Quail: Quantum Understanding and AI for Interpretability and Learning

Dr Jackie Tan, Team Quail

For submission to the
Future Leaders in Quantum (FLIQ) Hackathon

Ex-Associate Professor in Data Science (Nanyang Technological University), 3x serial entrepreneur

Research at Center of Quantum Technologies (National University of Singapore)

Broad areas of expertise = AI + web3

Broad domains = fintech + DS + edu

Agenda

- Problem
- Solution
- Approach
  - Classical approach
  - VQA approach
  - Dashboard
- Results
- Limitations
- Future Work
- Conclusion

Problem

Drug-induced allergies are life-threatening

- Drug-induced allergies (DIA) are hard to predict
- Active area of research
- A strong need for pharmacovigiliance

New research has emerged with useful data

Contents lists available at ScienceDirect

Toxicology

journal homepage: www.elsevier.com/locate/toxicol

InterDIA: Interpretable prediction of drug-induced autoimmunity through ensemble machine learning approaches

Lina Huang, Peineng Liu, Xiaojie Huang *

Department of Clinical Pharmacy, Jieyang People's Hospital 522000, China

New research has emerged with useful data

- Generated physicochemical properties of non-DIA and DIA-inducing drugs
- Separate train (477) + test (120) set

- 196 features, very high dimensional
- Imbalanced dataset, 25% DIA-inducing vs non-inducing

Solution

QML leverages expressivity of quantums states

QML models complex, high-order correlations in data:

- Feature map
- Ansatz
- Measurement + Optimization

QML might learn patterns that classical models might miss!

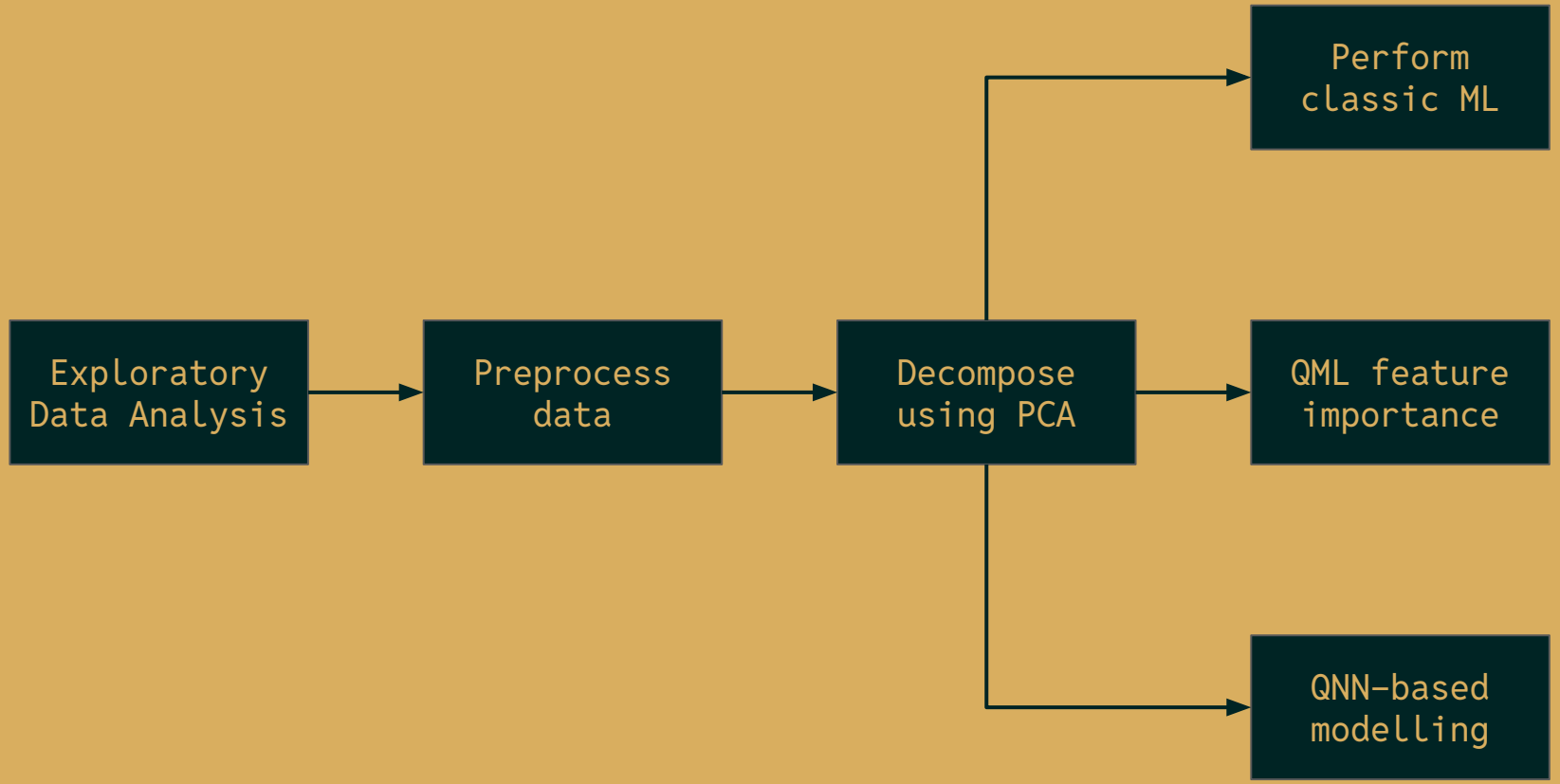
High-dim data -> PCA -> Feature map -> VQA -> Prediction

