GPT's Idea of Stock Factors

Yuhan Cheng and Ke Tang

School of Management, Shandong University
Institute of Economics, Tsinghua University *

September 3, 2023

Abstract

We amalgamate the capabilities of the GPT-4 computational model with the avant-garde methodology of autonomous factor generation, culminating in the synthesis of high-return factors within the equity investment milieu. Empirical outcomes elucidate that the factors conceptualized by ChatGPT attain a commendable Sharpe ratio peaking at 4.49, accompanied by an annualized return trajectory reaching 66.16%. Notably, the superlative excess returns garnered remain unaccounted for by the quintessential five-factor model. Through the implementation of an unembellished model averaging paradigm, the ensemble of 35 factors, conceived by ChatGPT, manifests an apex long-short annualized return of 88% and a Sharpe ratio registering at 2.46. In stark contrast to conventional data mining techniques, the temporal expenditure requisite for GPT's factor generation is minuscule. It relies on knowledge inference without the need for data input, and it can provide a thorough economic explanation for its factors.

Keywords: LLM; Asset Pricing; Automated Factor Generating;ChatGPT;Knowledge Inference
JEL Classification: Financial Engineering

*E-mails: Yuhan Cheng:chengyuhan@sdu.edu.cn and Ke Tang: ketang@tsinghua.edu.cn. Special thanks to ...

1 Introduction

Factor utilization in financial markets has its origins in the 1960s, stemming from the seminal work of the Capital Asset Pricing Model (CAPM) presented by Sharpe (1964) in "Capital asset prices: A theory of market equilibrium under conditions of risk." Within this framework, risk was singularly attributed to the market factor. This paradigm was subsequently expanded upon by Fama and French (1992), who incorporated two additional factors: size and book-to-market value. Their construct, known as the Fama-French three-factor model, has since become a cornerstone in the annals of quantitative finance. As the depth and breadth of financial market data grew, buttressed by advancements in computational prowess, there ensued a proliferation of proposed factor models aimed at elucidating the intricacies of asset returns. Notable inclusions to this factor catalog have been momentum, volatility, quality, and investment style. For example, the seminal work of Carhart (1997) augmented the existing paradigm with a momentum factor, whereas Fama and French (2015) introduced an investment style dimension.

In contemporary financial research, the ascendancy of machine learning has imbued the factor selection process with a newfound complexity. Numerous investigations have been oriented towards harnessing the potential of machine learning, both in factor genesis and selection. This has birthed methodologies such as deep learning for factor synthesis and reinforcement learning for factor curations. Notably, Kanter and Veeramachaneni (2015) presented an avant-garde algorithm, Deep Feature Synthesis, adept at autonomously gleaning features from intricate relational datasets. In a parallel vein, Lam et al. (2017) unveiled the One Button Machine, a system designed to replicate key data scientist functions, notably in automatic feature discovery. Further innovations were observed with Fang et al. (2019), who propounded the Alpha Discovery Neural Network – a neural architecture primed for crafting diverse financial indicators. This was furthered by Fang et al. (2020) and their introduction of NNAFC, a neural tool amalgamating financial acumen with neural networks, with the overarching aim of refining investment stratagems.

In their stride towards further innovation, Zhang et al. (2020b) showcased the AutoAlpha evolutionary algorithm. This methodology, while recognizing patterns within Alpha factors, offers a structured search mechanism adept at expeditiously exploring potential search space domains. The PCA-QD approach, nested within this, refines the search parameters by mitigating redundancy and amplifying optimization. Diverging slightly, Zhang et al. (2020a) introduced the DoubleEnsemble framework, employing learning trajectory re-weighting coupled with feature selection ensembles, thereby facilitating the identification of salient samples and features. Lastly, Cui et al. (2021) advanced the discourse with AlphaEvolve, an AutoML-infused Alpha mining apparatus. Unique to this is its capacity to meld relational domain expertise, paving the way for discerning intricate inter-stock relationships.

In recent financial research endeavors, the FETCH framework proposed by Li et al. (2022) has garnered attention for its dual capabilities in feature generation and selection. Particularly salient is its prowess in facilitating the seamless transfer of feature engineering across heterogeneous datasets. Complementing this, Zhang et al. (2023b) introduced OpenFE, an automated feature generation framework, which, when empirically tested on a myriad of benchmark datasets, exhibited competencies analogous to those of seasoned machine learning experts. In a novel approach, Yu et al. (2023a) posits a model rooted in reinforcement learning methodologies, advocating for a comprehensive exploration of the search space. Eschewing traditional filtration techniques, this innovative paradigm customizes the creation of alpha factors to optimize performance metrics.

The foundation for many contemporary factor generation methodologies is anchored in data mining. Here, a plethora of factors are stochastically generated, with only the apex ones being retained. Notwithstanding their empirical successes, the adoption of such techniques in the economic landscape has faced critiques, primarily attributed to the absence of a cogent economic rationale. Furthermore, the robustness of out-of-sample predictions of factors derived via data mining remains a topic of contention. In juxtaposition to these pure data mining methodologies, ChatGPT possesses the capability to elucidate these factors, rendering comprehensive insights underpinned by economic theories. Distinctly divergent from exhaustive exploration and traditional mining, ChatGPT asserts that the factors it manifests are the synthesis of financial scholarship, interdisciplinary knowledge integration, and mathematical rigor. A noteworthy observation is the expedited factor generation process facilitated by ChatGPT compared to traditional paradigms. This is attributed to ChatGPT's reliance on knowledge-based inference rather than data-intensive approaches, empowering it to construct factor calculation algorithms without the obligatory financial data.

In the contemporary research milieu, there has been an unmistakable ascendancy of Large Language Models (LLMs) such as ChatGPT. These linguistic behemoths have demonstrated unparalleled competencies across multifarious tasks, prompting scholars to explore their applicability in financial domains. In a seminal contribution, Araci (2019) delineated FinBERT—a financial lexicon-oriented model inspired by the BERT architecture. It set new benchmarks in financial sentiment analysis. Jiang, Kelly, and Xiu (2022) harnessed LLMs in forecasting financial returns, adeptly gleaning insights from a corpus of financial news articles. Breaking traditional bounds, Wu et al. (2023) introduced BloombergGPT, tailored specifically for financial tasks, registering remarkable proficiency.

Transitioning to the realm of personal finance, Lakkaraju et al. (2023) delineated 13 cardinal queries, encapsulating a spectrum of banking dimensions, from accounts to credit instruments, and extending to temporal deposits. This exploration also encompassed multifaceted interactions, such as high-value acquisitions and financial advisory engagements. Albeit their rigorous efforts, the resultant insights were somewhat muted. Yu et al. (2023b) highlighted the intricacies and impediments in leveraging machine learning for financial chronologies, shedding light on challenges such as data cross-referencing and historical data integration. Yet, their research underscored the latent potential of LLMs in decoding financial news timelines and asset pricing dynamics.

In a magisterial exposition, Xie et al. (2023) unveiled PIXIU—a confluence of FinMA (enhanced with LLaMA), a pioneering financial LLM, a curated fine-tuning dataset, and a multifarious evaluation benchmark. On a parallel note, Zhang et al. (2023a) presented a meticulous backtesting paradigm tailored to the nuances of the Chinese financial ecosystem. They evaluated a cohort of LLMs, including ChatGPT, for their aptitude in extracting sentiment indicators from Chinese news abstracts. The revelations ignited discourses on the underpinnings of LLM efficiency in sentiment analysis. Shah and Chava (2023) orchestrated a comparative study of ChatGPT against other generative LLMs, illuminating nuances in data annotation, performance deltas, and the broader implications of generative models in finance. Their observations accentuated that while ChatGPT exhibited resilience even without labeled data, models with fine-tuning typically demonstrated superior outcomes. Culminating the discourse, Son et al. (2023) delineated the vast potentialities of LLMs in financial applications, emphasizing the burgeoning prowess of astute financial inferencing, particularly when armed with refined instructional paradigms and voluminous data reservoirs.

This study pioneers the fusion of auto-generating factors with the sophisticated GPT-4 large language model, culminating in the generation of return-maximizing factors specific to the domain of stock investment. The procedural elegance lies in its simplicity: we provide ChatGPT with concise information regarding our data structure, encompassing variables such as opening prices, closing prices, and trading volume. This input serves as a directive for ChatGPT to architect novel factors and concurrently generate the requisite computational code. Notably, this resultant code can be executed to return profitable trajectories, obviating the need for conventional mutual information filtering paradigms.

The empirical results bear testimony to the exceptional caliber of factors contrived by ChatGPT. A salient observation is that over fifty percent of these factors engender alphas that remain inexplicable by both the three-factor and five-factor models. Remarkably, the apex factor registered an annualized return rate of 66.16%. A considerable proportion of these factors manifest significant alpha magnitudes, eclipsing those formulated by human experts anchored in financial scholarship. Furthermore, they markedly supersede factors extrapolated through data mining techniques. This underscores the unparalleled efficacy and latent potential of ChatGPT in the realm of factor generation, paving the way for novel paradigms in stock investment.

Nevertheless, a caveat warrants attention: GPT-4 exhibits limitations in its ability to precisely identify the prospective trajectory of its synthesized factors, particularly regarding the decision to invert the factor value. However, both static and dynamic multi-factor investment portfolios curated with these factors demonstrate robust returns and statistically noteworthy alphas. This attests to the general efficacy of GPT-4 generated factors, corroborating their capability to enhance investment returns.

In encapsulation, the present research marks a watershed moment in asset pricing academia. By adroitly intertwining GPT-4 with the art of factor generation, we have birthed factors of unparalleled quality. This augments the repertoire of tools available to investors and scholars, proffering an astute and efficient avenue for factor discernment and origination, thus expanding the horizon for stock investment pursuits.

The distinctiveness and innovative attributes of this investigation can be delineated as follows:

1.Efficacy in Factor Generation: The GPT model manifests an unparalleled proficiency in factor generation, returning factors with return metrics that outstrip those delineated in canonical financial literature. This prodigious performance enshrines the GPT as an instrumental asset, endowing both academicians and financiers with sophisticated methodologies for factor discernment and origination.

