هذا المشروع يهدف إلى **تصنيف أورام الثدي** إلى فئتين رئيسيتين:

* **حميد (Benign)**
* **خبيث (Malignant)**

وذلك بالاعتماد على بيانات **Breast Cancer Dataset** المدمجة في مكتبة **Scikit-learn**، والتي تحتوي على 569 عينة و30 خاصية إحصائية تصف شكل الخلايا (مثل نصف القطر، الملمس، المحيط…).

المشروع يستخدم خوارزمية **Naive Bayes (GaussianNB)** وهي إحدى خوارزميات التعلم الآلي البسيطة والفعالة، والتي تعتمد على **نظرية بايز** وتفترض استقلالية الخصائص. تم اختيارها لأنها سريعة في التدريب وتحقق نتائج جيدة مع البيانات الطبية.

**خطوات العمل في المشروع:**

1. تحميل بيانات سرطان الثدي من مكتبة Scikit-learn.
2. استكشاف البيانات (الخصائص والفئات المستهدفة).
3. تقسيم البيانات إلى مجموعة تدريب (67%) واختبار (33%).
4. تدريب النموذج باستخدام **Gaussian Naive Bayes**.
5. تقييم النموذج باستخدام **مقياس الدقة (Accuracy Score)**.
6. تجربة التنبؤ على عينات جديدة للتحقق من أداء النموذج.

**النتائج:**

* النموذج أعطى دقة جيدة في التنبؤ بنوع الورم.
* يمكنه تصنيف العينات الجديدة بشكل سريع وموثوق.

**مستقبل المشروع:**

* تحسين النموذج باستخدام خوارزميات أخرى مثل **SVM** أو **Random Forest**.
* تجربة تقنيات **اختيار الخصائص (Feature Selection)** لتحسين الأداء.
* توسيع المشروع ليشمل بيانات أكبر وأكثر تنوعًا.

هذا المشروع يوضح كيف يمكن للتعلم الآلي أن يُسهم في المجال الطبي، خصوصًا في **التشخيص المبكر لمرض السرطان** مما يساعد على إنقاذ الأرواح.