Text classification by means of quantum natural language processing (qNLP)

by Kevin Shen, Shruti Pistolwala, Clarissa Anjani, Sebastian Soldner, Philipp Zent and Friederike Butt

Acknowledgment

First of all, we would like to thank our partners at Merck Thomas Ehmer, Harsha Gurulingappa and Insa Mohr for their constant support, advise and inspiration.

Thanks also goes to the QEL team for giving us the opportunity to get together in this group and work on this interesting project.

We would also like to thank Bob Coecke and his team at CQC for giving us the chance to build on their research and to get first insights into this fascinating field.

Contents

1 We are qTex

We are qTex and as part of our course, Quantum Entrepreneurship Laboratory (QEL), we have developed a proof of concept on the use of quantum computing to classify drug adverse effects (AEs) for improving overall patient outcomes.

QEL is taught at the Technical University Munich in a joint teaching format with PushQuantum Munich and TUM Venture Labs, as well as industry partners. The goal of QEL is to examine various possible applications of quantum computing in the industry through a practical hands-on approach, by creating realistic, and comprehensive business models in interdisciplinary teams. Our partner for this project is Merck, a healthcare/pharmaceuticals company, who we have collaborated with to explore a sub-domain of Quantum Natural Language Processing (qNLP).

After reviewing various potential practical applications of this technology, we decided to focus on quantum computing in the area of drug development to enhance or highlight AE events with higher accuracy and greater speed. As AE studies are often extremely complex and handle a large variety of factors, classic calculations of large datasets are often manual and do not capture the full breadth of a patient's makeup such as their family history and existing medication. As quantum technology is a field which has seen rapid interest and continued investment from industry partners in recent years, we believe that we are for the first time at the intersection of a point where we could realistically enhance AE detection with quantum technology.

2 Business model

2.1 Problem and Product Scope

Healthcare companies like Merck develop drugs to deal with the world's most challenging diseases. One major issue for these companies is to make sure that patients are safe, while still benefiting from an effective drug.

After drugs are approved and launched, cases have arisen where patients suffer from rare adverse and severe side effects that have gone undetected in clinical development and comprehensive testing. In phase 4 trials or pharmacovigilance, drugs are continuously observed to evaluate the long-term patient care and safety as part of post-market surveillance practices. Once adverse effects are detected, the necessary intervention can be taken to minimize the risks and thus improve patient outcomes as a whole.

Today, drug adverse events are often detected too late which often results in lawsuits against pharmaceutical companies. For example, Merck & Co had to pay 4.85 billion dollars in settlements related to their painkiller Vioxx which was accused for causing heavy cardiovascular effects [1]. Other companies that have been charged with notable lawsuits include Takeda and GlaxoSmithKline [2, 3]. A study by the FDA found that in only the United States, around 2.2 million patients are hospitalized annually due to serious adverse drug reactions [4]. In addition, 100,000 deaths annually have been associated to adverse drug events [4].

Based on the aforementioned problem, we have determined the scope of our qTex analytics platform at this point in time to be the creation of a proof of concept to do the following:
1) Assess the commercial feasibility of using quantum computing for post-market drug surveillance
2) Lay groundwork for using quantum computing to classify adverse effects and not-adverse effects

2.2 Value Proposition & Key Benefits

Our unique selling point is the classification technology that uses quantum natural language processing. Three key benefits for pharmaceutical companies and their patients are as follows:
1) More peace of mind
2) More patient security
3) More hidden insights in drug adverse events

2.3 Target Customer Groups

The dichotomy between our target customer and our end user is important to understand for the context of marketing channels, customer feedback segments, and financial approach. Therein, we make distinct the two segments targeted below.

1) Paying Customer

Our paying customer are phase IV clinical trial providers and pharmacovigilance departments in the healthcare space. To solidify a core niche, we are particularly interested in solving customer problems related to classification of AE drug reactions. Our classification algorithms and web interface support patient safety analytics departments gain clearer insights on drug AE's, detect potential adverse events faster, discover adverse events that would have otherwise been hidden or have taken much longer to discover, as well as discover unintended cross-drug effects from users taking multiple drugs at once. To this end, we will be negotiating with large enterprises with seemingly larger R&D or analytics budgets. The pricing can be approached from a bulk or long-term perspective. In addition, pricing can be negotiated independently of individual trials.