2.Optimization of Research Workflow: Beyond sheer return augmentation, the GPT model offers a transformative approach that mitigates the labor-intensiveness inherent to traditional asset pricing endeavors. Conventional factor elucidation mandates sustained human effort characterized by repetitive analyses and validation cycles. In stark contrast, the GPT model's inherent autonomy curtails this prolonged sequence, curtailing human-induced latencies in factor investigation. This facilitates a more expedited transition from factor conceptualization to pragmatic deployment, allowing scholars to divert their attention to facets such as the nuance of trading algorithms.

3.Paradigmatic Shift in Asset Pricing Research: Our proposition to incorporate the GPT model into the tapestry of asset pricing scholarship heralds a radical departure from entrenched research archetypes. Predominant empirical asset pricing frameworks are anchored in a symbiotic fusion of human sagacity and financial erudition for factor postulation and its subsequent corroboration. The GPT model's advent, with its capacity to emulate extensive manual research hours in a temporally compressed manner, suggests an impending tectonic shift. Its capabilities in autonomous feature extraction and selection signal an expedited trajectory for unearthing avant-garde factors, thereby catalyzing innovations in asset pricing pedagogy.

4.Pioneering Trajectory in Factor Research: The facility to engage in discourse with ChatGPT and swiftly tap into its expansive financial knowledge repository positions it as a game-changer. By doing so, researchers can expeditiously derive predictive pricing factors, virtually obviating temporal inefficiencies inherent to human-centric factor ideation. This modus operandi is poised to redefine the fabric of factor research, catalyzing paradigmatic evolutions in both scholarly and industrial milieus.

This methodology represents a marked departure from extant investigations that juxtapose sizable models with the domain of finance. Whereas antecedent scholarship was predominantly circumscribed to sentiment analysis, the present exposition accentuates the expansive financial acumen inherent to the GPT, spotlighting its superlative capacity for factor origination. The salient advantages encompass:

1.Amplified Accessibility and Evolutionary Alignment: The research modality obviates the imperatives for localizing gargantuan language models, opting instead for interfacing directly with the model. In light of the fact that GPT-4 is emblematic of over a trillion neuronal parameters, the proposition of localizing colossal models becomes progressively untenable. Coupled with the recognition that the open-source ethos may not be universally espoused by all future models, finance-centric algorithms predicated on model localization confront substantial impediments in leveraging non-open-source or incrementally large-scale models. Over an extended temporal spectrum, algorithms that eschew model localization resonate more harmoniously with the evolutionary arc of massive language model advancements.

2.Expedited Deployment and Enhanced Efficacy: This stratagem engenders rapid deployment, mitigating computational latencies and economizing on modeling engineers' temporal commitments. Albeit the interface response latency of GPT might be perceived as protracted and incurring operational expenses, the algorithm circumvents inundating the model with copious data, restricting interactions to dialogues not exceeding 20,000 lexemes. This enterprise primarily enjoins researchers to architect the backtesting and visualization subroutines, with the factor computational logic being directly bequeathed by the GPT.

3.Elevated Economic Significance: Factors engendered pursuant to this methodology possess pronounced economic salience. Each factor is delineated with a lucid mathematical expression, complemented by an executable Python script conferred by the GPT, thereby streamlining the journey for scholars to extract economic insights.

To encapsulate, the integration of GPT in the realms of factor generation and asset pricing scholarship not only culminates in premium factor returns but also orchestrates a dramatic abatement in human latencies. This augments research efficacy, propelling asset pricing paradigms towards heightened efficiency and sagacity. The resultant innovations enhance the robustness and acumen of investment adjudications and the refinement of investment strategies.

2 Process of GPT's Factor Generation

In this study, we position ChatGPT as a surrogate financial savant, tasking it with the origination of unprecedented features (or factors) poised to elucidate prospective stock returns. To bolster the external validity of our findings while tempering computational overheads, ChatGPT remains privy solely to the schema of data storage, remaining oblivious to the substantive data therein. To obviate potential ambiguities inherent in textual delineations or mathematical notations proffered by ChatGPT, we solicit the direct extrapolation of the computational algorithm pertaining to the factors from ChatGPT. Subsequent to this, scholars are enjoined to transplant this computational algorithm onto their local computational environments for the actual factor derivation.

We refrain from commissioning ChatGPT for the actual computation of factor magnitudes. Contrarily, our mandate for ChatGPT is to elucidate the computational modality germane to the factor. Broadly, the factor magnitude vector, encompassing all equity instruments at temporal juncture t, is formulated as a function contingent on market idiosyncrasies spanning temporal juncture t and its antecedents: $z_{it} = f(X_t, X_{t-1}, \cdots)$. Here, X_t epitomizes the cross-sectional data at the temporal juncture t. ChatGPT, with an acumen of the dataframe storage of $(X_t, X_{t-1}, \cdots)$, discerns its columnar designations and the semantics inherent to each column's contents. Anticipations are vested in ChatGPT to proffer a potentially lucrative function f drawn from its expansive knowledge reservoir, ensuring the innovativeness of f. To further insulate against the peril of ChatGPT reverting to factors already chronicled in the extant literature, we impose a stipulation for the manifestation of a multifaceted factor. As a further safeguard, a meticulous manual scrutiny of the extrapolated factor is undertaken to affirm its novelty.

3 Factors and Data

Leveraging the computational prowess of GPT-4, we synthesized a cohort of factors, subsequently isolating 35 that flawlessly integrated within the coding framework. Conscious of the "forward-looking" dilemma and the temporal constraints of GPT's training data, which is delimited at September 2021, our retrospective analysis employed data from October 2021 to December 2022, sourced from the CRSP's U.S. equity market dataset.

4 Empirical Results

4.1 Single Factor Portfolio

Our empirical methodology comprised a daily rebalanced long-short factor return analysis from October 2021 through December 2022. In the purview of a long-only modus operandi, equity investments were channeled towards stocks manifesting factor values in the upper 50th percentile. Conversely, for the long-short paradigm, long allocations were directed towards the top 50% of stocks by factor value, juxtaposed with short positions in the lower 50%. Preliminary observations delineated that while GPT-conceived factors presented formidable stock return predictability, the model encountered predicaments in unambiguously discerning the sign of the factors, namely discerning the directionality of stock returns as a function of factor values. In the wake of this ambiguity, empirical outcomes of the factors were adopted as the evaluation benchmark.

Delving deeper, a performance assessment of each factor was executed spanning 2000 to September 2021. Factors demonstrating positive performance retained their original values during the factor return analysis, while negatively performing factors underwent value inversion prior to the analysis. The longitudinal performance is comprehensively elucidated in Table 1.

Under the long-short schema, merely 5 factors registered negative returns, while the lion's share of 30 factors championed positive returns. Factor28 emerged as the epitome of performance, realizing an astoundingly robust annualized return of 66.16%, closely trailed by factor24, boasting 55.347%. Factors 4, 30, and 35 also underscored their potency with returns transcending 40%. Cumulatively, the high-performing echelon of factors is characterized by 9 entities exceeding 20% in annualized returns, while 21 surpassed the 10% threshold.

Examining this through the lens of the Sharpe ratio, a predominantly direct correlation between return magnitude and Sharpe ratios is evident, albeit with outliers like factor11. The pinnacle of Sharpe ratios is represented by factor24, registering 4.4913, narrowly eclipsing factor28. Notably, this signifies factor24's marginally diminished volatility, a claim substantiated by its maximum drawdown rates. Factors exhibiting returns beyond the 40% benchmark consistently flaunted Sharpe ratios beyond 3.

In terms of drawdown metrics, a substantial 24 factors were characterized by drawdown rates beneath 10%. Of these, only 6 factors navigating the positive return space surpassed the 10% drawdown threshold. It's pivotal to highlight the peak drawdown rate, which, for the high performing cadre, was capped at 13.9%. Factor30 was the paragon of resilience, with a minuscule drawdown rate of 3.4246%, thereby bolstering its

Sharpe credentials. Several other factors also manifested commendable drawdown resistance, cementing their robustness.

The long-only strategy is dissected in Table 2, unveiling the annual returns, Sharpe ratios, and peak drawdowns corresponding to each factor. A striking observation is the ubiquitously positive returns across all factors. Factor6 manifests the least potent performance, with an annual return of just 8.7699%, ironically contriving negative returns under a long-short modus operandi. All other contenders transcend the 10% annualized return mark. Factor28 ascends as the performance titan, charting an annualized return of 53.899%. Nipping at its heels are factor4 and factor24, showcasing returns of 44.842% and 44.324%, respectively. A high-performing league emerges with seven factors surmounting 30% in annual returns. This elite cadre, in addition to the aforementioned, envelops factors 27, 30, 31, and 35. Notably, factors 30 and 35 vault beyond the 40% echelon. Zooming out, 24 factors within the long-only architecture eclipse the 20% annualized return threshold.

Navigating through the terrain of risk-adjusted returns, Factor28 remains unchallenged with its Sharpe ratio north of 1.9. Trailblazing behind are factors 30, 24, and 4, each brandishing Sharpe ratios soaring beyond 1.6. Barring factors that languish in the negative under the long-short framework, every other factor musters a Sharpe ratio above 0.7, with a laudable 18 breaching the 1.0 landmark.

Juxtaposing the long-only and long-short strategies reveals a predilection of the former towards elevated drawdowns. Factor28, while leading in returns and Sharpe metrics, also bears the brunt of the highest drawdown, peaking at 27.278% among its positively returning peers in the long-short realm. In stark contrast, factor20 emerges as the paragon of resilience, restricting its drawdown to a modest 17.993%.

A subsequent layer of nuance is introduced by evaluating alpha returns, showcased in Table 3. The analytical spectrum bifurcates into two temporal windows: an extended span from 2003 to 2022, using 2000-2002 returns as a directional reference, and a narrower window mirroring our earlier tests, anchored to 2000-September 2021 reference returns, spanning October 2021 through December 2022.

