Regularized Models FX V1

With Highly dense Features of day, hour
4 and time bound features

By https://github.com/Aliyansayz

Contents of the Research

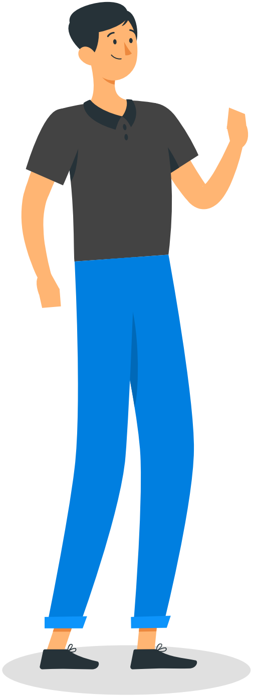
Here's what you'll find in this **Regularized Fx Models highly dense features** research:

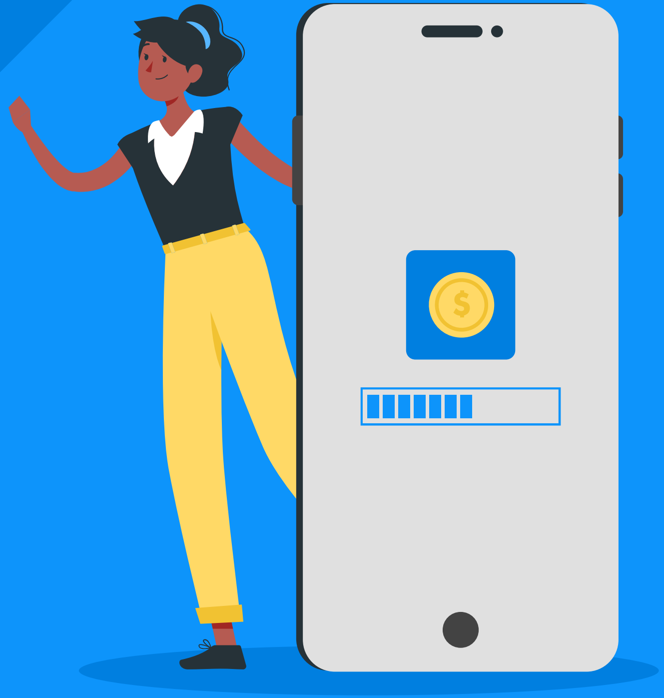
1. Importance of ML powered trading over traditional statistical methods
2. Why multiple currencies pairs upto 28 better in times of anomalies than few chosen
3. Features Used
4. Two Important Goals Tackling **Data Drift** Minimizing Losses And Scaling Gains
5. **Single Month Regularized vs Two Months Regularized Hyperparameters** Hyperparameters
6. All Months Results Why Exclude December
7. All 28 Forex Pairs Hyperparameters
8. Ending Note And Future Prospect:
    - Machine Learning library used **catboost**
    - Creating separate models for **BUY** and **S**ELL rather than using pips change negative sign for sell and positive value of predicted pips change as buy opportunity
    - **Training Data** used from June 2020 to September 2023, **Testing Data** from October 2023 to May 2024 .
    - Future Trajectory
    - Use of **Consistency Edge** , for longer periods atleast 6 months 8 months
    - Addition of **Rigidity** in **Hour4 Features** making hour 4 timeframe consistent to major price swings in favor of next day swing

By https://github.com/Aliyansayz

“It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong”

—RICHARD FEYNMAN



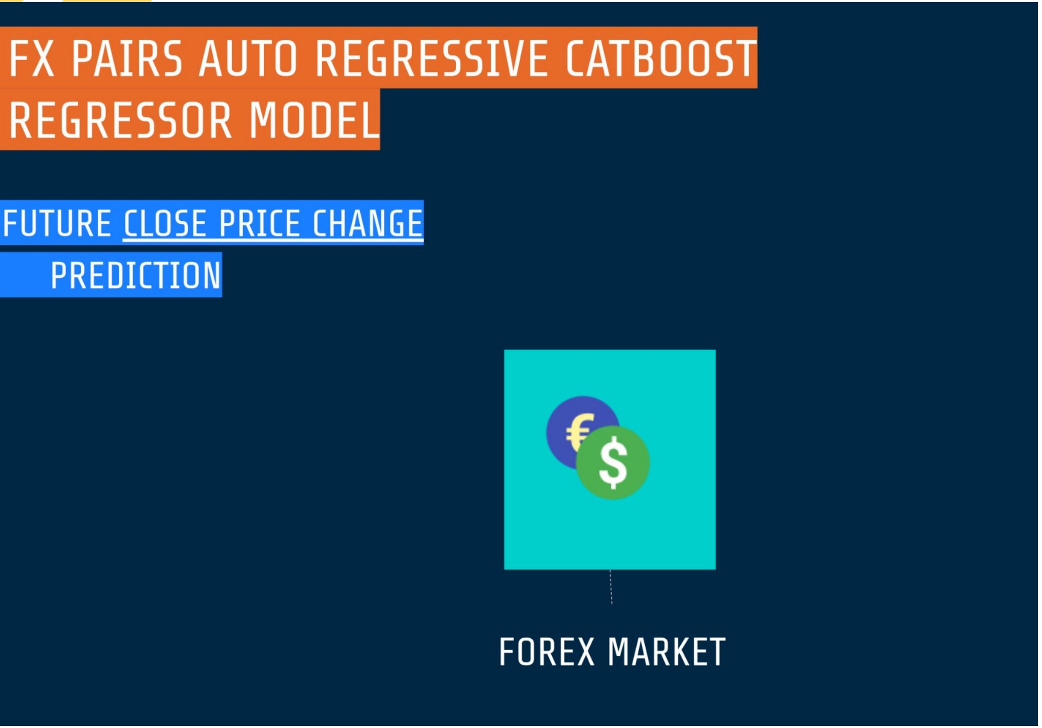


Importance Of ML Over Traditional Statistic Model

ML-Powered Trading vs. Statistical Methods

- ML models can analyze vast datasets and uncover hidden patterns.
- Traditional statistical methods may struggle to capture complex market dynamics.
- ML models can adapt to changing market conditions, leading to more robust predictions.
- ML models can provide more accurate and timely price predictions.
- ML-powered trading can automate decision-making, reducing human error.

SIGNIFICANCE Of Multi Currency ML Model Features v1



1. Densed amount of Features
2. Hour 4 features to add granularity
3. Seasonality features like day of week, month of year, week of month

By https://github.com/Aliyansayz

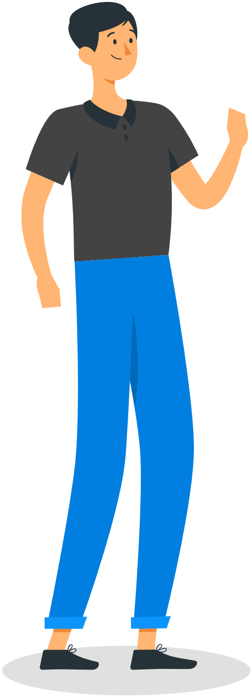
2

Diversity

By https://github.com/Aliyansayz

Why 28 Currency Pairs Are Better Than a Few

- Diversification reduces exposure to the impact of anomalies in individual currency pairs.
- A broader dataset improves model robustness and accuracy.
- Captures a wider range of market behaviors.
- Anomalies in individual pairs have less impact on overall performance.
- More data points lead to better model training and generalization.



3

Importance Of ML Model Features

PRICING MODEL FEATURES VS TARGET

Features Of Day

All Day Features(previous 9 days),

Features Of Hour 4

Heikin Ashi, relative range and candle features of hour4 groupby day index used previous 7 days.

Price Change Of Next Day

Using previous days features next day price change is ***predicted*** +ve means price rise -ve means price fall

Features Intro

1 — Relative Strength Index

Higher value indicate strong upward move

2 — Average True Range Filter

Supertrend, trailing band using ema

3 — Seasonality

Day of week, week of month, month of year

4 — Dispersion, Exponential Moving Average

Standard deviation, exponential moving average 3, 5, 7 period

PRICING MODEL FEATURES VS TARGET

Features X

All Day Features(previous 9 days), some hour-04 features (previous 7 days)

Target y

Daily Price Change, if negative then drop in price if positive then rise from the previous day price

MODEL TARGET VARIABLE

Price Change present day	If JPY NOT in Pair	Price Change * 10 ** 4 (expressed in pips)
Price Change present day	If JPY in Pair	Price Change * 10 ** 2 (expressed in pips)

DAY FEATURES Lag By 9

Day Of Week	Week Of Month	Month	Last 9 Days
OHLC	Relative Strength Index RSI	Standard Deviation STDEV	Last 9 Days
Heikin Ashi	Supertrend	Elastic Supertrend	Last 9 Days
Bollinger Bands	Price Range	Median	Last 9 Days
EMA-3-5-7-14			Last 9 Days

DAY FEATURES Period Factor Settings

Relative Strength Index RSI	14 Period	**Month**	**Last 9 Days**
Standard Deviation STDEV	5 Period	**Price Range**	**High - Low**
Heikin Ashi	simple	**Elastic Supertrend**	**atr_length=10, atr_multiplier=2.5 , ma_length=10**
Bollinger Bands	Period= 5 ,Factor= 2	**Median**	**High+Low/2**
EMA-3-5-7-14	Respective period	**Supertrend**	**period= 10, multiplier=0.66**

HOUR-4 FEATURES Used

Relative Range	9 Period	Close - ((Highest_high + Lowest_low) / 2)	Last 7 Days
Candle Type	By each candle		Last 7 Days
Heikin Ashi	By each candle		Last 7 Days

Hour-4 Features Lag By 7

Relative Range	Last 7 Days	Selected
Candle Type	Last 7 Days	Selected
Heikin Ashi	Last 7 Days	Selected
OHLC		Unselected
Standard deviation		Unselected

Hour 4 Lagged Features Addition Into Model

Group By Daily Date Index

Hour4 OHLC

Combine all 6 ohlc into a single index into one single day index

Lag By 7

Make Hour 4 Features Lag By 7 for each index

Hour4 Features

Make features of lagged ohlc Relative range, standard deviation, candle type etc

4

Dealing With Data Drift Issues

Data Drift

- Data drift is the change in model input data over time.
- This causes model performance degradation.
- Statistical Properties that change: mean, standard deviation, correlations, etc.
- Real World Example: Forex market pricing model trained on pre-pandemic data.
- Post-pandemic, model performance degrades due to changes in market dynamics.

Tackling The Challenge Of Data Drift

- Rather than using one month regularized hyperparameters we used two months for regularization in other words 44 days(22 days for each month) net gains is positive when a particular hyperparameter is used.

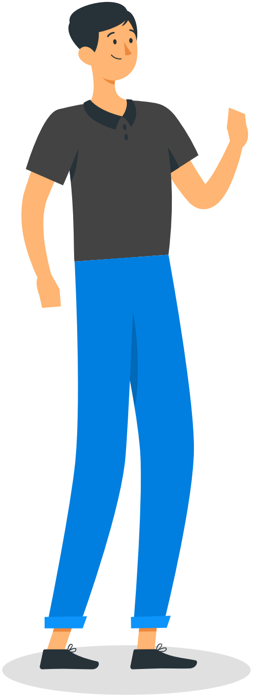
- Previously February was giving ***398 pips*** in net gains when May month only regularized hyperparameters used.

- Using April-May regularized hyperparameters, February was giving ***1288 pips***, model was able to counter data drift effectively.

Steps Taken For Tackling Data Drift

- Data drift is the change in model input data over time.
- This causes model performance degradation.

- Real World Example: Forex market pricing model trained on pre-pandemic data.
- It was impossible for us to keep every model for each pair in surplus for all months
- We need overall positive net gains by 60% like 16 to 20 forex pairs in positive gains. Out of 28 fx pairs.