2) End User

Our end users are the data analysts working for a given phase IV clinical trial provider's analytics or pharmacovigilance department. These users will often interface directly with our web analytics dashboard, as well as the classification input and output files. While our marketing and financial negotiations operate with a firm's overseeing department head or R&D budget personnel, our user experience and feedback must come from these respective end users. Additionally, we must consider that this data is pivotal to the success of [a] trial[s], and will be presented to upper level executives. This means that we should also come from the perspective that our end user is upper management of a given firm in some respects. Additionally, if data is published or released in a certain event, the end user can also be a federal agency or independent auditor interested in drug safety, and we should model our platform to anticipate these contingencies.

2.4 Market Size

Our market size(TAM/SAM/SOM), and derivation thereof, is key to understanding our scaling potential and competitive landscape. We break down our research below.

1) Total Available Market (TAM)

We consider our total addressable market for our initial product to be the pharmacovigilance market. We focus on the North American market for our business proposal as a proof of concept, and will move to other regions given the initial feedback from this rollout. The North American pharmacovigilance market is valued at $1.62 Billion USD in 2020 [5].

2) Serviceable Available Market (SAM)

Our goal is to break into the drug safety software market (a component of the pharmacovigilance market size), which is currently valued at $163.10 million USD [6]. This is a domain that is quickly growing, leveraging the power of big data and machine learning, where we hope to capitalize on the emergence of quantum computing to use our NLP technology to gain a foothold in this market.

3) Serviceable Obtainable Market (SOM)

We expect to capture a 5% market once our company has fostered its position in this market. This would give us an annual revenue stream of more than $8 million USD annually. The expected growth of this market is considerable - estimated to reach an annual growth rate of 7-8% with a market size of $220 million USD in 2025.

2.5 Financials

Revenue Streams

Our revenue streams can be broken down into two fixed components.

1) Monthly Pricing Our product will be our classification and web data analytics platform and we model prices accordingly. We have a three month free trial, and then we price per trial negotiated with individual firms depending on factors such as annual revenue, trial size and trial quantity. This consistent revenue will be needed for cash flow liquidity, server setup, research & development, and web platform enhancements. We consider the monthly pricing scheme to be the majority of our income, estimated to be 85-90% of our revenues. Our clients will negotiate a contract at the beginning of our partnership, and will be able to scale their monthly pricing every 6 months beginning after the trial period ends at a review period where we consider the aforementioned factors. The more trials a firm has, the more marginal savings they will benefit from, which depends on the contract with an individual firm.

2) Customer Benefit Based Pricing To offset the costs of our monthly pricing to consumers, we additionally propose customer benefit based pricing. This means that when finding statistically significant adverse event reactions, we generate a fee based on the savings generated from the avoidance of lawsuits/settlements/recalls. This fee is based off of historical settlements and other industry benchmarks. These events are rare, and we estimate them to be 10-15% of our revenue stream.

Pricing Strategy

As stated in the revenue streams section, we plan to use a freemium approach to entice potential customers to try out our product with the hopes of a longer term contract post trial period. Users will enjoy unlimited features, and can grow a preliminary understanding of their trials and needs. Once a client ends the trial period, we will negotiate an individual contract depending on business factors. We consider our pricing strategy to be low-cost, as we will price competitively due to aggressive price points and the customer benefit based pricing inclusion. Pricing will be equally favorable for small and larger clients (<$5M USD and >$500M USD annual revenue) due to this dynamically scaling pricing, as well as trial quantity marginal benefit based pricing.

Unit Economics

As with any software product, our unit economics scale exponentially in our favor with an nth increase in customers. As a SaaS platform, we consider our unit to be the "Software User". If we consider the R&D for a fully functional qTex product to be $1.2 million USD (based on salaries, research and development, backend infrastructure, etc.), and an estimate of acquiring 8 clients in our first year, we have a customer acquisition cost of $150,000 USD. As trials can last for years, we anticipate our lifetime value (LTV) of a customer to be 2.2 million USD. This means our unit profitability is $2.05M USD, scaling with each additional client acquired. Since our first year of estimated revenue at $8 million USD, we will only acquire a percentage of this LTV, given that clients will usually stay multiple years to realize this LTV.

Consideration of Profitability

The nature of our business is predicated on the ability of leveraging quantum technology. If quantum aspects of our business plan are unable to be implemented in a relevant timeframe due to hardware inaccessibility, or if quantum advantages are not as beneficial as to super-cede exciting classical implementations of AE detection, our profitability will be at risk or diminished. This means that extra care, time, and investment will be put into understanding the exact means to implement ZX calculus for our specific use cases on quantum hardware to achieve our desired business goals. Also, if the classification can not compete with future classical methods in terms of complexity of sentences or scale of output, additional risks to profitability can arise. The pivotal moment arises in the initial stages after seed capital investment, where we truly explore the potential of our technology. Once this becomes more clear, Series A funding and onwards can have serious investment to realize a top-class web interface and company infrastructure. Once we can achieve monthly cash flows through client acquisition, we will achieve cash flow liquidity and our profitability will be in good standing.

