University of Stuttgart
Institute of Industrial Automation
and Software Engineering

# Evaluation of performance of 1 foundation models for 2 semantics-based classification of standardized data properties for measurement data 3

Research Project Presentation

Yang Jiang

Study Program: Electromobility

# Motivation


Sensor Technologies in the Era of Smart Factory and Industry 4.0 [1]

- sensors and the measurement data they generate play an important role in smart factories
- sensor information should be understood smoothly between different scenarios or factories
- determination and classification of these equivalent data features

# Outline

Fundamentals

Conception of Classification System

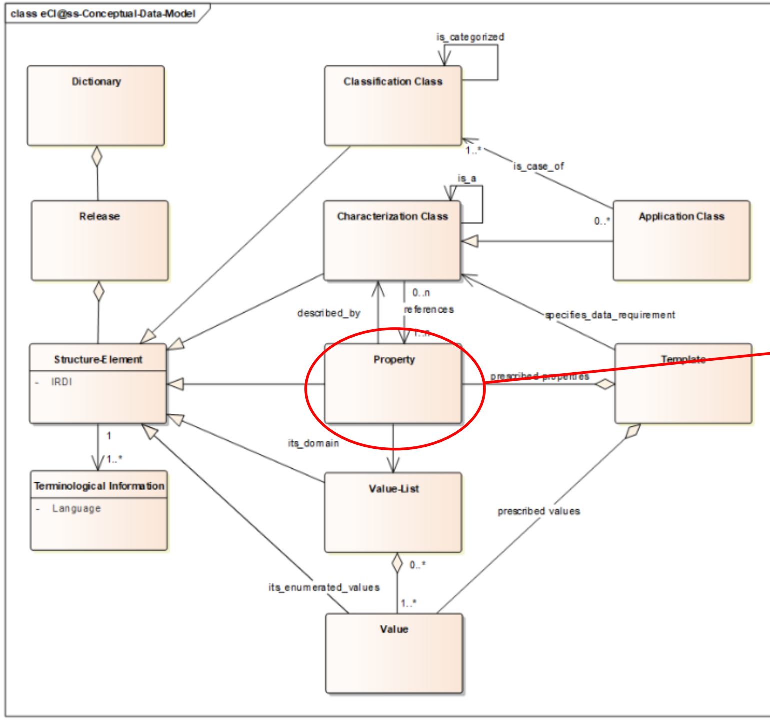
Implementation of Classification System

Evaluation

Conclusion and Future Work

# Fundamentals

## ECLASS




ECLASS-Conceptual-Data Model [2]

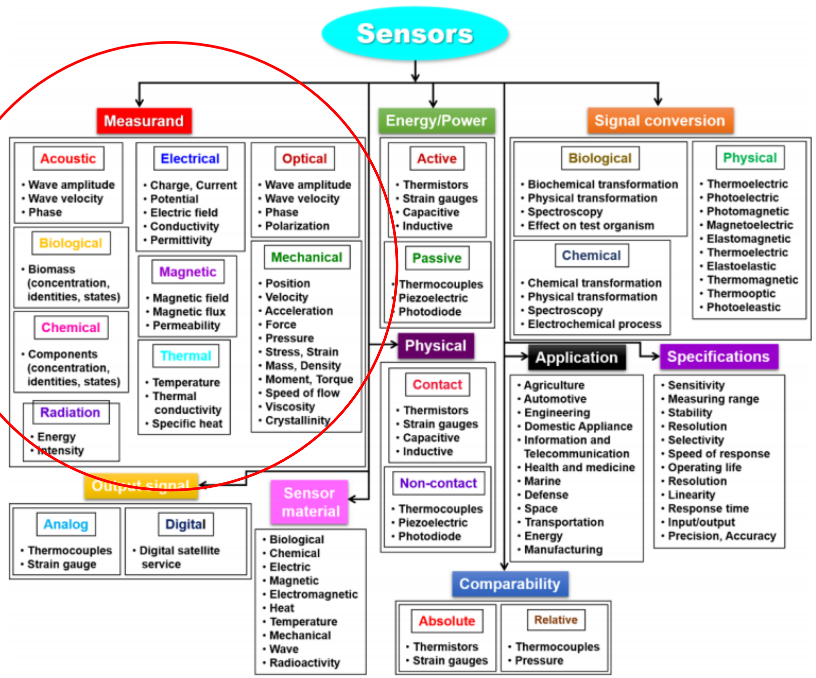
**Property**

| | |
|---|---|
| **Preferred name** | Output voltage |
| **Definition** | Voltage available at the output of an operating unit |
| **IRDI** | 0173-1#02-AAB427#007 |
| **Data type** | REAL_MEASURE |

