

Massenspektrometrische Interpretation mit maschinellem Lernen

Peter Rösner

DigiLab Software GmbH, 24161 Altenholz, Posener Straße 18; p.roesner@digilab-software.de

1. Einführung

Die Methoden der künstlichen Intelligenz gehören zu den erfolgversprechendsten Technologien der Gegenwart. Ein Teilgebiet der künstlichen Intelligenz ist das maschinelle Lernen. Hierunter versteht man Algorithmen, die in der Lage sind, in großen Datenmengen Muster und Zusammenhänge zu erkennen, ohne explizit dafür programmiert zu sein. Bekannte Beispiele maschinellen Lernens sind etwa die Schrift-, Stimm- und Gesichtserkennung, die Früherkennung von Krebs bei bildgebenden Verfahren und beeindruckende Chatbots.

In der forensischen Analytik wurden diese leistungsfähigen Methoden bislang lediglich in speziellen Fällen eingesetzt [1]. So berichten Koshute et al. [2] über die Extraktion von 24 massenspektrometrischen Peaks, die für ein maschinelles Lernen zur Identifikation Fentanyl-analoger Verbindungen eingesetzt wurden. Bonetti et al. [3] verwendeten maschinelles Lernen zur Differenzierung von Positionsisomeren psychoaktiver Verbindungen mit Direct Analysis in Real Time Mass Spectrometry (DART MS).

Das Peakmuster von Elektronenstoß (EI)-Massenspektren reflektiert strukturelle Eigenschaften des Moleküls und seiner Fragmente. In gewissem Rahmen erlaubt eine sorgfältige Analyse solcher Massenspektren daher eine Ermittlung von Partialstrukturen und in günstigen Fällen sogar der Struktur des Moleküls selbst. Dieser Prozess der massenspektrometrischen Interpretation ist zeitintensiv und erfordert umfangreiches Fachwissen über das Fragmentierungsverhalten von Molekülen. Ein einzelner Experte wird daher im Laufe seines Lebens nur einen Bruchteil der vielen unbekanntem Massenspektren interpretieren können, die bei der täglichen Arbeit in analytischen Laboratorien anfallen.

Es lag daher nahe, zu prüfen, ob maschinelles Lernen eine erfolgreiche Methode sein könnte, charakteristische Muster und Zusammenhänge in einer großen Mengen von Massenspektren automatisch zu erkennen und strukturelle Vorhersagen zu treffen, ohne zeitintensive massenspektrometrische Interpretationen oder eine Strukturanalyse durchzuführen zu müssen.

Der hier vorgestellte Ansatz ist allgemeinerer Natur als die oben erwähnten Verfahren, da er die originalen Daten der Massenspektren selbst verwendet und gestattet, eine Vielzahl von Strukturklassen und Partialstrukturen zu identifizieren.

2. Maschinelles Lernen zur Substanzklassenidentifikation

Ein einfaches Modell zur Lösung von Entscheidungsproblemen ist der binärer Entscheidungsbaum (Abb. 1). Er trifft an jedem Knoten eine ja/nein-Entscheidung bis das Resultat (Blatt) erreicht ist. Die Entscheidungsregel ist im Fall zweier Massenspektren der Vergleich ihrer Intensitäten. Binäre Entscheidungsbäume besitzen jedoch eine Reihe von Nachteilen wie eine hohe Streuung (Varianz) und ein hohes Risiko von Überanpassungen, die die Vorhersagegenauigkeit (Generalisierung) mindern.

Der im Jahr 2001 von Leo Breiman und Adele Cutler [4] eingeführte maschinelle Lernalgorithmus „Random Forest“ vermeidet viele diese Nachteile, indem er eine Reihe von binären Entscheidungsbäumen verwendet. Deren Vorhersagen werden bei diesem Modell gemittelt und

verglichen, um die endgültige Vorhersage zu erhalten (Abb. 1). Hierdurch wird die Genauigkeit und die Robustheit des maschinellen Lernmodells entscheidend verbessert.

