Machine Learning for Misinformation Containment: *A Candid Assessment of the State of the Art*

Dr. Vishnu S. Pendyala
San Jose State University

23rd NFIC Invited Talk

Synopsis— **Misinformation containment has been proven to be NP-hard more than a decade ago. It is undoubtedly a complex problem to solve and appropriately attracted plenty of attention from the research community. A wide variety of machine learning algorithms such as support vector machines and logistic regression, ensemble techniques like random forest and Adaboost, deep learning frameworks such as LSTM and GAN, language models like BOW / TF-IDF and BERT, and many more have been tried out in the attempts to solve the problem. In terms of feature engineering as well, no stone has been left unturned. Manual feature extraction, graph embeddings, and other approaches to representational learning have all been tried. Not just supervised and unsupervised learning, but various other types of learning such as few-shot learning, meta learning, transfer learning, self-supervised learning, semi-supervised learning, reinforcement learning, and active learning have been explored extensively for the problem. Despite the voluminous research literature purporting to solve the problem using machine learning methods, misinformation containment is largely unsolved and is in fact growing by the day. It is therefore pertinent to understand this huge disconnect between what is claimed in the literature and the actual reality. The talk provided insights into the current state-of-the-art solutions and analyzed why they are not helping enough. The talk presented some future directions that in the speaker's opinion hold the promise and explained why there is hope.**

Keywords— Deep Learning, Natural Language Processing, Self-supervised Learning

■ **Speaker** Dr. Vishnu S. Pendyala is a faculty member of the Department of Applied Data Science at San Jose State University and the chair of IEEE Computer Society, Silicon Valley Chapter. He has over two decades of experience with software industry leaders like Cisco and Synopsys in the Silicon Valley, USA. Dr. Pendyala served on the Board of Directors, Silicon Valley Engineering Council during 2018-2019. During his recent 3-year term as an ACM Distinguished speaker and before that as a researcher and industry expert, he gave numerous (50+) invited talks. He holds MBA in Finance and PhD, MS, and BE degrees in Computer Engineering from US and Indian universities. Dr. Pendyala taught a one-week course sponsored by the Ministry of Human Resource Development (MHRD), Government of India, under the GIAN program in 2017 to Computer Science faculty from all over the country and delivered the keynote in a similar program sponsored by AICTE, Government of India in 2022. Dr. Pendyala's book, "Veracity of Big Data: Machine Learning and Other Approaches to Verifying Truthfulness" is indexed in several libraries, including those of MIT, Stanford, CMU, and internationally.