



AG Chemometrie und multivariate Datenauswertung Jahresbericht 2020

Obmann: Prof. Dr. Philipp Weller

Die AG hatte Ende 2020 einen Mitgliederstand von 7 aktiven und 8 korrespondierenden Mitgliedern. Im Berichtsjahr konnte die Sitzung der AG aufgrund der Pandemie nur einmalig am 22.7.2020 online erfolgen.

Wahlen:

Am 22.7.2020 wurden im üblichen Turnus von 3 Jahren Obleute und Vertretung, sowie SchriftführerIn gewählt.

Als Obmann wurde wiedergewählt: Prof. Dr. Philipp Weller (HS Mannheim)

Als Stellvertreter wurde wiedergewählt: Dr. Dirk Lachenmeier (CVUA KA)

Als neue Schriftführerin wurde gewählt: Dr. Anja Bonte (MRI)

Weiterführung des Aufbaus der AG:

Die AG befindet sich im vierten Jahr nach ihrer Gründung weiterhin noch in einer Aufbauphase. Es besteht großes Interesse an der Mitarbeit in der AG. Schwerpunkte und Themen für die AG werden 2021 allem sein:

- Standardisierung von Prozessen und Strategien zur Datenanalyse und geeigneter Validierung
- Erarbeitung von Toolboxen für die Datenanalyse aus typischen Analysesystemen
- Multivariate Analyse in der Lehre: Einbindung in die Ausbildung von Studenten
- Vernetzung mit anderen Arbeitsgruppen

Zusammenarbeit mit anderen Arbeitsgruppen innerhalb der LChG:

Zentraler Aspekt der multivariaten Datenanalyse ist die Anwendung der Datenanalyse seitens der Anwender aus den anderen Fachgruppen, z. B. Isotopenanalytik, Aromen, etc., um die Anwendbarkeit der Werkzeuge in der Realität zu evaluieren und anzupassen. Die AG hat es sich daher zum Ziel gesetzt, die Interaktion mit anderen Fachgruppen zu intensivieren.

Zusammen mit der AG Hochschullehre wurde 2020 ein Positionspapier zum Bedarf an Ausbildung in Chemometrie in der Lehre erstellt.

Evaluierung von Software-Paketen für die Datenanalyse:

Im Rahmen der Sitzung der AG wurden 2020 weiterhin Open-Source-Softwarepakete zur Datenanalyse evaluiert und vorgestellt

- R (Grundpaket, sowie angepasste Distributionen, bspw. R-GUI von R. Leardi)
- Python
- Orange3
- KNIME

Das größte Potenzial haben hier Python und R, da eine große Community existiert, die diese Tools sehr aktiv weiterentwickeln. Insbesondere im Bereich des maschinellen Lernens sind bereits zahlreiche Pakete verfügbar, die auf die Bedürfnisse der Anwender angepasst werden können (z. B. Scikit-Learn). Hierbei ist jedoch zurzeit noch erheblicher Programmieraufwand erforderlich, was die AG als große Herausforderung für neue Anwender sieht. Aus diesem

Grund sollen 2021/22 erste Workshops für Basisaufgaben, wie PCA oder HCA organisiert werden.

Standardisierung von Datenanalyse-Strategien:

Eines der Kernprobleme der Chemometrie/MVA stellt für die AG die fehlende Standardisierung in der Anwendung dar. Im Hinblick auf das Ziel eines flächendeckenden Einsatzes der multivariaten Analyse sieht hier die AG großen Bedarf in der Erarbeitung von Vorschlägen für die Anwendung typischer Datenanalyse-Tools, wie PCA, LDA, PLS-DA etc. Ein Aspekt hierbei ist auch die Präzisierung der Terminologie.

Der Vorschlag der Arbeitsgruppe ist die Durchführung von Workshops, bei der Anwender einen in der AG erarbeiteten Workflow, z. B. mit IR- oder NMR-Daten selbst anwenden und nachvollziehen können. Dieser Workflow soll innerhalb der Fachgruppe LChG zugänglich gemacht werden.

Angebot von Workshops zur Datenanalyse:

Die Planung für Workshops im Jahr 2020 wurde angesichts Pandemie auf 2021/2022 verschoben. Es besteht weiterhin der Vorschlag, auf Regionaltagungen und voraussichtlich in Online-Formaten derartige Workshops für Doktoranden anzubieten.

Das Ziel ist hierbei vor allem, auch den Nachwuchs zeitig an die Anwendung solcher Werkzeuge zu bringen und das Bewusstsein für die Leistungsfähigkeit des maschinellen Lernens zu schärfen.