

Hogyan tervezzük redox-aktív molekulákat gépi tanulással

Hamza Andrea

TTK
Elméleti Kémiai Csoport

Szeged, 2024 márc. 7-8.

Célkitűzés



**Computer aided design for
next generation flow batteries**

Feladatunk

- új redox-aktív molekulák azonosítása
- nagyszámú molekula virtuális szűrése („high-throughput screening”)
- módszerek:

kvantumkémia számolások

gépi tanulás („machine learning”)

Mitől redox-aktív egy molekula?

Molekula

- fel tud venni
- el tudja tárolni (stabil)
- le tudja adni (reverzibilis)

1 elektront

Befolyásoló tényezők:

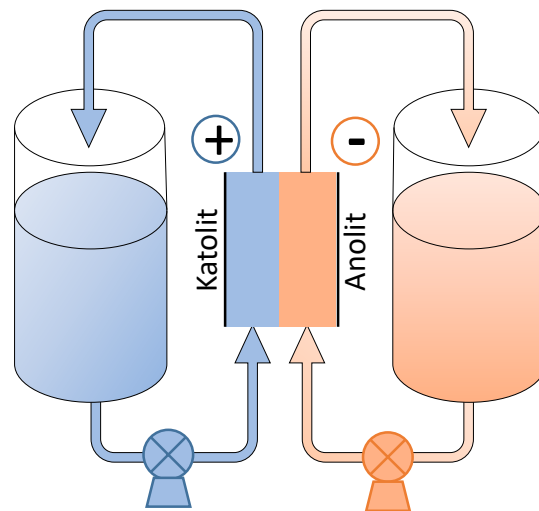
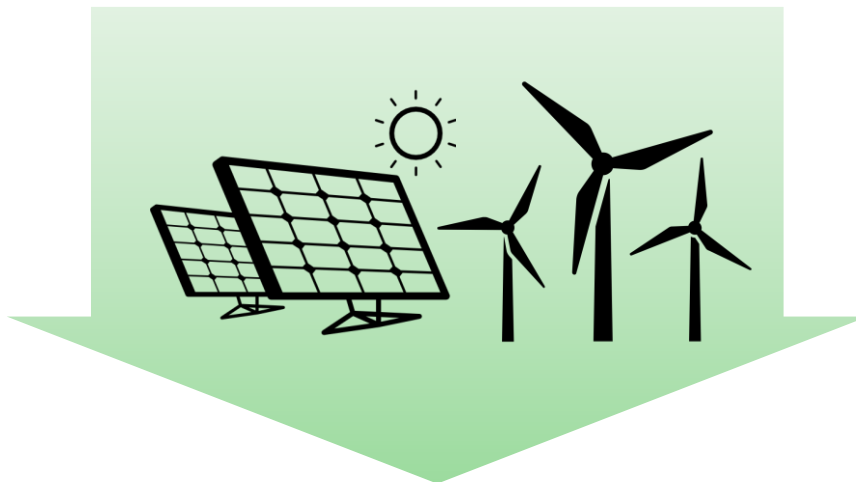
- elektronszerkezet
- térbeli elrendeződés

Új generációs redox folyadékáramos akkumulátorok

- ár
- stabilitás
- toxicitás

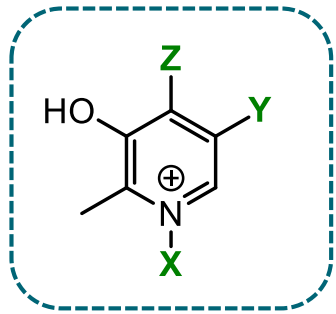
Szerves anyagok

- kémiai stabilitás
- **redox potenciál**
- reverzibilitás
- vízoldhatóság



A legígéretesebb molekula kiválasztása

B6-PYR adatbázis



B6 vitamér

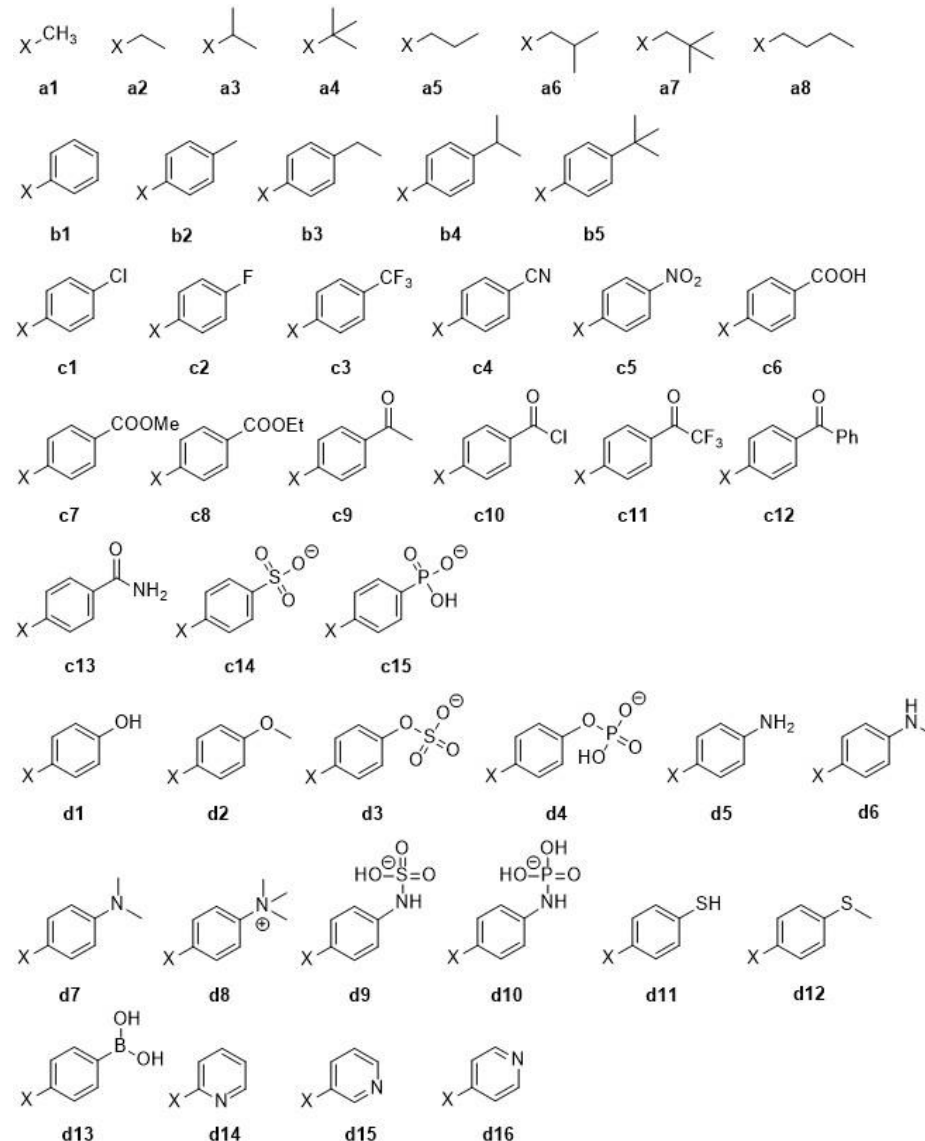
piridinium származékok

5 alapváz

118 szubsztituens



6712 molekula



Redox aktív ?

