

Exploring Quantum Natural Language Processing in Risk Analysis: A Comparative Study on Aviation Accident Reports

July Macedo^a, Plínio Ramos^{a*}, Caio Souto Maior^b, Márcio Moura^a, Isis Lins^a

^aUniversidade Federal de Pernambuco, Recife, Brazil

^bUniversidade Federal de Pernambuco - CAA, Caruaru, Brazil

ABSTRACT

Quantum computing, leveraging properties like superposition and entanglement, has transformative potential across diverse domains. However, Quantum Natural Language Processing (QNLP) applications in Risk Analysis (RA), particularly in aviation, remain understudied. This research bridges this gap by comparing QNLP models to classify aviation accidents, focusing on human error. Using Lambeq's pipeline to convert sentences into quantum circuits, we analyze accident narratives from the National Transportation Safety Board (NTSB) database. Our study tests three models: two quantum and one hybrid. The hybrid model, achieving 80.48% accuracy, outperforms purely quantum models and a classical TF-IDF NLP model in nuanced semantic relationships. This research highlights QNLP's potential to enhance accuracy and efficiency in accident classification, with future work aimed at validating these models on real quantum hardware.

KEYWORDS. Quantum Computing. Natural Language Processing. Aviation Accidents.

Paper topics (indicate in order of PRIORITY the paper topic(s))

1. Introduction

RA is a critical concern across various industries, it encompasses the identification of hazards, cause-and-effect analysis, and risk assessment, enabling experts to manage risks proactively and prevent potential accidents [Macêdo, Moura, Aichele, et al. 2022; Macêdo, Moura, Ramos, et al. 2022; P. M. Ramos et al. 2022; P. M. S. Ramos et al. 2022]. Specifically, the aviation industry is inherently exposed to significant risks, where accidents frequently result in significant financial losses and loss of life [Faulhaber et al. 2022]. Despite a declining trend in the number of annual accidents, incidents continue to occur and approximately 80% of all aviation accidents are primarily caused by human errors [Parnell et al. 2021].

Furthermore, the aviation industry acquires data from diverse sources and in multiple formats, including voice recordings from air traffic control and written accident investigation reports [Miyamoto et al. 2022]. An accident investigation aims to uncover and report the underlying causes of a particular accident. Regulatory agencies mandate companies to maintain a compilation of accident reports to enable safety specialists to undertake appropriate measures to mitigate or eliminate identified root causes following their analysis [Liu and Yang, 2023]. These reports are generally composed in natural language, and a comprehensive human review of the entire compilation of reports is nearly impractical given the sheer volume of reports produced by a company.

Quantum computing has garnered increasing attention due to its transformative potential to augment computational power [Stoos et al. 2023]. This technology harnesses quantum mechanics properties, such as superposition and entanglement, to manipulate and represent data. Researchers have leveraged quantum hardware and simulators, such as those released by IBM, to explore quantum applications across diverse domains [Bernal et al., 2022]. Despite the increasing use of quantum approaches in various fields, Quantum Natural Language Processing (QNLP) applications in Risk Analysis (RA) remain underexplored, particularly in the aviation industry, where risk analysis is critical.

Thus, our study endeavors to bridge the lack. By integrating QNLP with operational research methods, this study aims to improve RA and safety management in industrial settings. The ability to process and analyze large volumes of unstructured data from accident reports can lead to more accurate identification of risk factors and the development of effective preventive measures. Specifically, this research involves a comparative analysis of QNLP models to classify aviation accidents through binary classification, determining the presence or absence of human error as a causal factor. The proposed model is tested using a comprehensive public database of accident investigation reports conducted by the NTSB from 1982 to 2022 [NTSB 2022]. This advancement not only contributes to the academic field of industrial engineering but also offers practical implications for enhancing operational efficiency and safety protocols in high-risk industries such as aviation.

2. Methodology

Our approach involves adopting Lambeq's pipeline [Kartsaklis et al. 2021] to convert sentences into quantum circuits. This facilitates the development of a QNLP model tailored for determining causal factors behind aviation accidents, focusing on human error. The public database NTSB contains over 20,000 documents organized into spreadsheets, offering detailed information on accidents, including occurrence dates, injury severity, and distinctions between crew and passengers [NTSB 2022].

For this analysis, we randomly selected 200 narratives and employed OpenAI Summarization to condense each narrative to 20 words, addressing hardware and qubit limitations while retaining critical insights. The reports were randomly divided into training and test datasets, with 80% used for training (20% of which was set aside for validation) and 20% for testing.

The first quantum model configuration included TketModel as the Lambeq model, AerBackend, SPSSAOptimizer, and Quantum Trainer, using the IQPAnsatz type. The second quantum model used NumpyModel, Quantum Trainer, RotoSolveOptimizer, and StronglyEntangled Ansatz, with 24 circuits for training and 25 epochs. The hybrid model combined

the IQP-type Ansatz, Adam-type optimizer, and a custom Lambeq Model inheriting from the PennyLaneModel class.

