

مقاله master چی میگ و چرا کار میکنه :

گیت‌بندی راهنمایی بازار (Market-Guided Gating):

در این مرحله، ابتدا یک بردار m_t برای نمایش وضعیت فعلی بازار ایجاد می‌شود. سپس از یک مکانیسم گیت‌بندی برای اعمال این بردار به داده‌های ویژگی استفاده می‌شود. این کار باعث بهبود ویژگی‌های ورودی شده و تمایل به انتخاب ویژگی‌هایی که با وضعیت فعلی بازار مرتبط هستند، را دارد.

تجمیع داخلی سهام (Intra-Stock Aggregation):

در این مرحله، برای هر سهام، در هر زمان t ، اطلاعات از دیگر زمان‌ها جمع‌آوری شده و یک تعبیر محلی به نام $h_{u,t}$ ایجاد می‌شود. این تعبیر محلی اطلاعات مهم و مربوط به زمان‌های گذشته را حفظ می‌کند.

تجمیع میان سهام (Inter-Stock Aggregation):

در این مرحله، به ازای هر زمان t ، همبستگی بین سهام با استفاده از یک مکانیسم توجه محاسبه می‌شود و هر سهام اطلاعات جمع‌آوری شده از سهام دیگر را تجمیع می‌کند. این اطلاعات تجمیع شده شامل همبستگی‌ها و اطلاعات مربوط به زمان t می‌باشد.

تجمیع زمانی (Temporal Aggregation):

در این مرحله، برای هر سهام، از تمام تعبیرهای زمانی قبلی استفاده شده و یک تعبیر جامع برای هر سهام تولید می‌شود. این تعبیر جامع شامل تمام اطلاعات زمانی مربوط به سهام است.

پیش‌بینی (Prediction):

در مرحله نهایی، تعبیر جامع سهام به لایه‌های پیش‌بینی ارسال می‌شود تا برچسب پیش‌بینی شود. در این مرحله، از اطلاعات جامع تجمیع شده استفاده می‌شود تا قیمت آتی سهام پیش‌بینی شود.

در این بخش، ابتدا به ترکیب اطلاعات از دو جنبه برای ایجاد یک بردار m_t برای توصیف جامع وضعیت فعلی بازار پرداخته می‌شود.

قیمت شاخص بازار:

قیمت شاخص بازار میانگین وزنی از قیمت‌های یک گروه از سهام‌ها S است که بر اساس سهم سرمایه‌گذاری بازار آن‌ها است. معمولاً S از شرکت‌های برتر با بیشترین سرمایه‌گذاری بازار تشکیل شده است که نماینده‌ای از بازار یا بخش خاصی است و ممکن است با سهام مورد علاقه سرمایه‌گذاران در بازار S متفاوت باشد. در اینجا قیمت فعلی شاخص بازار در t و همچنین قیمت‌های تاریخی شاخص بازار که توسط میانگین و انحراف معیار در روزهای گذشته d توصیف می‌شود (برای نشان دادن نوسانات قیمت)، در نظر گرفته می‌شود.

حجم معاملات شاخص بازار:

حجم معاملات S نشان‌دهنده درگیری سرمایه‌گذاران است و فعالیت بازار را نمایش می‌دهد. ما میانگین و انحراف معیار حجم معاملات شاخص بازار در روزهای گذشته d را در نظر می‌گیریم تا اندازه واقعی بازار را نشان دهیم.

مکانیسم گیت‌بندی

مکانیسم گیت‌بندی، برای هر بعد ویژگی، یک ضریب اسکالینگ تولید می‌کند که برای بزرگ یا کوچک شدن مقدار ویژگی، از آن استفاده می‌شود و این کار باعث تاکید یا کاهش میزان اطلاعات از ویژگی می‌شود که به مازول‌های بعدی می‌رود. مکانیسم گیت‌بندی توسط آموزش مدل یادگرفته می‌شود و ضریب به اندازه کمی که ویژگی به بهبود عملکرد پیش‌بینی کمک می‌کند، بهینه‌سازی می‌شود و در نتیجه نشان‌دهنده موثر بودن ویژگی است.

