



ChemAxon's 7th European User Group Meeting

Budapest, May 17th 2011

Ahmed Abdelaziz

eADMET GmbH



- GmbH since Sep. 2010,
Plan: 7 People and first sales in 2011
- Background: Helmholtz Zentrum München
- Business model : software licensing and integration services, development of models (consultancy), training, tailoring of software to company needs
- USPs
 - Reliable compound property predictions
 - Web based, integrated system for R&D and risk assessment, tailored solutions
- License agreement (software, IP and right to opt)
 - Cooperation agreement with the HMGU

About OCHEM

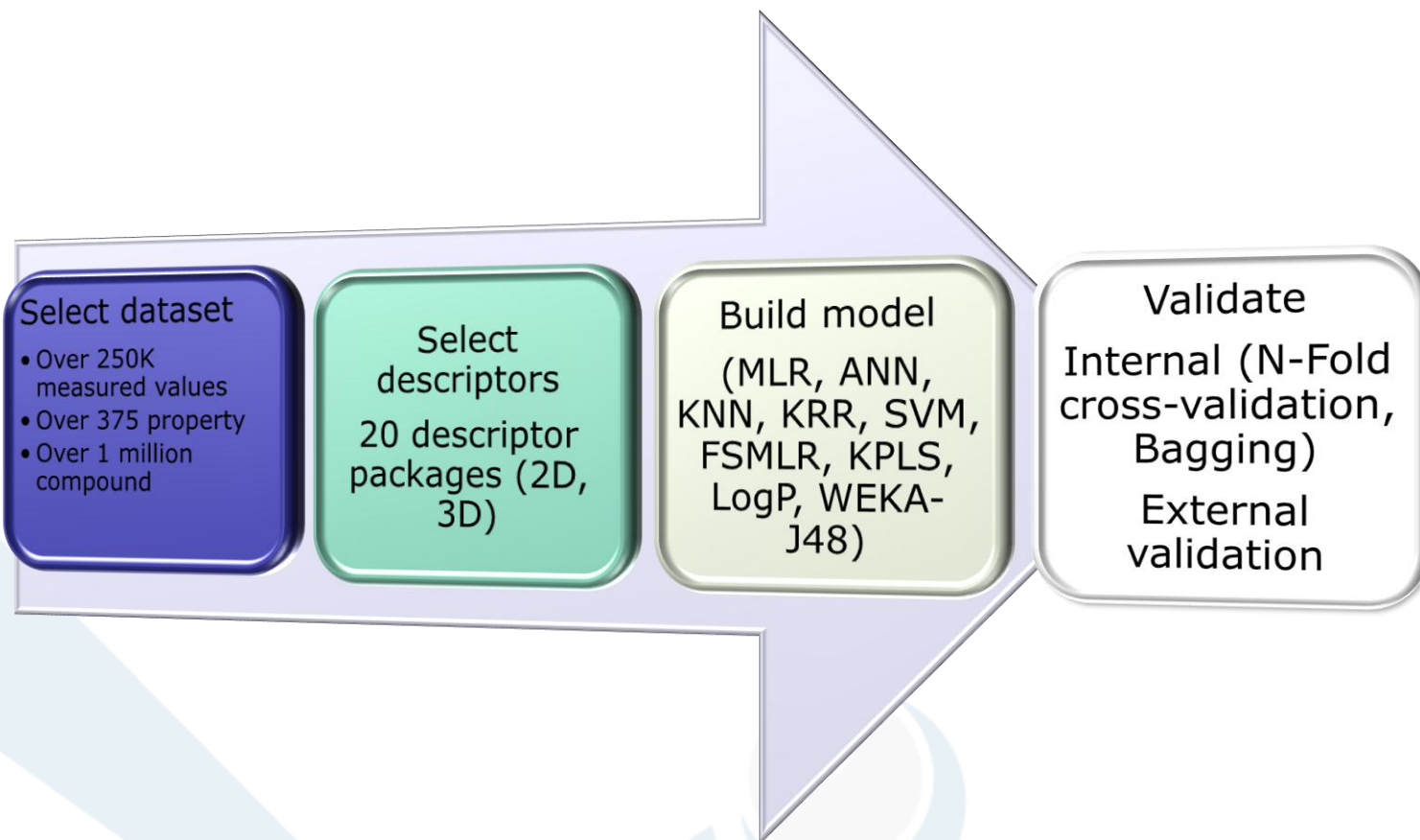
Strong Foundation: More than 3 years of development supported by the German government

Long Track Record: (Igor Tetko) the creator of VCCLAB

Free: for Academic use

High Sustainability: eADMET is committed to the maintenance of both free and commercial versions of the platform

