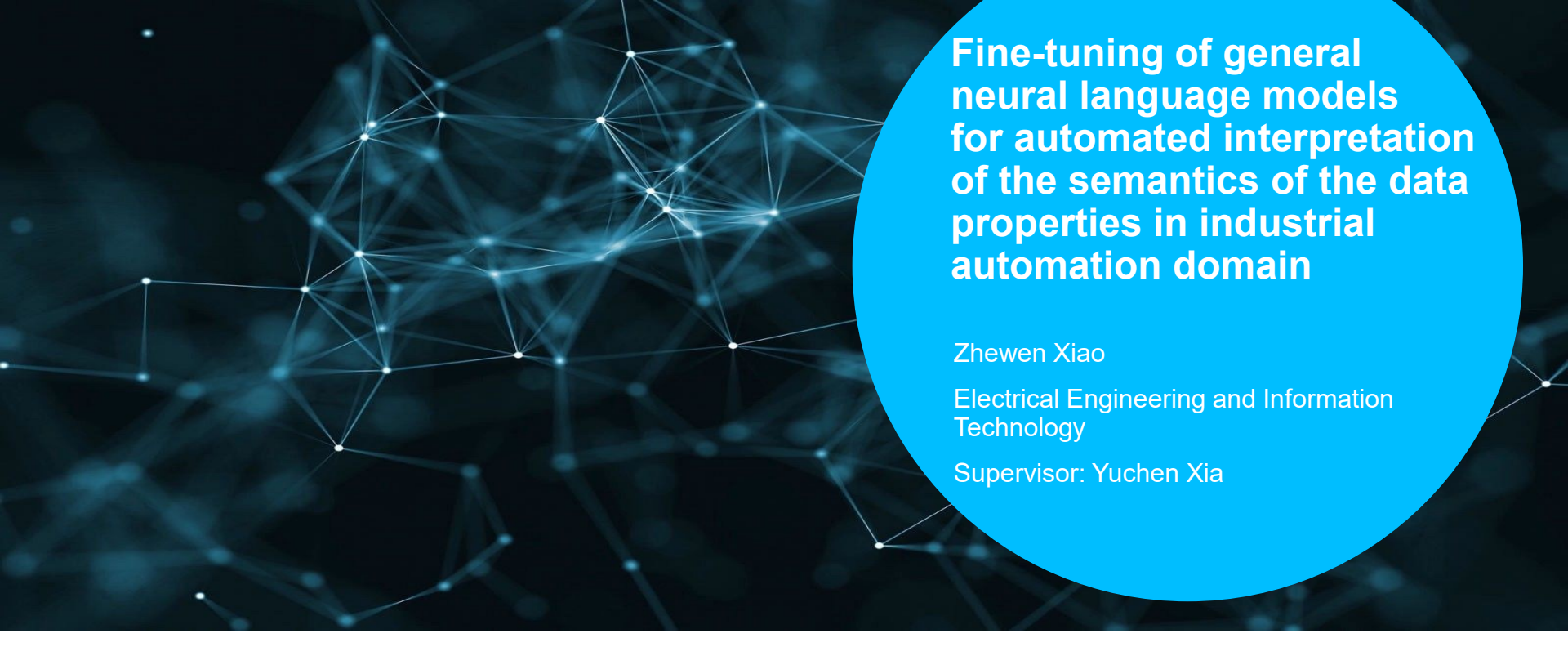
University of Stuttgart
Institute of Industrial Automation
and Software Engineering



Fine-tuning of general neural language models for automated interpretation of the semantics of the data properties in industrial automation domain

Zhewen Xiao

Electrical Engineering and Information Technology

Supervisor: Yuchen Xia

Motivation

Demand for technical data queries

[1]

[2]

- Equipment maintenance
- Asset checking
- Production process analysis

Time-consuming, different vocabularies with the same meaning

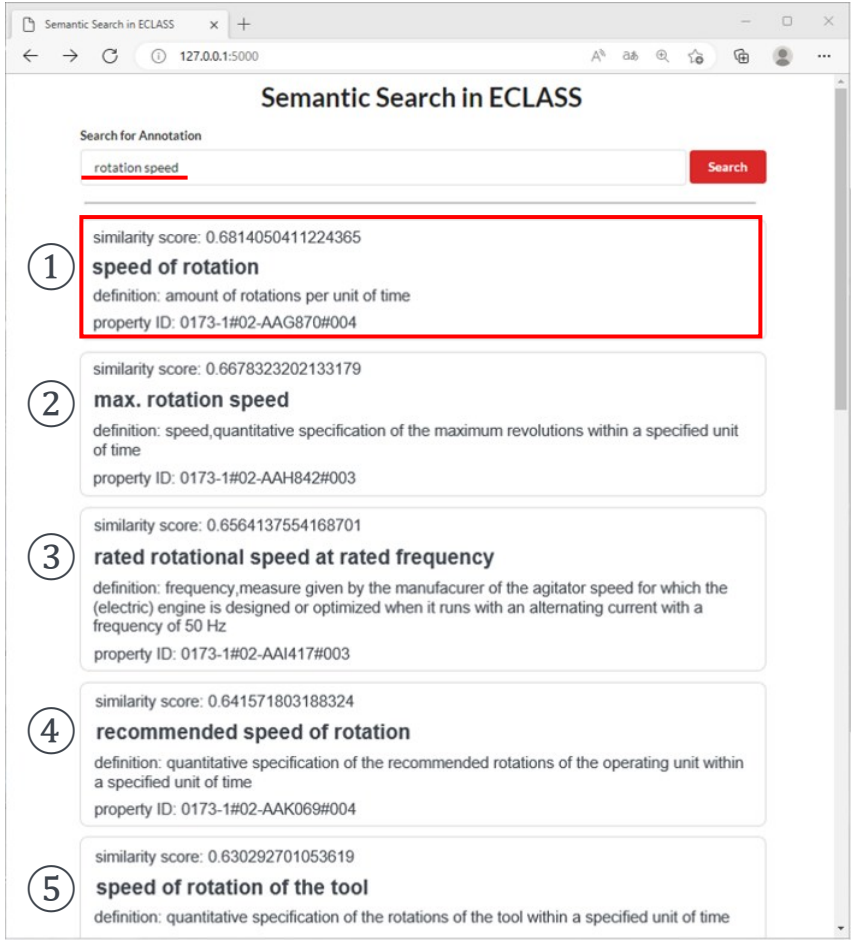
Semantic search engine:

- Based on natural language models
- Need a given database

Motivation

Semantic Search Engine

Natural Language Model → Indexing → CLASS (Dictionary entries)



- Can be used for **automated semantic annotation** (Digital Twin)

- **Accuracy of the ranking** is crucial

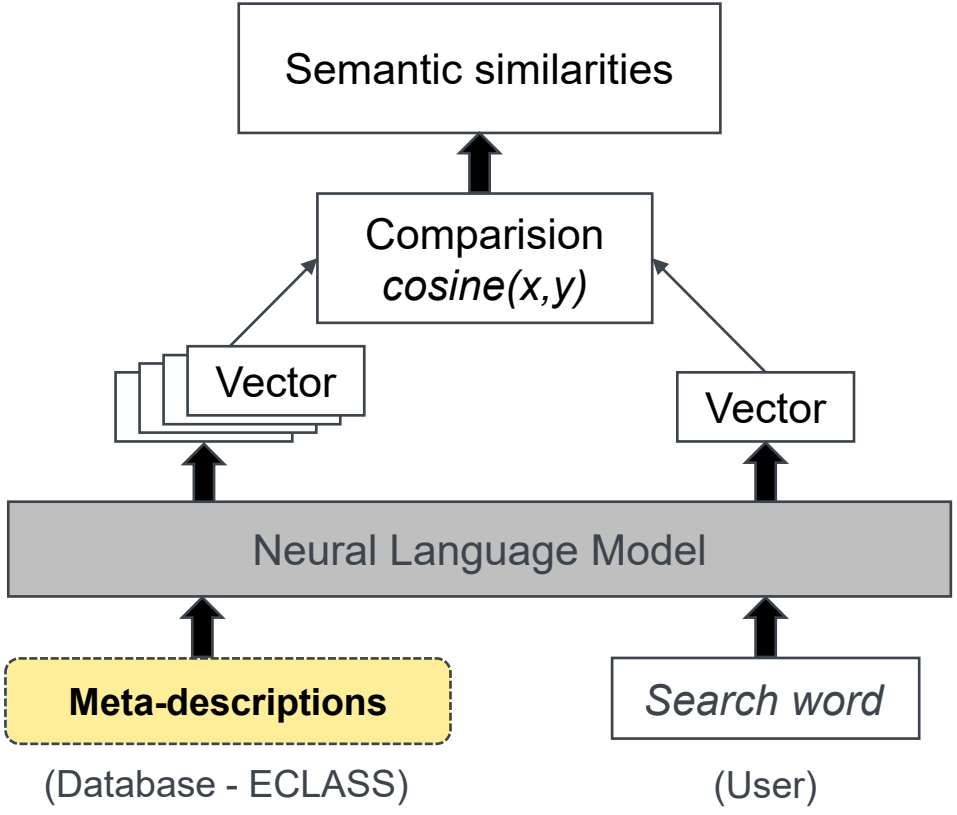
Deficiency of the ranking

- Models are trained by general corpora (news sources on the internet, Wikipedia, etc.)
- The ranking is **not optimal**

Search results for the keyword "*angular velocity*"		
Ranking	ECLASS	Similarity
Top1	min. angle range: …	0.434
Top2	max. rotation speed: …	0.428
Top3	circumferential speed: …	0.420
Top4	max. permissible circumferential speed: …	0.418
Top5	forerun speed: …	0.406

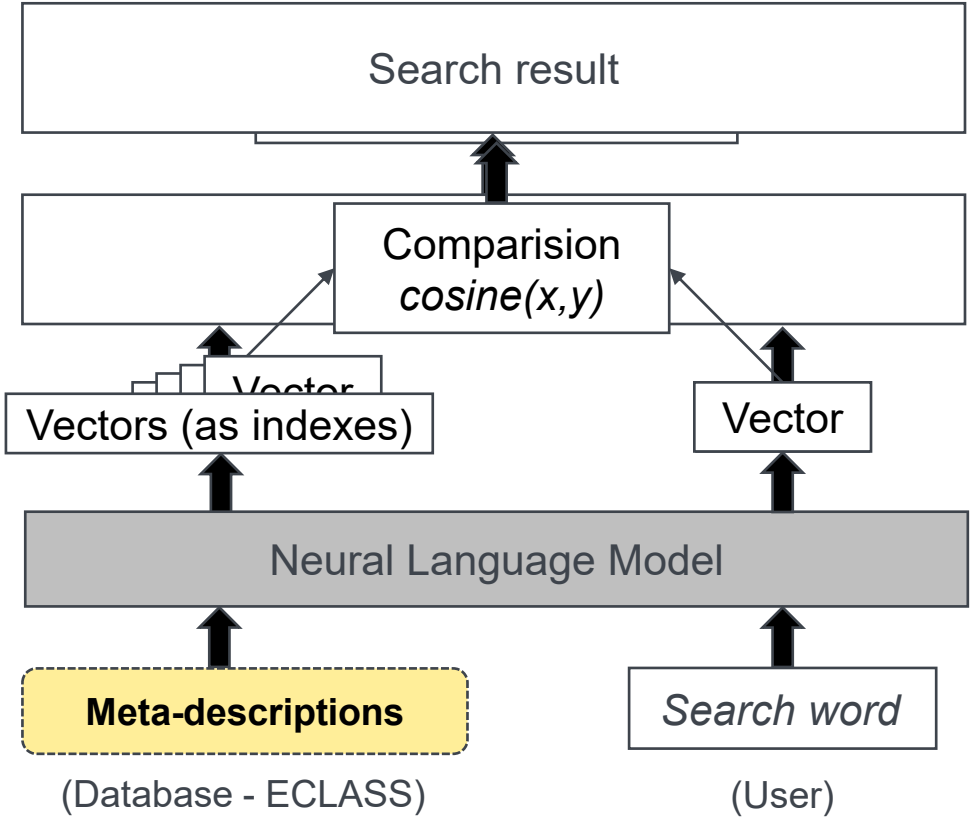
- **To optimize the ranking**, Models shall be **fine-tuned** to adapt to the domain of industrial automation, **better understand domain-specific vocabularies**.