Approach

Project in Parts

- Part I: Exploratory Data Analysis
- Part II: Data Processing
- Part III: Benchmarking using Classic ML
- Part IV: Performing VQA to Predict DIAs
- Part V: Quantum Feature Importance

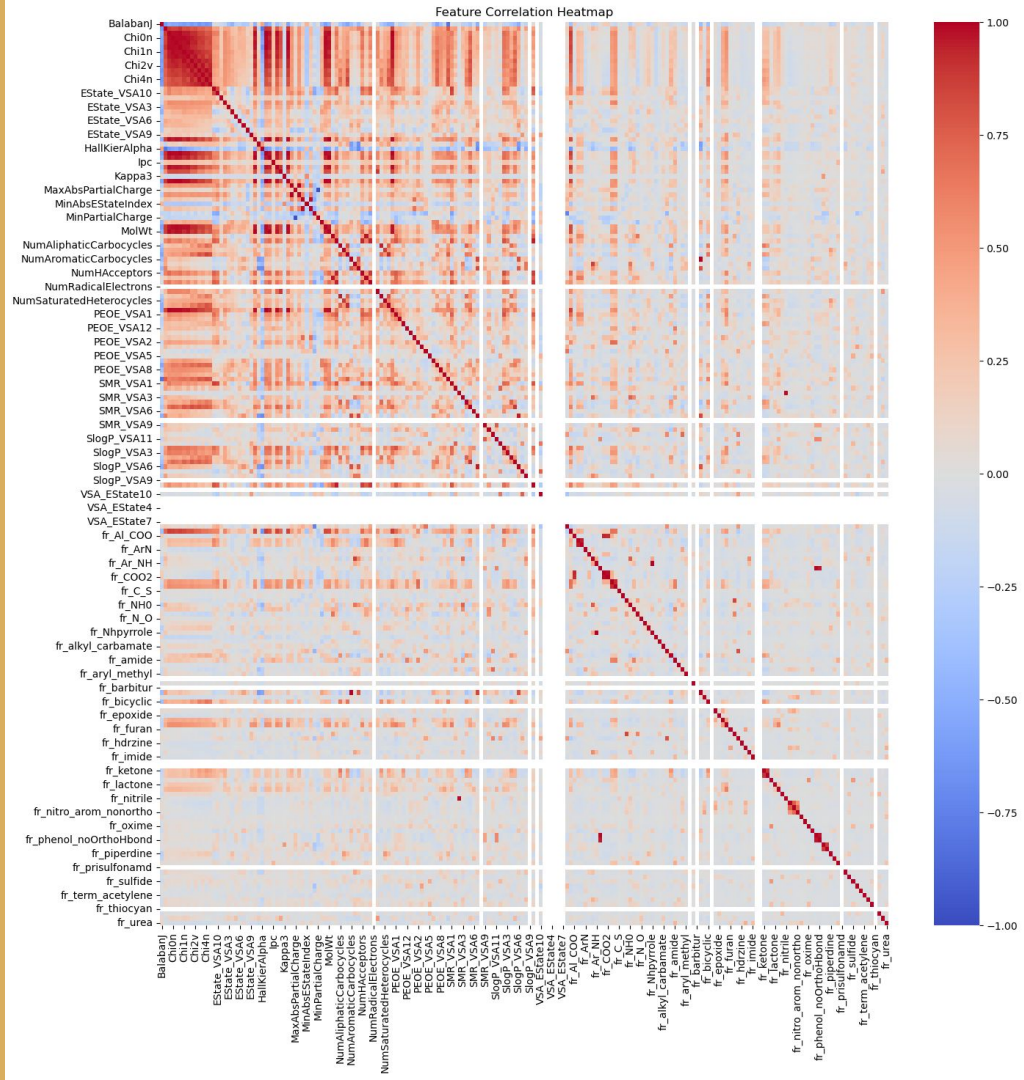
Workflow


Exploratory Data Analysis
Preprocess data