Example of the ECLASS Property

# Fundamentals

## Measurement data taxonomy




Classification of Sensors [3]

1 Messung nichtelektrischer Größen
- 1.1 Grundlegende Messgeräte
  - 1.1.1 Zeitmessung
  - 1.1.2 Längenmessung
  - 1.1.3 Zählen
- 1.2 Weitere Messgeräte elementarer Größen
  - 1.2.1 Flächeninhaltsmessung
  - 1.2.2 Volumenmessung
  - 1.2.3 Ortsbestimmung, Winkel- und Richtungsmessung
  - 1.2.4 Masse, Gewichtskraft, Dichte usw.
  - 1.2.5 Temperatur

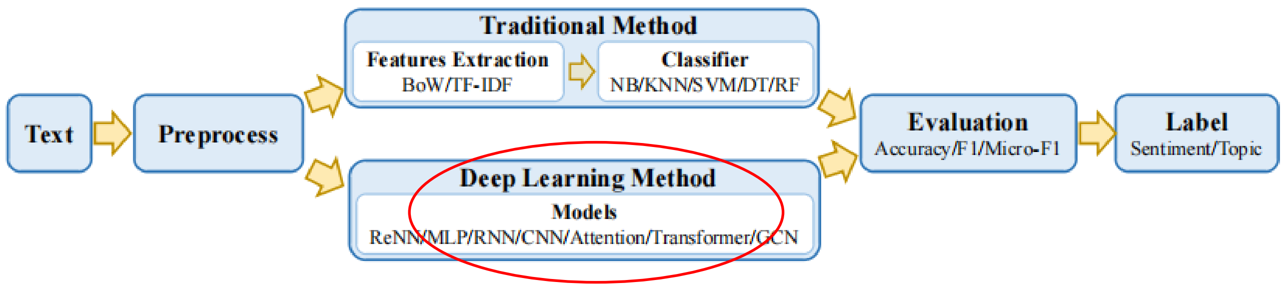
2 Messung elektromagnetischer Größen
- 2.1 Elektrische Größen
- 2.2 Magnetfeld
- 2.3 Radioaktivität und Strahlung

3 Abgeleitete Messgeräte
- 3.1 Geschwindigkeit
  - 3.1.1 Drehzahl
  - 3.1.2 Beschleunigung
  - 3.1.3 Zurückgelegter Weg
  - 3.1.4 Leistung
- 3.2 Messungen an Flüssigkeiten und Gasen
- 3.3 Messungen an Feststoffen
- 3.4 Meteorologische Instrumente
- 3.5 Messung der lichttechnischen Größen und Farbeigenschaften
- 3.6 Schall- und Schallpegelmessung
- 3.7 Kombinierte Geräte
- 3.8 Universelle Messgeräte für verschiedene elektrische Größen
  - 3.8.1 Qualität der Messungen

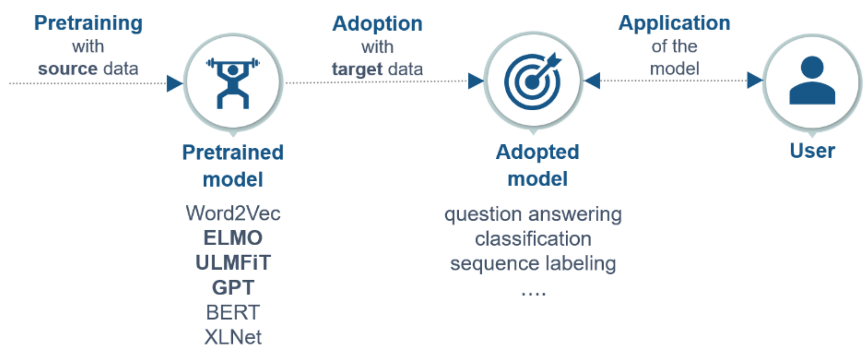
Classification of measuring instruments [4]

# Foundamentals

## Text Classification and transfer learning




Flowchart of the text classification [5]




Steps in transfer learning [6]

