**NAIVE BAYES**

O algoritmo Naive Bayes é um classificador probabilístico que se baseia no Teorema de Bayes, que é uma fórmula matemática que calcula a probabilidade de um evento ocorrer com base em evidências ou condições prévias. O Teorema de Bayes é usado para atualizar a probabilidade de um evento à medida que novas evidências são apresentadas. No contexto do algoritmo Naive Bayes, o Teorema de Bayes é usado para calcular a probabilidade de um dado pertencer a uma determinada classe, com base nas características do dado e nas probabilidades a priori das classes.

O Naive Bayes assume que todas as características são independentes umas das outras, dado a classe, por esse motivo é chamado de ingénuo ou naive em inglês. Isso significa que a presença de uma característica não influencia a presença de outra característica. Essa suposição é frequentemente irrealista, pois, em muitos casos, as características podem estar correlacionadas.

**Desempenho Prático**: Apesar da suposição ingênua, o Naive Bayes frequentemente apresenta um desempenho surpreendentemente bom em muitos problemas práticos, como filtragem de spam e análise de sentimentos. Isso se deve ao fato de que, mesmo que as características não sejam totalmente independentes, a simplicidade do modelo pode levar a boas generalizações.

**Prós do Naive Bayes**

1. **Simplicidade e Facilidade de Implementação**:
   * O algoritmo é fácil de entender e implementar, tornando-o uma boa escolha para iniciantes em aprendizado de máquina.
2. **Rapidez**:
   * O Naive Bayes é muito rápido tanto no treinamento quanto na previsão, o que o torna adequado para grandes conjuntos de dados e aplicações em tempo real.
3. **Desempenho em Alta Dimensionalidade**:
   * Funciona bem em problemas de alta dimensionalidade, como classificação de texto, onde o número de características (palavras) pode ser muito maior do que o número de amostras.
4. **Menos Dados Necessários**:
   * O algoritmo pode funcionar bem mesmo com um número relativamente pequeno de dados de treinamento, especialmente em comparação com modelos mais complexos.

**Contras do Naive Bayes**

1. **Suposição de Independência**:
   * A principal limitação do Naive Bayes é a suposição de que todas as características são independentes entre si, o que raramente é verdade na prática. Isso pode levar a previsões imprecisas quando as características estão correlacionadas.
2. **Desempenho em Dados Não-Normais**:
   * O Gaussian Naive Bayes, em particular, assume que as características seguem uma distribuição normal. Se os dados não seguirem essa distribuição, o desempenho do modelo pode ser comprometido.