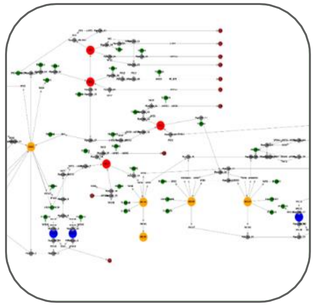
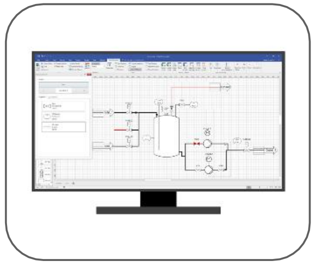
technische universität dortmund

Using artificial intelligence to support the engineering of piping and instrumentation diagrams: modeling & prototyping

J. Oeing[1], W. Welscher[2], K. Brandt[1], N. Kockmann[1]

1. TU Dortmund, Laboratory of Equipment Design, Dortmund/DE
2. X-Visual Technologies GmbH, Berlin/DE





Department of Biochemical and Chemical Engineering

KEEN

XVISUAL Engineering Software

equipment design

Overview

Data preprocessing
- Why using graph data?
- How to represent P&ID as graph?

Prototype - P&ID consistency check
- Implementation of GNNs in P&ID software
- How to interact with a GNN in P&ID software?

Motivation → Data → Modeling → Prototyping → **Conclusion & Outlook**

Machine-readable plant topology data
- Potential
- Accessability
- Availability

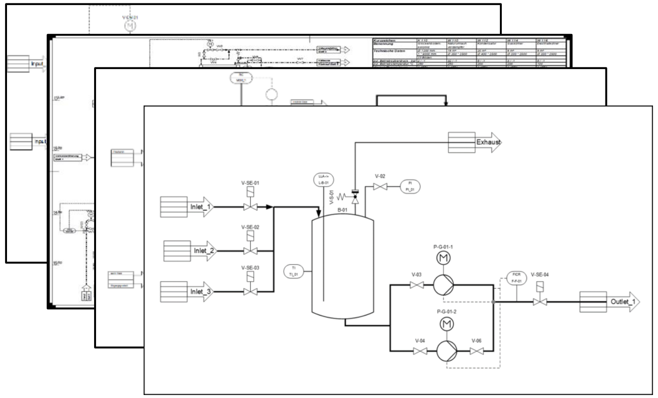
Graph Learning
- What is a Graph Neural Network?
- How to do consistency checks on graph based P&IDs using GNNs?

Conclusion & Outlook
- Overview – GNN consistency checks
- Next steps / Extensions

Motivation

- Digitization requires well-structured data[1]
- Machine-readable engineering data
  - CAEX[2]
  - AutomationML[3]
  - DEXPI[4]
  - …

➢ Combine machine-readable P&IDs and artificial intelligence (AI) to use stored knowledge



<AutomationML/>

CAEX

DEXPI
Data Exchange in
the Process Industry

[1] Wiedau et al., Chem.Ing.Tech., 2021
[2] IEC 62424, 2016
[3] automationml.org, 2022
[4] dexpi.org, 2022

Consistency checks in P&IDs

Hypothesis: *AI is able to learn patterns in P&IDs and thus can detect anomalies in P&IDs.*

- Added value through direct involvment in the drawing process of P&IDs

Aim of this talk:

- Using Graph Neural Networks (GNN) to learn from P&IDs
- Support the Engineering of P&IDs (real time consistency checks)
  - Avoid errors
  - Reduce costs

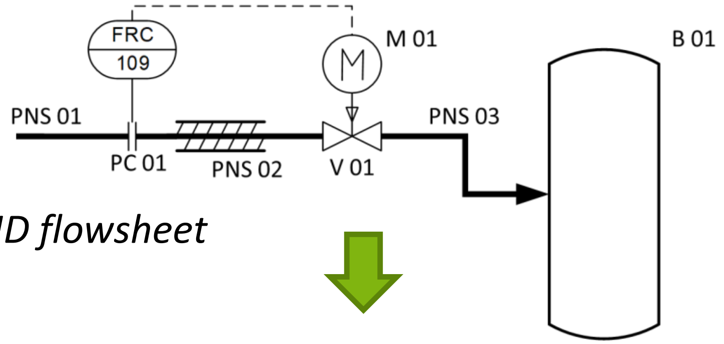
- Understanding machine-readable P&IDs as a broad information model