The tableau reflects compelling alpha generation for the September 2021-December 2022 stretch. Nine factors — 4, 13, 21, 24, 27, 28, 30, 31, and 35 — indomitably stand out, their alphas achieving statistical significance at the stringent 1% threshold. Bar factor13 and factor21, the cohort, under the long-only prism, breaches the 30% return barrier. We also discern alphas of factors 19, 22, and 25 attaining significance at the 5% level, and a spate of factors — 2, 3, 6, 8, 11, 17, and 22 — securing their place at the 10% significance echelon. Factor28 is particularly stellar, parading an alpha of 0.74355, emblematic of its prowess in consistently outmaneuvering market returns. Post value inversion, its long-short strategy return of 66.16% vastly outpaces its peers, consonant with its high alpha. Factor24's alpha prowess is also palpable at 0.61935, harmonizing with its formidable 55.347% return under the long-short schema. Factors 4, 30, 31, and 35 also engrave their alpha supremacy, each registering above 0.4, reflecting their long-short strategy returns that comfortably exceed 37%.

Notably, among the top five factors in terms of alpha values—factor28, 24, 35, 31, and 30—their annualized returns under daily rebalancing in the long-short strategy consistently exceed those in the long-only strategy, further highlighting the strong predictive capabilities of these factors. Among the factors with statistical significance at the 10% level, factor8 holds the lowest alpha value, yet it still reaches 0.13636. In absolute terms, only 6 factors exhibit three-factor and five-factor alphas below 0.5.

The alphas obtained from testing spanning from 2003 to 2022 are even more prominent. Notably, only 5 factors exhibit negative alphas, while a total of 23 factors display positive alphas that are statistically significant at the 1% level. Additionally, 3 factors exhibit alphas that are statistically significant at the 5% level, and 1 factor demonstrates significance at the 10% level. Factor28 maintains its position with the highest alpha at 0.39464, followed by factor24 at 0.35665. Factor11, 14, 23, 27, 29, and 30 also exhibit alphas exceeding 0.2. This demonstrates the enduring predictive power of the factors generated by GPT, as evidenced by their ability to consistently deliver significant market outperformance over a 20-year backtesting period.

4.2 Multi-factor Portfolio Investment

Among the quintessential factors delineated by their alpha magnitudes—namely factor28, 24, 35, 31, and 30—their annualized returns under a daily rebalancing regime in the long-short stratagem consistently outstrip those accrued under the long-only framework. This accentuates the potent prognosticative prowess of these specified factors. Delving into factors with a statistical significance demarcated at the 10% threshold, factor8 registers the most diminutive alpha magnitude, albeit clocking in at 0.13636. In absolute metrics, a mere sextet of factors manifest three-factor and five-factor alphas that fall below the 0.5 demarcation.

The alpha trajectories extrapolated from the temporal analysis encompassing 2003 to 2022 are even more conspicuous. A paucity of factors, totaling five, exhibit negative alphas. In stark contrast, a compendium of 23 factors unveil positive alphas achieving statistical significance at the rigorous 1% standard. Further dissection reveals 3 factors with alphas significant at the 5% criterion, while a singular factor exhibits significance at the 10% marker. Factor28 perpetuates its preeminence with an alpha cresting at 0.39464, with factor24 in proximate pursuit at 0.35665. Additional constituents, namely factor11, 14, 23, 27, 29, and 30, demonstrate alphas transcending the 0.2 threshold. This underscores the sustained predictive efficacy of the factors engineered by GPT, as corroborated by their consistent capability to outpace the market over two decades of backtesting.

An analysis of factors catalyzed by ChatGPT attests to their formidable predictive robustness. This intimates that an elementary aggregation of these factors can precipitate significant returns. As previously alluded, GPT grapples with discerning the polarity of a factor signal. Ergo, rudimentary aggregation may inadvertently amalgamate negative factors with affirmative coefficients. Notwithstanding this inherent foible potentially attenuating the model's efficacy, the returns from model aggregation persist in their impressiveness.

Beyond the rudiments of aggregation, an elementary dynamic model aggregation stratagem was also probed. The only deviation between dynamic and static aggregations hinges on the historical return trajectory dictating a factor's polarity. This elementary algorithm culminated in a model return of 0.37593.

For a comprehensive assessment of the efficacy of factors molded by GPT-4, we embarked on backtesting multi-factor investments anchored to historical datasets. On a diurnal basis, we normalized factor magnitudes based on antecedent data and forged both long-only and long-short portfolios grounded in the aggregation of these normalized values. The normalization leveraged was the Min-Max normalization method, explicated as:

$$\frac{x - min(x)}{max(x) - min(x)} \tag{1}$$

Initially, a static multi-factor portfolio was postulated. Explicitly, the polarity of each factor—either positive or negative—was deduced from their long-short returns spanning 2000 to September 2021. Subsequent aggregation returned $average_{t,stock}$. The term $sign$ epitomizes the vector that ascertains the positive or negative delineation of each factor. $sign_i$ resonates with the polarity of the returns for the i-th factor from 2000 to September 2021, oscillating between values of 1 or -1. The computation is elucidated as:

$$average_{t,stock} = \frac{1}{n} \sum_{i=1}^{n} factor^{i}_{t,stock} \times sign_i \tag{2}$$

Where $factor_{t,stock}$ symbolizes the factor magnitude of the stock on day t, and $factor^{i}_{t,stock}$ designates the i-th factor. Given the ensemble of 35 factors, n is consequently set at 35.

Following the articulated procedure for multi-factor portfolio configuration and its subsequent daily recalibration, the return dynamics are encapsulated in Table 4.

With respect to the long-only portfolio, it registered an annual return trajectory of 45.161%, complemented by a Sharpe ratio of 1.6295. When juxtaposed against single-factor portfolios, the cumulative return of the multi-factor ensemble may marginally lag behind the apex performer, factor 28. However, it eclipses the returns of all other individual factors. The maximum drawdown metric, recorded at 25.439%, aligns closely with the majority of the factors when assessed within the paradigm of long-only stratagems.

Transitioning to the long-short landscape, the annualized return underwent a decrement to 39.728% relative to its long-only counterpart. Yet, the Sharpe ratio soared to an impressive 2.4854, with the maximum drawdown receding to a mere 7.1795%. Such metrics insinuate that despite the attenuation in absolute returns relative to the long-only spectrum, the affiliated risk was substantially curtailed. Long-only portfolios inherently exhibit heightened susceptibility to market vicissitudes, paving the way for potential substantial capital erosions or appreciations during specific temporal windows. Conversely, long-short portfolios adeptly neutralize the capriciousness engendered by market fluctuations.

Our investigative trajectory further embraced an uncomplicated modus operandi for dynamically modulating the polarity of factor magnitudes. Explicitly, subsequent to the daily normalization of factor magnitudes predicated on their historical trajectories employing the Min-Max algorithm, the polarity (affirmative or negative) of every factor was anchored to the cumulative historical return up until the specified day.

To elucidate, assuming $return_t$ embodies the return vector for every discrete factor on day t, and $factor_{t,stock}$ epitomizes the Min-Max standardized factor magnitude of a specific stock on day t, we initially ascribe a polarity to the factor magnitude predicated on its return, culminating in the recalibrated factor magnitude, $newfactor_{t,stock}$:

$$newfactor_{t,stock} = factor_{t,stock} \times sign(return_t) \tag{3}$$

Subsequent to this, the diurnal multi-factor mean for the stock is deduced as:

$$average_{t,stock} = \frac{1}{n}\sum_{i=1}^{n} newfactor^{i}_{t,stock}, \tag{4}$$

with $newfactor^{i}_{t,stock}$ representing the i-th factor magnitude. This mean, $average_{t,stock}$, is then harnessed as the novel factor for portfolio orchestration.

In our empirical evaluations, we deployed both the individual factor returns stemming from daily recalibrated long-only and long-short portfolios as benchmarks for polarity determination. The resultant metrics are also encapsulated in Table 4.

A discerning observation in a dynamically modulated setting, benchmarked against the long-only returns of individual factors, unveils that the annualized return for the long-only stratagem reached 0.34044. The long-short stratagem, in contrast, touched a commendable 0.37593. Notably, both these values markedly transcend those observed in the static configuration.

The analytical insights gleaned from the long-short returns unveil that the circumvention of certain market nuances endows the GPT-4 generated factor with a capacity to consistently realize stable and appreciable returns. The dynamism in factor value adjustment, which astutely accounts for sporadic market inversions, does dual service—mitigating the risks associated with reversals while capitalizing on potential profit avenues. Such dynamics further underscore the potency and utility of the GPT-derived factor in portfolio management.

Drawing a parallel between employing the daily rebalanced long-short portfolio return and the long-only returns as the yardstick for discerning the polarity of return outcomes, the dynamism in returns under the former seems less accentuated. The plausible reason for this attenuation could be that long-short returns, to a degree, fall short in echoing the undulating market currents. This results in a tardy capture of the factor's inversion signals, rendering the factor value adjustments less agile and efficient. Yet, the annualized return for the portfolios—both long-only and long-short—stands at 0.30262 and 0.30036 respectively. Paired with Sharpe ratios of 1.2644 and 2.4478, these figures denote tangible advancements over their static counterparts.

Delving deeper, the alphas computed for both static and dynamic scenarios are encapsulated in table 5. With all alphas consistently surpassing the 0.3 mark, the alphas pertaining to static investment and dynamic investment (with long-only data as a reference) both exhibit 1% statistical significance. Contrarily, the alpha for dynamic investment, referenced against long-short data, manifests a 5% statistical significance—potentially skewed by some preliminary setbacks. In a three-factor regression context, the zenith of alpha is seen for dynamic investments using daily long-only data, registering at 0.38067, while the static counterpart is close on its heels at 0.37762. Conversely, in a five-factor regression landscape, the alpha pinnacle is associated with static investments, recording 0.37034, while the dynamic counterpart referencing daily long-only data posts 0.36583—a decrement roughly amounting to 0.015 vis-à-vis the three-factor backdrop.

The numerically robust and statistically significant alphas further embellish the overarching credibility of the GPT-4 derived factor. Such empirical evidence attests to the proposition that investment frameworks pivoted on GPT-4 generated factors are poised to accrue elevated returns.