Basis - how ranking works?

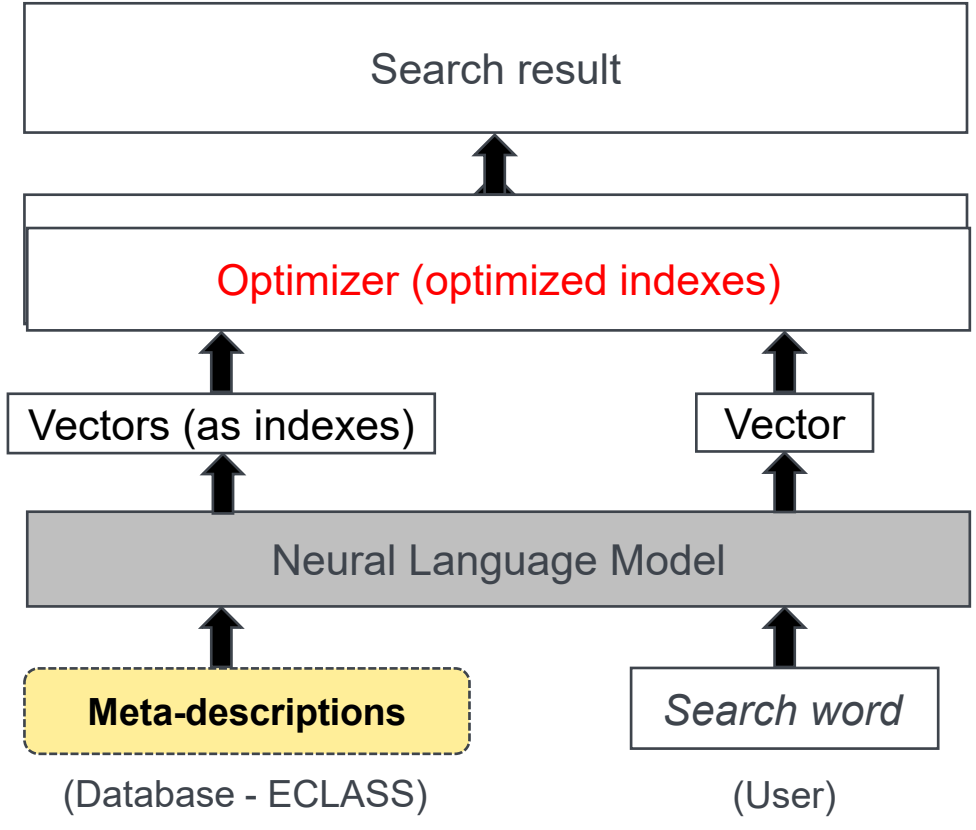


[3]

Basis - how ranking works?



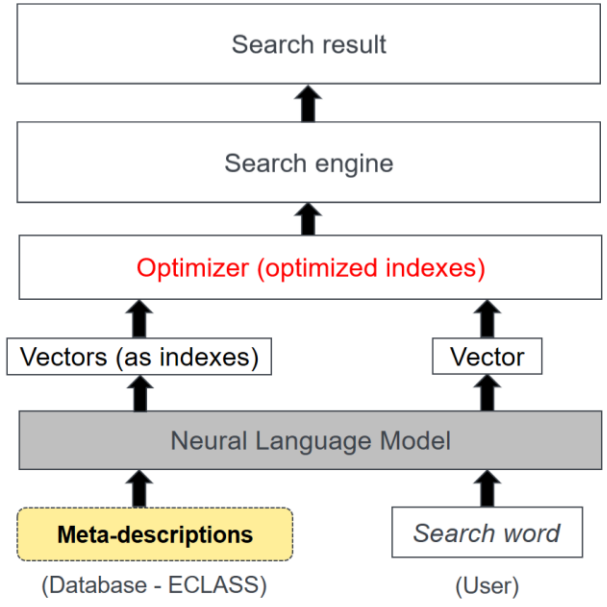
Basis - how ranking works?



How can we realise this optimizer?

2 approaches

- 1. Fine-tuning with added neural network
  - ➢ Building the neural network
  - ➢ Creating the training / test data
  - ➢ Training the neural network

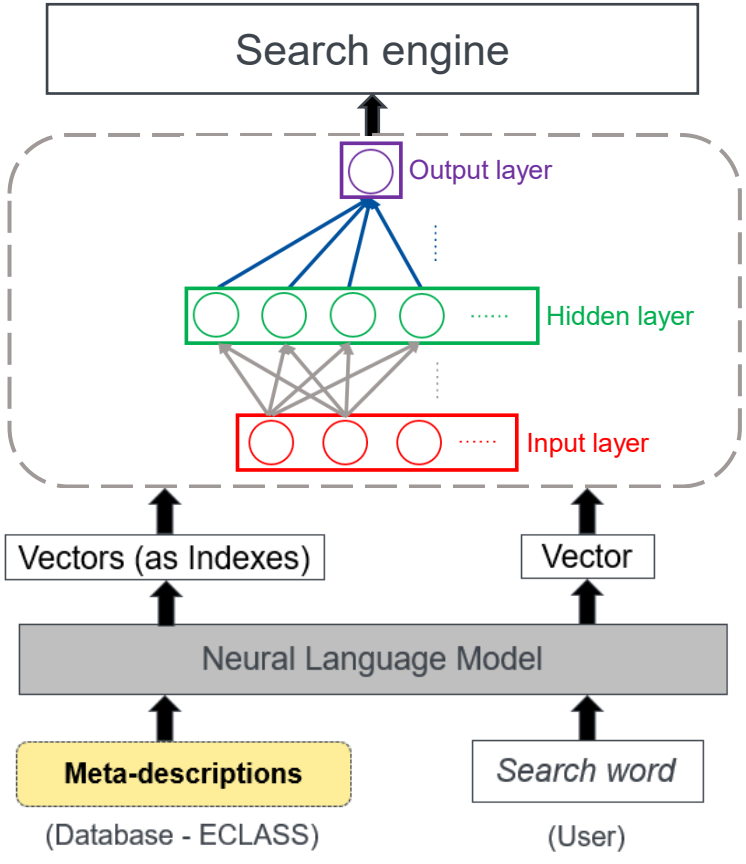


- 2. Fine-tuning by using nearest neighbour algorithm
  - ➢ Creating the history data

