

# 코오롱베니트 커스텀 트랙 주제 안내

## 트랙 소개

코오롱베니트와 함께하는 커스텀 트랙에서는 **산업 현장의 실제 문제를 바탕으로 한 두 가지 주제 중 하나를 선택**하여 참가합니다. (★ 표시된 문제 1은 코오롱베니트에서 특히 관심을 두고 있는 주제입니다.)

이번 트랙은 정답을 빠르게 맞히는 것보다, 양자 컴퓨팅을 산업 문제에 어떻게 적용할 수 있을지에 대한 아이디어와 논리를 발전시키는 데 목적이 있으며, 눈앞의 문제 해결 자체보다, 앞으로 산업 현장에 적용해볼 수 있는 세부 주제 영역을 함께 발굴하는 것을 더 중요하게 보고 있습니다. 문제를 어떻게 해석하고, 어느 단계에 양자 기술을 배치하며, 그 선택을 어떻게 설득력 있게 설명하는지가 핵심입니다.

### **하이브리드 접근도 인정됩니다.**

현재 기술적 한계로 양자 컴퓨터만으로는 결론을 내리기 어려운 경우에도 괜찮습니다. "왜 양자 이득이 크지 않은지"를 논리적으로 도출했다면 그 자체로 유의미한 결과로 인정하며, 양자와 AI를 결합한 하이브리드 형태의 현실적인 접근 방식도 적극 허용합니다.

## 코오롱베니트 특별상

코오롱 커스텀 트랙 문제를 가장 우수하게 해석하고 산업 적용 가능성을 잘 제시한 팀에게는 코오롱베니트 특별상이 별도로 수여됩니다. 수상팀에게는 PoC 연계 검토 및 공동 프로젝트 지원이 제공되며, 수상자가 대학생인 경우 코오롱베니트 실무진의 커리어 멘토링 기회도 함께 제공됩니다.

## ★ 주제 1. Quantum AI 기반 소재구조-물성 예측

---

### “Quantum AI 기반 소재구조-물성 예측 및 최적 후보 탐색”

#### 이런 분들에게 추천합니다

화학·재료공학·화학정보학(cheminformatics) 전공 또는 실무 경험이 있거나, 분자 구조 데이터를 다뤄본 경험이 있는 참가자에게 적합합니다. 여기에 양자컴퓨팅/머신러닝 구현이 가능한 개발자가 팀에 함께 있으면 이상적입니다. 코딩 실력보다 문제를 어떻게 재정의하고 설득력 있게 구조화하는지가 더 중요하므로, 도메인 지식과 구현 역량을 팀 내에서 나누어 맡는 구성을 권장합니다.

#### 배경

소재 개발에서는 분자 또는 고분자 구조로부터 원하는 물성을 빠르고 정확하게 예측하는 기술이 매우 중요합니다. 기존에는 실험이나 시뮬레이션을 통해 물성을 확인해야 했기 때문에 많은 시간과 비용이 소요되었고, 최근에는 SMILES와 같은 구조 표현과 머신러닝을 결합해 물성을 예측하는 연구가 활발히 진행되고 있습니다.

다만 소재 데이터는 다음과 같은 특성을 가지고 있어, 일반적인 머신러닝 접근만으로는 한계가 있습니다.

- 제한적인 데이터 수
- Descriptor 차원이 높음
- 구조와 물성 간 관계가 강한 비선형성을 가짐
- 여러 물성을 동시에 고려해야 함
- 최종적으로는 단순 예측을 넘어 최적 후보 소재 탐색까지 필요함

이러한 특성은 Quantum Machine Learning, Quantum Kernel, Variational Quantum Circuit, QUBO 기반 최적화 등과 결합하기 좋은 문제 구조를 가지고 있습니다.

#### 과제 목표

참가자는 공개 소재 데이터셋을 활용하여 양자 머신러닝 기반 물성 예측 모델을 개발하고, 예측 결과를 바탕으로 목표 물성 조건을 만족하는 후보 소재를 탐색하는 것을 목표로 합니다.

- SMILES 기반 소재 구조 데이터를 활용한 물성 예측 모델 개발
- Quantum ML 모델과 Classical ML 모델의 성능 비교
- 소량 데이터 조건에서 Quantum ML의 적용 가능성 탐색

- 목표 물성 조건을 만족하는 최적 후보 소재 선정
- 예측 문제와 최적화 문제를 결합한 Quantum AI 활용 시나리오 구현

## 문제 설명

참가자는 공개 데이터셋으로부터 제공되는 SMILES 문자열과 물성값을 이용합니다. SMILES로부터 RDKit 등의 도구를 활용해 molecular descriptor 또는 fingerprint를 생성하고, 이를 입력값으로 사용해 물성 예측 모델을 학습합니다.

이후 다음 중 하나 이상의 Quantum AI 접근법을 적용하여 문제를 해결합니다.

- Quantum Kernel Regression
- Quantum Support Vector Regression
- Variational Quantum Regressor
- Quantum Neural Network
- QUBO 기반 feature selection
- QUBO 기반 후보 물질 선택
- Quantum-inspired optimization

최종적으로 classical baseline 모델과 비교하여 예측 성능을 평가하고, 목표 물성 조건을 만족하는 후보 소재 리스트를 도출합니다.

## 활용 가능 데이터셋

아래 공개 데이터셋 중 선택하여 활용할 수 있으며, 저분자와 고분자 두 카테고리로 구분됩니다.

| 구분 | 데이터셋       | 대상  | 주요 물성          | 특징              |
|----|------------|-----|----------------|-----------------|
| 1  | PolyMetriX | 고분자 | 유리전이온도(Tg)     | 실험 기반 고분자 데이터   |
| 2  | RadonPy    | 고분자 | 밀도, 굴절률, 열전도율  | 시뮬레이션 기반 고분자 물성 |
| 3  | ESOL       | 저분자 | 수용해도           | 소규모 저분자 회귀 문제   |
| 4  | QM9        | 저분자 | 에너지갭, 원자화에너지 등 | 표준 분자 물성 벤치마크   |

빠르게 시작하고 싶다면 ESOL 또는 QM9의 일부 subset을, 고분자 산업 응용성을 강조하고 싶다면 PolyMetriX 또는 RadonPy를 선택하는 것을 권장합니다.

위 데이터셋은 모두 공개된 실제 데이터이며, 예선·본선 구분 없이 참가자가 직접 조달하여 동일하게 활

용하시면 됩니다.

#### [참고 데이터] SMILES 및 물성값 관련 공개 데이터셋

- 저분자(small molecule)와 고분자 두 카테고리로 구분

| No | 데이터셋                               | 사이트                                                                                                                                                                              | 물성(개수)                                   |
|----|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| 1  | PolyMetriX, 고분자                    | <a href="#">GitHub - lamalab-org/PolyMetriX: PolyMetriX is a comprehensive Python library that powers the entire machine learning workflow for polymer informatics. · GitHub</a> | 보정된<br>유리전이온도<br>(7,367)<br><br>실험       |
| 2  | RadonPy, 고분자                       | <a href="#">GitHub - RadonPy/RadonPy: RadonPy is a Python library to automate physical property calculations for polymer informatics. · GitHub</a>                               | 밀도, 굴절률,<br>열전도율(1,070)<br><br>시뮬레이션     |
| 3  | ESOL, 저분자                          | <a href="#">scikit-fingerprints/MoleculeNet_ESOL · Datasets at Hugging Face</a>                                                                                                  | 수용해도 (1,128)                             |
| 4  | QM9, 저분자<br>(가장 널리 쓰이는<br>표준 벤치마크) | DeepChem 라이브러리에서 로드 가능                                                                                                                                                           | 에너지갭, 원자화<br>에너지외 총 12 개<br>물성 (133,885) |

## 생각의 출발 및 확장

아래는 접근 방향과 완성도를 가늠하는 데 참고할 수 있는 힌트입니다. 구체적인 파이프라인 설계와 선택 근거는 참가팀이 직접 구성해야 합니다.

### 출발점

- SMILES로부터 어떤 정보를 추출해 입력값으로 쓸지부터가 이미 설계의 일부입니다.
- 물성 예측 하나만으로는 충분하지 않습니다. Classical 모델과 Quantum 모델의 성능을 나란히 비교하는 설계까지 포함해야, "왜 양자인가"를 설득력 있게 설명할 수 있습니다.

### 한 단계 더 나아간다면

- 예측에서 그치지 않고, 목표 조건을 만족하는 후보를 어떻게 선별하고 순위를 매길 것인지까지 설계에 포함하면 제안의 완성도를 높일 수 있습니다.
- 여러 물성 조건을 동시에 고려해야 하는 상황이라면, 이를 하나의 최적화 문제로 묶어 표현하는 방법도 고민해볼 만합니다.

## 참가 전 확인해 볼 점

**정답 하나를 맞히는 문제가 아닙니다.**

아래 질문에 대한 참가팀 자신만의 논리와 선택 근거를 결과물에 함께 제시해야 합니다.

- 어떤 데이터셋과 물성을 선택했고, 그 이유는 무엇인가?
- Quantum 접근이 실제로 어떤 강점 또는 한계를 보였으며, 그 결과를 어떻게 해석할 것인가?
- 이 접근을 실제 산업 현장에 적용한다면 어떤 비즈니스 가치로 이어질 수 있는가?

