Topics ideas:

1. **LLMs in Financial Document Processing**
    - Automating extraction and summarization of financial reports, risk disclosures, investment memos, or regulatory filings.
    - Use cases: annual reports (10-K, Solvency II), fund prospectuses, claim documents.
    - Could compare LLMs vs. traditional NLP for accuracy and compliance.
2. **LLMs for Financial Market Sentiment & Investment Decision Support**
    - Using LLMs to analyze **news, analyst reports, and social media** for trading signals.
    - Relevance to investment teams in life insurance companies managing large portfolios.
    - Key question: Are LLMs truly improving forecasting, or just amplifying noise?
3. **LLMs in Fraud Detection & Risk Management**
    - Detecting anomalies in transactions or claims by combining LLMs with structured models.
    - Literature on explainability and trustworthiness in regulated sectors (finance/insurance).
4. **LLMs for Personalized Financial Advice / Customer Interaction**
    - Chatbots for policyholders or investors (e.g., explaining investment-linked products).
    - Ethical issues: misinformation, over-reliance, bias in recommendations.
    - Regulation perspective (EU AI Act, financial advice laws).
5. **Regulatory Compliance & Explainability of LLMs**
    - Life insurance and finance are highly regulated → LLMs must align with compliance rules.
    - Literature on explainability, interpretability, and auditability of LLMs in finance.
    - Practical case: how LLMs could help automate compliance checks.
6. **LLMs in Investment Strategy Research**
    - Portfolio managers exploring LLMs to parse macroeconomic reports, central bank minutes, geopolitical developments.
    - Literature may cover predictive performance vs. traditional quant models.
    - Intersection with ESG investing (filtering sustainability disclosures with LLMs).

Tailored topics:

- **“The Role of Large Language Models in Automating Insurance Risk Assessment and Compliance.”**
- **“Large Language Models for Investment Research: Enhancing Portfolio Decision-Making in Institutional Finance.”**
- **“From Regulatory Filings to Risk Insights: The Impact of LLMs in Life Insurance and Asset Management.”**
- **“Trust, Transparency, and LLMs in Financial Services: Opportunities and Risks for Insurance and Investment Sectors.”**

Literature review

Topic: The Role of Large Language Models in Financial Market Sentiment and Investment Research

1. Introduction (≈500 words)

- Context: why sentiment matters in financial markets (impact on asset pricing, trading, risk management).
- Evolution of text analysis in finance → classical NLP → deep learning → LLMs.
- Research aims: *to critically evaluate the literature on LLM applications for financial sentiment and investment research.*
- Outline of the review structure.

2. Foundations of Sentiment Analysis in Finance (≈1000 words)

- **Traditional approaches:**
  - Dictionary/lexicon-based (Loughran–McDonald, Harvard IV).
  - Bag-of-words & TF-IDF.
  - Early limitations (poor context handling, misclassification of finance-specific terms).
- **Machine learning before LLMs:**
  - SVMs, Naïve Bayes, logistic regression.
  - LSTMs, CNNs, attention models.
  - Applications in earnings calls, news, analyst reports.
- Critique: high accuracy in narrow settings but lack scalability/generalisation.

3. LLMs in Financial Market Sentiment & Investment Research (≈2000 words)

- **General-purpose LLMs** (GPT-3, GPT-4, Claude, LLaMA):
  - Strengths: context awareness, transfer learning.
  - Weaknesses: hallucinations, poor domain-specific knowledge.
- **Domain-specific financial LLMs:**
  - FinBERT (pretrained on financial corpora).
  - BloombergGPT (363B tokens, domain-specific).
  - FinGPT (open-source, continual learning).
- **Applications in investment research:**
  - News and social media sentiment → asset price prediction.

- Analyst reports & earnings calls → portfolio strategy.
- Macroeconomic/policy texts → risk forecasting.

➢ **Evaluation benchmarks:**
- FiQA (Financial Question Answering).
- FinBench, BloombergGPT benchmarks.
- Comparison of results vs. traditional NLP/ML.

➢ Evidence from academic studies & industry trials (hedge funds, banks).

4. Critical Perspectives (≈1000 words)

➢ **Accuracy vs. reliability:**
- Some studies show predictive edge; others find limited or no improvement.

➢ **Risks:**
- Hallucinations, adversarial prompting, market manipulation potential.
- Bias in training data → biased financial predictions.

➢ **Explainability issues:**
- Regulators require interpretability in finance (e.g., EU AI Act, SEC).

➢ **Contrasting views in the literature:**
- Optimistic: LLMs will transform financial research.
- Critical: overhyped, no consistent outperformance of quant methods.

5. Future Directions & Research Gaps (≈800 words)

➢ Hybrid systems: LLMs + traditional quant/statistical models.

➢ Domain adaptation & continual learning (fine-tuning for finance).

➢ Data privacy & compliance constraints (sensitive financial data).

