Predicting Ratings of Indian IPOs from Red Herring Prospectus

Sohom Ghosh and Sudip Kumar Naskar

January 29, 2025

Predicting Ratings of Indian IPOs from Red Herring Prospectus

Sohom Ghosh*,[1] and Sudip Kumar Naskar*,[2]

[1]Jadavpur University, Kolkata, West Bengal, India, ORCiD: 0000-0002-4113-0958.
[2]Jadavpur University, Kolkata, West Bengal, India, ORCiD: 0000-0003-1588-4665.

Contributing authors: sohomg.cse.rs@jadavpuruniversity.in; sudipkumar.naskar@jadavpuruniversity.in;

Abstract

In recent years, Initial Public Offerings (IPOs) of Indian companies have emerged as a popular investment opportunity, with many investors seeking quick returns through listing gains. However, analysing lengthy prospectuses to make informed, data-driven investment decisions can be a cumbersome task. To address this challenge, we propose a task to mine red herring prospectuses of companies planning to go public and classify them into four categories: Apply, Neutral, May Apply, or Avoid. This method aims to streamline the decision-making process for investors by providing clear and concise recommendations based on the prospectus data.In addition to introducing two new datasets, we propose a novel method for predicting ratings of Indian IPOs that surpasses the performance of existing state-of-the-art Large Language Models.

Keywords: Financial Texts, Indian IPO, Natural Language Processing, Large Language Models

*Corresponding Authors

1 Introduction

An Initial Public Offering (IPO) is the process by which a private company first offers its shares to the public, transitioning to public ownership. This event enables the company to raise capital by selling ownership stakes to individual and institutional investors.

In the Indian context, IPOs are categorized into Mainboard (MB) IPOs and Small and Medium Enterprises (SME) IPOs, each serving distinct market segments. Mainboard IPOs are intended for larger, established companies that meet stringent regulatory requirements and are listed on major stock exchanges such as the BSE and NSE. These offerings typically attract a broad investor base, involve larger issue sizes, and provide higher liquidity and market recognition.

Conversely, SME IPOs cater to small and medium enterprises, which often have more relaxed eligibility criteria and are listed on specialized platforms like BSE SME and NSE Emerge. SME IPOs generally involve smaller investment amounts and a limited number of allottees, making them accessible to retail investors but associated with higher risks and lower liquidity compared to Mainboard IPOs. The IPO prospectus is a critical legal document that provides potential investors with comprehensive information about the company and its offering. It serves as a transparency tool, enabling informed investment decisions. There are two types of IPO prospectuses: the Draft Red Herring Prospectus (DRHP) and the Red Herring Prospectus (RHP). The DRHP is the initial document filed with the Securities and Exchange Board of India (SEBI) prior to launching an IPO. It outlines the company's business model, financial statements, risk factors, and intended use of raised funds, subject to SEBI's review and approval. Upon receiving SEBI approval, the company issues the RHP, which includes updated information such as the final offer price and number of shares offered. The RHP is made available to potential investors during the offer period, providing essential details for investment decision-making. Both DRHP and RHP are vital components of the IPO process, ensuring that investors have access to accurate and comprehensive information about the company's financial health. Typically, DRHPs range from 100 to 300 pages in length, while RHPs are usually between 80 and 250 pages. Reading these lengthy documents can be time-consuming and overwhelming for novice investors. Reviews and ratings provided by experts can often be subjective and biased. For popular IPOs, investors are frequently inundated with expert reviews. However, for lesser-known IPOs, expert reviews are rarely available.