# Foundamentals

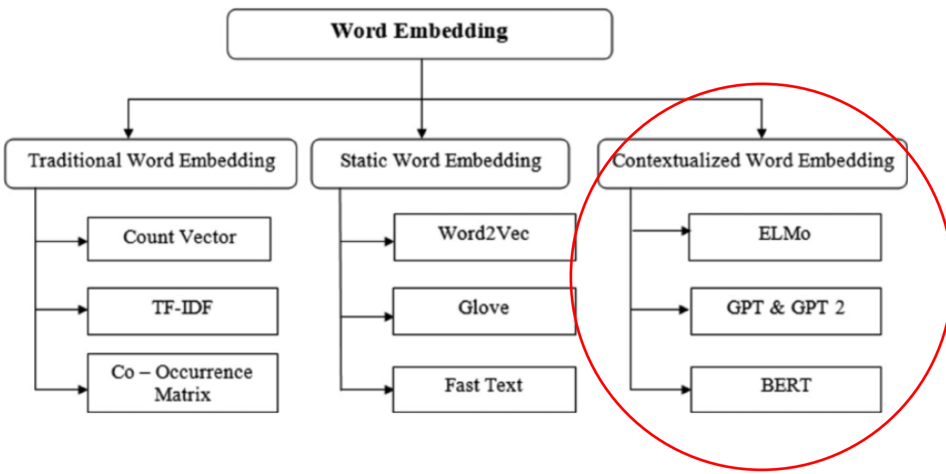
## Foundation Models

| Model Name (Release Time) | Model Architecture (#params) | Pre-training Tasks | Pre-training Data | Applications |
|---|---|---|---|---|
| BERT(10.2018) | En (0.1, 0.3) | MLM(token) +Next Sentence Prediction | BookCorpus + English Wikipedia (800M+2500M words) | Sentence classification or token classification |
| RoBERTa(07.2019) | En (0.1, 0.3) | MLM(token) | BookCorpus + CC-News + OPENWEBTEXT + STORIES (160GB) | |
| DeBERTa(06.2020) | En (0.1, 0.4, 0.7, 0.9, 1.5) | MLM(token) + Disentangled Attention | Wikipedia + BookCorpus + OPENWEBTEXT + STORIES (78GB) | |
| ELECTRA(03.2020) | En (0.1, 0.3) | MLM(token) + Replace Token Detection | Wikipedia + BookCorpus + Giga5 + ClueWeb + Common Crawl (32.89B words) | |
| ELMo(02.2018) | De (0.1) | LTR + RTL | 1 Billion Word Benchmark | Text generation |
| GPT-3(05.2020) | De (175) | LTR | Common Crawl + WebText2 + Book1 + Book2 + Wikipedia (75TB) | |
| CPM-1(12.2020) | De (2.6) | LTR | Chinese corpus (100GB) | |
| XLNET(06.2019) | De (0.1, 0.3) | LTR | Wikipedia + BookCorpus + Giga5 + ClueWeb + Common Crawl (32.89B words) | |
| Palm(04.2022) | De(540) | LTR | filtered webpages + 2 books +Wikipedia +news articles + source code+ social media conversations(780B tokens) | |
| BART(10.2019) | En-De (0.1, 0.4) | MLM(span) | BookCorpus + CC-News + OPENWEBTEXT + STORIES (160GB) | Translation, summarization and question answering |
| T5(10.2019) | En-De (0.1, 0.2, 0.7, 3, 11) | MLM(span) | C4 Dataset(750 GB) | |
| Switch transformers (01.2021) | En-De (385, 1600) | MLM(span) | C4 Dataset(750 GB) | |

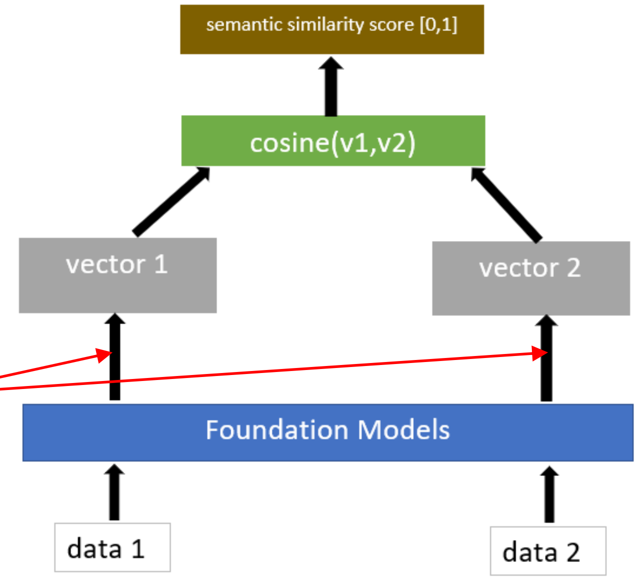
- trained on significant quantities of data
- adapted to a wide range of downstream tasks.