Adatbázis építése kvantumkémiai számításokkal

Comp² protocol
„Composite CompBat”

$$\Delta G_{comp} = \Delta E_{o,water}(DFT) + \Delta G_{RT,water}(xTB)$$

DFT: M06-2X/6-311+G**/SMD(water)
xTB: GFN2-xTB/GBS(water)

Teszt számolások:

- Kísérlet
- „full” DFT

MAE = 0.22 V

redukciós potenciál

$$\Delta E_{red}^{comp} = -\frac{\Delta G_{comp}}{n_e F} + \Delta E^{SHE}$$

Nernst egyenlet

megbízható módszer

HistPlotly webalkalmazás

B6-PYR adatbázis

- $1e^-$ redukciós potenciál vizes közegben
- xyz, png szerkezetek
- hisztogramok

Interaktív histogram
analizáló*

<http://histplotly.ttk.hu>

Redox Histogram Visualizer

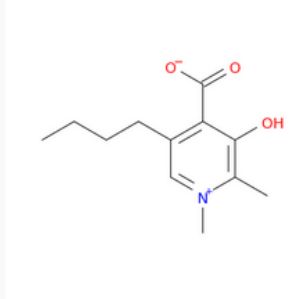
All PYR cores

About Usage Data **View**

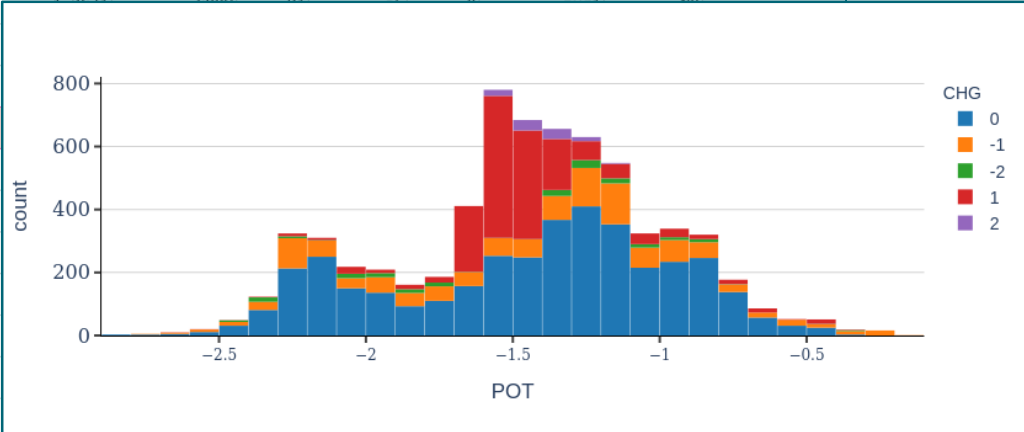
DOWNLOAD CSV

NR	POT	R3	R1	R2	CHG	CORE	RC
1	-1.3018	COOH	P1	a1	0	pyr1	N0
2	-1.2517	COOH	P1	a2	0	pyr1	N0
3	-1.3525	COOH	P1	a3	0	pyr1	N0
4	-1.3621	COOH	P1	a4	0	pyr1	N0
5	-1.3531	COOH	P1	a5	0	pyr1	N0
6	-1.3531	COOH	P1	a6	0	pyr1	N0
7	-1.3531	COOH	P1	a7	0	pyr1	N0
8	-1.3531	COOH	P1	a8	0	pyr1	N0
9	-1.3531	COOH	P1	a9	0	pyr1	N0
10	-1.3531	COOH	P1	a10	0	pyr1	N0
11	-1.3531	COOH	P1	a11	0	pyr1	N0
12	-1.3531	COOH	P1	a12	0	pyr1	N0
13	-1.3531	COOH	P1	a13	0	pyr1	N0
14	-1.3531	COOH	P1	a14	0	pyr1	N0
15	-1.3531	COOH	P1	a15	0	pyr1	N0
16	-1.0979	COOH	P1	c3	0	pyr1	N0

Molecule structure



[Download PNG](#)
[Download xyz](#)
[Download reduced xyz](#)



count

POT

CHG

- 0
- 1
- 2
- 1
- 2

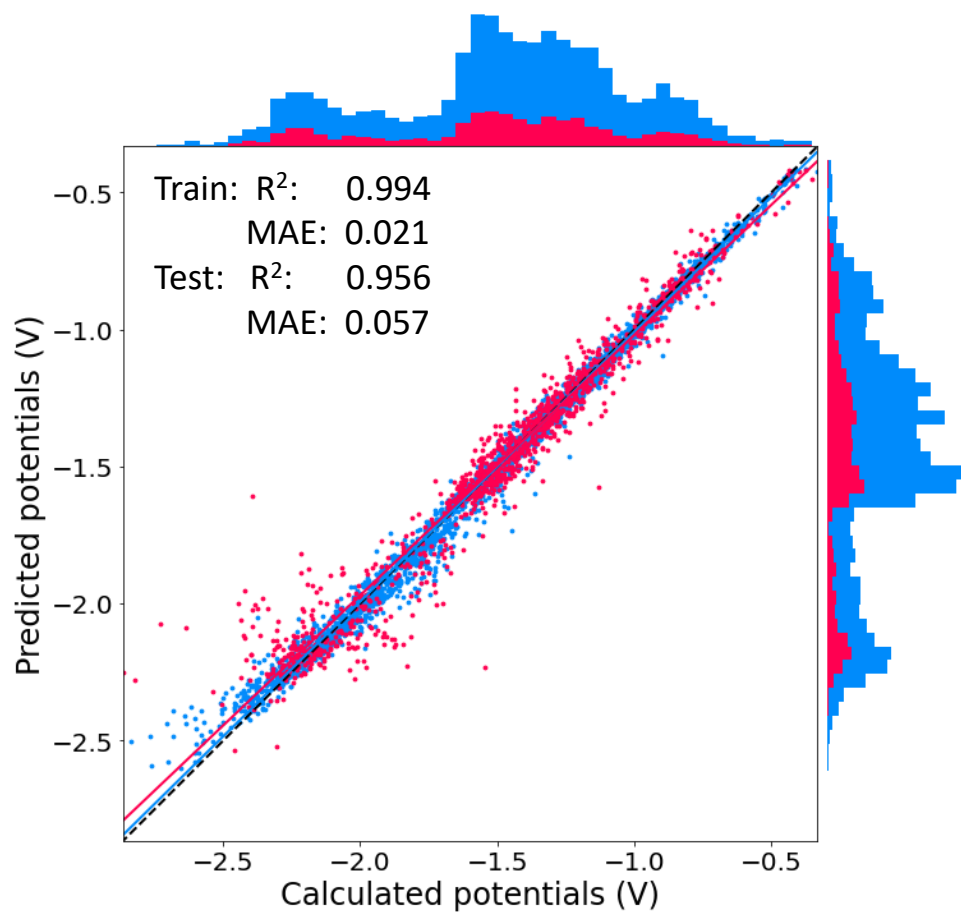
« < 1 / 34 > »

Histogram settings - select columns and resolution:

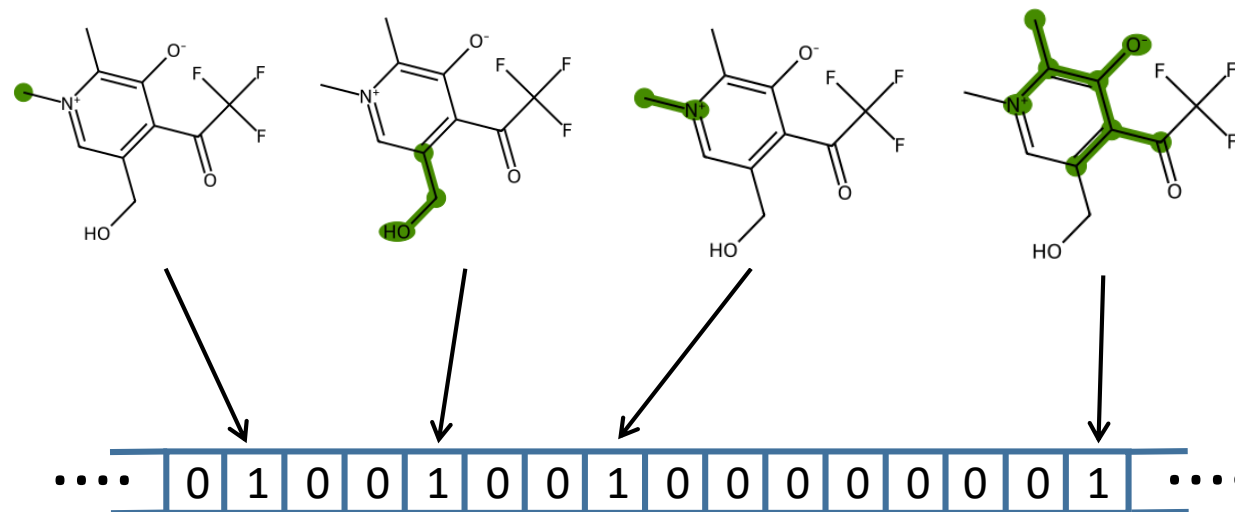
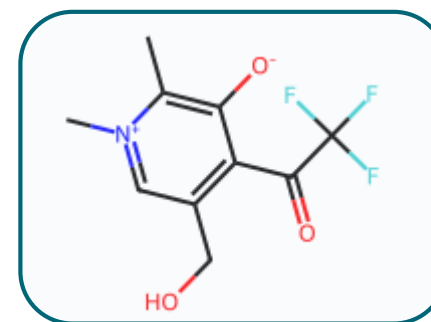
POT CHG × resolution SUBMIT

Becslés vagy/és értelmezhetőség?

