Theme and Sentiment Analysis for Energy Commodity Investment Strategies Using Large Language Models: Evidence from ERNIE Bot

Yixuan Chen

Wuhan Jingkai Foreign Language School, Wuhan, China

56379195@qq.com

Keywords: Energy Commodities; Oil and Natural Gas; Large Language Models; ERNIE Bot; Topic Modeling; Sentiment Analysis; Investment Strategy; Geopolitics

Abstract: Energy commodities, particularly oil and natural gas, exert systemic influence across industrial costs, international balances of trade, financial stability, and geopolitics. This study analyzes publicly available 2022–2024 sell-side research from two leading international investment banks and employs the large language model ERNIE Bot to extract topics and classify sentiments, thereby quantifying the relative importance of themes and mapping risk–opportunity profiles. Results indicate that market intelligence and price dynamics in oil and gas dominate the discourse; the Russia–Ukraine war is associated with strongly negative sentiment; policy and regulatory topics show mixed interpretations; and methodological themes such as trend factors, style integration, and the use of historical derivatives information are viewed positively. While sustainability and the green transition are increasingly present, they remain secondary to fossil-fuel fundamentals and event-driven factors. The findings demonstrate that large language models can convert unstructured financial text into structured insights that support strategy design, risk management, and forward-looking investment judgment in energy commodity markets.

1. Introduction

1.1 Research Background and Significance

The strategic management of energy commodity exposure is essential due to its multi-channel effects on real activity, trade, and financial stability. Oil and natural gas prices influence upstream investment cycles and downstream sectoral cost structures, while their interactions with exchange rates, interest rates, and credit conditions create feedback loops that can magnify shocks. In practice, cost pass-through from fuel inputs affects margins in transport, chemicals, and manufacturing, while capex plans in exploration and production adjust to shifting breakevens and financing costs, altering medium-term supply elasticity. Terms-of-trade movements driven by energy prices reshape current accounts and fiscal balances, especially for exporter and importer blocs, feeding back into sovereign risk premia and currency dynamics. Geopolitical events and policy interventions act as exogenous drivers that reshape short- and medium-term expectations, propagating through futures curves, implied volatilities, and cross-asset correlations; this makes it imperative for investors to access timely, interpretable, and systematically organized information derived from complex textual sources, so they can calibrate positions, hedges, and capital allocation with greater precision amid heightened uncertainty.

1.2 Literature Review

Traditional research emphasizes oil and gas supply–demand fundamentals, OPEC+ production policies, inventories, and price elasticities, alongside the transmission of geopolitical events and macro variables to energy demand. Empirical strands examine inventory-based pricing, convenience yields, and the role of storage and logistics constraints in shaping the term structure, while event studies quantify the impact of sanctions, conflicts, and policy announcements on returns and volatility. Text mining and machine learning have increasingly been applied to analyst reports, news, and policy documents to measure sentiment, identify themes, and assess event impacts,

  DOI: 10.25236/gemmsd.2025.075

demonstrating predictive value for short-horizon returns and positioning shifts. Large language models extend these capabilities through cross-genre understanding, long-context summarization, and weakly supervised labeling, enabling structured extraction from unstructured narratives and improving recall across heterogeneous report styles. Nonetheless, a comprehensive, LLM-based quantification of themes and sentiments in sell-side energy research and its operational linkage to investment strategy—bridging textual indicators to portfolio construction, risk overlays, and timing rules—remains underdeveloped and calls for standardized workflows and validation.

1.3 Research Content and Methods

This study constructs a theme–importance–sentiment framework on a corpus of sell-side energy and commodities research from two international investment banks between 2022 and 2024. Using standardized prompts, ERNIE Bot identifies the top themes in each report, estimates their relative importance based on mention frequency and narrative emphasis, and classifies sentiment into positive, negative, neutral, and neutral/positive, producing a consistent, schema-like output per document. Aggregating across documents, we quantify global topic importance and sentiment distributions, interpret the joint landscape of risks and opportunities by mapping themes to semantic domains, and derive investment implications that distinguish cyclical catalysts from structural trends, with attention to methodological robustness through importance-weighting, rolling subsample checks, and light human validation to ensure market relevance.

