Multiclass Logistic Regression: One-vs-Rest (OvR) vs Softmax

Logistic Regression is mainly used for **binary classification**, but it can be extended to **multiclass classification** using:

1. **One-vs-Rest (OvR)**: Trains **one logistic model per class**, treating it as "one vs all others."
2. **Softmax Regression (Multinomial Logistic Regression)**: Uses the **Softmax function** to assign probabilities across multiple classes in a single model.

1️⃣ One-vs-Rest (OvR) Approach

- For **each class**, a separate **binary logistic regression** model is trained.
- At prediction time, the model with the **highest probability** is selected.
- Good for **small datasets** but less efficient for large numbers of classes.

🔹 **Scikit-Learn handles this automatically using multi_class='ovr'**.

2 Softmax Regression (Multinomial)

- Uses a **single model** with the **Softmax function**:

$$P(y = j|x) = \frac{e^{z_j}}{\sum_k e^{z_k}}$$

- Converts raw scores (z_j) into **probabilities**.
- More computationally efficient for **many classes**.
- Implemented using `multi_class='multinomial'` with `solver='lbfgs'` or `newton-cg'`.