



PROYECTO FINAL

Santiago Pereira
Brandon Reyes
Nancy Mazariegos



CONTEXTO

- Problema: Predicción de movimientos intraday en forex GBPUSD (intervalos de 5 minutos)
- Objetivo: Sistema de meta-modelo ensemble de 10 componentes para trading cuantitativo
- Solución implementada: Temporal Fusion Transformer (TFT) - Primer componente del sistema



DATOS

- Variables de Entrada (32 features totales):
- Temporales Conocidas (13):
- Componentes: hour, day_of_week, minute
- Sesiones Forex: Tokyo, London, NY, Overlap
- Encoding cíclico: sin/cos de hora y día
- Variables de Mercado (19):
- Retornos logarítmicos: 1p, 3p, 6p periodos
- Precios OHLC: Close, High, Low, Open
- Indicadores técnicos: RSI, MACD, Bollinger Bands, ATR, Momentum, Volatilidad
- Variable Objetivo:
- target_log_return: Retorno logarítmico futuro a 5 minutos
- 7 cuantiles predichos: [2%, 10%, 25%, 50%, 75%, 90%, 98%]



MODELO Y ARQUITECTURA

Componentes Clave:

- Variable Selection: Aprende qué features son importantes
- LSTM Encoder/Decoder: Captura dependencias temporales
- Attention: Identifica timesteps relevantes (interpretabilidad)
- Quantile Output: Predice distribución completa, no solo media

Configuración:

- Parámetros: 60,769 ($\sim 0.1M$)
- Hidden size: 32 | Attention heads: 2 | Dropout: 0.2
- Cuantiles: [0.02, 0.10, 0.25, 0.50, 0.75, 0.90, 0.98]

RESULTADOS

Performance Test Set:

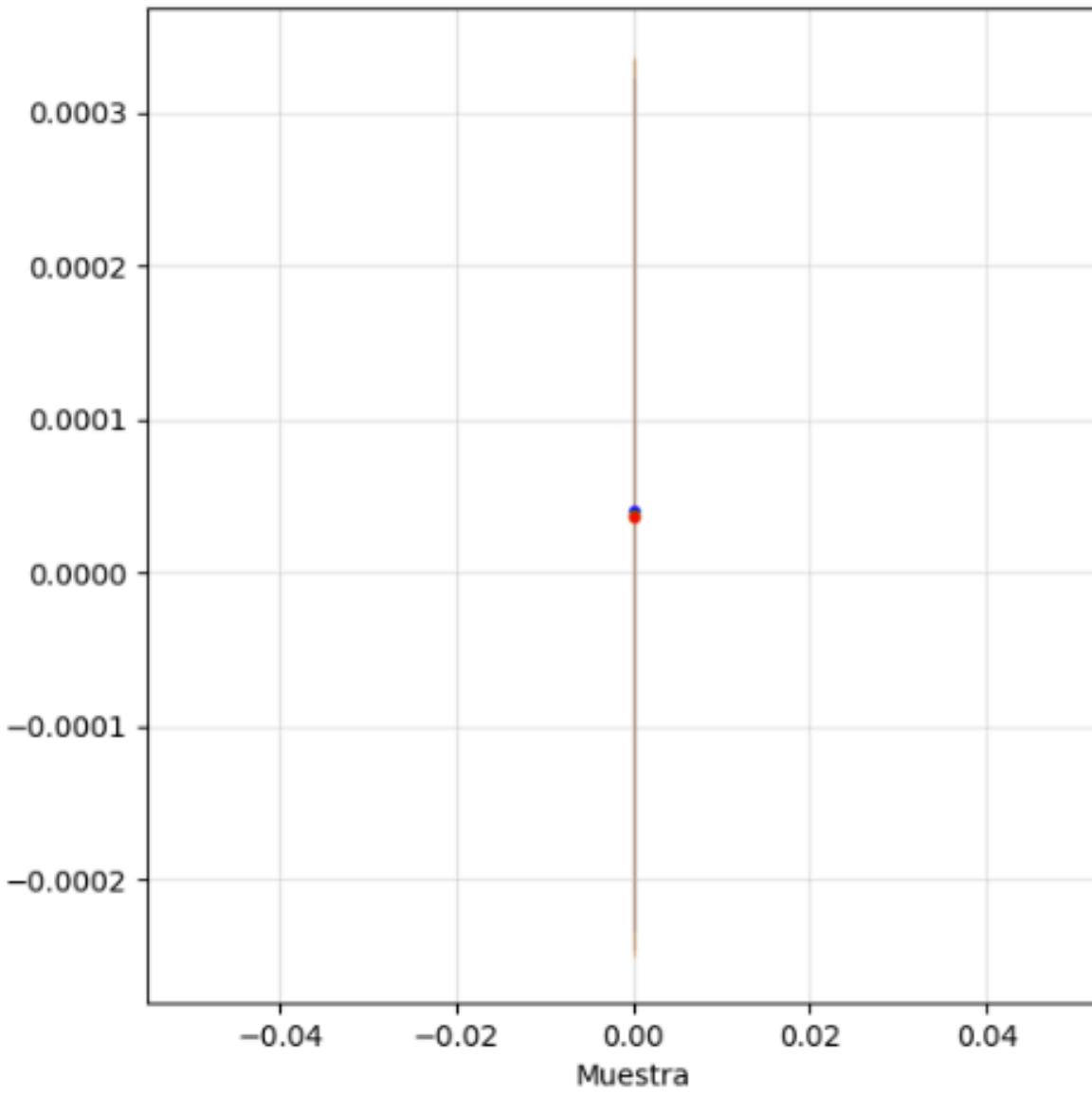
- Samples: 1,743 registros (8 días out-of-sample)
- Mediana predicción: -0.000020 ± 0.000001
- VaR 10%: -0.000156
- VaR 5%: -0.000287
- Incertidumbre promedio: 0.000581