## 주제 2. Knowledge Graph 기반 비즈니스 가치 조합 최적화

### “Knowledge Graph 기반 최대가치 기업·기술 조합 탐색을 위한 Quantum Optimization Challenge”

#### 이런 분들에게 추천합니다

전략기획·투자분석·비즈니스 애널리틱스에 관심이 있거나, 조합 최적화·그래프 이론에 익숙한 참가자에게 적합합니다. 여기에 QUBO/QAOA 등 양자 최적화 구현이 가능한 개발자가 팀에 함께 있으면 이상적입니다. 이 문제 역시 구현력보다 문제를 어떻게 구조화하고 현실적으로 설득력 있게 풀어내는지가 더 중요합니다.

#### 배경

기업은 새로운 사업 기회나 투자 대상을 찾을 때, 개별 항목의 점수만 보는 것이 아니라 여러 요소를 함께 고려해야 합니다.

- 기업 자체의 가치
- 기술 간 시너지
- 시장 성장성
- 특허 및 경쟁 우위
- 공급망 및 규제 리스크
- 예산 제약
- 중복 투자 방지

이 문제는 본질적으로 조합 최적화 문제이며, QUBO(Quadratic Unconstrained Binary Optimization) 또는 Ising 형태로 정식화하면 양자 최적화 알고리즘으로 실험할 수 있습니다.

#### 시나리오

참가자는 전략투자팀의 Quantum AI 분석가 역할을 맡습니다. 운영진은 기업, 기술, 시장, 특허, 리스크 정보가 담긴 synthetic Knowledge Graph를 제공하며, 참가자의 목표는 다음 조건을 만족하는 최적의 기업·기술 조합을 찾는 것입니다.

- 총 예산 이내
- 선택 개수 제한 충족
- 경쟁 또는 중복되는 조합은 회피

- 리스크는 낮추고, 가치와 시너지는 높이는 방향으로 선택

이 Knowledge Graph는 실존 기업 데이터가 아닌 가상의(synthetic) 데이터로, 예선 단계에서는 노드·관계 구성 방식과 변수 스키마 설명만 제공되며, 실제 데이터셋 파일은 본선 진출팀에게 별도로 제공될 예정입니다.

## 참가자 과제

주어진 그래프에서 후보 노드들 중 일부를 선택하여, 다음 목적을 최대화합니다.

- 기업 가치
- 기술 가치
- 시장 성장성
- 관계 시너지
- 특허 방어력

동시에 다음 항목은 최소화합니다.

- 비용
- 리스크
- 중복성
- 경쟁 충돌

풀이 방식은 Quantum Annealing, QAOA, Quantum-inspired Solver 중 선택하여 적용할 수 있으며, 양자 우위를 증명하는 것보다 문제를 양자 최적화 형태로 잘 모델링하고 푸는 것 자체가 목표입니다.

## 기획 의도

이 문제는 단순히 쿼텀 알고리즘을 구현하는 경연이 아니라, 현실의 복잡한 의사결정 문제를 최적화 문제로 전환하고 해결하는 능력을 평가하기 위해 설계되었습니다.

실무에서는 여러 후보 중 하나만 고르는 문제보다, 여러 후보를 함께 고려했을 때 가장 좋은 조합을 찾는 문제가 더 중요합니다. 다만 이 과정은 정보가 분산돼 있고 관계가 복잡하며 정량화하기 어려워, 사람이 직접 판단하기에는 한계가 있습니다.

즉 이 과제는 Knowledge Graph를 통해 기업·기술·시장·리스크의 관계를 구조화하고, 이를 QUBO 기반 조합 최적화 문제로 변환하여, 참가자가 양자 또는 하이브리드 방식으로 탐색하도록 설계되었습니다.

## 심사 포인트

단순한 정답 여부보다 다음에 중점을 둡니다.

- 문제를 얼마나 잘 구조화했는가
- 관계와 제약을 얼마나 잘 반영했는가
- 제안한 조합이 현실적으로 설득력 있는가
- Classical 방식 대비 어떤 장점이 있는가
- 설명과 시각화가 명확한가

## 생각의 출발 및 확장

양자 기술 활용이 이 문제에서 어떤 가치를 가지는지는, 아래 관점을 참고해 참가팀이 직접 설명해야 합니다.

- 복잡한 조합 문제를 어떤 형태로 표현할 것인가
- Classical baseline 과 비교 가능한 실험을 어떻게 설계할 것인가
- 제안하는 방식은, 산업계의 어떤 문제로 응용 및 확장 가능할 것인가

## 참가 전 확인해 볼 점

**정답 하나를 맞히는 문제가 아닙니다.**

아래 질문에 대한 참가팀 자신만의 논리와 선택 근거를 결과물에 함께 제시해야 합니다.

- Knowledge Graph 의 노드와 관계를 어떻게 정의했고, 그 근거는 무엇인가?
- 제안한 조합이 현실적으로 설득력이 있는가?
- Classical 방식과 비교했을 때 어떤 차이와 트레이드오프가 있었는가?

---

\* 문의 사항은 운영사무국으로 연락하거나, 데이콘 웹페이지의 'Quantum Reframing Challenge 2026' 게시판에 질문 주시기 바랍니다.