StockSage: Knowledge-Augmented Loser Stock Predictor

BUSN30135 AI and Financial Information

Team: Kirthic Vishnu Arun Kumaar, Goutham Baskaran, Ken Gan, Zacher Lewis

"We pledge our honor that we have not violated the Chicago Booth Honor Code during this assignment."

Impact of Losers

Why it matters:

✓ Loser stocks - companies on track to underperform the S&P 500 .

✓ Avoiding big losses is often more valuable than picking winners.

> ➤ 30% loss = 43% gain required to break even.

✓ Most stock prediction models used to rely only on **numbers** or only on **text.**

Who benefits?

✓ **Institutional Portfolio Managers:** Optimize investment strategies & mitigate losses.

✓ **Investment Bank Risk Teams:** Improve risk assessment models for better decision-making.

✓ **Individual Investors & Financial Advisors:** Enhance long-term portfolio performance.

✓ **Quantitative Funds:** Use multi-modal AI to improve trading signals.

The Problem With Existing Models

Why Current Approaches Fall Short:

✓ **Over-reliance on numbers:** Traditional models use financial ratios & earnings reports but ignore future risks.

✓ **Ignoring qualitative insights:** SEC filings (10-K, 8-K) contain risk factors, but traditional models don't extract insights.

✓ **Inability to contextualize information:** No real-time data from external sources like news, social media, or macroeconomic trends are synced with company information.

How We Aim to Improve It:

✓ Combine financial data, SEC filings (10-K, 8-K), sentiment analysis, and real time news articles.

✓ Enhance accuracy through an **Ensemble model.**

StockSage - Solution Overview

How We Solve This Problem:

✓ **Multi-layered prediction model**: Combines quantitative, qualitative, & contextual signals.

✓ **One-year prediction horizon**: Long-term risk detection.

✓ **XAI (Explainable AI)**: No "black box" decisions.

Model Architecture:

✓ **Foundation/Quantitative Layer:** Statistical/Learning-based financial model.

✓ **Qualitative Layer:** NLP extracts insights from SEC filings (10-K , 8-K).

✓ **Context Layer:** Web search retrieves market developments.

✓ **Reasoning Layer:** Reasoning model integrates all signals for predictions.

Economic & Financial Models

What We Tested:

✓ **Fama-French 5-Factor Model (Fama&French, 2014):**

➢ Limited by reliance on historical data, works well for diversified portfolios but doesn't react to real-time changes.

✓ **Cross-Sectional Forecasting (Rihtamo, Lof&Nyberg, 2024):**

➢ Accounts for some firm-specific predictors but ignores market inefficiencies & behavioral factors.
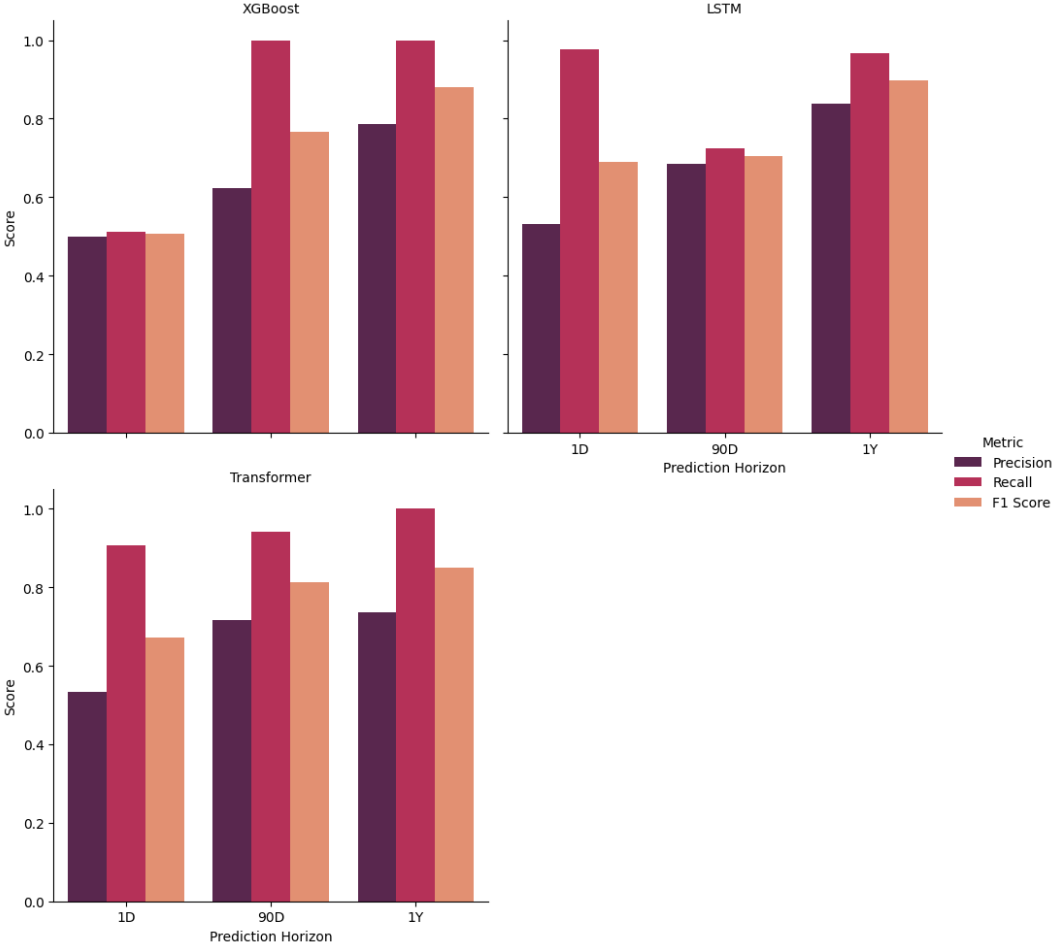
Results & Model Performance