P&IDs as Graphs

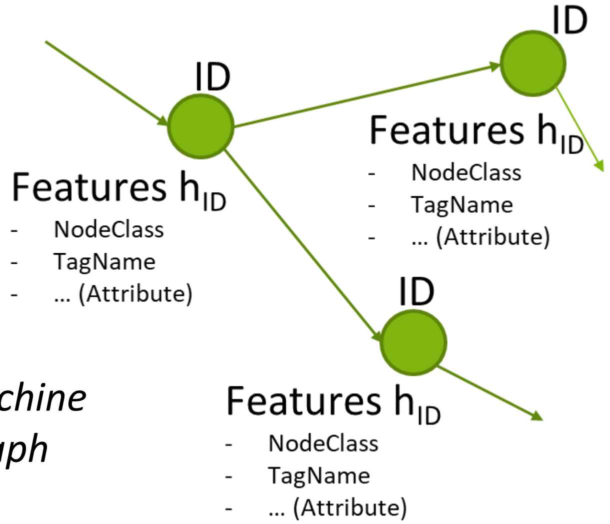
- Abstraction of P&IDs as directed graphs
  - Nodes → equipment, components
  - Edges → piping, signal lines
  - Direction → direction of the process streams

- Advantages
  - Mathematical topology discription
  - Data availability for AI methods

➢ How do we get the data we need?



P&ID flowsheet



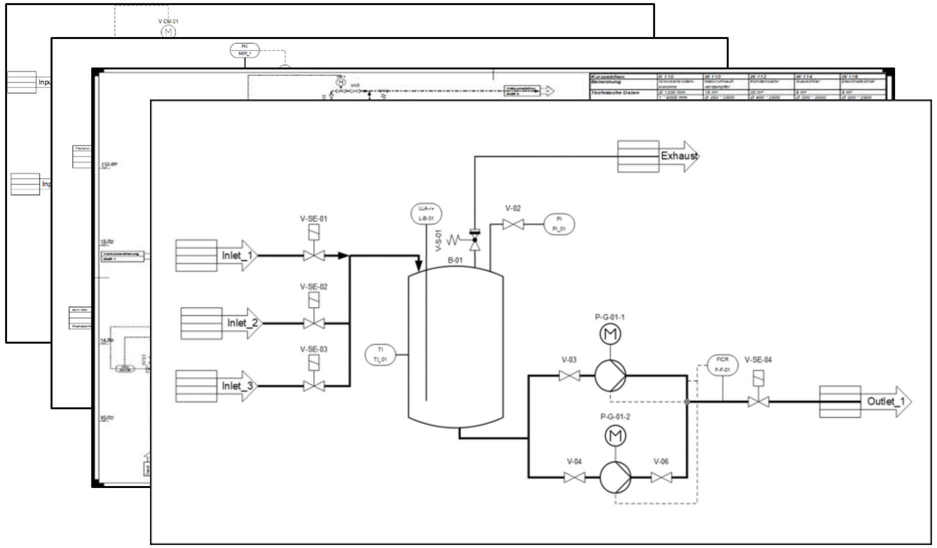
P&ID as machine readable graph

Data

- 25 P&IDs (DEXPI)
  - 8x distillation plants
  - 6x absorption desorption plants
  - 11x modular laboratory plants

- Directed graphs (GraphML)
  - 2504 nodes
  - 2842 edges (directed)

- Categories
  - 16 categories
  - Unbalanced distribution



Category	Number	Category	Number	Category	Number
Valves	725	Heat exchanger	82	Vessels	56
Ball valves	74	Separators	30	Pipe Tees	616
Globe valves	36	Temperature sensors	171	Piping elements	67
Check valves	47	Pressure sensors	115	Inlets / Outlets	226
Safety valves	68	Flow sensors	60		
Pumps	60	Other sensors	71		

