**Literature review - Methods to Convert Regression Outputs to Binary Values**

Converting regression outputs to binary values for applying DDM (Drift Detection Method) involves transforming continuous regression predictions or errors into a format suitable for binary classification. This is necessary because DDM is designed to detect concept drift based on binary error rates (correct vs. incorrect classifications).

**Methods to Convert Regression Outputs to Binary Values**

1. **Thresholding on Error**:
   * **Description**: Calculate the error between the predicted values and the actual values (e.g., Mean Absolute Error (MAE) or Mean Squared Error (MSE)). Define a threshold; if the error is above this threshold, it is considered an "error" (1); otherwise, it is a "correct" prediction (0).
   * **Application**:
     + **Absolute Error**: error = |y\_true - y\_pred|
     + **Threshold**: If error > threshold, assign 1 (error); otherwise, assign 0 (no error).
   * **Choosing Threshold**: The threshold can be selected based on domain knowledge, statistical properties of the errors (like setting it to a certain percentile of error distribution), or experimentally by evaluating drift detection sensitivity.
2. **Quantile-based Thresholding**:
   * **Description**: Use quantiles to set dynamic thresholds based on the distribution of errors. For example, if errors above the 95th percentile are considered significant, this threshold can dynamically adjust to the data's variability.
   * **Application**:
     + Compute error distribution quantiles.
     + Classify errors above a chosen quantile (e.g., 95th) as 1 (error) and those below as 0 (no error).
3. **Relative Change or Percentage Error**:
   * **Description**: Instead of absolute errors, use relative errors (percentage difference between predicted and actual values). Set a threshold on this relative error to determine if a prediction is "correct" or "incorrect."
   * **Application**:
     + **Relative Error**: relative\_error = |(y\_true - y\_pred) / y\_true|
     + **Threshold**: Similar to absolute error, apply a threshold to the relative error.
4. **Regression to Binary Classification**:
   * **Description**: Train a separate binary classifier to predict if the regression model's prediction is within an acceptable range. This meta-model uses regression outputs and actual values to classify predictions as correct or incorrect.
   * **Application**:
     + Use original regression outputs and true values to train a binary classifier.
     + The classifier outputs 1 for errors and 0 for correct predictions based on some criteria (e.g., error margins).

***Further Reading :***

1.Handling Concept Drifts in Regression Problems – the Error Intersection Approach Lucas Baier1 , Marcel Hofmann2 , Niklas Kühl1 , Marisa Mohr2,3 and Gerhard Satzger

* In this paper it is mentioned that EDDM can only be applied to classification problems but don’t discuss how.