SentimentPulse: Real-Time Consumer Sentiment Analysis with Custom Language Models

Lixiang Li, Nagender Aneja, Alina Nessen, Bharat Bhargava

PURDUE UNIVERSITY®

Background

- **Economic Significance of Consumer Sentiment:**
  - Reflects public perception of the economy’s health.
  - Influences market trends and informs policy decisions.
- **Limitations of Traditional Surveys:**
  - Resource-intensive (time and cost).
  - Collected infrequently, missing real-time dynamics.
- **Need for a Dynamic Approach:**
  - Real-time, cost-effective sentiment analysis is essential.
  - Supports and complements traditional survey methods.

Background

- **Leveraging Language Models with Continual Learning:**
  - Uses timestamped data (e.g., news and S&P 500) for dynamic sentiment tracking. Captures fluctuations in sentiment over time efficiently.
- **Overcoming Limitations of Foundation Models:**
  - Foundation models are trained on unspecific internet corpora without time stamps.
  - Our model is specifically tailored to handle time-sensitive economic data.
- **Innovative Contribution:**
  - First application of a language model focused on economic consumer sentiment analysis.
  - Designed to work without depending on large-scale foundation models.

Relevant Work

Consumer Sentiment:

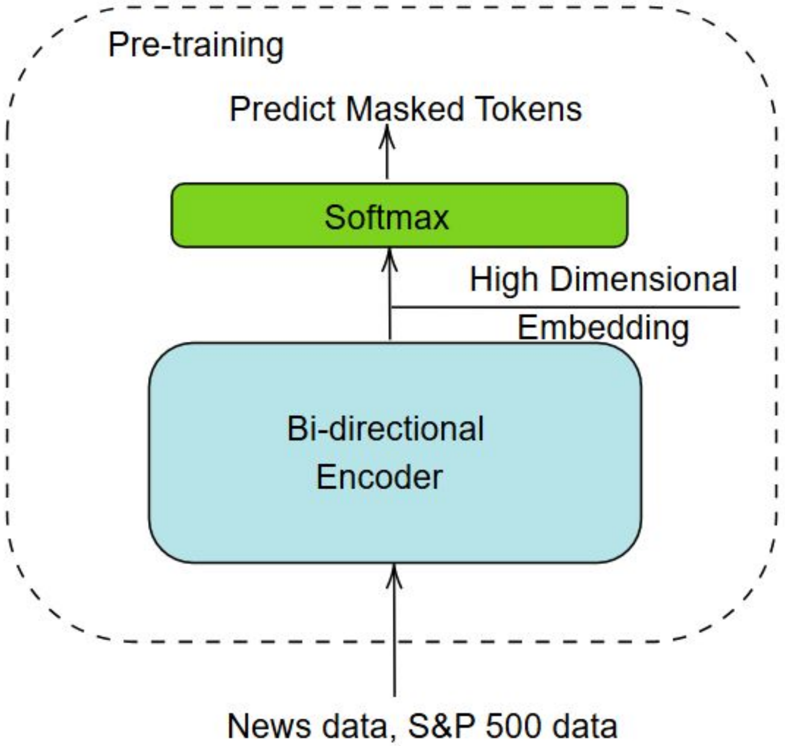
- **Kaur & Sharma (2023):**
  - LSTM-based model with hybrid feature extraction.
  - Transforms pre-processed reviews into feature vectors for sentiment analysis and summarization.
- **Han et al. (2023):**
  - Predicts consumer confidence index using machine learning.
  - Leverages web search keywords and Chinese consumer confidence data.

