VOLATILITY PREDICTION

OPTIVER GROUP 7

DATA3888 DATA SCIENCE CAPSTONE

Show UI

TAN_1CLAS_800095

1.03

RESEARCH HYPOTHESIS

HYPOTHESIS (H_1)

GIVEN EQUIVALENT DATA QUALITY, FEATURE ENGINEERING, AND PREPROCESSING CONDITIONS, A TRANSFORMER-BASED MODEL OUTPERFORMS TRADITIONAL ARCHITECTURES—SPECIFICALLY LSTM NETWORKS AND RANDOM FOREST REGRESSORS—IN SHORT-TERM VOLATILITY FORECASTING ON LEVEL-2 LIMIT ORDER BOOK (LOB) DATA.

NULL HYPOTHESIS (H_0)

UNDER IDENTICAL DATA, FEATURE ENGINEERING, AND PREPROCESSING CONDITIONS, A TRANSFORMER-BASED MODEL DOES NOT SIGNIFICANTLY OUTPERFORM LSTM NETWORKS OR RANDOM FOREST REGRESSORS IN FORECASTING SHORT-TERM VOLATILITY ON LEVEL-2 LIMIT ORDER BOOK DATA.

1334758950_SINCE_1932
XXL V7868376M

Show UI

VOLATILITY

VOLATILITY MEASURES HOW MUCH AN ASSET’S PRICE MOVES OVER TIME. IT REPRESENTS THE DEGREE OF UNCERTAINTY OR RISK IN RETURNS AND OFTEN CHANGES OVER TIME IN PATTERNS LIKE BURSTS OR CALM PERIODS.

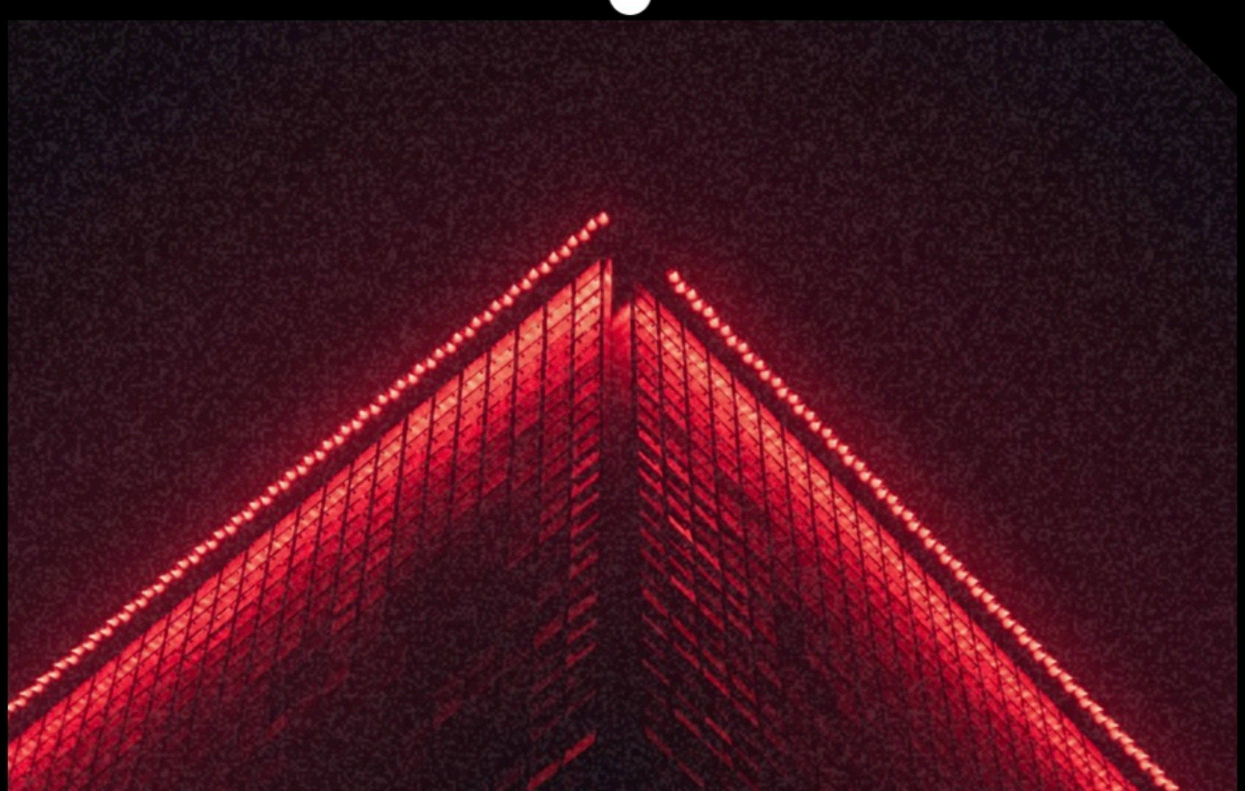
ALTHOUGH VOLATILITY BY ITSELF PROVIDES SIGNIFICANT MARKET INSIGHTS, WE HAVE ASSUMED THAT IT WILL BE USED AS AN INPUT SIGNAL IN A LARGER AND MORE COMPLEX MACHINE LEARNING MODEL.

MORE ON VOLATILITY

Show UI

THE FINANCE THEORY

VOLATILITY IS NOT JUST NOISE — IT IS THE MEASURABLE SIGNATURE OF RISK IN FINANCIAL SYSTEMS.



Show UI

KEY SECTIONS

WHY VOLATILITY MATTERS IN FINANCE?

ACCURATE SHORT-TERM VOLATILITY FORECASTING INFORMS POSITION SIZING, PORTFOLIO HEDGING, MARKET-MAKING SPREADS, AND HIGH-FREQUENCY ARBITRAGE. FOR DERIVATIVE MARKETS, VOLATILITY ASSUMPTIONS DIRECTLY INFLUENCE OPTION PRICING VIA MODELS LIKE BLACK-SCHOLES AND HESTON. IN A LIMIT ORDER BOOK SETTING, CORRECTLY ANTICIPATING VOLATILITY SURGES ENABLES TRADERS TO ADJUST ORDER PLACEMENT STRATEGIES DYNAMICALLY, BALANCING EXECUTION COST AND ADVERSE SELECTION RISK.

MORE ON VOLATILITY

EMPIRICAL VOLATILITY EXHIBITS WELL-DOCUMENTED STATISTICAL BEHAVIORS: CLUSTERING, MEAN REVERSION, LEPTOKURTOSIS, AND THE LEVERAGE EFFECT. THESE IMPLY THAT VOLATILITY IS A STOCHASTIC PROCESS, NOT A CONSTANT. CLUSTERING, FOR INSTANCE, MEANS HIGH-VOLATILITY PERIODS ARE LIKELY TO BE FOLLOWED BY FURTHER HIGH VOLATILITY — A PATTERN EXPLOITABLE BY MODELS LIKE GARCH OR TRANSFORMERS WITH ATTENTION MECHANISMS. YOUR TRANSFORMER LEVERAGES THIS BY MODELING LONG-RANGE DEPENDENCIES ACROSS ORDER BOOK SEQUENCES THAT SIMPLER MODELS LIKE RANDOM FORESTS CANNOT CAPTURE.

Show UI

KEY SECTIONS

MARKET INCOMPLETENESS AND VOLATILITY RISK

STOCHASTIC VOLATILITY MODELS (LIKE HESTON OR BATES) INTRODUCE NON-HEDGEABLE RISK DUE TO INDEPENDENT VARIANCE DRIVERS, LEADING TO INCOMPLETE MARKETS. IN PRACTICAL TERMS, THIS MEANS VOLATILITY ITSELF BECOMES A RISK FACTOR, PRICED THROUGH RISK PREMIA OR VIA INSTRUMENTS LIKE VIX DERIVATIVES. THIS MATTERS BECAUSE YOUR LOB-BASED VOLATILITY FORECASTS, IF ACCURATE, COULD IN THEORY SUPPORT THE HEDGING OF VOLATILITY RISK — NOT THROUGH REPLICATION, BUT THROUGH PREDICTIVE ANTICIPATION, WHICH IS HIGHLY VALUED IN BUY-SIDE STRATEGIES AND MARKET-MAKING.