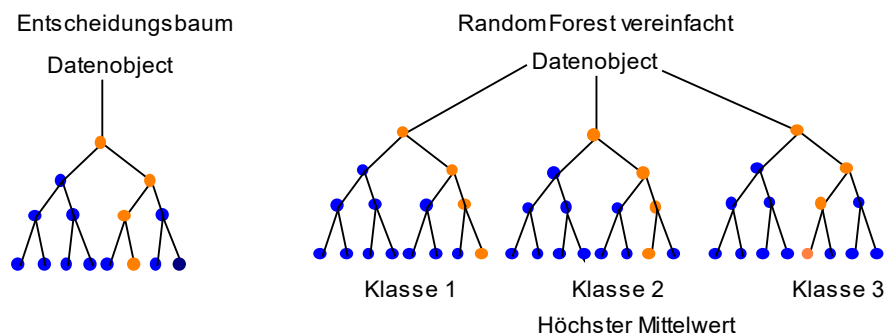


Abb. 1. Einfacher Entscheidungsbaum und Kombination dreier Entscheidungsbäume unter „Random Forest“.

„Random Forest“ ist eines der am meisten verwendeten Modelle des maschinell überwachten Lernens. Hierunter versteht man Algorithmen, die in der Lage sind, allein aus bekannten Eingabe- und Ausgabepaaren richtige Vorhersagen über die Eigenschaften einer Beobachtung zu machen. Überwacht bedeutet in diesem Zusammenhang die Markierung der Ausgabe für eine bestimmte Eingabe. Die Vorhersage von Eigenschaften einer Beobachtung wird Generalisierung genannt. „Random Forest“ besitzt einen hohen Generalisierungsgrad, d. h. er kann gut neue Daten voraussagen, die er vorher im Training nicht gesehen hat. Im Fall von EI-Massenspektren besitzt der Algorithmus daher eine hohe Fähigkeit, die Strukturklasse eines unbekanntes Spektrums vorherzusagen.

Massenspektrensammlungen sind eine hervorragende Grundlage für den Einsatz von Modellen des maschinell überwachten Lernens. Man kennt die Eingabedaten in Form der Intensitäten der Masse/Ladungs-Verhältnisse (m/z) der Massenspektren und die Ausgabe in Form der Namen der Verbindungen. Damit sind die Voraussetzungen für den Einsatz des maschinell überwachten Lernens erfüllt.

Große Massenspektrensammlungen enthalten implizit eine enorme Menge an Expertenwissen, dass durch maschinelles Lernen zur Vorhersage der Struktur unbekannter Massenspektren genutzt werden kann. Hierzu bedarf es keiner Beteiligung eines Menschen mehr. Auf der Basis von 33540 EI-Massenspektren der Massenspektrensammlung „Designer Drugs 2023“ [5] wurde das Modell „Random Forest“ mit über 100 verschiedenen Partialstrukturen und Strukturklassen wie Fentanylen, Cannabinoiden, Morphinen und Amphetaminen trainiert (Tab. 1).

Begriffe, die Wirkungen an biologischen Rezeptoren beschreiben, wie Fentanyle oder Cannabinoide, umfassen verschiedene chemische Strukturklassen. Daher wurde das Modell ebenfalls mit den zugrunde liegenden separaten Strukturklassen dieser übergeordneten Begriffe trainiert. Im Fall der Fentanyle wurden etwa die Strukturklassen der Fluorofentanyle, Thienylfentanyle, Furanylfentanyle, Benzylfentanyle und alpha-Methylfentanyle separat trainiert.

2.1. Leistungsfähigkeit und Grenzen des Verfahrens

Mit dem trainierten Modell „Random Forest“ ergaben sich bei der Vorhersage definierter Strukturklassen unbekannter Verbindungen in der Regel Genauigkeiten, die vom idealen statistischen Wert 1 nur gering abwichen (Tabelle 2). Das gleiche gilt für die Werte der Kreuzvalidierung.

