Patent model development

Develop patent model to check patent infringement for chemical compound

CALICI

Company | Calici Co., Ltd.

Bio & AI Team | Christine Le

Project Overview

Project Overview

Project Name	• Patent infringement detection model for chemical compound
Overall Objective	• Develop a ML model capable of identifying potential infringement by analyzing input chemical compounds against a database of patented molecules
Context	• This project will provide a function to add in Pharmaco-Net – a drug discovery platform to check whether lead compounds may infringe existing patents and provide guidance for structural modifications to avoid infringement during optimization
Requested by	• Calici Co.,

Scope of work

Scope of work

- Preprocess and curate a dataset from public and proprietary chemical patent databases (e.g., USPTO, Google Patents, etc.).
- Build an AI/ML model to recognize image (chemical compound, Markush Claim image) in the patent document and check the description to understand the whole protected parts.
- Input: Chemical compound (SMILES)
- Expected output:
  - Patented/Not patented
  - If patented, show the part infringing patent

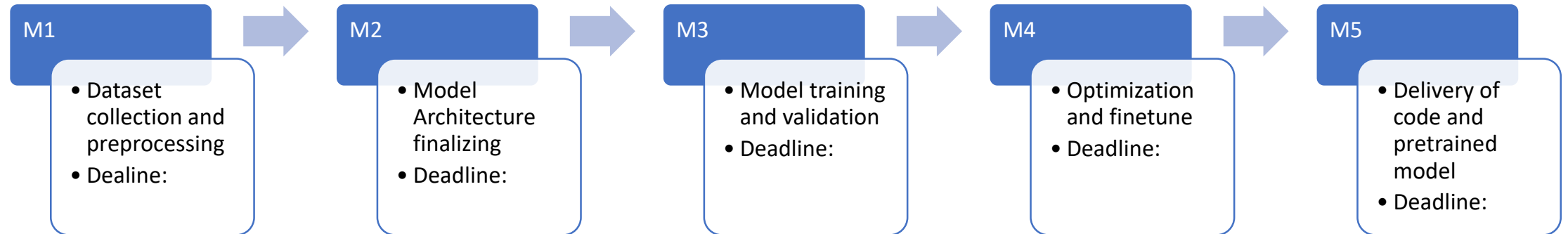|	
Technical	Language: Python 3.9+
	Libraries: Rdkit, Pytorch or TensorFlow
	Must support SMILES input and optionally InChL
Quality expectation	Precision:
	Inference time:

Timeline & Milestones

Timeline & Milestones

M1
- Dataset collection and preprocessing
- Dealine:

M2
- Model Architecture finalizing
- Deadline:

M3
- Model training and validation
- Deadline:

M4
- Optimization and finetune
- Deadline:

M5
- Delivery of code and pretrained model
- Deadline:

Deliverables

Deliverables

- All codes and dataset used in project with README file for reproduction
- Trained model
- Evaluation report (model performance metrics)
- Sample output on 100 test compounds

Budget & Payment

Budget & Payment

Estimated Budget:
Payment plan:
Payment method:

Communication & Contact

Communication & Contact

- Zoom meeting to report on progress at the end of each milestone/ Zoom meeting to report on progress every month
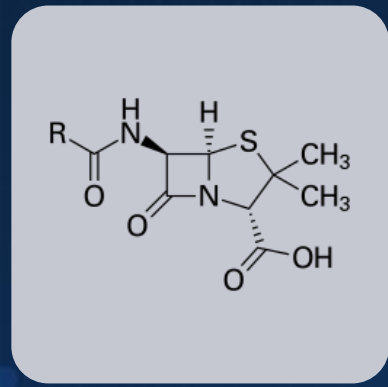- Any issues during work, contact with … via …

References

Similar works for references

CAS SciFinder: https://cas-product-help.zendesk.com/hc/en-us/articles/9838526793613-Find-Patent-Markush-Structures

PatentFinder: https://arxiv.org/html/2412.07819v1

감사합니다

T H A N K   Y O U

한국 |	· 대전광역시 유성구 대학로 99(궁동), 대전 TIPS타운 1.5층 디5 (34134)
	· 042 – 719 – 4319 | 070 – 8664 – 1859
	· info@calici.co
기업부설연구소 |	· 대전광역시 유성구 테크노4로 17(관평동), 대덕비즈센터 A415호
	· 070 – 4659 - 5510
미국 |	· 3003 N First St, San Jose, CA 95134
	· 2450 Holcombe Blvd, Suit j Houston TX 77024
파마코-넷 |	https://pharmaco-net.org/