# Conception of Classification System

## Word embedding and Semantic Similarity




Types of word embedding techniques

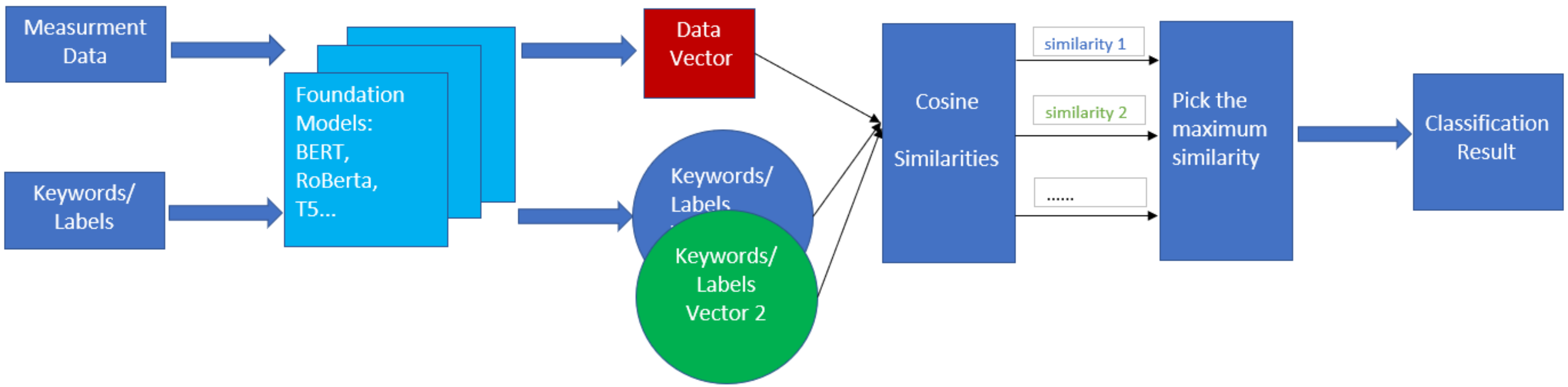



Semantic similarity score based on foundation models

representation of words in NLP
real-valued vector

# Conception of Classification System

## Conception System Diagram

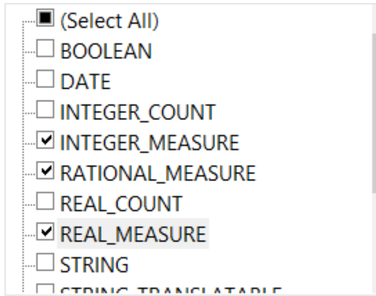



Conception System Diagram

# Implementation of Classification System

## Determination of Dataset

| PreferredName | Definition |
|---|---|
| Material of joint replacement (foot, ankle) | fabric or raw material the joint replacement (foot, ankle) is made of (having regard to its ou |
| step distance | Distance to highest level of a step drill |
| contact radius | radius of that part of the cutting item that can create a defined non replaceable position wi |
| plug length | length of that part of the cutting item that has uncompleted thread pitches that gets first in |
| plug angle | Angle of that part of the cutting edge that is closest to the workpiece, which centers the to |
| cutting edge centre count | number of cutting edges that are able to cut across the center of the tool item axis |
| countersunk depth of connection bore | depth of the cylindrical counter bore of a hole in the centre of a tool item or adaptive item |
| connection diameter, workpiece side, upper deviation | greatest permissible upper deviation from the nominal connection diameter |
| connection diameter, workpiece side, lower deviation | greatest permissible lower deviation from the nominal connection diameter |
| tool style code | attribute designation BLD |
| drilling diameter | cutting part of a combination milling tool that drills a hole |
| stock removal recommended | thickness of material that is recommended to be removed in a reaming operation |
| flange diameter | dimension between two parallel tangents on the outside edge of a flange |
| presetting torque | presetted torque where the transmission of power will be cut to avoid damage to the cutti |
| corner radius | nominal radius of a rounded corner measured in the XY-plane |
| insert included angle | angle between the major and the minor cutting edges of a cutting item |