Fama-French Model Metrics Across Time Horizons
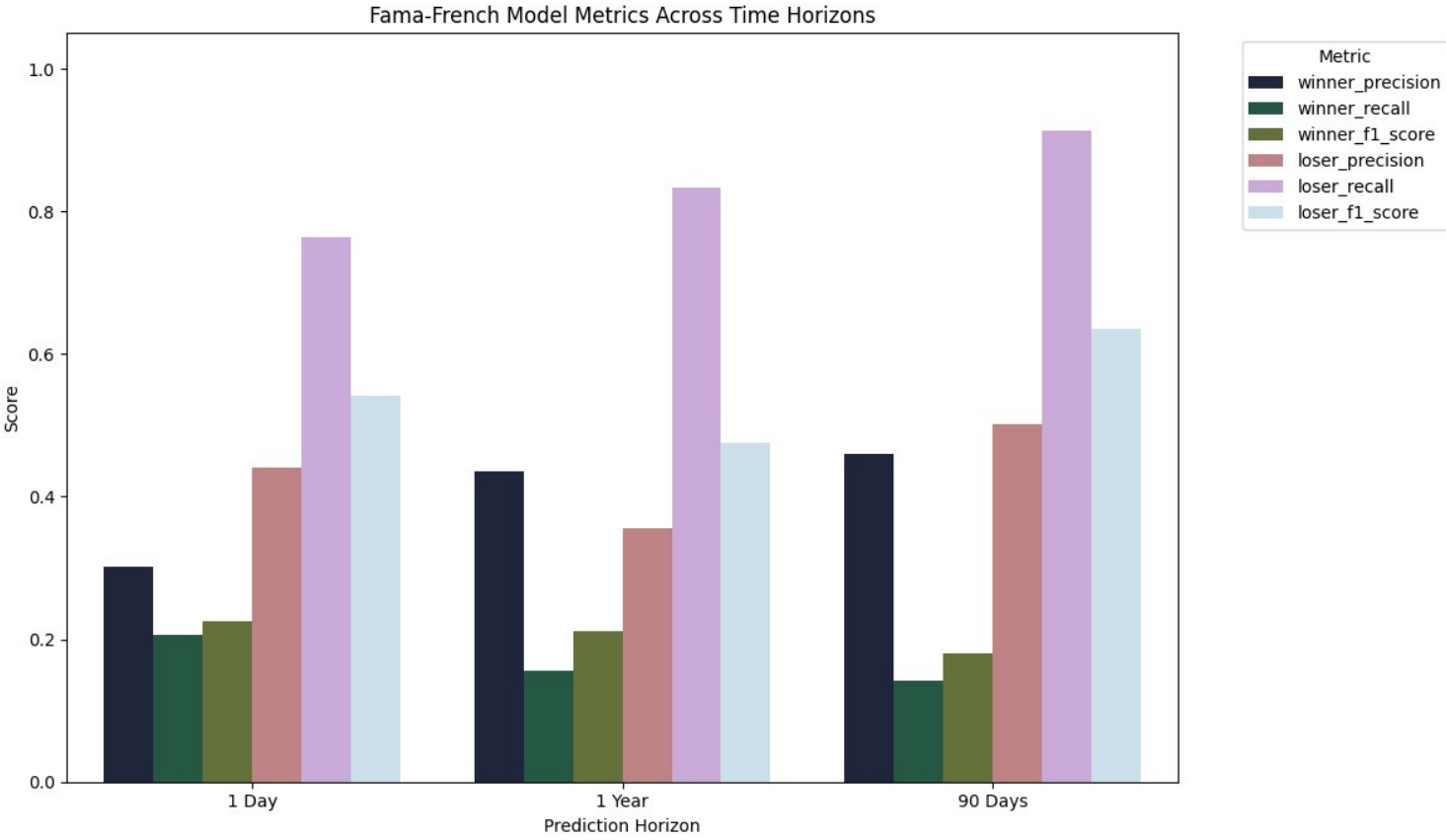
Learning Models