For both the static and dynamically calibrated multi-factor portfolios, an additional layer of analysis was undertaken. Here, portfolios were architecturally designed by ranking factor magnitudes in descending order and then uniformly allocating weights to the top decile. The associated annualized return, the Sharpe ratio, and the maximum drawdown metrics are tabled in 6.

In these tabulations, the return differential between the quintile with the apex factor magnitude and its antithetical counterpart with the nadir factor magnitude is showcased. This essentially reflects the strategy of adopting a long-only stance on the upper decile stocks as per factor magnitudes while simultaneously shorting the bottom decile. Remarkably, under the static paradigm, this long-short portfolio configuration records an astounding annualized return of 0.88097. Transitioning to the dynamic framework, referenced against long-only returns, the return metric dips to 0.41481. However, a dynamic strategy based on long-short returns registers an enhanced annualized return of 0.69870. Furthermore, the Sharpe ratio for the quintile housing stocks with the zenith factor magnitude consistently overshadows its counterpart with the nadir factor magnitude by a margin exceeding a unit. In the static construct, the long-short portfolio's Sharpe ratio, clocking 2.4662, testifies to its stellar performance credentials.

Overall, equities characterized by elevated factor values tend to manifest superior returns and Sharpe ratios relative to their counterparts with diminished factor values. Notably, the annualized return does not exhibit a strict inversely proportional relationship with diminishing factor values, with group 2 serving as a prominent outlier. The quintile labeled as "High" consistently boasts the pinnacle Sharpe ratio, registering 2.0391 under the dynamic investment modus operandi predicated on long-short returns. This translates to an annualized return of 0.73425, marking an approximate surplus of 0.163 over group 9, which posts an annualized return of 0.57093. A singular deviation from this trend is observed in the dynamic investment strategy referenced against long-only returns; here, the "High" quintile's annualized return stands at 0.50977, trailing both group 9's 0.52126 and group 8's 0.613. Nevertheless, this figure substantially eclipses the returns of other quintiles. In the static paradigm, the annual return of the "High" quintile is gauged at 0.66492, markedly surpassing other cohorts, with group 9 recording 0.57613 and the remainder falling below the 0.4 threshold.

The empirical outcomes lend credence to the proposition that factors engineered by GPT are adept at returning remarkable returns. These findings insinuate that such factors harbor the capacity for versatile integration into a diverse array of investment portfolio structures. It can be rationally posited that these economically cogent and efficacious factors, sculpted by GPT, are poised to consistently proffer optimized returns in pragmatic investment landscapes.

5 How GPT Constructs the Aforementioned Factors

5.1 Principles and Fundamental Logic of GPT's Factor Generation

In seeking to decipher the methodology through which GPT operationalizes the generation of these factors, our inquisitiveness led us to consult GPT directly. It articulated its factor formulation protocol in the subsequent manner:

1. Understanding the Requirements and Data: First, it's crucial to know the data we are working with (e.g., OHLCV data) and the requirements, such as the type of stocks, trading frequency, investment horizon, etc.

2. Exploring Existing Indicators: There's a rich library of existing technical and fundamental indicators in finance. Many factors are based on these well-known indicators, such as Moving Averages, Bollinger Bands, MACD, etc.

3. Combining Indicators: By combining different indicators, we can create complex factors that encapsulate different aspects of the market, such as trend, momentum, volume, and volatility.

4. Applying Financial Theories: Sometimes, the creation of factors is based on specific financial theories or investment philosophies. For example, some factors may focus on identifying value stocks, growth stocks, or market inefficiencies.

5. Experimenting with Mathematical Transformations: Applying mathematical or statistical transformations to price, volume, or other data can uncover hidden patterns or normalize the data in a useful way.

6. Grouping and Rolling Operations: Utilizing grouping by assets and applying rolling computations allows us to calculate factors in a way that respects the time-series nature of the data.

7. Iterative Testing and Refinement: Creating a factor often requires several iterations of testing, refinement, and validation. The effectiveness of a factor can be assessed through back-testing, out-of-sample testing, and statistical analysis.

8. Incorporating Domain Knowledge: Sometimes, unique insights from a particular industry, market conditions, or company specifics can inspire the creation of a novel factor.

9. Ensuring Practicality: Factors need to be computable from available data, robust to noise, and applicable in real-world trading scenarios.

The process through which GPT conjures these remarkably prognosticative factors remains an intricate conundrum in financial economics. Nonetheless, endeavoring to delineate the operational intricacies underpinning the intricate outputs emanating from such an expansive computational model poses substantial challenges. ChatGPT orchestrates its computations via a conspicuously nonlinear trajectory, and given that the GPT-4.0 iteration, which is the focal point of our study, is fortified with in excess of a hundred billion parameters, endeavors at interpretative dissection are inherently intricate. Providentially, the interactive capability of ChatGPT accords us the privilege to solicit its elucidation on the factor construction methodology. Through our exhaustive dialogues, we managed to crystallize four pivotal tenets.

Focusing on Mathematical Soundness: Since I don't have personal intuition, I ensure that the combinations I suggest have a sound mathematical basis. This might include ensuring that the indicators are compatible in terms of scale, distribution, or statistical properties.

Understanding Underlying Principles: I use the underlying mathematical and statistical properties of the indicators to think about how they might combine meaningfully. For example, if one indicator measures trend and another measures volatility, combining them might provide insight into the momentum of a trending market.

Considering Market Context: Based on historical data and industry practices, I can make educated guesses on how certain indicators might behave together in different market conditions.

Applying General Analytical Techniques: I can apply standard mathematical and statistical techniques, such as normalization, weighting, and statistical correlations, to combine indicators in a logically coherent way. Based on its self-interpretation, GPT benefits immensely from a vast Market Context and comprehends the underlying semantics and logic of the factors. It can rationally combine foundational factors based on its grasp of Underlying Principles. This combination adheres to basic Mathematical Soundness and general Analytical Techniques while being inspired by its vast Market Context.

This insight is simultaneously intuitive and revelatory. It underscores ChatGPT's prowess as a formidable financial analyst, furnished with an extensive repository of fiscal acumen. Its capability to autonomously synthesize and amalgamate this vast expanse of financial information without fatigue, and at an inconsequential expense, foregrounds the transformative implications of expansive linguistic models for both the financial sector and scholarly endeavors.

ChatGPT elucidates: "In the exemplars I've delineated, my approach amalgamates recognized technical markers in diversified configurations, occasionally integrating mathematical transmutations and statistical indices to engender novel factors. Such syntheses are predominantly steered by heuristics pertaining to potential efficacious signals (e.g., the conjunction of trajectory and volume data). However, these heuristic concoctions are customarily subjected to rigorous empirical validation to affirm their pragmatic efficacy."

To encapsulate, through the adroit exploitation of established technical indicators, integrating mathematical operations—like computing standard deviations and averages—and by evaluating market prices, oscillations, trajectories, and inversion cues, GPT meticulously engineers market-optimized factors with the potential to procure substantial returns.

5.2 Basic Factors and Strategies Referenced by GPT

To deeply probe the nexus between the factors manifested by GPT and canonical financial literature as well as pragmatic fiscal experiences, we embarked on an analytical journey scrutinizing the rudimentary operations and factors that GPT leverages. Predominantly, its foundational operations draw inspiration from the numpy and pandas libraries, embracing rudimentary functions such as groupby, rolling, std, kurt, while also extending its arm into the realm of sophisticated functionalities like Exponential Moving Average. As GPT elucidates, its elemental factors are meticulously curated from **Financial Literature** and **Known Strategies**.

Each factor GPT has unveiled is inherently a composite entity. Elucidated within Table 7 are the elemental factors GPT leverages to synthesize these composite structures, accompanied by their comprehensive explications. Each factor emblematically encapsulates treatments of various stock metrics such as open price, close price, highs, lows, and volumes—these treatments range from calculating standard deviations, momentum to rate of change. Within these factors, one can discern traditional market technical indicators, epitomized by the likes of MACD and daily averages. Some components, however, elude direct correspondence to prevalent factors within extant databases, emerging more as GPT's ingenious improvisations. For instance, the factor12 element, christened as Trading Range Expansion (TRE), gauges the differential between the current trading band and its counterpart from two days hence. This contrasts the established Range Expansion Index. Additionally, Momentum and Intraday Range Interaction (MIRI) from factor 5 stand apart from the pantheon of recognized financial markers; its digital footprint is virtually non-existent.

Contrasting these factors with ones forged purely from data mining or stochastic genesis, it's evident that GPT's creations are underpinned by robust economic logic and are more semantically transparent. They transcend mere data dependency, thus potentiating them as instruments of potentially lucrative investment returns.

At a macroscopic lens, GPT's sophisticated factors gravitate towards capturing market volatilities, price trajectories, and momentum. A significant majority of these composite factors bear at least one foundational element which acutely addresses price gyrations or trade volume oscillations. GPT's meticulousness extends to recognizing reversal cues, frequently employing indicators such as the Commodity Channel Index (CCI), price reversal (PR), and the Parabolic Stop and Reverse (PSAR)—all hallmarks of reversal prognosticators. It often marries the Volume Acceleration (VA) and Directional Movement Index (DMI) to discern price trends, augmenting the predictive prowess of the factor. Out of the rich tapestry of 31 factors proffered by GPT, a staggering 81 unique foundational components were identified. However, factoring in the nuances of code which rendered certain factors inoperable, this number burgeons to 105. This profound diversity underscores that, while GPT amalgamates a plethora of base factors, it seldom retreads the same pathway. Such a kaleidoscopic array is a testament to GPT's expansive financial lexicon and its proclivity for novel factorial synthesis.

5.3 GPT's Economic Analysis of Representative Factors

To more profoundly delineate the distinctive attributes inherent to the factors synthesized by ChatGPT, we shall dissect selected factors that exemplify potent representational capacity and unique attributes. We prompted GPT to deconstruct these factors through an economic lens. As per the insights drawn from GPT's elucidation, it unmistakably manifests a holistic comprehension of financial axioms pertinent to the devised factors, accompanied by nuanced rationales, notwithstanding the potential contention surrounding such interpretations.

Factor 4, an amalgamated entity, is engendered through the convolution of the Rate of Change (ROC), Accumulation/Distribution Line (ADL), and Commodity Channel Index (CCI). Concretely, ROC stands as a momentum-centric technical metric quantifying the percentage alteration juxtaposing the prevailing price to the antecedent price of a delineated period. It computes the rate of change grounded on a decacyclic timeline of closing valuations. In the economic paradigm, the rate of change serves as an instrument to gauge the tempo and magnitude of price oscillations, unmasking ephemeral market trajectories. Brisk price perturbations might resonate with market supply-demand dynamism or emergent contingencies. An augmented rate of change might hint at tempestuous market sentiments, potentially triggered by the dissemination of pivotal intel or unexpected market perturbations.

The Accumulation/Distribution Line (ADL) functions as a volume-anchored metric with an intent to quantify the aggregated monetary ingress and egress concerning a security. Its mathematical representation for a stock on the nth day is articulated as:

$$\sum_{t=1}^{n} \frac{(close_t - low_t) - (high_t - close_t)}{high_t - low_t} * volume_t \tag{5}$$

This metric plays a pivotal role in discerning acquisition or divestment thrusts, thereby unveiling the vector of capital trajectories in the marketplace. Economically construed, the capital trajectory epitomizes the collective sentiment of market stakeholders apropos a security. A superlative ADL delineation insinuates capital influx, potentially echoing amplified acquisition momentum, whilst an inferior ADL signifies capital egress, potentially resonating with intensified divestment momentum. Such capital maneuvers encapsulate market stakeholders' prognostications and assurance.

The Commodity Channel Index (CCI) quantitatively scrutinizes the discrepancy between the extant price and its chronological median. Within economic discourses, the CCI acts as an apparatus to ascertain whether a valuation deviates from its protracted median, pinpointing potential overbought or oversold market states. Its mathematical formulation is:

$$TP = \frac{high + low + close}{3} \tag{6}$$

$$CCI = \frac{TP – mean(TP)}{0.015 \times std(TP)} \tag{7}$$

CCI measures the extent of the current price's departure from its historical midpoint. Divergences amid the price and its chronological median might be attributed to supply-demand stasis and market appraisals. An elevated CCI metric might insinuate a notable divergence from the protracted average, potentially garnering market scrutiny and potentially heralding a market recalibration.

When amalgamating these tripartite indicators, this composite metric offers perspicacious revelations regarding the velocity and amplitude of price metamorphoses, the trajectory and vigor of capital streams, and the divergence of valuations from their historical benchmarks. On a day-trading echelon, should a security exhibit an elevated value for this composite element, it might insinuate a robust positive rate of change, capital accretion, and an elevated CCI. This could signify an overbought market milieu. Retaining such equities under these circumstances might portend amplified risk quotients. In contrast, when the factor metric is minimized, signaling a pronounced negative rate of change, capital drainage, and a diminutive or adverse CCI, the market might be navigating an oversold stratum. Harnessing such equities during this phase might herald a market reversion, culminating in substantial returns. This delineation proffers a rationale for the pronounced returns concomitant with investments grounded in the inverse values of Factor 4.

Factor 13 is delineated as an amalgamation of price reversal, volume rate of change, and intraday price volatility metrics. The Price Reversal (PR) metric quantifies the disparity between the extant closing price and the zenith closing price over the preceding 7-day interval. A negative PR typically portends an imminent price reversal trajectory. The Volume Rate of Change (VRC), derived from a 5-day cyclical assessment, stands as a testament to the dynamism in trading volume, acting as a mirror to the market participants' trading predilections and liquidity fluxes. A pronounced VRC might intimate heightened market enthusiasm, potentially triggered by salient events or pivotal information, presaging potential shifts in price orientations. Intraday Price Volatility (IPV) calibrates the ratio between the intraday price amplitude (highest to lowest price differential) and the culminating price. A soaring IPV underscores pronounced intraday price gyrations, the genesis of which can be attributed to factors like market liquidity, trading momentum, or significant news influx. A surging IPV might encapsulate frenetic market sentiments or amplified trading fervor, opening up avenues for strategic trading maneuvers. Securities characterized by diminished factor values might exhibit a negative PR, a pronounced and bullish VRC, and heightened IPV, hinting at a potential inflection from a bearish to a bullish trend. Thus, contrarian investment strategies tethered to the inverse of this factor's values might culminate in lucrative dividends.

Factor 21 is synthesized by weaving together the Exponential Moving Average Ratio (EMAR), Standard Deviation of Volume (SDV), and Price Range Breakout (PRB). The EMAR is articulated as a quotient of the 5-day and the 20-day exponential moving averages, serving as a barometer to gauge the velocity and trajectory of short-term trends vis-a-vis their long-term counterparts. An augmented EMAR typically signifies a robust recent price trend, whereas its diminished counterpart leans towards a prevailing long-term trend. The SDV, encapsulating the volatility of trading volume over a 14-day continuum, offers insights into trading volume dynamism. Economically construed, volume oscillations are emblematic of market stakeholders' trading zeal and engagement. An escalated SDV might be emblematic of pronounced shifts in market sentiment. The PRB metric, illustrating the chasm between the closing price and the apogee price over a 10-day cycle, acts as an indicator of price range ruptures. Such breakouts could be precursors to nascent market trajectories or pivotal price oscillations. Investing contrarily to the factor's orientation might proffer substantial returns, especially when accompanied by a declining PRB.

Factor 24 amalgamates the Parabolic Stop and Reverse (PSAR), Williams %R, and the Triple Exponential Moving Average (TEMA) metrics. The PSAR furnishes potential market inflection points, its placement beneath ascending prices and above descending ones. Elevated PSAR values could either be symptomatic of a robust price trajectory or herald impending reversals. Williams %R, a momentum-centric metric, demarcates overbought and oversold terrains, with its zenith values potentially presaging price recalibrations. The TEMA, with its predisposition to accentuate recent data, showcases heightened reactivity to price perturbations. Augmented TEMA values might resonate with robust price momentum, heralding lucrative trading prospects. A stock characterized by a soaring combined factor value might be nestled in a bearish market with burgeoning reversal potential, thereby signposting opportunistic buy signals.

Factor 27, synthesized by integrating the Relative Strength Index (RSI), Commodity Channel Index (CCI), and Directional Movement Index (DMI), provides a nuanced lens into market momentum and trends. Both the RSI and DMI proffer insights into market momentum and its trajectories, equipping investors with a strategic vantage point to decipher market oscillations. By amalgamating these metrics with the CCI, which offers a meticulous appraisal of price deviations in conjunction with the territory of overbought or oversold conditions, a more refined prognostication about potential price aberrations from their historical mean and imminent reversals becomes attainable. Elevated composite factor values, typified by surging

RSI, CCI, and DMI indicators, often portend an overbought market scenario, teetering on the brink of a substantive decline. Such conditions render stocks ripe for strategic short-selling endeavors.

Factor 35 harmonizes the intricacies of the Aroon Oscillator, Chaikin A/D Oscillator, and the Rate of Change (ROC) to offer a holistic understanding of market dynamics. The Aroon Oscillator, a trend-discerning metric, quantifies the vigor and potential persistence of prevailing trends by juxtaposing the Aroon Up and Aroon Down indicators. Concurrently, the Chaikin A/D Oscillator gauges the momentum embedded within the Accumulation/Distribution Line (ADL), serving as a barometer for market participants' sentiment and convictions regarding a particular security. The ROC, firmly ensconced in the momentum indicator echelon, calibrates the percentage differential between the contemporary prices and their historical counterparts from a predetermined temporal horizon. This confluence of indicators facilitates a panoramic evaluation of price fluctuations against the backdrop of overarching market forces. In scenarios where the composite factor value is ascendant, market perceptions of the stock gravitate towards optimism, bolstered by formidable buying and selling impetuses. Such conditions often hint at a propitious milieu for initiating long positions on securities exemplifying elevated Factor 35 values.

In light of the foregoing analysis, it becomes unequivocally manifest that ChatGPT possesses an astute grasp of intricate financial forecasting paradigms and is adept at crafting factors rooted in this nuanced understanding. Remarkably, these delineated factors exhibit a marked departure from their counterparts, engendered through conventional data mining techniques.

6 Conclusion

from conventional data-mining-derived factors, highlighting the potency of GPT's knowledge-based reasoning in the financial domain. Instead of relying solely on pattern recognition from vast datasets, the approach championed by GPT-4 stems from its extensive foundational knowledge of financial theories, models, and empirical research, which it ingeniously applies to generate factors with salient predictive efficacy.

The methodology embodied by GPT in factor generation delineates a paradigm shift. Where traditional methods might get mired in statistical noise, GPT's knowledge-driven factors foreground key economic underpinnings and sophisticated financial reasoning. Such an approach not only amplifies the reliability of the derived factors but also deepens the trust that financial practitioners might place in them.

Moreover, the GPT-driven approach's inherently interpretative nature accentuates its academic and practical significance. Unlike many black-box methodologies that return factors of nebulous origin and logic, GPT-4's process remains transparent, aligning closely with well-established financial principles. This interpretability fosters both confidence and comprehension, allowing researchers and practitioners alike to discern the underlying logic and the potential implications of each factor.

As financial markets continue to evolve in complexity, the quest for sophisticated, reliable, and actionable tools to understand and navigate this landscape becomes ever more pressing. Our findings underscore the value and versatility of leveraging advanced language models like GPT-4 in this quest. By amalgamating the computational prowess of such models with foundational financial knowledge, we stand on the cusp of a novel era in financial analysis and decision-making.

It's our fervent hope that this research catalyzes further exploration into the synergy between artificial intelligence and finance, driving innovations that hold the promise of enhanced market efficiency, investor utility, and a deeper understanding of the intricate dance of global capital markets.

References

Araci, D. 2019. Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063* .

Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of finance* 52:57–82.

Cui, C., W. Wang, M. Zhang, G. Chen, Z. Luo, and B. C. Ooi. 2021. Alphaevolve: A learning framework to discover novel alphas in quantitative investment. In *Proceedings of the 2021 International Conference on Management of Data*, 2208–16.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *the Journal of Finance* 47:427–65.

———. 2015. A five-factor asset pricing model. *Journal of financial economics* 116:1–22.

Fang, J., J. Lin, S. Xia, Z. Xia, S. Hu, X. Liu, and Y. Jiang. 2020. Neural network-based automatic factor construction. *Quantitative Finance* 20:2101–14.

Fang, J., S. Xia, J. Lin, Z. Xia, X. Liu, and Y. Jiang. 2019. Alpha discovery neural network based on prior knowledge. *arXiv preprint arXiv:1912.11761* .

Jiang, J., B. T. Kelly, and D. Xiu. 2022. Expected returns and large language models. *Available at SSRN* .

Kanter, J. M., and K. Veeramachaneni. 2015. Deep feature synthesis: Towards automating data science endeavors. In *2015 IEEE international conference on data science and advanced analytics (DSAA)*, 1–10. IEEE.

Lakkaraju, K., S. K. R. Vuruma, V. Pallagani, B. Muppasani, and B. Srivastava. 2023. Can llms be good financial advisors?: An initial study in personal decision making for optimized outcomes. *arXiv preprint arXiv:2307.07422* .

Lam, H. T., J.-M. Thiebaut, M. Sinn, B. Chen, T. Mai, and O. Alkan. 2017. One button machine for automating feature engineering in relational databases. *arXiv preprint arXiv:1706.00327* .

Li, L., H. Wang, L. Zha, Q. Huang, S. Wu, G. Chen, and J. Zhao. 2022. Learning a data-driven policy network for pre-training automated feature engineering. In *The Eleventh International Conference on Learning Representations*.

Shah, A., and S. Chava. 2023. Zero is not hero yet: Benchmarking zero-shot performance of llms for financial tasks. *arXiv preprint arXiv:2305.16633* .

Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance* 19:425–42.

Son, G., H. Jung, M. Hahm, K. Na, and S. Jin. 2023. Beyond classification: Financial reasoning in state-of-the-art language models. *arXiv preprint arXiv:2305.01505* .

Wu, S., O. Irsoy, S. Lu, V. Dabravolski, M. Dredze, S. Gehrmann, P. Kambadur, D. Rosenberg, and G. Mann. 2023. Bloomberggpt: A large language model for finance.

Xie, Q., W. Han, X. Zhang, Y. Lai, M. Peng, A. Lopez-Lira, and J. Huang. 2023. Pixiu: A large language model, instruction data and evaluation benchmark for finance. *arXiv preprint arXiv:2306.05443* .

Yu, S., H. Xue, X. Ao, F. Pan, J. He, D. Tu, and Q. He. 2023a. Generating synergistic formulaic alpha collections via reinforcement learning. *arXiv preprint arXiv:2306.12964* .

Yu, X., Z. Chen, Y. Ling, S. Dong, Z. Liu, and Y. Lu. 2023b. Temporal data meets llm–explainable financial time series forecasting. *arXiv preprint arXiv:2306.11025* .

Zhang, C., Y. Li, X. Chen, Y. Jin, P. Tang, and J. Li. 2020a. Doubleensemble: A new ensemble method based on sample reweighting and feature selection for financial data analysis. In *2020 IEEE International Conference on Data Mining (ICDM)*, 781–90. IEEE.

Zhang, H., F. Hua, C. Xu, J. Guo, H. Kong, and R. Zuo. 2023a. Unveiling the potential of sentiment: Can large language models predict chinese stock price movements? *arXiv preprint arXiv:2306.14222* .

Zhang, T., Y. Li, Y. Jin, and J. Li. 2020b. Autoalpha: An efficient hierarchical evolutionary algorithm for mining alpha factors in quantitative investment. *arXiv preprint arXiv:2002.08245* .

Zhang, T., Z. A. Zhang, Z. Fan, H. Luo, F. Liu, Q. Liu, W. Cao, and L. Jian. 2023b. Openfe: Automated feature generation with expert-level performance. In *International Conference on Machine Learning*, 41880–901. PMLR.

Appendix

Table 1: **Performances of Factors-long short**

The table displays the returns of factors generated by GPT when implemented in long-short portfolios. A long-short portfolio means going long on stocks ranked in the top 50% based on their factor values and shorting those in the bottom 50%. Our factor return test period is from October 2021 to the end of 2022. We use the positive and negative performance of long-short total returns for each factor from 2000 to September 2021 as the reference. If a factor's return is negative from 2000 to September 2021, we take the opposite of the factor values for the following factor return testing; otherwise, we conduct the testing using the original factor values.

Factor	Annualized Return	Sharpe Ratio	Max Drawdown
factor1	0.078983	0.78658	0.067298
factor2	0.13694	1.3725	0.041848
factor3	0.16409	1.6346	0.060855
factor4	0.41127	3.2263	0.06438
factor5	0.091482	0.96611	0.059092
factor6	-0.15742	-1.8293	0.23157
factor7	0.066023	0.61637	0.067433
factor8	0.14235	1.5274	0.037539
factor9	0.04314	0.32961	0.075877
factor10	0.028006	0.17803	0.11063
factor11	0.22168	1.3942	0.085803
factor12	0.052999	0.39502	0.10849
factor13	0.25392	2.5068	0.044595
factor14	0.10636	0.77598	0.070162
factor15	-0.071784	-1.0284	0.16745
factor16	0.05146	0.46792	0.082357
factor17	0.16762	1.6377	0.047413
factor18	-0.12006	-1.4221	0.18121
factor19	0.19857	1.801	0.066468
factor20	0.16044	1.0433	0.053437
factor21	0.17613	1.0999	0.11092
factor22	0.17991	1.8579	0.040127
factor23	0.11402	0.67006	0.12132
factor24	0.55347	4.4913	0.087487
factor25	0.19581	1.9152	0.039251
factor26	-0.0036311	-0.16865	0.09649
factor27	0.35218	2.819	0.065002
factor28	0.6616	4.4491	0.13316
factor29	0.035456	0.19077	0.10019
factor30	0.4428	3.9695	0.034246
factor31	0.37547	2.8379	0.088788
factor32	0.10318	0.92899	0.067779
factor33	-0.071541	-0.54518	0.22742
factor34	0.036353	0.18361	0.139
factor35	0.43463	3.9651	0.055648

Table 2: **Performances of Factors-long only**

This table shows the returns of factors created by GPT in long-only portfolios. For this strategy, long-only involves choosing stocks from the top 50% based on their factor values. Our factor return testing period spans from October 2021 to the end of 2022. We use the positive or negative sign of the long-short total returns of each factor from January 2000 to September 2021 as a reference. If a factor has a negative returns during this period, we conduct the factor return testing by taking the opposite of the factor values. Conversely, if a factor has a positive return during this period, we conduct the testing using the original factor values.

Factor	Annualized Return	Sharpe Ratio	Max Drawdown
factor1	0.20479	0.89204	0.20155
factor2	0.23443	1.0099	0.20684
factor3	0.24808	1.0758	0.18583
factor4	0.44842	1.6328	0.26959
factor5	0.21132	0.91828	0.20943
factor6	0.087699	0.37461	0.23565
factor7	0.20013	0.85673	0.22782
factor8	0.24238	1.0312	0.2245
factor9	0.18744	0.80099	0.23814
factor10	0.17976	0.76965	0.25036
factor11	0.2772	1.0771	0.23132
factor12	0.19183	0.80766	0.26495
factor13	0.29238	1.278	0.18683
factor14	0.21933	0.89947	0.22586
factor15	0.12955	0.56631	0.22303
factor16	0.19352	0.83108	0.24896
factor17	0.24964	1.1161	0.24227
factor18	0.10531	0.41937	0.25588
factor19	0.2671	1.0763	0.23943
factor20	0.24549	1.4548	0.17993
factor21	0.25458	0.88686	0.26787
factor22	0.25659	1.1862	0.22937
factor23	0.22314	0.83267	0.26979
factor24	0.44324	1.8314	0.24462
factor25	0.26593	1.1539	0.19595
factor26	0.16408	0.73379	0.25523
factor27	0.34354	1.3997	0.22706
factor28	0.53899	1.9339	0.27278
factor29	0.18552	0.77443	0.25873
factor30	0.38943	1.8877	0.184
factor31	0.35378	1.3401	0.24271
factor32	0.21809	1.0455	0.23394
factor33	0.13028	0.4832	0.29923
factor34	0.18621	0.76978	0.23569
factor35	0.38338	1.5258	0.23638

Table 3: **Alphas of Factors**

This table depicts the alpha of each GPT-generated factor within long-short portfolios. We present the portfolio alphas for two distinct time periods. "Short term" refers to the alphas of long-short factor portfolios from October 2020 to the end of 2022, with signs determined by the positive or negative sign of each factor's return from 2000 to September 2021. "long term" represents the alphas of long-short factor portfolios from 2003 to the end of 2022, with signs determined by the positive or negative sign of each factor's returns from 2000 to 2022. Statistical significance is denoted as: *** for 1% significance, ** for 5% significance, and * for 10% significance.