IPO grading is an evaluation process that assesses the fundamentals of a company's initial public offering (IPO) relative to its peers. This grading provides investors with an independent opinion on the quality and potential of the IPO, helping them make informed investment decisions. In India, IPO grading became mandatory in April 2007 for all new issues, as mandated by the Securities and Exchange Board of India (SEBI). This requirement aims to enhance transparency and encourage independent research in the equity market. However, The effectiveness of IPO grading was questioned, leading SEBI to make it optional starting February 4, 2014. [1]. Furthermore, IPO grading is

[1] https://www.angelone.in/knowledge-center/ipo/ipo-grading (accessed on 19$^{\text{th}}$ January, 2025)

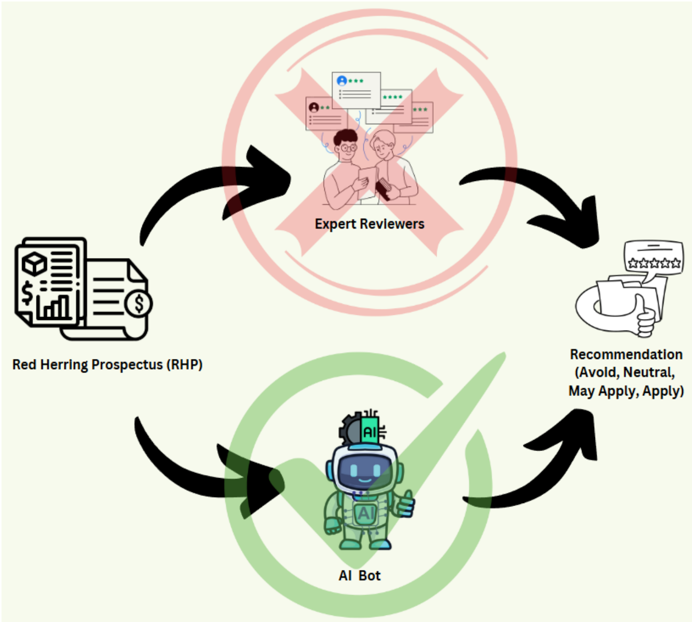

Fig. 1 IPO Rating Prediction

inherently subjective and can vary across different rating agencies. This variability raises questions about the consistency and reliability of the grades assigned. [2]

Therefore, an automated system capable of mining these prospectuses would facilitate the development of a decision-making tool to assist investors in determining whether to subscribe to an IPO. This is presented in Figure 1.

Our Contributions

Our contributions are as follows:

- We introduce two new India-specific datasets (one for Main Board IPOs and another for SME IPOs) along with a task focused on predicting the ratings of these IPOs.
- We propose a novel method for mining prospectus of these IPOs which consists of a Retrieval Augmented Generation framework along with a fine-tuned small encoder based language model. This method outperforms state-of-the-art Large Language Models (LLMs) under zero shot settings.

2 Related Works

The prediction of Initial Public Offering (IPO) performance has garnered significant attention in the financial literature, particularly due to its implications for investors

[2] https://www.motilaloswal.com/blog-details/what-is-ipo-grading-process-in-india/21319 (accessed on 19^{th} January, 2025)

and market efficiency. Various studies have investigated the determinants of IPO performance, emphasizing factors such as market conditions, investor behaviour, and corporate governance. A key aspect of this research is the phenomenon of IPO underpricing, which is crucial for understanding overall IPO performance. Most prior studies have concentrated on short-run underpricing [1] [2] or long-run underperformance [3].

Some researchers have explored the usefulness of IPO grading. The study [4] indicates that a substantial number of retail investors are familiar with the IPO grading process. However, perceptions of its effectiveness and influence on investment decisions vary. While IPO grading is considered a valuable tool for investors [5], its impact is not consistent across different segments of the investor population. As per [6], securities with higher IPO grades are observed to exhibit a lower degree of underpricing. Additionally, higher IPO grades are associated with an increase in subscription rates across all kinds of investors. The influence of credit ratings on IPO underpricing has been well-documented. Dhamija and Arora found that firms with credit ratings experience significantly less underpricing than those without, indicating that improved corporate governance and transparency can lead to better IPO valuations [7]. Jacob and Agarwalla [8] explored the effects of mandatory IPO grading in India. They concluded that such certifications can enhance demand of institutional investors, but their impact on overall pricing efficiency is limited. All of these studies highlight the significance of IPO grading; however, none of them propose automated methods for grading IPOs. On the contrary, automated methods for predicting ratings from texts [9] have been well-studied in several domains like e-commerce [10], local service [11], etc. Consequently, we present the task of predicting ratings based on the prospectuses of Indian companies that are preparing for IPOs. This task is similar to automated grading of IPOs and it would pave the way for a valuable tool that empowers investors with data-driven insights to make more confident and informed decisions regarding IPO subscriptions. To the best of our knowledge, the proposed task represents a novel contribution to this field.