Random Forest Regresszió

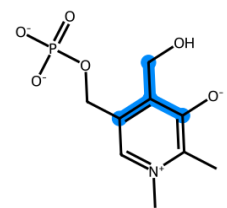
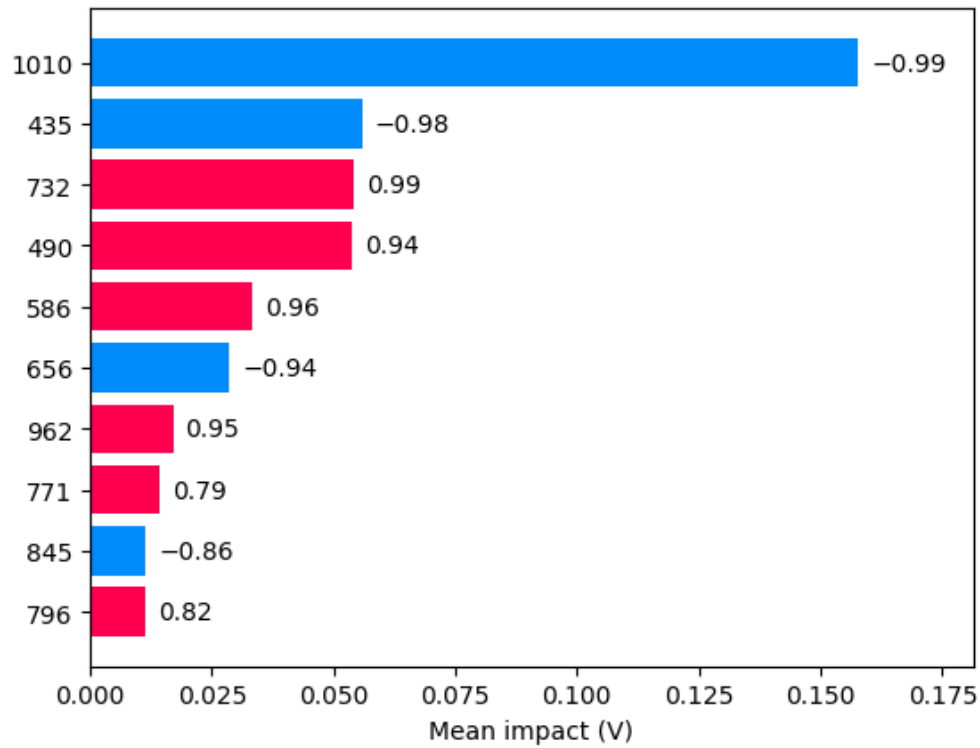


Morgan Fingerprints

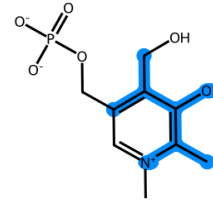


Egy lehetséges értelmezés

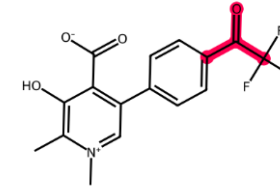
Globális jellemzők fontossága: SHAP analízis



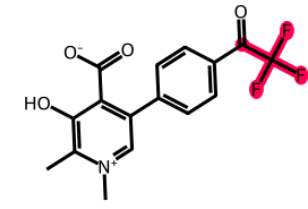
1010



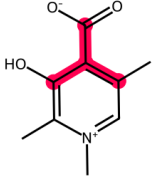
435



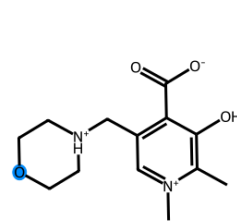
732



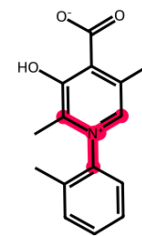
490



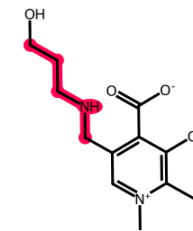
586



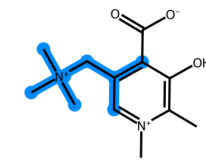
656



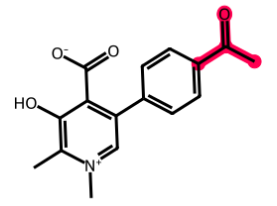
962



771



845



796

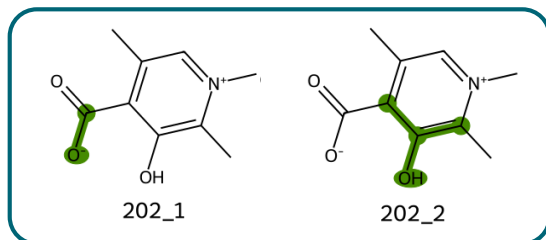
Morgan Fingerprints

Bizonytalanságok

„bit collision”

