

In our ML.NET pipeline, we are using the **SDCA (Stochastic Dual Coordinate Ascent) Maximum Entropy** classifier, which is a **multiclass classification algorithm**. Here's a breakdown of what this means:

### 1. The Model: SdcaMaximumEntropy

- **Type:** Linear multiclass classifier
- **Algorithm:** Stochastic Dual Coordinate Ascent (SDCA) with Maximum Entropy (softmax) loss
- **Best For:**
  - Text classification (like your banking query categorization)
  - Medium-sized datasets
  - Scenarios where you need probabilistic outputs (confidence scores)

### 2. Why This Model Was Chosen for Your Pipeline

Your pipeline:

```
Append(mlContext.MulticlassClassification.Trainers.SdcaMaximumEntropy(  
    labelColumnName: "RequestType",  
    featureColumnName: "Features"))
```

- **Text Features:** You first featurize text (FeaturizeText) into numerical vectors.
- **Multiclass Need:** Since RequestType has multiple categories (Account Services, Loans, Cards, etc.), a multiclass classifier is required.
- **SDCA Strengths:**
  - Efficient for medium-sized datasets
  - Supports L1/L2 regularization (helps prevent overfitting)
  - Outputs probabilities (confidence scores) for each class

### 3. Alternatives in ML.NET

If you wanted to experiment with other models:

Model (Trainers.*)	Best For	Pros	Cons
LbfgsMaximumEntropy	Small datasets	Precise, good for low-dimensional data	Slower on large data
NaiveBayes	Simple text classification	Fast training	Assumes feature independence
LightGbm	Large datasets	Handles non-linear patterns	Less interpretable
OneVersusAll	Pair with binary classifiers	Flexible	More complex setup

#### 4. Key Parameters You Could Tune

```
var options = new SdcaMaximumEntropyMulticlassTrainer.Options
```

```
{
    LabelColumnName = "RequestType",
    FeatureColumnName = "Features",
    L1Regularization = 0.01f, // Sparsity (feature selection)
    L2Regularization = 0.01f, // Prevents overfitting
    NumberOfThreads = 4      // Parallelism
};
```

```
.Append(mlContext.MulticlassClassification.Trainers.SdcaMaximumEntropy(options))
```

#### 5. How It Works in Your Pipeline

1. **Text → Features:** FeaturizeText converts words to numerical vectors (e.g., TF-IDF).
2. **Labels → Keys:** MapValueToKey encodes categories (e.g., "Cards" → 1, "Loans" → 2).
3. **SDCA Training:** Learns weights to separate classes in feature space.
4. **Keys → Labels:** MapKeyToValue converts predicted keys back to original labels.

#### 6. When to Consider Switching Models

- **If accuracy is low:** Try LightGbm or deeper featurization (e.g., TextFeaturizingEstimator with n-grams).
- **If training is slow:** Use LbfgsMaximumEntropy for smaller data.
- **If you need probabilities:** Stick with SdcaMaximumEntropy or LbfgsMaximumEntropy.

**For your banking query classification task, SdcaMaximumEntropy is a solid default choice! 🦄**