Team members:

Sandesh Puligundla, Pramod Nayak, Pramod Bykhovsky, Bruno Graca Coelho

Press release: Improving Financial Analysis: AI-Powered solution for Core Earnings Estimation

Chicago, 03/13 – A team of Chicago Booth students has unveiled an innovative AI-driven approach to estimating core earnings — providing investors with a clearer picture of a company's true financial health.

This initiative leverages **Large Language Models (LLMs)** to extract and analyze financial data from complex 10-K reports, bridging the growing gap between GAAP net income and core earnings. As financial disclosures become increasingly complex, this AI-powered workflow integrates financial knowledge and machine reasoning to identify abnormal financial adjustments and enhance earnings analysis.

AI-Powered Financial Insights

The solution employs **state-of-the-art AI technology and techniques**, including:

- **Advanced vector search with Pinecone** for retrieving relevant financial data
- **Retrieval-Augmented Generation (RAG)** for precise financial data extraction
- **LLM-based augmentation** for interpreting and structuring data
- **Machine learning models (XGBoost, Random Forest)** for core earnings projections

Breakthrough Accuracy & Future Development

Rigorous testing using data from 18 companies demonstrated that **optimized AI prompting significantly improves accuracy**, though challenges remain in handling complex financial adjustments. Complexity matters, as with complex financial events the solution misses or miscategorizes them.

As expected, GPT-4o outperformed 4o-Mini. Further enhancements are underway, including testing with Gemini LLM and developing a fully autonomous AI-powered financial agent.

"This breakthrough AI solution has the potential to transform financial analysis by providing investors with an additional tool to support them in calculating reliable earnings estimates," said the development team. *"While challenges remain, our approach is another contribution to financial analysis with AI supporting financial decision-making faster and at scale."*

For more details on the **Core Earnings Explorers** initiative, contact Sandesh Puligundla or access our UChicago Booth – 30135 class repository.

Frequently Asked Questions (FAQ)

1. What problem are you actually trying to solve?

Estimation of core earnings, i.e., a firm's persistent profitability from its core business activities is central to investors' assessments of economic performance and valuations. Quantifying core earnings requires judgment and integration of information scattered throughout financial disclosures contextualized with general industry knowledge.
This has become increasingly difficult as financial disclosures have become more "bloated" and accounting standards have increased non-recurring impacts on GAAP net income. The chasm between GAAP earnings and what investors consider "core" earnings has widened, and bridging it has become more challenging.
Our solution aims to standardize, facilitate and improve the financial analysis and reporting on this key financial metric.

2. How do you test/ evaluate your solution?

a) We used multiple approaches to testing the data - leveraging both LLM and manual evaluation process.
b) The first phase was to upload the data into LLM, instruct it to compare output and evaluate quality
c) That phase would tend to fall short, thinking that most, if not all, out put is fine
d) The next phase would be to leverage LLM to support manual review.
e) For example: flagging potential flaws in the original analysis, and guiding the LLM to evaluate it, or manually go to the source file to review.
f) One of the conclusions of the exercise is data control at scale, particularly for a commercial grade output to be used in investment decisions, is a major challenge

3. Can you walk us through the technical implementation?

The technical implementation follows a structured Retrieval-Augmented Generation (RAG) workflow that extracts and analyzes financial disclosures:

a) Data Processing & Embedding
   - 10-K reports are split into 1024-token chunks with a 100-token overlap.
   - Each chunk is vectorized using OpenAI's text-embedding-3-small model (1536-dimension).
   - Vectorized data is stored in Pinecone for efficient similarity-based retrieval.

b) Financial Data Retrieval
   - Pinecone's vector search retrieves the top relevant text chunks related to financial queries.
   - Example: Retrieving effective tax rate information from multiple 10-K sections.

c) Extraction of Key Financial Metrics
   - The model extracts net income, tax rate, shares outstanding, and other financial data.
   - These values are structured into a FinancialMetrics model for processing.

d) Abnormal Expense & Revenue Detection
   - The system identifies non-recurring financial events (e.g., restructuring costs, legal settlements, asset impairments).

e) Core Earnings Calculation
   - Adjustments are made to remove abnormal financial events from net income.
   - Core earnings formula:

