NPS-MS: A Highly Accurate MS/MS Prediction Model for Novel Psychoactive Substances

Fei Wang[1,3], Daniel Pasin[7], Michael Skinnider[5], Jaanus Liigand[2,5], Leonard J. Foster[4], Petur Dalsgaard[7], Russell Greiner[1,3], David S. Wishart[1,2]

1.Department of Computing Science, 2.Department of Biological Science, University of Alberta, 3 Alberta Machine Intelligence Institute (AMII), Edmonton
4.Department of Biochemistry 5.Michael Smith Laboratories University of British Columbia
6.Institute of Chemistry, University of Tartu,Tartu, Estonia 7.Section of Forensic Chemistry, Department of Forensic Medicine,University of Copenhagen

Introduction

New psychoactive substances (NPS), also referred to as designer drugs, 'legal highs' ,or bath salts, have emerged as significant public health and drug control issues worldwide. These molecules possess the same psychoactive effects as conventional illicit drugs but with sufficiently different chemical structures that circumvent existing drug control legislation

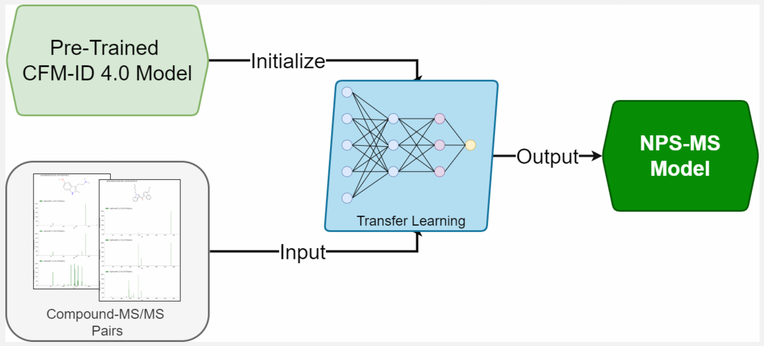
This work describes a learned model:

- It can accurately predict Electrospray Ionization Tandem Mass Spectrometry (ESI-MS/MS) for a given NPS compound structure.
- These predicted NPS ESI-MS/<S spectra enables more accurate identification than existing tools for MS/MS spectrum interpretation.

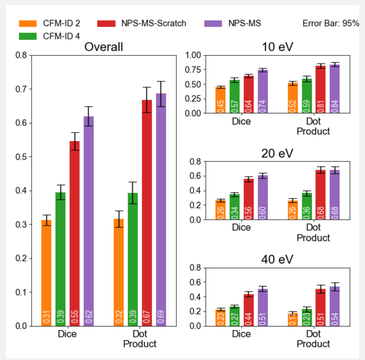
Such a model is learning from a pretrained CFM-ID 4.0 model and fine-tuned through NPS specific dataset.

NPS to Mass Spectra Prediction

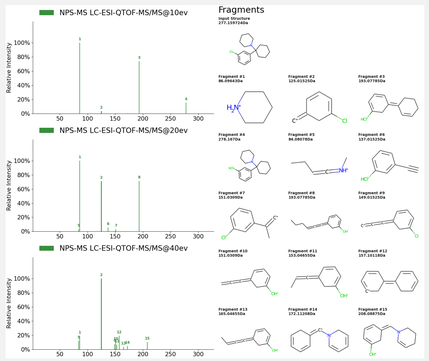
- NPS-MS model is transfer-learning from a pre-trained CFM-ID 4.0 model parameters trained on standard metabolite data and then fine-tuned on NPS specific dataset.
- Training dataset contains 494 unique NPS compounds and their associated MS/MS spectra, Hold out test set has130 unique NPS compounds and their associated MS/MS spectra.