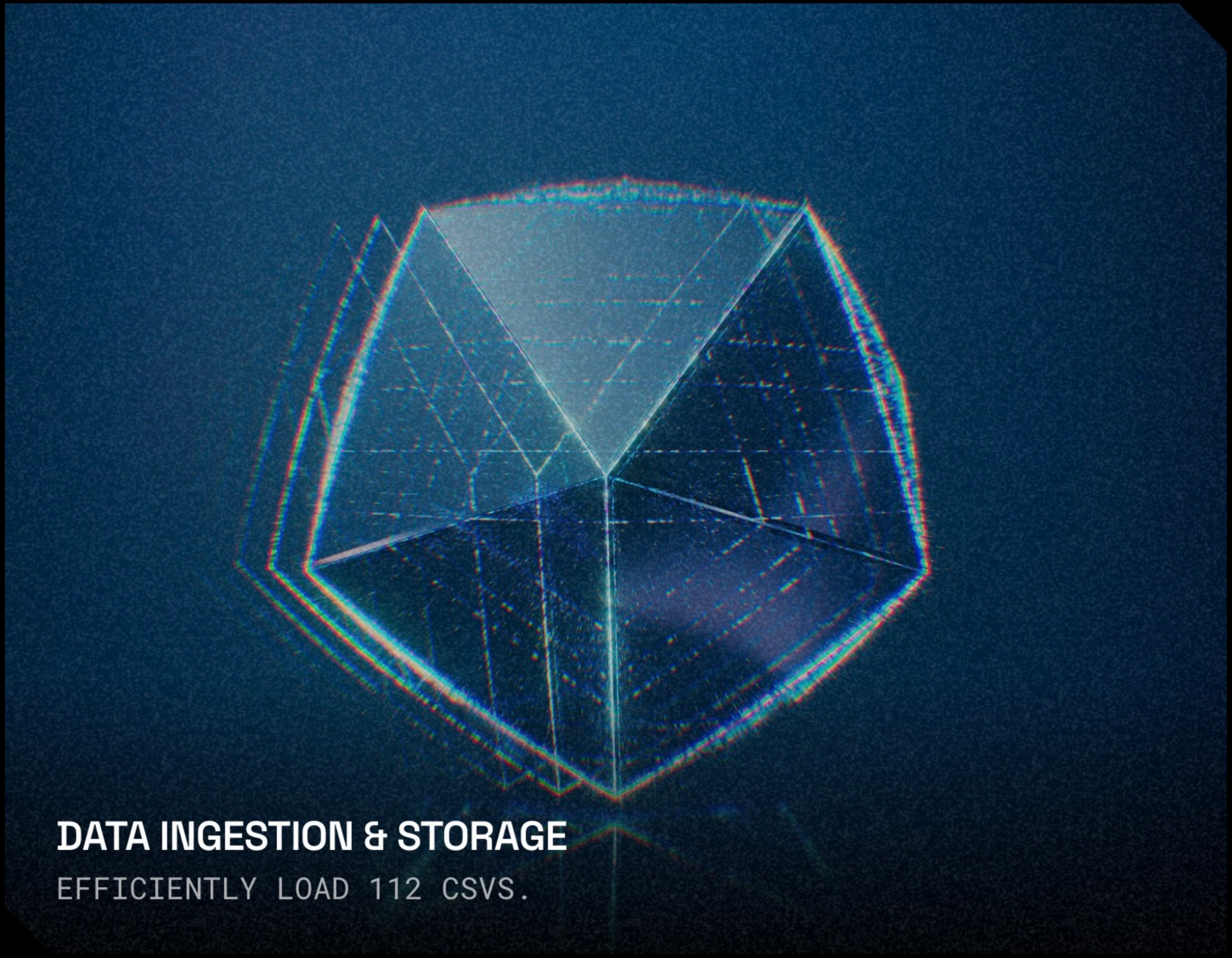
MACHINE LEARNING AND MODERN VOLATILITY MODELING

MACHINE LEARNING SURROGATES, LIKE TRANSFORMERS, OFFER A DATA-DRIVEN ALTERNATIVE TO CLASSICAL STOCHASTIC MODELS BY LEARNING THE NON-LINEAR STRUCTURE OF VOLATILITY DIRECTLY FROM MICROSTRUCTURE FEATURES. LEVEL-2 LOB DATA ENCODES RICH INFORMATION ABOUT MARKET INTENT AND LIQUIDITY SHIFTS. TRANSFORMERS' SELF-ATTENTION MECHANISMS ARE UNIQUELY CAPABLE OF IDENTIFYING TEMPORALLY SPARSE BUT INFORMATIONALLY RICH PATTERNS IN THIS STREAM, POTENTIALLY OUTPERFORMING TRADITIONAL MODELS IN BOTH ACCURACY AND RESPONSIVENESS TO CHANGING MARKET REGIMES.

Show UI

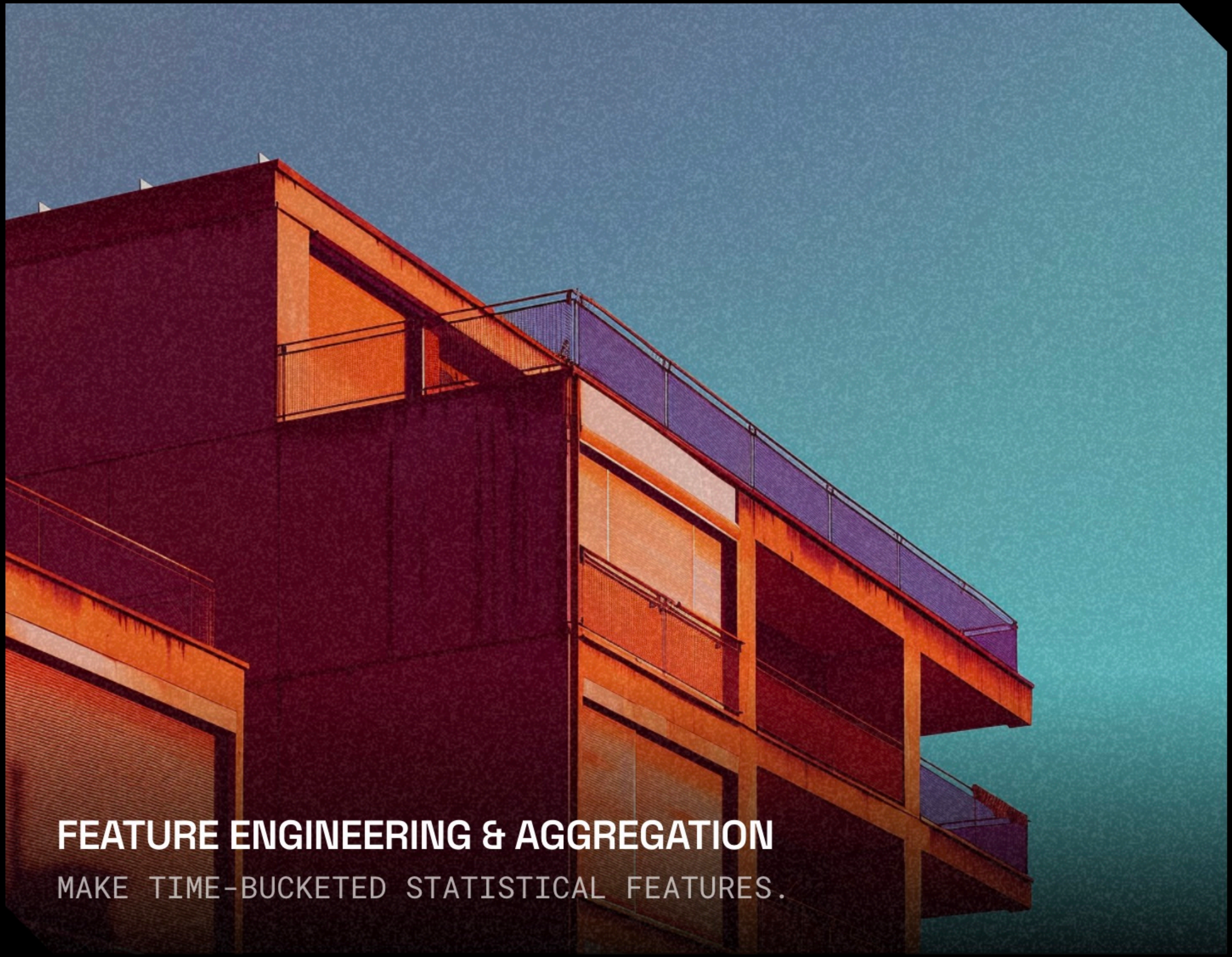
THE DATASET

LEVEL 2 LIMIT ORDER BOOK (LOB) SHOWS ALL BID AND ASK ORDERS AT MULTIPLE PRICE LEVELS, NOT JUST THE BEST BID AND ASK. IT PROVIDES FULL MARKET DEPTH, SHOWING HOW MUCH LIQUIDITY EXISTS AT EACH PRICE LEVEL.



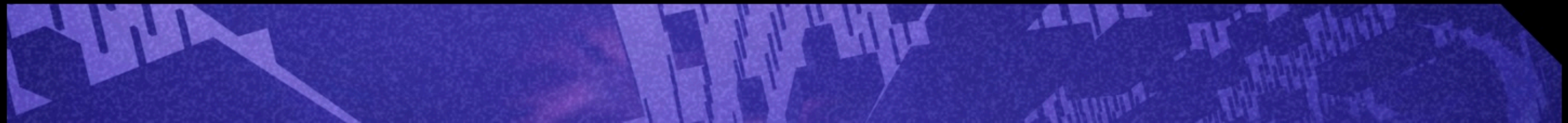
DATA INGESTION & STORAGE

EFFICIENTLY LOAD 112 CSVS.



FEATURE ENGINEERING & AGGREGATION

MAKE TIME-BUCKETED STATISTICAL FEATURES.

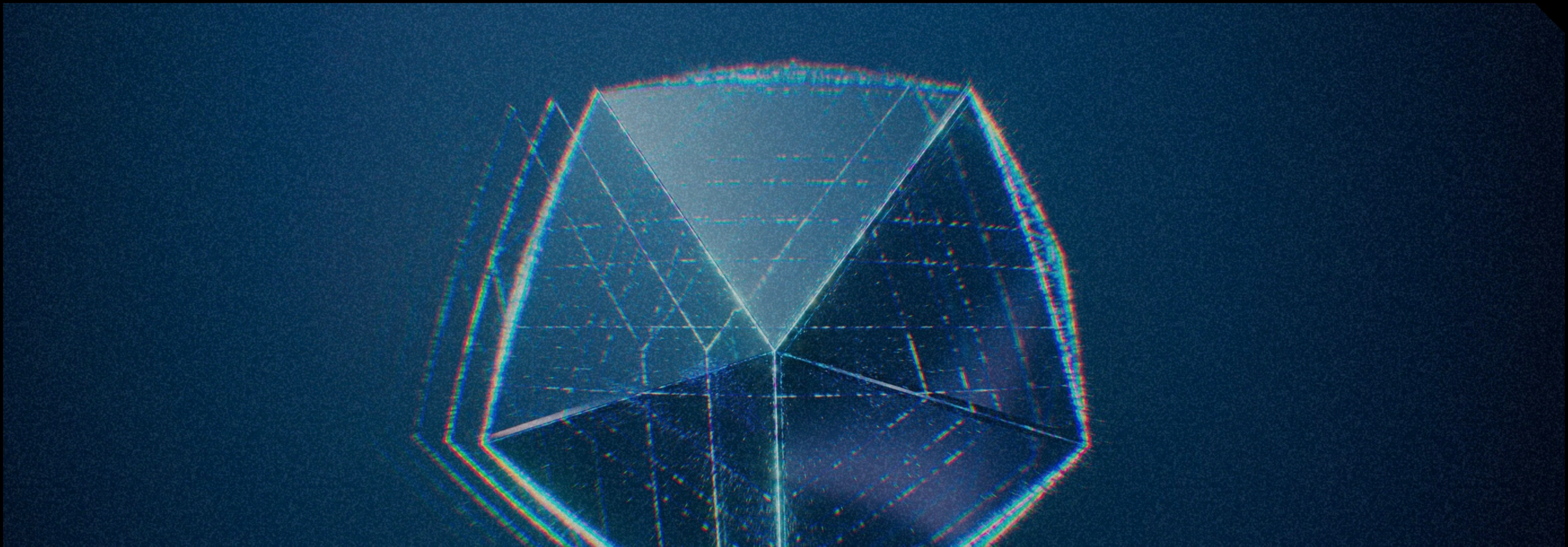


Show UI

DATA INGESTION & STORAGE

TAN_1CLAS_800095

TO EFFICIENTLY PREPARE OUR 112 INDIVIDUAL CSVS FOR ANALYSIS, WE FIRST CONSOLIDATE THEM INTO A SINGLE COLUMNAR DATASET. THIS ENSURES FAST I/O AND DOWNSTREAM PROCESSING.



