

NOVÉ TECHNIKY UČENIA SA BEZ UČITEĽA PRE KLASIFIKÁCIU A PREDIKCIU MOLEKULÁRNYCH VLASTNOSTÍ

Samuel Baran

Zadávatel' práce:

Tachyum s.r.o.

Vedúci práce (Tachyum):

Ing. Slavomír Mat'ášovský

Vedúci práce (UPJŠ):

RNDr. Ľubomír Antoni, PhD.

MOTIVÁCIA

Dáta popisujúce molekulárne vlastnosti

Heterogénne

rôzna veľkosť molekúl

variabilný počet komponentov

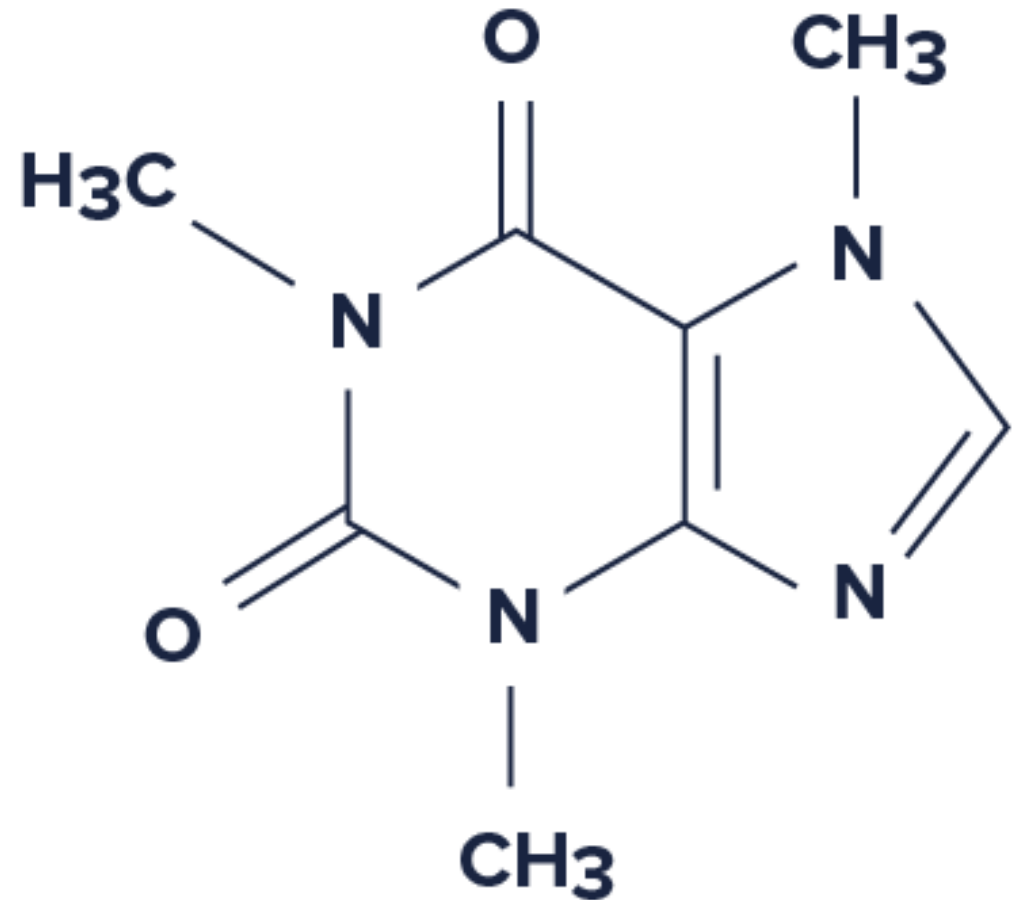
rôzne väzby

priestorové rozmiestnenie

Náročné na zber

špecializované nástroje

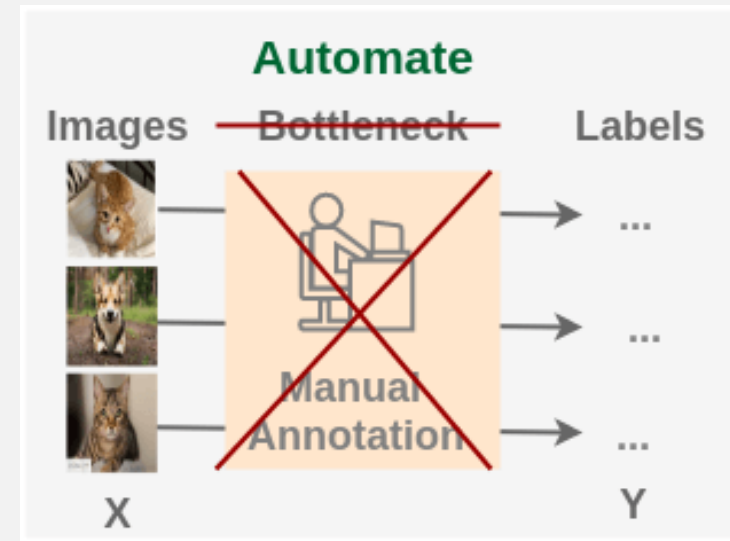
dohľad expertov pri zbere dát



MOTIVÁCIA

Metódy učenia bez učiteľa

- využívajú neoznačené dáta
- hľadajú štruktúry, vzory, podobnosti v dátach
- výstupom sú klastre dát bez anotácií
- dotrénovanie anotácií klastrov
 - tréning s učiteľom
 - menšie množstvo dát



CIELE

- Spracovať prehľad metód klasifikácie a predikcie vlastností molekúl a najnovších techník učenia sa bez učiteľa alebo čiastočného učenia bez učiteľa, ktoré sú využívané v oblasti počítačového videnia.
- Navrhnuť novú metódu klasifikácie a predikcie vlastností molekúl, ktorá bude inšpirovaná existujúcimi technikami učenia sa bez učiteľa alebo čiastočného učenia bez učiteľa z oblasti počítačového videnia.
- Implementovať a experimentálne overiť navrhnuté riešenie testovaním na dátach pre klasifikáciu a predikciu molekulárnych vlastností.
- Porovnať dosiahnuté výsledky navrhnutého riešenia s doteraz najúspešnejšími technikami vyvinutými v oblasti molekulárnych vlastností

AKTUÁLNY STAV

Štúdium literatúry
popisujúcej problémy
