DLP Final Project Prposal:
Market Guided Stock Transformer

Group 7
110705009 陳重光、313551047 陳以瑄、313554043 戴明貴

Outline

- Intoduction - Market Guided
- MASTER - AAAI'24
- Motivation & Innovation
- Problem Definition

Introduction - Market Guided

Stock prediction features can be divided into two types:

1. Individual Stock Features:
    - Open price, close price, etc.
    - Trading volume
2. Shared market features:
    - Market index
    - Macroeconomic indicators, e.g. interest rate

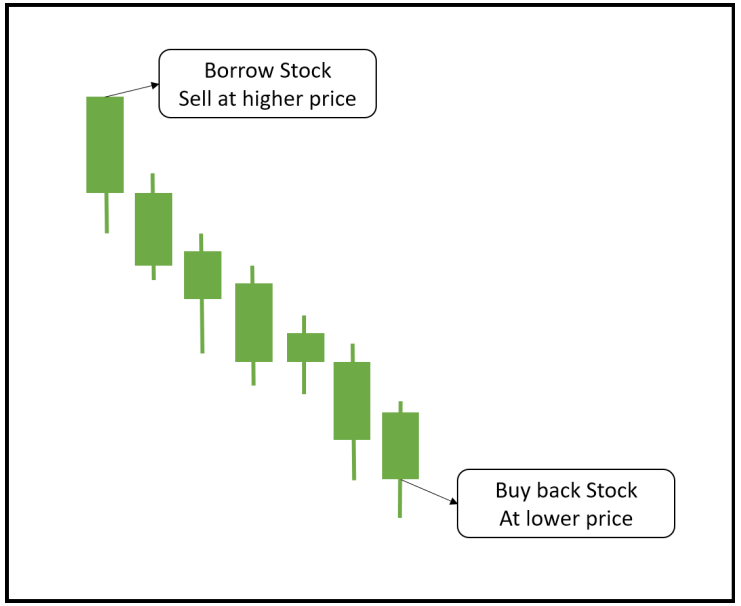
Introduction - Market Guided

The market feature impacts the effectiveness of other features.

Example: Short Selling

When investors believe a stock is overvalued.

1. Borrow stock, sell at high price.
2. Buy back at lower price when it falls.
3. Return to owner.



Short selling interest:
the amount of stocks being short.

Introduction - Market Guided

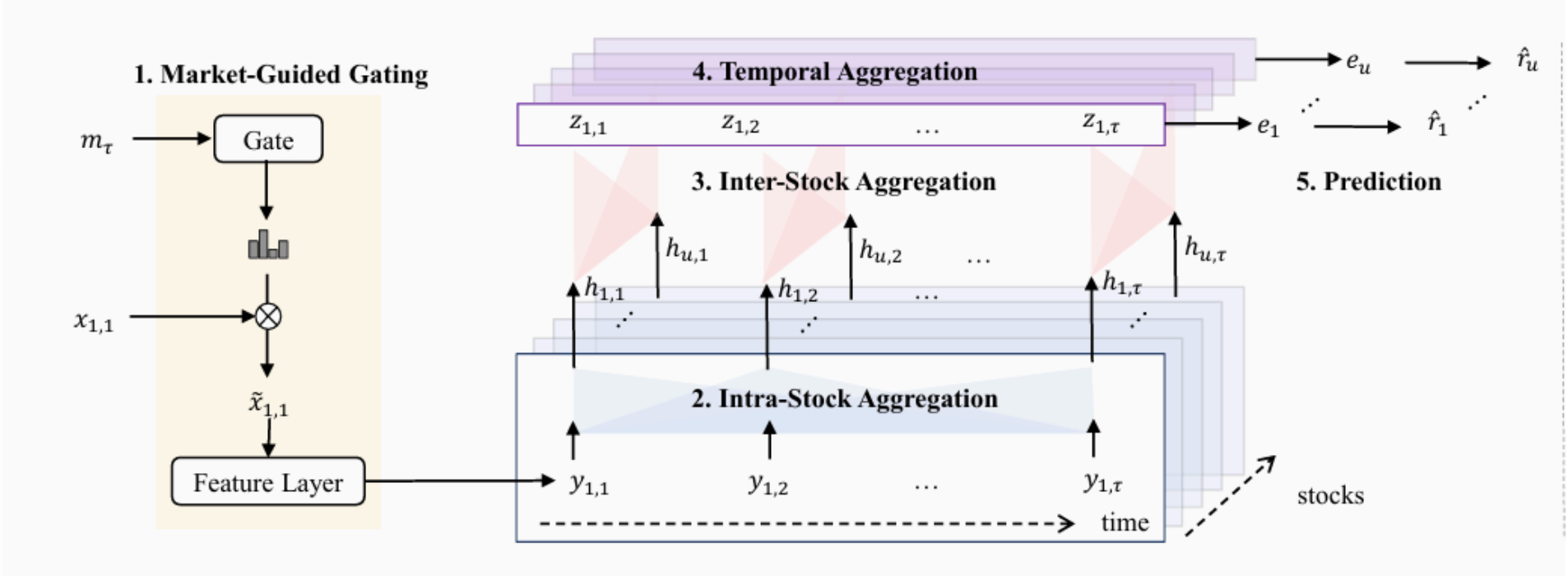
The market feature impacts the effectiveness of other features.