$$\alpha(m_\tau) = F \cdot \text{softmax}_\beta(W_\alpha m_\tau + b_\alpha),$$

که در آن، b_α و W_α ماتریس وزن و بایاس قابل یادگیری هستند، و β پارامتر فرایپارامتر دمایی است که کنترل می‌کند که توزیع خروجی چقدر تیز باشد. Softmax رقابتی بین ویژگی‌ها را فراهم می‌کند تا ویژگی‌های مؤثر را از ویژگی‌های نامآمیز تمیز دهد. در اینجا، یک دمای کوچکتر β باعث می‌شود تا توزیع بیشتر به بعضی از ابعاد تمرکز کند و تأثیر گیت‌بندی قوی‌تر باشد، در حالی که یک β بزرگتر باعث می‌شود تا توزیع به تساوی نیز تمایل پیدا کند و اثر گیت‌بندی ضعیف‌تر باشد. لازم به ذکر است که ما از عملیات ضرب در هادامارد، مقدار را در هر بعد F برای بار اول بزرگ می‌کنیم. این عملیات توزیع تولید شده را با یک توزیع یکنواخت که در هر بعد برابر با $F/1$ است، مقایسه می‌کند تا تعیین شود که آیا مقدار باید بزرگتر یا کوچکتر شود. ایده اصلی پشت تولید ضریب‌ها از m_τ این است که مؤثر بودن ویژگی‌ها تحت تأثیر وضعیت بازار قرار می‌گیرد. به عنوان مثال، اگر مدل یاد بگیرد که فاکتور میانگین متحرک (MA) در دوره‌های بازار پرتنش مفید است، هنگامی که بازار دوباره پرتنش می‌شود، میانگین متحرک را تأکید می‌کند.

داده :

ما چارچوب کاری خود را روی بازار سهام چین با مجموعه‌های سهام CSI300 و CSI800 ارزیابی می‌کنیم. CSI800 و CSI300 دو مجموعه سهام هستند که شامل ۳۰۰ و ۸۰۰ سهم با بالاترین ارزش سرمایه در بورس شانگهای و بورس شنژن می‌باشند. مجموعه داده شامل اطلاعات روزانه از سال ۲۰۰۸ تا ۲۰۲۲ از CSI300 و CSI800 می‌باشد. ما از داده‌ها از فصل اول ۲۰۰۸ تا فصل اول ۲۰۲۰ به عنوان مجموعه آموزش، داده‌ها از فصل دوم ۲۰۲۰ به عنوان مجموعه اعتبارسنجی و ده فصل آخر یعنی از فصل سوم ۲۰۲۰ تا فصل چهارم ۲۰۲۲ به عنوان مجموعه آزمون استفاده می‌کنیم. ما از شاخص‌های عمومی Alpha158 (Yang et al. 2020) برای استخراج ویژگی‌های سهام از داده‌های جمع‌آوری شده استفاده می‌کنیم

. برای نمایندگی از بازار، ما ۶۳ ویژگی با شاخص‌های بازار CSI300، CSI500 و CSI800 ایجاد کرده‌ایم

معماری:

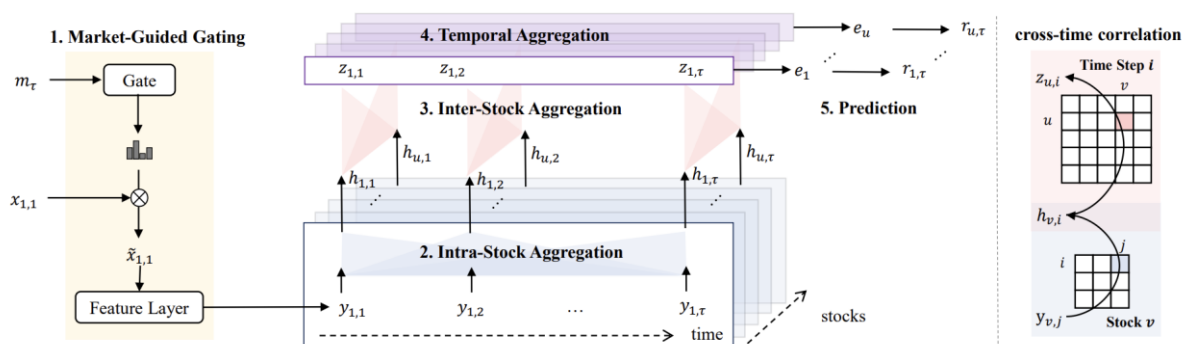


Figure 2: Overview of the MASTER framework.