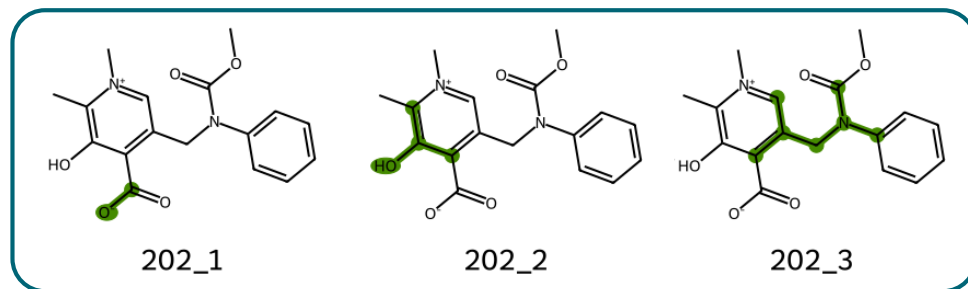
PYR-1

Bit: 202

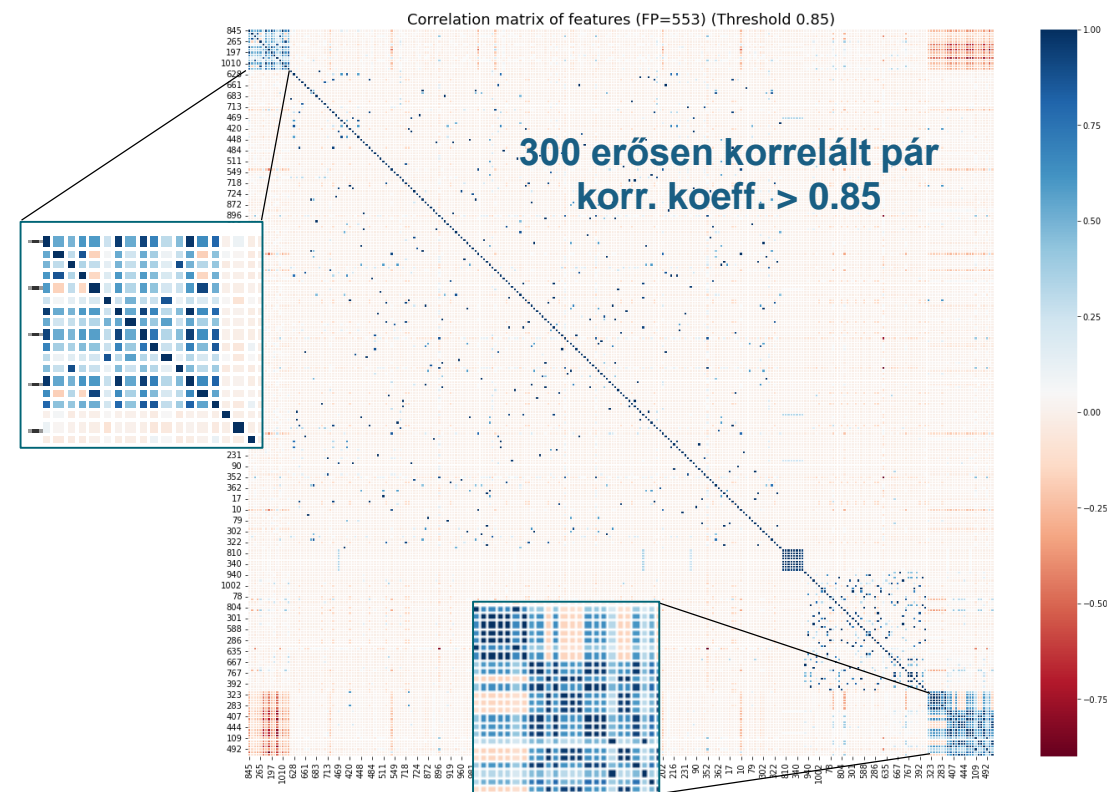


PYR-2720

Bit: 202



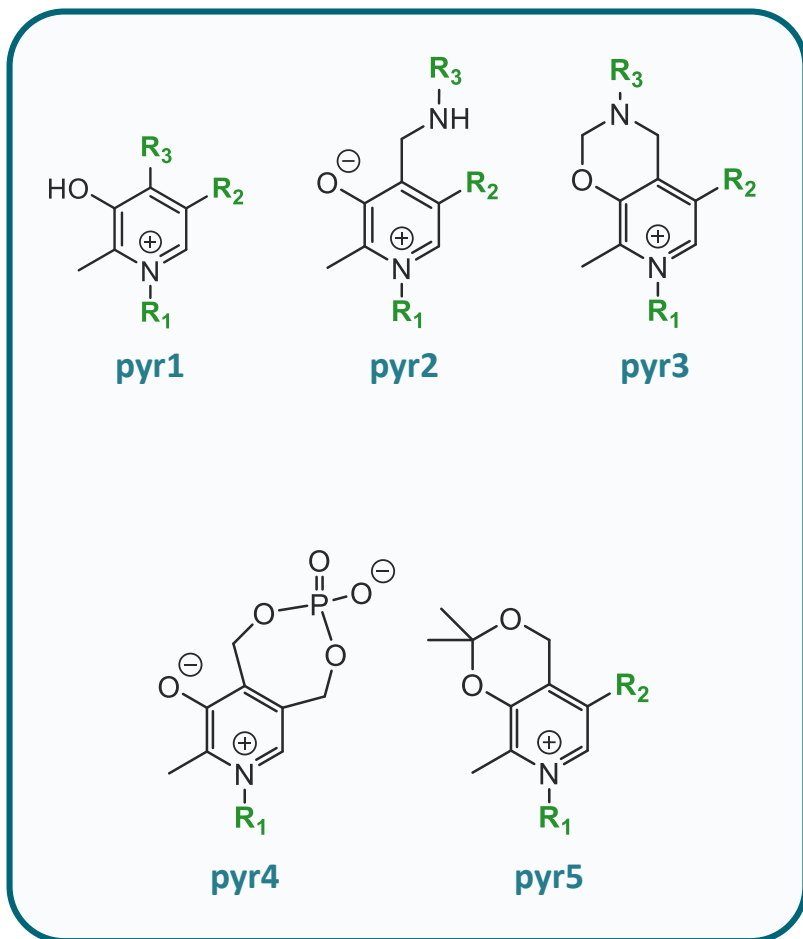
Korrelációs mátrix



+ nem mindig funkciós csoportokhoz köthető jellemzők

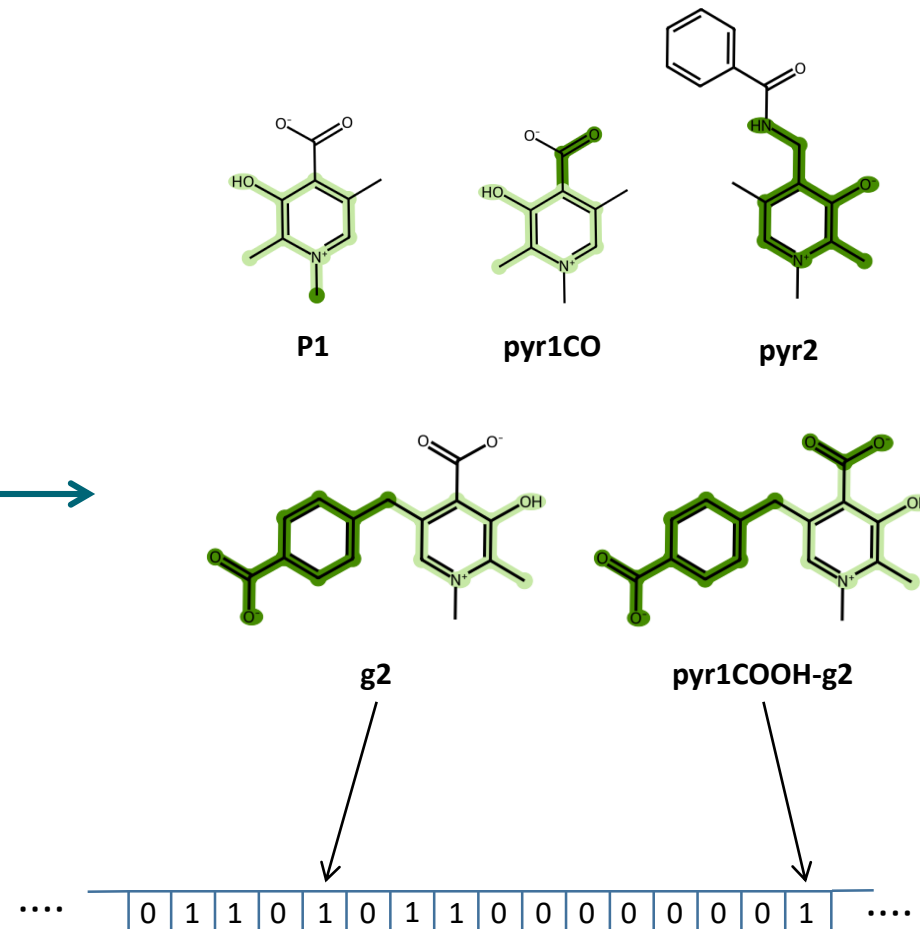
Új fingerprintek bevezetése: RGD (R-group decomposition)