Decompose using PCA
Perform classic ML
QML feature importance
QNN-based modelling

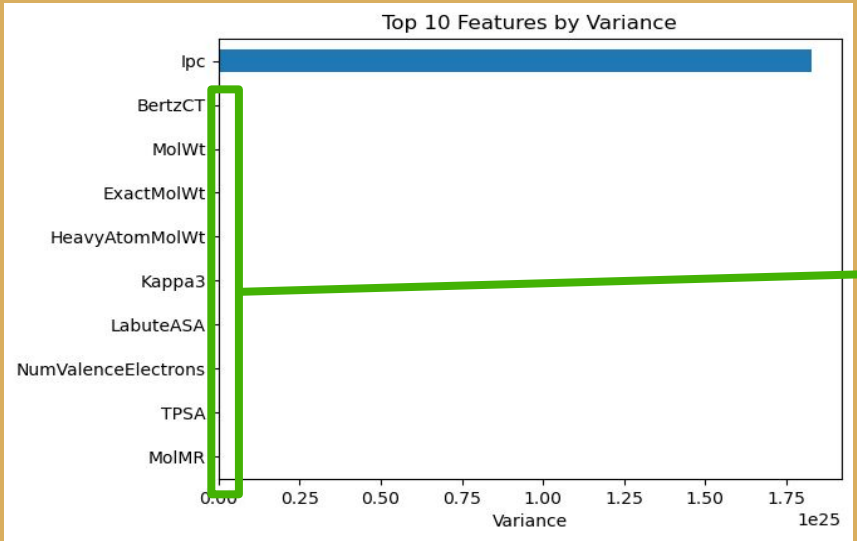
Results

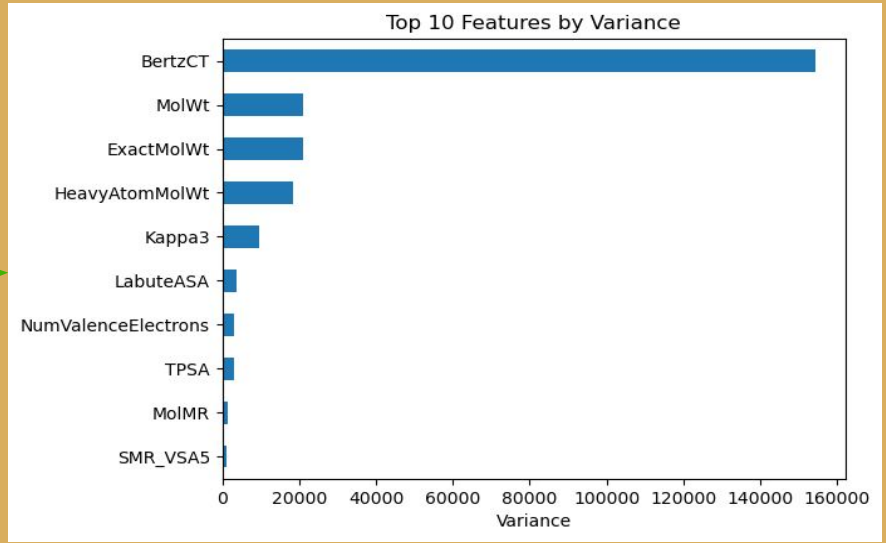
High number of correlated features

- Possibility of performing dimension reduction
- Dimension reduction = we can use a small number of qubits



