# 🧬 AI4Trade Final Execution Checklist (China + USA, Oct 2025 Forecast)

**Objective:** Predict bilateral trade flows (USD) at the HS6 level → aggregate to HS4 for submission, outperforming the naive latest-value baseline.

## 🗂️ A. Directory Layout

| Folder | Purpose | Key Contents |
| --- | --- | --- |
| /data/raw/ | Source Parquets (harmonized 2023–2025) | USA\_2023.parquet, CHN\_2024.parquet, etc. |
| /data/features/ | Engineered feature datasets & splits | features\_\*\_\*.parquet |
| /models/ | Model artifacts (optional dumps) | Per-model folders |
| /predictions/oof/ | Out-of-fold predictions | {model}\_{segment}\_h{2,3}\_final.parquet |
| /predictions/forecast/ | Raw model forecasts (Oct 2025) | {model}\_{segment}\_h{2,3}\_final.parquet |
| /predictions/merged/ | Merged model results for blending | forecast\_all\_models\_h{2,3}\_final.parquet |
| /predictions/final/ | Ensembles & submission files | blend\_weights\_final.csv, final\_forecast\_hs4\_final.parquet, submission\_final.csv |
| /logs/ | Validation metrics, JSON run logs | {run\_id}\_cv\_scores\_final.csv |
| /docs/ | Documentation | All strategy and schema files |

## 🧩 B. Phase-Wise Workflow Summary

| Phase | Notebook | Segment(s) | Horizon | Description | Input | Output (with \_final suffix) |
| --- | --- | --- | --- | --- | --- | --- |
| **1. Feature Engineering (done)** | 00\_feature\_engineering.ipynb | CHN + USA, Export/Import | — | Create HS6-level lag, MA, cross-flow, macro features | Harmonized Parquet files | /data/features/features\_CHN\_export.parquet, etc. |
| **2. Train/Test Split Creation** | 01\_make\_splits.ipynb | All 4 segments | h=2 for CHN, h=3 for USA | Build y\_target = y(t+h), split into train (non-null) & test (context @ t = Aug for CHN / Jul for USA) | Feature Parquets | 8 files total, e.g. /data/features/features\_CHN\_export\_train\_h2\_final.parquet |
| **3. Model Training — XGB-Tweedie** | 12\_xgb\_tweedie.ipynb | Each segment separately | h=2 (CHN), h=3 (USA) | Direct-horizon forecasting using Tweedie objective; validation per Horizon Policy folds (C1–C6, U1–U5) | Train/Test splits | /predictions/oof/xgb\_tweedie\_{segment}\_h{2,3}\_final.parquet, /predictions/forecast/xgb\_tweedie\_{segment}\_h{2,3}\_final.parquet, /logs/xgb\_tweedie\_{segment}\_h{2,3}\_cv\_scores\_final.csv |
| **4. Model Training — XGB-Log1p** | 13\_xgb\_log1p.ipynb | Same | Same | log1p-scaled regression (reg:squarederror) for smoother series | Train/Test splits | Analogous \_final OOF, forecast, log files |
| **5. Model Training — LGBM-RMSE** | 10\_lgbm\_rmse.ipynb | Same | Same | Standard regression (RMSE objective), leaf-wise splits | Train/Test splits | Analogous \_final files |
| **6. Validation + Weight Derivation** | 40\_blend\_weights.ipynb | Per segment | h=2/h=3 | Compute sMAPE per model (HS4 aggregate), compare vs naive baseline, derive inverse-sMAPE weights (floor = 0.10). | All 3 OOF files for a segment | /predictions/final/blend\_weights\_final.csv |
| **7. Segment-Level Ensemble** | 41\_make\_ensemble.ipynb | Each segment | Same | Weighted blend of 3 models → ensemble at HS6 → aggregate to HS4 | Model forecast files + weights | /predictions/final/ensemble\_{segment}\_h{2,3}\_hs6\_final.parquet, /predictions/final/ensemble\_{segment}\_h{2,3}\_hs4\_final.parquet |
| **8. Combine All Segments** | 42\_concat\_segments.ipynb | CHN E/I + USA E/I | h=2, 3 | Concatenate 4 HS4-level ensembles into single national-level file | 4 ensemble HS4 files | /predictions/final/final\_forecast\_hs4\_final.parquet |
| **9. Submission File** | 50\_make\_submission.ipynb | All | — | Format per OEC: "Country1","Country2","ProductCode","TradeFlow","Value"; include only top 20 partners ≥ 200 HS4 lines | Final forecast file | /predictions/final/submission\_final.csv |
| **10. Documentation & Review** | — | — | — | Verify sMAPE < naive, check logs, archive run metadata, and write 2-page method summary | Logs + outputs | /docs/Method\_Description\_Final.docx |

## 📊 C. Validation Fold Templates (Direct Horizon Setup)

### 🇨🇳 China (h = 2)

| Fold | Train ≤ | Gap (1 mo) | Validate (t+2) | Comment | Weight |
| --- | --- | --- | --- | --- | --- |
| C1 | 2024-05-01 | Jun 2024 | 2024-08-01 | Early H2 anchor | 1.0 |
| C2 | 2024-06-01 | Jul 2024 | 2024-09-01 | Late Q3 cadence | 1.0 |
| C3 | 2024-07-01 | Aug 2024 | 2024-10-01 | Mirror final target | 1.25 |
| C4 | 2024-09-01 | Oct 2024 | 2024-12-01 | Peak seasonality | 1.0 |
| C5 | 2025-03-01 | Apr 2025 | 2025-06-01 | Near submission | 1.5 |
| C6 | 2025-04-01 | May 2025 | 2025-07-01 | Pre-cutoff check | 1.5 |

### 🇺🇸 USA (h = 3)

| Fold | Train ≤ | Gap (1 mo) | Validate (t+3) | Comment | Weight |
| --- | --- | --- | --- | --- | --- |
| U1 | 2024-05-01 | Jun 2024 | 2024-08-01 | Early H2 anchor | 1.0 |
| U2 | 2024-07-01 | Aug 2024 | 2024-10-01 | Mirror final target | 1.25 |
| U3 | 2024-08-01 | Sep 2024 | 2024-11-01 | Peak seasonality | 1.0 |
| U4 | 2025-02-01 | Mar 2025 | 2025-06-01 | Close to submission | 1.5 |
| U5 | 2025-03-01 | Apr 2025 | 2025-07-01 | Last observable target | 1.5 |

## 📞 D. Deliverables Checklist (End-of-Run)

| Category | Expected Files (all suffixed \_final) |
| --- | --- |
| **Data Splits** | 8 Parquet files |
| **OOF Predictions** | 12 Parquet files |
| **Forecast Predictions** | 12 Parquet files |
| **CV Logs** | 12 CSV files |
| **Blend Weights** | 1 CSV file |
| **Ensemble Outputs (HS6 + HS4)** | 8 Parquet files |
| **Final Forecast (Merged)** | 1 Parquet file |
| **Submission CSV** | 1 CSV file |
| **Method Description** | 1 Word document |

## 📦 E. Final Submission Package to OEC

**Contents:** 1. submission\_final.csv 2. Method\_Description\_Final.docx (≤ 2 pages, reproducible method) 3. *(Optional)* GitHub repository link for transparency

**Note:** All models must beat the naive carry-forward baseline (July/Aug 2025 values) on sMAPE to be eligible. Each artifact must include the \_final suffix to distinguish this definitive run set.