2. Theoretical Foundations and Technical Approach

2.1 Systemic Properties of Energy Commodities

Oil and natural gas underpin energy systems and industrial production, with price shocks affecting upstream capital formation, supply elasticity, and downstream cost pass-through across transport, chemicals, and manufacturing [1]. The transmission is multi-channel: capex cycles adjust with expected breakevens and financing conditions; supply responses are constrained by reservoir decline rates, drilling lead times, and infrastructure bottlenecks; and demand exhibits heterogeneous elasticities across sectors and regions, amplifying or dampening price moves. Energy prices interact with macro-financial variables, including exchange rates, interest rates, and liquidity conditions, creating channels through which policy shifts and geopolitical events can propagate rapidly via trade balances, inflation expectations, and risk premia embedded in asset prices. Contract structures and market microstructure—term structures, storage economics, and inventory visibility—further shape volatility and the pace of information assimilation. Effective investment analysis must integrate fundamentals, regulatory context, and event risk across multiple time scales, balancing cyclical sensitivity with structural trends such as decarbonization, electrification, and technology-driven efficiency gains, while remaining alert to regime shifts that alter historical correlations and factor loadings.

2.2 Large Language Models for Financial Text Mining

Compared with traditional bag-of-words and topic models, large language models provide advantages in cross-domain comprehension, long-document summarization, semantic clustering, and low-friction annotation via prompt engineering [2]. Their contextual embeddings capture domain-specific nuances—distinguishing, for example, between inventory draws as bullish near-term signals versus structural supply tightness—and handle co-referential language and nested arguments common in analyst prose. By eliciting theme lists, approximate importance shares, concise summaries, and sentiment labels in a unified pass, LLMs generate structured outputs suited for quantitative synthesis, with retrieval augmentation and chain-of-thought prompting improving coverage and consistency when reports are lengthy or stylistically varied. This capability supports scalable surveillance of research corpora, enabling timely extraction of market-relevant signals that complement numerical data streams such as price curves, positioning, and macro releases, and it facilitates downstream tasks like clustering reports by thesis, tracking narrative drift over time, and

aligning textual insights with tradable hypotheses.

2.3 Workflow and Measurement Framework

The workflow consists of corpus assembly, standardized prompting, model output validation, and statistical aggregation. Reports are processed with a uniform prompt that requests the five most salient themes, their relative importance, brief summaries, and sentiment categorization, while instructing the model to consolidate synonymous constructs to ensure label stability across documents. After light-quality control to correct obvious ambiguities and harmonize theme nomenclature, themes are pooled across documents to compute global relative importance by frequency and to summarize sentiment distributions by theme, with importance-weighted statistics mitigating bias from unusually long or narrow reports. The interpretive layer maps themes onto semantic domains—traditional fundamentals, policy/regulation, geopolitics, methodology/strategy, and green transition—to form an integrated risk–opportunity view that is actionable for portfolio design, linking prominent negative-skew domains to hedging priorities and highlighting positive-skew, high-importance domains as candidates for directional exposure or optionality-based strategies; periodic re-estimation on rolling windows is used to monitor regime changes and adjust the mapping as market narratives evolve.

3. Data and Methods

3.1 Corpus and Tooling

The dataset comprises publicly available 2022–2024 energy and commodities research from Goldman Sachs (Commodities Insights) and J.P. Morgan (Commodities Research), curated through their public-facing portals and archived releases to ensure document authenticity and temporal coverage across macro cycles and event windows. Reports span weekly and monthly market updates, thematic deep-dives, policy analyses, and quantitative methodology notes, yielding a heterogeneous corpus in length, style, and technical density. Each document was converted to machine-readable text, with light preprocessing to remove boilerplate legal disclaimers, figure captions, and extraneous headers, while preserving analytical sections, executive summaries, and appendices that contain substantive commentary. ERNIE Bot, a large language model trained with supervised fine-tuning and reinforcement learning from human feedback, is employed to perform summarization, topic extraction, and sentiment classification; its knowledge-enhanced and retrieval-augmented capabilities help stabilize outputs across stylistic variation, while long-context handling reduces truncation risk for lengthy reports [3]. To promote consistency, the model was invoked with fixed temperature and decoding parameters to limit stochastic variance, and outputs were logged with metadata including report source, date, and inferred subsector (oil, gas, power, metals) to support transparent auditability and reproducibility.