Bei jedem Training eines Modells wurde eine zehnfache Kreuzvalidierung durchgeführt, um das Ausmaß der Vorhersagefähigkeit des Modells zu überprüfen.

Tabelle 1. Strukturklassen und Anzahl ihrer Massenspektren (MS) in „Designer Drugs 2023“.

Strukturklasse	Anzahl MS	Strukturklasse	Anzahl MS
Acetamide-compounds	437	Esters	2581
Acylpiperazine-opioids	6	Fentanyles	1699
Alcohols	31	Fluorobenzenes	3075
Alkanes	85	Fluorobenzyl	1463
Alkylaromatics	51	Fluorofentanyles	907
Alkylphenols	29	FuranylFentanyles	265
Alcohols	31	HalogenoBenzenes	43
Alkanes	85	Hydrocabons	401
Alkylaromatics	51	Indane-2-amines	2193
Alkylphenols	29	Indazolecannabinoids	147
Alkylamines	89	Indolecarboxamidecannabinoids	599
Aminocyclohexanol-opioids	10	Indoles	2061
Amphetaminehomologs	83	Iodobenzenes	1641
Amphetamines	1715	Iodobenzyl	1113
Aromaticmonocyclic	7191	Isatinhydrazone-cannabinoids	29
Aromatic-compounds	5029	Ketamines	21
AroylIndoleCannabinoids	1417	Ketones	53
Arylbutan-2-amines	1701	Methoxyindoles	837
Arylcyclohexylamines	49	Methylenedioxy-compounds	147
Arylpropan-2-amines	1217	Methylphenyl-compounds	1701
Azaindolcarboxamides	33	Morphines	45
Benzaldehydes	809	N1-compounds	14883
Benzamide-opioids	809	N2-compounds	9085
Benzimidazole-cannabinoids	51	N3-compounds	743
Benzodiazepines	105	N-AlkylPiperidine	79
Benzoyl-compounds	895	Naphthalenes	115
Benzylfentanyles	771	Naphthoyl-compounds	197
Benzyl-compounds	4419	Nitazene-opioids	37
Benzylamines	663	Nitro	4001
BiAromatic-compounds	663	P1-compounds	9223
Br1-compounds	19475	Phencyclidines	15
Br2-compounds	207	PhenethylFentanyles	751
Br3-compounds	283	Phenethylamines	1091
Bromobenzenes	1287	Phenothiazine	99
Bromobenzyl	931	Phenyl-compounds	12335
Cannabinoids	1271	Phenylaminofentanyles	3463
Carbolines	31	Phthalate-compounds	33
Carbonicacids	111	Phytocannabinoids	1431
Carbonylindoles	271	Piperazine-compounds	721
Cathinones	15679	Piperidine-opioids	19
Chlorobenzenes	195	Pyridoyl-compounds	12383
Chlorobenzyl	505	S1-compounds	1969
Cl1-compounds	18061	Steroids	259
Cl2-compounds	565	TMS1-compounds	4147
Cl3-compounds	57	Thienylfentanyles	435
Coumarins	43	Trifluoroacetyl	739
Cycloaliphatic-compounds	17	Trimethoxyphenyl-compounds	3137
Diazepines	1009	Trimethylsilyl-compounds	2661
Diazinane-compounds	793	Tropines	355
Dimethoxyphenyl-compounds	2799	Tryptamines	489
Ergolines	1467	3-Amidoindazole-cannabinoids	2083
alpha-MethylFentanyles	275	tert-Butyl	1329

Die hierbei ermittelten Abweichungen vom statistisch maximalen Wert von 1 lagen bei den strukturell definierten Strukturklassen im Promillebereich. Partialstrukturen wie Phenyl sind in einer Vielzahl verschiedener Strukturklassen enthalten. Der für das Training verwendete Phenyl Datensatz war daher sehr heterogen. Diese Heterogenität führte zu einer Minderung der Genauigkeit und Vorhersagefähigkeit, die zu Abweichungen vom statistisch maximalen Wert im Prozentbereich führte (Tabelle 2). Diese Minderung der Genauigkeiten und Fähigkeiten auf neue, unbekannte Daten zu verallgemeinern, wurde auch bei anderen Partialstrukturen beobachtet. Die insbesondere bei strukturell eng definierten Strukturklassen hohen Werte der Kreuzvalidierung belegen die hohe Generalisierungsfähigkeit des Modells.