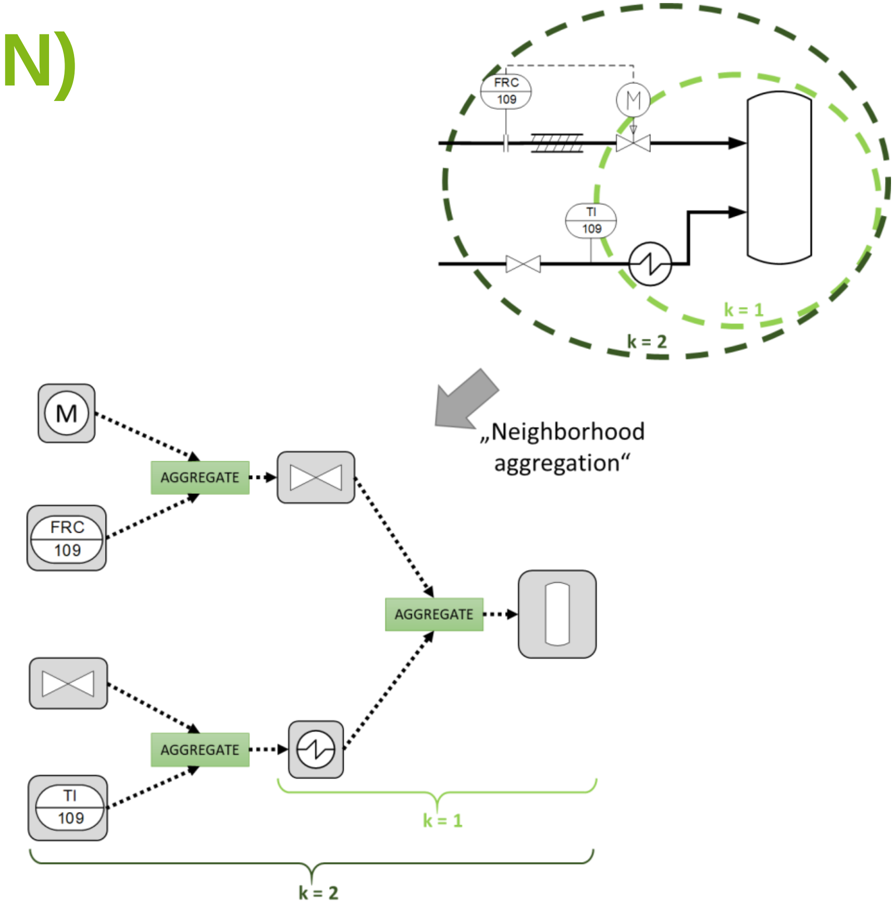
Modelling – Graph Neural Networks (GNN)

- k-layer GNN‘s generate embeddings h_u of node u based on their neighborhood structure[1]

- Message passing in GNNs[2]

$$h_u^{(k+1)} = Update^{(k)}\left(h_u^{(k)}, Aggregate^{(k)}\left(\left\{h_v^{(k)}, \forall v \in N(u)\right\}\right)\right)$$

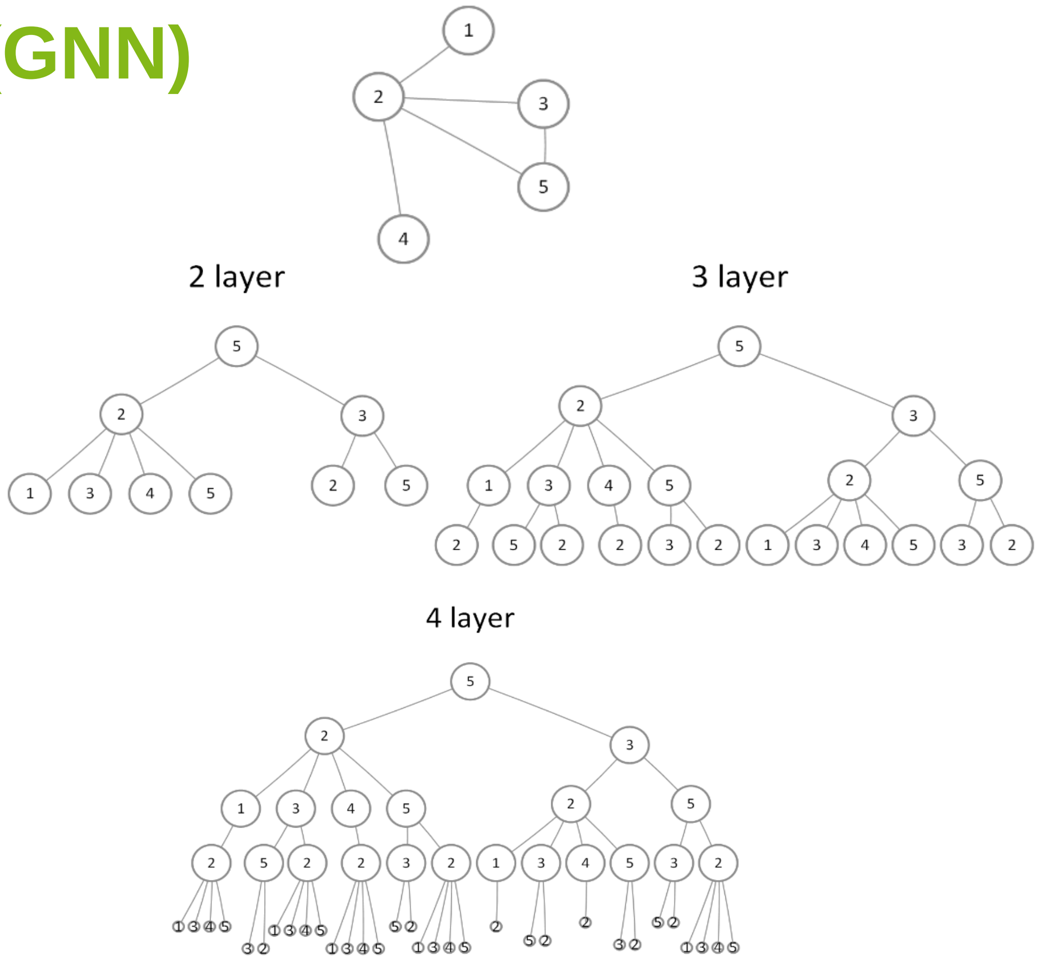
  - <u>Aggregation</u>: collects neighborhood information
  - <u>Update</u>: introduces a non-linearity into the output of a neuron



