Supporting Information

Molecule Property Prediction based on Spatial Graph Embedding

Xiaofeng Wang†, Zhen Li*,†, Mingjian Jiang†, Shuang Wang†, Shugang Zhang† and Zhiqiang Wei†

*Correspondence:

Zhen Li:

E-mail: lizhen@ouc.edu.cn

†College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China

CHEMICAL DESCRIPTORS

Below, we list the molecular fingerprints that were extracted by ChemoPy to describe molecules in the Physical Chemistry and Biophysics tasks.

Table S1. List of molecular fingerprints

Constitutional descriptors		Tsch	Schiultz index
Weight	Molecular weight	Tigdi	Graph distance index
nhyd	Count of hydrogen atoms	Platt	Platt number
nhal	Count of halogen atoms	Xu	Xu index
nhet	Count of hetero atoms	Pol	Polarity number
nhev	Count of heavy atoms	Dz	Pogliani index
ncof	Count of F atoms	Ipc	Ipc index
ncocl	Count of Cl atoms	BertzCT	BertzCT
ncobr	Count of Br atoms	GMTI	Gutman molecular topological index based on simple vertex degree
ncoi	Count of I atoms	ZM1-2	Zagreb index with order 1-2
ncarb	Count of C atoms	MZM1-2	Modified Zagreb index with order 1-2
nphos	Count of P atoms	Qindex	Quadratic index
nsulph	Count of S atoms	diametert	Largest value in the distance matrix
noxy	Count of O atoms	radiust	radius based on topology
nnitro	Count of N atoms	petitjeant	Petitjean based on topology
nring	Number of rings	Sito	the logarithm of the simple topological index by Narumi
nrot	Number of rotatable bonds	Hato	harmonic topological index proposed by Narnumi
ndonr	Number of H-bond donors	Geto	Geometric topological index by Narumi
naccr	Number of H-bond acceptors	Arto	Arithmetic topological index by Narumi
nsb	Number of single bonds	ISIZ	Total information index on molecular size
ndb	Number of double bonds	TIAC	Total information index on atomic composition

ntb	Number of triple bonds	DET	Total information index on distance equality
naro	Number of aromatic bonds	IDE	Mean information index on distance equality
nta	Number of all atoms	IVDE	Total information index on vertex equality
PC1-6	Average molecular weight	Sitov	Logarithm of the simple topological index by Narumi

Topological descriptors

		Hatov	Logarithm of the simple topological index by Narumi
W	Weiner index	Hatov	Logarithm of the simple topological index by Narumi
AW	Average Wiener index	Getov	Geometric topological index by Narumi
J	Balaban's J index	Gravto	Gravitational topological index based on topological distance
Thara	Harary number	MREF	Molar refractivity

Basak descriptors

		GMTIV	Gutman molecular topological index based on valence vertex degree(log10)
IC0-6	Information content with order 0-6 proposed by Basak		

Molecular property descriptors

SIC0-6	Complementary information content with order 0-6 proposed by Basak	logP	LogP value based on the Crippen method
CIC0-6	Complementary information content with order 0-6 proposed by Basak	logP2	Square of LogP value based on the Crippen method

Burden descriptors

		TPSA	Topological polarity surface area
bcutp1-16	Burden descriptors based on polarizability	UI	Unsaturation index
bcutm1-16	Burden descriptors based on atomic mass	Hy	Hydrophilic index
bcute1-16	Burden descriptors based on atomic electronegativity	MREF	Molar refractivity

MOE-type descriptors

bcutv1-16	Burden descriptors based on atomic vloumes		

Molecular property descriptors

		PEOEVSA	MOE-type descriptors using partial charges and surface area contributions surface area contributions
logP	LogP value based on the Crippen method	SMRVSA	MOE-type descriptors using MR contributions and surface area contributions
logP2	Square of LogP value based on the Crippen method	VSAEstate	MOE-type descriptors using surface area contributions and Estate indices
TPSA	Topological polarity surface area	SLOGPVSA	MOE-type descriptors using SLogP contributions and surface area contributions
UI	Unsaturation index	PEOEVSA	MOE-type descriptors using partial charges and surface area contributions surface area contributions
Hy	Hydrophilic index	SMRVSA	MOE-type descriptors using MR contributions and surface area contributions