Modeling Workflow



Molecules preprocessing by chemaxon

- Aromatize/dearomatize structures for compatibility with certain descriptors
- Standardize molecular structure by given molecular templates
- Remove counter ions prior to descriptor calculation
- Neutralize compounds

Chemaxon descriptors

- 77 descriptors integrated
- Descriptors are grouped into 7 categories: Elemental Analysis, Charge, Geometry, Partitioning, Protonation, Isomers, Others
- User has the ability to specify a pH value or range of values.
- Descriptor (and model) calculations are distributed on more than 100 calculation servers
- User can set timeout for single molecule processing
- Ring count descriptors are calculated for 3-, 4-, 5-, 6-, 7- and 8-member rings

Post processing of descriptors

- Eliminate descriptors with less than (x) unique values
- Group descriptors, that have correlation factor more than certain value
- Perform principal component analysis
- After filtering, user can also select necessary descriptors manually

Model statistics

Overview

Applicability domain

Model name: log(IGC50-1), 8381

Public ID is [4124075](#)

Predicted property: [log\(IGC50-1\)](#)

Training method: ANN

measured in $-\log(\text{mmol/L})$

[ChemaxonDescriptors (0.0)]

Correl. limit: 0.95

Supersab, 1000 iterations, 3 neurons

ensemble=64

5-fold cross-validation

418 filtered descriptors

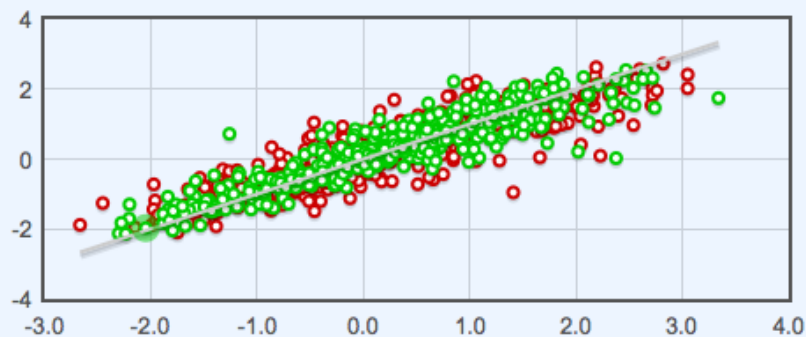
Supersab, 1000 iterations, 3 neurons

ensemble=64

Calculated in 530 seconds

Size: 750 Kb

| Data Set | # | R2 | q2 | RMSE | MAE |
|---|-----------------------------|------|------|------|------|
| ● Training set: T. pyriformis train | 644 records | 0.82 | 0.82 | 0.45 | 0.32 |
| ● Test set: T. pyriformis test | 449 records | 0.83 | 0.83 | 0.43 | 0.31 |



 [Download model statistics in Excel format](#)

 [View configuration XML](#)

Applicability domain

Overview **Applicability domain**

Williams plot **W** with *ASNN-CORREL* used as a distance to model.

Distance to model

ASNN-CORREL

Averaging type

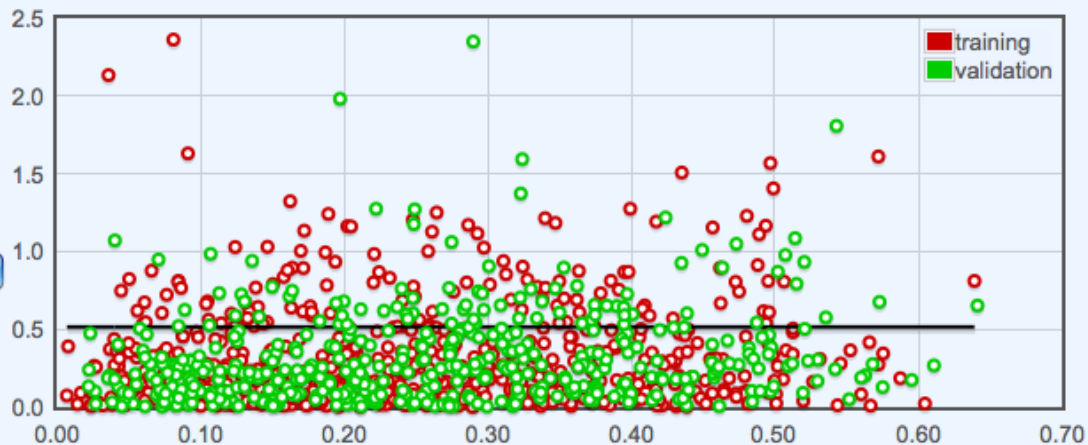
Default

X axis

ASNN-CORREL

Show negative residuals

add distance to model





Feel Free to Try it

www.ochem.eu



Thank you

www.eadmet.com