A quantum-inspired framework for video sentiment analysis

10 September 2020, by Ingrid Fadelli

Textual Modality	Visual Modality	Acoustic Modality	Sentiment
	😐	Neutral Voice	Neutral
The pizza is OK	🙂	Passionate Voice	Positive
	🙁	Disappointing Voice	Negative

Example of a simple sentiment analysis task. Credit: Li et al.

Automatically identifying the overall sentiment expressed in a video or text could be useful for a wide range of applications. For instance, it could help companies or political parties to screen large amounts of online content and gain insight on what the public thinks about their products, services, campaigns or initiatives.

Researchers at University of Padua, the Open University and University of Copenhagen have recently introduced a new framework for video sentiment analysis that is based on quantum physics theory. This framework, presented in a paper published in Elsevier's *Information Fusion* journal, fuses different types of video features and analyzes the complex interactions between them.

The study carried out by this team of researchers is part of the [QUARTZ (Quantum Information Access and Retrieval Theory) project](#), a broader research endeavor funded by the European Union's Marie Sk?odowska Curie Actions program. The key aim of the QUARTZ project is to explore the potential use of quantum mechanics for designing effective information retrieval systems, such as web search engines.

"As part of the QUARTZ project, we have been addressing many problems ranging from purely theoretical issues to experimental investigations," Qiuchi Li, one of the researchers who carried out the study, told TechXplore. "Our recent study is rooted in both these areas, as it considers the foundational issues of representing multimodal information and tests the effectiveness of the proposed methods by means of scientific experimentation."

Over the past decade or so, quantum theoretical frameworks have been found to have a number of advantages over classical probability theories for explaining some aspects of human cognition. Despite these advantages, researchers have so far only developed a few effective quantum-inspired models that can achieve [human performance](#) on cognitive tasks.

In their study, Li and his colleagues set out to develop a quantum-inspired model that performs well on sentiment analysis tasks, which are typically completed by humans. Their overall objective was to determine whether such a framework could attain human-comparable performance.

"We previously built models for text-based sentiment analysis, but this paper focuses on sentiment analysis in a multimodal context, and a quantum-driven approach is proposed to tackle the crucial task of multimodal fusion," Li explained. "Just like other multimodal sentiment analysis systems, the proposed model accepts multimodal features of an utterance and predicts its sentiment as the output."

The framework developed by Li and his colleagues analyzes videos based on a number of features, including textual, visual and acoustical elements. In contrast with conventional state-of-the-art models that rely on [artificial neural networks](#) (ANNs), the new framework analyzes video content from a perspective that is aligned with quantum physics theory.
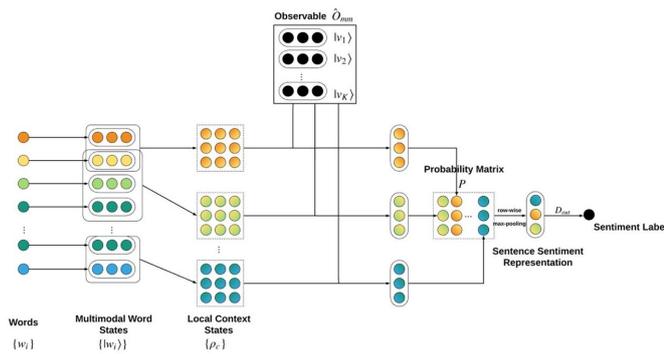
Figure outlining how the framework developed by the researchers works. Credit: Li et al.

The model views words as pure states in a composite system, maps utterances onto a quantum mixture of the words of which it is composed, and makes predictions about the sentiment expressed in an utterance via a quantum measurement process. The maps generated by the model are then used to build a complex-valued [neural network](#) that can produce multimodal representations of both written words and utterances. These representations are then used to carry out sentiment analyzes on videos.

"To facilitate such a quantum view, we designed a neural network that operates like a classical neural network, but with complex-valued, numerically constrained layers to faithfully implement quantum concepts and processes," Li said. "The complex-valued neural network performs comparably to state-of-the-art models on two benchmarking datasets, but its quantum view brings about a unique superiority in understanding the contributions to the predicted sentiment."

Li and his colleagues evaluated their framework on two datasets and found that it achieved highly promising results. While its performance and accuracy were similar to those achieved by more conventional neural network-based models, it presented unique qualities that set it apart from frameworks introduced in the past.

In fact, the quantum-inspired model uses a theoretical concept known as reduced density matrix. This element allowed it to make predictions about part of the data (i.e., features related to only one or two sensory modalities) without continuous retraining, which is very difficult to achieve using almost all conventional neural network-based models.

This unique characteristic allows the model to determine with high precision how much textual and visual features, or combinations of the two, contribute to the final sentiment expressed in a video. Classical ANN-based models, on the other hand, often need to interrupt their analyses when focusing on different modalities or modality combinations, as their structure needs to be retrained to map the intermediate representations onto the output label. Retraining models in this way, however, can create biases that affect their real performance due to the inherent randomness in the training process.

"We believe that this work has far-reaching implications for the applications of quantum theoretical frameworks," Li said. "We have demonstrated that quantum frameworks are capable of handling multimodal data equally well as the state-of-the-art classical models, and further identified its unique advantages in understanding the contributions to the sentiment judgment."

The study could inspire additional investigations into the potential of quantum-theoretical frameworks for multimodal [sentiment](#) analysis and natural language processing (NLP). Moreover, the quantum-inspired [model](#) they developed could eventually be used to analyze videos and automatically determine their overall affective qualities.

"The main focus of our next studies will be to further increase the performance of quantum-inspired models so that they could beat classical deep-learning models," Li said. "To achieve this, we plan to follow the most successful paradigm of pretraining in NLP and design quantum-inspired pre-training models for textual and multimodal data. In the meantime, we also plan to explore substitutions of advanced neural network components with quantum-inspired elements, such as the recurrent

layers and attention component. More specifically, we plan to seek help from the concept of quantum evolution to construct a quantum-theoretic version of these components and integrate them into the framework of this paper, for finer-grained modeling and better performance."

APA citation: A quantum-inspired framework for video sentiment analysis (2020, September 10) retrieved 28 September 2021 from https://techxplore.com/news/2020-09-quantum-inspired-framework-video-sentiment-analysis.html