Neighborhood aggregation of a P&ID - GNN[3]

[1] Leskovec, J., Inductive Representation Learning on Large Graphs, 2017
[2] Hamilton, W., Graph Representation Learning, 2020
[3] Oeing, J. et al., Dig. Chem. Eng., 2022

Challenges – Graph Neural Networks (GNN)



<u>2-layers:</u>

- Aggregation of information two steps away
  - No modelling of larger context

<u>Multiple-layers:</u>

- Aggregation gets to many information
  - Features of different nodes very similar („Oversmoothing“)

➢ Jumping Knowledge Networks (JKN)

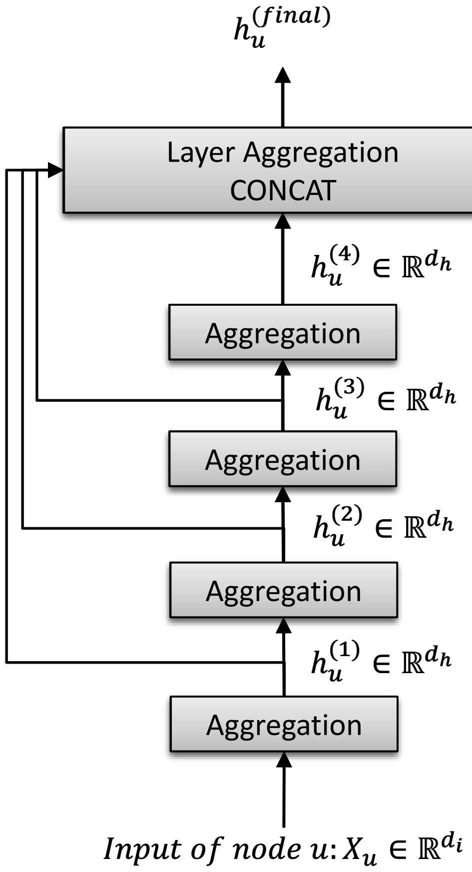
Jumping Knowledge Networks (JKN)

GNN:

- nodes features of one layer used for calculation of features of next layer

JKN:

- save features of all layers and use it again at the end
  - concatenation of all hidden feature vectors as input of last layer
  - nodes only receive information from neighbours

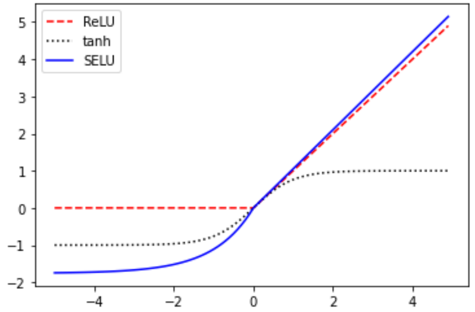


4-layer Graph Neural Network (GNN) / **Jumping Knowledge Network**[1]