3 Problem Statement

Given a company's IPO prospectus, our objective is to comprehend its content and categorize it into one of four classifications: Apply, Neutral, May Apply, or Avoid, providing a concise and informed assessment of the investment opportunity. As this a classification problem with class imbalances, we will use Micro, Macro, and weighted F1 score for evaluation.

4 Dataset

We gathered data on MB and SME IPOs separately from the chittorgarh website.[3] The MB data is available from 2011, while SME data starts from 2012. Our collection of this data continued until November 7, 2024, and includes the following information: Review Title (this contains name of the company as well), Name of the Author / Organization who wrote the review, Year of the IPO, Link to access the review, Link

[3]https://chittorgarh.com/ (accessed on 19th January, 2025)

to a webpage containing comprehensive details about the IPO, Key (Unique identifier of each row), Link to access the (D)RHP in PDF format, Name of the JSON file having text contents extracted from (D)RHP, Text content of the review, Recommendation (Apply, Neutral, May Apply, or Avoid).

To ensure data quality, we excluded entries without reviews or recommendations. Notably, mainboard IPOs often have multiple reviews; in such cases, we retained only those reviews that aligned with the majority recommendation. For example, if a company has five reviews—three recommending "Apply" and two recommending "Avoid"—we would keep only the three "Apply" reviews. Conversely, 97% of SME IPOs have reviews authored by a single individual, leading us to discard the remaining 3% of data. For reviews provided in PDF format, we utilized PyPDF [4] to extract text. The Draft Red Herring Prospectuses (DRHP) and Red Herring Prospectuses (RHP), were available in PDF format. In instances where both DRHP and RHP were present, we prioritized the RHP. To further ensure the quality of our data, we compared IPO ratings with their actual opening prices. For Main Board IPOs, we found that in 82.17% of cases, an 'Apply' recommendation corresponded to an opening price higher than the issue price. For SME IPOs, this figure was 83.49%. In total, we collected 1,830 instances for mainboard IPOs and 1,131 for SME IPOs. Data up to 2023 was used for training purposes, while data from 2024 was reserved for testing. We present the data distribution with respect to year and recommendations in Figures 2 and 3 respectively.

The copyright for this content belongs to its respective owners, and we do not claim any copyright rights over this data. This dataset has been released under the CC-BY-NC-SA-4.0 licence for non-commercial research purposes only. We are not liable for any monetary loss that may arise from the use of these datasets and model artefacts.

5 Experiments and Results

In this section, we describe the experimented we conducted and discuss the corresponding results.

Due to budget limitations and computational constraints, we were unable to use the entire prospectus in PDF format into the LLMs at once. Additionally, as noted in [12], larger context sizes can lead to a decrease in the performance and reasoning capabilities of LLMs. Therefore, it was essential for us to extract specific sections from the prospectus that were most relevant to determining the ratings of the IPOs. Thus, we conducted a randomized selection of 200 reviews for both MB and SME IPOs separately. The limitation to 200 reviews was necessitated by the rate limit of Groq [5] API's free tier. The selected reviews were processed using the Llama-3 8B model [13], from which we extracted questions utilizing the prompt specified in Section A. Subsequently, these questions were submitted to Perplexity.ai Pro [6] to compile a comprehensive list of distinct questions, which are presented in Section B. The rationale for employing two different large language models (LLMs) stemmed from the superior capabilities of the Perplexity Pro model in handling complex tasks, albeit

[4] https://pypi.org/project/pypdf/ (accessed on 19th January, 2025)
[5] https://console.groq.com/docs/overview (accessed on 19th January, 2025)
[6] https://www.perplexity.ai/ (accessed on 19th January, 2025)