**21600 terms**

Filter

Data Type

INTERGER_MEASURE
RATIONAL_MEASURE
REAL_MEASURE

**7798 terms**

Filter

dynamic data

measurement data collected by sensors

**144 terms**

# Implementation of Classification System

## Determination of Ctategories,Keywords,Foundation Models

```
dic_clusters = {}
dic_clusters["mechanical"] = ['speed, force, pressure, mass, torque']
dic_clusters["electrical"] = ['electrical, current, potential, voltage']
dic_clusters["thermal"] = ['thermal, temperature, heat']
dic_clusters["magnetic"] = ['magnetic, magnet, inductivity']
dic_clusters["acoustic"] = ['acoustic, noise, sound, ultrasonic']
```

Mechanical: 53
Electrical: 30
Thermal: 33
Magnetic: 8
Acoustic: 20

Final Categories and Keywords

```
model = SentenceTransformer('sentence-t5-base')

# msmarco-bert-base-dot-v5
# msmarco-roberta-base-v2
# sentence-t5-base
```

BERT

RoBERTa

T5

Code to use different foundation models

# Implementation of Classification System

## Concrete Example

```
dic_clusters = {}
dic_clusters["mechanical"] = ['length, width, area, position, velocity, speed, '
                              'acceleration, force, pressure, stress, mass, moment, torque']
dic_clusters["electrical"] = ['current, voltage, potential, electrical']
dic_clusters["thermal"] = ['temperature, heat, thermal']
```

| Measurement data (Name: Definition) | Dim 1 | Dim 2 | … | Dim 768 |
|---|---|---|---|---|
| Rotation speed:... | 0.016 | 0.038 | … | 0.023 |
| Battery potential:... | 0.022 | 0.026 | … | 0.058 |
| … | … | … | … | … |
| Temperature range:... | 0.017 | 0.090 | … | 0.052 |

- 200×768

| Keywords/Labels | Dim 1 | Dim 2 | … | Dim 768 |
|---|---|---|---|---|
| Mechanical | 0.039 | 0.020 | … | 0.057 |
| Electrical | 0.027 | 0.002 | … | 0.016 |
| Thermal | 0.026 | 0.006 | … | 0.036 |

- 3×768

Cosine similarity

| | Mechanical | Electrical | Thermal |
|---|---|---|---|
| Rotation speed:... | 0.5 | 0.1 | 0.1 |
| Battery potential:... | 0.0 | 0.9 | 0.2 |
| … | … | … | … |
| Temperature range:... | 0.3 | 0.5 | 0.7 |

- 3×768

Normalization

| | Mechanical | Electrical | Thermal |
|---|---|---|---|
| Rotation speed:... | 0.71 | 0.14 | 0.14 |
| Battery potential:... | 0.0 | 0.82 | 0.18 |
| … | … | … | … |
| Temperature range:... | 0.20 | 0.33 | 0.47 |

- 3×768

Classification in the perspective of vectors and matrixes

# Evaluation

## Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1 score

```
## Accuracy, Precision, Recall
accuracy = metrics.accuracy_score(y_test, predicted)
auc = metrics.roc_auc_score(y_test, predicted_prob,
                            multi_class="ovr")
print("Accuracy:",  round(accuracy,2))
print("Auc:", round(auc,2))
print("Detail:")
print(metrics.classification_report(y_test, predicted))
```

Code to generate classification report

- Confusion Matrix

```
## Plot confusion matrix
cm = metrics.confusion_matrix(y_test, predicted)
fig, ax = plt.subplots()
sns.heatmap(cm, annot=True, fmt='d', ax=ax, cmap=plt.cm.Reds,
            cbar=False)
ax.set(xlabel="Pred", ylabel="True", xticklabels=classes,
       yticklabels=classes, title="Confusion matrix")
plt.yticks(rotation=0)
fig, ax = plt.subplots(nrows=1, ncols=2)
```