Large variance in top 10 features necessitates scaling





QML performance is comparable to trad ML

	Traditional ML (Original)		
	Logistic regression	Random Forest	6-component PCA + RF
Accuracy	25%	81%	76%
F1-score	0.4	0.48	0.36
Recall	0.62	0.82	0.68
Precision	0.5	0.66	0.60

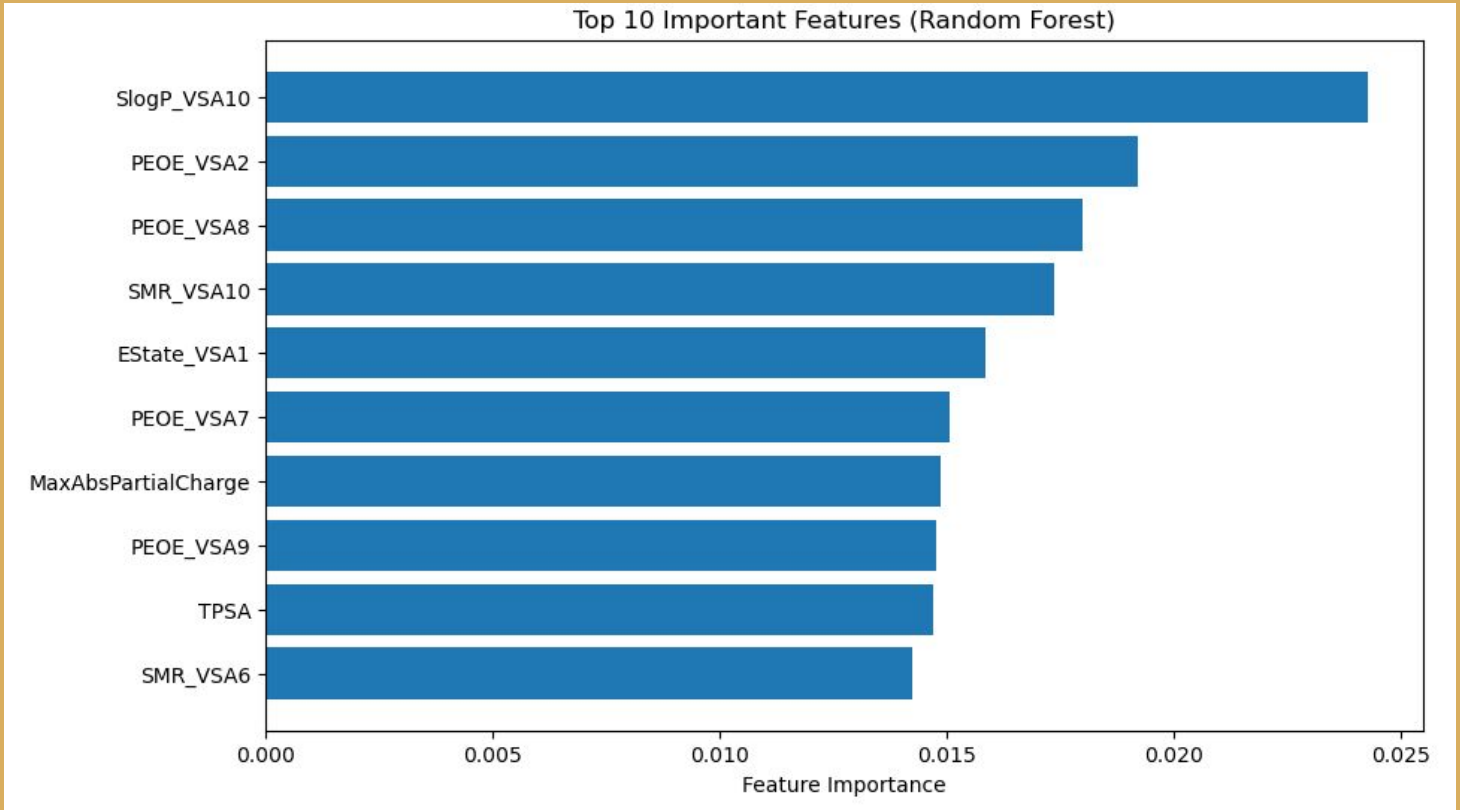
QML performance is comparable to trad ML

	Traditional ML (Scaled)			Quantum ML (Scaled, RealAmplitudes)		
	Logistic regression	Random Forest	6-component PCA + RF	2 qubits	4 qubits	6 qubits
Accuracy	74%	81%	82%	60%	50%	56%
F1-score	0.31	0.5	0.48	0.43	0.29	0.36
Recall	0.62	0.80	0.82	0.58	0.47	0.53
Precision	0.57	0.67	0.66	0.60	0.47	0.54