2.6 Go-To-Market Strategy

Due to the tech-intensity of our business, we will invest the vast majority of our resources and energy on research and development of our qNLP models. As the development of this deep tech requires a significant amount of time, we are not planning to enter the market before 2023. We will acquire first customers via multiple channels. An obvious and easiest channel for our business will be to utilize the partnership with our industry partner Merck. They have their own trials for which our product will be of use. In addition to our good relationship with Merck, we aim to continuously expand our network within

the pharma and healthcare industry until market entry. We expect a high credibility when approaching potential customers based on our ongoing partnership with Merck and the founders' association to excellent universities (TUM, RWTH) and institutes (TUM Venture Labs, PushQuantum). The planned late market entry in 2023 enables us to not only grow the aforementioned network, but also to thoroughly prepare and carry out marketing operations. On the one hand we will approach every trial provider within our network. On the other hand we will also perform many cold outreaches to trial providers outside our network and thereby ensure to have a solid amount of customers once our product is ready to be used. Besides relying on the described classical outreaches, we identified several pharma trade fairs as a viable opportunity to present our company on a large scale. We expect to generate a large amount of additional leads by investing into booths and pitching at trade fairs, such as PharmaTech Expo, Pharmac India Trade Fair, GPCE, BioProcess International, Expopharm, Pharma Asia, Chemspec Eruope, Austropharm and many more. Once we have gained a solid customer basis, we will shift part of our focus on retaining those customers. However we will be constantly thriving to extend our market share by continuously performing the described customer acquisition activities. We seek to extend our business beyond the pharma and healthcare space in the long run through continued research and product optimization, which will eventually yield possible applications in other industries.

2.7 Competitor Analysis

We could not identify a direct competitor working on qNLP use cases within the pharma and healthcare industry. Cambridge Quantum Computing (CQC) is leading in the research field of qNLP, but they are not targeting the pharma and healthcare market, as well as not purely focusing on qNLP.

When examining service providers for phase IV trials in more detail, it was found that many companies offer add-on services to improve the quality of phase IV trials. One company with the most similarity is Semalytix who uses classical NLP and suggests a similar value proposition. As they already have experience in analyzing phase IV trial data by using NLP, we identify Semalytix as our main competitor. The main differentiator is our focus on qNLP as opposed to the Semalytix' use of classical approaches. As NLP is a quantum native problem, we aim to achieve a competitive advantage over Semalytix in terms of accuracy and speed-up once our technology is mature enough.

3 User Journey

One of the main end users that would benefit from the qTex analytics platform are patient safety departments in large pharmaceutical companies, as discussed in section 2.3. This section describes an exemplary user journey related to how the user is doing, thinking, and feeling as each stage is completed.

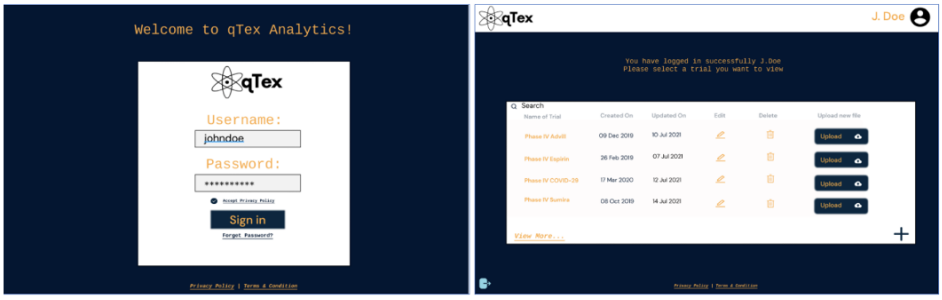
John Doe, Patient Safety Data Analyst

John Doe is a patient safety data analyst in a large pharmaceutical company. He is responsible for monitoring various post-market or phase 4 clinical trials. John Doe has heard of patients complaining about Espirin and would like to understand better which adverse effect should be a point of concern.

Expectations

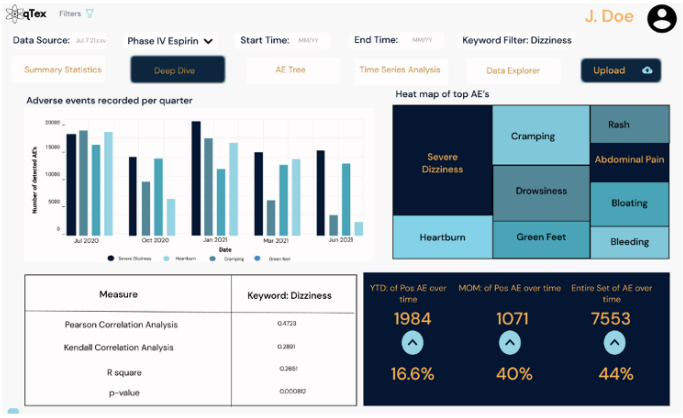
John Doe wants to

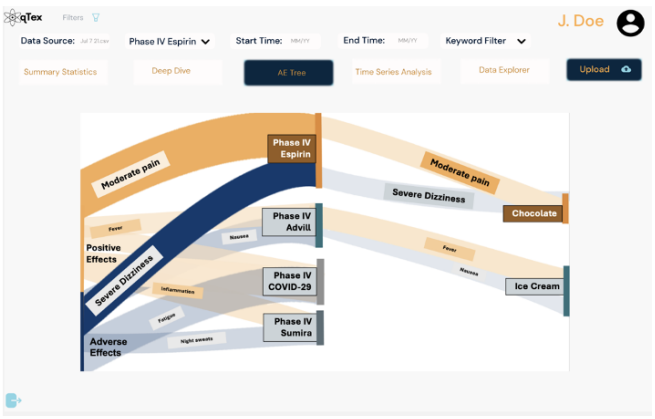
- get an overview of clinical trials and their adverse effects,
- get a better understanding of adverse events that need to be closely monitored and eventually communicated to affected parties,
- be able to review a dataset and check if it is sound evidence to draw a particular conclusions.

Stage of Journey	**1. Intend to look up for information on adverse effects related to Espirin**
Figma	
Doing	• Log in to qTex analytics • Click on the clinical trial Espirin in the page showing the overview of clinical trials
Thinking	• What have patients been complaining about when they take Espirin?
Feeling	• Curious about the complaints related to Espirin • Concerned with any unwanted harm that was inflicted on patients

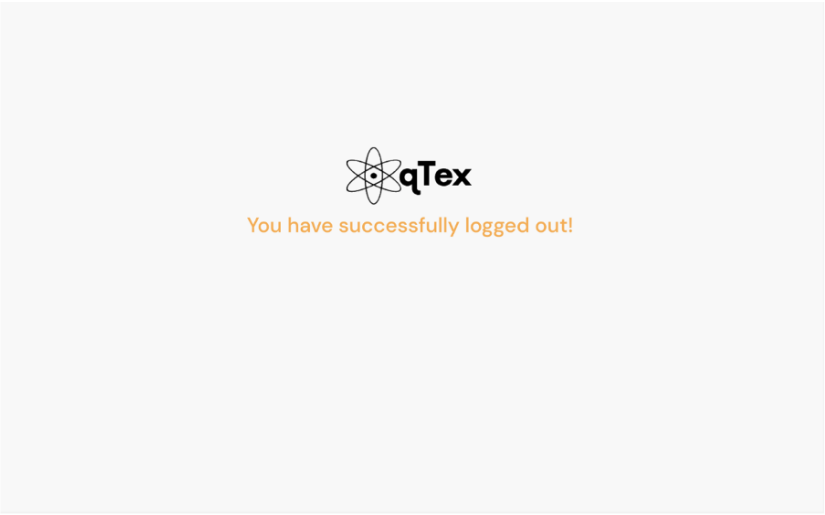
Stage of Journey	**2. Assess high-level visualizations on Espirin clinical trial**
Figma	
Doing	• Review the classifications of adverse events
Thinking	• What is the progress of the sentences that have been logged and classified throughout time? • How have the classifications of adverse events varied throughout time? • What is the proportion of classifications of adverse events out of all the total logged sentences? • I notice that severe dizziness has the highest number of hits
Feeling	• Relieved that there is a high-level overview of adverse event classifications • Interested in finding more details on how the adverse events varied over time

Stage of Journey	2. Assess high-level visualizations on Espirin clinical trial
Figma	
Doing	• Review the classifications of adverse events
Thinking	• What is the progress of the sentences that have been logged and classified throughout time? • How have the classifications of adverse events varied throughout time? • What is the proportion of classifications of adverse events out of all the total logged sentences? • I notice that severe dizziness has the highest number of hits
Feeling	• Relieved that there is a high-level overview of adverse event classifications • Interested in finding more details on how the adverse events varied over time

Stage of Journey	3. Review further information on adverse event classifications
Figma	
Doing	• Evaluate additional information related to adverse events
Thinking	• How do the top adverse event classifications compare against each other through time? • How statistically significant is the dataset that was used? • I see again that severe dizziness shows as being considerably higher than the other adverse events
Feeling	• How do the top adverse event classifications compare against each other through time? How statistically significant is the dataset that was used?