Jól definiált
váz
+
szubsztituensek



B6-PYR

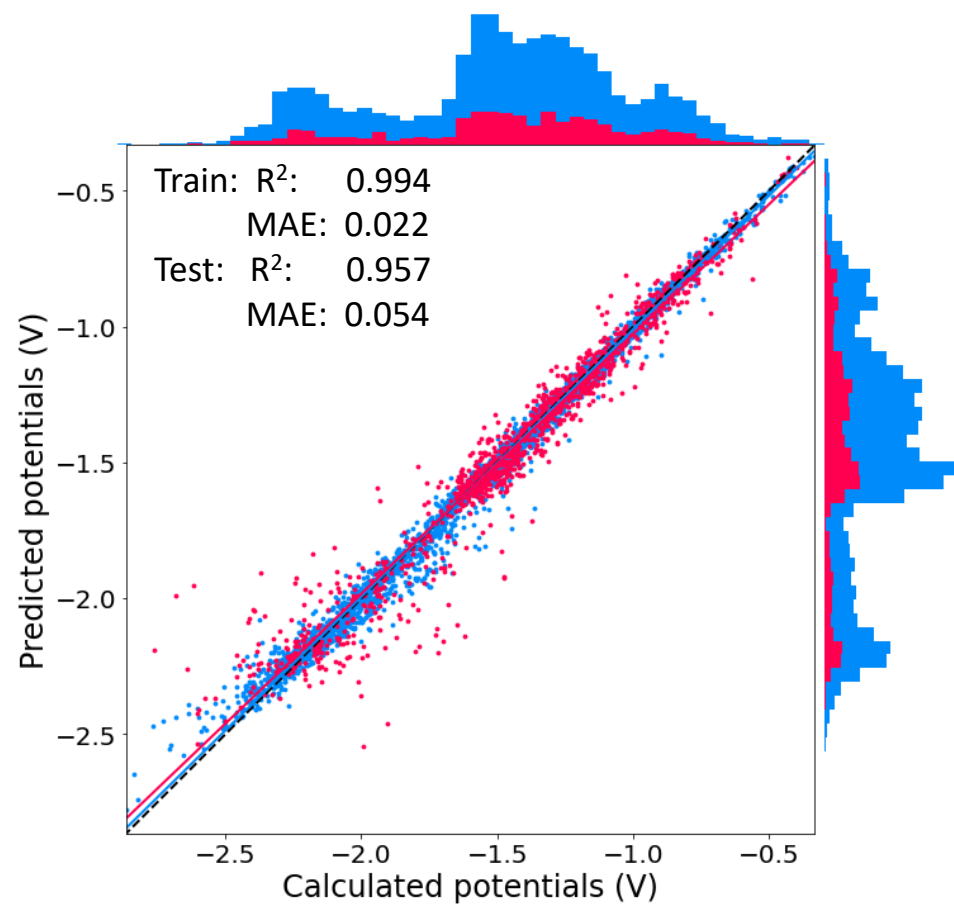
RGD



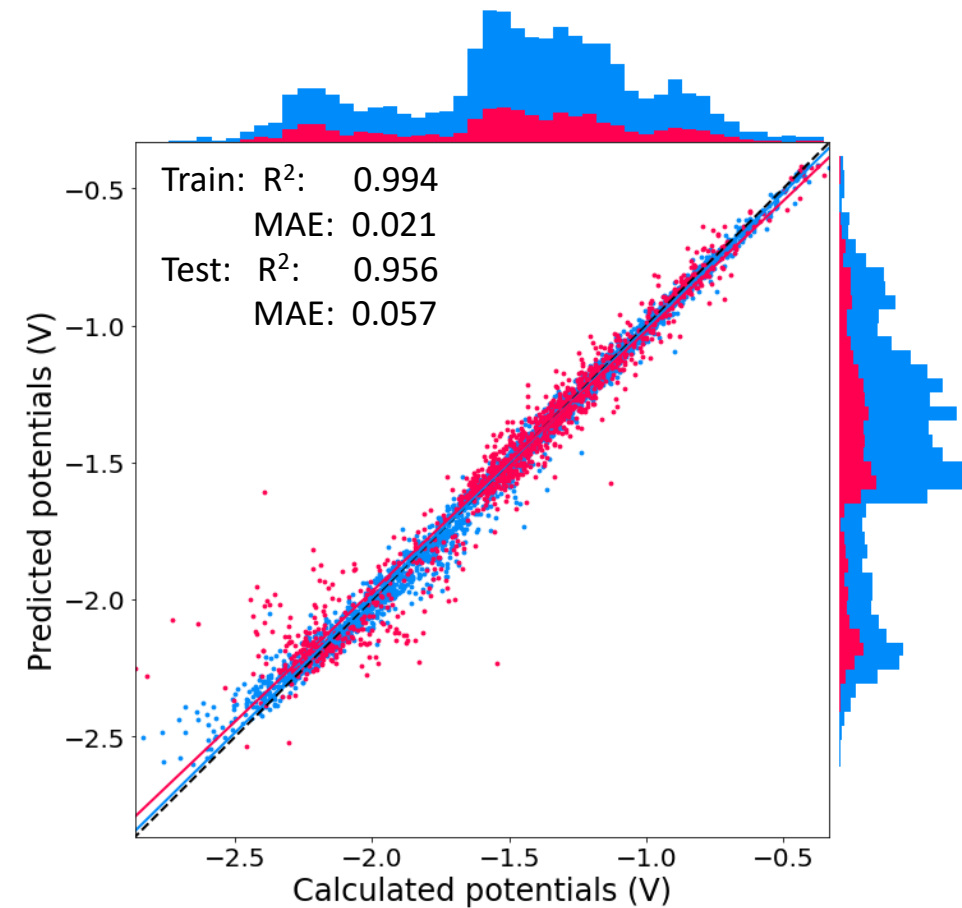
RGD Fingerprints

Random Forest regresszió

RGD Fingerprints



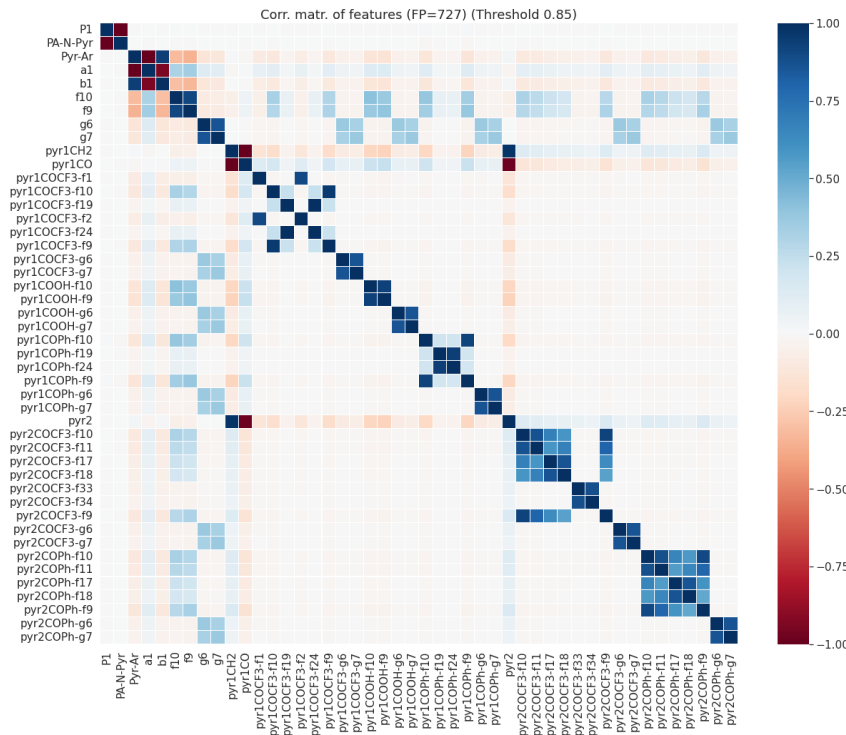