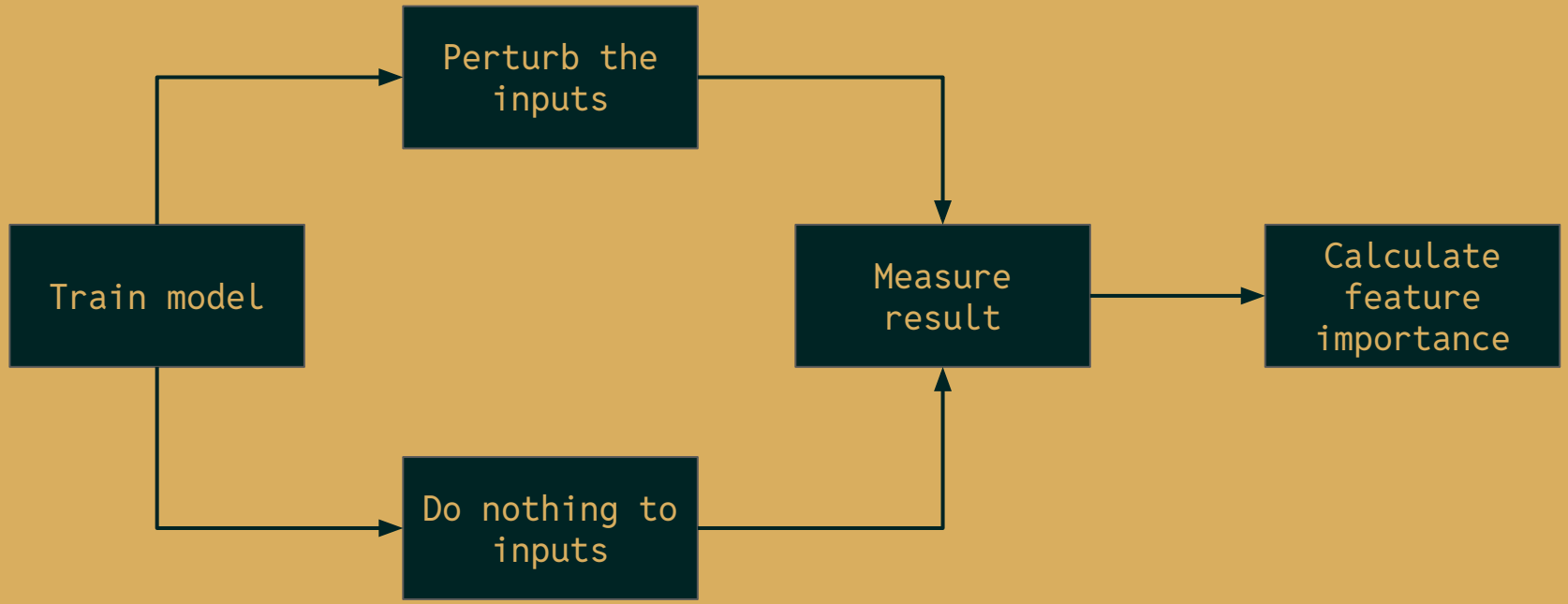
QML performance is comparable to trad ML

	Traditional ML (Scaled)			Quantum ML (Scaled, EfficientSU2)		
	Logistic regression	Random Forest	6-component PCA + RF	2 qubits	4 qubits	6 qubits
Accuracy	74%	81%	82%	53%	53%	54%
F1-score	0.31	0.5	0.48	0.38	0.30	0.38
Recall	0.62	0.80	0.82	0.53	0.49	0.54
Precision	0.57	0.67	0.66	0.54	0.48	0.55