[1] K. Xu et al., 2018

Graph neural Networks (GNN) – Training

- Activation functions
  - ReLU, SeLU, tanh

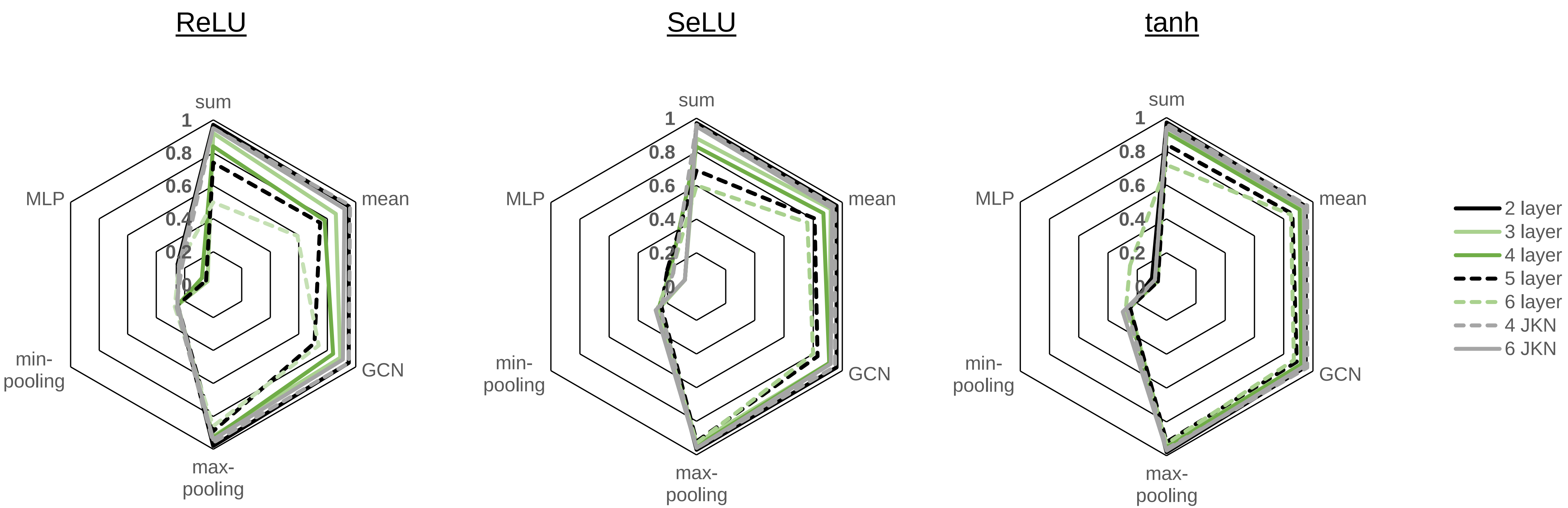


- Train/Test split
  - 0.7/0.3
- Python
  - Deep Graph Library (DGL) / pytorch
- Loss Calculation
  - Optimizer: Adam (learning rate = 0.01)
  - Loss: Cross-entropy

- Aggregation functions
  - Sum
  - Mean
  - Graph Convolution Network (GCN)
  - Max-Pooling
  - Min-Pooling
  - Multilayer-perceptron (MLP)
- Accuracy