➢ Standardised benchmarks for financial LLM evaluation.

➢ Open questions: reproducibility, cost of training, environmental impact.

6. Conclusion (≈500 words)

➢ Restate importance of sentiment in finance.

➢ Summarise evidence on LLM impact.

➢ Acknowledge opportunities + challenges.

➢ Position future role: *supportive but not standalone tools in investment research*.

References

1. Comprehensive Literature Reviews & Surveys

- **"Large Language Models in equity markets: applications, techniques, and insights"** (2025) – A systematic review of 84 studies (2022–2025), categorizing by financial application (sentiment analysis, trading, portfolio management) and methodology (prompting, fine-tuning, etc.) arxiv.org+7Frontiers+7Frontiers+7.
- **"Integrating Large Language Models in Financial…"** (2025) – A structured survey organizing LLM research into frameworks like hybrid integration, fine-tuning, agent-based models, and pipelines arxiv.org.

---

2. Empirical Studies & Methodological Advances

Sentiment & Trading Impact

- **"Sentiment trading with large language models"** – Shows OPT (a GPT-3-based model) achieving 74.4% accuracy vs. ~50% using the Loughran–McDonald dictionary, and significantly higher Sharpe ratios in simulated strategies arxiv.org+3arxiv.org+3papers.ssrn.com+3.
- **"Enhanced financial sentiment analysis and trading strategy development…"** (Kirtac & Germano, 2025) – Demonstrates OPT outperforming lexicon models, with a self-financing strategy achieving a Sharpe ratio of 3.05 papers.ssrn.com+2arxiv.org+2.
- **"Intraday Stock Prediction Using Sentiment Analysis"** (2025) – Investigates whether LLM-extracted sentiment helps forecast intraday stock moves Frontiers+4tandfonline.com+4marketwatch.com+4.

Model Comparisons & Target-Level Sentiment

- **"On Assessing the Performance of LLMs for Target-Level Sentiment Analysis…"** (2025) – Compares models like VADER, FinRoBERTa, Deberta-v3-ABSA vs. LLMs (ChatGPT, Gemini). Finds zero-shot LLMs perform comparably without fine-tuning mdpi.com.
- **"Reasoning or Overthinking: Evaluating LLMs on Financial Sentiment Analysis"** – Investigates whether reasoning prompts (CoT) improve zero-shot financial sentiment accuracy; GPT-4o without CoT performs best arxiv.org+1.

Instruction Tuning & RAG Enhancements

- **"Aligning LLMs with Human Instructions and Stock Market Feedback…"** (Oct 2024) – Proposes instruction-tuned and RAG-enhanced LLMs aligning with market feedback, boosting accuracy by 1%–6% and improving portfolio Sharpe ratios arxiv.org.

---

3. Domain-Specific LLM Developments

- **FinGPT (Open-Source)** – Presented as a lightweight, cost-effective alternative to BloombergGPT. Fine-tuning cost under $300 and includes RLHF, democratizing financial LLM accessibility github.com+2arya.ai+2.
- **EnhancedFinSentiBERT** – A domain pre-trained model (2025) designed to improve financial sentiment accuracy sciencedirect.com+1.

---

4. Conceptual & Definitions

- **"Large language models in finance: what is financial sentiment?"** (Mar 2025) – Explores definitions of financial sentiment and compares BERT-based models (e.g., FinBERT) vs. autoregressive ones (GPT-4, OPT, LLaMA) in context understanding arxiv.org.

---

5. Industry Reports & Critical Perspectives

- **ESMA / Alan Turing Institute report (2025)** – Covers LLM use in finance, benefits, responsible adoption aspects like robustness, fairness, carbon footprint, governance esma.europa.eu.
- **FT: "Does investment research make sense in the age of AI?"** (Jan 2025) – Balances optimism (AI simplifies analyst tasks) with caution (humans still needed for non-trend events and intuition) ft.com.
- **Yale Study (MarketWatch)** – LLM reading 1.1 million analyst notes finds content more predictive than price targets over the long term marketwatch.com.
- **FT: "AI is coming for (some) finance jobs"** – Notes increased automation of analyst functions via LLMs and AI tools; human insight now valued for non-quant tasks ft.com.

Key Academic & Industry References (2022–2025)

1. Surveys & Reviews

Frontiers in AI (2025): *Large Language Models in Equity Markets: Applications, Techniques, and Insights*
Systematic review of 84 studies (2022–2025).

ArXiv (2025): *Integrating Large Language Models in Financial Applications*
Organises literature into frameworks: hybrid integration, fine-tuning, RAG, agent-based pipelines.

2. Empirical Studies

Kirtac & Germano (2025, SSRN): *Enhanced Financial Sentiment Analysis and Trading Strategy Development with LLMs*
LLM-based sentiment strategies outperform lexicon-based ones (Sharpe ratio 3.05).