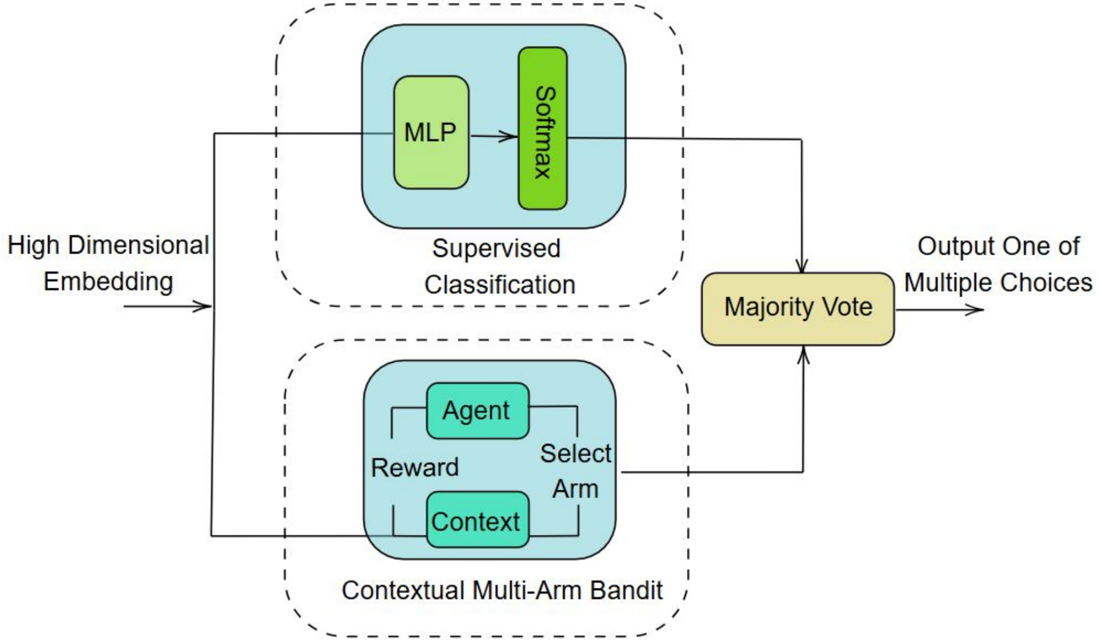
Relevant Work

Multiple Choice QA Methodology:

- **Transformer & CNN Approaches:**
  - **Huang et al. (2022):** Transformer encoder-decoder generates clue text for MCQA.
  - **Chaturvedi et al. (2018):** CNN captures embeddings, with attention layers scoring answer options.
- **Two-Stage & Hybrid Models:**
  - **Jin et al. (2019):** Combines coarse tuning (via NLI) with multi-task fine-tuning.
  - **Chen et al. (2019):** Uses Bi-LSTM and convolutional spatial attention for enriched representations.
- **Retriever-Reader Framework & LLM Studies:**
  - **Huang et al. (2021):** Employs a retriever (with novel word weighting) and reader fusion for scenario-based QA.
  - **Robinson et al. (2022):** Demonstrates LLMs' competitiveness in MCQA across 20 datasets with answer order-insensitive prompting.

Proposed Model Framework





Data

Pre-training Dataset:

- **Sources:**
  - New York Times News API (e.g., Politics, Economy, Business Day, etc.)
  - Guardian News API (e.g., Money, Politics, Business, Society under USA-News)
  - S&P 500 data
- **Key Features:**
  - Filtered by economic categories and divided by timestamp (e.g., January 2014 snippet)
  - Custom pre-training ensures inclusion of timestamp details missing in existing encoders

Fine-tuning Dataset – UMCSI Survey:

- **Source:**

  University of Michigan Consumer Sentiment Index (since 1978, monthly reports)

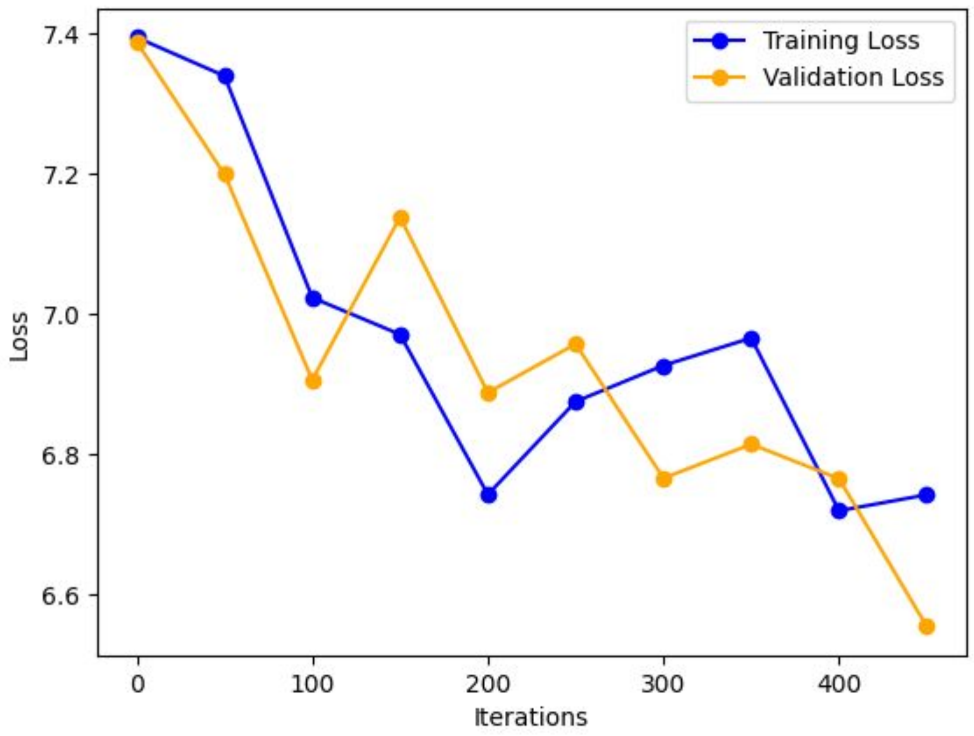
- **Components:**
  - 5 survey questions addressing personal finances, business conditions, and buying power
  - Demographic details (income, residence, political affiliation, education, household composition)