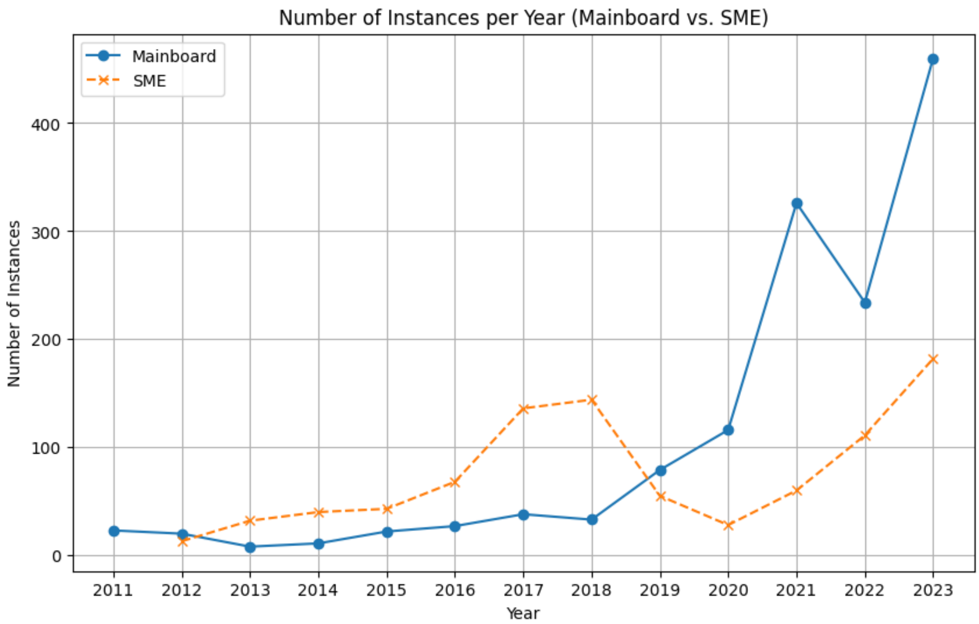

Fig. 2 Data Distribution up to year 2023

limited to two queries per day under the free tier. In contrast, the smaller Llama-3 8B [13] model allowed for multiple queries. We utilized the expert reviews solely for extracting the questions mentioned above and did not use them in any other steps of the process.

Following the methodology outlined in [14], we extracted text from the prospectus (RHP) which were present in PDF format. Optical character recognition (OCR) was performed using Tesseract to extract text from images within the documents. Each page was converted into embeddings utilizing Nomic [15]. Employing a Retrieval-Augmented Generation (RAG) framework, for each of compiled questions mentioned in Section B, we identified the two most pertinent pages based on two criteria: first, through cosine similarity for semantic matching, and second, via BM25 [16] for syntactic similarity. The retrieved pages, along with their corresponding questions, were then passed into the Llama-3.2 3B [13] model to generate answers. Details relating to the prompt we used is mentioned in section A. This process yielded a total of 16 answers for each instance, corresponding to the 16 questions posed.

We employed a zero-shot approach by prompting the Gemma-2 9B, Llama 3.1 70B, and Llama-3.2 3B models to classify the aggregate of 16 answers into one of four categories: Apply, Neutral, May Apply, or Avoid. Details of the prompts are provided in section A. We then repeated these experiments by substituting the aggregate of answers with a single summary. These summaries were generated using Llama-3.2 3B [13]. Prompt details are presented in A. We observe, this change led to improved model