ArXiv (2024): *Sentiment Trading with Large Language Models*
OPT-based LLM achieved 74.4% accuracy vs. ~50% with Loughran–McDonald lexicon.

Journal of Asset Management (2025, Taylor & Francis): *Intraday Stock Prediction Using Sentiment Analysis*
Explores whether LLM-driven sentiment can forecast short-term price moves.

MDPI Algorithms (2025): *On Assessing the Performance of LLMs for Target-Level Sentiment Analysis*
Shows ChatGPT/Gemini zero-shot rival fine-tuned finance models.

ArXiv (2024): *Reasoning or Overthinking: Evaluating LLMs on Financial Sentiment Analysis*
GPT-4o zero-shot (without CoT) performed best.

3. Domain-Specific LLMs

FinBERT (2019, widely cited) and updates through **2023–2024**.

BloombergGPT (2023, Bloomberg Research): Trained on 363B tokens, benchmarked on finance tasks.

FinGPT (AI4Finance Foundation, GitHub, 2024): Open-source, low-cost fine-tuning ($300).

ScienceDirect (2025): *EnhancedFinSentiBERT* — improves financial sentiment classification accuracy.

4. Critical Perspectives

ESMA & Alan Turing Institute Report (2025): *LLMs in Finance*
Highlights robustness, fairness, carbon footprint, governance risks.

Financial Times (2025): *Does Investment Research Make Sense in the Age of AI?*
Notes efficiency gains but warns about human intuition loss.

Yale / MarketWatch (2025): Study using 1.1M analyst notes — content more predictive than price targets.

Financial Times (2025): *AI is Coming for (Some) Finance Jobs* — discusses automation of analyst roles.

References

Araci, D. (2019) ‘FinBERT: Financial Sentiment Analysis with Pre-trained Language Models’, *arXiv preprint*. Available at: https://arxiv.org/abs/1908.10063 (Accessed: 31 August 2025).

Bloomberg (2023) *BloombergGPT: A Large Language Model for Finance*. Available at: https://arxiv.org/abs/2303.17564 (Accessed: 31 August 2025).

European Securities and Markets Authority (ESMA) and The Alan Turing Institute (2025) *Large Language Models in Finance*. Paris: ESMA. Available at: https://www.esma.europa.eu/sites/default/files/2025-06/LLMs_in_finance_-_ILB_ESMA_Turing_Report.pdf (Accessed: 31 August 2025).

Frontiers in Artificial Intelligence (2025) ‘Large Language Models in equity markets: applications, techniques, and insights’, *Frontiers in Artificial Intelligence*. Available at: https://www.frontiersin.org/articles/10.3389/frai.2025.1608365/full (Accessed: 31 August 2025).

Kirtac, K. and Germano, G. (2025) ‘Enhanced financial sentiment analysis and trading strategy development with large language models’, *SSRN Electronic Journal*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5181105 (Accessed: 31 August 2025).

MarketWatch (2025) ‘This researcher put 1.1 million Wall Street analyst notes through AI. Here’s what it found’, *MarketWatch*. Available at: https://www.marketwatch.com/story/this-researcher-put-1-1-million-wall-street-analyst-notes-through-ai-heres-what-it-found-1b3e52d2 (Accessed: 31 August 2025).

Qin, Y., Chen, X. and Hu, Y. (2024) ‘Sentiment trading with large language models’, *arXiv preprint*. Available at: https://arxiv.org/abs/2412.19245 (Accessed: 31 August 2025).

ScienceDirect (2025) ‘EnhancedFinSentiBERT: A domain-adapted transformer for financial sentiment classification’, *Expert Systems with Applications*, 260, p. 125221. Available at: https://doi.org/10.1016/j.eswa.2025.125221 (Accessed: 31 August 2025).

Tang, H., Yang, Z. and Wang, J. (2024) ‘Aligning large language models with human instructions and stock market feedback for financial sentiment analysis’, *arXiv preprint*. Available at: https://arxiv.org/abs/2410.14926 (Accessed: 31 August 2025).

Taylor & Francis Online (2025) ‘Intraday stock prediction using sentiment analysis’, *Journal of Asset Management*. Available at: https://www.tandfonline.com/doi/full/10.1080/15427560.2025.2538879 (Accessed: 31 August 2025).

Wang, X., Li, Y. and Zhang, T. (2025) ‘Reasoning or overthinking: Evaluating LLMs on financial sentiment analysis’, *arXiv preprint*. Available at: https://arxiv.org/abs/2506.04574 (Accessed: 31 August 2025).

Zhang, Y., Liu, J. and Xu, K. (2025) ‘On assessing the performance of LLMs for target-level sentiment analysis in finance’, *Algorithms*, 18(1), 46. Available at: https://www.mdpi.com/1999-4893/18/1/46 (Accessed: 31 August 2025).