Continual Pretraining + SFT

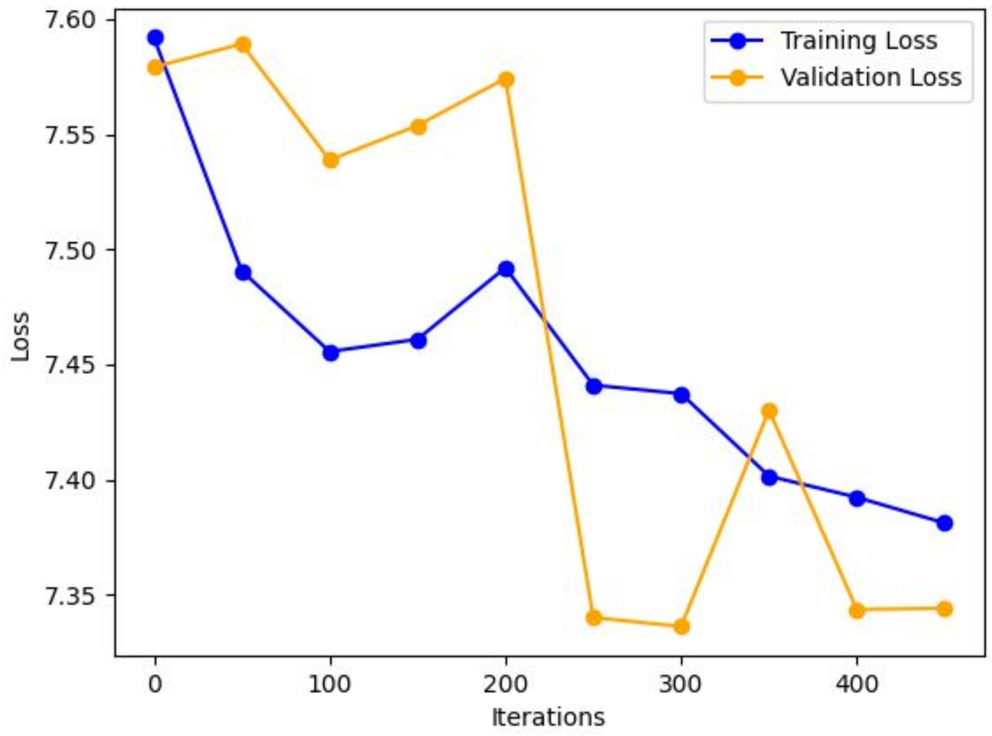
Algorithm 1 Continual Learning on News corpus and S&P 500, and fine-tuning on Survey Data

```
for data in (2014 – 2015, 2015 – 2016, 2016 – 2017, 2017 – 2018, 2018 – 2019) do
    encoder = pre-train(encoder, data)
    model1 = MLP(encoder, classifier)
    model2 = ContextualBandit(encoder)
    for each surveyQuestion do
        Context = GenerateContext(encoder, surveyData)
        for each in (Supervisedclassification, UCB, EG, AG) do
            Supervised_classifier(model1, Context)
            UCB(model2, Context)
            EG(model2, Context)
            AG(model2, Context)
        end for
    end for
end for
```

Pretraining Loss



Number of Parameters: 732 million; Attention block dimension: 160; Max input token allowed: 150; Batch size: 16



Number of Parameters: 369 million; Attention block dimension: 80; Max input token allowed: 150; Batch size: 16

Pretraining loss vs. number of iterations between the training set and validation set with two different parameter settings of the encoder

