Going public with Chat-GPT: Can Large Language Models price U.S. IPO’s

Keywords: Natural Language Processing, Machine Learning in Finance, Large Language Models, Textual analysis

**Project description**

The decision to go public in an Initial Public Offering (IPO) is an attractive opportunity for firms to raise capital in exchange for shares of the company. While this is a lengthy process, one of the key aspects is the issuance of an IPO-prospectus. This prospectus gives investors the possibility to evaluate the future performance of the firm. When investors evaluate the outlook of the firm, they traditionally focus on key financial indicators presented in this document. However, when investors apply this approach, relevant contextual information is neglected (Loughran, Mcdonald, 2013).

Recent research by Lopez-Lira and Tang (2023), along with Breitung and Müller (2023), highlight the potential of large language models to capture additional information disclosed in firm reports. By integrating large language models, this project aims to innovate upon existing methodologies for assessing corporate performance, offering a novel approach to investment analysis.

In this project, the student will process the information presented in IPO-prospectus. To do so, the student will engineer a system that exploits the capabilities of large language models (e.g., GPT-4). Initially, the student will collect various IPO-prospectus from the database of the U.S. Security and Exchange Comission (SEC). In a second step, the student will implement a parser to dissect the documents and extract relevant information so that it can be stored in a well-structured format, suitable for further analysis. In the next stage, a large language model like GPT-3 or GPT-4, is used to extract relevant information from the documents. Here, the student will develop a method to price the firm and gauge the optimism regarding its future prospects. The pricing method will be developed in close collaboration with the project advisor employing prompt engineering methods. Lastly, the resulting metric is to be benchmarked with real-time data to assess the performance of the Large-Language Model.
What we are looking for

- Strong analytical and project management skills
- Determination and passion for your areas of expertise
- Good Python programming skills
- Interest to work at the intersection of finance and IT
- 1 or 2 persons

What we offer

- Knowledge in quantitative finance, corporate finance and machine learning
- Kick-off session including introduction to relevant finance and/or business topics
- Experience with IDPs
- Open dialogue and support
- Access to prime capital markets databases (Bloomberg, Datastream, Thomson Reuters, etc)
- Potential for publication and/or evaluation of future use cases
- Both single and group projects are possible

Interested?

Please send an e-mail with CV, academic transcript and your preference for this project to sebastian.mueller.hn@tum.de

Questions?

In case of any (e.g. topic related) questions, please contact Sebastian Müller (sebastian.mueller.hn@tum.de)

References