predikcie a klasifikácie
molekulárnych vlastností

- MoleculeNet
- DeepChem

Štúdium metód učenia bez
učiteľa

- kontrastívne učenie

ODPORÚČANÁ LITERATÚRA

- Zhenqin Wu and Bharath Ramsundar and Evan N. Feinberg and Joseph Gomes and Caleb Geniesse and Aneesh S. Pappu and Karl Leswing and Vijay S. Pande, . "MoleculeNet: A Benchmark for Molecular Machine Learning". CoRR abs/1703.00564. (2017). <https://arxiv.org/pdf/1703.00564.pdf>
- DeepChem - <https://deepchem.readthedocs.io/en/latest/index.html>
- Mathilde Caron and Hugo Touvron and Ishan Misra and Hervé Jégou and Julien Mairal and Piotr Bojanowski and Armand Joulin, . "Emerging Properties in Self-Supervised Vision Transformers". CoRR abs/2104.14294. (2021). <https://arxiv.org/abs/2104.14294>
- Mahmoud Assran and Mathilde Caron and Ishan Misra and Piotr Bojanowski and Armand Joulin and Nicolas Ballas and Michael G. Rabbat, . "Semi-Supervised Learning of Visual Features by Non-Parametrically Predicting View Assignments with Support Samples". CoRR abs/2104.13963. (2021). <https://arxiv.org/abs/2104.13963>
- Yuyang Wang and Jianren Wang and Zhonglin Cao and Amir Barati Farimani, . "MolCLR: Molecular Contrastive Learning of Representations via Graph Neural Networks". CoRR abs/2102.10056. (2021). <https://arxiv.org/abs/2102.10056>
- Seyone Chithrananda and Gabriel Grand and Bharath Ramsundar, . "ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction". CoRR abs/2010.09885. (2020). <https://arxiv.org/abs/2010.09885>

ĎAKUJEM ZA POZORNOST

Otázky?