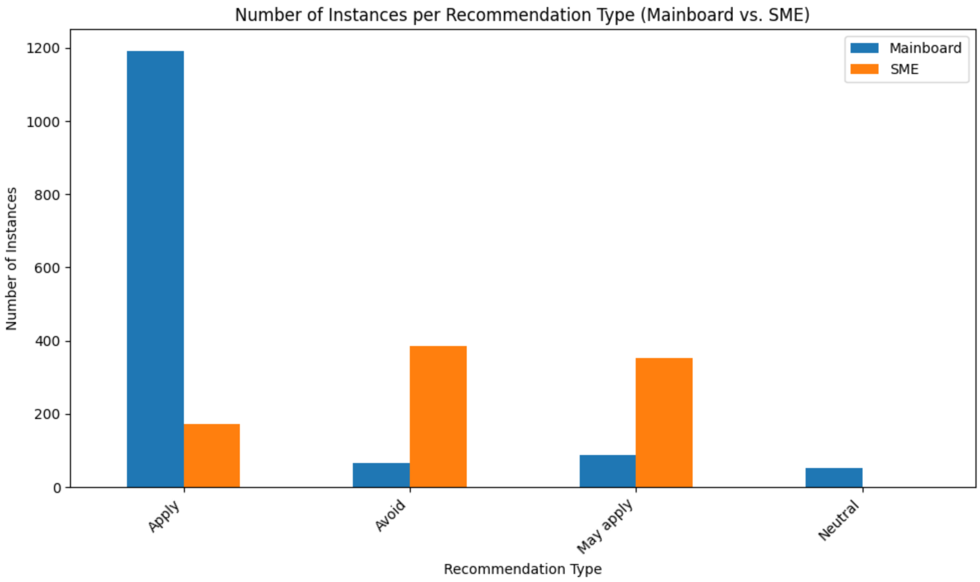

Fig. 3 Distribution of Recommendations

performance in most cases. Subsequently, we fine-tuned Llama-3.2 3B and Gemma-2 9B using supervised fine-tuning methods.

Finally, we trained three encoder based models (RoBERTa [17], LongFormer RoBERTa [7], and DeBERTa [18]) with the summaries for classification. The hyper-parameters are mentioned in Appendix C.

We observed that for MB IPOs, the LongFormer RoBERTa outperformed all other models in terms of micro, macro, and weighted F1 scores. In contrast, for SME IPOs, the Gemma-2 9B model excelled in micro F1 scores, while the Llama 3.1 70B model achieved the highest macro F1 scores. Additionally, the RoBERTa model demonstrated superior performance in terms of the macro FA score.

We present the overall flow in Figure 4 and results in Table 1.

6 Conclusion

In this paper, we introduce the task of mining the prospectuses of Indian companies preparing for IPOs to predict their ratings. To support this task, we propose two new datasets: one for SME IPOs and another for Main Board IPOs. Additionally, we present a novel framework that utilizes Retrieval-Augmented Generation (RAG) to extract relevant sections from the prospectus, summarize them, and employ fine-tuned encoder-based small language models to predict the final ratings. Our approach

[7] https://huggingface.co/markussagen/xlm-roberta-longformer-base-4096 (accessed on 19$^{\text{th}}$ January, 2025)

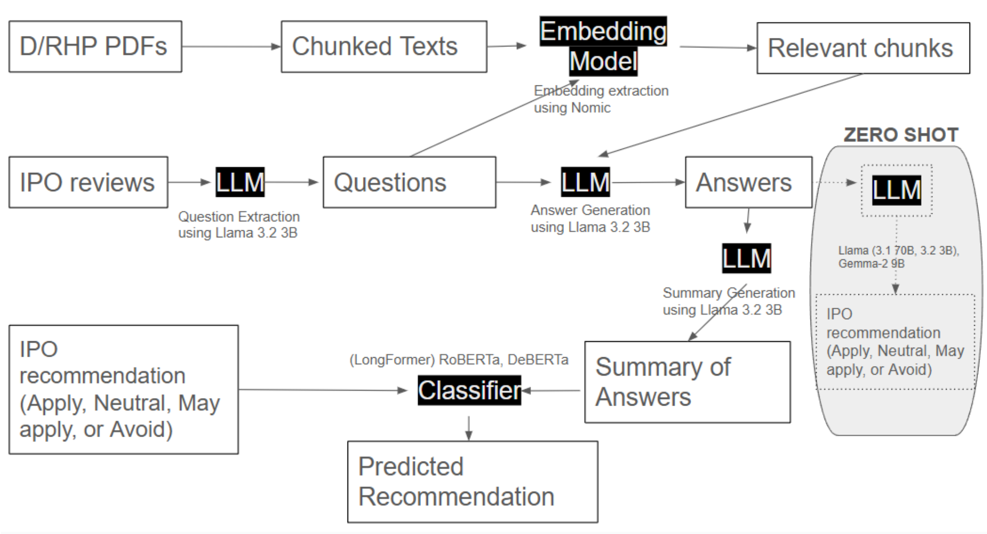

Fig. 4 Detailed Flowchart narrating our methodology

demonstrates superior performance compared to existing state-of-the-art large language models, such as Llama 3.1 70B and Gemma-2 9B, when evaluated under zero-shot settings.

