

Textual Analysis in Finance and Accounting

Supervisors: [Dr Rodion Skovoroda](#), [Professor Ali Ataullah](#), [Professor Tomasz Piotr Wisniewski](#) and [Dr Cristiana Bernadi](#), Department of Accounting and Finance, The Open University Business School

Project Description:

The finance and accounting research is increasingly using modern textual analysis techniques for extracting valuable insights from vast repositories of unstructured data. This PhD project call seeks to explore the innovative intersection of textual analysis, machine learning, and large language models (LLMs) within the domains of finance and accounting. The primary objective is to harness advanced computational techniques to analyse financial documents, corporate disclosures, earnings reports, and social media content, thereby unveiling patterns, trends, and predictive indicators that remain obscured in traditional numerical data analyses.

The ability to efficiently process and interpret textual information has been shown to generate novel results in accounting (Bochkay *et al.*, 2023; Dong *et al.*, 2023), finance (Bybee *et al.*, 2023; Huang *et al.*, 2023; Loughran and McDonald, 2020; Yang *et al.*, 2020) and management (Bellstam *et al.*, 2021) with potentially profound implications for financial analysts, investors, corporate finance managers, accountants, and regulatory bodies.

This PhD project call invites driven candidates who are eager to contribute to advanced research at the nexus of computational linguistics, finance, and accounting. The project aims to develop novel methodologies and novel applications of existing methodologies that may relate to a range of different topics in accounting and finance, for instance:

- Topic modelling of company sustainability (ESG) reports and associated climate-related risks
- ESG-linked CEO compensation analysis and the link to financial performance
- Sentiment analysis of financial news
- Automated fraud detection

Expectation: The PhD student is expected to develop specific research questions and gradually develop three working papers over the period of three years that are potentially publishable in internationally leading academic journals.

It is hoped that the prospective candidate will demonstrate an interest in developing both a theoretical/conceptual and empirical contribution within this field.

Applicant Specification: The candidate must express a keen interest in modern textual analysis techniques. The candidate must be proficient in the Python programming language. Prior knowledge and experience in quantitative research methodologies within the broad domains of finance, accounting, and economics is required. A master's degree in finance, accounting, economics, or equivalent is required.

About the Supervisors:

Rodion Skovoroda : My research interests currently include novel applications of Machine Learning methods to the analysis of dynamic narratives with applications to (sustainable) finance and management. My research on Executive Pay and managerial incentives with applications to Corporate Finance and Corporate Governance emphasises the need for further improvements in the way executive pay is disclosed, measured and analysed. My other research interests include topics in International/Applied Economics such as Foreign Direct Investment and Political Risks. My research has been published in leading international journals such as Journal of International Business Studies, British Journal of Management, British Journal of Industrial Relations, Long Range Planning, Economics Letters, etc.

Ali Ataullah: My research focuses on two important themes: (1) corporate finance and societal challenges (e.g. environmental justice and workers' welfare), and (2) financial well-being of people affected by health conditions. My expertise in corporate finance will be very relevant to this PhD project. In particular, my recent published and ongoing work explores the impact of firms' toxic waste releases on corporate financial policies. The textual analysis in the PhD project will benefit from and contribute to this line of research.

Tomas Piotr Wisniewski: My primary research interests revolve around asset pricing and investment strategies. I am interested in whether stock prices reflect available information efficiently or merely follow the behavioural tendencies of traders. More recently, I have concentrated my research efforts on examining the interplay between stock markets and political developments. Several of my publications also investigate the use of technology in payment systems and banking. I have used automated content analysis in my work to extract information from the text of annual reports and other disclosures to inform stock market trading decisions,

Cristiana Bernardi: My main research interest lies within the area of financial reporting, particularly in the information content of mandatory and voluntary disclosures concerning non-financial performance. Her current research is on climate change risk disclosure and sustainability within the cocoa supply chain. Another area of research she is focusing on concerns COVID-19 regulatory news communications.

References:

- Bellstam, G., Bhagat, S. and Cookson, J.A. (2021). "A text-based analysis of corporate innovation". *Management Science*, 67(7), pp. 4004-4031.
- Bochkay, K., Brown, S.V., Leone, A.J. and Tucker, J.W. (2023). "Textual analysis in accounting: What's next?". *Contemporary Accounting Research*, 40(2), pp.765-805.
- Bybee, L., Kelly, B. and Su, Y. (2023). "Narrative asset pricing: Interpretable systematic risk factors from news text". *The Review of Financial Studies*, 36(12), pp.4759-4787.
- Dong, M.M., Stratopoulos, T.C. and Wang, V.X. (2023). "A scoping review of ChatGPT research in accounting and finance". (December 30, 2023). Available at: <http://dx.doi.org/10.2139/ssrn.4680203>.
- Huang, A.H., Wang, H. and Yang, Y. (2023). "FinBERT: A large language model for extracting information from financial text". *Contemporary Accounting Research*, 40(2), pp.806-841.

Loughran, T. and McDonald, B. (2020). "Textual analysis in finance". *Annual Review of Financial Economics*, 12, pp.357-375.

Yang, Y., Uy, M.C.S. and Huang, A. (2020). "Finbert: A pretrained language model for financial communications". Available at: <https://doi.org/10.48550/arXiv.2006.08097>.