Tab. 2. Strukturklassen mit Genauigkeiten und Werte der Kreuzvalidierungen.

Strukturklasse	Genauigkeit	Kreuzvalidierung
Fentanyl	0.996	0.998
Benzylfentanyl	0.997	0.999
Furanylfentanyl	1.000	0.999
Thienylfentanyl	1.000	0.999
Fluorofentanyl	0.999	0.999
alpha-Methylfentanyl	1.000	0.999
Phenylaminofentanyl	0.997	0.998
Morphine	0.997	0.999
Phytocannabinoide	0.999	1.000
Cannabinoide	0.997	1.000
Arolyndolcannabinoide	0.990	1.000
Indazolcannabinoide	0.997	0.999
Indoles	0.998	0.997
Amphetamine	0.989	0.994
Phenethylamine	0.997	0.996
Ketamine	0.999	1.000
Ergoline	0.999	1.000
Phenethylfentanyl	0.999	0.999
Phenyl-Compounds	0.940	0.957
Aromatic-Compounds	0.966	0.973
Aromaticmonocyclic	0.946	0.958
N1-Compounds	0.938	0.923
N2-Compounds	0.984	0.983

Das Ionenmuster von EI-Massenspektren reflektiert nicht in jedem Fall alle strukturellen Merkmale eines Moleküls. Einige Verbindungen erleiden schnelle Fragmentierungen, wobei große Molekülteile als Neutralteilchen abgespalten werden können. Diese nichtionischen Fragmente werden im Massenspektrum nicht dargestellt und entziehen sich somit einer Identifikation. In diesen Fällen kann auch die strukturelle Vorhersage des Modells fehlschlagen, da sich die Strukturvorhersagen des Algorithmus lediglich auf die im MS verbleibenden Kationen stützen können. Ein Beispiel sind langkettige verzweigte Brom- oder Iodalkane, die ihre Halogenatome durch eine schnelle induktive-Spaltung verlieren. Das nach der Fragmentierung verbleibende Alkylkation kann dann nur noch die typischen Fragmente eines Alkans zeigen (Abb. 2).

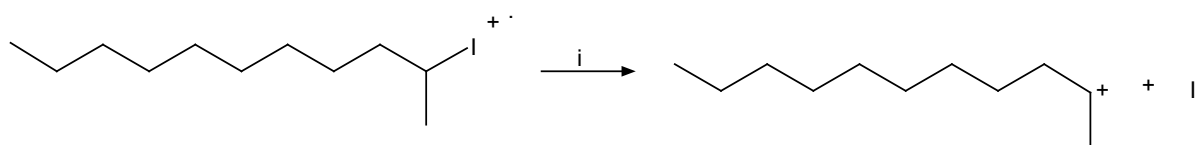


Abb. 2. Induktiver Bindungsbruch bei verzweigten langkettigen Iodalkanen.

Dies gilt in analoger Weise auch für eine Reihe anderer Substanzklassen. In der Praxis muss man daher mit fehlerhaften Voraussagen rechnen. Sie können höher sein als die über das Modell errechneten statistischen Werte. Diese Abweichungen sind in der Begrenztheit der Methode der Elektronenstoß-Massenspektrometrie und nicht im Algorithmus des Modells begründet.

