



Aine

Tietojenkäsittelytieteen kandiohjelma

**koneoppimisen menetelmät
lääkkeiden/kemiallisessa syntetisoinnin
mallennuksessa**

Heikki Pulli

12.10.2021

MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA
HELSINGIN YLIOPISTO

Ohjaaja(t)

Tarkastaja(t)

Yhteystiedot

PL 68 (Pietari Kalmin katu 5)
00014 Helsingin yliopisto

Sähköpostiosoite: info@cs.helsinki.fi

URL: <http://www.cs.helsinki.fi/>

Sisällys

1	Johdanto	1
2	Uusien lääkkeiden löytäminen	2
3	Kehitettyjä koneoppimismalleja	3
4	Lääkkeiden syntetisointi ja retrosyntetisointi	4
5	Lääkkeen retrosyntetisoinnin ongelma	5
6	Koneoppimisen käytön hyödyt	6
7	Kehitettyjä apuvälineitä	7
8	Tulevaisuuden koneoppimisen mallit ja kehitys	8
	Lähteet	9

1 Johdanto

Lyhyt johdanto mitä aineessa käsitellään (de novo lääketutkimus ja retrosynteesi)

2 Uusien lääkkeiden löytäminen

Miksi uusien lääkkeiden löytäminen on vaikeaa? [3]

Ongelman kuvaaminen koneoppimisongelmana [4, 5]

3 Kehitettyjä koneoppimismalleja

- Prototype based drug design [4]
- Deep generative autoencoder [5]

4 Lääkkeiden syntetisointi ja retrosyntetisointi

Mitä tarkoittaa lääkkeen syntetisointi? [1]

5 Lääkkeen retrosyntetisoinnin ongelma

Miksi lääkkeiden retrosyntetisointi on hankalaa? [2, 1]

6 Koneoppimisen käytön hyödyt

Miten koneoppiminen voi auttaa uusien lääkkeiden retrosyntetisoinnissa? [8]

7 Kehitetyjä apuvälineitä

- 3N-MCTS [6]
- expert knowledge aided neural networks [7]

Miten koneoppimista hyödynnetään tällä hetkellä lääkkeiden syntetisoinnissa? [6, 4, 7]

8 Tulevaisuuden koneoppimisen mallit ja kehitys

Miten lääkkeiden kehitys tulee hyötymään tulevaisuuden koneoppimisesta? [2]

Lähteet

- [1] A. F. de Almeida, R. Moreira ja T. Rodrigues. "Synthetic organic chemistry driven by artificial intelligence". eng. *Nature reviews. Chemistry* 3.10 (2019), s. 589–604. ISSN: 2397-3358.
- [2] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev ja A. Walsh. "Machine learning for molecular and materials science". eng. *Nature (London)* 559.7715 (2018), s. 547–555. ISSN: 0028-0836.
- [3] S. Ekins, A. C. Puhl, K. M. Zorn, T. R. Lane, D. P. Russo, J. J. Klein, A. J. Hickey ja A. M. Clark. "Exploiting machine learning for end-to-end drug discovery and development". eng. *Nature materials* 18.5 (2019), s. 435–441. ISSN: 1476-1122.
- [4] S. Harel ja K. Radinsky. "Accelerating Prototype-Based Drug Discovery Using Conditional Diversity Networks". Teoksessa: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '18. London, United Kingdom: Association for Computing Machinery, 2018, s. 331–339. ISBN: 9781450355520. DOI: 10.1145/3219819.3219882. URL: <https://doi-org.libproxy.helsinki.fi/10.1145/3219819.3219882>.
- [5] A. Kadurin, S. Nikolenko, K. Khrabrov, A. Aliper ja A. Zhavoronkov. "druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico". eng. *Molecular pharmaceutics* 14.9 (2017), s. 3098–3104. ISSN: 1543-8384.
- [6] M. H. S. Segler, M. Preuss ja M. P. Waller. "Planning chemical syntheses with deep neural networks and symbolic AI". eng. *Nature (London)* 555.7698 (2018), s. 604–610. ISSN: 0028-0836.
- [7] B. Shin, S. Park, J. Bak ja J. C. Ho. "Controlled Molecule Generator for Optimizing Multiple Chemical Properties". Teoksessa: *Proceedings of the Conference on Health, Inference, and Learning*. CHIL '21. Virtual Event, USA: Association for Computing Machinery, 2021, s. 146–153. ISBN: 9781450383592. DOI: 10.1145/3450439.3451879. URL: <https://doi.org/10.1145/3450439.3451879>.

- [8] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer ja S. Zhao. "Applications of machine learning in drug discovery and development". eng. *Nature reviews. Drug discovery* 18.6 (2019), s. 463–477. ISSN: 1474-1776.