Code to plot confusion matrix

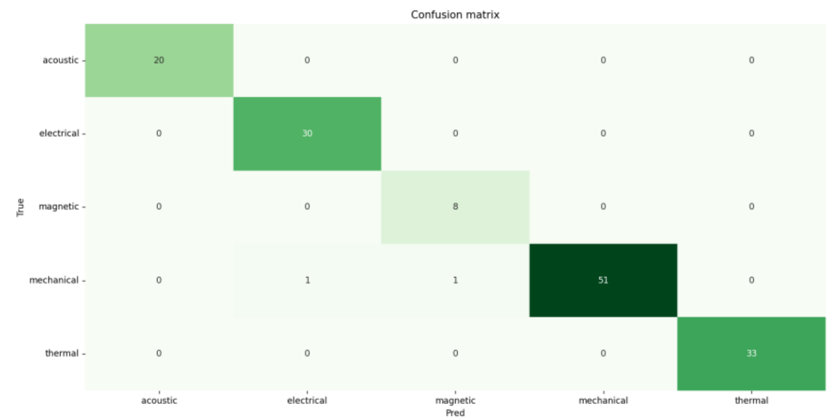
# Evaluation

## Evaluation Results of BERT

Detail:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| acoustic | 1.00 | 1.00 | 1.00 | 20 |
| electrical | 0.97 | 1.00 | 0.98 | 30 |
| magnetic | 0.89 | 1.00 | 0.94 | 8 |
| mechanical | 1.00 | 0.96 | 0.98 | 53 |
| thermal | 1.00 | 1.00 | 1.00 | 33 |
| accuracy | | | 0.99 | 144 |
| macro avg | 0.97 | 0.99 | 0.98 | 144 |
| weighted avg | 0.99 | 0.99 | 0.99 | 144 |

Accuracy, precision, recall, and the F1 score of BERT




Confusion Matrix of BERT

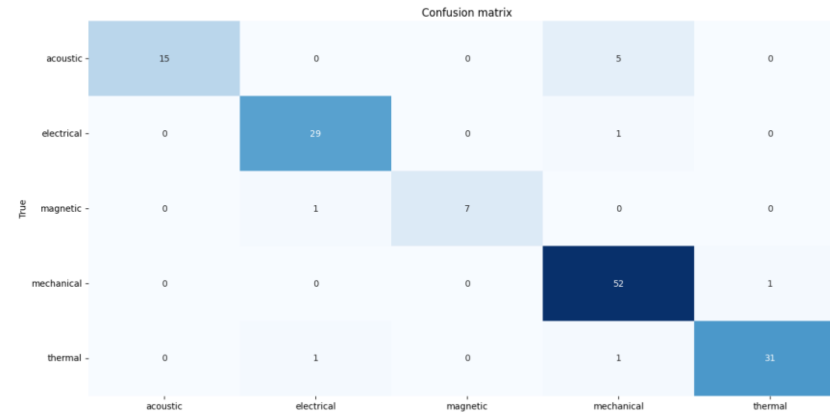
# Evaluation

## Evaluation Results of RoBERTa

Detail:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| acoustic | 1.00 | 0.75 | 0.86 | 20 |
| electrical | 0.94 | 0.97 | 0.95 | 30 |
| magnetic | 1.00 | 0.88 | 0.93 | 8 |
| mechanical | 0.88 | 0.98 | 0.93 | 53 |
| thermal | 0.97 | 0.94 | 0.95 | 33 |
| accuracy | | | 0.93 | 144 |
| macro avg | 0.96 | 0.90 | 0.92 | 144 |
| weighted avg | 0.94 | 0.93 | 0.93 | 144 |