Conception of the 1st approach

Added neural network as optimizer

- Structure



(Designed neural network as optimizer)

Conception of the 1st approach

Added neural network as optimizer

- Building the training / test set

Search results for the keyword "rotation speed"			
Ranking	ECLASS	Similarity	
Top1	speed of rotation: …	0.681	S_1
Top2	max. rotation speed:…	0.668	S_2
Top3	rated rotation speed: …	0.656	
Top4	recommended speed of rotation: …	0.642	S_4
Top5	speed of rotation of the tool: …	0.630	

$$\Delta S = \frac{(S_1 + S_2)}{2} = 0.0325$$

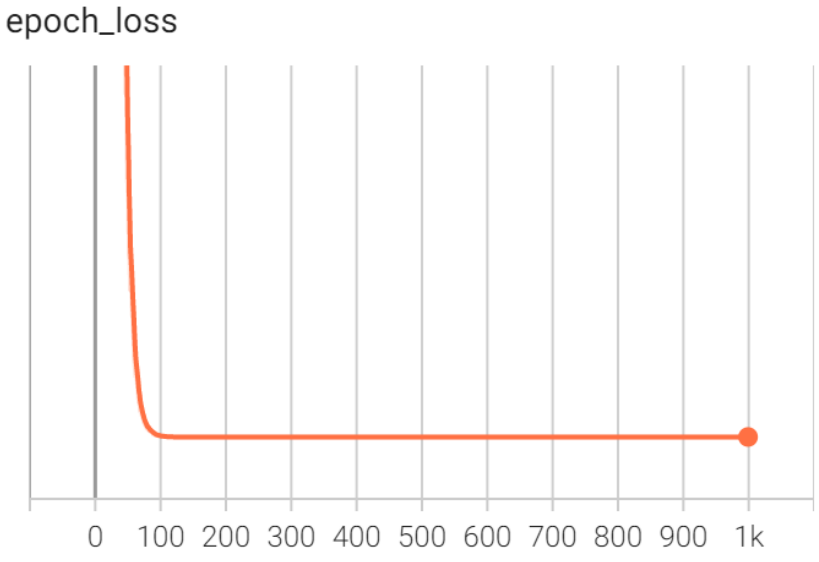
[Vector of *"Rotation speed"* + Vector of *"recommended speed of …"*] $\Delta S = 0.0325$

(Search word) (Search result, need to be reranked) (Label, loss function value)

Implementation of the 1st approach

Added neural network as optimizer

```
inputs = tf.keras.Input(shape=(1536,))
x = tf.keras.layers.Dense(3000, use_bias=True, name='Dense1')(inputs)
outputs = tf.keras.layers.Dense(1, use_bias=True, activation='relu', name='Output')(x)
m = tf.keras.Model(inputs, outputs)
m.compile(tf.keras.optimizers.SGD(learning_rate=0.01), 'mse')
```



Evaluation of the 1st approach

Added neural network as optimizer

Added neural network	≤	Unoptimized search engine
Accuracy: 58%		**Accuracy: 84%**

26% decreased

Possible reasons:

- **Structure too simple** (one hidden layer only), distorts output information from previous model (12 hidden layers), negatively optimized.
- **Traingset too small** (≤ 50 pieces of data), also a difficulty in the domain.

Conception of the 2st approach

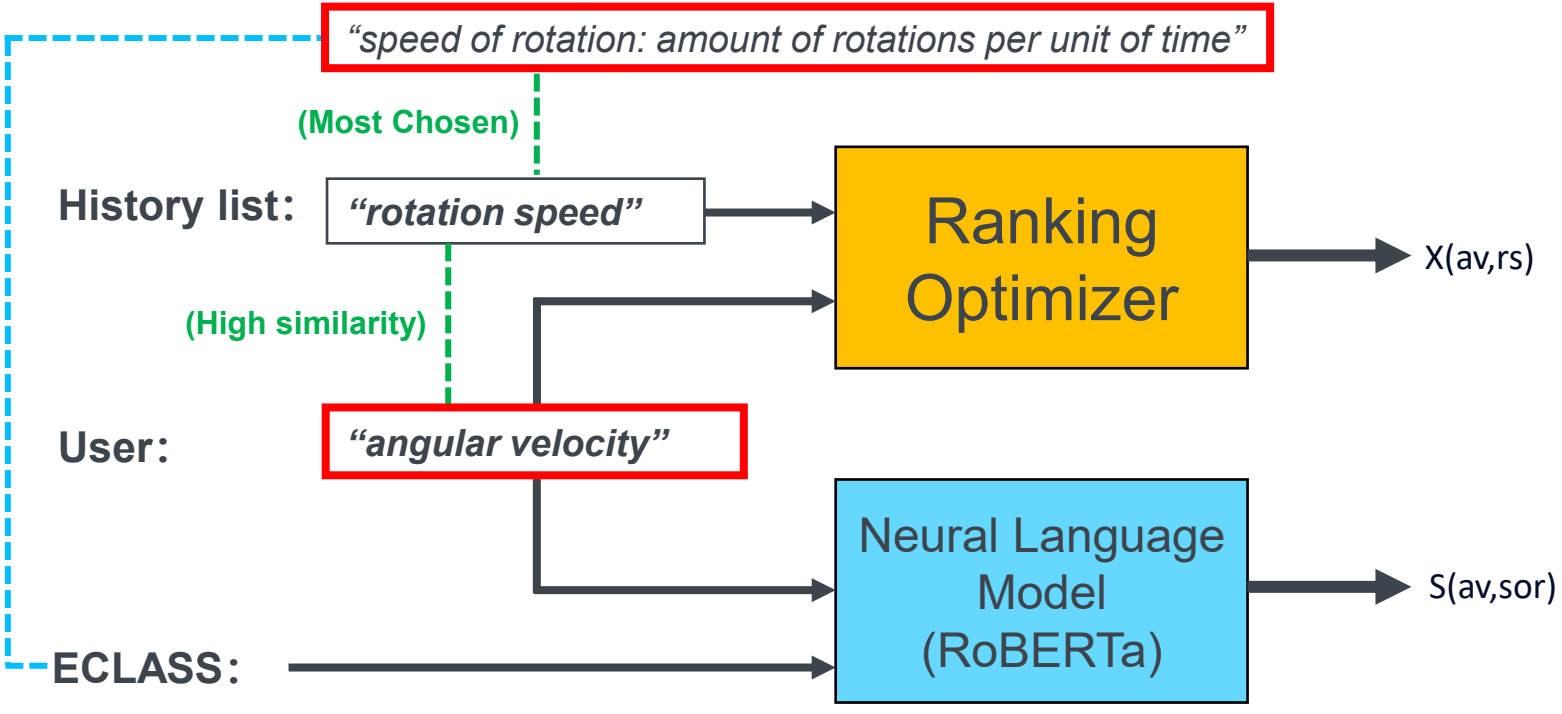
Ranking Optimizer, based on nearest neighbour algorithm