Show UI

COMBINING CSVS

WE USE **POLARS LAZY SCAN** WITH AN EXPLICIT SCHEMA TO DEFER PARSING UNTIL COLLECTION. THIS AVOIDS EAGER MEMORY ALLOCATION AND SPEEDS UP INITIAL INGESTION.

PARQUET SERIALIZATION

ONCE COLLECTED, WE WRITE THE COMBINED DATAFRAME TO PARQUET WITH SNAPPY COMPRESSION. THE COLUMNAR FORMAT AND LIGHT COMPRESSION YIELD FASTER RELOADS AND REDUCED DISK USAGE.

PERFORMANCE BENCHMARKS & NEXT STEPS

MEASURE AND PRESENT CSV-VS-PARQUET LOAD TIMES AND STORAGE FOOTPRINTS. GOING FORWARD, CONSIDER DELTA LAKE FOR VERSIONING AND INCREMENTAL UPDATES TO ACCOMMODATE LIVE DATA FEEDS.

CTION VOLATILITY PREDICTION V

GROUP 07

AYUSH SINGH | KYLIE HARYONO | CHRISTY LEE | TOBIT LOUIS | ZICHUN HAN

GO BACK TO TOP

Show UI

FEATURE ENGINEERING & AGGREGATION

TAN_1CLAS_800095

TO TRANSFORM RAW ORDER-BOOK TICKS FOR OUR SELECTED STOCKS INTO MODELING-READY INPUTS, WE COMPUTE A COMPREHENSIVE SET OF PRICE, VOLUME, AND STATISTICAL FEATURES, THEN AUTOMATICALLY REFINE THEM.

Show UI

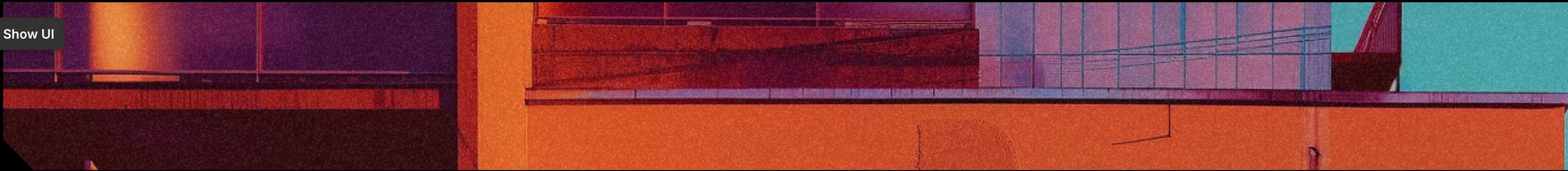
SNAPSHOT-LEVEL & DERIVED PRICE/VOLUME FEATURES

WE CALCULATE INSTANTANEOUS METRICS SUCH AS MID_PRICE, SPREAD, IMBALANCE, BOOK_PRESSURE, NORMALIZED_SPREAD, OBI_L2, LOB_ENTROPY (AND ITS NORMALIZED FORM), WAP, AND LOG_WAP_RETURN. THESE CAPTURE THE REAL-TIME STATE OF LIQUIDITY AND PRICE PRESSURE AT EACH TICK.

ROLLING & FUTURE-TARGETED STATISTICS

WITHIN EACH SECOND BUCKET, WE DERIVE LOG_RETURN AND REALIZED_VOLATILITY OVER A 30-TICK WINDOW, BIPOWER_VAR, AND A 30-TICK AHEAD TARGET RV_FUTURE. WE ALSO ADD ROLLING_VOL_30 AND ROLLING_IMBALANCE_MEAN_30 TO SUMMARIZE RECENT DYNAMICS, PLUS SIN/COS TRANSFORMS OF SECONDS_IN_BUCKET TO ENCODE INTRADAY SEASONALITY.

Show UI



AUTOMATED FEATURE SELECTION & REDUCTION

AFTER TYPE-CASTING TO FLOAT32/INT32, WE APPLY `VARIANCETHRESHOLD` TO DROP ZERO-VARIANCE COLUMNS AND THEN REMOVE ONE OF ANY PAIR WITH SPEARMAN > 0.98. THIS PRUNES ABOUT 30% OF FEATURES, LEAVING ~20 HIGH-INFORMATION PREDICTORS IN `FE30STOCKS.PARQUET`.O4-MINI-HIGH

ON VOLATILITY PREDICTION VO

GROUP 07

AYUSH SINGH | KYLIE HARYONO | CHRISTY LEE | TOBIT LOUIS | ZICHUN HAN

GO BACK TO TOP

Show UI

DATA PREPROCESSING & K-MEANS CLUSTERING

FILTER OUT STOCKS & GROUP SIMILAR TRADING PATTERNS.

PREDICTIVE EVALUATION & STOCK SELECTION

RANK STOCKS BY COMBINED METRICS.

REDUCED ON-DISK FOOTPRINT BY ~80% USING SNAPPY COMPRESSED PARQUET.

GENERATED 28 FEATURES, PRUNING VIA VARIANCE THRESHOLDING AND SPEARMAN FILTERING TO RETAIN ~20 ROBUST PREDICTORS.

FILTER STOCKS WITH IQR FILTERING, ACHIEVED A 0.62 SILHOUETTE SCORE AT K=5, AND SELECTED TOP 30.

DATA PREPROCESSING & K-MEANS CLUSTERING

TO FOCUS ON REPRESENTATIVE STOCKS AND UNCOVER LATENT GROUPINGS, WE FILTER OUT VOLATILITY OUTLIERS AND THEN CLUSTER ON PER-STOCK META-FEATURES.

Show UI

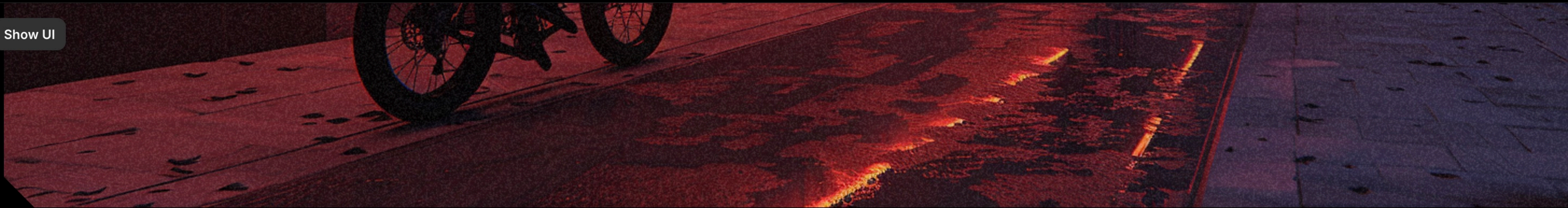
VOLATILITY-BASED FILTERING

COMPUTE EACH STOCK’S MEAN REALIZED VOLATILITY, DERIVE IQR BOUNDS WITH A MULTIPLIER OF 1.5, AND EXCLUDE THOSE BEYOND THESE LIMITS. THIS REMOVES ANOMALOUS INSTRUMENTS WITH EXTREME VARIABILITY.

META-FEATURE STANDARDIZATION

AGGREGATE THE TIME-BUCKET FEATURES BY STOCK INTO A META-FEATURE TABLE, THEN APPLY STANDARDSCALER. STANDARDIZING ENSURES THAT HIGH-VARIANCE METRICS DON’T DOMINATE THE CLUSTERING GEOMETRY.

Show UI



K-MEANS EXECUTION & INSIGHTS

RUN K-MEANS WITH FIVE CLUSTERS (RANDOM_STATE=42). ANALYZE CLUSTER CENTROIDS TO LABEL GROUPS (E.G., “HIGH-VOLATILITY, LOW-LIQUIDITY”). VISUALIZE VIA PCA TO VALIDATE SEPARATION AND GUIDE FURTHER CLUSTER-SPECIFIC MODELING.

CTION VOLATILITY PREDICTION

GROUP 07

Show UI

PREDICTIVE EVALUATION & STOCK SELECTION

FINALLY, WE FORECAST VOLATILITY USING PAST FEATURES, BENCHMARK AGAINST A NAÏVE MEAN FORECAST, AND RANK STOCKS BY A COMBINED METRIC FOR SELECTION.

Show UI

EXPANDING-WINDOW REGRESSION

FOR EACH STOCK WITH AT LEAST 10 PERIODS, WE REGRESS REALIZED VOLATILITY ON ITS LAGGED FEATURE SET IN AN EXPANDING-WINDOW MANNER. THIS SIMULATES A ROLLING REAL-TIME FORECASTING REGIME.

QLIKE BASELINE

USE THE HISTORICAL MEAN OF REALIZED VOLATILITY AS A BENCHMARK FORECAST. WE COMPUTE THE QLIKE LOSS TO MEASURE HOW THIS SIMPLE PREDICTOR PERFORMS COMPARED TO OUR REGRESSION MODELS.

Show UI

COMBINED SCORING & TOP PICKS

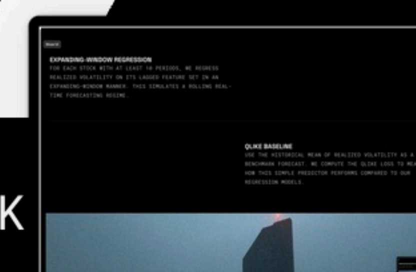
DEFINE A SCORE = 0.5·R^2 − 0.5·QLIKE FOR EACH STOCK. SORTING BY THIS METRIC YIELDS OUR TOP 30 CANDIDATES. THIS BALANCES EXPLANATORY POWER WITH BASELINE FORECASTING ERROR, ENSURING ROBUST SELECTION.