مقاله CHATGPT INFORMED GRAPH NEURAL NETWORK FOR STOCK MOVEMENT PREDICTION

تو این مقاله میاد میگ هر چقدرم میخوان سعی کنن بفهمن ارتباط بین سهام ها دقیقا چطوریه تا حالا نتونستن پس ما بیایم از gpt استفاده کنیم سرتیتر خبر ها رو میدی جی پی تی و ازش میپرسی چه شرکت هایی در ارتباط هستن و با چه احساساتی بعد از این یک گراف میکشه گرافه رو میدی به gnn تبدیل به بردار میکنه برداره رو ترکیب میکنه با داده مالی میدی و اینکه ی بارم داده مالی رو تنها میدی lstm و این دوستان رو با یک mlp ترکیب میکنه

حالا ایده ها :

یک اینکه موقع ساخت گراف بیام از شاخص های مالی استفاده کنیم خیلی کمک میکنن ارتباط رو بهتر بفهمیم همچنین میتونن در اینکه ما یک وزنی بدیم به گراف کمک کنند

ایده دوم اینکه مکانیزم توجه همیشه جوابه شاید با اون و شاخص های مالی بتونیم ارتباط رو بهتر مدل کنیم و نیاز به گراف نباشه شایدم شاخص های مالی رو بدیم به توجه بعد استفاده کنیم تو گراف خلاصه ساخت گرافش جا داره هنوز

نکته بعدی اینکه یک از این lstm ساده استفاده کرده که قضیه انقدر ساده خوب نی شاید بهتره از transformer ها و شبکه های پیچیده تر استفاده کنیم همچنین می تونیم شاخص ها را اینجا هم اضافه کنیم

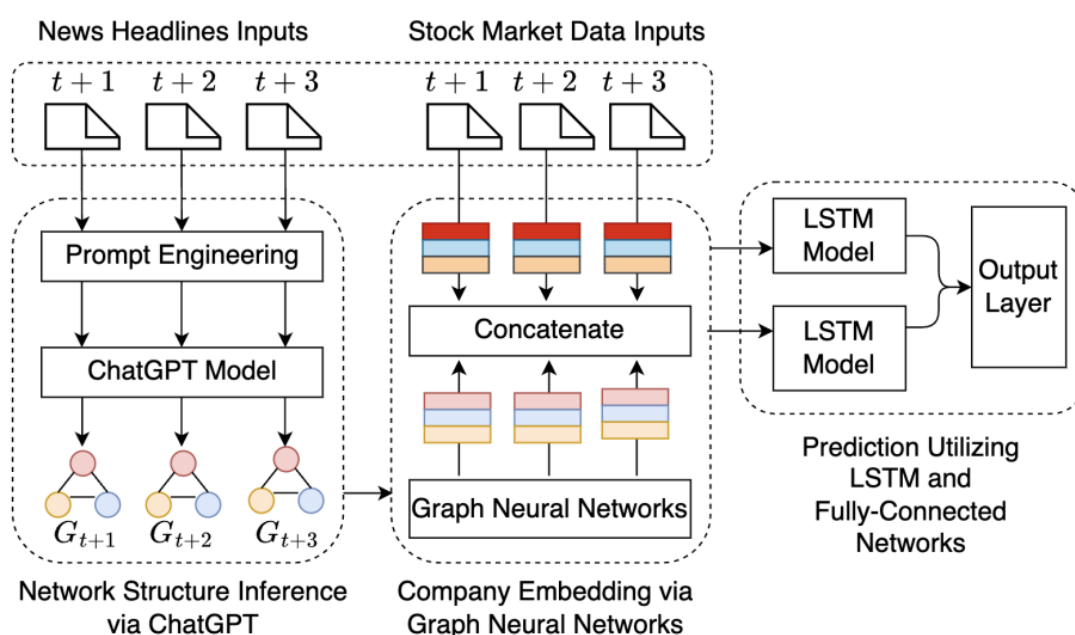


figure 1: Framework Overview: Combining Graph Neural Network and ChatGPT to predict stock movements.

