



The
University
Of
Sheffield.



Imputation of *in vitro* toxicity data

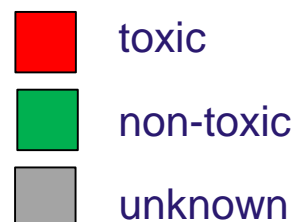
30/09/2020

Moritz Walter



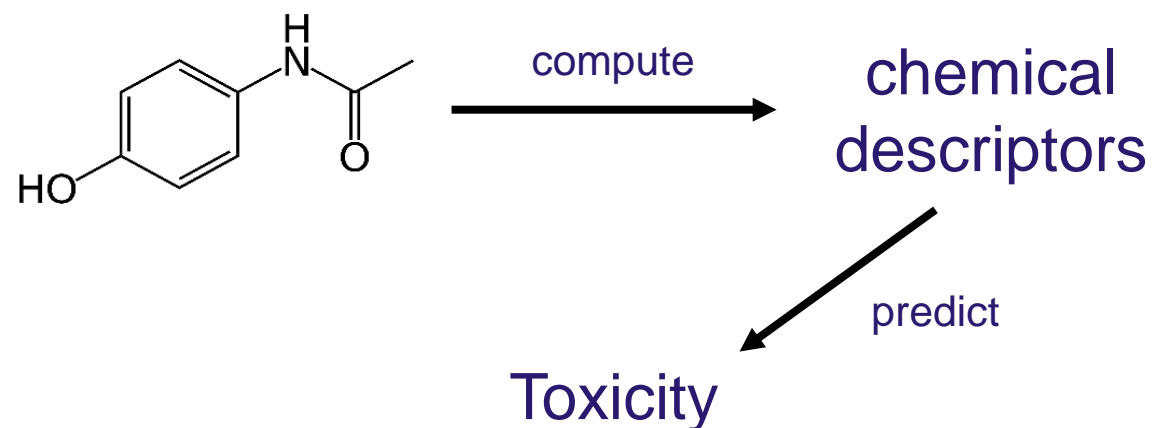
Imputation vs. QSAR prediction

| | A1 | A2 | A3 | A4 |
|----|-----------|-----------|-----------|-----------|
| C1 | toxic | unknown | unknown | non-toxic |
| C2 | unknown | toxic | unknown | toxic |
| C3 | non-toxic | toxic | non-toxic | unknown |
| C4 | unknown | non-toxic | toxic | non-toxic |
| C5 | unknown | unknown | toxic | non-toxic |



A: assay
C: compound

- Feature Net
- Macau

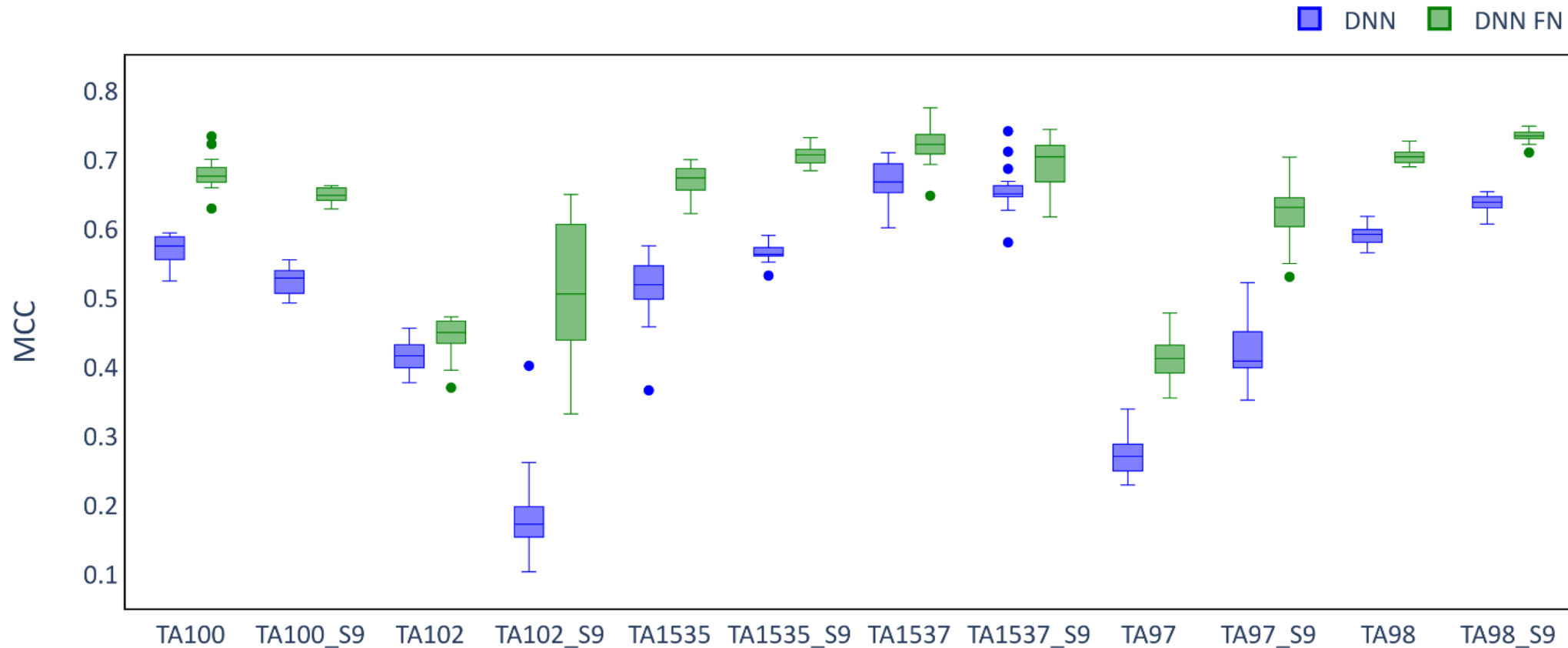


- Random Forest
- XGBoost
- Deep Neural Networks



The University Of Sheffield.

Results Imputation



Data: Ames mutagenicity test dataset (6 bacteria strains \pm S9 mix for metabolic activation)
Box: 20 independent runs with different random seed on the same test set



The
University
Of
Sheffield.



Highlights Poster

- Train/Test splitting for imputation
- Imputation model details (Feature Net + Macau)
- Results (Ames + Tox21)



The
University
Of
Sheffield.

Acknowledgement



Prof. Dr. Val Gillet
Dr. Antonio de la Vega de León



Dr. Sam Webb