5

Regularization 1-Month vs 2-Month

One Month vs Two Month Regularization

- With passage of time one month May regularized model showed poor performance
- April performance was very high 3955 pips in net gains
- February performance showed a drop with disappointing 398 pips net gain only
- In comparison to one month , two months april-may regularized model showed 4 times better performance in february with 1287 pips net gain.

- Overall using **two months** regularized model is safer than single month in longer runs

7

Fine Tuning Hyperparameters

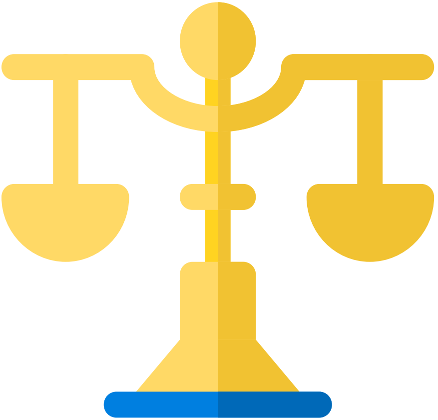
Fine Tuning Hyperparameters

Previous Months

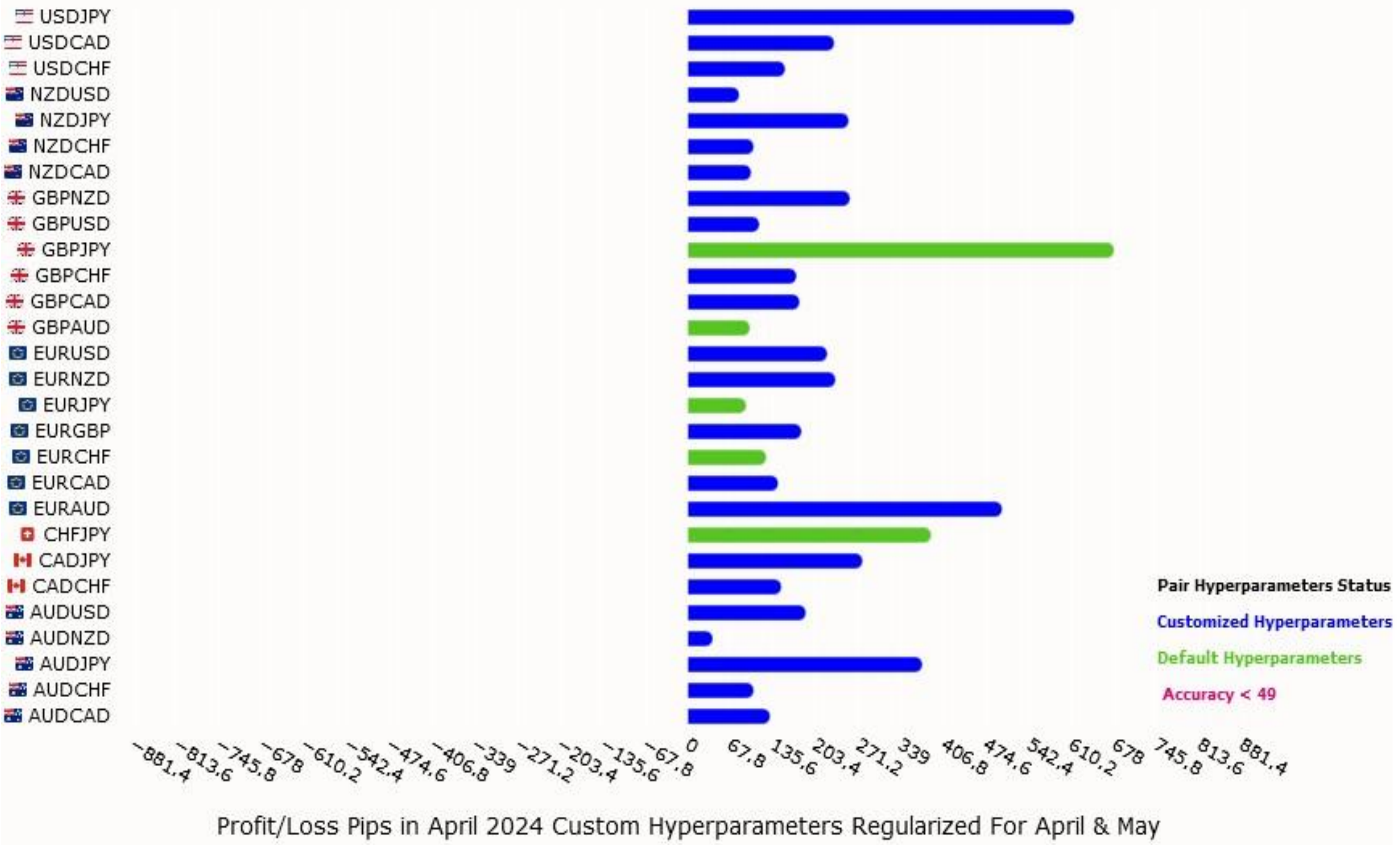
Previous points related to previous months are 50% accurate by days overall and net gain accuracy 50 to 54% overall