Example: Short Selling
Effectiveness in different market status:

- Bull Market: Short selling loses money, less concern.
- Bear Market: Short selling signals pessimism, more significant.

→ **Using market status to select relevant features.**

MASTER:Market-Guided Stock Transformer for Stock Price Forecasting [1]



[1] MASTER:Market-Guided Stock Transformer for Stock Price Forecasting -AAAI 24

Limitation

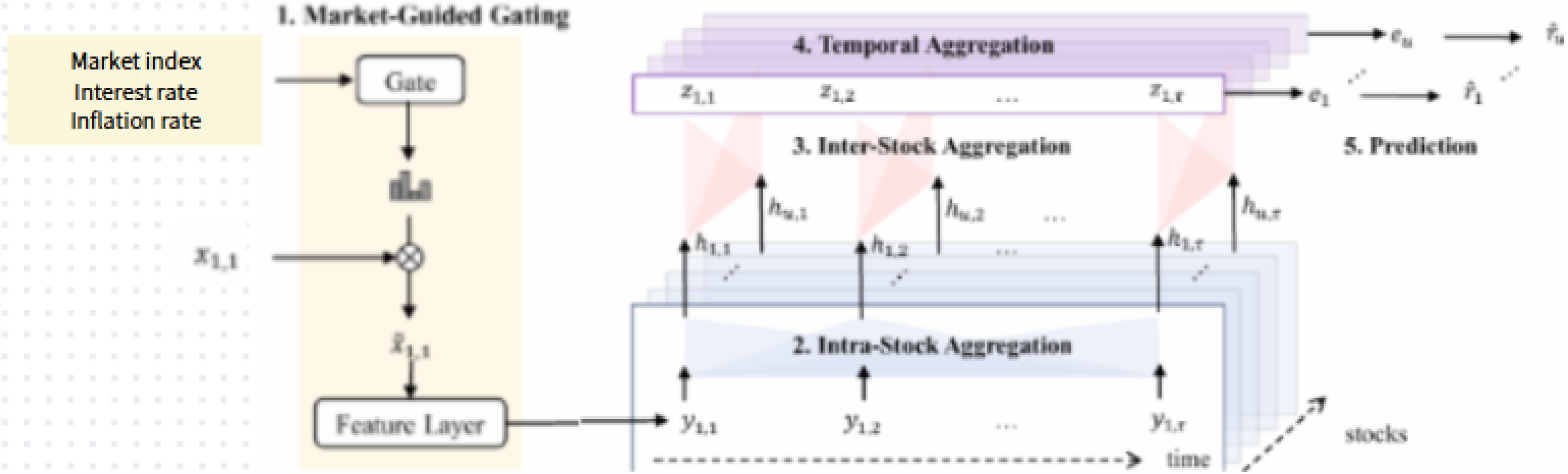
Simple Representation of Market Status:
Only market index prices and trading volumes are used as inputs.

