Notice for the PhD Viva-Voce Examination

Mr Nagendra B V, Registration Number: 1981906, PhD Scholar at the Department of Statistics and Data Science, School of Sciences, CHRIST (Deemed to be University) will defend his PhD thesis at the public viva-voce examination on Wednesday, 06 August 2025 at 11.00 am in Room No. 05, Ground Floor, R & D Block, CHRIST (Deemed to be University), Bengaluru - 560029, Karnataka, India.

Title of the Thesis	:	**Linguistic Discovery Using Lexicon and Machine Learning Approaches in Financial Earnings Calls**
Discipline	:	**Data Science**
External Examiner - I	:	**Dr Sandeep Kumar Panda** Professor Department of Artificial Intelligence and Data Science Faculty of Science and Technology ICFAI Foundation for Higher Education Hyderabad - 501203 Telangana
External Examiner - II	:	**Dr Senthil Kumar T** Professor School of Computing Amrita University Coimbatore - 641112 Tamil Nadu
Supervisor	:	**Dr Kumar Chandar S** Professor School of Business and Management CHRIST (Deemed to be University) Bengaluru - 560029 Karnataka

The members of the Research Advisory Committee of the Scholar, the faculty members of the Department and the School, interested experts and research scholars of all the branches of research are cordially invited to attend this open viva-voce examination.

Registrar (Academics)

Place: Bengaluru
Date: 02 August 2025

ABSTRACT

Finance is the backbone of a country's economic development and prosperity. A good financial system is needed to boost capital and money markets of a country. The financial domain, offlate has received a good traction due to spurring growth in economies. The application of text mining techniques, a sub-filed of datascience is increasingly becoming popular in the field of finance and other domains as well. The limitation in terms of information available in quantitative structured data can be overcome through qualitative data. The hidden knowledge available in unstructured corpus like documents is very useful in decision making. Studies have shown how tonality present in unstructured data be used to study change in stock price movements, exchange rate movements and other macro economic indicators. This knowledge is helpful to the investor community in short to medium term decision making which constitutes the technical analysis, however long term decisions are supported by fundamental analysis. Opinion mining from unstructured corpus enable organizations develop business strategies and improve customer experience with better product features and services. In this thesis, the novel concept of "inverse effect" which is the percentage disagreement between the current period sentiment and the changes in the stock price movement in the subsequent period is presented, which helps investors in devicing short to medium term trading strategies and ensure protection of principal.

The study further led to the development of a financial lexicon called FNBLex based on ten years period Earnings Call Transcripts (ECT) collected over fourteen Information Technology (IT) stocks, using naive bayesian engine. The FNBLex provided a better lift over other lexicons like VADER in detecting the inverse effect. Finally the cross domain lexicon transfer was performed on the ECTs collected from the Indian banks with FNBLex. The FNBLex showed a better discriminatory power over VADER and Loughran-McDonald, a widely used financial lexicon, developed using 10-K reports. The study also aims to compare the performance of domain specific lexicon over language models. More specifically, the financial variant of BERT namely the FinBERT, currently the gold standard, is compared with the domain specific lexicon. The summary of the research contributions is provided below.

- Earnings call textual corpora as a source of information to detect anomalies in the changes in the stock price movements
- Development of financial lexicon using historical ten years earnings call using naïve Bayesian generative model and classification of expert label using restrictive and flexible models
- Application of financial lexicon to validate its efficacy in the banking domain, which is termed as cross-domain lexicon transfer
- Validation of the efficacy of language models over domain specific lexicons

Keywords: *Inverse Effect, Earnings Call Transcripts, Sentiments, FNBLex, Lexicon, VADER, Loughran-McDonald, SentiEcon, BERT, FinBERT, naive Bayesian, SVM, biLSTM, cross-domain lexicon transfer*

Publications:

1. **Nagendra BV**, Kumar Chandar S, Simha J B, Yash Kaushal- Extracting Linguistic Tones in Earnings Call using Transformer Model and Performance Comparison with Lexicon-based Approaches, Journal of Trends in Computer Science and Smart Technology 7, no. 1 (2025): 84-99, doi: 10.36548/jtcsst.2025.1.006
2. **Bv, N.,** Simha, J.B. and Abhi, S.- Deploying NLP Techniques for Earnings Call Transcripts for Financial Analysis: A Reverse Phenomenon Paradigm, Proceedings of the7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) held during 11 October 11 - October 13, 2023 hosted jointly by IEEE Inc, doi:10.1109/I-SMAC58438.2023.10290494
3. **Nagendra, B.V.**, Chandar, S.K., Simha, J.B. and Bazil, J.J.- Financial Lexicon based Sentiment Prediction for Earnings Call Transcripts for Market Intelligence, Proceedings of the 5th International Conference on Image Processing and Capsule Networks (ICIPCN) held during July 3 to July 4, 2024, jointly hosted by IEEE Inc, doi:10.1109/ICIPCN63822.2024.00103
4. **Nagendra, B.V.,** Simha, J.B., Manu, K.S., Kirubanand, K.V. and Kaushal, Y.- Cross Domain Lexicon Transfer-A Case Specific to Application in Banking Domain, Proceedings of the 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI) held during January 7 – January 8, 2025, hosted jointly by IEEE Inc, doi: 10.1109/ICMCSI64620.2025.10883579