TILITY PREDICTION VOLATILITY P

GROUP 07

AYUSH SINGH | KYLIE HARYONO | CHRISTY LEE | TOBIT LOUIS | ZICHUN HAN

GO BACK



Show UI

REDUCED ON-DISK FOOTPRINT BY ~80% USING SNAPPY COMPRESSED PARQUET.

GENERATED 28 FEATURES, PRUNING VIA VARIANCE THRESHOLDING AND SPEARMAN FILTERING TO RETAIN ~20 ROBUST PREDICTORS.

FILTER STOCKS WITH IQR FILTERING, ACHIEVED A 0.62 SILHOUETTE SCORE AT K=5, AND SELECTED TOP 30.

CANDIDATE MODELS

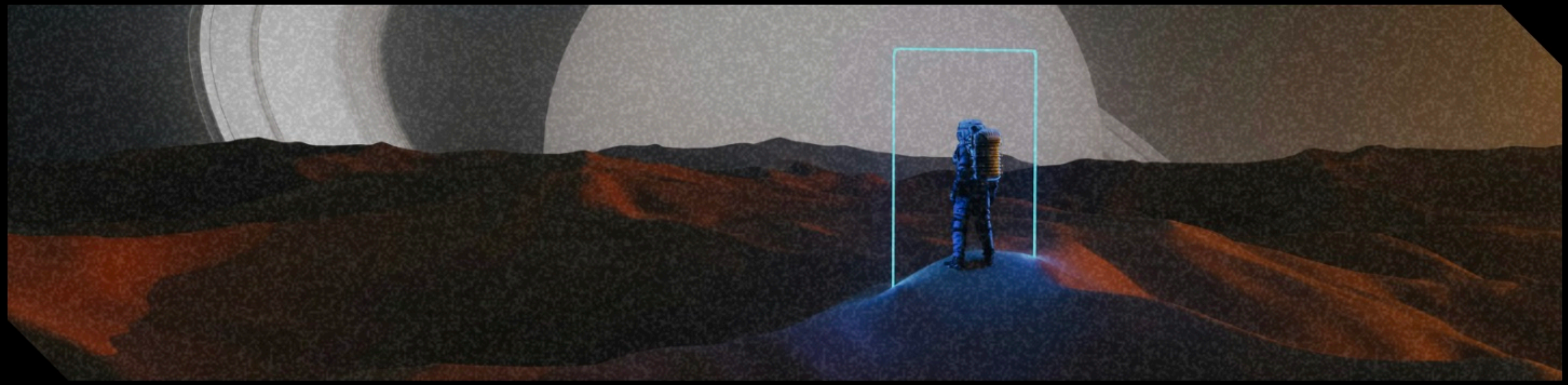
WE DEVELOPED OVER 15 MODELS AND HAVE STRATEGICALLY SELECTED THE FOLLOWING FOUR FOR DETAILED ANALYSIS. THIS SELECTION INCLUDES OUR TWO TOP-PERFORMING MODELS—TRANSFORMER AND LSTM—ALONG WITH A BASELINE MODEL DEMONSTRATING SOLID FOUNDATIONAL PERFORMANCE (WEIGHTED LEAST SQUARES, WLS), AND A HIGH-PERFORMING MODEL WITHIN ITS COMPLEXITY CLASS (RANDOM FOREST), WHICH CONSISTENTLY OUTPERFORMED OTHER MODELS OF COMPARABLE COMPLEXITY.

Show UI

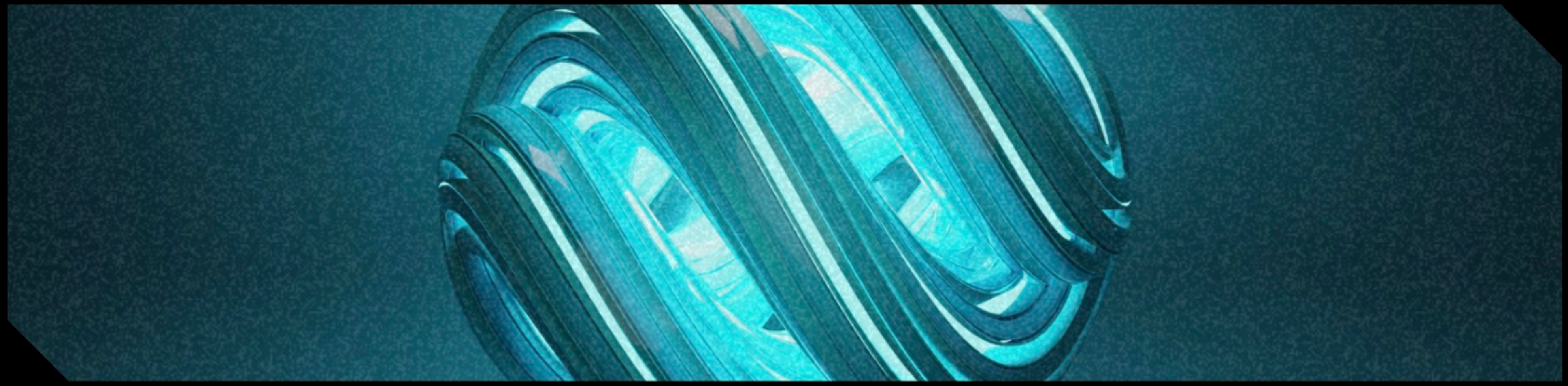
.1 **WEIGHTED LEAST SQUARES**

.2 **RANDOM FOREST**

.3 **LSTM**



.4 **TRANSFORMER**



Show UI

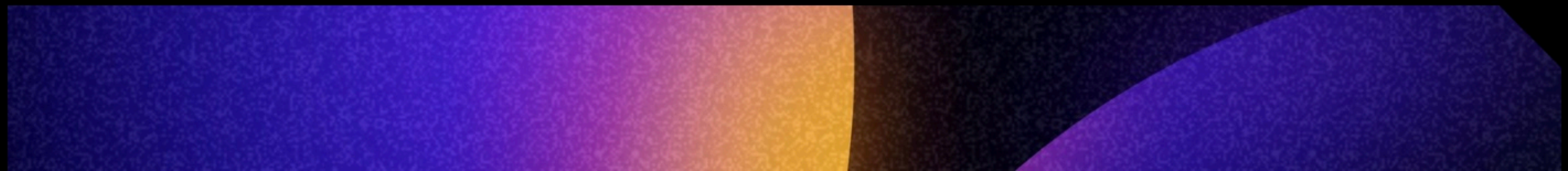
CANDIDATE MODELS

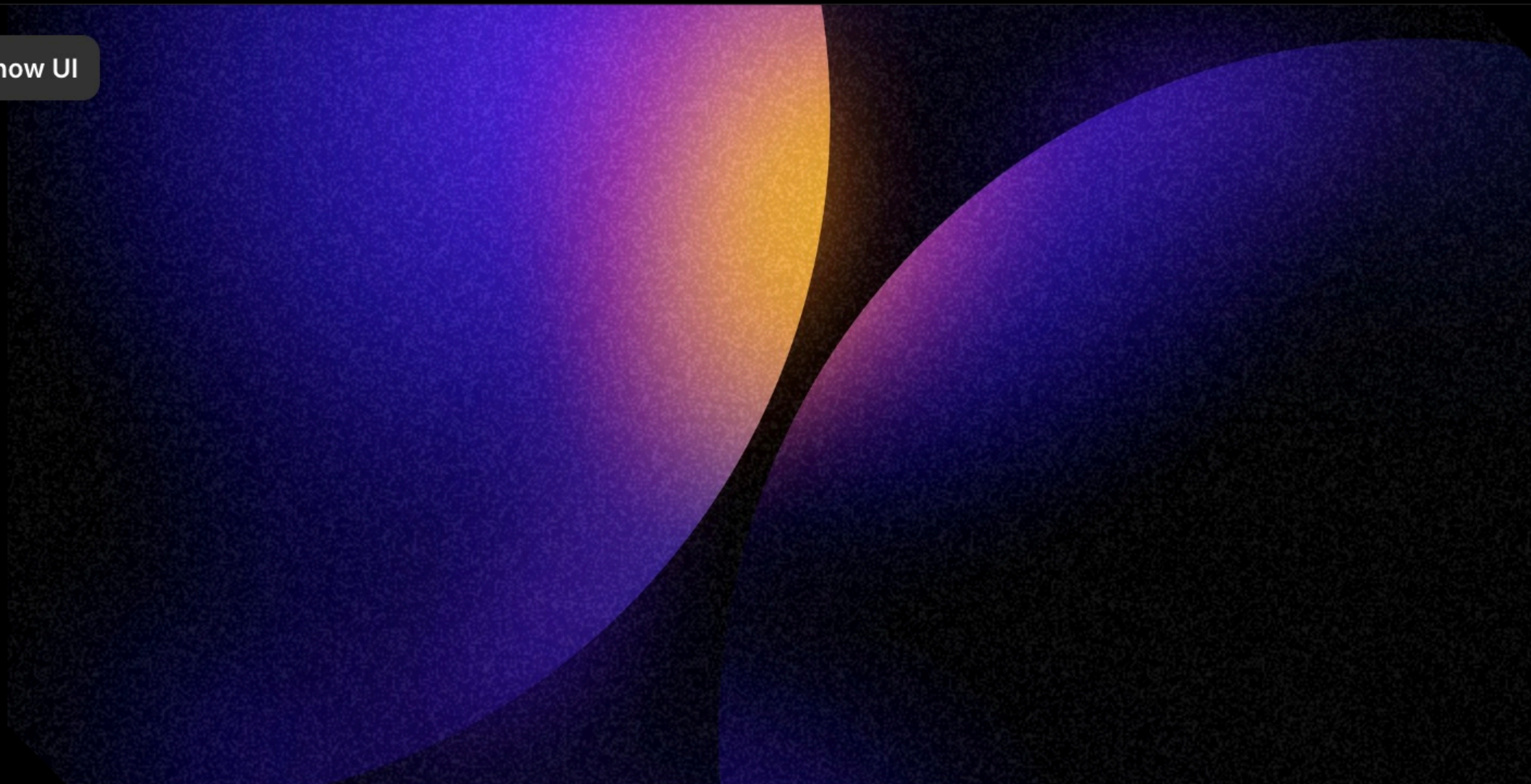
Show UI

TRANSFORMER

WE IMPLEMENTED A LIGHTWEIGHT TRANSFORMER ENCODER FOR VOLATILITY FORECASTING BY PROJECTING EACH 30-STEP LOB SEQUENCE INTO A **64-DIMENSIONAL EMBEDDING**, THEN STACKING **TWO ENCODER BLOCKS—EACH WITH FOUR-HEAD SELF-ATTENTION** (KEY/QUERY DIM=64), A **4× FEED-FORWARD EXPANSION TO 256 UNITS**, AND **RESIDUAL + LAYER-NORM CONNECTIONS**. AFTER **GLOBAL AVERAGE POOLING**, A SINGLE LINEAR OUTPUT HEAD PRODUCES THE FORECAST, TOTALING ≈**100 K PARAMETERS**. INPUTS AND LOG-TRANSFORMED TARGETS WERE MINMAX-SCALED; **TRAINING USED ADAM (1×10^{-3}) WITH MSE LOSS ON A STRICT 80/10/10 CHRONOLOGICAL SPLIT, BATCH SIZE 32, AND EARLY STOPPING (PATIENCE = 15) OVER 50 EPOCHS.** THE MODEL REQUIRED ROUGHLY **500 GPU-HOURS TO TRAIN** AND DELIVERED STRONG OUT-OF-SAMPLE RMSE, R^2, AND QLIKE PERFORMANCE.