3. Results and Discussion

To compare the results of the quantum models, we trained a classical model based on string diagrams, achieving 73% accuracy in around one hour. The first quantum model showed 63% accuracy, taking over 8 hours to run, indicating high computational cost. The second quantum model, with RotoSolveOptimizer, took over 12 hours to train on 24 instances, achieving 52.5% accuracy, demonstrating poor cost-performance balance.

The hybrid model achieved 80.48% accuracy in just over one hour, surpassing both quantum and classical models in computational efficiency. While the hybrid model shows promising results, challenges remain in developing scalable quantum computing systems. A comparison with a traditional TF-IDF NLP model, which attained 93.12% accuracy, revealed that the hybrid QNLP model excels in discerning nuanced semantic relationships, such as distinguishing between communication failure and pilot error.

4. Conclusion

This study highlights the transformative potential of QNLP in risk analysis, particularly in aviation. While all purely quantum model analyses were conducted using simulation tools, future implementations on quantum hardware are anticipated to offer faster and more accurate results. QNLP offers unique benefits, including semantic extraction while preserving grammar and mitigating biases found in large language models. Moving forward, our research will focus on enhancing model robustness, validating performance against human expertise, and testing results on real quantum hardware, aiming to advance QNLP's application in risk analysis and enhance safety measures.

Acknowledgements

The authors thank the following Brazilian research funding agencies for the financial support: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), 310892/2022-8, 402761/2023-5; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Finance Code 001. We also thank the Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE), APQ-0394-3.08/22 and Human Resources Program 38.1 (PRH 38.1) entitled Risk Analysis and Environmental Modeling in the Exploration, Development and Production of Oil and Gas, financed by Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) and managed by Financiadora de Estudos e Projetos (FINEP), 044819.

References

- Bernal, D. E., Ajagekar, A., Harwood, S. M., Stober, S. T., Trenev, D., & You, F. (2022). Perspectives of quantum computing for chemical engineering. *AIChE Journal*, 68(6), 1–14. <https://doi.org/10.1002/aic.17651>
- Faulhaber, A. K., Friedrich, M., & Kapol, T. (2022). Absence of Pilot Monitoring Affects Scanning Behavior of Pilot Flying: Implications for the Design of Single-Pilot Cockpits. *Human Factors*, 64(2), 278–290. <https://doi.org/10.1177/0018720820939691>
- Kartsaklis, D., Fan, I., Yeung, R., Pearson, A., Lorenz, R., Toumi, A., de Felice, G., Meichanetzidis, K., Clark, S., & Coecke, B. (2021). lambeq: An Efficient High-Level Python Library for Quantum NLP. *ArXiv Preprint*, arXiv:2110, 1–19. <http://arxiv.org/abs/2110.04236>
- Liu, C., & Yang, S. (2023). A text mining-based approach for understanding Chinese railway incidents caused by electromagnetic interference. *Engineering Applications of Artificial Intelligence*, 117. <https://doi.org/10.1016/j.engappai.2022.105598>

Macêdo, J. B., Moura, M. das C., Aichele, D., & Lins, I. D. (2022). Identification of risk features using text mining and BERT-based models: Application to an oil refinery. *Process Safety and Environmental Protection*, 158, 382–399. <https://doi.org/10.1016/j.psep.2021.12.025>

Macêdo, J. B., Moura, M. das C., Ramos, M., Lins, I. D., & Zio, E. (2022). Machine learning-based models to prioritize scenarios in a Quantitative Risk Analysis: An application to an actual atmospheric distillation unit. *Journal of Loss Prevention in the Process Industries*, 77, 104797. <https://doi.org/10.1016/j.jlp.2022.104797>

Miyamoto, A., Bendarkar, M. V., & Mavris, D. N. (2022). Natural Language Processing of Aviation Safety Reports to Identify Inefficient Operational Patterns. *Aerospace*, 9(8). <https://doi.org/10.3390/aerospace9080450>

NTSB. (2022). Accident Data. Download the Census of US Civil Aviation Accidents. <https://data.nts.gov/avdata>

Parnell, K. J., Banks, V. A., Plant, K. L., Griffin, T. G. C., Beecroft, P., & Stanton, N. A. (2021). Predicting Design-Induced Error on the Flight Deck: An Aircraft Engine Oil Leak Scenario. *Human Factors*, 63(6), 938–955. <https://doi.org/10.1177/0018720819872900>

Ramos, P. M., Macedo, J. B., Maior, C. B., Moura, M. C., & Lins, I. D. (2022). Combining BERT with numerical variables to classify injury leave based on accident description. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 1748006X2211401. <https://doi.org/10.1177/1748006X221140194>

Ramos, P. M. S., Maior, C. B. S., Moura, M. C., & Lins, I. D. (2022). Automatic drowsiness detection for safety-critical operations using ensemble models and EEG signals. *Process Safety and Environmental Protection*, 164, 566–581. <https://doi.org/10.1016/j.psep.2022.06.039>

Stoos, V., Ulmke, M., & Govaers, F. (2023). Quantum Computing for Applications in Data Fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 59(2), 2002–2012. <https://doi.org/10.1109/TAES.2022.3212026>