

General

ID ¹		
Use case name	AI solution to calculate amount of contained material from mass spectrometry measurement data	
Application domain	Manufacturing	
Deployment model	Embedded systems	
Status	PoC	
Scope ²	Calculating amount of contained material from mass spectrometry measurement data using chromatography	
Objective(s)	To find an accurate and efficient solution to calculating amount of contained material without dependence on individuals	
Narrative	Short description (not more than 150 words)	An AI solution was developed that could automatically pick the peak related to the contained material from measurement data through deep learning. Compared with manual results by an experienced operator, the automated peak picking results using AI had a false detection rate of 7% and an undetected rate of 9%. The peak picking operation time using AI was estimated to be about one fifth.
	Complete description	<p>The technology was developed that utilizes AI (artificial intelligence) to process the vast amounts of data used in analyzing the measurement results, which are essential to analytical processes, acquired from mass spectrometers.</p> <p>Mass spectrometers are used for research and quality control in various areas such as the establishment of early detection techniques for diseases and the measurement of residual pesticides in foods, and because of improvements in sensitivity and speed, the amount of data acquired is enormous. As a result, the data analysis step called "peak picking" has become the bottleneck in the workflow. Complete automation is difficult and to some extent manual adjustments are required. Therefore, there are differences in analysis accuracy depending on each operator and there is a possibility that analytical results might be affected by each operator's practices and data alterations. In recent years, automated data analysis with high accuracy that eliminates this kind of dependence on individuals is now demanded in the fields of healthcare and new drug development.</p> <p>To solve this issue using AI, the three companies investigated the application of deep learning, a neural network technology that imitates brain neurons. Arising to confront this process were two problems: 1) insufficient training data; and 2) learning could not proceed when analytical equipment output data was input, as is, into the deep learning network. The technologies to produce extra data to compensate for the lack of training data and to convert the analysis equipment output features into images were developed. Moreover, the companies</p>

	<p>developed the feature extraction technology to learn the analytical skills of experienced analysts. By doing this, the deep learning network was able to learn from the over 30,000 items of generated training data. Compared with manual peak picking results by an experienced operator, the automated peak picking results using AI had a false detection rate of 7% and an undetected rate of 9%. These results indicate that an automated peak picking can compare favorably with a peak picking by an experienced operator.</p>			
Stakeholders ³				
Stakeholders' assets, values ⁴				
System's threats and vulnerabilities ⁵				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
	1	Recall	Proportion of the true positive to positive results by an experienced operator	Improve accuracy
	2	Precision	Proportion of the true positive to positive results by AI	Improve accuracy
	3	Operation time	Ratio of operation time using AI to the conventional one	Improve efficiency
AI features	Task(s)	Recognition		
	Method(s) ⁶	Deep Learning		
	Hardware ⁷			
	Topology ⁸			
	Terms and concepts used ⁹	Deep Learning, Data Augmentation		
Standardization opportunities/ requirements				
Challenges and issues	<p>Challenges: Achieve the same level as experienced operators for peak picking. Issues: 1) Lack of training data per contained material, 2) how to create good images for deep learning from mass spectrometry measurement data</p>			

Societal concerns	Description	
	SDGs ¹⁰	(Select from pull-down menu)

Data (optional)

Data characteristics	
Description	Mass spectrometry measurement data
Source ¹¹	Mass spectrometry
Type ¹²	Numerical data
Volume (size)	
Velocity (e.g. real time) ¹³	Batch
Variety (multiple datasets) ¹⁴	Single
Variability (rate of change) ¹⁵	Static
Quality ¹⁶	High

Process scenario (optional)

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition ¹⁷	Post-condition ¹⁸
1	Training	Train a model (deep neural network) with training samples			
2	Evaluation	Evaluate whether the trained model can be deployed			
3	Execution	Pick peaks using the trained model and calculate the amount of contained material			
4	Retraining	Retrain a model with training samples			

Training (optional)

Scenario name	Training				
Step No.	Event ¹⁹	Name of process/Activity ²⁰	Primary actor	Description of process/activity	Requirement

Specification of training data ²¹	
--	--

Execution (optional)

Scenario name	Execution				
Step No.	Event ²⁶	Name of process/Activity ²⁷	Primary actor	Description of process/activity	Requirement

Input of Execution ²⁸	
Output of Execution ²⁹	

Retraining (optional)

Scenario name	Retraining				
Step No.	Event ³⁰	Name of process/Activity ³¹	Primary actor	Description of process/activity	Requirement

Specification of retraining data ³²	
--	--

Footnote

¹ Leave this cell blank.

² The scope defines the limits of the use case.

³ Stakeholder involved in the scenario - examples are: type of organization; customers, 3rd parties; end users; humans; environment; negative stakeholders (attackers, criminals, etc).

⁴ Assets and values that are valuable to the stakeholders and at the risk of being compromised by the AI system deployment – examples can include competitiveness; reputation or trust; fairness; safety; privacy; stability; etc.

⁵ Threats and vulnerabilities can compromise the assets and values above. Examples are: different sources of bias; incorrect AI system use; new security threats; challenges to accountability; new privacy threats (hidden patterns).

⁶ AI method(s)/framework(s) used.

⁷ Hardware system used.

⁸ Topology is the study of geometric forms differentiated by intersection and bifurcation. The term is used for the graphic aspects network architectures.

⁹ Terms and concepts listed here can be used to extend the work of WG 1 (AWI 22989 and AWI 23053) as necessary.

¹⁰ The Sustainable Development Goals (SDGs), otherwise known as the Global Goals, are a collection of 17 global goals set by the United Nations General Assembly. SDGs are a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity. See URL for more details: <http://www.undp.org/content/undp/en/home/sustainable-development-goals.html>

¹¹ Origin of data, which could be from instruments, IoT, web, surveys, commercial activity, or from simulations.

¹² Structured/unstructured Images, voices, text, gene sequences, and numerical. Composite: time-series, graph-structured

¹³ The rate of flow at which the data is created, stored, analysed, or visualized.

¹⁴ Data from a number of domains and a number of data types. The wider range of data formats, logical models, timescales, and semantics complicates the integration of the variety of data.

¹⁵ Changes in data rate, format/structure, semantics, and/or quality.

¹⁶ Completeness and accuracy of the data with respect to semantic content as well as syntactical of the data (such as presence of missing fields or incorrect values)

¹⁷ Describe which condition(s) should have been met before this scenario happens.

¹⁸ Describe which condition(s) should prevail after this scenario happens. The post-condition may also define "success" or "failure" conditions.

¹⁹ The event that triggers the step. This might be completion of the previous event.

²⁰ Action verbs should be used when naming activity.

-
- ²¹ Training data can be further specified.
- ²² The event that triggers the step. This might be completion of the previous event.
- ²³ Action verbs should be used when naming activity.
- ²⁴ Specify input of evaluation.
- ²⁵ Specify output of evaluation.
- ²⁶ The event that triggers the step. This might be completion of the previous event.
- ²⁷ Action verbs should be used when naming activity.
- ²⁸ Specify input of evaluation.
- ²⁹ Specify output of evaluation.
- ³⁰ The event that triggers the step. This might be completion of the previous event.
- ³¹ Action verbs should be used when naming activity.
- ³² Retraining data can be further specified.