Stage of Journey	4. Compare Espirin with other clinical trial in AE Tree
Figma	
Doing	• Observe the adverse events that have been reported in the other clinical trials
Thinking	• Do other clinical trials have patients suffering from severe dizziness?
Feeling	• At ease that Espirin is still associated to pain relief and severe dizziness is not a complaint in the other clinical trials • Aware of other clinical trials and their related adverse events

Stage of Journey	5. Analyse time series data to see variation of adverse events across time
Figma	
Doing	• Find the pattern of adverse events in the time series analysis
Thinking	• How do the adverse event classifications vary throughout the data that has been logged?
Feeling	• Excited that there is a possibility to identify patterns monthly and potentially find explanations for it

Stage of Journey	7. Log out of the qTex analytics platform
Figma	
Doing	• Log out of the qTex analytics platform
Thinking	• I need to let the team know that there is proof that severe dizziness could be caused by taking Espirin • What should I do with knowledge that severe dizziness shows a considerable association to Espirin? • How should our company communicate the risks to patients?
Feeling	• Calmer that there is empirical evidence of the adverse events that could be caused by Espirin

For a more detailed view, we recommend to read the section 'User journey' on https://qtex.trudawnsolutions.com/.

4 Quantum Solution

4.1 Task

The goal is to perform a binary classification of sentences into the classes of adverse effects and not-adverse effects using quantum natural language processing (qNLP). A full treatment of the underlying theory and detailed descriptions of the individual steps are beyond the scope of this report. For further information we recommend to read two recent papers [7] and [8], which we used to build our model.

4.2 Quantum Natural Language Processing (qNLP)

The foundation of qNLP is diagrammatic reasoning. Using DISCOCAT and categorial quantum mechanics (CQM), language can be interpreted as a quantum process [7]. This quantum representation can be translated into a quantum circuit using the ZX-calculus. The resulting circuit model contains gates that depend on parameters β_i. We can train the model with classical optimization algorithms [7] and find a model that can perform the desired classification. The general pipeline is depicted in Fig. 1.

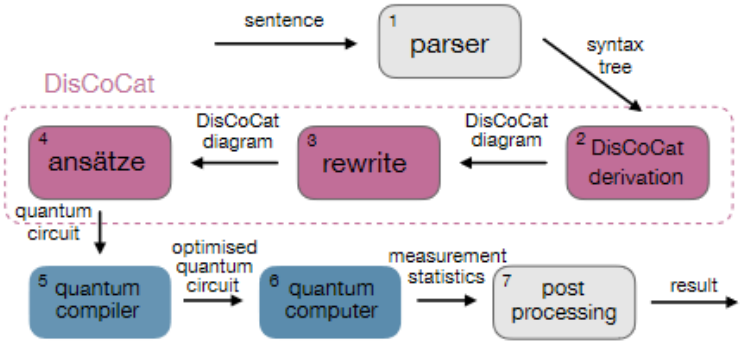


Figure 1: Schematic overview of the general pipeline for qNLP [8].

In our solution, we built on the approach depicted in figure 1, described in 'QNLP in Practice: Running Compositional Models of Meaning on a Quantum Computer' [8], and adapted specific steps within this pipeline. In the following, we briefly explain each step for the exemplary sentence 'Aspirin reduces severe headache' and outline which changes were done.

0) Pre-Parser

For the processing of sentences, we need to know which type of words have which dimension (also see next section). In a pre-parsing step, the words are tagged corresponding to their type of word, for example 'ADJ' for an adjective or 'N' for a noun, to make the dataset grammar independent.

1) Parser

The meaning of words is represented by tensors with their dimension given by the type of word. For example, nouns, adjectives and transitive verbs would have the types n, $n \cdot n^l$ and $n^r \cdot n \cdot n^l$, respectively. n, describes a noun or a noun phrase and l, r denote that an n is expected on the left of right (left/right adjoint) [8]. For a type s sentence, the sentence 'Aspirin reduces severe headache' is parsed as

$$n \cdot (n^r \cdot s \cdot n^l) \cdot (n \cdot n^l) \cdot n \longrightarrow (n \cdot n^r) \cdot s \cdot (n^l \cdot n) \cdot (n^l \cdot n) \longrightarrow 1 \cdot s \cdot 1 \cdot 1 \longrightarrow s. \qquad (1)$$

So, words are represented by states which are connected by grammar reductions [7], [9]. This approach goes back to the formalism of categorial grammar as a natural language syntax [10].