What We Tested:

✓ **XGBoost:** Good for structured data but lacks ability to detect long-term dependencies in stock trends.

✓ **Temporal Fusion Transformer:** Designed for time-series, but did not outperform LSTM due to its reliance on additional input features which we did not have

✓ **LSTM (Long Short-Term Memory Network):** Best for capturing complex temporal dependencies in stock price movements

Results & Model Performance

Deep Learning Models:

Results & Model Performance

Final Decision:

- ✓ **LSTM:** Best model for precision & accuracy.

    - ➢ Remembers patterns over time, making it better for long-term trend forecasting.

- ✓ **LSTM Accuracy:** ~84% (Best among tested).

- ✓ **Precision:** 83% (Reliable predictor of winners/losers).

- ✓ **Recall:** 96% (Captured most actual losers).

Key Insights:

- ✓ Quanttitative models provide a baseline but **miss key behavioral & contextual factors**.

- ✓ Behavioral inefficiencies create opportunities for arbitrage, which traditional models miss.

- ✓ LSTM model **does not capture qualitative insights**

Augmentations to Quantitative model

Usage of Large Language Models (LLMs) to extract and structure qualitative data :

1. Qualitative and contextual signals from SEC filings

- ✓ Extracts risk & opportunity signals (8-K, 10-K).

- ✓ Key insight : 8-Ks contain more change-driving information (Uncovering Information - Prof. Levy).

2.Deterministic Impact Scoring

- ✓ Developed a scoring system to calculate impact of risk and opportunity signals.

3. Analyzing Trends and Assessing Regulatory Risks

- ✓ Positive/Negative/Neutral Trend identified as a ratio of risk/opportunity impact.

- ✓ Deterministic Numerical Score Assigned to Regulatory Risks.

4.Sentiment Analysis

- ✓ Used Sentiment Intensity Analyzer from **vaderSentiment** to assign polarity and compound scores to risk and opportunity signals.

Augmentations to Quantitative model

Bringing it all together:

5. **LLM-Based Prediction**

   ✓ Generate final prediction using compiled data from SEC filings.

6. **Conduct Market Research**

   ✓ Incorporates real-time external context like news, stock performance, macroeconomic trends, competitive position of industry, and analyst opinions.

7. **Reasoning Layer**

   ✓ Synthesizes quantitative insights from LSTM, qualitative signals from SEC filings, and external news data and classifies the stock as winner/loser based on reasoning.