پرامپت:

Forget all your previous instructions. I want you to act as an experienced financial engineer. I will offer you financial news headlines in one day. Your task is to:

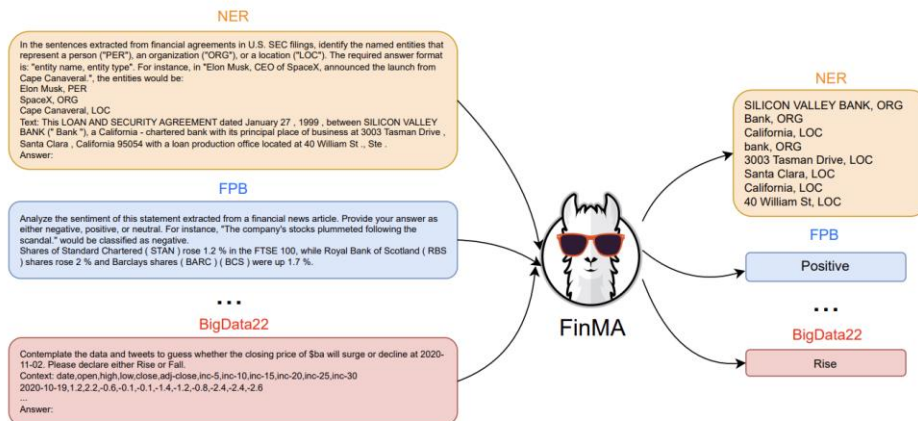
1. Identify which target companies will be impacted by these news headlines. Please list at least five of them.
2. Only consider companies from the target list.
3. Determine the sentiments of the affected companies: positive, negative, or neutral.
4. Only provide responses in JSON format, using the key "Affected Companies".
5. Example output: {"Affected Companies": {Company 1: "positive", Company 2: "negative"}}
6. News Headlines are separated by "\n"

News Headlines: ...

بعدی سری ایده زد به ذهنم که شاید خوب باشه مثلاً ما برای هر سهام یک اکسپرت داشته باشیم

یا اینکه مثلاً اینکه شاخص های مختلف را داشته باشیم مدل multi task طور مثلاً یکی اتفاقات مهم رو شناسایی کنه و تأثیرش رو سهام ها بسنجه یکی بیاد ریسک رو شناسایی کنه یکی بیاد قیمت رو در بیاره و اینا کنار هم کار کنن

PIXIU: A Large Language Model, Instruction Data and Evaluation Benchmark for Finance



<https://github.com/The-FinAI/PIXIU?tab=readme-ov-file>

اومده instruct tuning کرده روی چندین تسک مختلف و دیتاست و مدل رو گذاشته کامل دمش گرم
 کلا به نظرم کار باحالیه که این تسک های مختلف رو از هم جدا کرد و جدا جدا به مدل آموزش بدی
 و خب ریزالت هاش همچین خوب بود میشه ازش استفاده کرده برای predict
 ی حرف خوبی که زده بود این بود این بنچ مارک ها همشون اومده هی فقط روی مدل های زبانی مالی خودشون رو
 evaluate کردن ولی خوبه که ببینیم تو stock prediction چند چندن داده:

Data	Task	Raw	Instruction	Data Types	Modalities	License
FPB	sentiment analysis	4,845	48,450	news	text	CC BY-SA 3.0
FiQA-SA	sentiment analysis	1,173	11,730	news headlines,tweets	text	Public
Headline	news headline classification	11,412	11,412	news headlines	text	CC BY-SA 3.0
NER	named entity recognition	1,366	13,660	financial agreements	text	CC BY-SA 3.0
FinQA	question answering	8,281	8,281	earnings reports	text,table	MIT License
ConvFinQA	question answering	3,892	3,892	earnings reports	text,table	MIT License
BigData22	stock movement prediction	7,164	7,164	tweets,historical prices	text,time series	Public
ACL18	stock movement prediction	27,053	27,053	tweets,historical prices	text,time series	MIT License
CIKM18	stock movement prediction	4,967	4,967	tweets,historical prices	text,time series	Public

Trade the Event: Corporate Events Detection for News-Based Event-Driven Trading

<https://github.com/Zhihan1996/TradeTheEvent>

روش های مبتنی بر ویژگی های متنی:
 این روش ها پیش بینی قیمت سهام را به عنوان یک مسأله طبقه بندی متن در نظر می گیرند.
 به طور مستقیم، بر اساس ویژگی های استخراج شده از اخبار، افزایش یا کاهش قیمت سهام را پیش بینی می کنند.