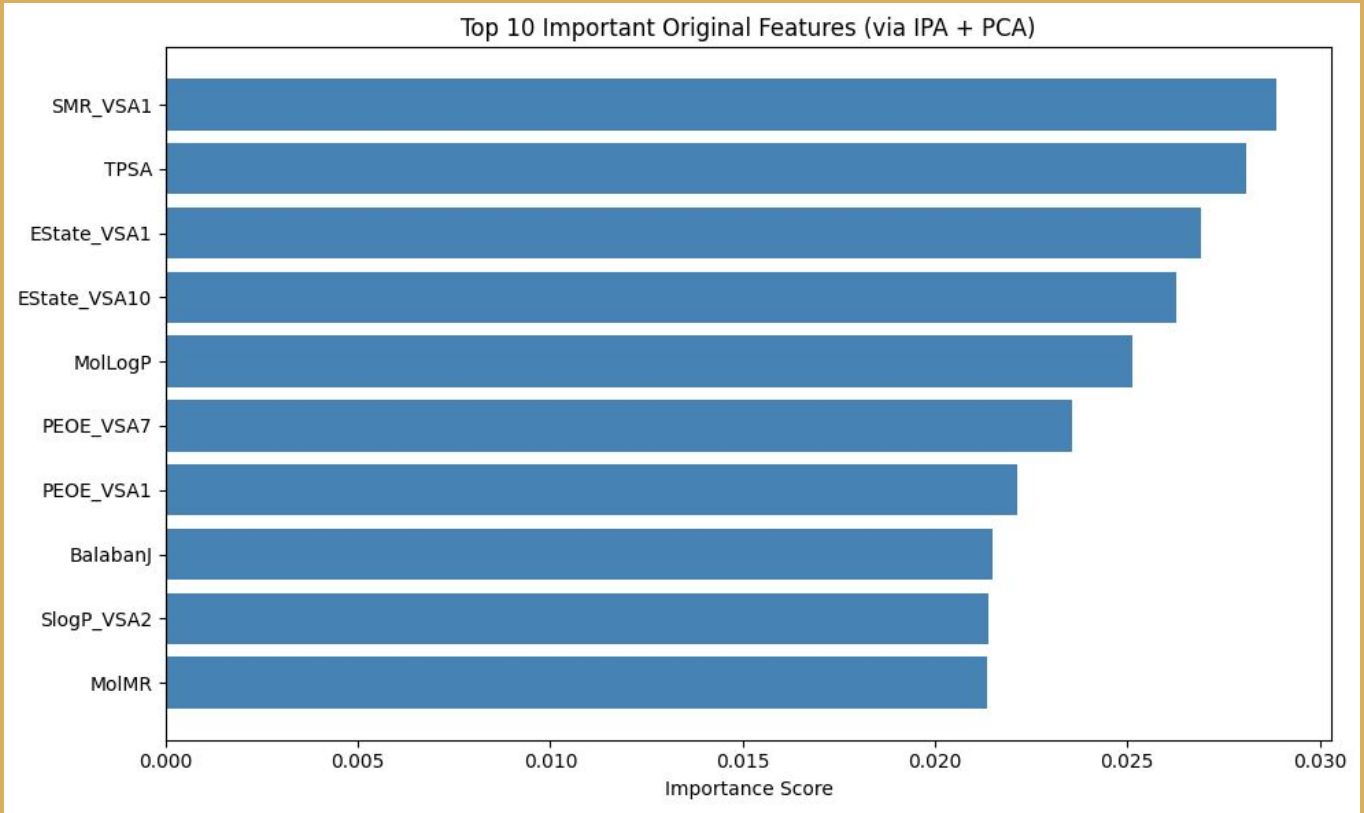
VQA models are explainable with input perturbations



VQA models are explainable with input perturbations



VQA models are explainable with input perturbations



Conclusion

VQAs hold immense potential for the future of ML

- Quail did a great job under time and manpower pressure
- Room for improvement for VQA performance
- VQAs are able to handle decompositions from large dimensions
- We used a novel perturbation-based method to do quantum feature importance
- Promising performance with additional qubits

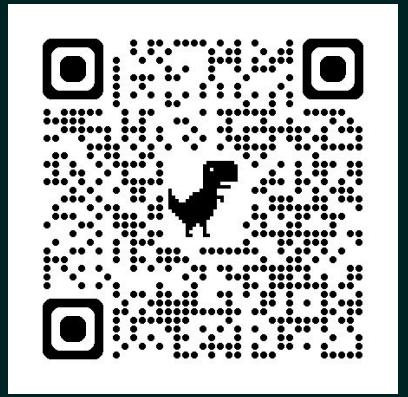
Let me win and send me to Geneva

- I’m a good scientist with great ideas
- I have the right mix of knowledge, expertise, and experience
- I’ll rep us well, you won’t regret it!

Thanks

@jackietanyen

LinkedIn



Github Repo