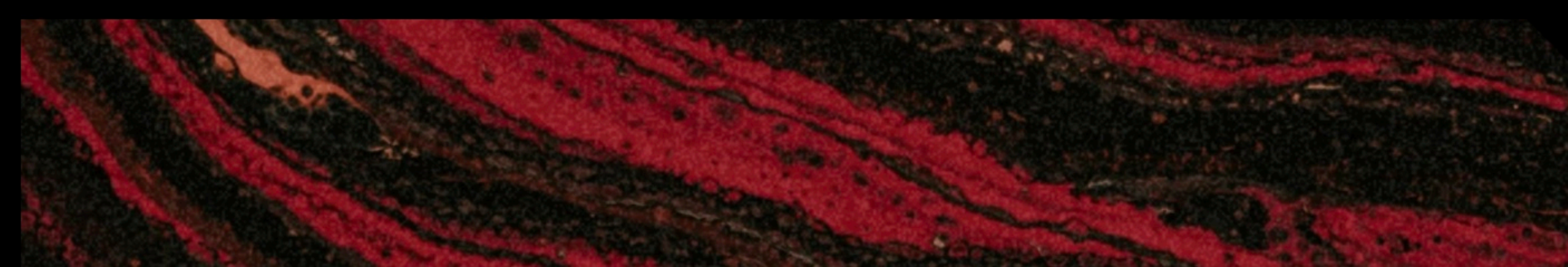
1334758950_SINCE_1932
XXL V7868376M





LSTM

WE BUILT A **TWO-LAYER STACKED LSTM**—FIRST WITH **64 UNITS RETURNING** ALL 30 STEPS (**~25 088 PARAMETERS**), FOLLOWED BY 20 % DROPOUT, THEN A **32-UNIT LSTM** (**~12 416 PARAMETERS**) WITH ANOTHER **20 % DROPOUT**—TOPPED BY A COMPACT FEED-FORWARD HEAD (**16-NODE DENSE + SINGLE LINEAR OUTPUT**, ~545 PARAMETERS), FOR **~38 K TRAINABLE WEIGHTS**. INPUTS WERE MIN-MAX–SCALED 30-STEP SEQUENCES OF ENGINEERED LOB FEATURES AND THE TARGET WAS LOG-TRANSFORMED VOLATILITY. WE ENFORCED A STRICT CHRONOLOGICAL SPLIT **(80 % TRAIN, 10 % VAL, 10 % TEST)** TO ELIMINATE LOOK-AHEAD BIAS, **TRAINED WITH ADAM AT A 1 × 10^{-4} LEARNING RATE, 128-SAMPLE BATCHES, AND EARLY STOPPING (PATIENCE = 5) FOR UP TO 50 EPOCHS.** OUT-OF-SAMPLE RMSE, R^2, AND QLIKE METRICS DEMONSTRATE STRONG SEQUENCE LEARNING AND STABLE VOLATILITY FORECASTS.