PURDUE UNIVERSITY

Different Finetuning Strategies

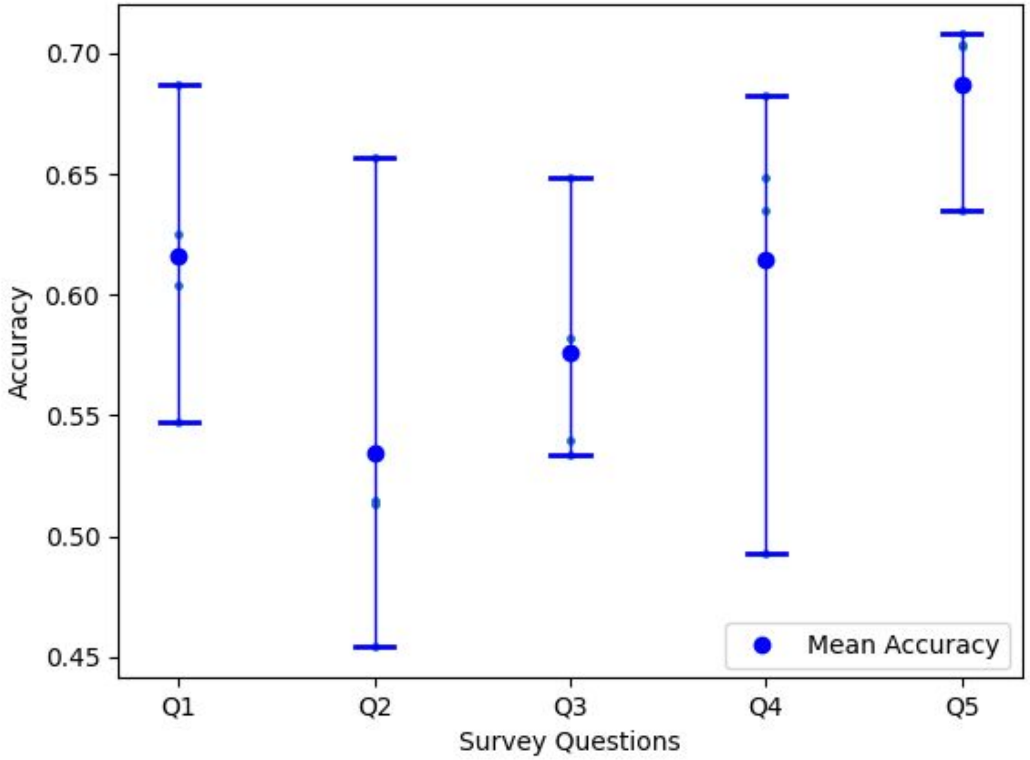
4 Finetuning Strategies:

- Supervised Classification (SC)
- Upper Confidence Bound (UCB)
- Epsilon Greedy (EG)
- Adaptive Greedy (AG)

Fine Tuning Methods	1 year	2 years	3 years	4 years	5 years
SC(Q1)	0.4458	0.5432	0.5543	0.6082	0.6875
SC(Q2)	0.5435	0.5242	0.5239	0.6143	0.6574
SC(Q3)	0.5389	0.5525	0.5356	0.5579	0.6485
SC(Q4)	0.5053	0.5342	0.5425	0.5932	0.6485
SC(Q5)	0.4564	0.5456	0.5982	0.6352	0.7034
UCB(Q1)	0.3821	0.4348	0.4854	0.5822	0.6252
UCB(Q2)	0.3245	0.3934	0.4354	0.5150	0.5152
UCB(Q3)	0.4023	0.4381	0.5208	0.5423	0.5396
UCB(Q4)	0.3831	0.4287	0.4929	0.5823	0.6349
UCB(Q5)	0.4564	0.5034	0.5723	0.6583	0.7083
EG(Q1)	0.3356	0.4345	0.4967	0.5242	0.5475
EG(Q2)	0.3113	0.392	0.4203	0.4345	0.4543
EG(Q3)	0.3564	0.3953	0.4422	0.4453	0.5334
EG(Q4)	0.4243	0.4035	0.4534	0.4563	0.4930
EG(Q5)	0.4564	0.5034	0.4835	0.5732	0.6359
AG(Q1)	0.3345	0.3852	0.4425	0.5435	0.6045
AG(Q2)	0.3054	0.3367	0.4035	0.4564	0.5135
AG(Q3)	0.3356	0.4253	0.4593	0.5103	0.5823
AG(Q4)	0.4501	0.4462	0.5024	0.6325	0.6823
AG(Q5)	0.4691	0.5409	0.5923	0.6832	0.7035
Average(Q1)	0.3745	0.4494	0.4947	0.5645	**0.6162**
Average(Q2)	0.3711	0.4116	0.4458	0.5051	**0.5351**
Average(Q3)	0.4083	0.4528	0.4894	0.5139	**0.5759**
Average(Q4)	0.4407	0.4531	0.4978	0.5661	**0.6146**
Average(Q5)	0.4596	0.5233	0.5616	0.6375	**0.6878**