Observing Month To Finetune

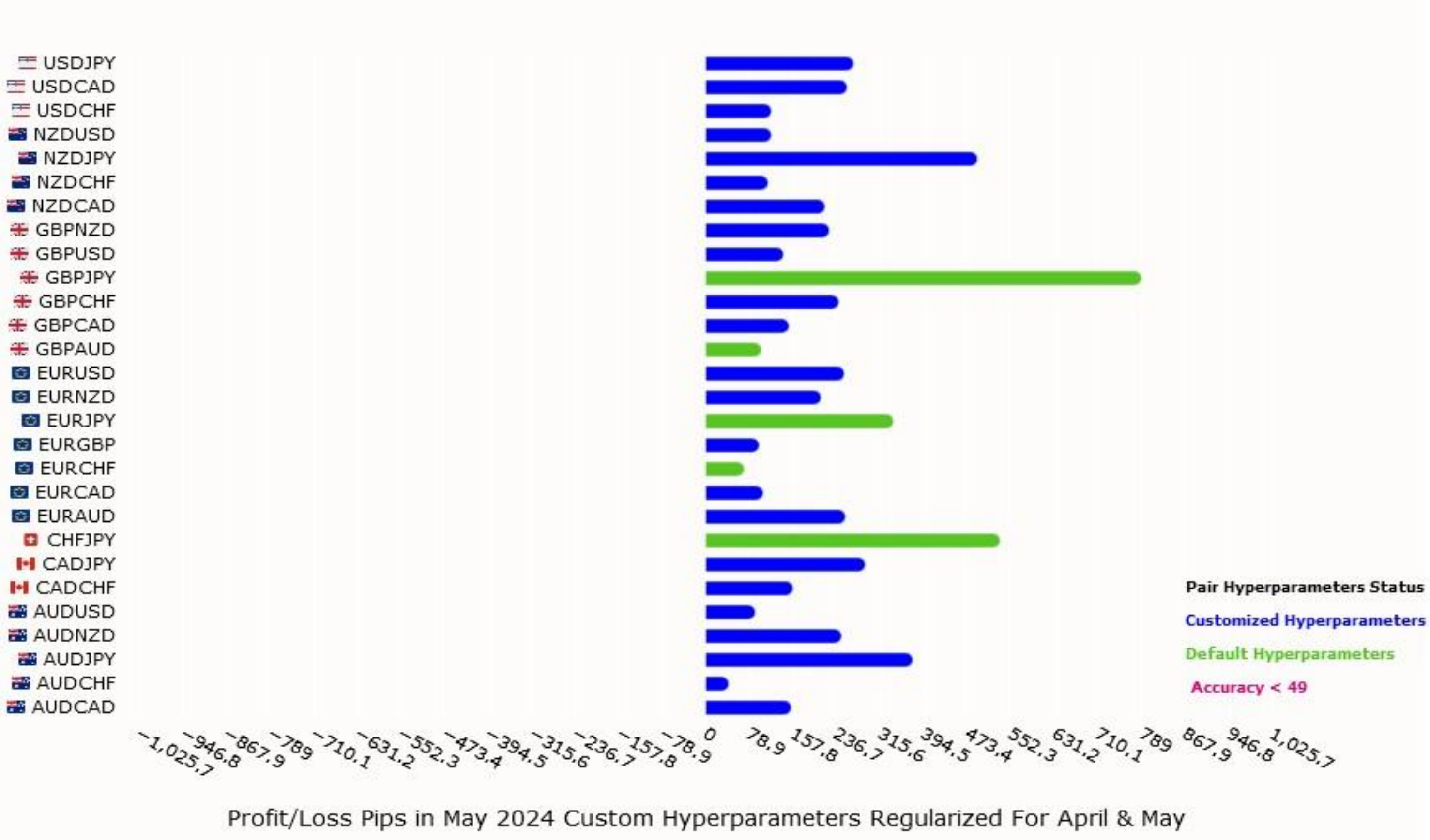
Observing present month must be giving us accuracy by days and net gains higher than 50%



Forex ML Model Predict Market Custom Hyperparameter 📈📉 April 2024



Forex ML Model Predict Market Custom Hyperparameter 📈📉 May 2024



Symbols Hyperparameters That Worked

SYMBOL	ITERATIONS	LEARNING RATE	DEPTH
AUDCAD	9	0.67	7
AUDCHF	7	0.71	7
AUDJPY	5	0.01	6
AUDNZD	15	0.68	7
AUDUSD	5	0.7	7
CADCHF	7	0.06	8
CADJPY	5	0.01	6

Window size 9 for all pairs..

Symbols Hyperparameters That Worked

SYMBOL	ITERATIONS	LEARNING RATE	DEPTH
🇨🇭 CHFJPY	300	0.01	6
🇪🇺 EURAUD	7	0.06	8
🇪🇺 EURCAD	20	0.039	7
🇪🇺 EURCHF	300	0.01	6
🇪🇺 EURGBP	188	0.001	7
🇪🇺 EURJPY	300	0.01	6
🇪🇺 EURNZD	188	0.55	7