Je mehr Massenspektren einer Strukturklasse im Datensatz existieren, desto bessere Voraussagen wird das Modell machen. So weist die Gruppe der Fentanyle Massenspektren von über 1600 verschiedenen Verbindungen auf und entsprechend zahlreich werden die Voraussagen des Algorithmus in dieser Strukturklasse sein.

Dies gilt auch für die Gruppe der Cannabinoide mit 1272 Einträgen und die Gruppe der Amphetamine mit 1715 Einträgen. Neben der Güte der Massenspektren bestimmt insbesondere der Typ der Datenbasis und dessen Größe die Vorhersagefähigkeit eines Modells. Ein Modell, trainiert mit einer spezialisierten Massenspektrensammlung von Drogen, wird naturgemäß bessere Vorhersagen für die Massenspektren unbekannter Drogen machen, als eine Massenspektrensammlung, die überwiegend Arzneimittelwirkstoffe repräsentiert.

Wie weit darf das Massenspektrum einer unbekanntes Verbindung von allen Einträgen des Trainingsdatensatzes (70 % des Gesamtdatensatzes) abweichen, um vom Algorithmus erkannt zu werden? Die Einführung eines Heteroatoms in ein Molekül führt im Regelfall zu einer drastischen Änderung des Fragmentierungsverhaltens. Heteroatome mit ihren freien Elektronenpaaren ermöglichen zusätzliche alpha-Spaltungen, die den Charakter eines Massenspektrums erheblich verändern können. Ein unbekanntes Massenspektrum, dessen Strukturklasse ein zusätzliches Heteroatom aufweist, kann daher die Vorhersagefähigkeit des Modells überfordern. Die Einführung eines zusätzlichen C-Atoms wird dagegen zu einer eher moderaten Änderung des Fragmentierungsmusters führen. Die Aussicht, solche unbekanntes Verbindungen über das Massenspektrum korrekt vorherzusagen wird damit größer.

2.2. Bedeutung für die Praxis

Das Modell „Random Forest“, trainiert mit leistungsfähigen Massenspektrenbibliotheken, ist in der Lage, im Sekundenbereich Hinweise auf die Strukturklasse eines unbekanntes Massenspektrums zu geben. Hierfür wurde der Algorithmus mit geeigneten Parametern so trainiert, dass die Anzahl falsch-positiver Voraussagen für definierte Strukturklassen in der Regel im Promillebereich lag. Die hohe Genauigkeit in vielen vermessenen Strukturklassen zeigt, dass maschinelles Lernen das Potential hat, eine nützliche Methode zur strukturellen Klassifikation unbekanntes Massenspektren zu werden.

Die Vorhersage der Strukturklasse eines unbekanntes Massenspektrums ermöglicht dem Analytiker, seinen Arbeitsaufwand zu reduzieren, indem er auf eine vertiefte Strukturauflärung verzichten kann, wenn das Programm eine irrelevante Substanzklasse indiziert. Gibt das Programm dagegen einen Hinweis auf eine relevante Drogenkategorie, ist eine weitere Strukturauflärung erforderlich. „Random Forest“ eignet sich auch, um zu überprüfen, ob eine durch Bibliothekssuche zugeordnete Struktur in das Raster der entsprechenden Strukturklasse fällt. Auf diese Weise ergibt sich ein zusätzlicher Beitrag zur Strukturabsicherung.

Das auf dem Modell „Random Forest“ basierende Programm „Mass Spectral AI Interpreter“ ergänzt die übliche Bibliothekssuche, indem es die nicht identifizierbaren, unbekanntes Massenspektren klassifiziert. Sie bietet selbst dann noch eine Hilfe, wenn alle Möglichkeiten einer spektrenbasierten Bibliothekssuche ausgeschöpft sind. Die Fähigkeit des Modells, auch Hinweise auf die Anzahl von Heteroatomen und andere Partialstrukturen zu geben, geht dabei über den einfachen Vorschlag der Strukturklasse hinaus. So wird etwa ein unbekanntes bromhaltiges Fentanyl neben der Klassifikation Phenethylfentanyl die Zusatzinformationen Bromobenzene, Br1- und N2-compounds aufweisen.