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

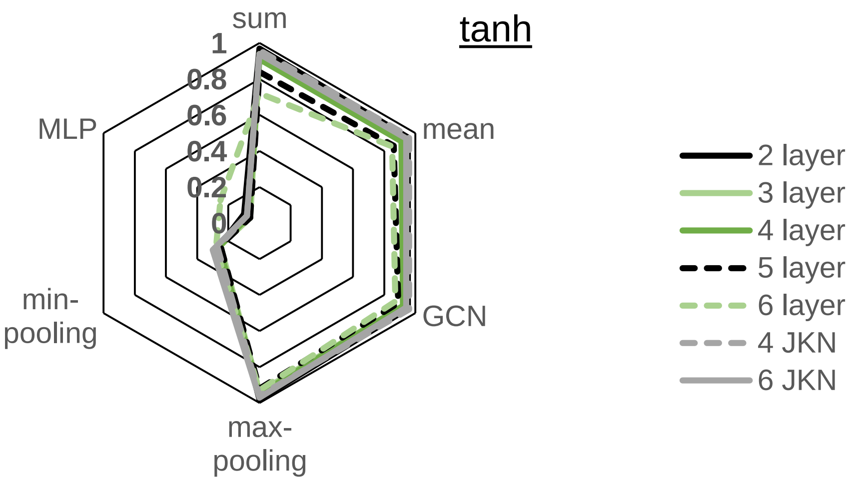
Results

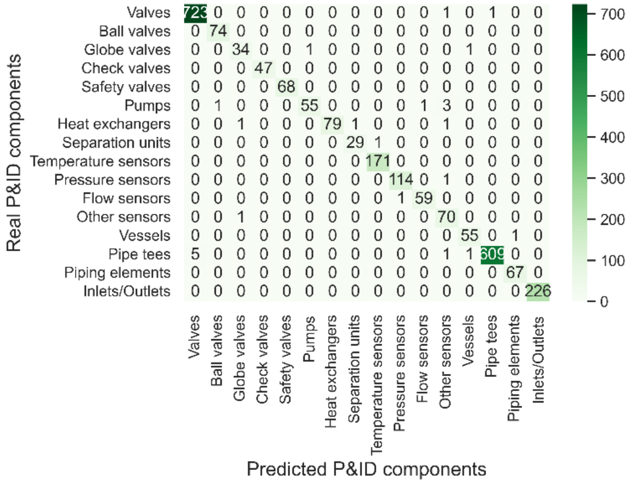

ReLU
SeLU
tanh
sum
mean
GCN
max-pooling
min-pooling
MLP
1
0.8
0.6
0.4
0.2
0
2 layer
3 layer
4 layer
5 layer
6 layer
4 JKN
6 JKN

Results

- Increase of the number of layer results in decrease of the accuracy
- **Max-Pooling** → high accuracy for all different numbers of layers
- **JKN** → high accuracy for all different aggregation functions

➢ Best results: 98 % test accuracy
  - 6-layer JKN
  - Max-pooling
  - tanh



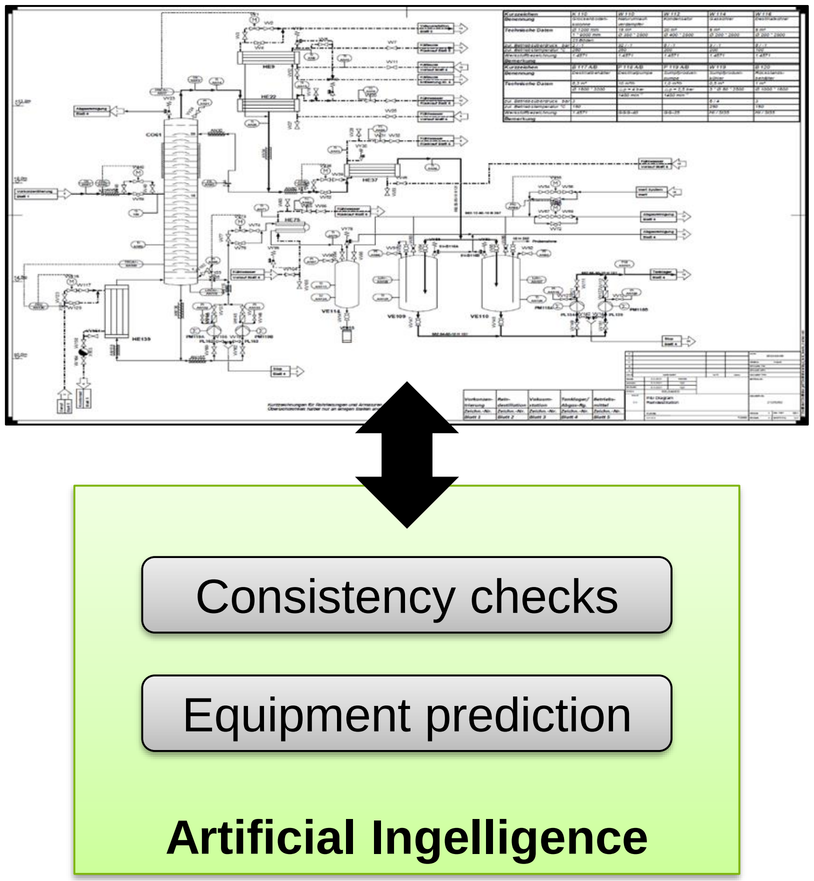


Confusion matrix (6-layer JKN, tanh, max-pooling, 35 hidden neurons)

Prototyping

- Use AI to support the drawing of P&IDs
  - Use stored knowledge from historical P&ID data
  - Reduce errors
  - Improve the quality of P&IDs

- Implementation of AI models in X-Visual PlantEngineer
  - GNN-based model for consistency checks
  - RNN-based model for prediction of P&ID equipment[1,2]

