AI Predictive Analytics Formula in Healthcare

In predictive analytics for healthcare, a common model might predict health outcomes Y based on several input variables (features) such as patient history, environmental factors, and clinical data.

$$Y = f(X_1, X_2, \ldots, X_n) + \epsilon$$

Where:

- Y represents the predicted health outcome (e.g., risk of disease, recovery time).

- $X_1, X_2, \ldots, X_n$ represent the input features, such as:

    - X_1: Patient medical history (e.g., past illnesses).
    - X_2: Environmental factors (e.g., pollution, living conditions).
    - X_3: Lifestyle (e.g., smoking, diet).
    - X_4: Clinical data (e.g., blood pressure, cholesterol).

- $f(\cdot)$ is the machine learning model function (e.g., linear regression, decision trees, neural networks).

- ϵ is the error term, accounting for noise or unmeasured factors.

The model seeks to minimize the error between actual outcomes Y and predicted outcomes $\hat{Y}$:

$$\min_{\theta} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Where θ represents the model parameters to be optimized.