Symbols Hyperparameters That Worked

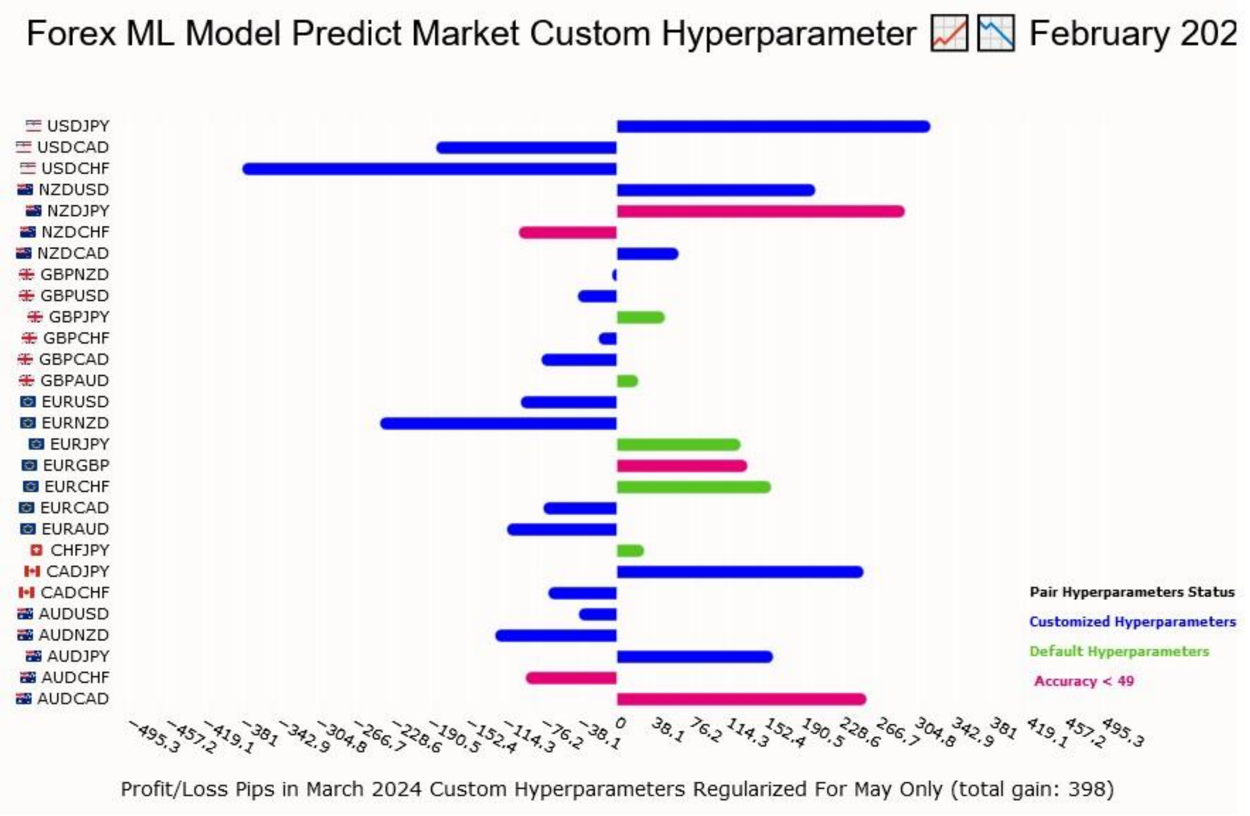
SYMBOL	ITERATIONS	LEARNING RATE	DEPTH
🇪🇺 EURUSD	450	0.025	5
🇬🇧 GBPAUD	13	0.77	8
🇬🇧 GBPCAD	5	0.001	7
🇬🇧 GBPCHF	15	0.1	8
🇬🇧 GBPJPY	300	0.01	6
🇬🇧 GBPNZD	15	0.1	7
🇬🇧 GBPUSD	13	0.77	7

Symbols Hyperparameters That Worked

SYMBOL	ITERATIONS	LEARNING RATE	DEPTH
🇳🇿 NZDCAD	17	0.38	7
🇳🇿 NZDCHF	27	0.35	7
🇳🇿 NZDJPY	15	0.001	6
🇳🇿 NZDUSD	5	0.77	7
🇺🇸 USDCAD	20	0.039	7
🇺🇸 USDCHF	188	0.67	7
🇺🇸 USDJPY	5	0.01	6

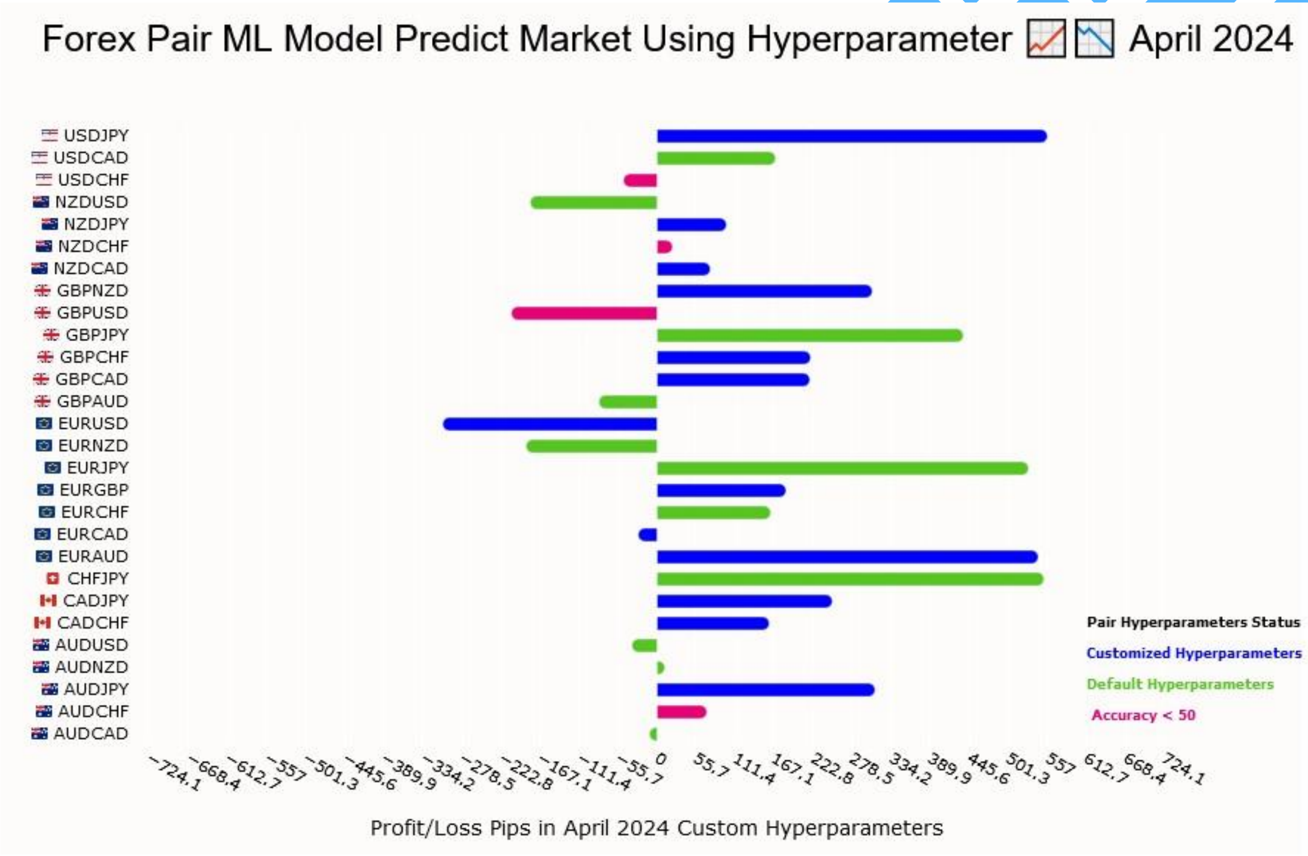
Month May Only Regularized Results In FEBRUAURY 2024