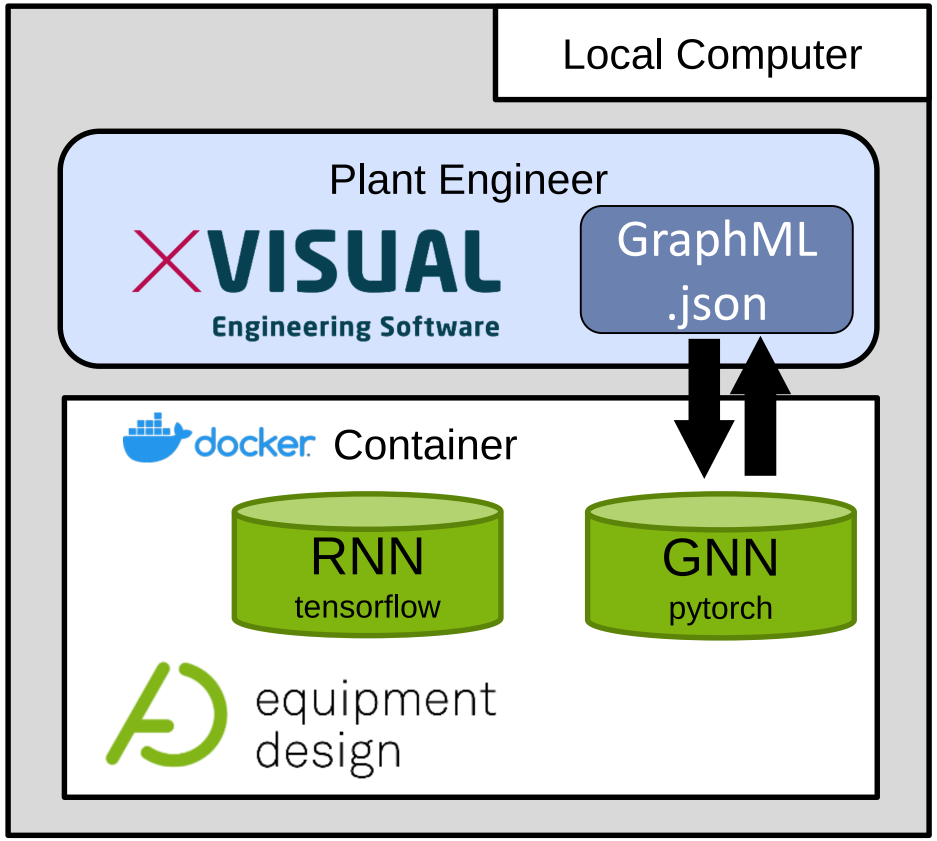


[1] J. Oeing et al., Dig. Chem. Eng., 2022 [2] J. Oeing, JT PAAT, 2021

Implementation

- **Plant Engineer** (X-Visual Technologies)
  - Data Exchange via exchange directory
  - Data Exchange using GraphML / .json

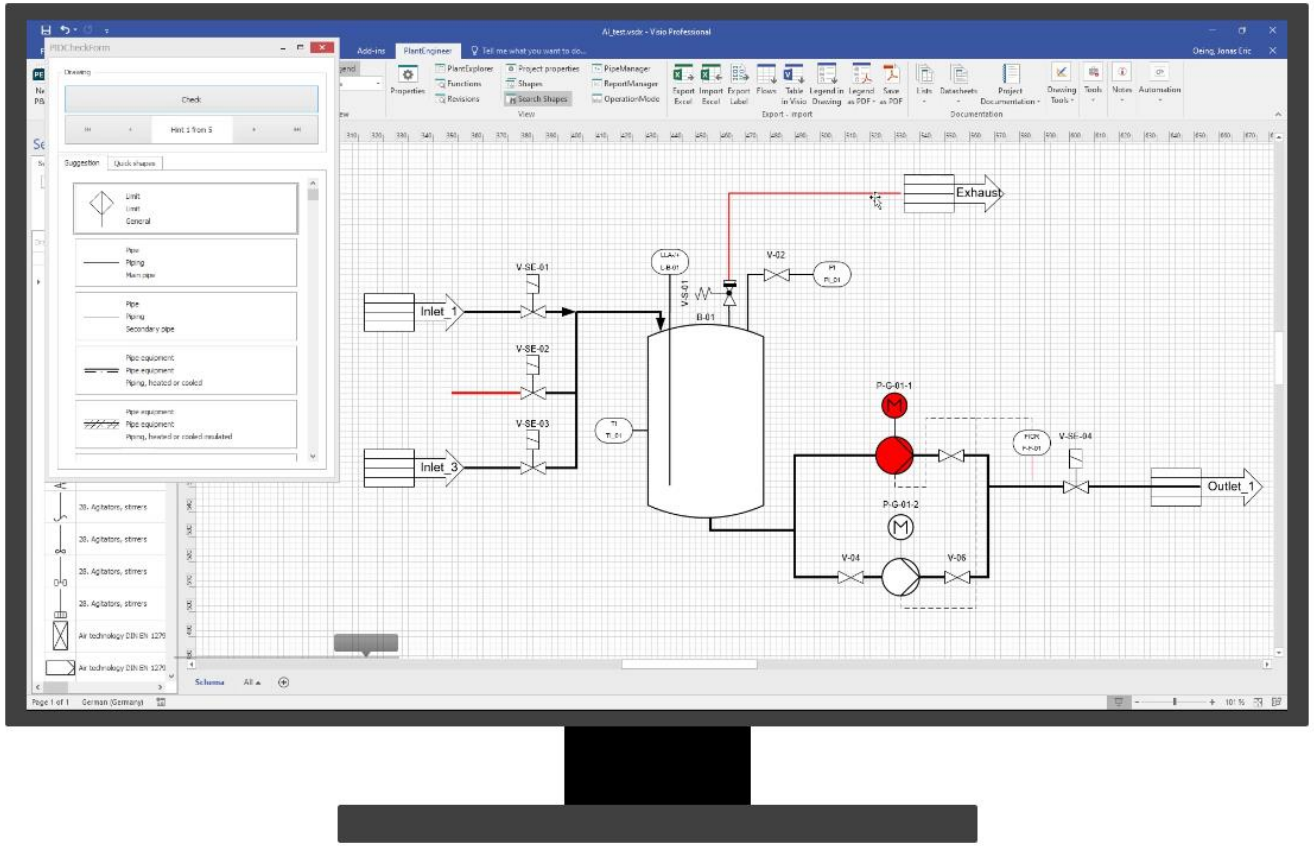
- **Docker Container** (TU Dortmund AD)
  - Python 3.9
  - GNN (pytorch)
    - Output: Inconsistent nodes/edge
  - RNN (tensorflow)
    - Output: Prediction of suitable equipment