- Creating a history data (history list)

Search words	ECLASS	Serial number	Selection
operating voltage	Operating voltage: voltage which, together with	BAB415#008	15
laser type	laser type: describes the kind of laser	AAK714#003	10
storage robot	container volume robot: specification of the capa	AAJ750#002	13
rotation speed	speed of rotation: amount of rotations per unit d	AAG870#004	20
belt drive	belt drive present: specification whether the unit	AAG555#004	11
power switch	switch-off present: wheter the product is equipp	AAL091#004	9
control unit	integrated control unit: specification whether th	AAF935#004	16
power supply	power supply: describes the source of electrical e	AAK788#003	10

Conception of the 2st approach

Ranking Optimizer, based on nearest neighbour algorithm



Conception of the 2st approach

Ranking Optimizer, based on nearest neighbour algorithm

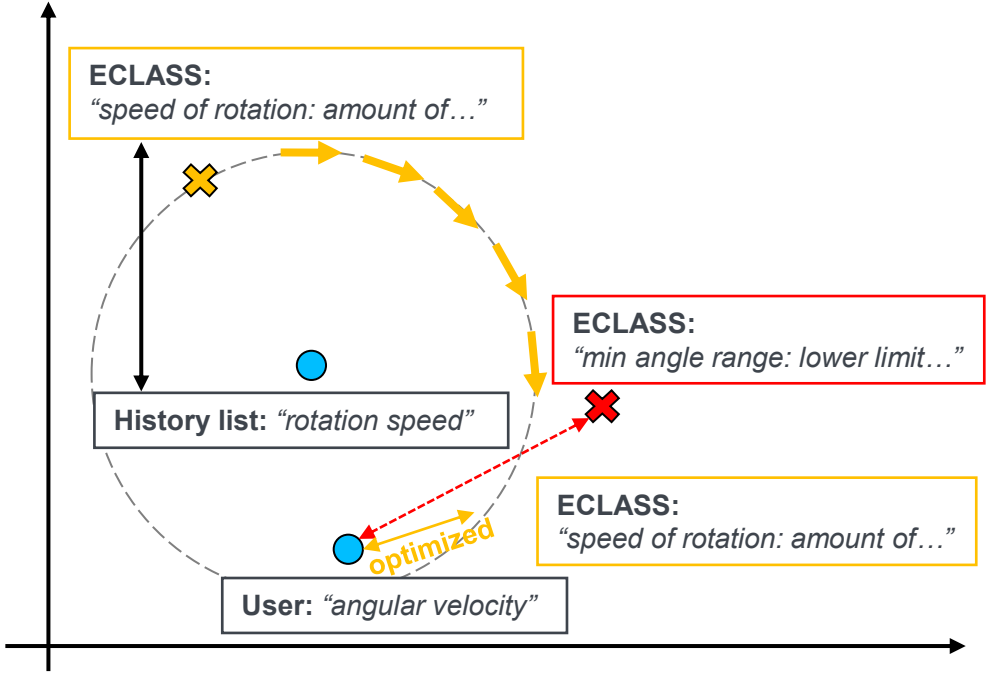
- Results before optimization

```
Please enter keywords:  angular velocity
Processing...
Top 10 results:
0.4335521459579468    min. angle range    low
0.4283602237701416    max. rotation speed
0.41960379481315613    Circumferential speed
0.41761070489883423    Max. permissible circu
0.4059376120567322    forerun speed    by the
0.4057466387748718    Maximum velocity at rat
0.4005143940448761    Nominal speed of the dr
0.3942490816116333    Angle    directional di
0.38980352878570557    speed of rotation    a
0.38961654901504517    max. rotation speed
```

Top 1, yet a useless search result

Top 9, however, it should be the best matched result

- Visualize high dimensional vectors into 2 Dimensions



Implementation of the 2st approach

Ranking Optimizer, based on nearest neighbour algorithm

Optimization factor

$$S(av,sor)_{optimized} = \frac{S(av,sor) + X(av,rs)}{S(av,sor) + (1 - S(av,rs))} > S(av,sor) \quad \text{with} \quad X(av,rs) = (1 - S(av,rs))/4$$

```
score_U_H = (util.dot_score(UserWord, HistoryListWord))
Optimization_factor = (1 - score_U_H) * 0.25
```

