INGENII + DATA MINE

Academic year 2022-2023

Abstract

Quantum Machine Learning methods to accelerating and improving protein-ligand binding affinity predictions in drug discovery.

Table of Contents


ingenii
ACCELERATING DATA ENGINEERING
About
Ingenii

Mission

Our mission is to solve the engineering challenges around adopting quantum computing and advanced analytics within the conventional data science stack.

Through this, we hope to accelerate solutions to high-impact problems that classical computing cannot solve across life & environmental sciences

What do we do?

We provide engineering automation and tooling that enables businesses to

- ✓ Quickly deploy enterprise-level data infrastructure
- ✓ Automate data engineering
- ✓ Accelerate quantum development
- ✓ Access a marketplace of easy-to-use data and quantum engineering modules

ingenii
ACCELERATING DATA ENGINEERING

Who is Ingenii?

Ingenii is a team of technologists with a passion for solving data and quantum engineering challenges. We believe we can have an impact by applying these skills towards making the adoption of quantum computing as easy as possible.

By simplifying the data and quantum engineering, we enable businesses to get started with quantum and other advanced analytics faster and cost effectively, enabling them to solve real-world problems more creatively and quickly than ever before.

With raging pandemics, climate change, inefficient precision medicines and countless other social and environmental challenges, this is our contribution to the wide-scale solution.

ingenii
ACCELERATING DATA ENGINEERING

About
The Project

Background

The idea behind this project is an iteration on one of the previous Data Mine projects, which you can read about here.

Some more specific examples:

- The final PDF
- The presentation recording

The idea was further developed by discussing with LLNL/ATOM and focusing on their 3D-CNN machine learning from this project:

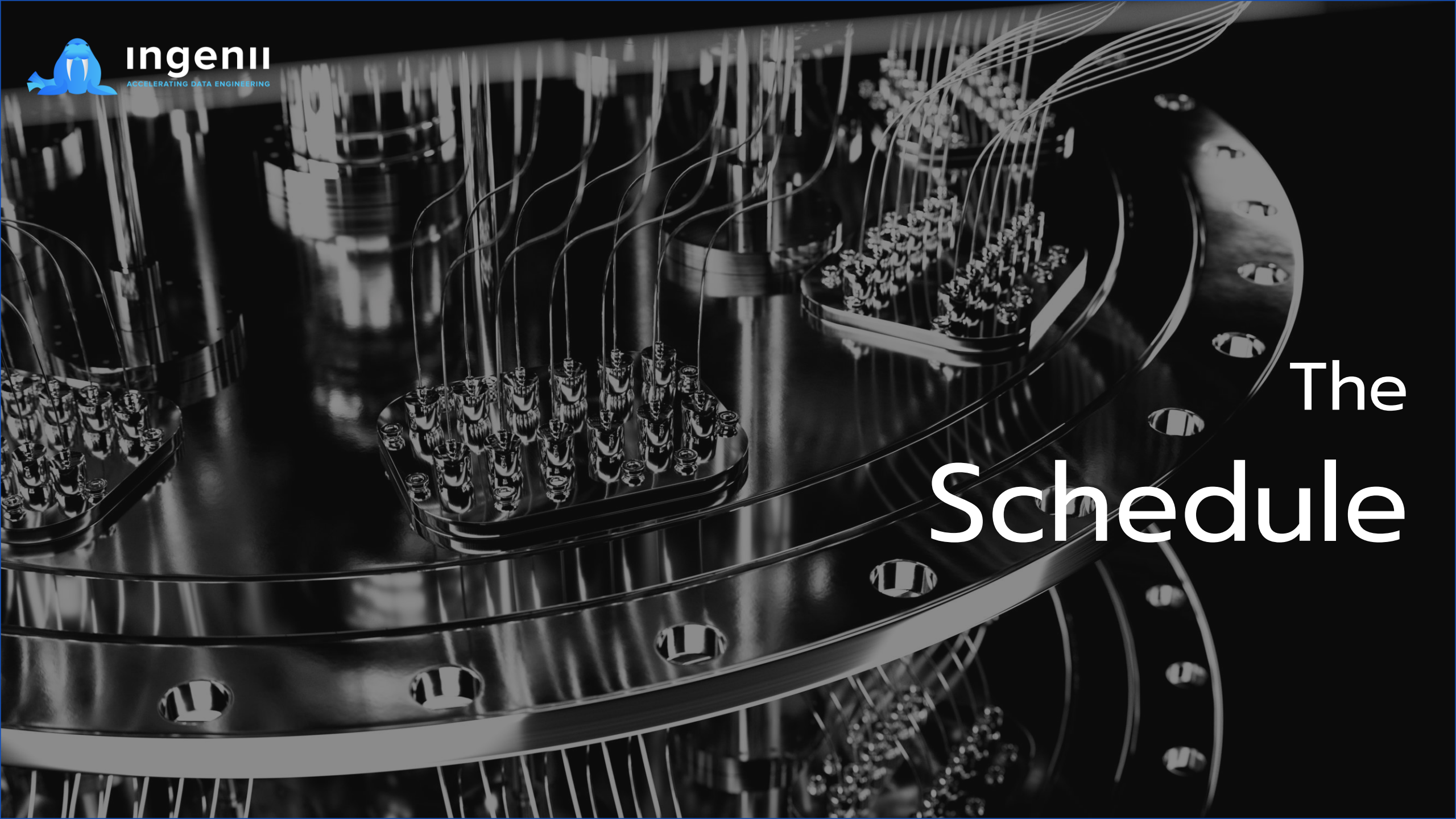
- Paper
- Git repository
- We use this pipeline