Improvements: Expanding Shared Market Features

1. Macroeconomic features
2. Industry-level features
3. News-based features

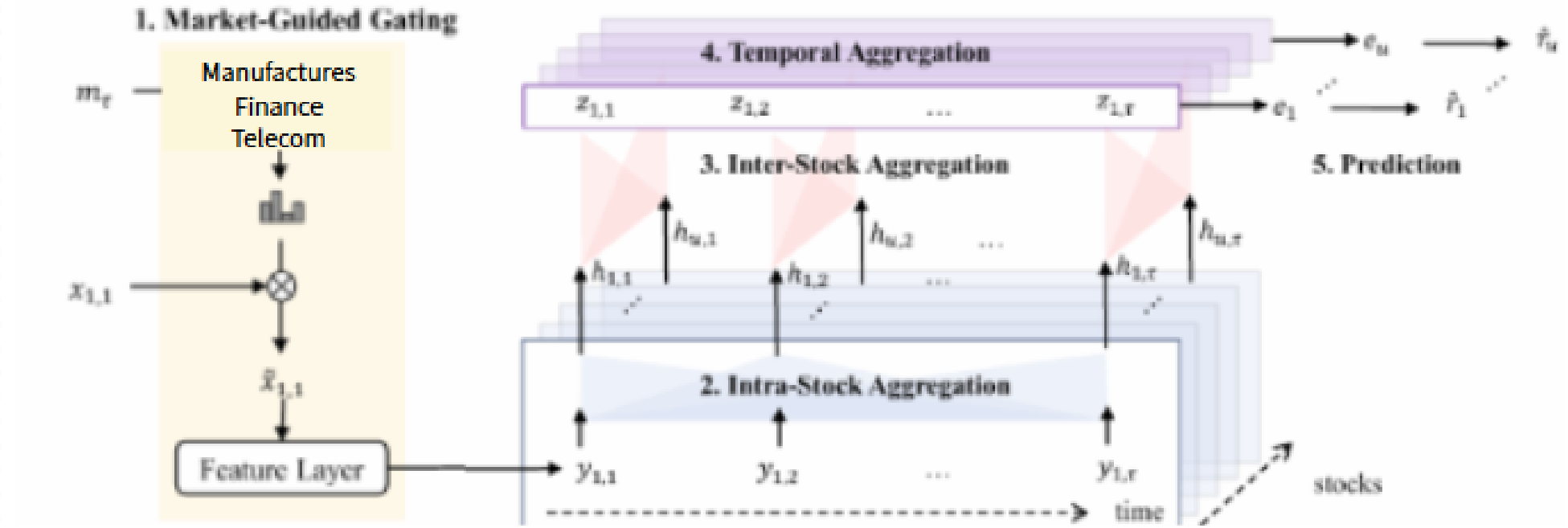
Idea 1

Market-guided indicators expansion for richer market dynamics



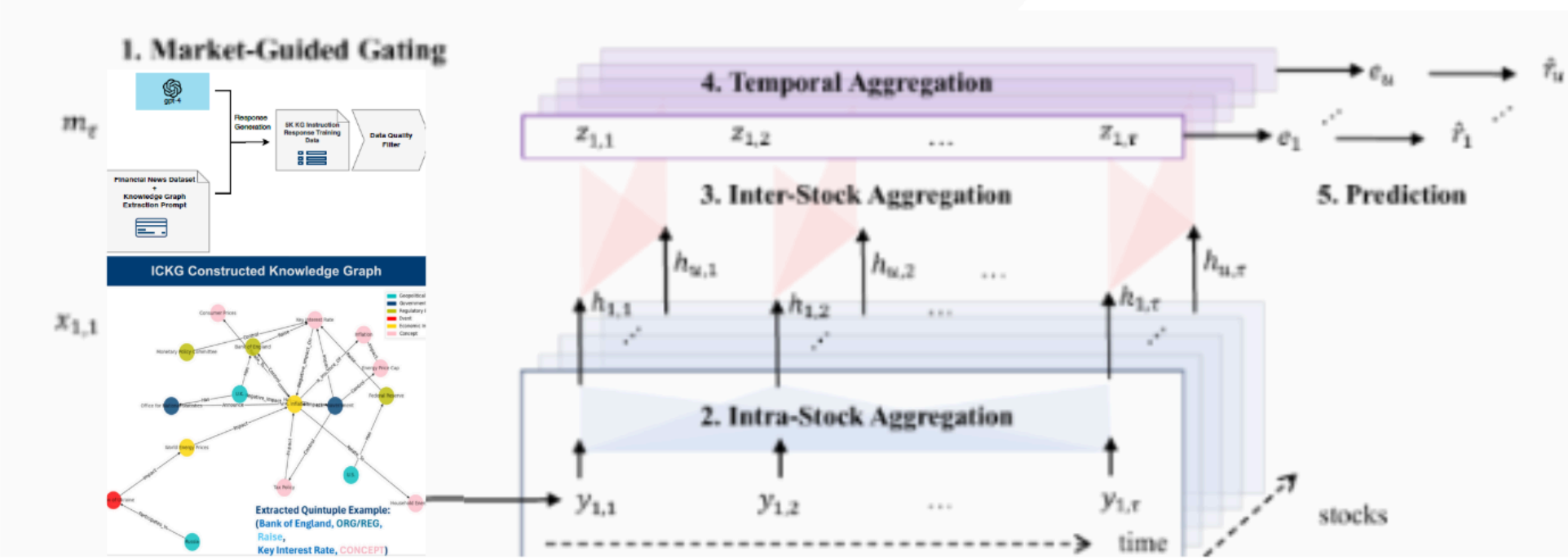
Idea 2

Industry-level features for sector-specific behavior capture



Idea 3

Supply chain and news features for enhanced forecasting



- **MASTER**: dynamic stock correlations with market-guided feature
- **FinDKG**: company relationships to identify business partners

Problem Definition

Given a set of stocks S with features $x_{u,t} \in \mathbb{R}^F$ collected at time steps $t \in [1, \tau]$:

For each stock, we consider:

- Individual stock features (price, volume)
- Shared market features(market index, macroeconomic indicators)
- Industry-Level feature (return)
- News-derived features (company and related party)

Output: The return ratio $r_u = \mathrm{Norm}_S((c_{u,\tau+d} - c_{u,\tau+1})/c_{u,\tau+1})$

Data Description

The dataset for input of this study consists of the following data:

- **Stock prices**
- **Industry**
- **Market index**
- **Economic indicator**
- **Sentimental scores**

Data Description (cont.)

- **Stock Price**:
  Using the S&P 500 constituents as our base, we classify stocks into 12 categories following the Fama-French industry classification [3].

  From each category, 8 companies are selected based on market capitalization, resulting in a total of 96 firms.

  For each selected company, we collect daily stock data, including open, high, low, and close prices and trading volume.

Data Description (cont.)

- **Industry**:
  The industry returns are derived from the Fama-French 12 industry classification dataset.
- **Market Index**:
  S&P 500 market index
- **Economic Indicators**:
  We use the VIX, interest rates, and other economic indicators
- **Sentimental Scores**:
  Daily news sentiment from RavenPack [4] is used to measure its impact on the market and stocks.

Data Description (cont.)

NAME	NUMBER	TRAINING	TEST	SOURCE
Stock Price	96 * 5	2010 - 2022	2023	WRDS - CRSP
Industry	12	2010 - 2022	2023	Fama-French
Market Index	1	2010 - 2022	2023	CRSP
Economic Indicators	1	2010 - 2022	2023	VIX, FRED
Sentimental Scores	6	2010 - 2022	2023	Ravenpack

Expected result

Due to the change in the set of stocks used (from Chinese stocks to U.S. stocks), prior related works are no longer directly applicable.

We aim to compare performance between:

1. The original version of MASTER
2. MASTER with three improvements

References

[1] Li, T., Liu, Z., Shen, Y., Wang, X., Chen, H., & Huang, S. (2024). MASTER: Market-Guided Stock Transformer for Stock Price Forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, 38(1), 162-170.

[2] Xie, J., Zhang, Y., Gong, X., Huang, J., Li, Z., Qin, B., & Liu, T. (2023). CausalStock: Deep End-to-end Causal Discovery for News-driven Stock Movement Prediction. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23) (pp. 2320-2329). ACM.