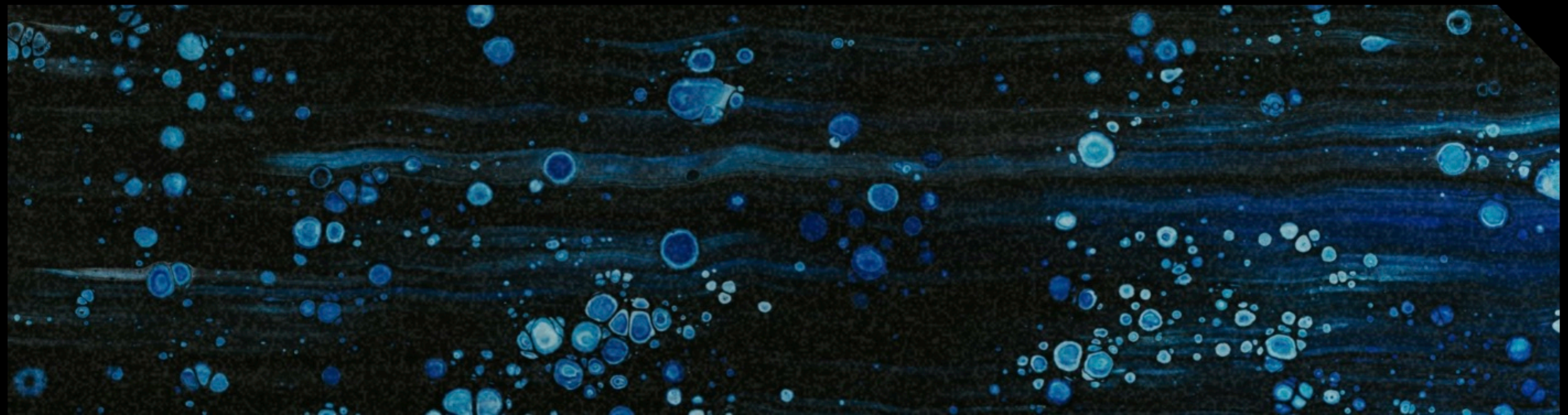


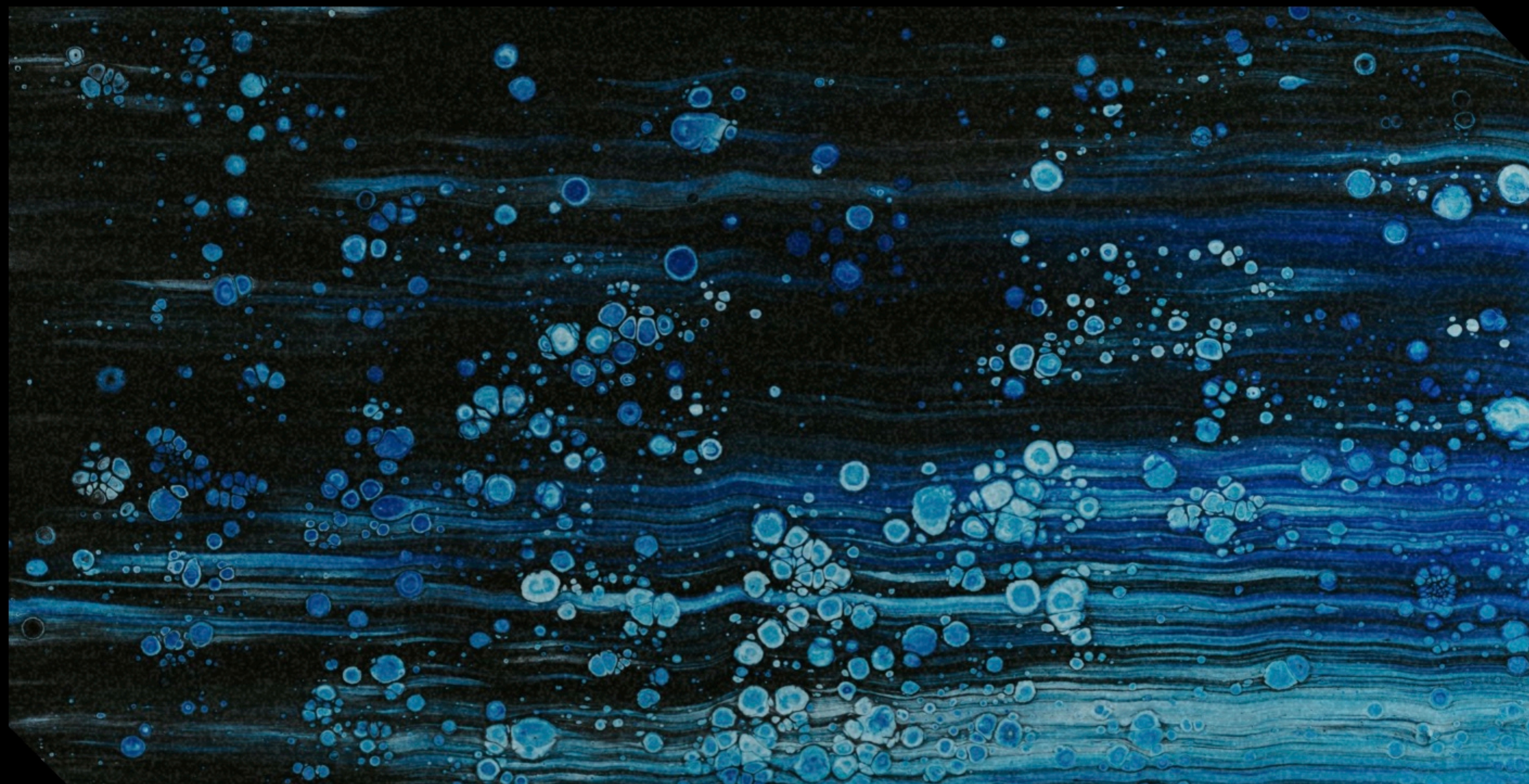
Show UI

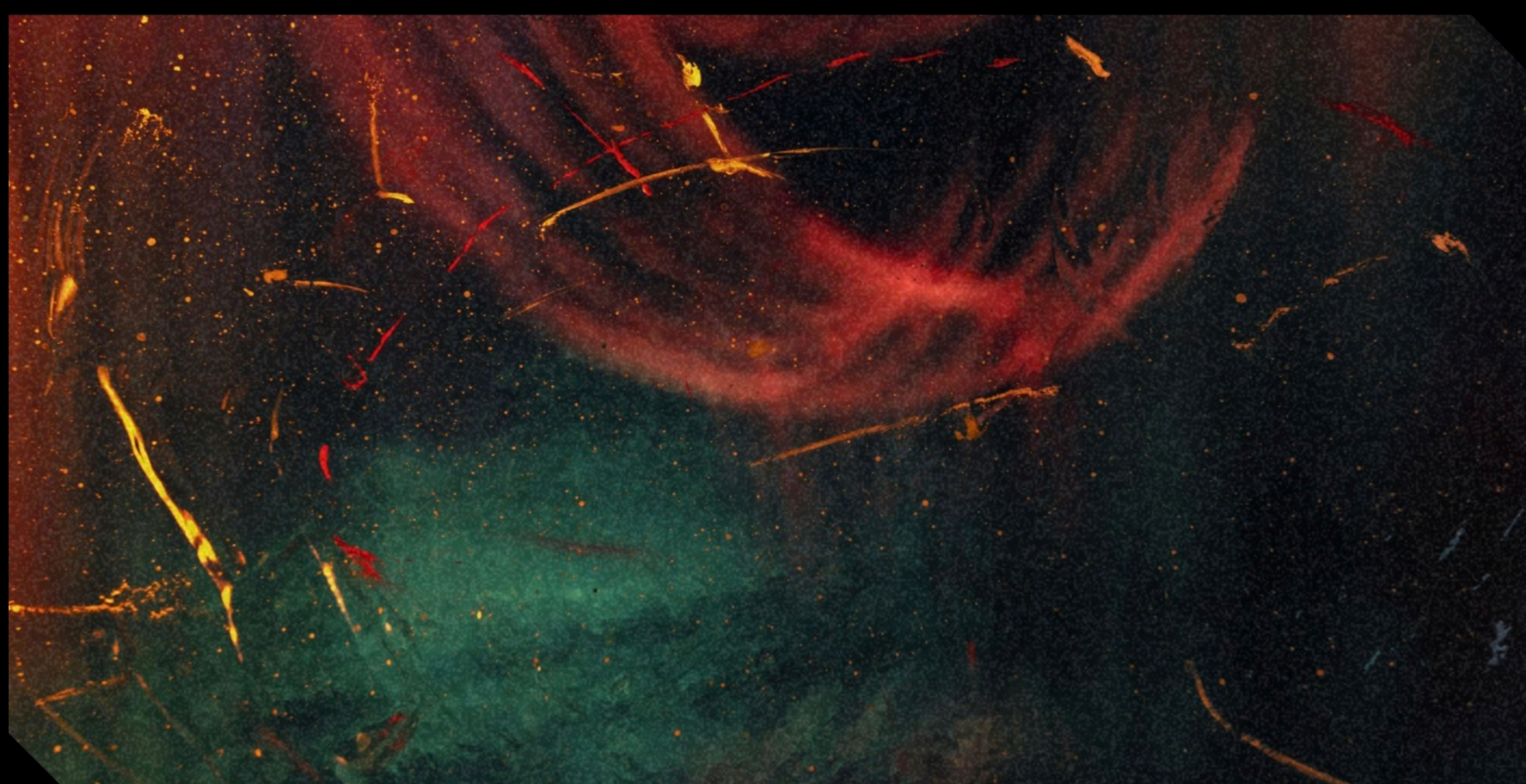
RANDOM FOREST

WE USED AN ENSEMBLE OF **500 TREES** GROWN TO **FULL DEPTH**, SAMPLING THE **SQUARE ROOT** OF FEATURES AT EACH SPLIT AND ENFORCING **A MINIMUM OF THREE SAMPLES PER LEAF**. **BOOTSTRAP AGGREGATION** WAS USED TO DECORRELATE TREES AND REDUCE VARIANCE. THE VOLATILITY TARGET WAS LOG-TRANSFORMED TO STABILIZE ITS HEAVY-TAILED DISTRIBUTION, AND WE **PARTITIONED THE DATA CHRONOLOGICALLY—80% FOR TRAINING, 10% FOR VALIDATION, 10% FOR TESTING**—TO ELIMINATE LOOK-AHEAD BIAS. TRAINING RAN IN **PARALLEL** ACROSS ALL CPU CORES FOR EFFICIENCY, AND WE MONITORED PROGRESS IN REAL TIME. OUT-OF-SAMPLE PERFORMANCE WAS MEASURED WITH **ROOT-MEAN-SQUARE ERROR** AND **QLIKE LOSS** (WITH CLIPPING), DEMONSTRATING THE MODEL'S ABILITY TO CAPTURE COMPLEX NONLINEAR INTERACTIONS IN ORDER-BOOK DYNAMICS.

TAN_1CLAS_000095







WEIGHTED LEAST SQUARES (WLS)

WE IMPLEMENTED A **HETEROSKEDASTICITY-AWARE** WEIGHTED LEAST SQUARES MODEL IN STATSMODELS TO FORECAST 30-TICK-AHEAD REALIZED VOLATILITY. **OBSERVATION WEIGHTS** WERE SET AS THE INVERSE OF A LONG-WINDOW ROLLING VARIANCE TO COUNTER TIME-VARYING NOISE. USING A **CHRONOLOGICAL 80/20 TRAIN-TEST SPLIT** ACROSS ALL 30 STOCKS, THE MODEL DELIVERES A STRONG BASELINE OUT-OF-SAMPLE R^2 AND QLIKE LOSS.

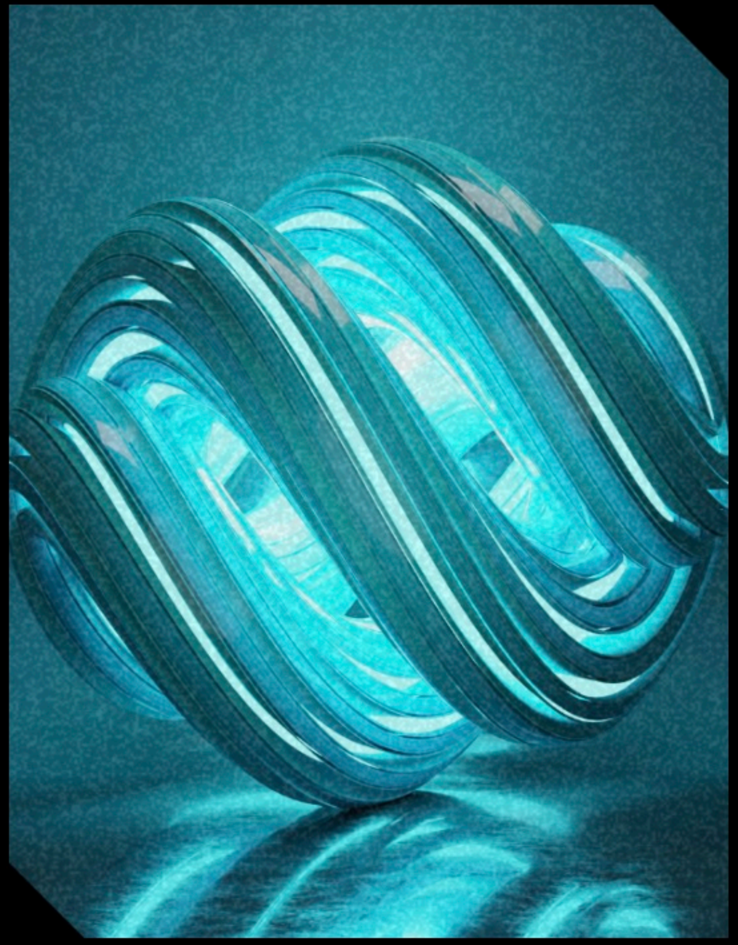
Show UI

TAN_1CLAS_800095

OUR RECOMMENDED MODEL: TRANSFORMER

Show UI

ABOUT THE MODEL



TRANSFORMER

MODEL STRUCTURE

THE TRANSFORMER PROCESSES EACH 30-STEP WINDOW OF ORDER-BOOK FEATURES THROUGH TWO ATTENTION LAYERS THAT AUTOMATICALLY FOCUS ON THE MOST INFORMATIVE MOMENTS. AFTER EACH ATTENTION PASS, A COMPACT NEURAL BLOCK REFINES THOSE SIGNALS, AND FINALLY A GLOBAL AVERAGE POOLS ACROSS TIME TO PRODUCE A SINGLE VOLATILITY FORECAST. THE MODEL TRAINS ON MORE THAN 100,000 PARAMETERS AND REQUIRES A MINIMUM OF 300HRS OF GPU TRAINING TIME.

PREDICTIVE ACCURACY

OUR MODEL CONSISTENTLY OUTPERFORMS ALTERNATIVES, ACHIEVING AN **OUT-OF-SAMPLE R^2 OF 0.75** AND A **QLIKE LOSS OF 0.06**. THESE METRICS MEAN IT EXPLAINS 75% OF VOLATILITY MOVEMENTS AND DELIVERS HIGHLY RELIABLE FORECASTS, OUTPERFORMING OUR LSTM AND WEIGHTED REGRESSION MODELS.