ingenii
ACCELERATING DATA ENGINEERING

End of Fall Milestone

Outputs

- Have a presentation to show at the end of each sprint, example and template will be shared
- We are aiming to publish the research and findings of this project as a published paper.
- A production-ready algorithm that can begin being applied to commercial engagements


ingenii
ACCELERATING DATA ENGINEERING
The
Schedule

Sprint 1
22/Aug – Week 1 + 2

Spring 2023 Project Timeline

Sprint 1
9/Jan – Week 1 + 2


ingenii
ACCELERATING DATA ENGINEERING
Preparations

Background Reading

- Pharmaceutical/quantum articles
  - https://quantumzeitgeist.com/big-pharma-working-with-quantum-computing-titans/
  - https://www.dwavesys.com/media/kjof1cdh/dwave_menten-ai_case_story-2_v4.pdf
  - https://www.accenture.com/us-en/case-studies/life-sciences/quantum-computing-advanced-drug-discovery
  - https://fortune.com/2021/01/28/roche-to-use-quantum-computing-for-drug-discovery/
- More generic:
  - https://venturebeat.com/2021/04/27/how-merck-works-with-seeqc-to-cut-through-quantum-computing-hype/
  - https://www.mckinsey.com/industries/life-sciences/our-insights/pharmas-digital-rx-quantum-computing-in-drug-research-and-development
  - https://www.ibm.com/downloads/cas/EVBKAZGJ

References

- Paper:
  - https://pubs.acs.org/doi/abs/10.1021/acs.jcim.0c01306
  - Supporting information: https://pubs.acs.org/doi/suppl/10.1021/acs.jcim.0c01306/suppl_file/ci0c01306_si_001.pdf
- Qiskit:
  - Qiskit textbook: https://qiskit.org/textbook/
  - Tutorial: https://qiskit.org/learn/course/introduction-course
- Classical ML models:
  - https://arxiv.org/abs/1703.10603
  - https://pubs.acs.org/doi/10.1021/acs.molpharmaceut.7b01134
  - https://pubs.acs.org/doi/10.1021/acscentsci.8b00507
- Quantum CNNs:
  - https://www.researchgate.net/publication/339539070_Quanvolutional_neural_networks_powering_image_recognition_with_quantum_circuits
  - https://arxiv.org/abs/2108.00661
  - https://arxiv.org/abs/2009.09423
  - https://arxiv.org/abs/1907.02085
- Quantum GNN:
  - https://arxiv.org/abs/2201.05158
  - https://pavanjayasinha.medium.com/quantum-graph-neural-networks-applied-1f5b37922425
  - https://arxiv.org/abs/2107.03257

ingenii
ACCELERATING DATA ENGINEERING

Getting Started

Code

- We will be using GitHub and the following repository: https://github.com/ingenii-dev/DataMine-2022
  - If you don't have access to the repository, just email: support@ingenii.dev
  - If you need any help with using Git or GitHub, check out the GitHub Quickstart pages
  - If you just need a quick cheat-sheet, try this one.
- Our IDE will be VScode and Jupyter Notebooks
- We will be using Qiskit for the quantum development: https://qiskit.org/

Computing Resources

- We have two resources available for the research:
  - Purdue Anvil: https://the-examples-book.com/data-engineering/rcac/anvil
  - Ingenii Azure Data Platform and Azure Quantum


ingenii
ACCELERATING DATA ENGINEERING
We look forward to working together!
Ingenii.io