This research has a few limitations that highlight opportunities for future work. Firstly, the list of questions we curated to extract relevant portions from the prospectuses of Main Board IPOs is not exhaustive and may not encompass all the critical details necessary for informed decision-making. In the future, we aim to make this list dynamic, adapting it based on various factors such as industry, profitability, and other relevant criteria.

Due to budget constraints, we were unable to process the entire text corpus (RHP) at once. While we are focused on rating predictions, the actual opening price will ultimately reveal the true performance. Analyzing the actual opening price data will provide more meaningful insights. Additionally, we utilized APIs from various service providers, including Groq [8], Cerebras [9], and OpenRouter [10], to evaluate the performance of different LLMs, such as Llama 3.1 70B [13] and Gemma-2 9B [19], under zero-shot settings. However, the same LLM may produce slightly varying results depending on the service used.

Declarations

Funding. Not Applicable.

[8] https://groq.com/ (accessed on 19$^{\text{th}}$ January, 2025)
[9] https://cerebras.ai/ (accessed on 19$^{\text{th}}$ January, 2025)
[10] https://openrouter.ai/ (accessed on 19$^{\text{th}}$ January, 2025)

		MB			SME		
Model	**Input**	**F1 (m)**	**F1 (M)**	**F1 (w)**	**F1 (m)**	**F1 (M)**	**F1 (w)**
Gemma-2 9B (Zero Shot)	All Answers	0.009	0.007	0.005	0.411	0.189	0.368
Llama-3.1 70B (Zero Shot)	All Answers	0.039	0.021	0.054	0.374	0.176	0.355
Llama-3.2 3B (Zero Shot)	All Answers	0.484	0.184	0.348	0.076	0.038	0.114
Gemma-2 9B (Zero Shot)	Summary	0.023	0.108	0.012	**0.516**	0.256	0.416
Llama-3.1 70B (Zero Shot)	Summary	0.115	0.044	0.191	0.457	0.281	**0.423**
Llama-3.2 3B (Zero Shot)	Summary	0.162	0.077	0.255	0.429	0.163	0.361
Llama 3.2 3b (SFT)	Summary	0.836	0.228	0.883	0.361	0.299	0.347
Gemma 2 9B (SFT)	Summary	0.716	0.233	0.814	0.402	0.298	0.349
RoBERTa	Summary	0.769	0.219	0.846	0.406	**0.335**	0.377
LongFormer RoBERTa	Summary	**0.968**	**0.246**	**0.952**	0.224	0.126	0.090
DeBERTa	Summary	0.912	0.239	0.925	0.457	0.319	0.383

Table 1 Model Performances. m = micro, M = Macro, w = weighted, SFT = Supervised Fine-tuning. Best performing models are highlighted in bold.

Conflict of interest. On behalf of all authors, the corresponding author states that there is no conflict of interest. All the opinions expressed here are those of the authors and do not reflect the views of their affiliations. This work is entirely non-commercial and is conducted for academic research purposes.

Ethics approval. This research did not involve any human participants and/or animals. There was no need for informed consent.

Availability of data and material. The datasets used in this paper can be obtained from https://huggingface.co/datasets/sohomghosh/indian_ipo_rating_prediction.

Code availability. Our code base is available at https://huggingface.co/datasets/sohomghosh/indian_ipo_rating_prediction.

Author contributions. Sohom Ghosh conducted the experiments and prepared the manuscript. Sudip Kumar Naskar re-examined it. All authors reviewed the manuscript.