Total Gains 398 pips



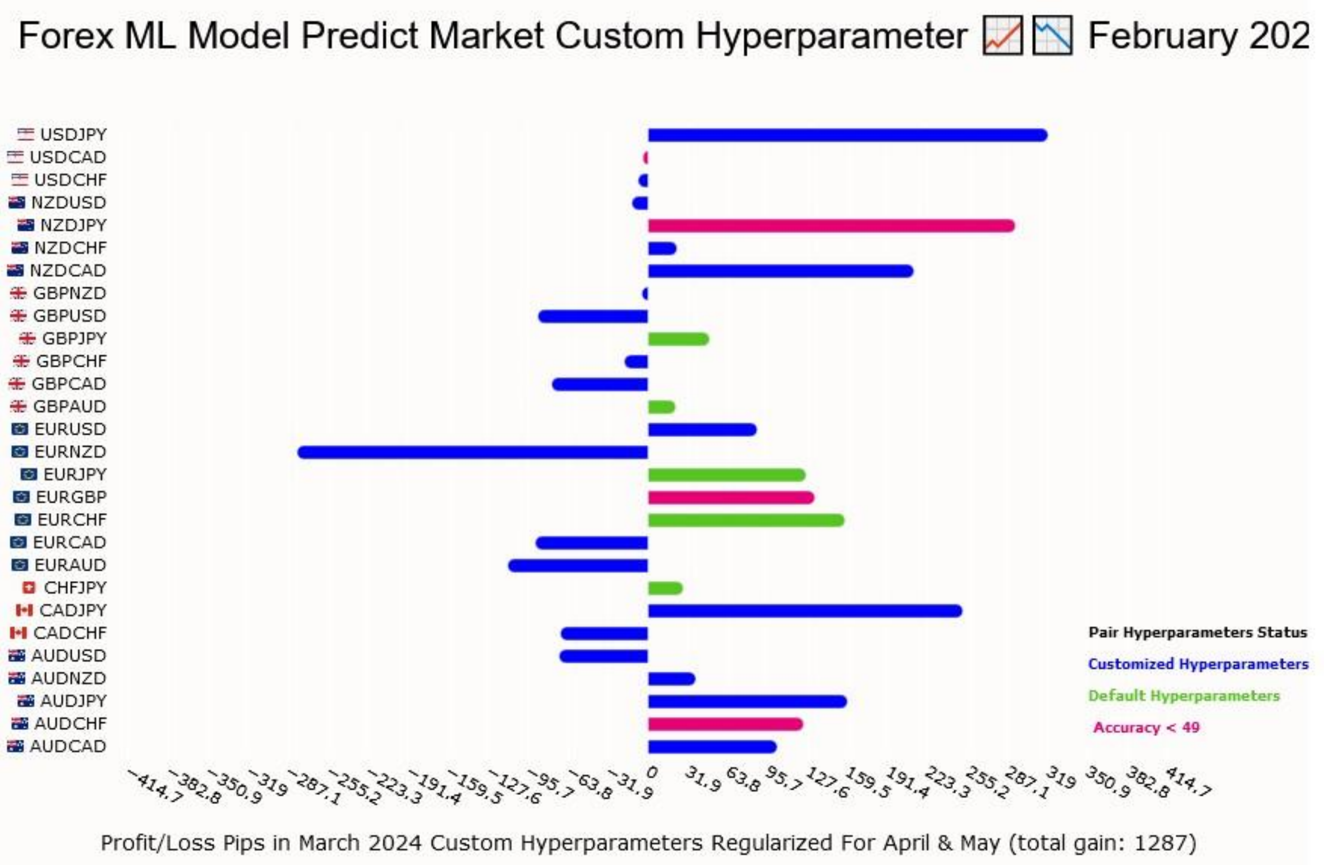
Month May Only Regularized Results In APRIL 2024

Total Gains 3955 pips



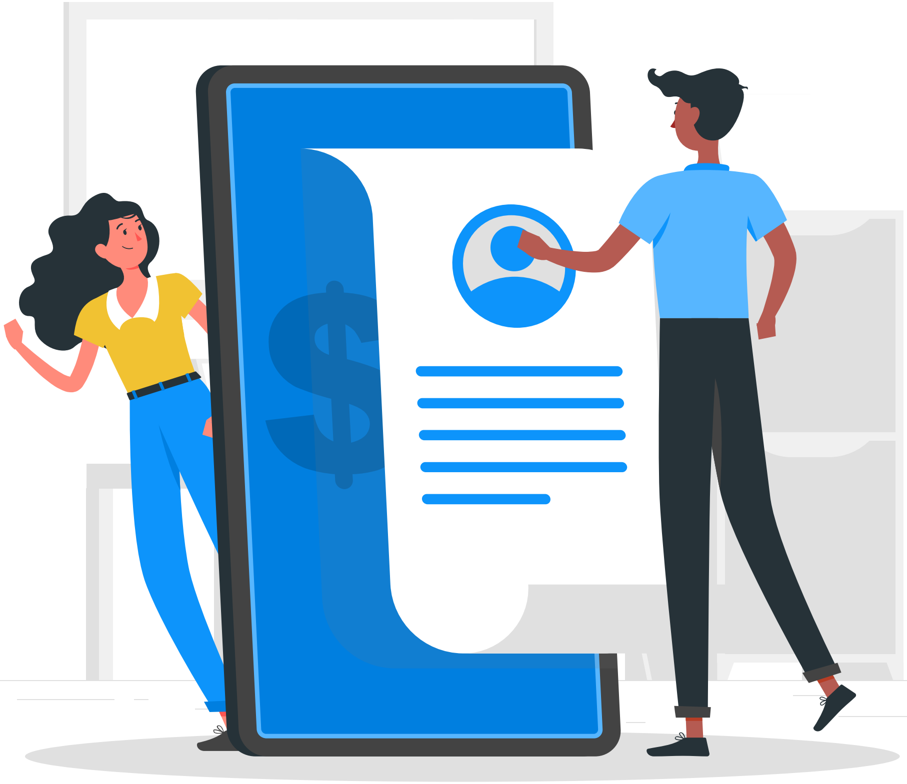
Two Months April May Only Regularized Results In FEBRUAURY 2024