Morgan Fingerprints



Korrelációs mátrix

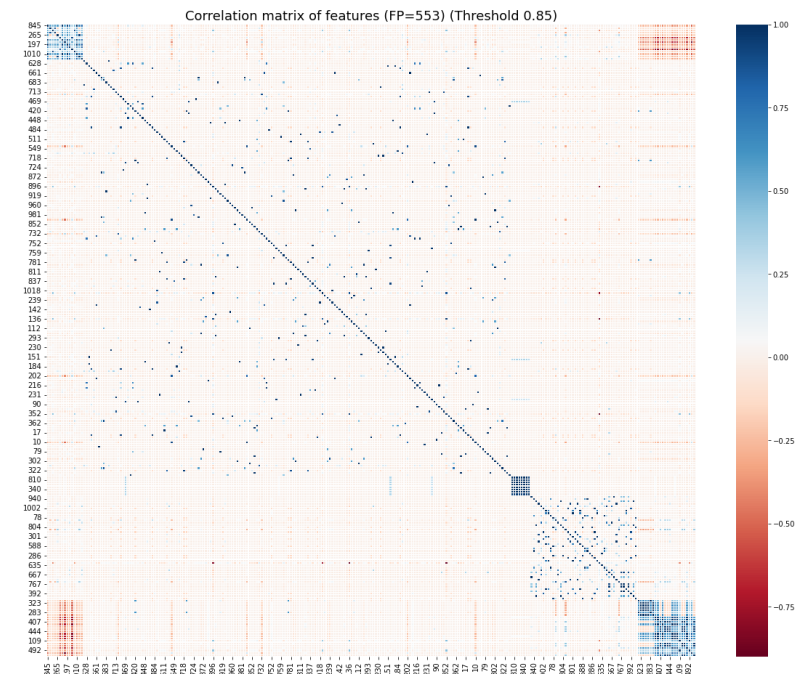
RGD Fingerprints

27 erősen korrelált pár
korr. koeff. > 0.85



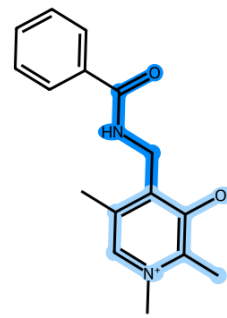
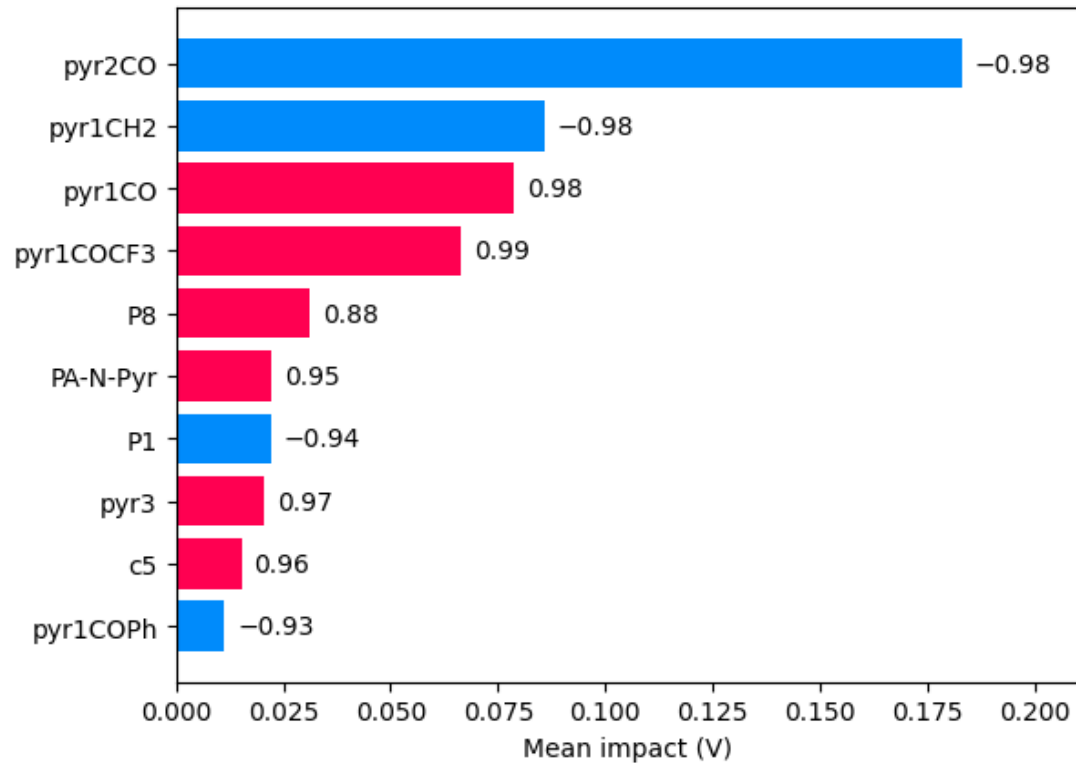
Morgan Fingerprints