Test Accuracy Using Different Training Strategies

Accuracy Variance of Finetuning Methods



Accuracy variance of four different fine-tuning methods across five survey questions

GPTs Answers Accuracy on Five Survey Questions

	Q1(PAGO)	Q2(PEXP)	Q3(BUS12)	Q4(BUS5)	Q5(DUR)
GPT-3.5-Turbo	0.2218	0.3687	0.2268	0.1843	0.3724
GPT-4	0.2710	0.5143	0.0691	0.1625	0.2778
SentimentPulse	0.6162	0.5351	0.5759	0.6146	0.6878

Comparison with GPT3.5 and GPT4

Summary

Model & Data Scale:

- Custom model with 732M parameters—small compared to modern GPTs
- Trained on a relatively small news dataset and S&P500 data

Contribution:

- A custom, temporally-trained language model outperforms GPT-3.5-Turbo and GPT-4 in real-time consumer sentiment prediction
- Demonstrates that cost-effective, small-scale models can achieve fine-grained performance improvements via continual learning
- Establishes a strong baseline for future research in economic sentiment analysis

Reference

- Chen, J., Chi, L., Peng, B., & Yuan, Z. (2024). HLLM: Enhancing Sequential Recommendations via Hierarchical Large Language Models for Item and User Modeling. arXiv:2409.12740v1 [cs.IR]
- Gagandeep Kaur and Amit Sharma. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, 10(1), 2023. doi: 10.1186/s40537-022-00680-6.
- Huijian Han, Zhiming Li, and Zongwei Li. Using machine learning methods to predict consumer confidence from search engine data. *Sustainability*, 15(4), 2023. ISSN 2071-1050. doi: 10.3390/su15043100. URL https://www.mdpi.com/2071-1050/15/4/3100.
- Zixian Huang, Ao Wu, Jiaying Zhou, Yu Gu, Yue Zhao, and Gong Cheng. Clues be- fore answers: Generation-enhanced multiple-choice QA. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 3272–3287, Seattle, United States, July 2022. Asso- ciation for Computational Linguistics. doi: 10.18653/v1/2022.naacl-main.239. URL https://aclanthology.org/2022.naacl-main.239.
- Joshua Robinson, Christopher Michael Rytting, and David Wingate. Leveraging large language models for multiple choice question answering. *ArXiv*, abs/2210.12353, 2022. URL https://api.semanticscholar.org/CorpusID:253098700.
- Akshay Chaturvedi, Onkar Pandit, and Utpal Garain. CNN for text-based multiple choice question answering. In *Proceedings of the 56th Annual Meeting of the Association for Com- putational Linguistics (Volume 2: Short Papers)*, pp. 272–277, Melbourne, Australia, July 2018. Association for Computational Linguistics. doi: 10.18653/v1/P18-2044. URL https://aclanthology.org/P18-2044.
- Di Jin, Shuyang Gao, Jiun-Yu Kao, Tagyoung Chung, and Dilek Hakkani-tur. Mmm: Multi- stage multi-task learning for multi-choice reading comprehension, 2019.
- Zhipeng Chen, Yiming Cui, Wentao Ma, Shijin Wang, and Guoping Hu. Convolutional spatial attention model for reading comprehension with multiple-choice questions. *Pro- ceedings of the AAAI Conference on Artificial Intelligence*, 33(01):6276–6283, jul 2019. doi: 10.1609/aaai.v33i01.33016276. URL https://doi.org/10.1609%2Faaai.v33i01.33016276.
- Zixian Huang, Ao Wu, Yulin Shen, Gong Cheng, and Yuzhong Qu. When retriever-reader meets scenario-based multiple-choice questions. In *Conference on Empirical Methods in Natural Language Processing*, 2021. URL https://api.semanticscholar.org/CorpusID: 237364132.
- Doe, J. 2025. The Future of Computing. *Communications of the ACM* 69, 10 (Oct. 2025), 123–456. https://doi.org/10.1145/xxxxxxx.xxxxxxx.