Accuracy, precision, recall, and the F1 score of RoBERTa




Confusion Matrix of RoBERTa

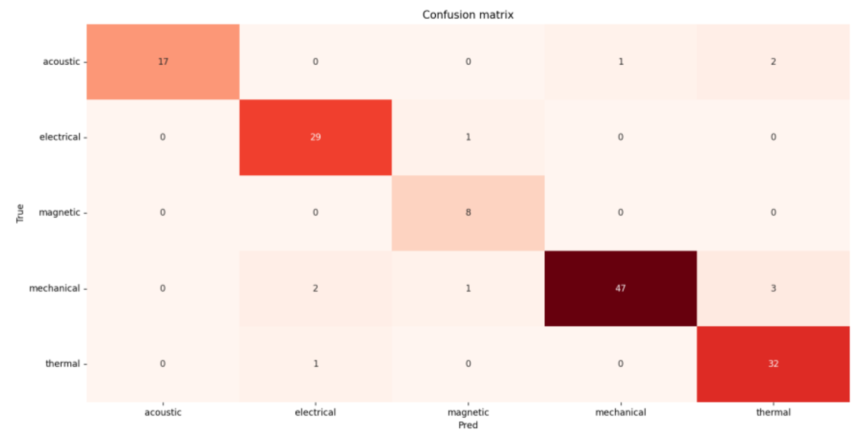
# Evaluation

## Evaluation Results of T5

Detail:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| acoustic | 1.00 | 0.85 | 0.92 | 20 |
| electrical | 0.91 | 0.97 | 0.94 | 30 |
| magnetic | 0.80 | 1.00 | 0.89 | 8 |
| mechanical | 0.98 | 0.89 | 0.93 | 53 |
| thermal | 0.86 | 0.97 | 0.91 | 33 |
| accuracy | | | 0.92 | 144 |
| macro avg | 0.91 | 0.93 | 0.92 | 144 |
| weighted avg | 0.93 | 0.92 | 0.92 | 144 |

Accuracy, precision, recall, and the F1 score of T5




Confusion Matrix of T5

# Evaluation

## Results Comparision

| | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| BERT | 0.99 | 0.97 | 0.99 | 0.98 |
| RoBERTa | 0.93 | 0.96 | 0.90 | 0.92 |
| T5 | 0.92 | 0.91 | 0.93 | 0.92 |

| | Run Time |
|---|---|
| BERT | 17.9 s |
| RoBERTa | 18.3 s |
| T5 | 19.4 s |

BERT model performs better than the other two models in this classification task

# Evaluation

## Evaluation results in different conditions

```
dic_clusters = {}
dic_clusters["time"] = ['time, date, timestamp']
dic_clusters["length"] = ['length, width']
dic_clusters["area"] = ['area, volume']
dic_clusters["mass"] = ['mass, gravity, density']
dic_clusters["temperature"] = ['temperature, heat']
dic_clusters["electrical"] = ['voltage, current, power']
dic_clusters["speed"] = ['speed, rotation, acceleration']
```

| | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| BERT | 0.89 | 0.91 | 0.89 | 0.89 |
| RoBERTa | 0.90 | 0.91 | 0.90 | 0.90 |
| T5 | 0.85 | 0.87 | 0.85 | 0.85 |

```
dic_clusters["mechanical"] = ['mechanical, position, speed, force, pressure, mass, torque']
dic_clusters["electrical"] = ['electrical, current, potential, voltage']
dic_clusters["thermal"] = ['thermal, temperature, heat']
dic_clusters["magnetic"] = ['magnetic, magnetic field, magnetic flux, permeability']
dic_clusters["acoustic"] = ['acoustic, wave amplitude, wave velocity, phase']
```

| | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| BERT | 0.82 | 0.81 | 0.81 | 0.79 |
| RoBERTa | 0.76 | 0.73 | 0.75 | 0.72 |
| T5 | 0.88 | 0.86 | 0.90 | 0.88 |

T5 has strong robustness

Factors affecting classification results:
1. Foundation models
2. Determination of dataset
3. Determination of categories
4. Determination of keywords

# Conclusion and Future Work

- Conclusion:

- Designed a classification system to classify the measurement data in ECLASS
- Selected datasets, categories and keywords to accomplish classification task
- Evaluated the performance of three foundation models: BERT, RoBERTa and T5

- Future work:

- fine-tuning the foundation models
- Increase the number of words in the dataset

University of Stuttgart
Institut of Industrial Automation
and Software Engineering

# Thank you!

**Yang Jiang**

e-mail st170646@stud.uni-stuttgart.de
phone +49 (0) 711 685-
fax +49 (0) 711 685-

University of Stuttgart

# References

[1] T. Kalsoom, N. Ramzan, S. Ahmed, and M. Ur-Rehman, “Advances in Sensor Technologies in the Era of Smart Factory and Industry 4.0,” Sensors, vol. 20, no. 23, p. 6783, Nov. 2020.
[2] ECLASS, “Conceptual data model,” ECLASS.
[3] A. A. Ensafi, “An introduction to sensors and biosensors,” Electrochemical Biosensors, pp. 1–10, 2019.
[4] Autoren der Wikimedia-Projekte, “Gerät zur Bestimmung physikalischer Größen,” Wikipedia.org, Dec. 11, 2002.
[5] Li Q, Peng H, Li J, et al. “A survey on text classification: From shallow to deep learning.” arXiv preprint arXiv:2008.00364, 2020.
[6] Modern Approaches in Natural Language Processing. https://slds-lmu.github.io/seminar_nlp_ss20/