300 erősen korrelált pár
korr. koeff. > 0.85

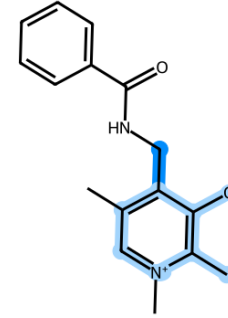


Értelmezhetőség az RGD-vel: SHAP analízis

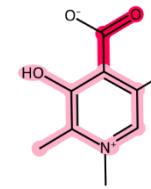
Globális jellemzők fontossága



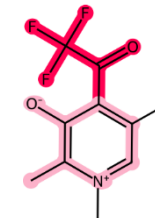
pyr2CO



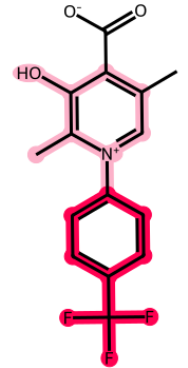
pyr1CH2



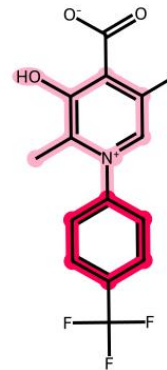
pyr1CO



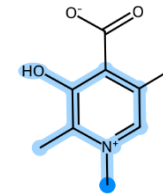
pyr1COCF3



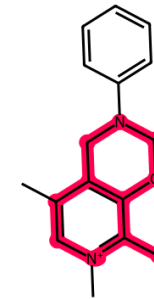
P8



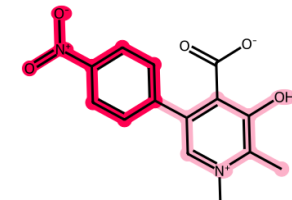
PA-N-Pyr



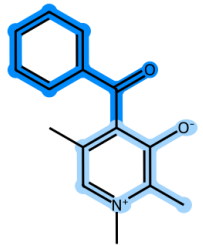
P1



pyr3

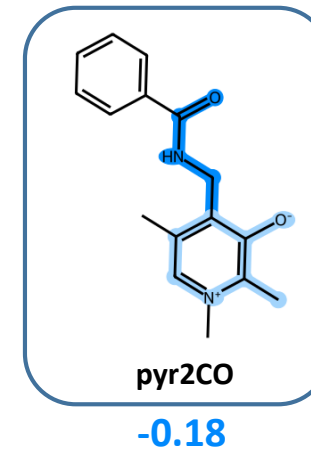
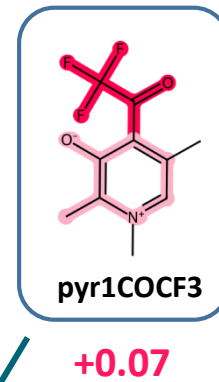
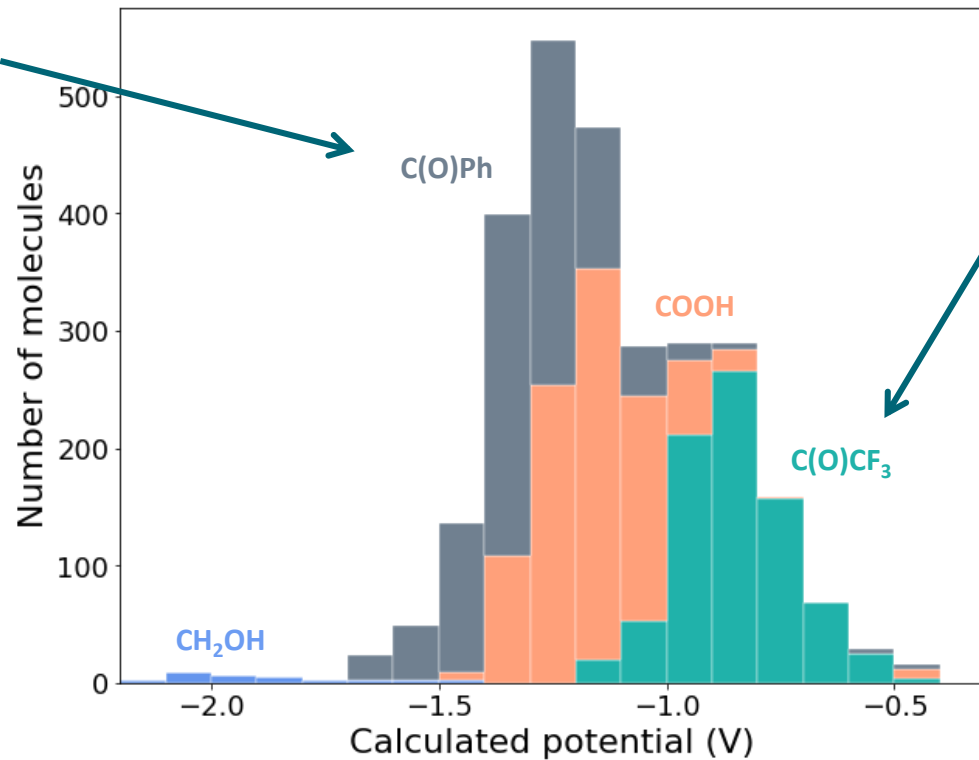
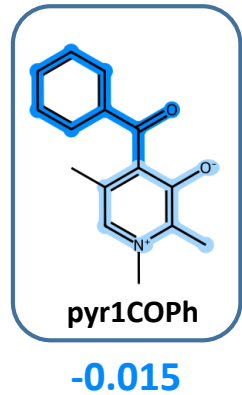


c5

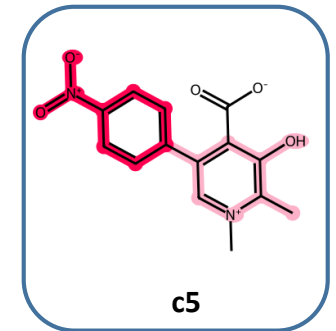


pyr1COPh

Értelmezhetőség az RGD-vel: SHAP analízis



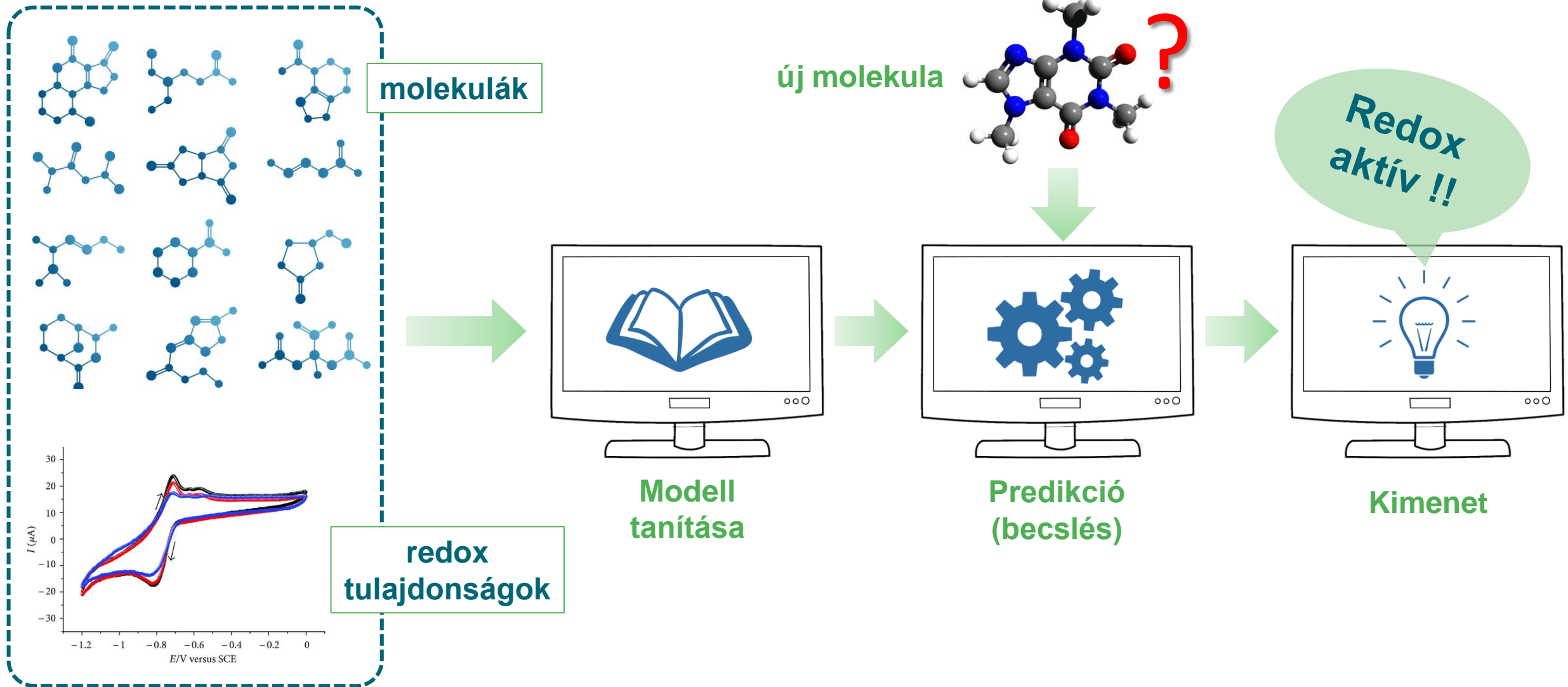
POT(average): -2.0 V



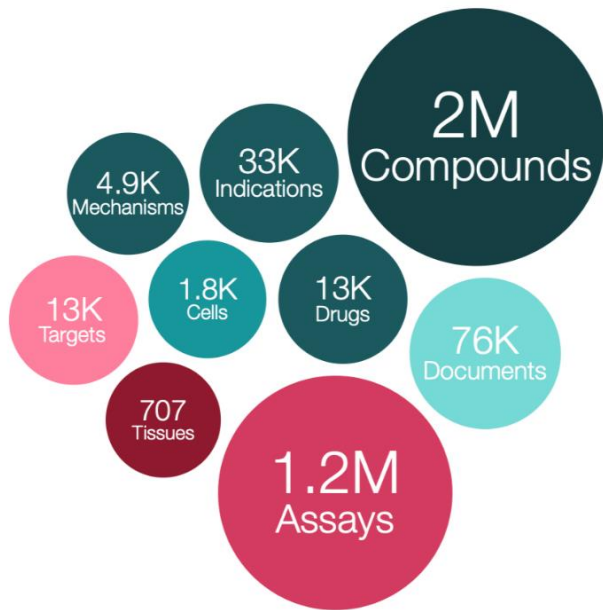
56 molecules

POT(average): -0.5 V

Gépi tanulás algoritmus



RP-ChEMBL adatbázis



Válogatás

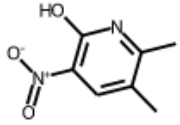
M < 200, nehéz atomok > 6
nagy diverzitás

Comp² protocol

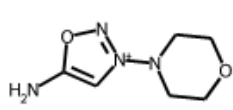
RP-ChEMBL adatbázis
(8000 molekula)

SMILES
(molekula repr.) + **1e⁻** redukciós
potenciál
(számolt)