Total Gains 1287 pips



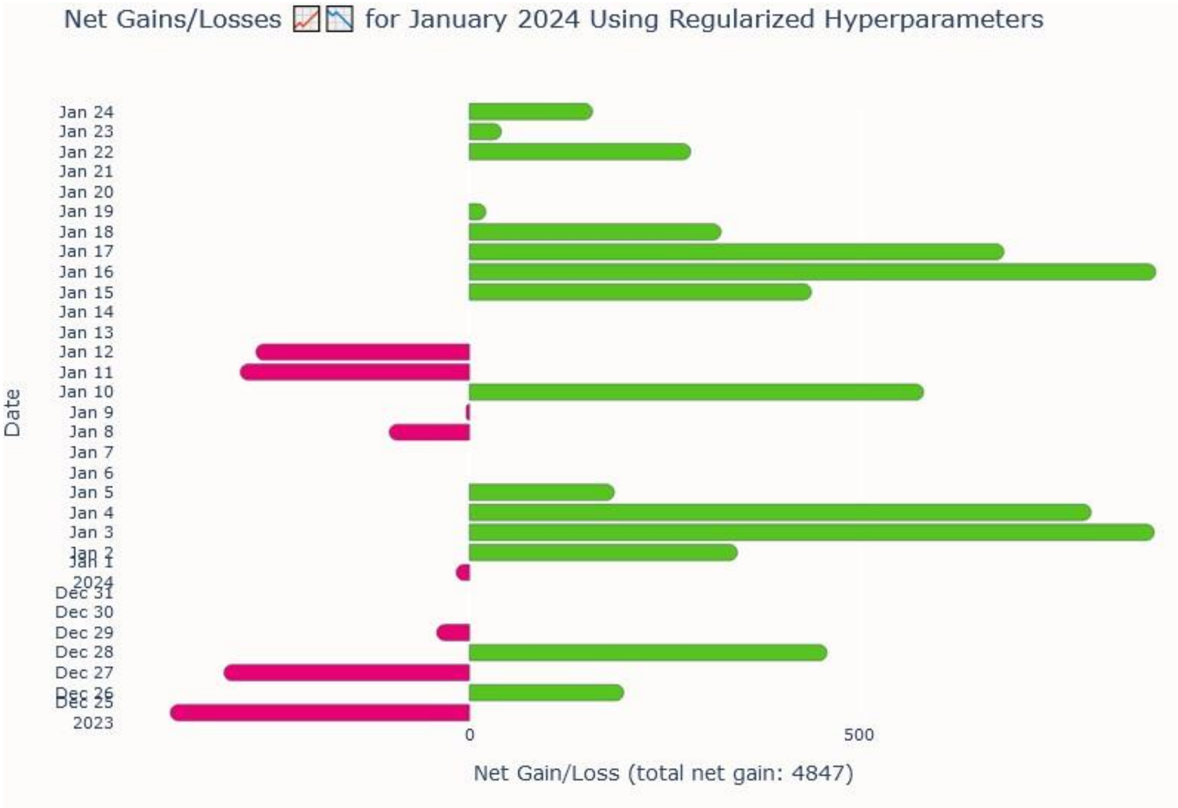
Ending Notes

8

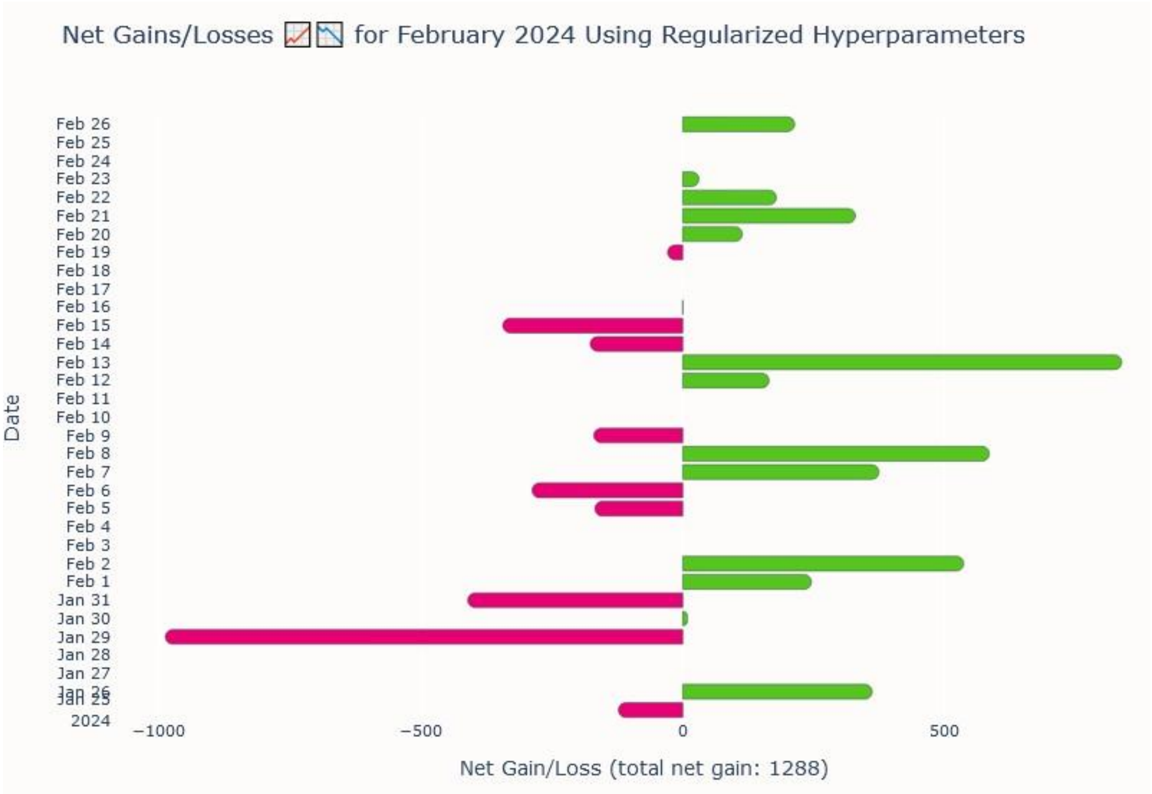


By https://github.com/Aliyansayz

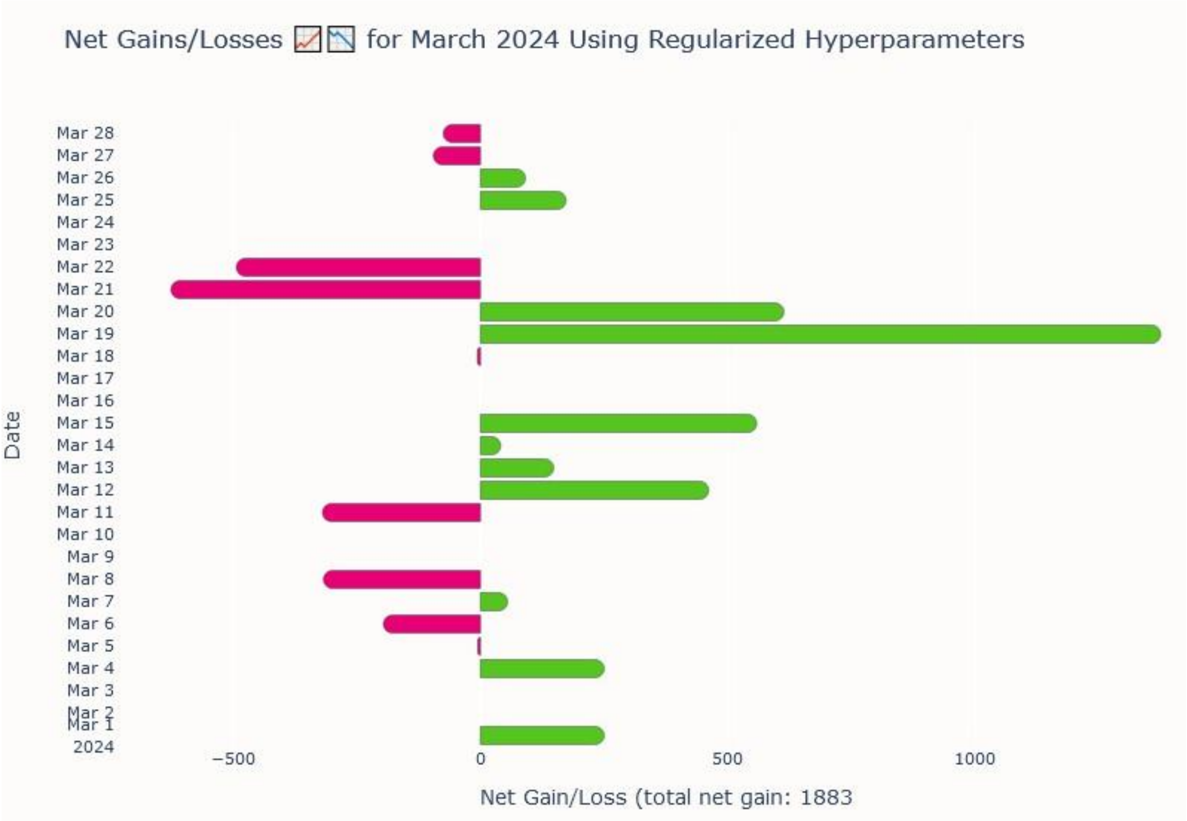
Average Possible Scenarios In Two Months Regularized Forex Pairs Trading



Average Possible Scenarios In Two Months Regularized Forex Pairs Trading



Average Possible Scenarios In Two Months Regularized Forex Pairs Trading



Future Trajectory

- **Trade off** Profitability over Consistency
- Use of **Consistency Edge** , for longer periods atleast 6 months 8 months to avoid negat

- Addition of **Rigidity** in **Hour4 Features** making hour 4 timeframe consistent to major price swings for primary purpose to favor with next day 6 candles
- Eliminating seasonal features of time like week of month, month, day of week observed increase in accuracy

Notebook Conclusion

Version 1 Regularized On ***Two Months April May*** On Two Conditions :

1. Hyper parameters chosen should give positive gains for any currency pairs for those two months
2. Previous months accuracy should be 50% and net gains accuracy close to 50% or 52%

Results(pips) ✅October 2167 ✅ November 2498 ❌ December -3080 ✅ January 4844 ✅ February 1288 ✅ March 1883

- 📉 Not all models were in positive gains in previous month despite all performing good on april may with positive gains

- ⭐ But parameters performing good on two different months gave them ability to keep majority of currency pairs model in positive gain

- 💢 December is also not recommended for trading by trading community so we might ignore that month outcomes

Conclusion

Regularization Of Parameters that helps in better accuracy across more months, thus ensures more positive gains on unseen months

Next version of research will be focused onavoiding collinearity and making *models good enough* over as many months at least 6 months they are consistent in profit

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by Freepik and illustrations by **Storyset**.

Please keep this slide for attribution.

By https://github.com/Aliyansayz