CHRIST
(DEEMED TO BE UNIVERSITY)
BANGALORE | DELHI NCR | PUNE

Notice for the PhD Viva Voce Examination

Mr Jimmy Jose, Registration Number: 2090204, PhD Scholar at the Department of Computer Science, School of Sciences, CHRIST (Deemed to be University) will defend his PhD thesis at the public viva-voce examination on Thursday, 21 August 2025 at 11.30 am in Room No. 044, Ground Floor, R & D Block, CHRIST (Deemed to be University), Bengaluru - 560029, Karnataka, India.

Title of the Thesis : **Semantic Relevant Sequential Keyphrase Generation Using Deep Reinforcement Learning**

Discipline : **Computer Science**

External Examiner - I : **Dr B E Manjunathswamy**
Professor
Department of Computer Science and Engineering
School of Studies of Engineering and Technology
Guru Ghasidas Vishwavidyalaya
Koni, Bilaspur - 495009
Chhattisgarh

External Examiner - II : **Dr Sajimon Abraham**
Professor
Department of Computer and IT
School of Management and Business Studies
M G University, Kottayam - 686560
Kerala

Supervisor : **Dr Beaulah Soundarabai P**
Associate Professor
Department of Computer Science
School of Sciences
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: 07 August 2025

ABSTRACT

In today's data-driven world, extracting or generating keyphrases is crucial for Natural Language Processing (NLP). Automated Keyphrase Extraction (KPE) and Keyphrase Generation (KPG) are essential for handling vast digital content, benefiting fields like search engine optimization (SEO), social media analysis, and text generation. The existing KPE and KPG methods struggle in dynamic environments and fail to effectively distinguish present and absent keyphrases, leading to poor semantic diversity. This research introduces three novel methodologies to address these issues:

GoD-BERT enhances PLM performance by integrating goal-specific preprocessing, a goal-knowledge graph, Graph Neural Networks (GNN), and adapter-based BERT fine-tuning. This model extracts goal-aware dialogue information and incorporates factual knowledge for better dialogue response generation. Performance evaluation shows GoD-BERT consistently outperforms BERT, GPT-2, and ERNIE across various datasets.

GWebPositionRank improves KeyBERT by incorporating spatial analysis at both local (document-level) and global (web-level) scales. It ranks keyphrases using graph-based methods and WebPositionRank, followed by diversity computation. Results on the SemEval2017 dataset confirm that GWebPositionRank outperforms existing unsupervised KPE techniques by leveraging external web sources and semantic diversity scoring.

MADeGen addresses the challenge of distinguishing present and absent keyphrases using reinforcement learning. It employs a multi-agent framework with an extraction agent and a generation agent, integrating Wikipedia for enriched keyphrase generation. A specialized reward function assesses the generated keyphrases against ground truth, triggering actor-critic-based policy optimization. Performance evaluations show MADeGen significantly improves recall, precision, and F1-score, demonstrating its efficiency in fine-tuning keyphrase generation.

Overall, the proposed methods outperform baseline models, proving the effectiveness of knowledge-aware learning, semantic diversity scoring, and external web sources. These contributions enhance dialogue response generation and KPE/KPG with notable improvements in precision and recall.

Keywords: *Keyphrase Extraction, Semantic similarity measure, Wikipedia source, KeyBERT, Pre-trained BERT model, Keyphrase Generation, Reinforcement Learning, Multi-agent, Extraction agent, Generation agent, Adaptive Reward, and Actor-critic*

Publications:

1. **J. Jose** and B. P. Soundarabai, " MADeGen: Multi-Agent based Deep Reinforcement Learning for Sequential Keyphrase Generation," *International Journal of Intelligent Engineering and Systems, Vol.17, No.6, 2024,* pp. 69-85, doi: 10.22266/ijies2024.1231.07.
2. **J. Jose** and B. P. Soundarabai, "GWebPositionRank: Unsupervised Graph and Web-based Keyphrase Extraction from BERT Embeddings," *IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)*, Bangalore, India, 2024, pp. 45-52, doi: 10.1109/ICWITE59797.2024.10503070.
3. **J. Jose** and B. P. Soundarabai, (2023). Text Preprocessing and Enrichment of Large Text Corpus-Based Keyphrase Generation for Goal-Oriented Dialogue Systems. In: Shukla, P.K., Mittal, H., Engelbrecht, A. (eds) Computer Vision and Robotics. CVR 2023. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-99-4577-1_14