Appendix A Prompts

Question Extraction Prompt:
The prompt used for extracting questions is:
You are an expert financial analyst who have extensive experience of participating in Initial Public Offerings (IPOs) of Indian companies. You are given a review about an Indian company going for IPO. Extract a list of key questions which have been answered in the given review and which would help in determining whether to apply

for the IPO. Return just a list of questions which can be answered from the review. Do not return anything other than the list of questions. Review: {review content}
Response:

Answer Generation Prompt:
This prompt was used for each of the 16 questions to generate the corresponding answer.
You are an expert financial analyst who have extensive experience of participating in Initial Public Offerings (IPOs) of Indian companies. Relevant contents from Red Herring Prospectus (RHP) of an Indian company going for IPO is given to you. Your task is to analyse and answer the given question in less than 300 words as free text. Use just the content provided to you to answer the question and not anything else. If the contents are not relevant, just return the word 'None'.
CONTENT-1: {semantically relevant content }
CONTENT-2: {syntactically relevant content}
Question: {question}
Response:

Summary Generation Prompt:
The prompt used for generating summary from answers is as follows:
You are an expert financial analyst who have extensive experience of participating in Initial Public Offerings (IPOs) of Indian companies. You are provided with various facts about a company going for IPO in the form of answers. Your task is to analyse these answers and generate a summary comprising of key points that investors needs to know to decide if they should subscribe for the IPO or not. If you are not confident answer nan. Just return the summary in 300 words and nothing else. Facts about the company's IPOs are as follows: {answers of 16 questions}.
Response:

Rating Prediction Prompt:
The prompt used for zero shot classification is:
"You are an expert financial analyst who has extensive experience of participating in Initial Public Offerings (IPOs) of Indian companies. You are given various facts of a company. Your task is to analyse these facts and decide whether an investor should 'Avoid', 'May apply', 'Apply', or, be 'Neutral' for the IPO. Your answer should be in a JSON structure with two keys, 'prediction' and 'justification'. The value corresponding to 'prediction' key should be 0,1,2, or, 3 only where 0 represents 'Avoid', 1 represents 'Neutral', 2 represents 'May apply', and 3 represents 'Apply'. The value corresponding to 'justification' key should be the explanation behind the prediction. Facts: {answers of 16 questions concatenated side by side}.
Response:"

Appendix B Questions

The list of 16 extracted questions is presented here.

- What is the price band and issue price of the IPO?
- What is the issue size and how many shares are being issued as part of the IPO?
- What is the implied market capitalization of the company after the IPO?
- How will the company utilize the funds raised through the IPO, and what is the purpose of the IPO?
- What is the company's revenue growth rate over recent financial years, and how has its financial performance been historically (including revenue, EBITDA, and net profit trends)?
- What are the key financial ratios, such as net profit margin, return on equity (RoE), return on capital employed (RoCE), and total debt?
- What is the shareholding pattern before and after the IPO, and who are the promoters?
- Are there any regulatory issues or conflicts of interest affecting the company?
- What are the company's plans for expansion and future growth, and how does it position itself in terms of competition within its industry?
- Who are the company's major customers, what is the revenue breakdown by sector, and is there a dependency on large institutional customers?
- What are the potential risks associated with increasing raw material costs, and what other risks does the company face?
- How does the company's valuation compare to its peers, and is the issue priced aggressively compared to industry standards?
- What is the competitive landscape of the industry in which the company operates?
- Has the company declared any dividends in the past, and what is its dividend policy?
- Who are the lead managers and registrar for the IPO, and what is their track record in terms of past IPO listings?
- Are there any concerns regarding transparency or missing details in the offer document?

Appendix C Hyper-parameters

Encoder based models
learning_rate=2e-5, per_device_train_batch_size=1, per_device_eval_batch_size=1, num_train_epochs=5, gradient_accumulation_steps=4, weight_decay=0.01

Sample code: https://huggingface.co/datasets/sohomghosh/indian_ipo_rating_prediction/blob/main/ipo-review-longformerroberta-classify-summarised.ipynb

Decoder based models
max_seq_length = 204, load_in_4bit = True, lora_alpha = 16, lora_dropout = 0, bias = "none", use_gradient_checkpointing = "unsloth", random_state = 3407, use_rslora = False, dataset_num_proc = 2, packing = False, per_device_train_batch_size = 2, gradient_accumulation_steps = 4, warmup_steps = 5, num_train_epochs=5, learning_rate = 2e-4, optim = "adamw_8bit", weight_decay = 0.01, lr_scheduler_type = "linear"