27 Features Generadas para Meta-Modelo:

- Cuantiles (7): Promedios de cada cuantil
- Incertidumbre (3): IQR, amplitud, skew
- Direccionales (5): Prob. alcista/bajista, strong signals
- Riesgo (4): VaR 10%, VaR 5%, CVaR
- Confianza (3): Por horizonte temporal
- Volatilidad (3): Short/medium/long
- Momentum (2): Tendencia y aceleración

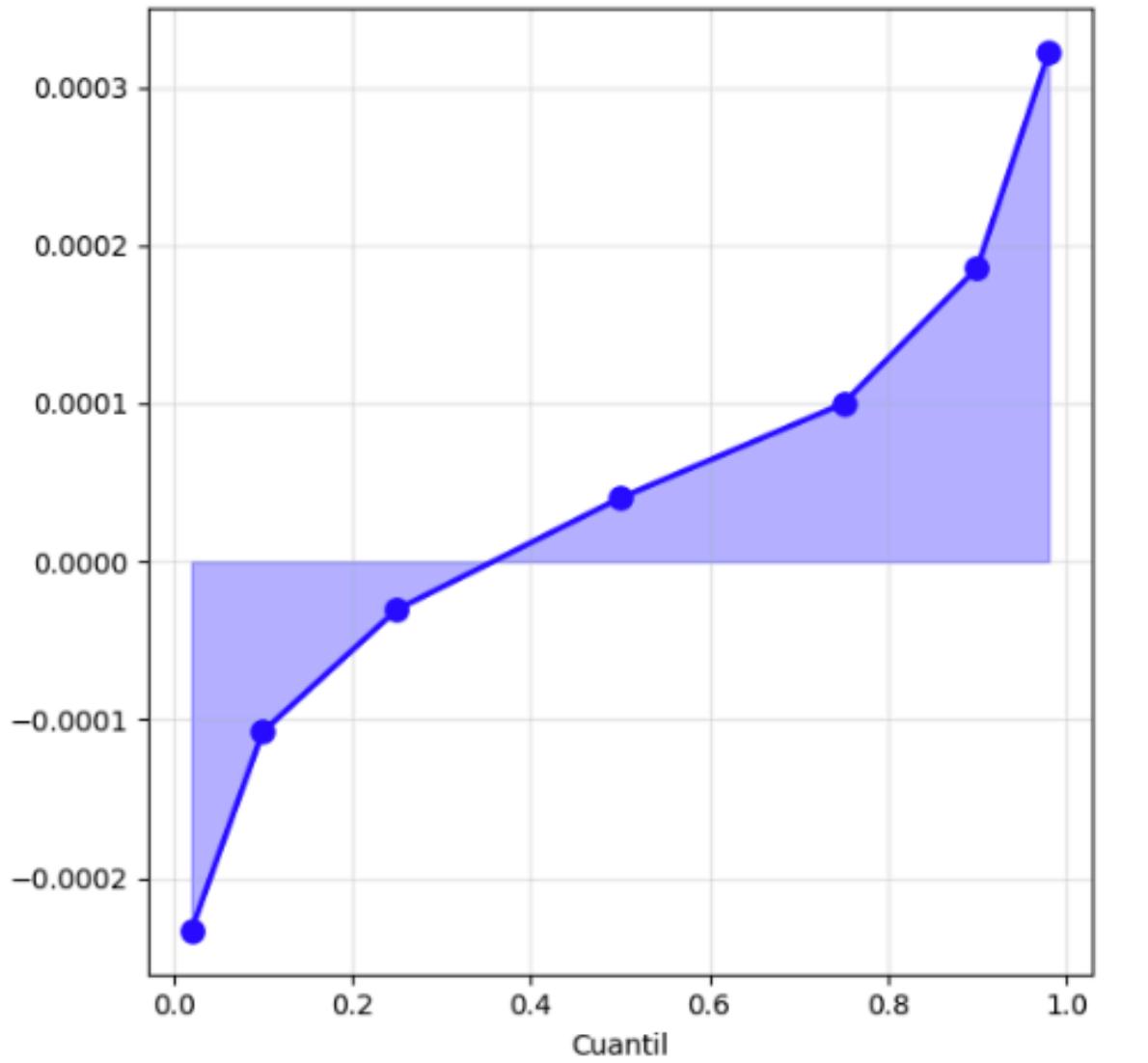
Predicciones Multi-Horizonte

Log Return Predicho

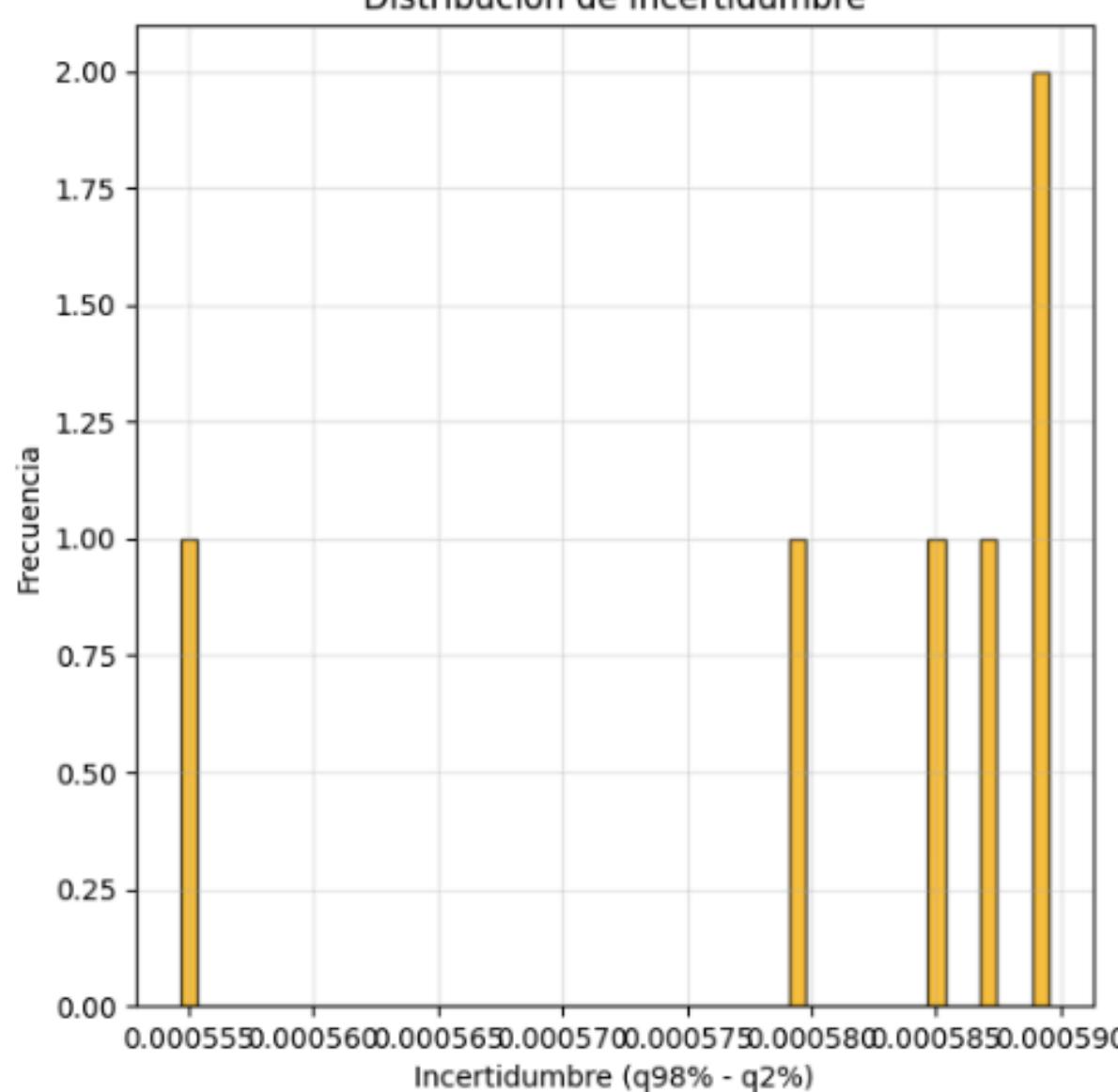


Distribución de Cuantiles
(Muestra 1, H1)

Valor Predicho



Distribución de Incertidumbre



DISCUSION

Insights del Mercado:

- Patrones fuertes durante overlaps de sesiones
- Encoding cíclico temporal es crucial
- Features técnicos mejoran performance significativamente

Validación del Enfoque:

- Arquitectura de complejidad apropiada (60K params)
- Balance entre performance e interpretabilidad
- Base sólida para sistema completo



RECOMENDACIONES Y MEJORAS

Con más tiempo se optimizarían hiperparámetros, añadirían features adicionales (Ichimoku, volumen, microestructura) y profundizaría el análisis de interpretabilidad. El siguiente paso crítico es completar el meta-modelo ensemble implementando los 8 modelos restantes: XGBoost_GPU, LSTM_Attention, CNN_1D_LSTM, FinBERT_Local, TabNet, VAR, Volume_CNN y GCN, culminando con un Neural Network Stacking como meta-learner final. Para producción se requiere pipeline en tiempo real (<100ms latency), backtesting extensivo con costos de transacción, y sistema integrado de risk management con position sizing basado en VaR.

GRACIAS

