**

Para a entrega do segundo desafio proposto na aula de Processamento de Linguagem Natural, foi definido o artigo *“The Strength of Weak Learning”* proposto por Robert E. Schapire em 1990 dentro do Laboratório de Ciência da Computação do MIT.

Um dos principais motivos para a escolha desse artigo é a apresentação de algo realmente revolucionário dentro do universo de Aprendizado de Máquina e Inteligência Artificial: a capacidade de aumentar a performance de modelos (ou saídas funções hipótese) de algoritmos conhecidos como *weak learning*.

Em um problema envolvendo o treinamento de um algoritmo para a construção de função hipótese que gere, como saída, a probabilidade de instâncias pertencerem a determinada classe, a missão de um modelo é dada pela determinação da probabilidade mais assertiva possível de que cada instância pertença a classe correta. Um modelo considerado de forte aprendizado *(strongly learnable)* apresenta baixos erros de predição com alta confiabilidade. Um modelo de fraco aprendizado *(weak learnable)* apresenta resultados levemente melhores que uma abordagem randômica. A ideia do artigo é mostrar que esses algoritmos de fraco aprendizado podem também ser uma ferramenta poderosa, desde que trabalhados corretamente.

Essa ideia pode ser considerada a base para modelos de *bagging, pasting* ou *boosting*, uma vez que estes são algoritmos que combinam diversos outros algoritmos para combinar as predições e gerar um modelo ainda melhor. Utilizando como referência o livro *HandsOn Machine Learning with Scikit-Learn & TensorFlow* de *Aurélien Géron*, é possível identificar algumas definições destes modelos combinados.

Suponha que você faça uma pergunta complexa para milhares de pessoas aleatórias. Após isso, você agrega todas as respostas. Talvez você perceba que a resposta agregada vale mais que a resposta de um único expert. A isto dá-se o nome de Wisdom of the Crowd (sabedoria do povo). Analogamente, se você agregar as predições de um grupo de preditores, provavelmente isto irá te levar a predições melhores que o melhor preditor analisado de forma individual. Um grupo de preditores é chamado ensemble e, a esta técnica, dá-se o nome de Ensemble Learning. Um algoritmo de Ensemble Learning é chamado Ensemble Method.

Por exemplo, é possível treinar um grupo de classificadores Decision Trees, cada um em um subset aleatório do training set. Para realizar predições, basta obter as predições de cada um dos classificadores individualmente e então predizer a classe que ganhar mais votos. Um conjunto de Decision Trees é chamado de Random Forest e, apesar da aparente simplicidade, é um dos algoritmos mais poderosos de Machine Learning disponíveis hoje.

Os modelos Ensemble são geralmente discutidos no final de cada projeto, uma vez que a solução já foi construída e já foram testados e levantados bons modelos para solução do problema. Neste ponto, possível realizar combinações destes modelos com o objetivo de criar um modelo ainda mais poderosos para aquele problema em questão. Soluções vencedoras em competições de Machine Learning geralmente envolvem métodos Ensemble.