2) DisCoCat derivation

The DisCoCat algorithm takes the meaning of parts, establishes a wire diagram that represents grammar, and connects the meaning with this wire diagram. This results in the meaning of the whole [7]. So, each word is represented as a state and 'wired up' by drawing a cup for each grammar reduction. For the above example, this results in the DisCoCat diagram depicted in Fig. 2.

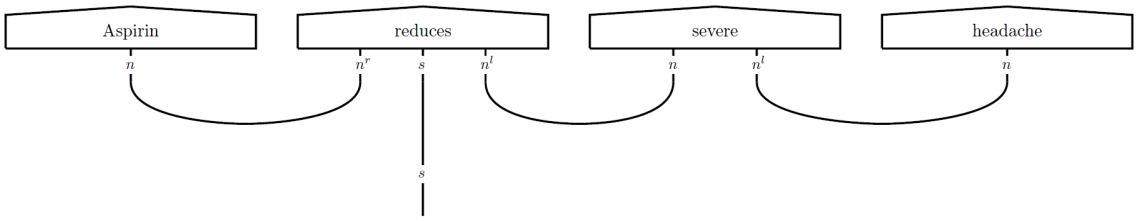


Figure 2: DISCOCAT diagram for 'Aspirin reduces severe headache'.

For more examples and to get to know this step in more detail, we refer to the 'DISCOCAT generator' [11].

3) Rewrite

The diagrams can be rewritten into equivalent diagrams which can be less costly. So far, there is no universal set of rules to rewrite diagrams into advantageous diagrams. In general, cups are costly so one effective transition is to 'bend the nouns around', i.e. replacing the cup and the noun's state $1 \longrightarrow n$ with an effect $n^{r/l} \longrightarrow 1$. For the above example, this would give a diagram as depicted in Fig. 3.

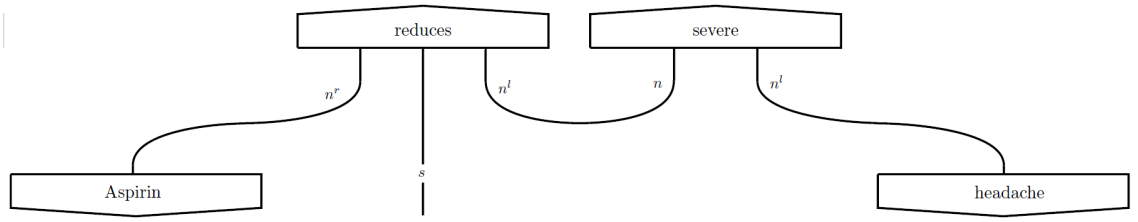


Figure 3: Equivalent DISCOCAT diagram after bending the nouns around to get a computational advantage.

4) Ansätze

The DISCOCAT diagram is mapped to a quantum circuit using a specific ansatz. The choice of ansatz on the one hand determines the number of parameters for the word representations, i.e. the number q_n and q_s for qubits representing wires of type n and s, respectively. Here, we want to keep the number of qubits as low as possible in light of the noise on NISQ devices, so we choose

$$q_n = 1 \quad \text{and} \quad q_s = 1. \tag{2}$$

On the other hand, we need to set concrete parameterized quantum states that replace the word states. We choose an Euler decomposition for effects (n type word states) like [8]

$$\langle w| = \langle 0|\, Rx(\theta_1)Rz(\theta_2)Rx(\theta_3). \tag{3}$$

So, $p_n = 3$ parameters describe each noun type word. In total, we choose

$$\{q_s, q_n, p_n\} = \{1, 1, 3\}, \tag{4}$$

where q_s and q_n are the number of qubits per type s- and n- type wire and p_n is the number of parameters for each effect. For the above example, we obtain a circuit as depicted in Fig. 4.