Factor	short term	long term	Factor	short term	long term
factor1	0.081366 (1.05)	0.037995** (2.38)	factor19	0.22016** (2.50)	0.078177*** (4.37)
factor2	0.14555* (1.75)	0.048151*** (3.08)	factor20	0.064678 (0.86)	0.15804*** (9.43)
factor3	0.16033* (1.89)	0.074474*** (4.59)	factor21	0.24133*** (2.96)	-0.0021094 (-0.13)
factor4	0.41524*** (3.80)	0.11757*** (4.63)	factor22	0.13958* (1.77)	-0.0058011 (-0.38)
factor5	0.10443 (1.39)	0.055657*** (3.60)	factor23	0.13558 (1.08)	0.23501*** (9.00)
factor6	-0.15809* (-1.92)	0.0085058 (0.56)	factor24	0.61935*** (6.51)	0.35665*** (17.97)
factor7	0.052289 (0.65)	0.026176* (1.68)	factor25	0.19768** (2.26)	0.038568** (2.18)
factor8	0.13276* (1.74)	0.050085*** (3.27)	factor26	-0.050981 (-0.65)	0.025383 (1.52)
factor9	0.030934 (0.35)	0.068305*** (4.02)	factor27	0.36163*** (3.40)	0.20942*** (10.04)
factor10	0.019998 (0.24)	0.041312** (2.43)	factor28	0.74355*** (6.80)	0.39464*** (17.96)
factor11	0.22122* (1.71)	0.23327*** (8.84)	factor29	0.025872 (0.23)	0.22233*** (10.11)
factor12	0.05269 (0.56)	0.068138*** (3.47)	factor30	0.42559*** (4.83)	0.20825*** (11.87)
factor13	0.27217*** (3.13)	0.060105*** (3.50)	factor31	0.43864*** (4.52)	0.14078*** (7.56)
factor14	0.11712 (1.09)	0.20045*** (9.36)	factor32	0.037085 (0.47)	-0.012 (-0.75)
factor15	-0.070954 (-0.98)	-0.0065781 (-0.44)	factor33	-0.062677 (-0.48)	0.17787*** (6.97)
factor16	0.065698 (0.85)	-0.043872*** (-2.65)	factor34	-0.0051649 (-0.04)	0.11937*** (4.88)
factor17	0.16279* (1.89)	0.053537*** (3.22)	factor35	0.48099*** (5.62)	0.17422*** (11.15)
factor18	-0.11913 (-1.45)	0.019631 (1.31)			

Table 4: **Multi-Factor Portfolio Returns**
The table demonstrates the returns from static and dynamic multi-factor portfolios, both for long-only and long-short combinations, using factors generated by GPT. For the static method, we determine the positive or negative sign of each factor based on their returns from 2000 to September 2021, then calculate the average of these signs to create a new factor for testing. For the dynamic approach, the sign of the factor values is adjusted daily based on the cumulative positive or negative returns of either the long-only or long-short strategies associated with each factor.

Static Multi-Factor Portfolio Returns			
Strategy	Annualized return	Sharpe ratio	Max drawdown
long-only	0.45161	1.6295	0.25439
long-short	0.39728	2.4854	0.071795

Dynamic Multi-Factor Portfolio Returns - Referenced by long-only Returns			
Strategy	Return	Sharpe ratio	Max drawdown
long-only	0.34044	1.3822	0.22944
long-short	0.37593	3.0931	0.061103

Dynamic Multi-Factor Portfolio Returns - Referenced by Long-Short Returns			
Strategy	Return	Sharpe ratio	Max drawdown
long-only	0.30262	1.2644	0.29177
long-short	0.30036	2.4478	0.12983

Table 5: **Alphas of Multi-Factor Strategies**
The table presents the alpha of static and dynamic multi-factor portfolios using factors generated by GPT in a long-short combination. "long-only base" refers to using the factor's long-only returns as a positive or negative reference, and "long-short base" means referencing the factor's long-short returns. *** denotes 1% statistical significance, ** denotes 5% statistical significance, and * indicates 10% statistical significance.

Factor	3factor-dayls	5factor-dayls
static	0.41809***	0.40985***
	(3.01)	(2.97)
$\text{dynamic}_{\text{long-only base}}$	0.38067***	0.36583***
	(2.97)	(2.86)
$\text{dynamic}_{\text{longshort base}}$	0.32770**	0.32256**
	(2.40)	(2.38)

Table 6: **Multi-Factor Portfolio Returns - 10 Bins**

The table displays the annualized returns, Sharpe ratio, and maximum drawdowns derived from both static and dynamic multi-factor portfolios based on factors generated by GPT. All stocks are categorized into ten groups based on the computed mean factor values, with a long-only strategy applied to each group. Specifically, the static approach first adjust the signs of factors according to the their long-short returns from 2000 to September 2022, while the dynamic method adjusts the sign of the factor values daily based on the cumulative returns' positive or negative direction of each factor for either long-only or long-short positions.

Static Multi-Factor Portfolio Returns - 10 Bins			
Group	Annualized return	Sharpe ratio	Max drawdown
Low	0.90252	2.2864	0.30424
2	0.74673	2.0861	0.22777
3	0.20703	0.65923	0.28248
4	0.1914	0.58776	0.23074
5	0.21317	0.72800	0.26801
6	0.18109	0.53500	0.34179
7	-0.10458	-0.48012	0.31533
8	-0.09856	-0.46569	0.34026
9	0.27494	0.90091	0.21767
High	0.021541	0.042082	0.34141
H-L	0.88097	2.4622	0.1335

Dynamic Multi-Factor Portfolio Returns - Referenced by long-only Returns - 10 Bins			
Group	Annualized return	Sharpe ratio	Max drawdown
Low	0.094964	0.33951	0.39546
2	0.20934	0.68301	0.19487
3	-0.036588	-0.19603	0.3092
4	0.067688	0.19416	0.32577
5	0.011829	0.0011659	0.30439
6	0.34726	1.1524	0.21421
7	0.20533	0.63581	0.25718
8	0.61300	1.7134	0.22975
9	0.52126	1.5392	0.26295
High	0.50977	1.5949	0.27727
H-L	0.41481	1.4839	0.20772

Dynamic Multi-Factor Portfolio Returns - Referenced by Long-Short Returns - 10 Bins			
Group	Annualized return	Sharpe ratio	Max drawdown
Low	0.035556	0.090087	0.37699
2	0.23402	0.77437	0.17019
3	0.017148	0.021575	0.20962
4	-0.017581	-0.11558	0.23718
5	0.22608	0.6264	0.30014
6	0.22253	0.7258	0.27082
7	0.18375	0.56628	0.3168
8	0.34612	1.1166	0.30783
9	0.57093	1.6698	0.29316
High	0.73425	2.0391	0.29997
H-L	0.69870	2.0195	0.14967

Table 7: **Description of Constituent Factors Used in the Complex Factors Created by GPT.** The table showcases all 81 factors used by GPT along-only with their explanations. Some factors share the same name but differ in calculation methods. The number in parentheses after the name of some factors indicates the total number of times the factor has been used; those without numbers have been used only once.

