

# Market Data Regime Detection

Project for **detecting market regimes using machine learning models** (mainly HMM but will be expanded). This repository includes scripts for preparing features, training, running experiment grids, evaluating models, and visualizing states on market data (EURUSD 15M or else).

Configuración

Seleccionar Modelo HMM:  
grid\_best\_st3\_full\_ret\_log-range\_hi-log\_volume-dist\_min20-macd\_sign...

Rango de Fechas  
Seleccionar rango:  
2024/07/24 - 2024/12/31

Información del Modelo  
Estados detectados: 3  
Estados: [0, 1, 2]

Configuración de Precio  
Columna de precio:  
close



Analyzes market data to classify the current state into different regimes (e.g., bullish, bearish, ranging). The main workflow in this repository allows you to:

- Build feature sets from a CSV of prices.
- Train HMMs with splits by size (sliding train/val/test).
- Run parallelizable grids on combinations of features / number of states / covariance type.
- Evaluate models and export states and probabilities in CSV/JSON.
- Visualize states over the price series with a Streamlit app ([app.py](#)).

Observaciones en el período

10,883

Cambio de precio

-0.0444

↓ -4.09%

Volatilidad (período)

0.0246

## Resumen por Estado (período seleccionado)

	Estado	Observaciones	Porcentaje	Precio Promedio	Volatilidad	Precio Min	Precio Max
0	0	2448	22.5%	1.0831	0.0250	1.0379	1.1206
1	1	4263	39.2%	1.0840	0.0239	1.0348	1.1206
2	2	4172	38.3%	1.0822	0.0250	1.0340	1.1200

## Main structure

- [app.py](#) — Viewer for HMM evaluation results.
- [src/](#):
  - [train\\_hmm.py](#) — Trains HMMs using splits by size. Saves models in [src/data/models/](#) and results in [src/data/results/](#).
  - [experiment\\_hmm.py](#) — Runs experimental grids (parallelizable with processes or threads) on sets of features, states, and covariances.

- `evaluate_hmm.py` — Recalculates states/probabilities for saved models and generates summaries (transitions, durations, conditional statistics).
- `preprocessing.py`, `build_feature_sets.py` — (support scripts) to generate and validate CSVs of features used by experiments.
- `data/` — Data:
  - default dataset: `dataset_raw/DUKASCOPY_EURUSD_15_2000-01-01_2025-01-01.csv`
  - feature CSVs and JSONs with subsets: `features/`
  - training and evaluation outputs: `models/`, `results/`

## Authors

- **SilverDelgado** — [GitHub](#)
- **Sauvageduck24** — [GitHub](#)

---

★ If you like this project, star would be appreciated hehe! ★



\*See you in the markets 💎\*