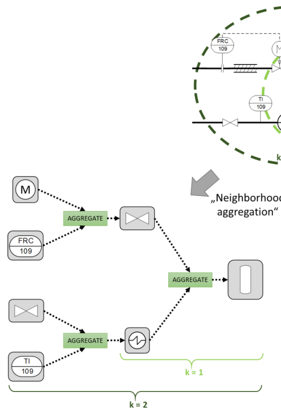
Graphical User Interface

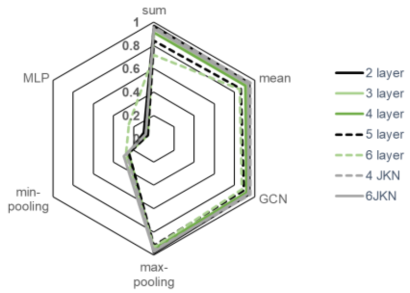
- AI - P&ID consistency check
  - Linking errors
  - Inconsistent equipment
- AI – P&ID equipment prediction
  - Selection of equipment by the cursor
  - Suggestion of suitable equipment in the surrounding area of the selection

➢ Faster drawing of P&IDs

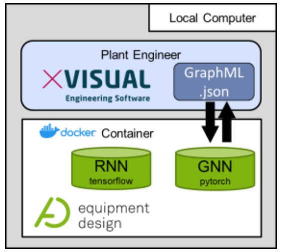


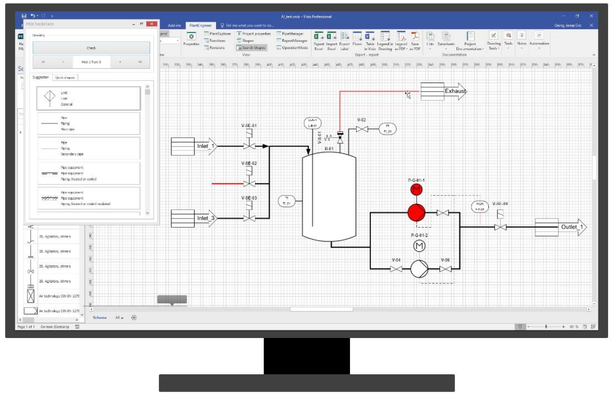
Conclusion & Outlook





- Reliable conistency checks with GNNs possible
  - High accuracies up to 98 % test accuracy
  - JKNs are able to check consistency in different depth of the P&ID





- Prototype for AI supported drawing of P&IDs
  - Real-time consistency checks using GNNs
  - Real-time prediction of Equipment using RNNs

- Outlook:
  - Combination of different GNN models to improve perfomance
  - Retrain models with higher amount of P&ID data

Thank you for your attention

KEEN platform

KE-EN

www.keen-plattform.de

www.tu-dortmund.de

Acknowledgments

DLR Projektträger

Bundesministerium
für Wirtschaft
und Klimaschutz