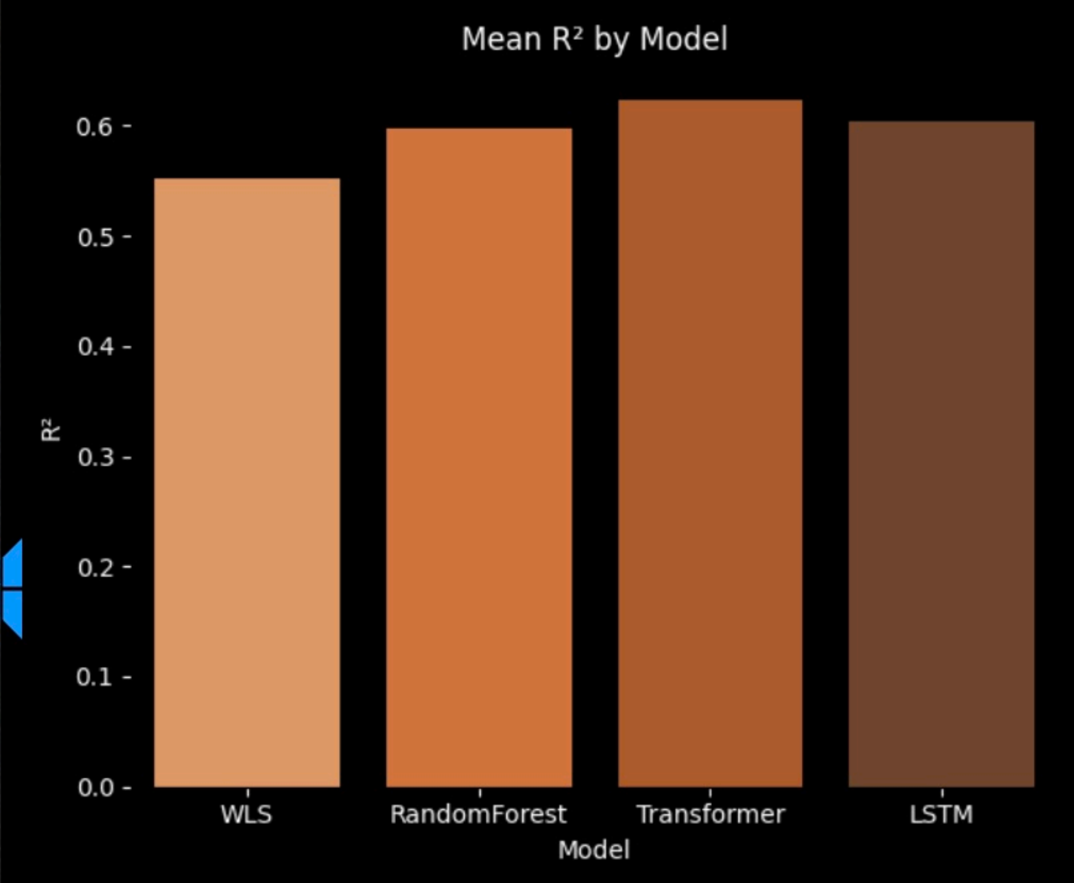
Show UI

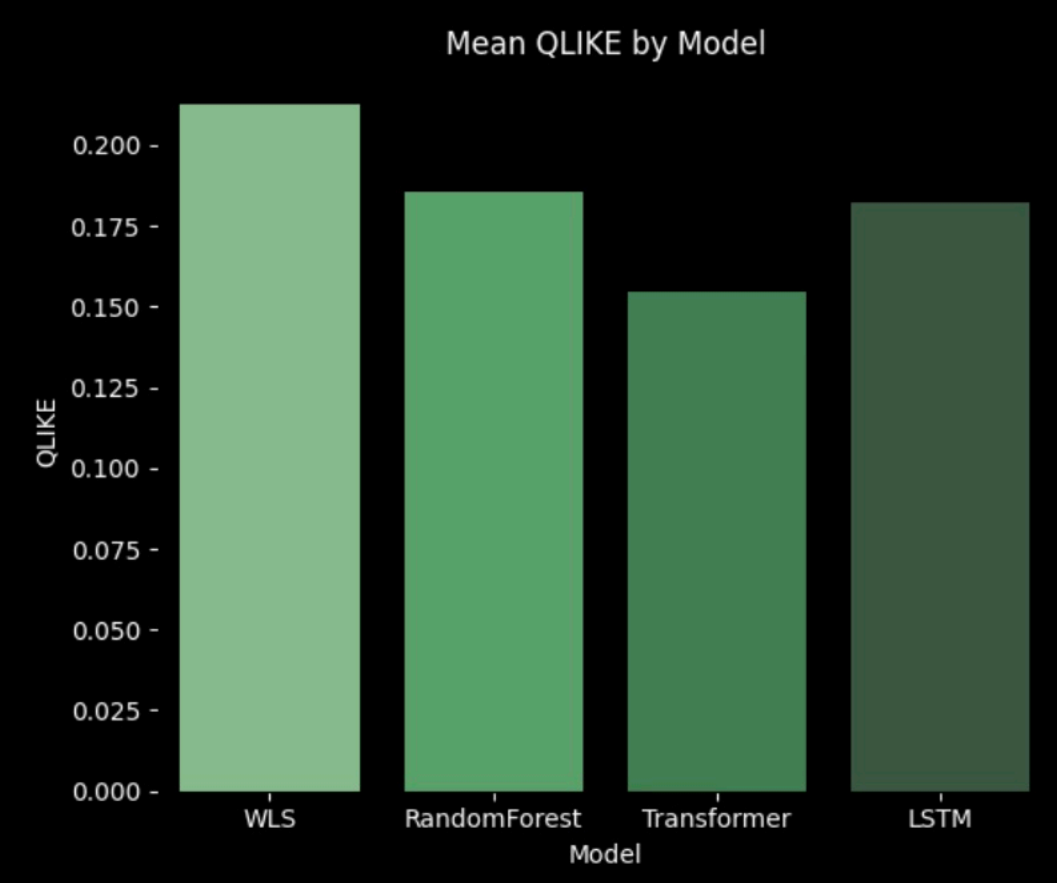
COMPARISON METRICS

WE EVALUATED THE MODELS USING OUT-OF-SAMPLE R-SQUARED, QLIKE LOSS, AND MSE. ADDITIONALLY, WE TRAINED THE MODELS WITH A CUSTOM QLIKE LOSS FUNCTION TO FACILITATE HYPERPARAMETER TUNING.

R-SQUARED (R^2)

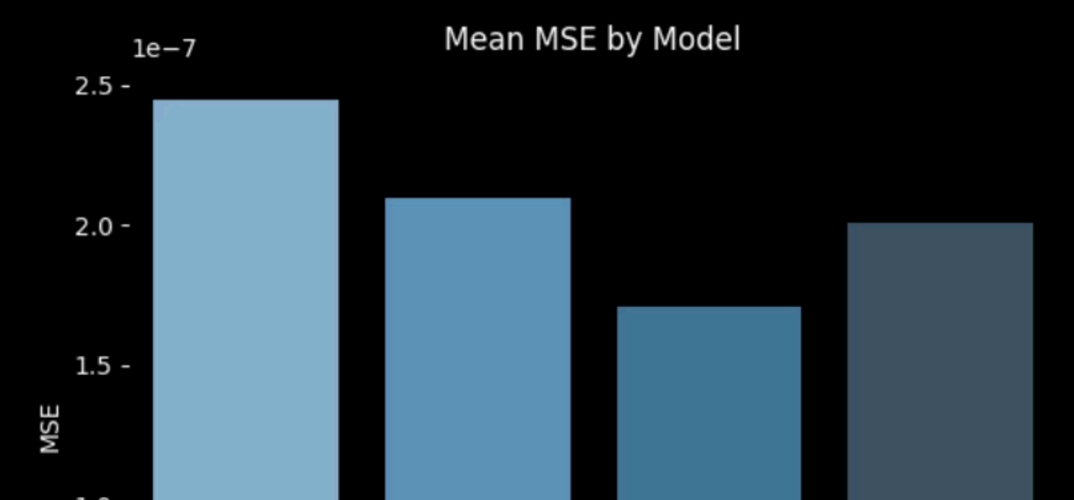
MEASURES HOW WELL THE MODEL PREDICTS UNSEEN DATA COMPARED TO A NAIVE BASELINE.





QLIKE LOSS

PENALIZES INACCURATE VOLATILITY FORECASTS, ESPECIALLY WHEN PREDICTED VARIANCES ARE TOO LOW.



