Chair of Financial Management and Capital Markets
TUM School of Management
Technical University of Munich

Modelling Business Segments using Language Models

Keywords: International Stock Market Data, Patents, NLP

Project description

The recent trend of specialization in global economies has resulted in companies offering highly specialized products even in similar industries. Using business descriptions, Breitung and Müller (2025) show that companies in the same industry can have a unique network of affiliated companies. However, as many firms nowadays operate globally with a comprehensive product portfolio it is useful to investigate business segments of firms in more detail. Patents can be an important indicator of a firm's product portfolio and their future endeavors. Further, high patenting activity is a strong driver of firm's future success (Hirshleifer et al, 2013).

This project aims to construct detailed business segments for a global dataset of firms. The business segments will be represented using the patenting activity of a firm by analyzing the content and text of patent filings. Due to the nature of patent data, this method allows for a time-varying analysis of all market segments companies operate in.

The student will begin by familiarizing herself with the global patent data. At the start the student will use various methods to describe and cluster the patents of one chosen firm. Using modern NLP methods, the student will model this firm's business segments as accurately as possible. As a ground truth business description and other sources should be used to compare the accuracy of the newly created business segments.

If the business segments can be accurately modelled for one firm in the next step the student will scale the previously developed methodology to a large set of public companies. Doing so the student will create a detailed global description of firm's business segments. Using patent data these segments can be analyzed further, as the data allows us to investigate the markets that specific firm segments operate in or the intensity with which firms invest into respective segments.

Throughout the project, the student will work closely with the supervisor, receiving guidance on project milestones, and best practices in financial data extraction.

Chair of Financial Management and Capital Markets
TUM School of Management
Technical University of Munich

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 Müller via yannik.gabelmann@tum.de

Questions?

In case of any (e.g. topic related) questions, please contact Yannik Gabelmann (yannik.gabelmann@tum.de)

References

Breitung, Christian, and Müller, Sebastian. (2025). “Global Business Networks“, In Journal of Financial Economics, (166), 104007.

Hirshleifer, David, Hsu, Po-H and Li, Dongmei. (2013). “Innovative efficiency and stock returns“, In Journal of Financial Economics, (107), 632-654.