3.2 Prompt Design and Output Structure

A standardized prompt requests ERNIE Bot to identify the five principal themes per report, estimate each theme's importance within the document, summarize its content, and assign a sentiment label among positive, negative, neutral, and neutral/positive, explicitly instructing the model to avoid duplicative phrasing and to treat closely related concepts under a unified theme label where appropriate. The prompt also specifies that importance should reflect intra-document salience proxied by mention frequency and narrative emphasis rather than absolute page count, and that sentiment classification should be grounded in evaluative language, directional statements, and risk–reward framing present in the analyst text. Returned outputs are reviewed for coherence, redundancy, and coverage, with minor normalization of theme titles—such as harmonizing "OPEC policy," "OPEC+ discipline," and "production cuts" to a single canonical label—to facilitate aggregation without altering original meaning; any ambiguous or compound themes are split or merged based on context to maintain one-to-one mapping between labels and underlying concepts [4]. The final structured record per report consists of a five-row schema capturing theme label,

importance score, one- to two-sentence summary, and sentiment class, supplemented by document-level identifiers to enable downstream grouping and cross-sectional comparisons.

3.3 Statistical Aggregation and Synthesis

Global topic importance is computed as the frequency share of each normalized theme across the entire corpus, with ties resolved by weighting themes by their within-document importance to avoid overcounting infrequent but highly emphasized topics; results are sensitivity-checked by excluding outlier reports with unusually narrow scope to ensure robustness [5]. Sentiment distributions are summarized by counting labels per theme and calculating the importance-weighted average sentiment intensity to reflect both prevalence and emphasis, while bootstrapped confidence intervals provide a sanity check against sampling noise from small theme cohorts. A joint view of theme prominence and sentiment intensity is constructed by mapping importance on one axis and the net sentiment balance on the other, highlighting clusters where market focus and directional conviction coalesce; this synthesis informs strategy mapping by indicating where positioning may be crowded or contrarian and guides risk prioritization by flagging themes with high prominence but skewed negative sentiment that could propagate into volatility. To validate stability, the analysis is repeated in rolling subperiods—early 2022 conflict onset, mid-2023 policy recalibration, and 2024 normalization—to detect regime shifts in topic weight and sentiment tone, ensuring that derived insights remain relevant across evolving market conditions.

4. Results

4.1 Topic Importance

Market intelligence and pricing dynamics in oil and natural gas dominate the research landscape, with themes related to supply–demand balances, inventory trajectories, and price behavior occupying the largest aggregate share. Geopolitical shocks, led by the Russia–Ukraine war, form the next tier of prominence, followed by OPEC/OPEC+ production policies and discipline, as well as macroeconomic and regulatory drivers influencing demand and cost of capital [6]. Methodological and strategy-oriented themes—including trend factors, style integration frameworks, and the economic value of long-horizon derivatives and term-structure information—feature meaningfully, reflecting a growing interest in systematic approaches. Sustainability, the energy transition, and green metals demand are present but carry lower relative weight compared with fossil-fuel fundamentals and event-driven topics.