Calculation of new similarity

Reranking and new output

Implementation of the 2st approach

Result example

Before

After

```
Please enter keywords:  angular velocity
Processing...
Top 10 results:
0.4335521459579468    min. angle range    lower limit of
0.4283602237701416    max. rotation speed    maximum poss
0.41960379481315613    Circumferential speed    Motion sp
0.41761070489883423    Max. permissible circumferential s
0.4059376120567322    forerun speed    by the device per
0.4057466387748718    Maximum velocity at rated value
0.4005143940448761    Nominal speed of the drive    Quoti
0.3942490816116333    Angle    directional difference bet
0.38980352878570557    speed of rotation    amount of rot
0.38961654901504517    max. rotation speed    indicates t
>>>>>>>>>>>>>>>After Boosting:>>>>>>>>>>>>>>>
0.5866344906973051    speed of rotation    amount of rota
0.4335521459579468    min. angle range    lower limit of
0.4283602237701416    max. rotation speed    maximum poss
0.41960379481315613    Circumferential speed    Motion sp
0.41761070489883423    Max. permissible circumferential s
0.4059376120567322    forerun speed    by the device per
0.4057466387748718    Maximum velocity at rated value
0.4005143940448761    Nominal speed of the drive    Quoti
0.3942490816116333    Angle    directional difference bet
0.38961654901504517    max. rotation speed    indicates t
```

The best **semantic-matched result** and as **Top 1**.

Evaluation

Models	Number of relevant entries from Top-20 Output				Average accuracy	Efficiency		Robustness
	Search word 1	*Search word 2*	*Search word 3*	…		Time Complexity	Space Complexity	
RoBERTa	80% (16)	95% (19)	80% (16)	…	84.0%	$O(n)$	$O(n)$	+
RoBERTa +Neural Network	55% (11)	60% (12)	45% (9)	…	58.6%	$O(n^2)$	$O(n)$	-
RoBERTa +Nearest Neighbour	90% (18)	95% (19)	85% (17)	…	91.1%	$O(n)$	$O(n)$	+

- **RoBERTa+Neural Network:** Accuracy decreased by 26%, inefficient, poor robustness.

✓**RoBERTa+Nearest Neighbour:** Accuracy improved by about **7%**, efficiency and robustness acceptable;

high requirements for History List, superiority mainly based on the user selection records (Big Data).

Summary and Outlook

Research results:

➢ Neural Language Model – **RoBERTa** fine-tuned by two methods:

- **Added neural network** unsatisfied due to its structure
- **Ranking Optimizer** improves the accuracy of semantic search engine by about 7%, semantic search engine **optimized** in the domain of **industrial automation**

Deficiencies:

- Manual judgment of “relevant” entries
- Limited records in history list

Outlook:

- Try out more advanced language models
- Integrate this engine into the data-annotation process

University of Stuttgart
Institut of Industrial Automation
and Software Engineering

Thank you!

Zhewen Xiao

e-mail st165303@stud.uni-stuttgart.de
phone +49 (0) 711 685-
fax +49 (0) 711 685-

University of Stuttgart

Reference

1. https://haus-der-dokumente.de/Services/Scan-Service/
2. https://www.thum.de/loesungen/branchen/industrie/
3. Y. Xia, N. Jazdi, M. Weyrich, “ Automated generation of asset administration shell: A transfer learning approach with neural language model and semantic fingerprints,” 2022.