Die Qualität einer statistischen Vorhersage einer Strukturklasse hängt auch vom Informationsreichtum der Spektren der verwendeten Spektrenbibliothek ab. Elektronenstoß-Massespektren weisen einen weit größeren Fragmentierungsreichtum auf, als Flüssigkeitschromatographie-Massespektrometrie (LC-MS)-Spektren. Daher bieten mit LC-MS-Spektren trainierte Modelle geringere Aussichten einer guten Strukturklassenvorhersage. Den Vorteil einer Substanzidentifikation wie bei einer „klassischen“ Bibliotheksuche bieten auf statistischen Verfahren beruhende Algorithmen wie „Random Forest“ dagegen nicht.

2.3. Programmtechnische Realisierung

Die Realisierung des Modells „Random Forest“ erfolgte in der Programmiersprache „Python“. Diese und die Programmierumgebung „Pycharm“ sind kostenlos im Internet zugänglich [6,7]. Dies gilt auch für umfangreiche Statistikpakete, die auch leistungsfähige Modelle des maschinellen Lernens zur Verfügung stellen [8].

Als Zeilen der Eingangsmatrix X für das Modell dienten die ganzzahligen originalen Intensitätswerte der Massenspektren, die auf den m/z -Bereich 30-800 normiert wurden. Die untere Grenze von m/z 30 erlaubt noch, die intensiven Immoniumionen der Phenethylamine zu registrieren. Diese sind wichtige Schlüsselionen der Amphetamine und Phenethylamine. Die obere Grenze von m/z 800 wurde gewählt, um auch die Massenspektren von schwereren bromhaltigen Verbindungen erfassen zu können.

Als Zeilenwerte des Ausgangsvektors Y dienten die Label 0 und 1, um ein einfaches binäres Klassifikationsmodell implementieren zu können. Jeder Zeile der Eingangsmatrix X wurde ein Zeilenwert des Y -Vektors zugeordnet, je nachdem ob die Intensitätszeile der Eingangsmatrix X zur untersuchenden Strukturklasse gehört (1) oder nicht (0). Die Matrizes X mit den korrelierten Y -Vektoren wurden dem Modell „Random Forest“ übergeben. Hierbei wurde der gesamte Datensatz zunächst in einen Trainingsdatensatz (70 %) und einen Testdatensatz (30 %) aufgeteilt. Mit dem Trainingsdatensatz wurde das Modell trainiert. Die Teilmenge des Testdatensatzes diente dagegen zur Vorhersage der Leistung des Modells. Dies betrifft die Vorhersage statistischer Werte, wie der Genauigkeit und Präzision, die die Verlässlichkeit der Daten kennzeichnen. Die Werte der Kreuzvalidierung geben dagegen Auskunft darüber, wie gut das Modell in der Lage ist, die Strukturklasse eines unbekanntes Massespektrums vorherzusagen.

Beim Training des Modells wurden die Parameter so gewählt, dass die Anzahl falsch-positiver Vorhersagen im Promillebereich blieb (Präzision > 0.99). Lediglich im Fall von Partialstrukturen mit ihren zahlreichen heterogenen Strukturklassen wurden falsch-positive Vorhersagen im Prozentbereich zugelassen (Präzision > 0.9). Die Anzahl der falsch-negativen Vorhersagen lag ebenfalls in diesen Bereichen.

Für das Training wurden folgende Hyperparameter (Parameter des Modells) und Techniken verwendet:

1. Die Anzahl der Entscheidungsbäume ($n_estimators$) des Modells „Random Forest“ wurde auf 200 festgelegt, um gute Ergebnisse in akzeptablen Rechenzeiten zu erhalten. Je mehr Entscheidungsbäume verwendet werden, desto genauer sind die Werte des Modells aber die Rechenzeit nimmt dabei stark zu. Die gewählte Anzahl der Entscheidungsbäume stellt daher einen Kompromiss zwischen Genauigkeit und akzeptabler Rechenzeit dar.
2. Der Parameter $max_features$ bestimmt die maximale Anzahl der Merkmale, die bei der Aufteilung eines Knotens in einem Entscheidungsbaum berücksichtigt werden. Die höchste Mittelwertbildung ihrer Wahrscheinlichkeiten bestimmt die Label Kategorie (0 oder 1).