به دلیل عدم توجه به دلایل تغییر قیمت سهام، در تصمیم‌گیری معاملاتی ضعیف عمل می‌کنند.

روش‌های مبتنی بر تحلیل احساسات:

این روش‌ها احساسات موجود در اخبار را به عنوان شاخصی برای حرکت سهام در نظر می‌گیرند.

اما، احساسات اخبار می‌تواند ذهنی باشد و به طور قابل توجهی تحت تأثیر دیدگاه و سبک نگارش نویسنده قرار گیرد.

راه حل : جمله‌ها را توکن کنیم برچسب بزنیم

گرفتاری: شناسایی رویدادها تنها در سطح کلمه (واژه به واژه)، ممکن است درک کل مقاله را نادیده بگیرد.

مدل تشخیص رویداد دو سطحی:

برای غلبه بر این مشکل، از یک مدل تشخیص رویداد دو سطحی استفاده می‌شود.

سطح پایین:

وظیفه شناسایی زیرمجموعه‌هایی از متن (توالی کلمات) را بر عهده دارد که رویدادهای خاصی را توصیف می‌کنند.

این کار با طبقه‌بندی تک تک کلمات (واژه‌ها) در متن انجام می‌شود.

سطح بالا:

نتایج پیش‌بینی شده از سطح پایین را دریافت می‌کند.

این نتایج را با اطلاعات کلی متن کل مقاله ادغام می‌کند.

با در نظر گرفتن این ترکیب، احتمال وقوع هر رویداد را پیش‌بینی می‌کند.

مزایای مدل دو سطحی:

این مدل می‌تواند هم به کلمات کلیدی خاص (سطح پایین) و هم به چارچوب کلی متن (سطح بالا) توجه کند.

این امر منجر به درک عمیق‌تر از رویدادهای ذکر شده در اخبار می‌شود.

گرفتاری دیگر : یکی دیگر از مشکلات موجود در مدل‌های فعلی این است که زمان‌بندی مقالات خبری را نادیده

می‌گیرند. بیشتر آن‌ها از مقالات خبری برای پیش‌بینی افزایش/کاهش اوراق بهادار مرتبط در روز(های) معاملاتی آینده

استفاده می‌کنند. با این حال، قیمت‌های سهام احتمالاً به‌طور فوری در پاسخ به اخبار قابل توجه تغییر می‌کنند.

بنابراین، حرکت سهام در روز(های) معاملاتی آینده ممکن است نتواند تأثیر مقاله خبری را به دقت بازتاب دهد. برای

حل این مشکل، ما پیش‌بینی‌های سهام را همان لحظه‌ای که یک مقاله خبری منتشر می‌شود انجام می‌دهیم و با

سیاست‌های معاملاتی پیشنهادی در همان لحظه معاملات انجام می‌دهیم.

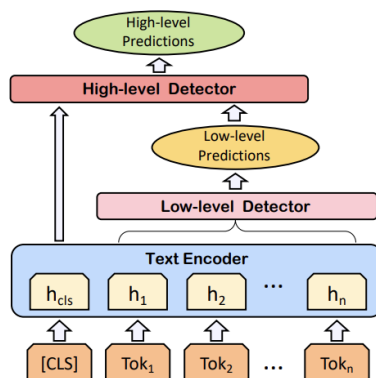


Figure 2: Model overview. The low-level detector identifies corporate events from each token, while the high-level detector summarizes the low-level predictions and the input representation to detect events at the article-level

این text encoder عه transformer عه

کار باحال دیگه: علاوه بر event بیایم اسم شرکت های موجود در رشته ها را در بیاریم بر اساس ticker شون سیاست خریدوفروش در انتها (Trade-At-End):

سهام شرکت "X" در اولین زمان ممکن پس از انتشار خبر خریداری می شود.

سهام به مدت k روز معاملاتی (مثلا 10 روز) نگه داشته می شود.

در پایان روز k ام، سهام به هر قیمتی که در بازار وجود دارد، فروخته می شود.

سیاست خریدوفروش با بهترین قیمت (Trade-At-Best):

سهام شرکت "X" در اولین زمان ممکن پس از انتشار خبر خریداری می شود.

در بازه زمانی k روز معاملاتی (مثلا 10 روز)، سهام به طور مداوم رصد می شود.

اگر در این بازه زمانی، قیمت سهام به بالاترین حد خود برسد، سهام فروخته می شود.

در غیر این صورت، سهام در روز k ام به قیمت بازار فروخته می شود.

راه :

(۱) شناسایی سیگنال های معاملاتی (مانند رویدادهای شرکتی یا احساسات) از مقالات خبری؛

(۲) برای هر مقاله که سیگنال های معاملاتی شناسایی شده اند، شرکت (تیکر) مربوطه را تشخیص دهید؛

(۳) داده های قیمت سهام شناسایی شده را در اطراف زمان انتشار اخبار به دست آورید؛

(۴) معاملات براساس سیاست های معاملاتی انجام دهید.

معیارهای ارزیابی:

نرخ سود (Winning Rate):

برای تراکنش های خرید، نرخ سود به صورت زیر محاسبه می شود: (قیمت فروش - قیمت خرید) تقسیم بر قیمت خرید
 ضریدر ۱۰۰.

برای تراکنش‌های فروش استقرای (Short-Selling)، نرخ سود به صورت زیر محاسبه می‌شود: (قیمت فروش - قیمت خرید) تقسیم بر قیمت فروش ضریر ۱۰۰.

نرخ سود بالا (Big Win Rate):

تعداد تراکنش‌هایی که سود آن‌ها حداقل ۱٪ باشد، بر کل تراکنش‌ها تقسیم می‌شود.

میانگین بازدهی هر تراکنش (Average Return per Transaction):

میانگین سود/زیان تمامی تراکنش‌ها محاسبه می‌شود.

بازده مازاد بر بازار (Excess Return):

عملکرد مدل با در نظر گرفتن عملکرد کلی بازار مقایسه می‌شود.

برای سنجش عملکرد بازار، از شاخص S&P 500 به عنوان معیار استفاده می‌شود.

بازده مازاد بر بازار برای هر مدل، برابر است با کل بازدهی مدل منهای بازدهی بازار.

فرضیه‌ها:

سرمایه‌ی اولیه برای هر مدل، ۱۰۰۰۰ دلار در نظر گرفته می‌شود.

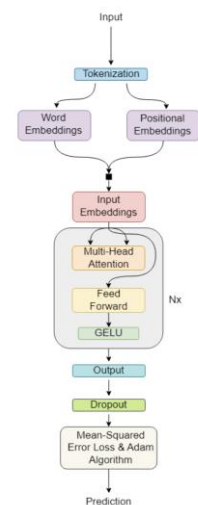
برای هر سیگنال معاملاتی، ۲۰۰۰ دلار سرمایه‌گذاری می‌شود.

در صورتی که میزان نقدینگی موجود کمتر از ۲۰۰۰ دلار باشد، ۲۰٪ از کل موجودی برای سیگنال معاملاتی جدید اختصاص داده می‌شود.

کارمزد ۰.۳٪ برای هر تراکنش در نظر گرفته می‌شود.

مقاله Natural Language Processing and Multimodal Stock Price Prediction

یک سهم ارزش 50 دلار کاهش 5 دلاری ارزش بیشتری دارد نسبت به یک سهم ارزش 6000 دلار که 5 دلار کاهش داشته باشد.



مقاله Integrating Stock Features and Global Information via Large Language Models for Enhanced Stock Return Prediction

استفاده ناکافی از اطلاعات معنایی: مدل‌های فعلی به طور کامل از اطلاعات معنایی غنی که LLMs ارائه می‌کنند، استفاده نمی‌کنند.