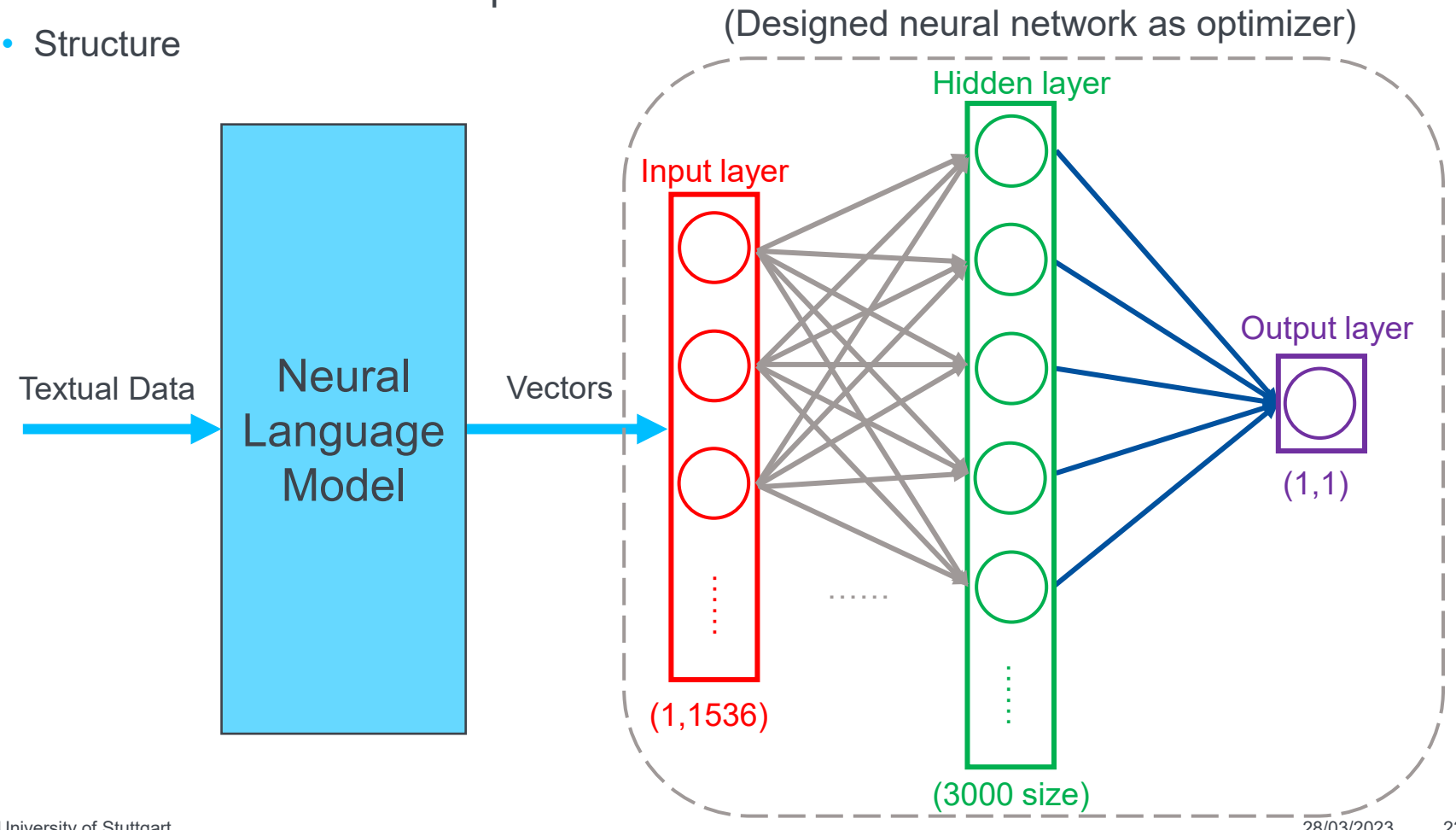
Data structure of ECLASS

ID	Name	Definition
0173-1#02	Material of joint replace	fabric or raw material the joint replacement (foot, ankle) is m
0173-1#02	step distance	Distance to highest level of a step drill
0173-1#02	contact radius	radius of that part of the cutting item that can create a define
0173-1#02	plug length	length of that part of the cutting item that has uncompleted t
0173-1#02	plug angle	Angle of that part of the cutting edge that is closest to the wc
0173-1#02	cutting edge centre cou	number of cutting edges that are able to cut across the center
0173-1#02	countersunk depth of cc	depth of the cylindrical counter bore of a hole in the centre o
0173-1#02	connection diameter, w	greatest permissible upper deviation from the nominal conn
0173-1#02	connection diameter, w	greatest permissible lower deviation from the nominal conn
0173-1#02	tool style code	attribute designation BLD
0173-1#02	drilling diameter	cutting part of a combination milling tool that drills a hole
0173-1#02	stock removal recomme	thickness of material that is recommended to be removed in
0173-1#02	flange diameter	dimension between two parallel tangents on the outside edg
0173-1#02	presetting torque	presetted torque where the transmission of power will be cu
0173-1#02	corner radius	nominal radius of a rounded corner measured in the XY-plane
0173-1#02	insert included angle	angle between the major and the minor cutting edges of a cu
0173-1#02	inscribed circle diamete	A circle inscribed in a surface (does not go beyond the surface
0173-1#02	clamping length maximu	greatest portion of the connection feature that is necessary t
0173-1#02	clamping length minimu	smallest portion of the connection feature that is necessary t
0173-1#02	cutting edge condition c	identifier for the design of the cutting edge of burrs
0173-1#02	guide diameter	diameter of that part of a collet that participates to centre the
0173-1#02	functional diameter	diameter that determines if different components can be mc
0173-1#02	thread form type	identifier for the cross section shape of a thread profile

