RESEARCH

Substructure-based Neural Machine Translation for Retrosynthetic Prediction

Umit V. Ucak¹, Taek Kang², Junsu Ko^{3*} and Juyong Lee^{1*}

*Correspondence: junsuko@arontier.co; juyong.lee@kangwon.ac.kr ³Arontier co., Seoul, South Korea ¹Division of Chemistry and Biochemistry, Department of Chemistry, Kangwon National University, Chuncheon, South Korea Full list of author information is

available at the end of the article

Abstract

Keywords: retrosynthesis planning; machine neural translation; seq-to-seq; attention

Additional Files as Figures.

Please find the supporting materials as **figures** within the "Additional Files" section of the BMC article.

Ucak et al. Page 2 of 3

Author details

¹Division of Chemistry and Biochemistry, Department of Chemistry, Kangwon National University, Chuncheon, South Korea. ²Center for Neuro-Medicine, Korea Institute of Science and Technology, Seoul, South Korea. ³Arontier co., Seoul, South Korea.

References

Additional Files

Additional File 7 : Figure S5B

File name : Supplementary Figure S5B

Title of data : Bioactively similar reactions

File format : Standard Latex figure, formatted as PNG.

Description of data: Depictions of ten bioactively similar reactant candidates (6-10)

Ucak et al. Page 3 of 3

Ground truth

6

Prediction

$$T_c = 0.91$$
 $T_c = 0.90$
 $T_c = 0.90$
 $T_c = 0.89$

Ground truth

8

Prediction

 $T_c = 0.90$
 $T_c = 0.83$

Ground truth

9

Prediction

 $T_c = 0.90$
 $T_c = 0.78$
 $T_c = 0.78$
 $T_c = 0.92$

Figure S5B: Ten reactions lie in the bioactively similar region used to assess the quality of retrosynthesis (6-10).