هماهنگی اطلاعات LLM با داده‌های موجود: ترکیب اطلاعاتی که LLMs از متن یاد می‌گیرند با داده‌های سنتی بازار سهام دشوار است.

embeddings تولید شده توسط LLMs و ویژگی‌های سهام به صورت ذاتی همان مدلی نیستند و در یک فضای معنایی یکپارچه هم‌تراز نیستند

در این مدل، ما دو بخش مهم داریم: مدل محلی و مدل جهانی. مدل محلی برای مدل کردن اطلاعات ویژگی‌های خاص هر سهم به کار می‌رود، مانند حجم، قیمت و ویژگی‌های فنی دیگر، برای پیش‌بینی بازدهی سهم. در عین حال، مدل جهانی اطلاعات جهانی را مدل می‌کند، به این معنی که تأثیر بازارها، صنایع و سیاست‌ها بر بازدهی سهم را مدل می‌کند. در این مدل، بازدهی هر سهم به صورت ترکیبی از اطلاعات محلی و جهانی محاسبه می‌شود. از طریق تابع F_{global} ، این دو نوع اطلاعات به هم ترکیب می‌شوند تا بازدهی پیش‌بینی شود. سپس با استفاده از تابع هزینه، که می‌تواند مثلاً هزینه میانگین مربعات باشد، ما سعی می‌کنیم که اختلاف بین بازدهی واقعی و پیش‌بینی شده را کمینه کنیم.

$$r_{t,i} = \alpha_{t,i} + \beta_{t,i} f_t, \quad (1)$$

where f_t represents the function that models the alignments between stock features and the global information at time t . We denote the Local and Global components as follows:

$$\alpha_t = F_{local}(M_{t-1}) \in R^{n_{t-1},1}, \quad (2)$$

$$\beta_t = F_{beta}(M_{t-1}) \in R^{n_{t-1},D}, \quad (3)$$

$$f_t \in R^{D,1}. \quad (4)$$

Here, $F_{local}(M_{t-1})$ and $F_{beta}(M_{t-1})$ represent the models that predict the α and β components, respectively. D is the desired dimensionality of the aggregated vector (i.e., the output dimension of f_t). Please note that $F_{local}(M_{t-1})$, $F_{beta}(M_{t-1})$, and f_t can be implemented using machine learning or deep learning models, such as Multi-Layer Perceptron (MLP). Based on the above descriptions, our proposed Local-Global model that estimates stock returns at time t is defined as follows:

$$\hat{r}_t = F_{local}(M_{t-1}) + \underbrace{F_{beta}(M_{t-1}) \cdot f_t}_{F_{global}(\cdot)}. \quad (5)$$

نگ دو بخش داریم بخش اول استفاده از اطلاعات محلی سهام برای predict قیمت سهام بخش دوم میشه اطلاعات خارجی طور حالا اطلاعات خارجی رو میتونیم با اطلاعات سهام های دیگ بدست بیاریم یا با llm ها یا با ترکیب هر دو

برای استفاده از سهام های دیگه از مکانیزم توجه میتونیم استفاده کنیم

$$K = M_{t-1} \cdot W_{\text{key}} \in \mathbb{R}^{n_{t-1} \times D}, V = M_{t-1} \cdot W_{\text{value}} \in \mathbb{R}^{n_{t-1} \times D}. \quad (6)$$

Next, we introduce a query vector $q \in \mathbb{R}^{1 \times D}$ and compute the attention weights a_{att} using the dot product between the query vector and the key matrix, normalized by their Euclidean norms:

$$a_{att} = \max \left(0, \frac{q \cdot K^T}{\|q\|_2 \cdot \|K\|_2} \right), \quad (7)$$

where the max function ensures that the attention weights remain non-negative. We then normalize the attention weights by dividing them by their sum to ensure that the attention weights sum up to 1:

$$a_{att} = \frac{a_{att}}{\text{sum}(a_{att})}. \quad (8)$$

Finally, the alignment function of Model 1 is defined as follows:

$$f_t^{\text{stock}} = (a_{att} \cdot V)^T \in \mathbb{R}^{D \times 1}. \quad (9)$$

برای llm میتونیم از لاما استفاده کنیم و بگیم که بیا اینم متن اخبار حالا توکن بعدی را پیش بینی کن و بردار رو در بیاریم بعدم با ی مدلی تبدیلش کنیم به فضایی که میخوایم مثلاً یک لایه خطی

Example of Prompt:

China implements a fishing ban period system in key water areas of the Yellow River Basin

Local development of culture, finance, nighttime economy, intellectual property protection, and others will receive priority support

Thailand has reported 63 cases of Omicron variant importation

...