Die Suche nach dem besten Split wurde so gewählt, dass bei eng definierten Strukturklassen die Anzahl falsch-positiver Vorhersagen im Promillebereich blieb. Die ermittelten Werte reichten dabei vom Standardwert $\sqrt{770}=28$ für Rezeptoren beschreibende Oberbegriffe wie Fentanyl oder Cannabinoide bis zur maximalen Anzahl der Intensitätswerte (770).

3. Der Parameter `min_samples_split` bestimmt die minimale Anzahl der Beobachtungen, die „Random Forest“ aufweisen muss, um sich in einem Knoten weiter aufzuteilen. Sein Wert wurde für alle Strukturklassen auf den Standardwert 2 gesetzt.
4. Um negative Effekte eines unausgeglichene Datensatz zu vermeiden wurde generell mit der Technik des Oversamplings gearbeitet. Ein unausgeglichener Datensatz liegt dann vor, wenn die Anzahl der mit dem Label 1 versehenen Massenspektren sehr viel kleiner ist als die mit dem Label 0 versehenen Massenspektren. Hierbei werden Beispiele aus der Minderheitsklasse (Label 1) modifiziert und solange dem Trainingsdatensatz hinzugefügt, bis die Anzahl der Minderheitsklasse so hoch ist, wie die Anzahl der Klassen in der Mehrheitsklasse (Label 0). Dies stellt sicher, dass das Modell eine optimale Datenunterstützung erfährt.
5. Die Fähigkeit der Generalisierung des Modells wurde über eine 10-fache Kreuzvalidierung bewertet und lieferte bei eng gefassten Strukturklassen Werte, die vom optimalen statistischen Wert 1 lediglich im Promillebereich abwichen.

Das zu analysierende unbekannte Massenspektrum kann dem Programm „Mass Spectral AI Interpreter“ in Form eines JCAMP-Formates übergeben werden. Neben der Übergabe eines einzelnen Massenspektrums können auch die Massenspektren ganzer Totalionenchromatogramme in einem proprietären Datenformat übergeben werden. Die Vorhersage eines einzelnen unbekanntes Massenspektrums über 100 unterschiedliche Strukturklassen ist innerhalb eines Zeitraums von 10 Sekunden abgeschlossen.

Das Programm ist unter Windows oder unter der „Python“-Programmierungsumgebung „Pycharm“ lauffähig.

3. Literatur

- [1] Lucas B, Ayres et al. Taking the leap between analytical chemistry and artificial intelligence: a tutorial review. *Analytical Chimica Acta* 2012;1161:338403.
- [2] Koshute P, Hagan N, Jameson NJ. Machine learning model for detecting fentanyl analogs from mass spectra. *Forensic Chemistry* 2022;27:100379.
- [3] Bonetti JL, Semaniour S, van Asten AC. Utilization of machine learning for the differentiation of positional NPS isomers with direct analysis in real time mass spectroscopy. *Analytical Chemistry* 2022;94:5029-5040.
- [4] Breiman L. Random forests. *Machine Learning* 2001;45(1):5-32.
- [5] Mass Spectra of Designer Drugs 2023, <https://www.designer-drugs.de/>
- [6] <https://www.jetbrains.com/pycharm/>; zuletzt eingesehen am 29.12.2023.
- [7] <https://www.python.org/downloads/>; zuletzt eingesehen am 29.12.2023.
- [8] <https://anaconda.org/r/r-randomforest/>; zuletzt eingesehen am 29.12.2023.