4.2 Sentiment Patterns

Negative sentiment concentrates in geopolitical conflict and oil and gas risk narratives involving supply–demand imbalances, tighter financial conditions, and recession risk, indicating a cautious stance on near-term stability. Policy and regulatory topics display mixed sentiment, balancing potential opportunities from incentives and industrial policy against compliance costs and timing uncertainty, while OPEC+ decisions elicit nuanced interpretations contingent on discipline and market context. Positive sentiment clusters around methodology and strategy, where trend factors deliver strong returns with lower downside risk, style integration—particularly Bayesian Optimization Integration—improves Sharpe and robustness, and historical price information demonstrates tangible economic value; additional positivity attaches to LNG investment upcycles, industry consolidation, and selected diversification strategies [7]. Neutral/positive nuances emerge around medium- to long-term natural gas demand resilience and certain market–strategy alignments that suggest structural opportunities amid cyclical volatility.

4.3 Structured View of Risks and Opportunities

Key risks include the propagation of geopolitical shocks into energy and broader commodity complexes, the dual drag from tighter financial conditions on demand and funding costs, uncertainty around OPEC+ policy calibration, and vulnerabilities stemming from elongated supply chains and low inventories. Opportunities center on the multi-year LNG investment cycle,

systematic strategies leveraging trend factors and diversified style integration, structural growth in green metals demand supported by policy and electric vehicle and renewable deployment, and efficiency gains from consolidation that can improve cash flow profiles and capital discipline.

5. Discussion

5.1 Dominance of Fossil-Energy Fundamentals and Event-Driven Catalysts

The joint distribution of topic importance and sentiments underscores that current sell-side discourse prioritizes traditional oil and gas fundamentals—supply, demand, inventories, and pricing—alongside event-driven catalysts such as geopolitical conflicts and OPEC+ policy shifts, reflecting their high investability and rapid transmission into market prices over short to medium horizons. In practice, analysts repeatedly anchor forecasts on observable, high-frequency indicators like inventory draws, refinery utilization, rig counts, and crack spreads, which offer timely signals that map cleanly into valuation frameworks and trading heuristics. Geopolitical flashpoints and coordinated production announcements add discrete event risk that is readily tradeable through futures curves and options skew, reinforcing a research bias toward catalysts with near-term payoff profiles. This configuration explains the sentiment bifurcation observed in oil and gas intelligence—neutral to negative tones cluster around volatility and downside tail risks—while also highlighting a structural preference for themes where price discovery is swift, liquidity is deep, and macro transmission channels are well understood.

5.2 The Dual-Edged Role of Policy and Regulation

Policy and regulatory developments act as conditional amplifiers of risk and reward, where incentives and industrial strategies may unlock growth while compliance burdens and implementation uncertainty complicate timing and execution, implying that investors should align exposures with policy calendars, funding windows, and measurable milestones to capture asymmetry [8]. The mixed sentiment pattern reflects how subsidy design, trade measures, carbon pricing, and permitting regimes can simultaneously catalyze capex cycles and introduce execution risk via supply bottlenecks or shifting eligibility criteria. Effective positioning therefore hinges on scenario-based mapping of policy credibility and duration, translating legislative timelines into position sizing and optionality overlays. Practically, investors may favor structures that benefit from convexity—such as call spreads or calendar exposure to compliance deadlines—while maintaining stop-loss discipline around legal challenges, political turnover, or cap adjustments that can invert expected payoffs.

5.3 Transferability of Methodological and Strategy Insights

Evidence of strong performance from trend factors during the financialized era, robustness gains from Bayesian style integration under realistic frictions, and the economic value of long-horizon historical information supports incorporating systematic overlays into energy portfolios, provided that transaction costs, scoring conventions, and estimation windows are thoroughly stress-tested to ensure durability across regimes. The portability of these insights rests on careful feature engineering that respects market microstructure—for example, aligning signal horizons with contract liquidity, roll schedules, and seasonality in energy curves—and on ensemble methods that mitigate model error by combining partially orthogonal predictors. Robust implementation further requires parameter stability checks, out-of-sample validation across volatility regimes, and sensitivity to tail behavior, with risk controls that cap leverage during regime shifts and adjust execution tactics to slippage conditions. When integrated with discretionary views on geopolitics and policy, such systematic components can enhance Sharpe ratios and reduce drawdowns through disciplined diversification.