System Architecture
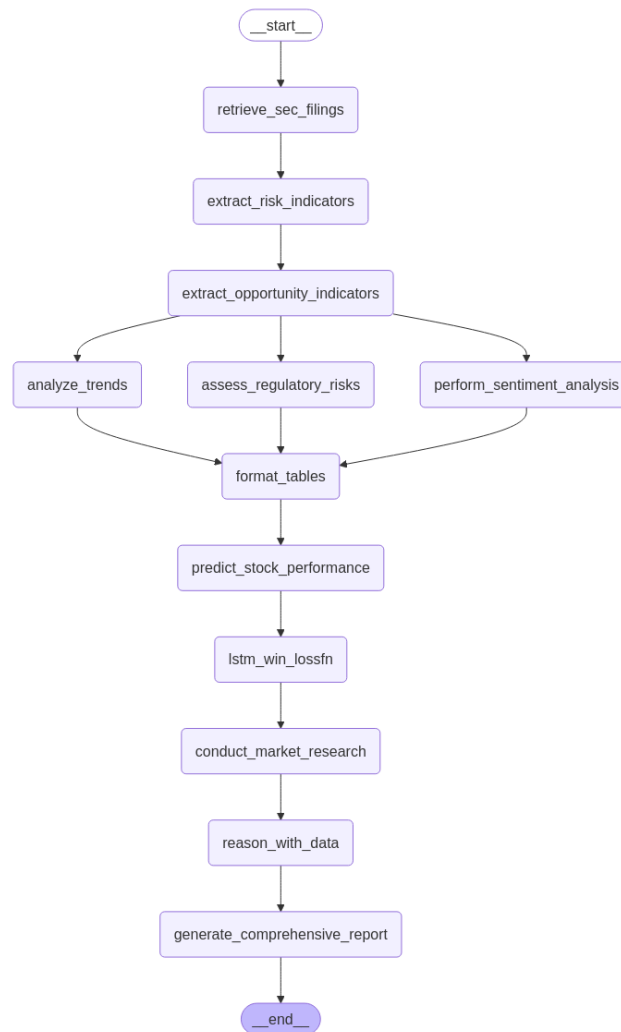
Conclusion and Limitations

Key Takeaways:

- ✓ Multi-modal approach explains traditional finance models' predictions and with back testing we could prove it beats them.

- ✓ Explainability: AI provides clear, justifiable predictions.

- ✓ With further tweaking and testing, investors can use this to minimize downside risk.

Limitations:

- ✓ **Market Uncertainty:**

  - ➢ Stock markets are influenced by many unpredictable external events. Our model provides probabilistic guidance, not certainty.

- ✓ **False Positives/Negatives:**

  - ➢ We mitigate this risk by integrating confidence scores and impact-weighted signals to avoid low-certainty recommendations.

- ✓ Web search component lacks historical back testing to avoid look ahead bias (API Limitations).

- ✓ One-year prediction horizon limits applicability for shorter-term strategies.

Future Work

Future Enhancements:

- ✓ Sector-specific models: Tech vs. Retail vs. Energy.

- ✓ Proprietary web archive: Incorporate historical web data for better backtesting.

