======================================================================

📊 PROGRESS UPDATE - 12:51:01

======================================================================

Phishing URLs:

• Collected: 15,046

• Checked: 0

• Accessible: 0 (0.0%)

• Features extracted: 0

Benign URLs:

• Collected: 14,997

• Checked: 23,344

• Accessible: 4,685 (20.1%)

• Features extracted: 4,675

======================================================================

[BENIGN-PROC-4] Stopping

[BENIGN-PROC-23] Stopping

[BENIGN-PROC-1] Stopping

[BENIGN-PROC-14] Stopping

[BENIGN-PROC-6] Stopping

[BENIGN-PROC-19] Stopping

[BENIGN-PROC-2] Stopping

[BENIGN-PROC-11] Stopping

[BENIGN-PROC-15] Stopping

[BENIGN-PROC-17] Stopping

[BENIGN-PROC-3] Stopping

[BENIGN-PROC-5] Stopping

[BENIGN-PROC-10] Stopping

[BENIGN-PROC-20] Stopping

[BENIGN-PROC-8] Stopping

[BENIGN-FEEDER] All URLs queued

[BENIGN-PROC-13] Stopping

[BENIGN-PROC-12] Stopping

[BENIGN-PROC-18] Stopping

[BENIGN-PROC-0] Stopping

[BENIGN-PROC-7] Stopping

[BENIGN-PROC-9] Stopping

[BENIGN-PROC-22] Stopping

[BENIGN-PROC-21] Stopping

[BENIGN-PROC-16] Stopping

[BENIGN-PROC-24] Stopping

======================================================================

✅ COLLECTION COMPLETE!

======================================================================

Time: 17.8 minutes (0.3 hours)

Phishing URLs: 15,046

Benign URLs: 15,016

Total: 30,062

======================================================================

TRAIN & TEST MODEL WITH FRESH DATASET

======================================================================

Dataset: 30,062 fresh URLs with real features

Expected: 85-92% accuracy (vs old 68%)

======================================================================

======================================================================

LOADING FRESH DATASETS

======================================================================

Loading training data from: streaming\_dataset/train\_dataset.pkl

✓ Training set loaded: 24,049 samples

• Features shape: (24049, 30)

• Label distribution: [12011 12038]

Loading validation data from: streaming\_dataset/val\_dataset.pkl

✓ Validation set loaded: 6,013 samples

• Features shape: (6013, 30)

• Label distribution: [3005 3008]

======================================================================

TRAINING MODEL ON FRESH DATA

======================================================================

Scaling features with RobustScaler...

Building Enhanced Stacking Ensemble Model...

Base models:

1. Gradient Boosting (200 estimators)

2. Random Forest (300 estimators, balanced)

3. MLP Neural Network #1 (256→128→64)

4. MLP Neural Network #2 (128→64→32)

Meta learner: Logistic Regression (balanced)

🚀 Training model (this may take 2-5 minutes)...

[Parallel(n\_jobs=-1)]: Using backend ThreadingBackend with 12 concurrent workers.

[Parallel(n\_jobs=-1)]: Using backend ThreadingBackend with 12 concurrent workers.

[Parallel(n\_jobs=-1)]: Using backend ThreadingBackend with 12 concurrent workers.

[Parallel(n\_jobs=-1)]: Using backend ThreadingBackend with 12 concurrent workers.

[Parallel(n\_jobs=-1)]: Done 5 out of 5 | elapsed: 5.6s finished

[Parallel(n\_jobs=-1)]: Done 5 out of 5 | elapsed: 8.2s finished

[Parallel(n\_jobs=-1)]: Done 5 out of 5 | elapsed: 12.6s finished

[Parallel(n\_jobs=-1)]: Done 5 out of 5 | elapsed: 14.3s finished

✓ Training completed in 22.1s (0.4 minutes)

💾 Model saved:

• fresh\_model\_results\fresh\_phishing\_model.pkl

• fresh\_model\_results\fresh\_phishing\_scaler.pkl

======================================================================

MODEL EVALUATION ON VALIDATION SET

======================================================================

Making predictions...

📊 PERFORMANCE METRICS:

Accuracy: 0.9915 (99.15%)

F1 Score: 0.9915

Precision: 0.9923

Recall: 0.9907

📋 Confusion Matrix:

Predicted

Benign Phishing

Actual Benign 2982 23

Phishing 28 2980

📈 Detailed Metrics:

True Positives: 2,980 (Phishing correctly detected)

True Negatives: 2,982 (Benign correctly detected)

False Positives: 23 (Benign wrongly flagged)

False Negatives: 28 (Phishing missed)

Specificity: 0.9923

🎯 Detection Rates:

Phishing detection: 99.1% (2980/3008)

Benign detection: 99.2% (2982/3005)

📄 Classification Report:

precision recall f1-score support

Benign 0.9907 0.9923 0.9915 3005

Phishing 0.9923 0.9907 0.9915 3008

accuracy 0.9915 6013

macro avg 0.9915 0.9915 0.9915 6013

weighted avg 0.9915 0.9915 0.9915 6013

======================================================================

TESTING WITH SAMPLE URLs

======================================================================

Testing 5 legitimate URLs...

✅ Google

Expected: BENIGN | Predicted: BENIGN

Confidence: 99.18%

Probabilities: Benign=99.18%, Phishing=0.82%

✅ PayPal

Expected: BENIGN | Predicted: BENIGN

Confidence: 99.24%

Probabilities: Benign=99.24%, Phishing=0.76%

✅ GitHub

Expected: BENIGN | Predicted: BENIGN

Confidence: 99.25%

Probabilities: Benign=99.25%, Phishing=0.75%

✅ Amazon

Expected: BENIGN | Predicted: BENIGN

Confidence: 99.11%

Probabilities: Benign=99.11%, Phishing=0.89%

✅ Microsoft

Expected: BENIGN | Predicted: BENIGN

Confidence: 99.03%

Probabilities: Benign=99.03%, Phishing=0.97%

Sample test accuracy: 5/5 (100.0%)

======================================================================

COMPARISON: OLD vs NEW MODEL

======================================================================

| Metric | Old Model | New Model | Improvement |

|-------------|--------------|--------------|-------------|

| Dataset | 47K dead URLs| 30K fresh | Quality ✓ |

| Accessible | ~20% | 100% | +400% ✓ |

| Features | Defaults | Real HTML | Much better✓|

| Accuracy | 68.49% | [SEE ABOVE] | Expected +15-25% |

| Recall | 68.55% | [SEE ABOVE] | Expected +10-20% |

======================================================================

🎉 TRAINING & TESTING COMPLETE!

======================================================================

📊 Final Results:

Training samples: 24,049

Validation samples: 6,013

Training time: 22.1s

✓ Accuracy: 0.9915 (99.15%)

✓ F1 Score: 0.9915

✓ Precision: 0.9923

✓ Recall: 0.9907

🚀 Improvement over old model:

Accuracy: +30.66 percentage points

(68.49% → 99.15%)

🏆 EXCELLENT! Achieved target accuracy of 85%+

📁 Model saved to: fresh\_model\_results/

• fresh\_phishing\_model.pkl

• fresh\_phishing\_scaler.pkl

• training\_summary.pkl