Factor	Description
Volatility-Momentum Oscillator (VMO)	First calculate the 5-day momentum, which is the difference between the current closing price and the closing price 5 days ago. Then divide this by the 5-day standard deviation of the close prices to normalize for volatility.
Volume Delta to Volatility (VDV)	For each day, calculate the change in volume from the previous day (Volume Delta). We then divide this by the 5-day standard deviation of the volume to normalize for volatility in the volume.
Volatility and momentum interaction (VMI)	First, compute the 7-day and 14-day momentum of the closing prices. Then, divide the 7-day momentum by the 14-day standard deviation of the close prices, and the 14-day momentum by the 7-day standard deviation. Multiply these two resulting factors together.
Volume and time decay (VTD)	Compute the difference between today's volume and the volume from 3 days ago. Then, apply an exponential decay to this difference, where the decay factor decreases as we go back in time.
Relative Price Oscillator (RPO)	Calculate the 10-day and 20-day simple moving averages of the closing price. Then, subtract the 10-day from the 20-day and divide the result by the 20-day average to get the oscillator.
Volume Change Rate (VCR)	This is calculated as the rate of change of a 5-day simple moving average of the volume.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Discrete Fourier Transform (DFT)	Compute the magnitude of the DFT of the close price over a 30-day window.
Mean Reverting Scaled Oscillator (MRSO)	This factor takes into consideration the mean-reverting characteristic of financial data. Create an oscillator using the 7-day and 21-day moving averages and scale it using the standard deviation over the past 21 days.
Volume Impact (VI)	Calculate the volume change compared to the average volume over the past 10 days. Then, apply a sigmoid function to this ratio to keep the impact within a reasonable range.
Autocorrelation Measure (AM)(2)	This measure will capture the autocorrelation of the closing prices over the past 5 days.
Momentum and Intraday Range Interaction (MIRI)	Calculate the 5-day and 10-day momentum of the closing prices. Then, compute the intraday price range (high - low) and average it over the past 5 and 10 days. The MIRI is the 5-day momentum times the 5-day average range divided by (the 10-day momentum times the 10-day average range).
Volume and Temporal Decay (VTD)	Calculate the difference between today's volume and the volume from 3 days ago. Then, apply a linear decay factor to this difference, which reduces the weight of older volume data.
Weighted Close Price (WCP)	Calculate this as (close * volume) / volume. It gives a sense of how much the closing price is "supported" by the volume.
Volatility Impact (VI)	We calculate the standard deviation of the past 10 days' closing prices as a measure of volatility. Then, we apply a logarithm to this measure to limit its impact.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Price-Volume Trend (PVT)	measures the cumulative volume-adjusted price change. It is calculated as the cumulative sum of the product of the price change rate and volume change rate.
Normalized Intraday Volatility (NIV)	This is the standard deviation of the intraday returns (close-to-close) over the past 10 days, normalized by the 20-day moving average of the closing prices.
Trading Range Divergence (TRD)	This is the divergence between the current intraday range (high - low) and its 5-day average, normalized by the 5-day average.
Volume Weighted Change (VWC)	Multiply the percentage change in price with the percentage change in volume.
Range Rate of Change (RRC)	Calculate this as the percentage change in the daily range (high - low).
Intraday Volatility Scaling (IVS)	Calculate this as the daily range (high - low) divided by the standard deviation of the daily range over the past 14 days.
Volume Acceleration (VA)	Calculate the percentage change in volume compared to 3 days ago.
Historical Price Movement (HPM)	Calculate the percentage change in the current closing price compared to the closing price 10 days ago.
Momentum Score (MS)	Calculate the current close price minus the close price 5 days ago.
Volume Spike (VS)	Calculate the percentage change in volume over the past 3 days.
Trading Range Fluctuation (TRF)	Compute the ratio of the current day's trading range (high - low) to the average trading range over the past 10 days.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Trend Score (TS)	Calculate the difference between the close price and its 10-day moving average. This score helps us detect the recent price trend.
Volume Dynamics (VD)	Calculate the ratio of the current volume to the 5-day moving average volume. A significant increase or decrease in this ratio indicates possible changes in trading interest.
Volatility Index (VI)	Calculate the standard deviation of the daily returns (close-to-close percentage change) over the past 20 days. Higher values indicate greater volatility.
Price-Volume Trend (PVT)	This measures the cumulative volume-adjusted price change. It is calculated as the cumulative sum of the product of price change rate and volume change rate.
Normalized Intraday Volatility (NIV)	This is the standard deviation of the intraday returns (close-to-close) over the past 10 days, normalized by the 20-day moving average of the closing prices.
Trading Range Divergence (TRD)	This is the divergence between the current intraday range (high - low) and its 5-day average, normalized by the 5-day average.
Volume Weighted Change (VWC)	Multiply the percentage change in price with the percentage change in volume.
Range Rate of Change (RRC)	Calculate this as the percentage change in the daily range (high - low).
Intraday Volatility Scaling (IVS)	Calculate this as the daily range (high - low) divided by the standard deviation of the daily range over the past 14 days.
Volume Acceleration (VA)	Calculate the percentage change in volume compared to 3 days ago.
Historical Price Movement (HPM)	Calculate the percentage change in the current closing price compared to the closing price 10 days ago.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Momentum Score (MS)	Calculate the current close price minus the close price 5 days ago.
Volume Spike (VS)	Calculate the percentage change in volume over the past 3 days.
Trading Range Fluctuation (TRF)	Compute the ratio of the current day's trading range (high - low) to the average trading range over the past 10 days.
Trend Score (TS)	Calculate the difference between the close price and its 10-day moving average. This score helps us detect the recent price trend.
Volume Dynamics (VD)	Calculate the ratio of the current volume to the 5-day moving average volume. A significant increase or decrease in this ratio indicates possible changes in trading interest.
Volatility Index (VI)	Calculate the standard deviation of the daily returns (close-to-close percentage change) over the past 20 days. Higher values indicate greater volatility.
Price Acceleration (PA)	This is the difference between the current price change rate (current close price to previous close price) and the price change rate from 2 days ago (close price from 2 days ago to 3 days ago). A higher PA value might indicate an acceleration in price movement.
Trading Range Expansion (TRE)	This is the difference between the current trading range (high - low) and the trading range from 2 days ago. An increase in this value might signal an expansion in the trading range.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Price Reversal (PR)	Compute the difference between the current close price and the highest close price in the past 7 days. A negative value might indicate a potential price reversal.
Volume Rate of Change (VRC)(2)	Calculate the percentage change in volume over the past 5 days. This metric can signal changes in trading activity.
Intraday Price Volatility (IPV)	Compute the ratio of the intraday price range (high - low) to the closing price. A higher value can indicate higher intraday price volatility.
Twisted Price Momentum (TPM)	Calculate the difference between the current close price and the median close price over the past 10 days. This can help us identify unusual price changes.
Volume Variability (VV)	Compute the coefficient of variation (standard deviation divided by mean) of the volume over the past 5 days. A higher value might indicate significant changes in trading volume.
Intraday Price Efficiency (IPE)	Calculate the absolute value of the open price minus the close price divided by the day's high minus the low. A lower value indicates a more efficient intraday market.
Relative Price Change (RPC)	Compute the percentage change in closing price from the previous day divided by the percentage change in opening price from the previous day. This gives an idea about how the price is moving relative to the opening of the day.
Volume Momentum (VM)	This is the current volume divided by the maximum volume over the past 10 days. A higher value can indicate momentum in trading volume.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Daily Range Expansion (DRE)	Calculate the percentage change in the daily range (High - Low) from the previous day. This factor captures if the intraday price range is expanding or contracting.
Average True Range Rate of Change (ATRRC)	Compute the average true range (ATR) for the last 14 days and find the percentage change from the previous day. ATR is a measure of market volatility.
Standardized Closing Price (SCP)	This is the closing price standardized by the mean and standard deviation over the last 30 days. It shows how much the closing price deviates from the recent average.
Weighted Close Price (WCP)	This is the average of high, low, and double the close price for the day, giving more weight to the closing price.
Rate of Change in Volume (RCV)	This is the percentage change in volume compared to the previous day. This indicates a shift in trading activity.
Normalized Intraday Range (NIR)	This is the ratio of the difference between the day's high and low prices to the closing price, normalized by the maximum range observed over the past 10 days.
Cumulative Return (CR)	This is the running total of the daily return. It shows the overall profit or loss over a period of time.
Volume Change Momentum (VCM)	This is the difference between today's volume and the average volume over the past 5 days. This provides an indication of any sudden interest in the stock.
Daily Price Oscillation (DPO)	This is the difference between the highest and lowest prices divided by the closing price. This gives a sense of the relative volatility of the stock price during the day.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Volatility (VOL)	Compute the standard deviation of daily returns over the last 14 days. It provides a measure of risk associated with price changes.
Normalized Price Deviation (NPD)	This is the deviation of the closing price from its 30-day moving average, normalized by the 30-day standard deviation of the price. It shows how much the current price is deviating from the recent average.
Volume Adjusted Momentum (VAM)	This is the product of the 5-day momentum of the closing price and the 5-day average volume. It indicates the direction and intensity of the stock's movement.
Daily Range Percent (DRP)	This is the ratio of the difference between the day's high and low prices to the closing price. It indicates the day's volatility relative to the closing price.
Deviation from Simple Moving Average (DSMA)	This is the difference between the closing price and its 10-day simple moving average. It shows how much the current price deviates from the recent trend.
Exponential Moving Average Ratio (EMAR)	This is the ratio of the 5-day exponential moving average to the 20-day exponential moving average. It helps in identifying the short-term trend relative to the longer-term trend.
Standard Deviation of Volume (SDV)(2)	Compute the standard deviation of volume over the last 14 days. It provides a measure of volatility in trading volume.
Price Range Breakout (PRB)	This is the closing price minus the maximum high price of the last 10 days. It shows if the stock is breaking out of its recent range.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Relative Strength Index (RSI)(3)	This is a momentum oscillator that measures the speed and change of price movements. It ranges from 0 to 100 and is often used to identify overbought and oversold conditions.
On Balance Volume (OBV) Ratio	This is the ratio of the difference between today's OBV and the previous day's OBV to the total OBV. It measures cumulative buying and selling pressure by adding the volume on up days and subtracting the volume on down days.
Price Change Acceleration (PCA)	This is the second derivative of the price, which measures how quickly the price change is accelerating or decelerating.
Rate of Change (ROC)(5)	This is a momentum-based technical indicator that measures the percentage change in price between the current price and the price a certain number of periods ago.
Accumulation/Distribution Line (ADL)	This is a volume-based indicator designed to measure the cumulative flow of money into and out of a security. It can help to identify buying or selling pressure.
Commodity Channel Index (CCI)(3)	This is a technical indicator that measures the difference between the current price and the historical average price.
Bollinger Band Width (BBW)	This is a technical analysis indicator derived from the standard Bollinger Bands. BBW is calculated by subtracting the lower band from the upper band and then dividing the result by the middle band.
Volume-price Trend (VPT)	This is a cumulative indicator that integrates volume and price data to depict the direction of a trend.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Detrended Price Oscillator (DPO)	This is an indicator in technical analysis that seeks to eliminate the long-term trends in prices by using a displaced moving average so it does not react to the most current price action.
Parabolic Stop and Reverse (PSAR)(2)	This is a method proposed by J. Welles Wilder Jr. to find potential reversals in the market price direction.
Williams %R(2)	This is a momentum indicator that measures overbought/oversold levels. It moves between 0 and -100, and it is typically used to identify buying or selling signals.
Triple Exponential Moving Average (TEMA)	It is a technical indicator that applies an increased weight to recent data points, making it more reactive to price changes.
Average True Range (ATR)(2)	This is a volatility indicator introduced by Welles Wilder that measures the degree of price volatility.
Moving Average Convergence Divergence (MACD)(3)	This is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price.
On Balance Volume (OBV)(2)	It is a technical trading momentum indicator that uses volume flow to predict changes in stock price.
Stochastic Oscillator	This is a momentum indicator comparing a particular closing price of a security to a range of its prices over a certain period of time.
Directional Movement Index (DMI)	This is an indicator developed by J. Welles Wilder for identifying when a definable trend is present in an instrument. That is, the DMI tells whether an instrument is trending or not.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Chaikin Money Flow (CMF)	This is an oscillator that measures the amount of Money Flow Volume over a specific period. Money Flow Volume forms the basis for the Accumulation Distribution Line. Instead of a cumulative total of Money Flow Volume, Chaikin Money Flow simply sums Money Flow Volume for a specific look-back period, typically 20 or 21 days.
Ultimate Oscillator	This is a technical indicator invented by Larry Williams that uses the weighted average of three different time periods to reduce the volatility and false transaction signals that are associated with many other indicators that mainly rely on a single time period.
Average Directional Movement Index Rating (ADXR)	It measures the strength of the trend, irrespective of its direction. The trend strength is indicated by rising ADXR levels, while non-trending periods are characterized by falling ADXR.
Balance of Power (BoP)	It is a price-based indicator that estimates the balance of market power. Positive BoP values signal buyers' dominance, while negative BoP values indicate sellers are in control.
Aroon Oscillator(2)	This indicator is used to measure the strength of a trend by comparing the time it takes for the price to reach the highest and lowest points over a given time period.
Money Flow Index (MFI)	This is a momentum indicator that measures the flow of money into and out of a security over a specified period of time. It is related to the Relative Strength Index (RSI) but incorporates volume, whereas the RSI only considers price.

Continued on next page

Table 7 – Continued from previous page

Factor	Description
Chande Momentum Oscillator (CMO)	Developed by Tushar Chande, this is a technical momentum indicator similar to other momentum indicators such as the Relative Strength Index (RSI). However, the CMO aims to capture the inherent volatility of a security by oscillating between -100 and +100, with 0 as the baseline.
Bollinger Bands(2)	A technical analysis tool defined by a set of lines plotted two standard deviations (positively and negatively) away from a simple moving average (SMA) of the security's price.
Chaikin A/D Oscillator	It measures the momentum of the Accumulation Distribution Line using the MACD formula. This makes it an indicator of an indicator. The Chaikin A/D Oscillator fluctuates above/below the zero line.