



MACD + LSTM Trading Strategy (M5)

Direction par MACD — Filtrage par LSTM — Validation Walk-Forward



Objectif du projet

Ce projet vise à construire une **stratégie de trading algorithmique robuste et réaliste** sur le Forex en **timeframe M5**, combinant :

- une **logique directionnelle explicable** (MACD),
- un **modèle de Machine Learning (LSTM)** utilisé comme **filtre de qualité des trades**,
- une **validation walk-forward stricte** afin d'éviter tout biais temporel.

L'objectif n'est **pas** de prédire le marché, mais de **sélectionner les setups les plus favorables** parmi des signaux techniques déjà définis.



Principe général de la stratégie

La stratégie repose sur une **séparation claire des responsabilités** entre l'indicateur technique et le modèle ML.

1 Direction du trade (logique déterministe)

La **direction (LONG / SHORT)** est déterminée **exclusivement par le MACD** :

- **LONG**
Lorsque l'histogramme MACD passe de négatif à positif
- **SHORT**
Lorsque l'histogramme MACD passe de positif à négatif



Le modèle LSTM **ne décide jamais de la direction du trade**.

2 Filtrage des trades (Machine Learning – LSTM)

Un **LSTM** est entraîné pour répondre à une seule question :

Ce setup MACD a-t-il une probabilité élevée d'être gagnant dans les prochaines bougies ?

Le modèle :

- apprend à reconnaître des **patterns favorables** dans les données passées,
- produit une **probabilité de succès (hit_proba)**,
- autorise le trade uniquement si cette probabilité dépasse un **seuil**.

Le LSTM agit donc comme un **filtre de qualité**, et non comme un modèle directionnel.

3 Définition du succès (label HIT)

Un trade est considéré comme **gagnant (HIT = 1)** si, dans les **5 bougies suivantes** :

- **LONG** → le **prix maximum futur** dépasse le prix d'entrée
- **SHORT** → le **prix minimum futur** est inférieur au prix d'entrée

Sinon, le trade est considéré comme **perdant (HIT = 0)**.

Aucun SL/TP n'est utilisé à ce stade.

Le modèle apprend la **qualité intrinsèque du setup**, indépendamment du money management.

Modèle 1 — Entraînement sur dataset complet (~10 000 bougies M5)

Résultats ML sur le jeu de test

```
Accuracy : 0.982
Precision : 0.000
Recall : 0.000
F1-score : 0.000
ROC AUC : 0.545
```

Analyse

- Le dataset est **fortement déséquilibré** (peu de trades réellement gagnants).
- L'accuracy seule est **peu informative**.
- La **ROC AUC > 0.5** montre néanmoins que le modèle capte un **signal exploitable**.

Distribution des probabilités prédites

```
Hit probability:
min ≈ 0.23
mean ≈ 0.25
max ≈ 0.28
```

 Le seuil optimal sur ce modèle global se situe autour de **0.24**

(Ce modèle sert de référence, mais n'est pas utilisé tel quel en production).

Modèle 2 — Validation Walk-Forward (approche réaliste)

Paramétrage Walk-Forward

```
TRAIN_SIZE = 1000 bougies (~3-4 jours M5)
TEST_SIZE = 300 bougies (~1 jour M5)
STEP_SIZE = 300 bougies
```

👉 Le modèle est **réentraîné à chaque itération**, simulant un usage réel.

📊 Résultats Walk-Forward (seuil LSTM = 0.55)

Start index	Trades	Accuracy
60	16	0.69
360	15	0.67
660	22	0.73
...
8460	14	0.79

- Nombre de trades suffisant sur chaque fenêtre
- Performances **stables dans le temps**
- Accuracy moyenne :

≈ 0.727

☑ Conclusions principales

- Le **LSTM n'est pas un modèle directionnel**, mais un **filtre probabiliste**.
- La direction reste **100 % explicable** via le MACD.
- Le walk-forward confirme :
 - l'absence de sur-apprentissage,
 - la robustesse temporelle du modèle,
 - la cohérence entre entraînement global et incrémental.
- Le seuil élevé (0.55) agit comme un **filtre conservateur**, privilégiant la qualité des trades.

🚀 Prochaines étapes possibles

- Optimisation **SL / TP** sur les trades filtrés
- Analyse du **drawdown**, de l'expectancy et du profit factor
- Calibration dynamique du seuil LSTM
- Passage en **paper trading / live trading**

🔗 Philosophie du projet

Le modèle ne cherche pas à prédire le marché,
mais à **éviter les mauvais trades**.

Cette approche hybride (technique + ML) privilégie :

- l'explicabilité,
- la robustesse,
- la reproductibilité en conditions réelles.