Conception of the 1st approach

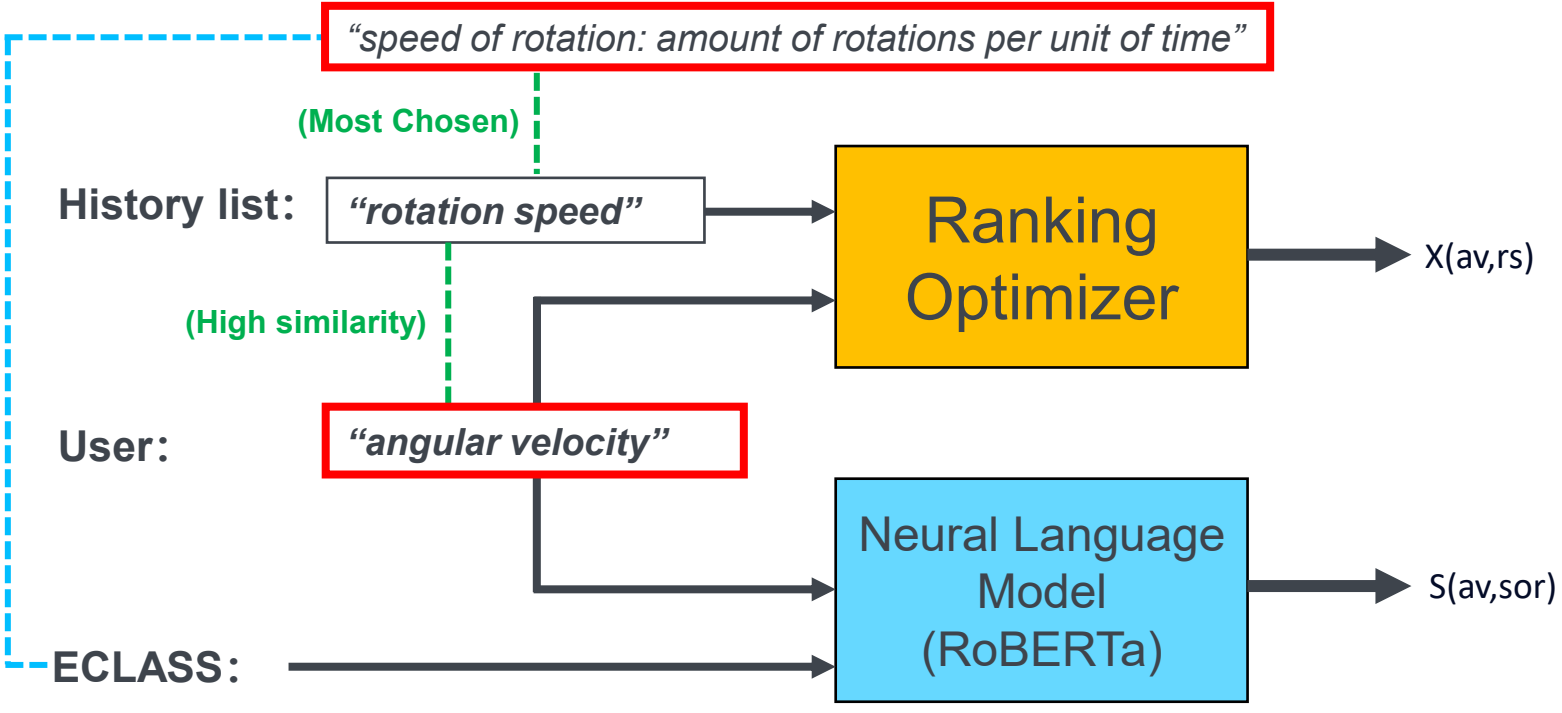
Added neural network as optimizer

- Structure



Conception of the 2st approach

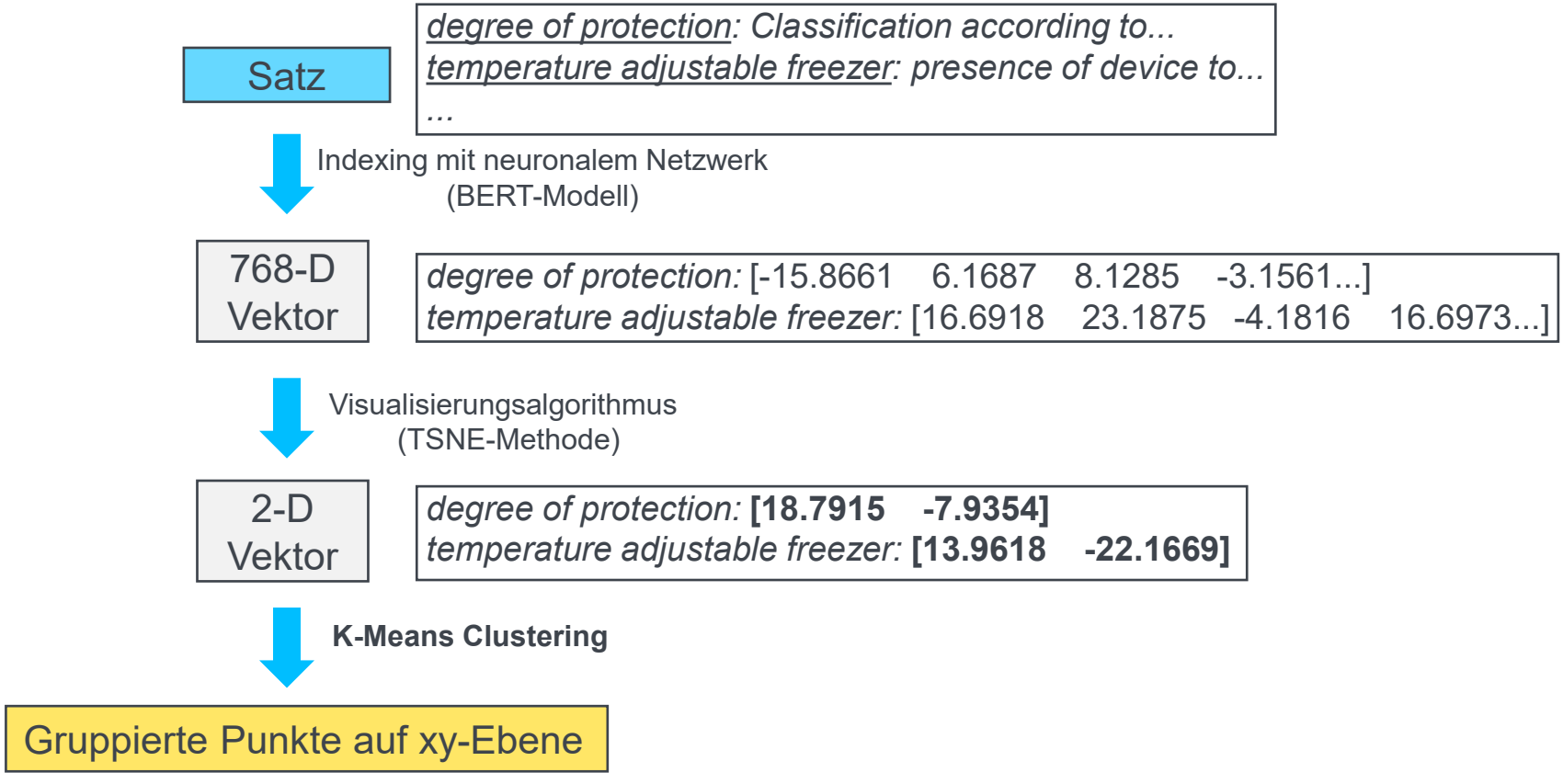
Ranking Optimizer, based on nearest neighbour algorithm



$$S(av,sor)_{optimized} = \frac{S(av,sor) + X(av,rs)}{S(av,sor) + (1 - S(av,rs))} > S(av,sor) \quad \text{with} \quad X(av,rs) = (1 - S(av,rs))/4$$

Hilfsfunktionen der Datenbank (3/3)

3. K-Means Clustering



(TSNE: **T**-distributed **S**tochastic **N**eighbor **E**mbedding)

Hilfsfunktionen der Datenbank (1/3)

1. Funktion „Semantische Suche“: Evaluierung der Genauigkeit

Annotierte Daten	Anzahl relevanter Einträge von Top-20 Output	Genauigkeit
serial_number	16	80.0%
measurement_method	17	85.0%
corrosion_resistance_class	16	80.0%
degree_of_protection	19	95.0%
test_conditions	15	75.0%
ambient_temperature	19	95.0%
degree_of_pollution	15	75.0%
max_output_current	16	80.0%
product_weight	18	90.0%
max_switching_frequency	17	85.0%

$$\frac{16}{20} = 80.0\%$$

Durchschnittliche Genauigkeit: **84.0%**