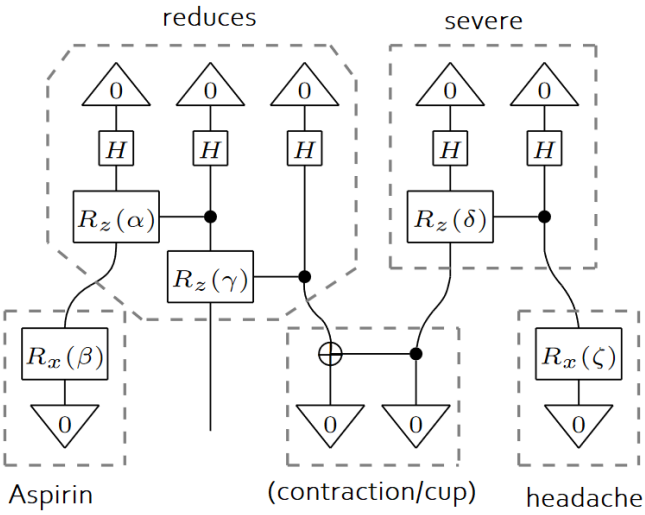


Figure 4: Possible circuit for 'Aspirin reduces severe headache' for chosen ansatz [8].

5) Quantum Compiler

The quantum compiler translates the quantum circuits into machine instructions.

6) Quantum Computer

The circuit is run n_{shot} times. In each run, the initial states are prepared, the gates are applied and all qubits are measured at the end, so we get out measurement statistics.

7) Post Processing

As the outcomes of the given circuits are only obtained with a certain probability, we have to postselect for a specific outcome. For example, for the circuit in Fig. 4, we need to postselect on the 0-outcome [8], [12].

So, we obtain the inputs for further post processing, as for example the optimization of the circuit-parameters.

Optimization

The ansatz described above, was shown to perform better than other approaches [8], so we took this ansatz and focused on varying the optimization process in order to increase accuracy further. The hyperparameters $q_s = 1$ and $p_n = 3$, give a total number of k parameters, depending on the vocabulary at hand. The goal is to find parameters that define the circuit in such a way that new sentences can be processed and classified accurately [13]. In other words, we want to train our model and find parameters that can reliably perform classifications. We define the objective function as the standard cross-entropy and minimize this cost [8]. For the minimzation of the cross-entropy, an adapted version of coordinate descent and parameter-shift rule [14] are used. All gate parameters are initialized at $\frac{\pi}{4}$. During each iteration, a random parameter θ_i is chosen and the circuit is evaluated twice at $\theta_i \pm \frac{\pi}{4}$, keeping the other parameters fixed. For each case, the circuit output is computed by averaging over N, usually >1024, shots. Then the results are compared to determine the descend direction. The parameter θ_i will then be updated accordingly with a step size decreasing with the number of iterations.

4.3 Results

Simulating the above described approach on a classical device for a set of 7 nouns, 5 verbs, 3 adjectives, amounting to 34 parameters, with 1024 shots for averaging, we obtain the following results.

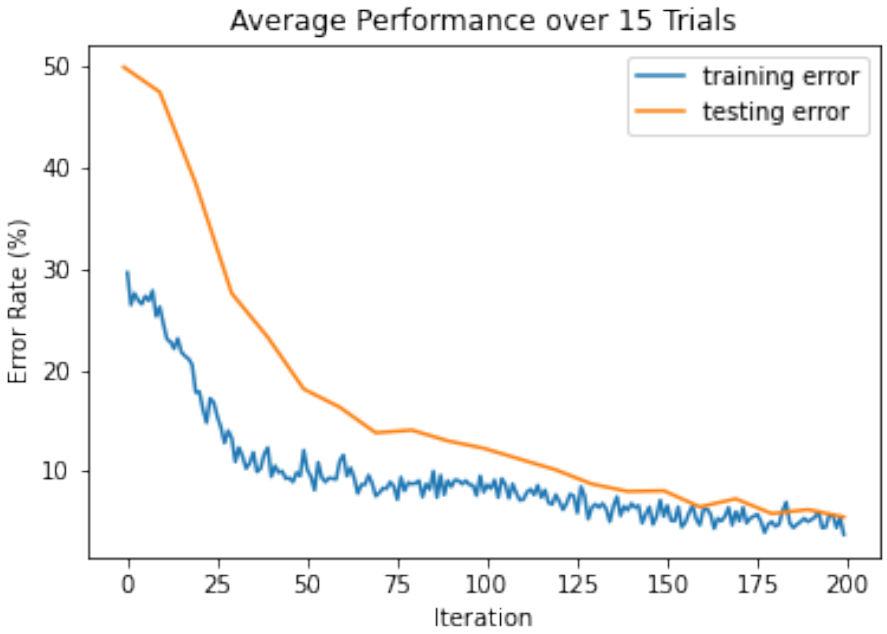


Figure 5: Resulting training and testing errors for averaging over 15 trials.