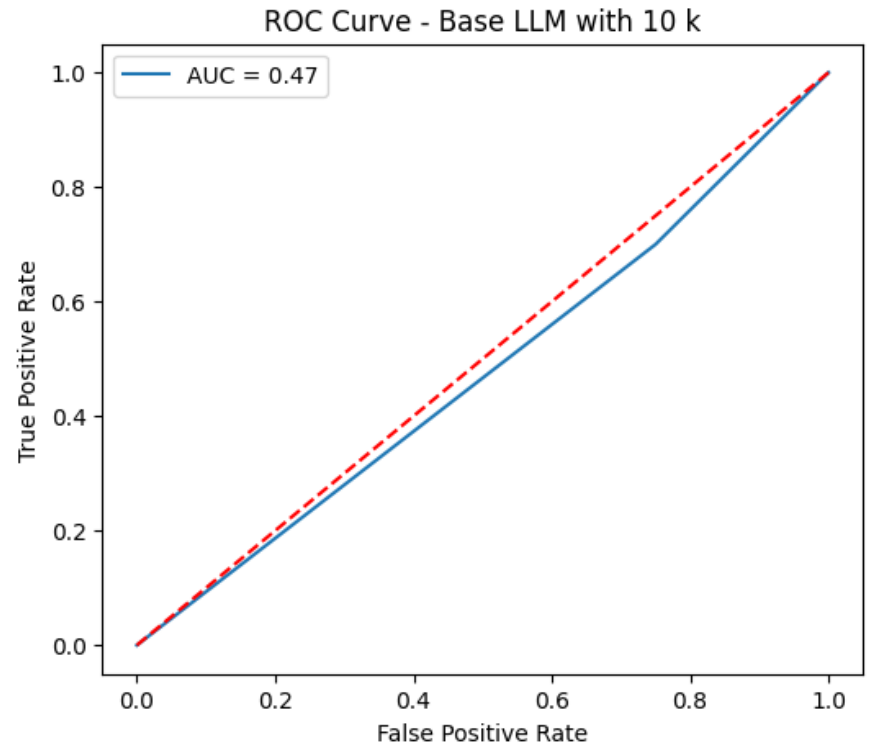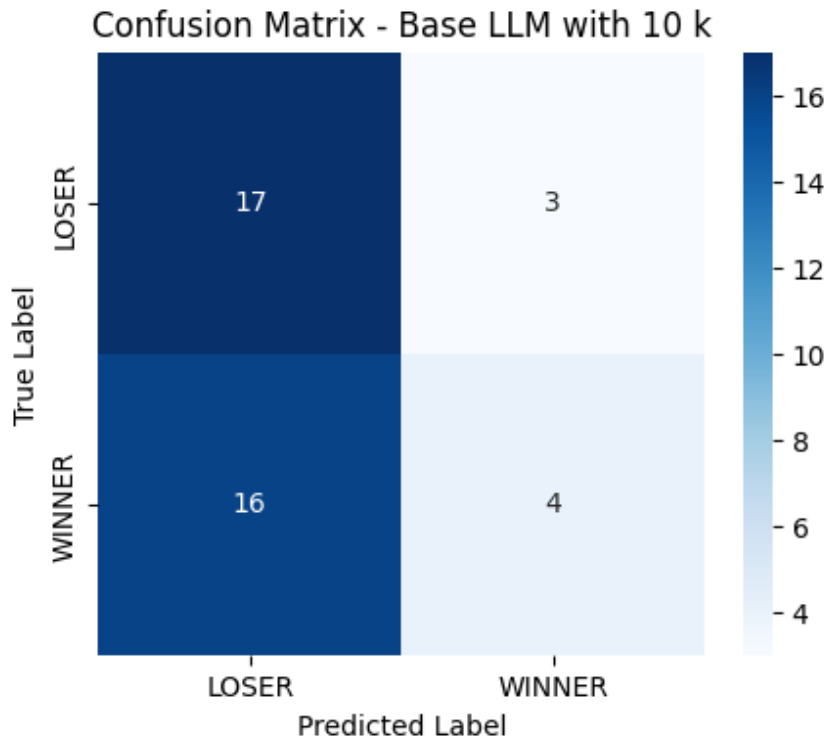- ✓ Expand to different time horizons (3 months, 6 months, etc.).

Demo/Q&A

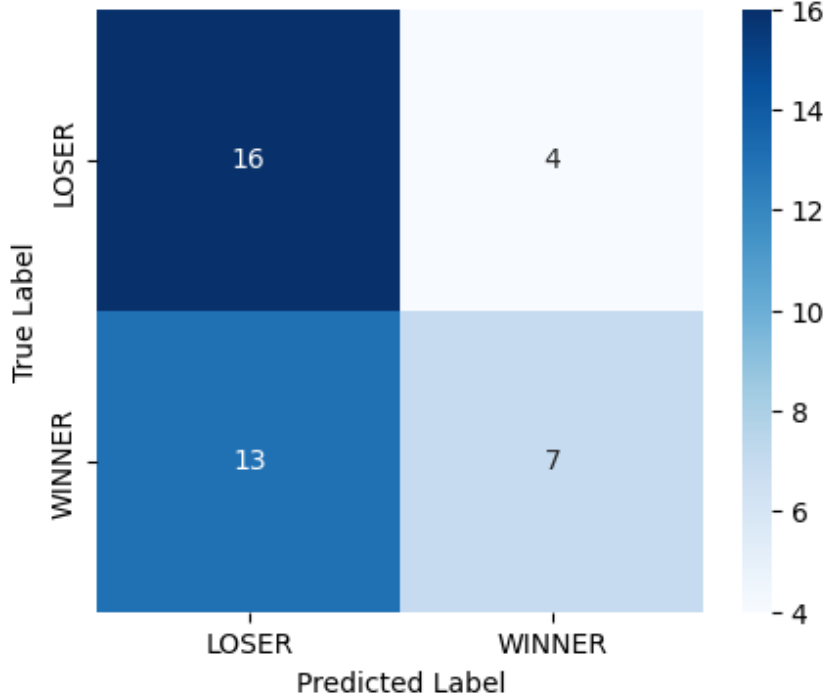# Thank You!!!