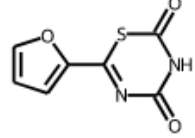
RP-ChEMBL molekula példák



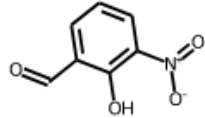
-0.63 V



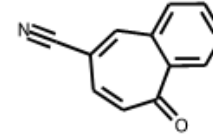
-0.69 V



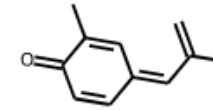
-0.84 V



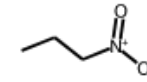
-0.9 V



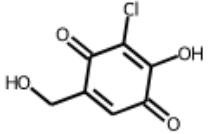
-0.95 V



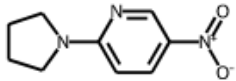
-0.77 V



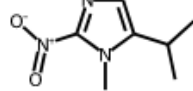
-0.91 V



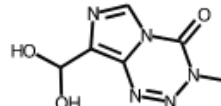
-0.53 V



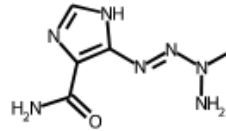
-0.71 V



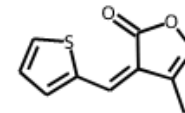
-0.54 V



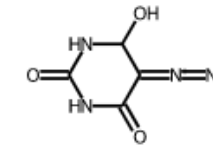
-0.69 V



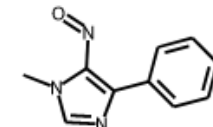
-0.88 V



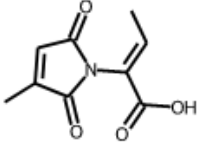
-0.7 V



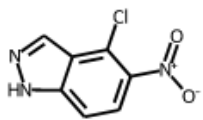
-0.97 V



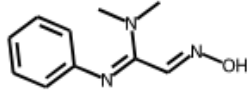
-0.64 V



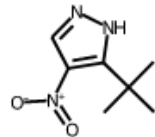
-0.91 V



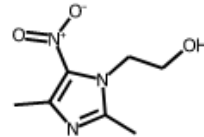
-0.54 V



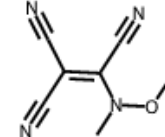
-0.83 V



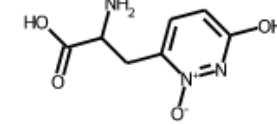
-0.76 V



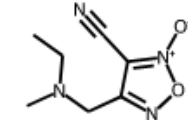
-0.72 V



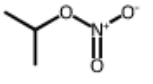
-0.65 V



-0.97 V



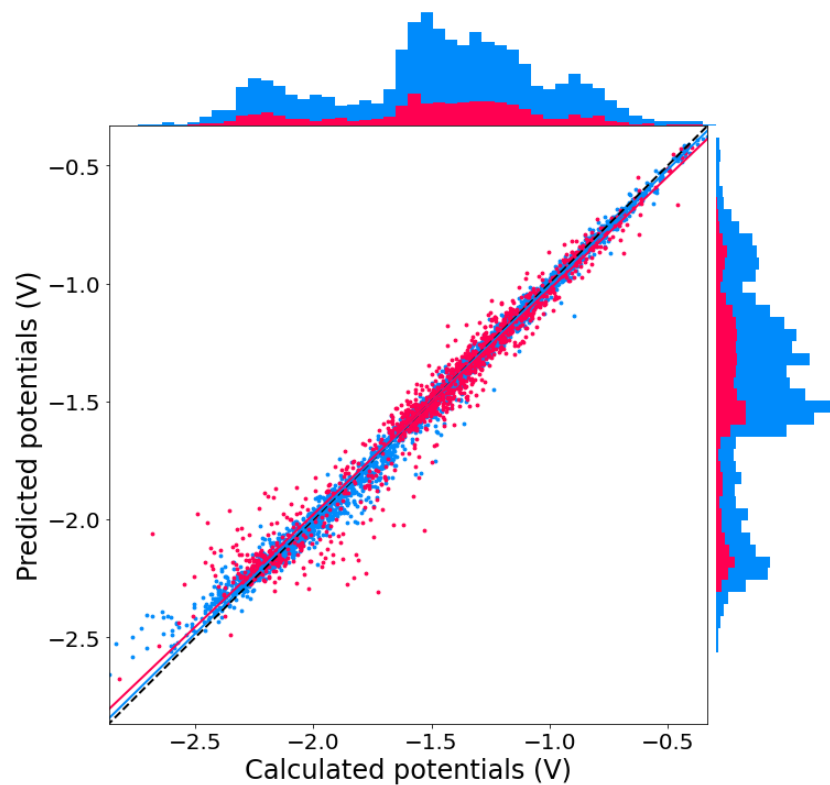
-0.76 V



-0.74 V

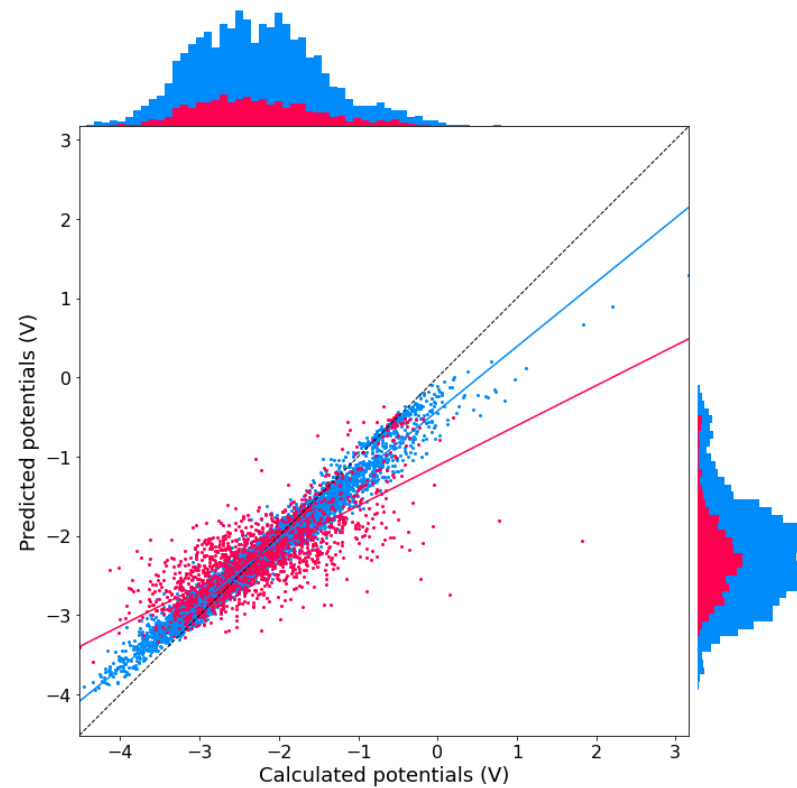
Random forest

B6-PYR



	MAE	RMSD	R ²
train	0.208	0.0342	0.9942
test	0.0566	0.0882	0.9578

RP-ChEMBL



	MAE	RMSD	R ²
train	0.148	0.198	0.962
test	0.378	0.510	0.579

Összegzés

„Fekete doboz” (HTS)

Tetszőleges molekula
redukciós potenciáljának
gyors becslése

SMILES kód

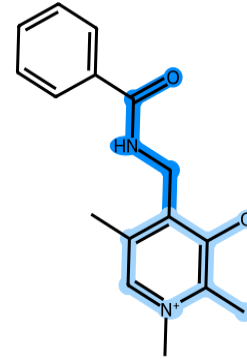


$1e^-$ redukciós potenciál
vizes közegben

„Értelmezhetőség”

pyr2CO

-0.18



Átlagos
potenciál: -2.0 V

Tervezett kiterjesztések:

- Nem csak vizes közegű potenciálok
- $2e^-$ potenciálok
- Hidrid affinitás

Résztvevők



Elméleti Kémiai Csoport

Imre Pápai
Madarász Ádám
Németh Flóra



Automatizálási és Alkalmazott Informatikai Tanszék

Szegletes Luca
Mátyás Pelle
Tugyi Beatrix
El-Ali Maya
Marcell Mészáros
Marcell Csikós



H-2020



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 875565. The content in this document represents the views of the authors, and the European Commission has no liability in respect of the content.



Köszönöm a figyelmet!