Due to the randomness in our optimization algorithm, the results vary over trials but converge well in most cases. We've repeated the experiment on the same data set 15 times and the averaged result is shown in Figure 5. With $17/18$ sentences being correctly labeled, the testing accuracy reaches 94.4% for 200 iterations. For more details, we refer to the GITHUB repository handed in together with this report.

4.4 Complexity

There are two general approaches to scale up the experiments. One ansatz would be to scale up the number of sentences. This results in a scale up of the time needed to get the statistics from a quantum computer for a sentence circuit, thus a prefactor on the time-cost [8]. Large-scale tasks could be tackled with parallelised ansätze, as is done in NLP problems. However, this is limited by the available hardware: long runtimes of single circuits, limited number of shots and submitted number of circuits per job and a limited number of available high-fidelity quantum processors lead to substantial queuing times. Another approach to scale up the experiment would be to extend the vocabulary. This would increase the dimension of the parameter space. The number of parameters and qubits depends on the number and type of words contained in the vocabulary as explained in Sec. 4.2.

#nouns	#verbs	#adjectives	k
9	5	3	40
20	5	3	73
100	5	3	313

Table 1: Number of parameters k for different vocabulary sets for the above chosen ansatz $\{q_s, q_n, p_n\} = \{1, 1, 3\}$, see Sec. 4.2.

4.5 Comparison with Classical Solutions

Classical solutions in the field of identifying adverse drug reactions take various different approaches ([15], [16], [17]) and do not yet exceed 95% accuracy for this kind of task. A detailed analysis or comparison of the complexity is difficult since the approaches are very different and this kind of comparison would fail to be meaningful. Therefore, adverse events detection is often done manually, since missing 5% of adverse events has a severe impact. Even though we worked with a smaller set of vocabulary and sentences, an accuracy of 94.4% was reached in our quantum approach.

The quality of hardware is constantly improving, promising scalability to large datasets in the future. On the other hand, our approach is practically independent of the specific dataset, so flexibly applicable to the problem at hand. In contrast to classical approaches, qNLP processes the meaning of words and lets these interact instead of merely 'counting statistics'. This promises more reliable predictions and extremely flexible applications. For example, it should be easy to adjust the model to different languages, since it is independent of the specific grammar.

5 Future plans

We provided a proof of concept with the above described solution and expect a further increase in accuracy of the results with continuing research. With additional resources and expected available high-fidelity quantum processors in the future, we plan to increase accuracy above $94\,\%$ within the next years and outperform available classical solutions.

References

[1] https://www.cbsnews.com/news/merck-to-fund-485b-vioxx-settlement/.

[2] https://www.drugwatch.com/news/2015/04/29/takeda-actos-settlement-2-billion/.

[3] https://www.justice.gov/opa/pr/glaxosmithkline-plead-guilty-and-pay-3-billion-resolve-fraud-allegations-and-failure-report.

[4] https://www.fda.gov/drugs/drug-interactions-labeling/preventable-adverse-drug-reactions-focus-drug-interactions.

[5] https://www.marketsandmarkets.com/Market-Reports/healthcare-lifesciences-nlp-market-131821021.html.

[6] https://www.mordorintelligence.com/industry-reports/pharmacovigilance-and-drug-safety-software-market.

[7] B. Coecke, G. de Felice, K. Meichanetzidis, and A. Toumi, *Foundations for near-term quantum natural language processing*, https://arxiv.org/abs/2012.03755.

[8] R. Lorenz, A. Pearson, K. Meichanetzidis, D. Kartsaklis, and B. Coecke, *Qnlp in practice: running compositional models of meaning on a quantum computer*, https://arxiv.org/abs/2102.12846.

[9] M. Baroni and R. Zamparelli, "Nouns are vectors, adjectives are matrices: representing adjective-noun constructions in semantic space",

[10] J. Lambek, "The mathematics of sentence structure", https://doi.org/10.1080/00029890.1958.11989160.

[11] https://qnlp.cambridgequantum.com/generate.html.

[12] S. Aaronson, "Quantum computing, postselection, and probabilistic polynomial-time", https://doi.org/10.1098/rspa.2005.1546.

[13] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models",

[14] G. E. Croocks, "Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition",

[15] T. Zhang, H. Lin, and et. al., "Identifying adverse drug reaction entities from social media with adversarial transfer learning model",

[16] A. Kormilitzin, N. Vaci, and et. al., "Med7: a transferable clinical natural language processing model for electronic health records",

[17] E. drissiya El allaly, M. Sarrouti, and et. al., "Mttlade: a multi-task transfer learning-based method for adverse drug events extraction",