Appendix

Results & Model Performance

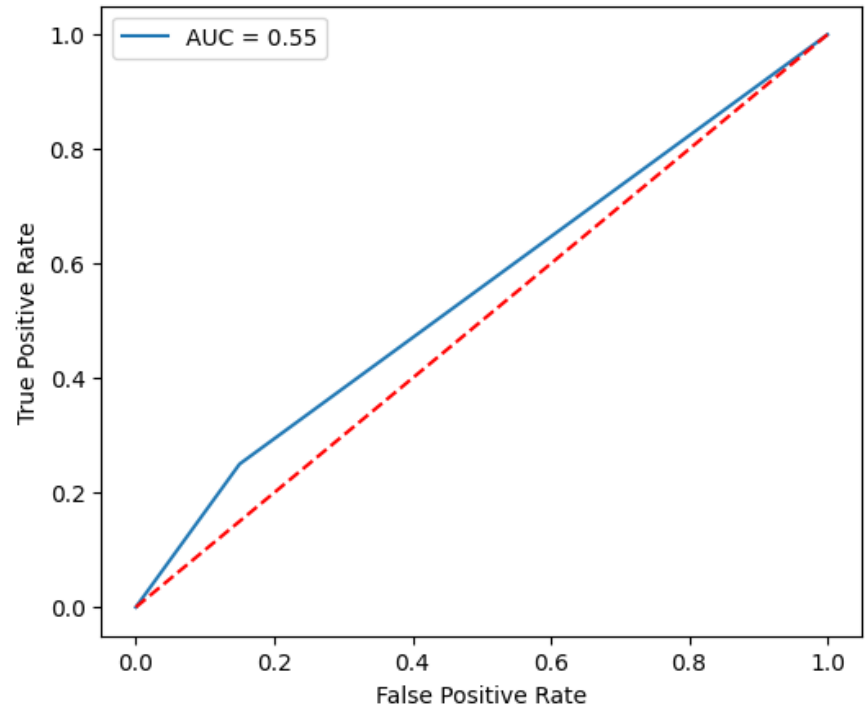**Large Language Models based predictions(Base LLM with 10k):**
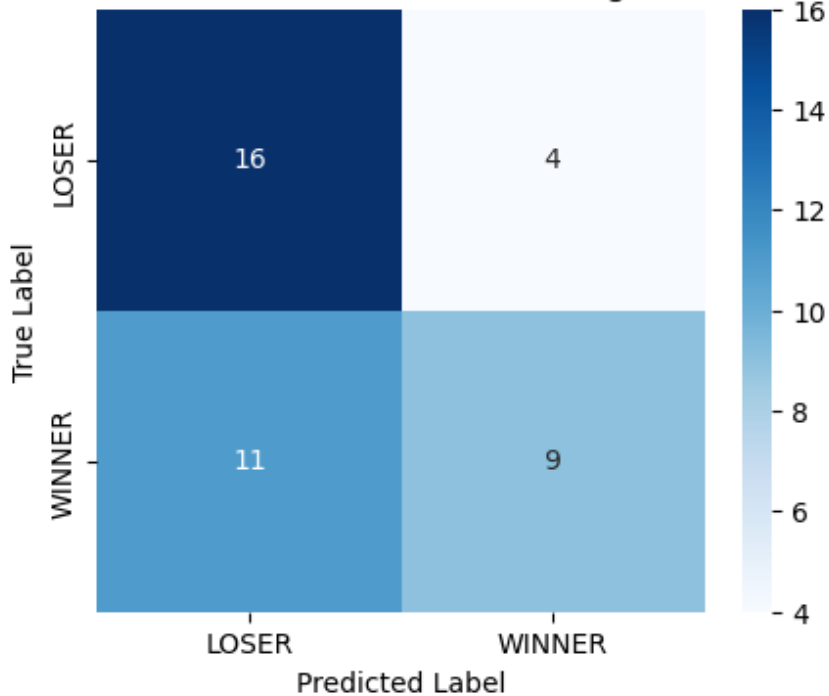
Results & Model Performance

Large Language Models based predictions(Base LLM with 10k and 8k):

Results & Model Performance

Large Language Models based predictions(LLM with reasoning and LSTM):