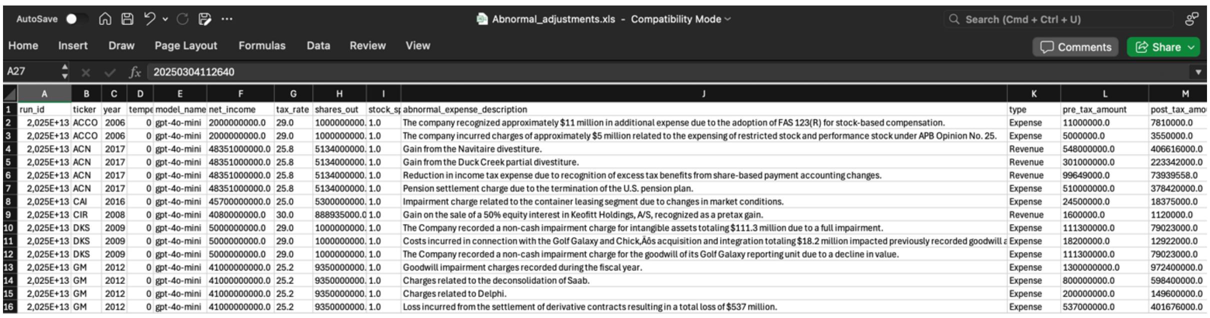
$$\text{Core Earnings} = \text{Net Income} - \sum(\text{Abnormal Expenses Adjusted for Tax})$$

f) Machine Learning for Earnings Forecasting
   - XGBoost and Random Forest models predict next-year core earnings per share (EPS).
   - These predictions provide forward-looking financial insights.

This implementation ensures accurate financial reporting and automated earnings analysis at scale.

4. Can I see an example of the solution's output?

Our solution, not only calculates core earnings by identifying abnormal adjustments (vide Figure 1), it also takes into account number of shares issued to calculate core earnings per share.



run_id	ticker	year	tempe	model_name	net_income	tax_rate	shares_out	stock_sp	abnormal_expense_description	type	pre_tax_amount	post_tax_amo
2,025E+13	ACCO	2006	0	gpt-4o-mini	2000000000.0	29.0	1000000000.	1.0	The company recognized approximately $11 million in additional expense due to the adoption of FAS 123(R) for stock-based compensation.	Expense	11000000.0	7810000.0
2,025E+13	ACCO	2006	0	gpt-4o-mini	2000000000.0	29.0	1000000000.	1.0	The company incurred charges of approximately $5 million related to the expensing of restricted stock and performance stock under APB Opinion No. 25.	Expense	5000000.0	3550000.0
2,025E+13	ACN	2017	0	gpt-4o-mini	48351000000.0	25.8	5134000000.	1.0	Gain from the Navitaire divestiture.	Revenue	548000000.0	406616000.0
2,025E+13	ACN	2017	0	gpt-4o-mini	48351000000.0	25.8	5134000000.	1.0	Gain from the Duck Creek partial divestiture.	Revenue	301000000.0	223342000.0
2,025E+13	ACN	2017	0	gpt-4o-mini	48351000000.0	25.8	5134000000.	1.0	Reduction in income tax expense due to recognition of excess tax benefits from share-based payment accounting changes.	Revenue	99649000.0	73939558.0
2,025E+13	ACN	2017	0	gpt-4o-mini	48351000000.0	25.8	5134000000.	1.0	Pension settlement charge due to the termination of the U.S. pension plan.	Expense	510000000.0	378420000.0
2,025E+13	CAI	2016	0	gpt-4o-mini	45700000000.0	25.0	5300000000.	1.0	Impairment charge related to the container leasing segment due to changes in market conditions.	Expense	24500000.0	18375000.0
2,025E+13	CIR	2008	0	gpt-4o-mini	4080000000.0	30.0	888935000.0	1.0	Gain on the sale of a 50% equity interest in Keofitt Holdings, A/S, recognized as a pretax gain.	Revenue	1600000.0	1120000.0
2,025E+13	DKS	2009	0	gpt-4o-mini	5000000000.0	29.0	1000000000.	1.0	The Company recorded a non-cash impairment charge for intangible assets totaling $111.3 million due to a full impairment.	Expense	111300000.0	79023000.0
2,025E+13	DKS	2009	0	gpt-4o-mini	5000000000.0	29.0	1000000000.	1.0	Costs incurred in connection with the Golf Galaxy and Chick‚Äôs acquisition and integration totaling $18.2 million impacted previously recorded goodwill a	Expense	18200000.0	12922000.0
2,025E+13	DKS	2009	0	gpt-4o-mini	5000000000.0	29.0	1000000000.	1.0	The Company recorded a non-cash impairment charge for the goodwill of its Golf Galaxy reporting unit due to a decline in value.	Expense	111300000.0	79023000.0
2,025E+13	GM	2012	0	gpt-4o-mini	41000000000.0	25.2	9350000000.	1.0	Goodwill impairment charges recorded during the fiscal year.	Expense	1300000000.0	972400000.0
2,025E+13	GM	2012	0	gpt-4o-mini	41000000000.0	25.2	9350000000.	1.0	Charges related to the deconsolidation of Saab.	Expense	800000000.0	598400000.0
2,025E+13	GM	2012	0	gpt-4o-mini	41000000000.0	25.2	9350000000.	1.0	Charges related to Delphi.	Expense	200000000.0	149600000.0
2,025E+13	GM	2012	0	gpt-4o-mini	41000000000.0	25.2	9350000000.	1.0	Loss incurred from the settlement of derivative contracts resulting in a total loss of $537 million.	Expense	537000000.0	401676000.0

Figure 1: Abnormal adjustments output list.