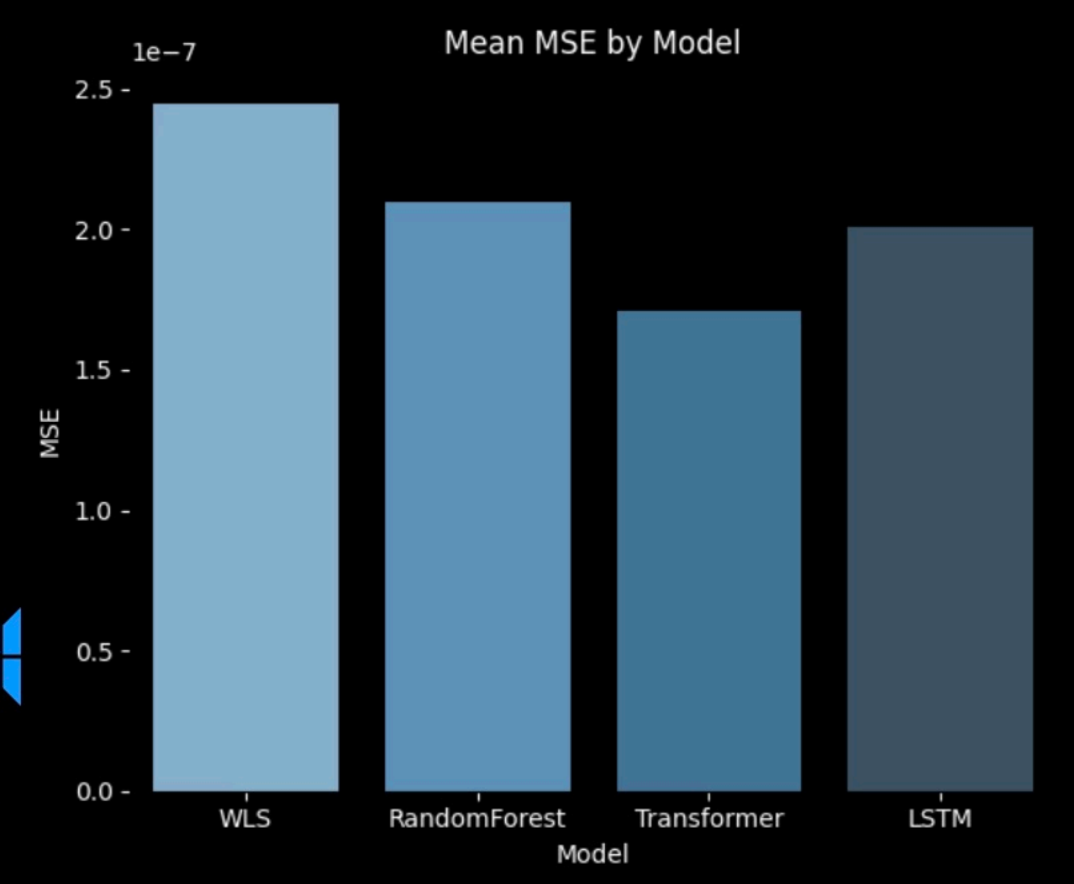
Show UI

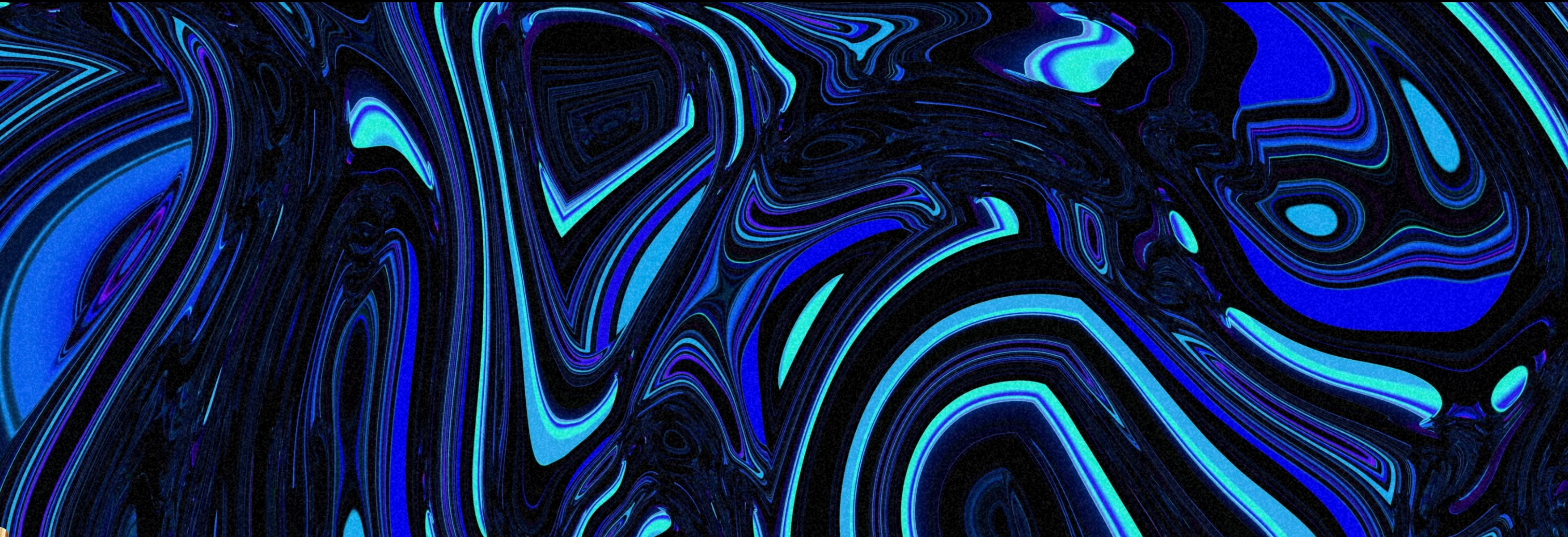
COMPARISON METRICS

WE EVALUATED THE MODELS USING OUT-OF-SAMPLE R-SQUARED, QLIKE LOSS, AND MSE. ADDITIONALLY, WE TRAINED THE MODELS WITH A CUSTOM QLIKE LOSS FUNCTION TO FACILITATE HYPERPARAMETER TUNING.

MSE

CALCULATES THE AVERAGE SQUARED DIFFERENCE BETWEEN PREDICTED AND ACTUAL VALUES.





Show UI

LIMITATIONS
AND FUTURE IMPROVEMENT

LIMITED STOCK COVERAGE

WE DID NOT USE ALL 112 STOCKS, SO THE MODELS MAY SUFFER SAMPLE BIAS AND POOR GENERALISATION ACROSS ASSETS. IMPROVEMENT: TRAIN ON THE FULL UNIVERSE WITH STRATIFIED SAMPLING OR ASSET-AWARE MINI-BATCHES.

NON-STATIONARITY OF VOLATILITY

STANDARD FEATURE TRANSFORMS ASSUME A FIXED DISTRIBUTION, YET MARKET VOLATILITY SHIFTS OVER TIME. IMPROVEMENT: EMPLOY ONLINE NORMALIZATION OR ADAPTIVE DIFFERENCING TO TRACK EVOLVING DISTRIBUTIONS.

AGGREGATION INFORMATION LOSS

BUCKETING TICKS INTO FIXED INTERVALS DISCARDS FINE-GRAINED MICROSTRUCTURE DYNAMICS. IMPROVEMENT: INTRODUCE MULTI-RESOLUTION INPUTS OR CONTINUOUS-TIME EMBEDDINGS TO PRESERVE DETAIL.

TRANSFORMER INDUCTIVE BIAS

TRANSFORMER INDUCTIVE BIAS

A VANILLA TRANSFORMER TREATS ALL POSITIONS EQUALLY, LACKING TEMPORAL BIAS FOR FINANCIAL TIME SERIES. IMPROVEMENT: ADD LEARNABLE POSITIONAL EMBEDDINGS OR DECAY-WEIGHTED ATTENTION TO PRIORITISE RECENT DATA.

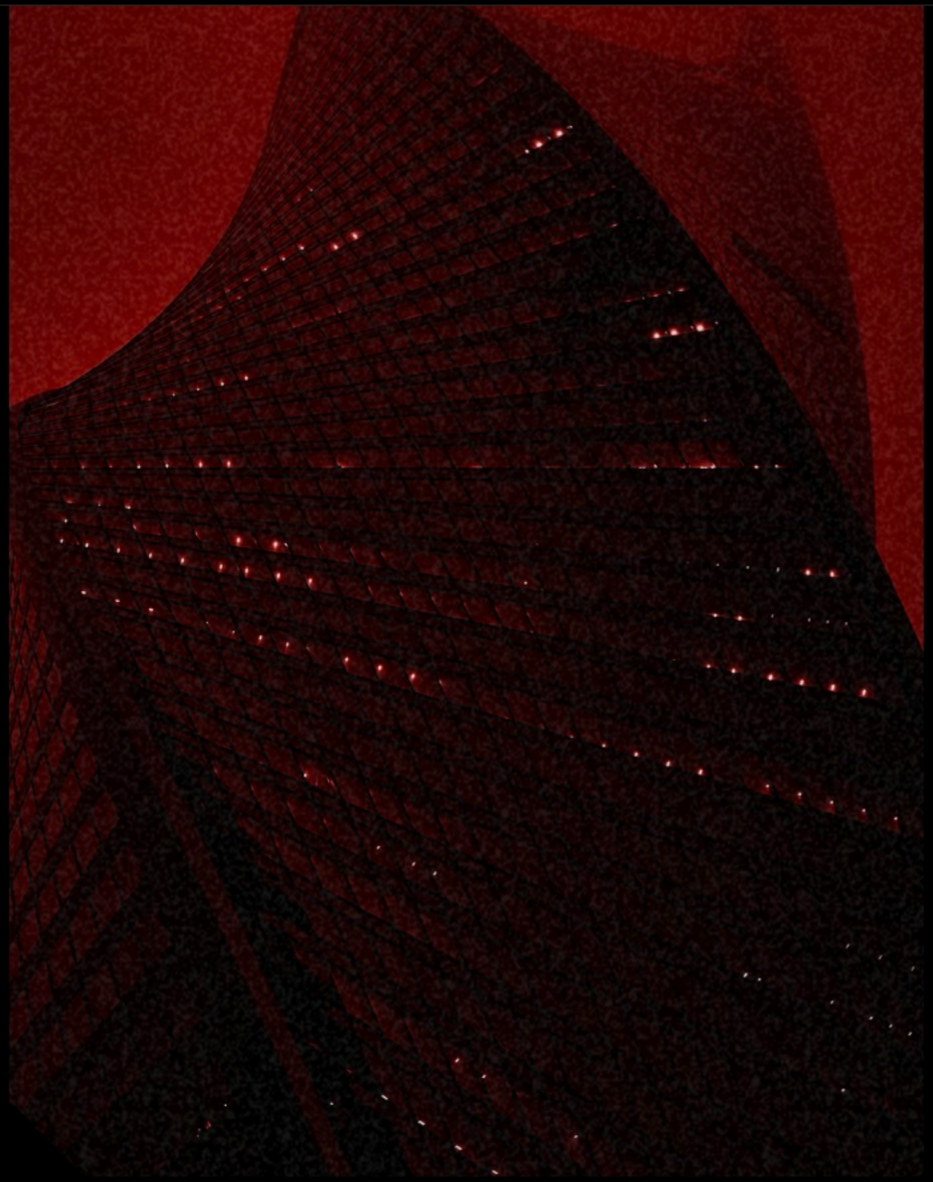
QUADRATIC ATTENTION COMPLEXITY

SELF-ATTENTION SCALES AS $O(T^2)$, LIMITING SEQUENCE LENGTH AND MODELLING OF LONG-TERM DEPENDENCIES. IMPROVEMENT: ADOPT EFFICIENT ATTENTION VARIANTS (E.G. LINFORMER, PERFORMER OR SPARSE ATTENTION).

CONCEPT DRIFT UNADDRESSED

POST-TRAINING REGIME SHIFTS DEGRADE PERFORMANCE OVER TIME WITHOUT ANY ADAPTATION MECHANISM. IMPROVEMENT: IMPLEMENT DRIFT DETECTION TRIGGERS AND ONLINE FINE-TUNING OR META-LEARNING UPDATES.

Show UI



PRACTICAL MARKET-MAKER FOCUS:
PRECISE VOLATILITY FORECASTS FOR PRICING,
HEDGING & INVENTORY MANAGEMENT.

LSTM & TRANSFORMER ROBUSTNESS
UNDER VOLATILE, STRESSED
MARKET CONDITIONS

**TRANSFORMER'S SUPERIOR FORECASTS ENABLE OPTIVER TO
PRICE OPTIONS PRECISELY, ADJUST QUOTES DYNAMICALLY, CAPTURE MORE SPREAD
& SUSTAIN LOW-LATENCY EXECUTION**

GROUP 07

AYUSH SINGH | KYLIE HARYONO | CHRISTY LEE | TOBIT LOUIS | ZICHUN HAN

GO BACK TO TOP