{All the news headlines of the day}

Each day we generate one V_{llm} . It is worth noting that we have also explored various strategies for constructing few-shot prompts. For instance, we attempted to incorporate instructions such as analyzing market trends, predicting stock returns, and labeling each piece of information with industry and market-related tags. However, calculating the correlation between the daily V_{llm} generated by these different prompt strategies revealed a remarkably high correlation. Consequently, the performance was unsatisfactory. On the other hand, the prompt construction method we proposed exhibited a lower correlation and yielded superior results. This phenomenon suggests that when the correlation between daily V_{llm} is lower, there is increased information diversity, which is expected to supplement more global information.

To align the stock features and V_{llm} , we define parameters $W_{llm} \in \mathbb{R}^{D, d_{llm}}$ that can be learned by minimizing the loss function l , and the alignment function of Model 2 becomes

$$f_t^{llm} = W_{llm} V_{llm}^T \in \mathbb{R}^{D, 1}. \quad (10)$$

اگ بخوایم دوتا باهم

information:

$$f_t = f_t^{\text{stock}} \odot V_{\text{sparse}}, \quad (11)$$

where f_t^{stock} is the alignment function defined in Model 1, and V_{sparse} is a sparse vector (e.g., $[0, 0, 1, \dots, 1, 0]$) transformed from f_t^{llm} . The transforming processes will be presented in section 2.2. The element-wise multiplication (\odot) is performed to incorporate the sparse information into the global representation. Now, the complete estimation of stock returns \hat{r}_t becomes

$$\hat{r}_t = F_{\text{local}}(M_{t-1}) + F_{\text{beta}}(M_{t-1}) \cdot (f_t^{\text{stock}} \odot V_{\text{sparse}}). \quad (12)$$

حاصل ضرب بین مال سهام و مال llm رو میایم sparse هم میکنیم

اما این دوتا اطلاعات اونقد باهم همراستا نیستن پس میاد ی ri میزنه که احتمالات اینا را یکی کنه

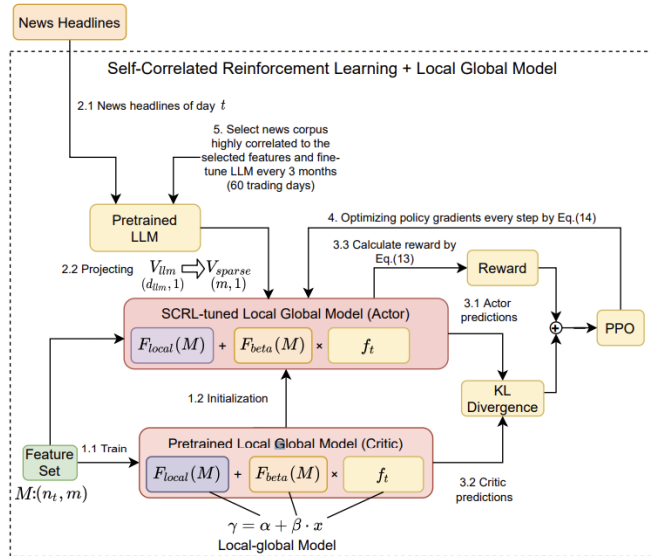


Figure 1: The Self-Correlated Reinforcement Learning with Local Global Model.

همزمان با اخبار انتخاب شده هم مدل llm اش رو finetune میکنه

ARIMA

Auto Regressive Integrated Moving Average is a statistical model class that is used for time series analysis and forecasting. It is a widely used approach in the field of econometrics for modeling and predicting economic data such as stock prices, GDP, and inflation rates. ARIMA models assume that time series data can be described as a combination of autoregressive (AR), moving average (MA), and differencing terms. The basic mathematical equation for an ARIMA(p,d,q) model:

$$y(t) = c + \varphi(1)y(t-1) + \dots + \varphi(p)y(t-p) - \theta(1)\varepsilon(t-1) - \dots - \theta(q)\varepsilon(t-q) + \varepsilon(t) \quad (1)$$