Finally, the solution uses machine learning forecast models to predict next year core earnings (vide Figure 2).

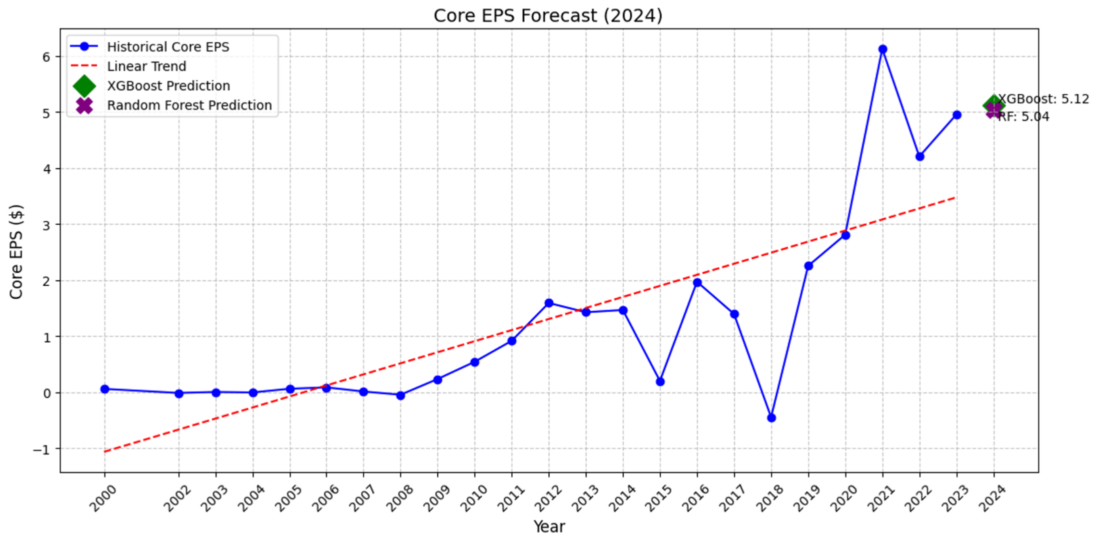


Figure 2: Calculated Core EPS and 2024 forecast.

5. Do you have any validation tests to show that your solution accurately calculates core earnings?
   a) Our focus was on flagging the pre-tax adjustments, and as our analysis showed the issues with flagging the right adjustments
   b) Multiple layers of LLM lead to pretty good accuracy around the numbers
   c) We know that the mistakes above prevent the LLM from accurately calculating the number
   d) Finally, we took a simplified approach to effectuating the post-tax adjustments - commercial grade solution would require materially more sophisticated system, requiring significant manual research to underpin it
6. Do you use any models provided by third-parties such as OpenAI?

   Yes. We used LLM models from OpenAI, ChatGPT 4o and ChatGPT 4o-Mini
7. What were the major issues observed in the current solution?
   a) LLM doesn't have sufficient ability to reason net income adjustments (e.g. one of the most common errors was including OCI figures)
   b) At times, the LLM would identify adjustments, but fail to account that these need to be capitalized, so the actual impact, should it be included (not always), would be much smaller
   c) Supervisory function doesn't lead to the desired outcome
   d) Quality control at scale is very challenging
8. Can the current technical solution be improved?

   Yes, in the future we would like to improve our AI-driven workflow and LLM-based financial analysis by :

   a) Improve Retrieval-Augmented Generation (RAG) Performance

    1. Optimize document chunking strategies (e.g., dynamic chunking) instead of fixed-size chunks to preserve financial context.e.g. Chunk by XBRL sections - Items or individual items.
    2. Enhance retrieval ranking algorithms by incorporating hybrid search (vector + keyword matching).

b) Enhance LLM’s Financial Context Understanding
    1. Fine-tune financial-specific prompts to ensure accurate identification of abnormal earnings adjustments.
    2. Implement LLM chain-of-thought reasoning to better assess whether financial adjustments impact core earnings.

c) Reduce Hallucination and Improve Fact Verification
    1. Introduce multi-step verification where the LLM cross-checks extracted data with multiple retrieved sources.
    2. Implement a supervisory LLM layer that re-evaluates financial adjustments for accuracy.

d) Optimize LLM-Driven Abnormal Expense Classification
    1. Use LLM-powered entity recognition (NER) to extract abnormal financial events with higher precision.
    2. Implement a confidence-scoring system where LLM outputs are validated based on structured financial datasets.

e) Improve AI Workflow for Core Earnings Forecasting
    1. Replace static prompt-based predictions with a multi-model ensemble using LLMs and traditional time-series forecasting models.
    2. Utilize LLM-generated feature engineering to improve the accuracy of XGBoost and LSTM-based EPS projections.