**3.2.1. Phân tích hồi quy tuyến tính đơn biến (Linear Regression with One Variable)**

Hồi quy tuyến tính đơn biến là một phương pháp thống kê để dự đoán giá trị của một biến phụ thuộc (y) dựa trên giá trị của một biến độc lập (x). Phương trình hồi quy tuyến tính có dạng: y=mx+b*y*=*mx*+*b*, trong đó:

* ( y ) là biến phụ thuộc.
* ( x ) là biến độc lập.
* ( m ) là hệ số góc (slope).
* ( b ) là hằng số (intercept).

**3.2.2. Phân tích hồi quy đa thức (Polynomial Regression)**

Hồi quy đa thức là một dạng mở rộng của hồi quy tuyến tính, trong đó mối quan hệ giữa biến độc lập và biến phụ thuộc được mô hình hóa như một đa thức bậc ( n ). Phương trình có dạng: y=anxn+an−1xn−1+...+a1x+a0*y*=*an*​*xn*+*an*−1​*xn*−1+...+*a*1​*x*+*a*0​.

**3.2.3. Phân tích sử dụng Cây quyết định (Decision Tree)**

Cây quyết định là một mô hình dự đoán sử dụng cấu trúc cây để biểu diễn các quyết định và hậu quả của chúng. Mỗi nút trong cây đại diện cho một thuộc tính, mỗi nhánh đại diện cho một giá trị của thuộc tính đó, và mỗi lá đại diện cho một kết quả dự đoán.

**3.2.4. Phân tích sử dụng Rừng ngẫu nhiên (Random Forest)**

Rừng ngẫu nhiên là một tập hợp của nhiều cây quyết định, được huấn luyện trên các tập con khác nhau của dữ liệu. Kết quả dự đoán cuối cùng được xác định bằng cách lấy trung bình (đối với hồi quy) hoặc bỏ phiếu đa số (đối với phân loại) từ các cây trong rừng.

**3.2.5. Phân tích sử dụng K láng giềng gần nhất (K-Nearest Neighbors)**

K láng giềng gần nhất là một thuật toán không tham số dùng để phân loại hoặc hồi quy. Đối với một điểm dữ liệu mới, thuật toán sẽ tìm ( K ) điểm dữ liệu gần nhất trong tập huấn luyện và dự đoán giá trị dựa trên các điểm này.

**3.2.6. Phân tích sử dụng Máy vector hỗ trợ (Support Vector Regression)**

Máy vector hỗ trợ là một thuật toán học máy sử dụng các siêu phẳng trong không gian nhiều chiều để phân loại hoặc hồi quy. Đối với hồi quy, mục tiêu là tìm một siêu phẳng sao cho các điểm dữ liệu nằm trong một khoảng cách nhất định từ siêu phẳng đó.