Sample code: https://www.kaggle.com/code/danielhanchen/fixed-kaggle-llama-3-2-1b-3b-conversation (accessed on 22$^{\text{nd}}$ January, 2025)

References

[1] Anand, R., Singh, B.: Effect of composition of board and promoter group retained ownership on underpricing of indian ipo firms: An empirical study. Indian Journal of Corporate Governance **12**(1), 21–38 (2019) https://doi.org/10.1177/0974686219836539

[2] Bajo, E., Raimondo, C.: Media sentiment and ipo underpricing. Journal of Corporate Finance **46**, 139–153 (2017) https://doi.org/10.1016/j.jcorpfin.2017.06.003

[3] Sahoo, S., Rajib, P.: After market pricing performance of initial public offerings (ipos): Indian ipo market 2002–2006. Vikalpa **35**(4), 27–44 (2010) https://doi.org/10.1177/0256090920100403

[4] Sarin, V., Sidana, N.: A study of perceptions of investors towards ipo grading in india. International Journal of Economic Research **14**(20), 757–770 (2017)

[5] Deb, S.S., Marisetty, V.B.: Information content of ipo grading. Journal of banking & Finance **34**(9), 2294–2305 (2010)

[6] Poudyal, S.: Grading Initial Public Offerings (IPOs) in India's Capital Markets: A Globally Unique Concept. Indian Institute of Management, ??? (2008)

[7] Dhamija, S., Arora, R.K.: Impact of quality certification on ipo underpricing: Evidence from india. Global Business Review **18**(2), 428–444 (2017) https://doi.org/10.1177/0972150916668611 https://doi.org/10.1177/0972150916668611

[8] Jacob, J., Agarwalla, S.K.: Mandatory ipo grading: does it help pricing efficiency? Vikalpa **40**(2), 132–144 (2015)

[9] Khan, Z.Y., Niu, Z., Sandiwarno, S., Prince, R.: Deep learning techniques for rating prediction: a survey of the state-of-the-art. Artificial Intelligence Review **54**, 95–135 (2021)

[10] Qu, L., Ifrim, G., Weikum, G.: The bag-of-opinions method for review rating prediction from sparse text patterns. In: Huang, C.-R., Jurafsky, D. (eds.) Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010), pp. 913–921. Coling 2010 Organizing Committee, Beijing, China (2010). https://aclanthology.org/C10-1103/

[11] Lei, X., Qian, X., Zhao, G.: Rating prediction based on social sentiment from textual reviews. IEEE transactions on multimedia **18**(9), 1910–1921 (2016)

[12] Fraga, N.: Challenging llms beyond information retrieval: Reasoning degradation with long context windows. Preprints (2024) https://doi.org/10.20944/preprints202408.1527.v1

[13] AI@Meta: Llama 3 Model Card (2024). https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md

[14] Ghosh, S., Maji, A., Vardhan, N.H., Naskar, S.K.: Experimenting with Multimodal Information to Predict Success of Indian IPOs (2024). https://arxiv.org/abs/2412.16174

[15] Nussbaum, Z., Morris, J.X., Duderstadt, B., Mulyar, A.: Nomic Embed: Training a Reproducible Long Context Text Embedder (2024)

[16] Lù, X.H.: BM25S: Orders of magnitude faster lexical search via eager sparse scoring (2024). https://arxiv.org/abs/2407.03618

[17] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: RoBERTa: A Robustly Optimized BERT Pretraining Approach (2019). https://doi.org/10.48550/arXiv.1907.11692

[18] He, P., Liu, X., Gao, J., Chen, W.: Deberta: Decoding-enhanced bert with disentangled attention. arXiv preprint arXiv:2006.03654 (2020)

[19] Team, G., Riviere, M., Pathak, S., al.: Gemma 2: Improving Open Language Models at a Practical Size (2024). https://arxiv.org/abs/2408.00118