5.4 Relative Secondary Role but Medium-Term Promise of the Green Transition

Although sustainability and green-transition themes currently carry lower research weight than fossil-fuel fundamentals, policy reinforcement, supply-chain localization, and persistent demand for

green metals indicate medium- to long-term allocation value, likely unfolding at a steadier cadence than event-driven trades yet sustained by credible policy support and technology diffusion. The investment expression of this trajectory tilts toward multi-year capex cycles in LNG, grid modernization, and battery and renewable supply chains, where ramp-up dynamics and learning curves gradually compress costs and expand addressable demand [9]. Given the uneven pace of permitting, infrastructure buildout, and critical-mineral refining capacity, investors may emphasize staged exposure—scaling positions into milestone achievements and favoring balance sheets that can bridge policy gaps—while employing hedges against cyclical slowdowns in global manufacturing. Over time, as transparency in subsidy pathways improves and localized ecosystems mature, these exposures can transition from thematic bets to core allocations supported by stable cash flows and declining policy dependency.

6. Conclusion

Using ERNIE Bot to analyze sell-side energy and commodities research from 2022 to 2024, this study quantifies themes and sentiments to reveal a landscape centered on oil and gas market intelligence, price dynamics, and geopolitical events, complemented by mixed readings on policy and regulation and positive appraisals of systematic methodologies. Green transition topics, while rising, remain secondary to near-term drivers rooted in fossil-fuel markets and geopolitics. The results affirm that large language models can efficiently transform unstructured financial text into structured, decision-relevant intelligence for identifying trends, diagnosing sentiment, and informing energy investment strategies. Future work should expand standardized corpora across institutions and languages, refine model alignment for domain specificity, and strengthen causal identification to tighten the link between textual signals and tradable outcomes.

References

[1] Dahl C .Commodity Markets: Evolution, Challenges, and Policies[J].The energy journal, 2024(2):45.DOI:10.1177/01956574241244446.

[2] Xiao Y, Liang H, Qin P, Zhang Y, Lei W. SCOP: Evaluating the comprehension process of large language models from a cognitive view[C]//Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025). Vienna, Austria: Association for Computational Linguistics, 2025: 17407–17431. DOI:10.18653/v1/2025.acl-long.852.

[3] Leiyun H , Jinghan H , Qingjin C ,et al.The performance evaluation of artificial intelligence ERNIE bot in Chinese National Medical Licensing Examination[J].Postgraduate Medical Journal, 2024. DOI:10.1093/postmj/qgae062.

[4] Group O M I .Exuberant Supply Creates Surplus, Pressuring Opec Policy[J].Oil Market Intelligence, 2023, 28(11)

[5] Xi R. Statistical Aggregation: Theory and Applications[M]. Singapore: Springer, 2022.

[6] Ferchichi M, Talbi M, Ismaalia F, Samil S. Economic and geopolitical shocks and their influence on the Saudi stock market, Saudi Aramco, and Bitcoin: Evidence from ARDL and VAR models[J]. International Journal of Economics and Finance, 2025, 17(4): 31. DOI:10.5539/ijef.v17n4p31.

[7] He L, Hou Y , Zhao G ,et al.Parameters optimization of WOFOST model by integration of global sensitivity analysis and Bayesian calibration method[J].Transactions of the Chinese Society of Agricultural Engineering, 2016.DOI:10.11975/j.issn.1002-6819.2016.02.025.

[8] Siebert P C .Clean Air Act Regulatory and Policy Developments[C]//111th Air and Waste Management Association annual conference and exhibition: charting the future: environment, energy & health, vol. 2: 111th Air and Waste Management Association annual conference and exhibition (ACE 2018), 25-28 June 2018, Hartford, Connecticut, USA.2018.

[9] Yuan X T , Shao C Y , Zhang Z Z ,et al.Comparing the performance of ChatGPT and ERNIE Bot in answering questions regarding liver cancer interventional radiology in Chinese and English contexts: A comparative study[J].Digital Health, 2025.DOI:10.1177/20552076251315511.