Spectra Prediction Performance



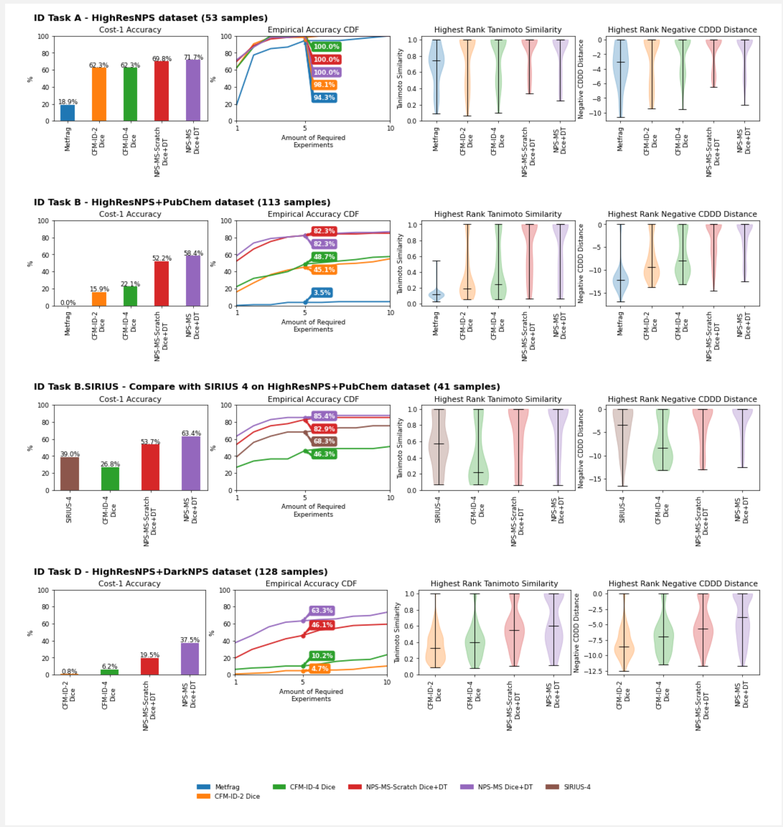
A Spectra Prediction Example（3-CI-PCP）



Mass Spectra to NPS Identification

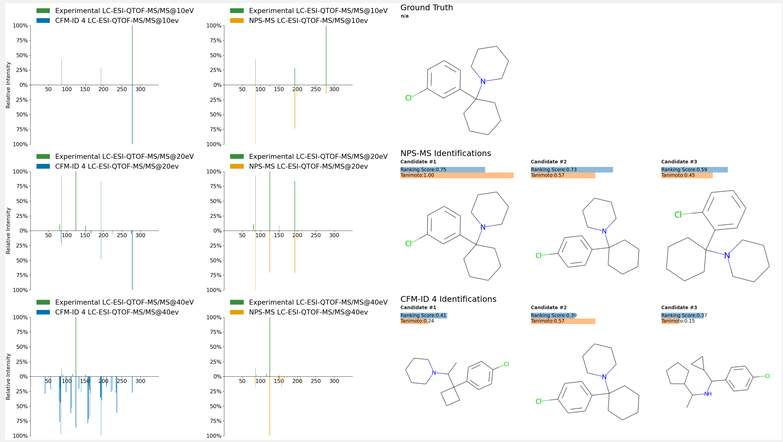
NPS Identification Workflow

- Retrieved a list of candidate compounds from the respective reference database by the precursor ion mass and a mass filter set to 10ppm.
- Ranking such candidate lists based on NPS-MS predicted spectra.



An NPS Identification Example（3-CI-PCP）

CFM-ID Predicted MS/MS Spectra | NPS-MS Predicted MS/MS Spectra | Identification Rankings



Conclusions

- NPS-MS can accurately predict ESI-QToF-MS/MS from NPS compound structure at 10eV, 20eV and 40eV.
- By leveraging transfer learning from a more generic model, NPS-MS overperformed CFM-ID 2.0, CFM-ID 4.0 and NPS-MS-Scratch model in the NPS to mass spectra prediction tasks.
- NPS-MS also overperformed CFM-ID 2.0,CFM-ID 4.0, Metfrag, and SIRIUS 4 in mass spectra to NPS identification tasks..
- This opens the possibility of screening for emerging NPSs, for which reference mass spectra may not yet be available. In addition, such a transfer learning approach could be also be used to train other domain specific mass spectra prediction model.

Acknowledgements

UNIVERSITY OF ALBERTA
TMIC The Metabolomics Innovation Centre
THE UNIVERSITY OF BRITISH COLUMBIA
UNIVERSITY OF COPENHAGEN
CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada
amii
NSERC CRSNG
GenomeCanada
NIH National Institutes of Health
INNOVATION.CA CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION