

Klassifizierung von E-Mails



Klassifizierung von E-Mails

Mithilfe Künstlicher Intelligenz (KI)

Betrieb: Sylvenstein Media GmbH



Betrieb: Sylvenstein Media GmbH



- Kleines IT-Unternehmen



Betrieb: Sylvenstein Media GmbH



- Kleines IT-Unternehmen
- Ohne Verwaltung



Betrieb: Sylvenstein Media GmbH



- Kleines IT-Unternehmen
- Ohne Verwaltung
- Modul zur Anbindung eines Mail-Servers



Betrieb: Sylvenstein Media GmbH



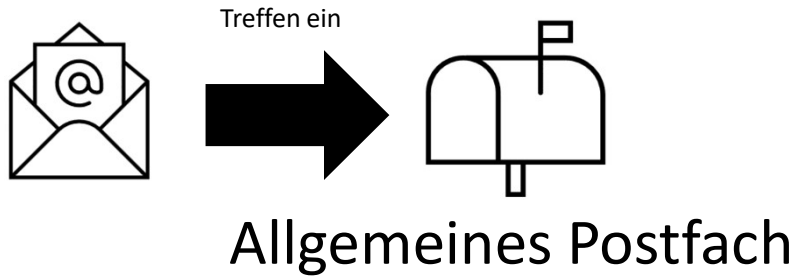
- Kleines IT-Unternehmen
- Ohne Verwaltung
- Modul zur Anbindung eines Mail-Servers



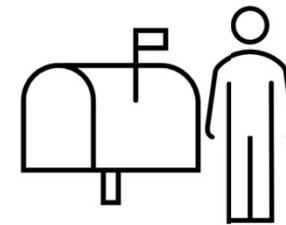
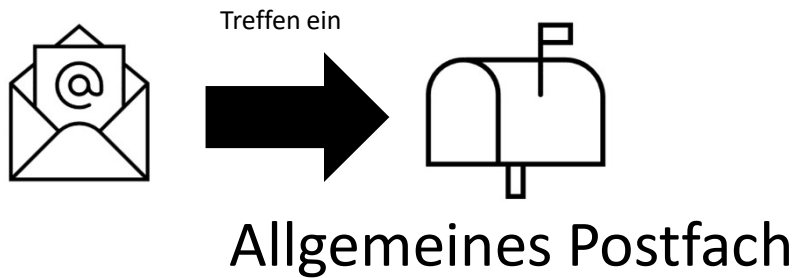
Projektbegründung:



Projektbegründung:



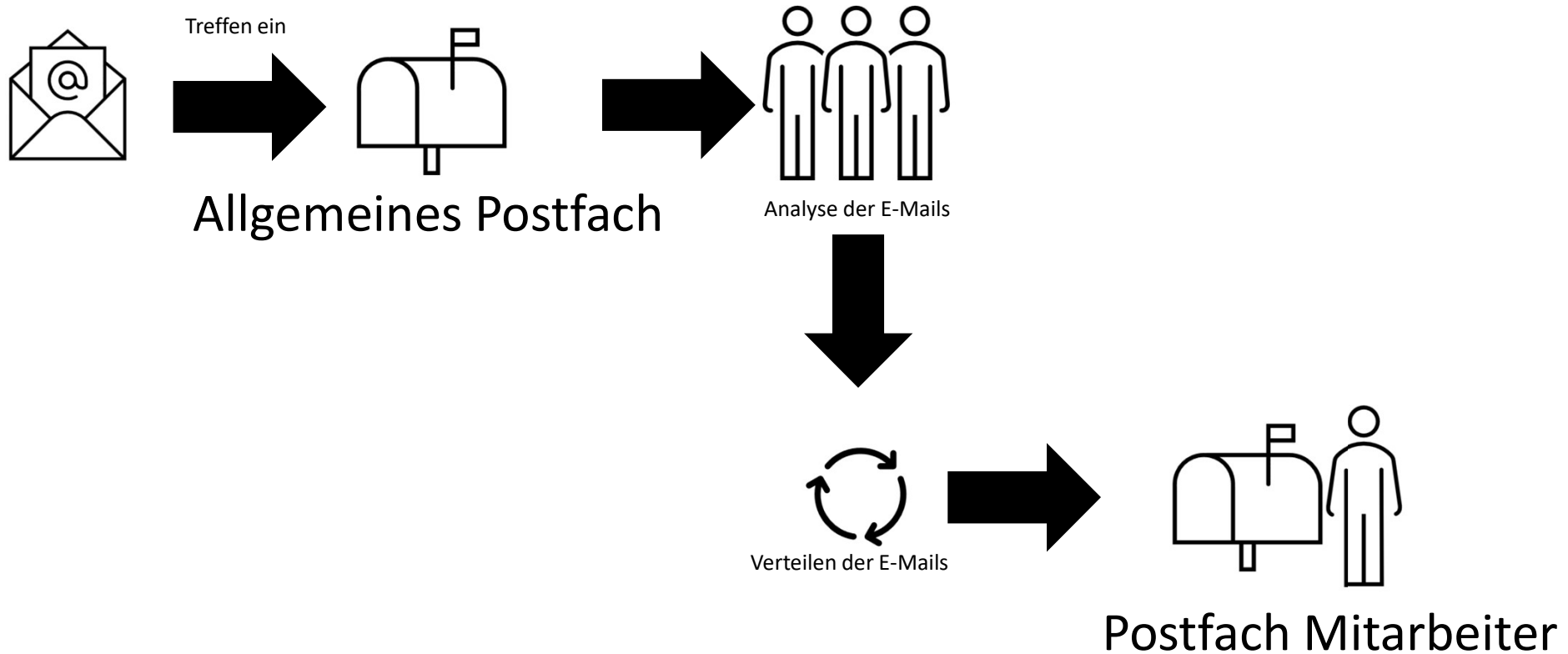
Projektbegründung:



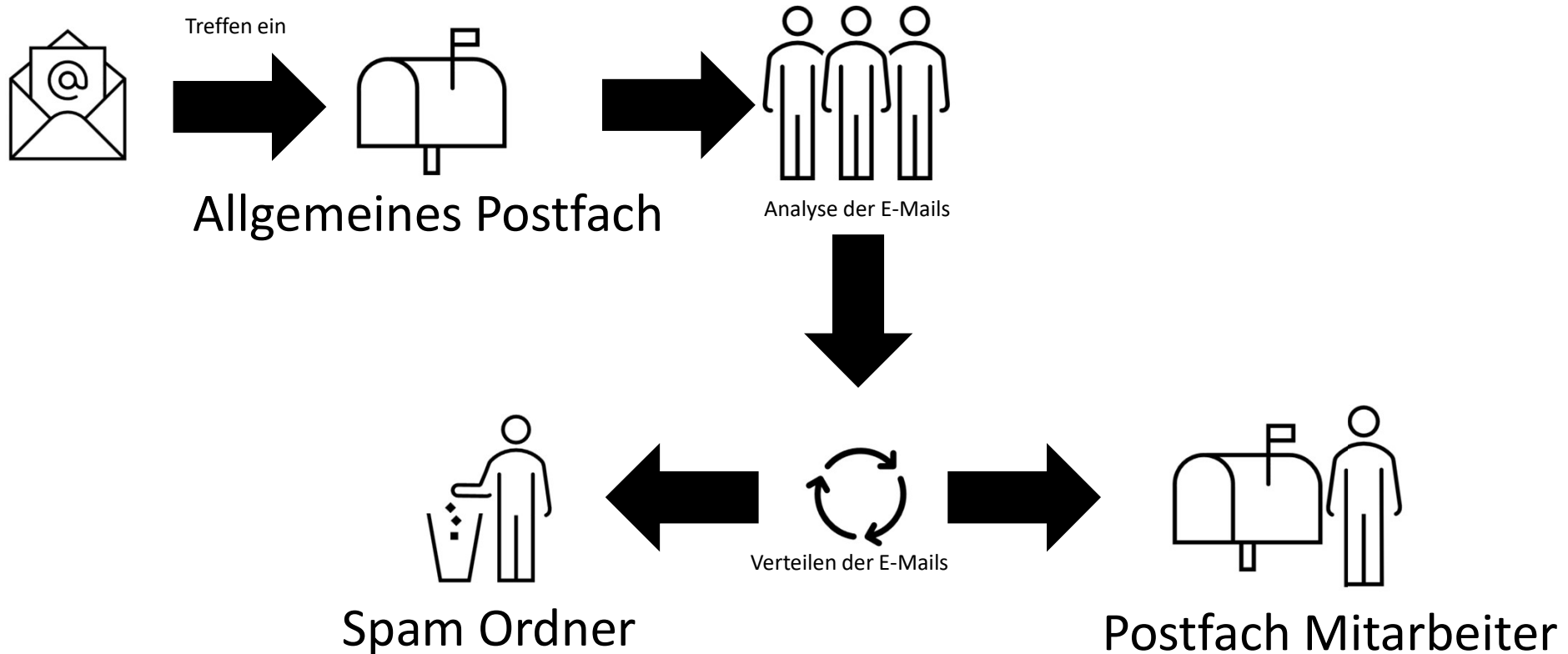
Postfach Mitarbeiter



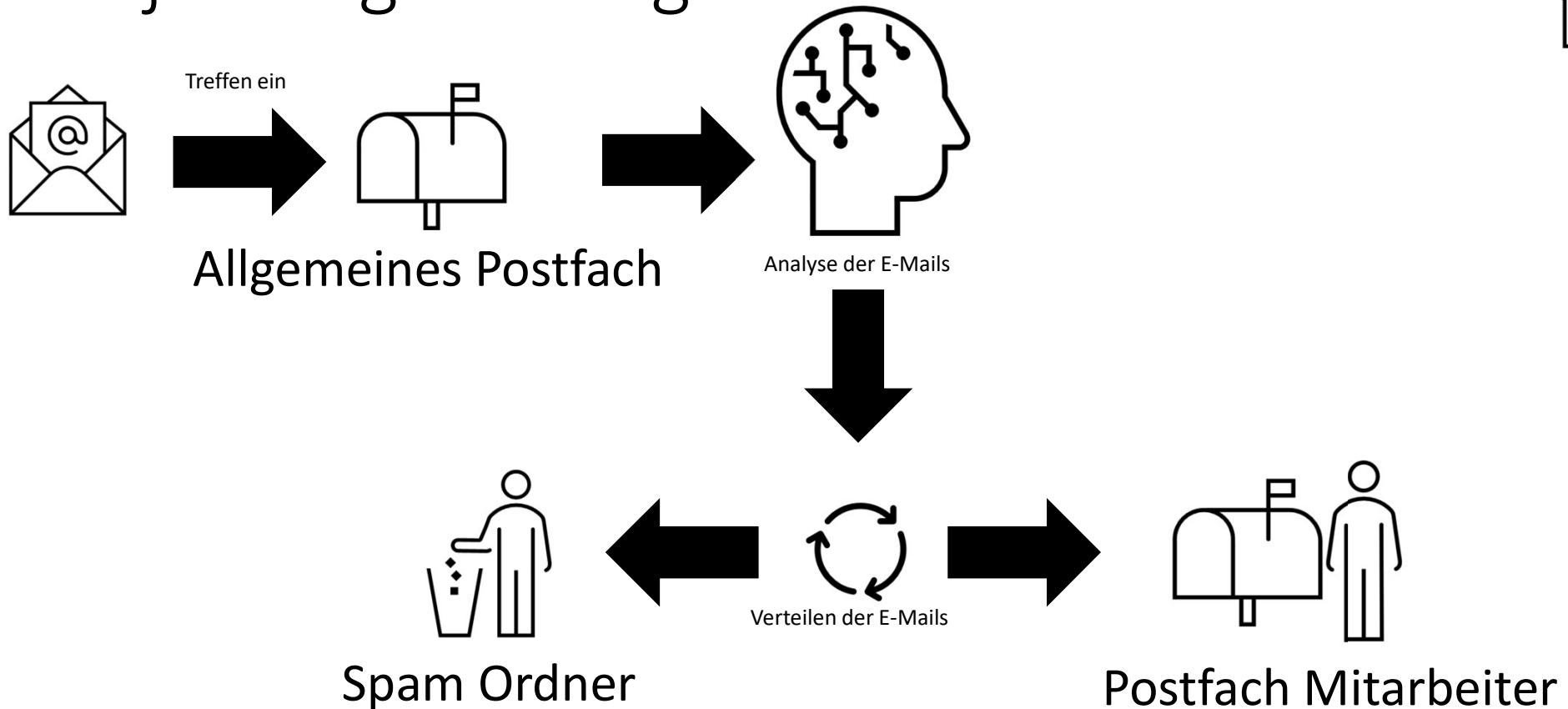
Projektbegründung:



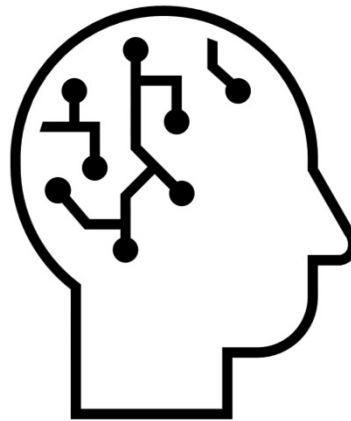
Projektbegründung:



Projektbegründung:



Projektbegründung:



Wissenserwerb

Planung des Projektes



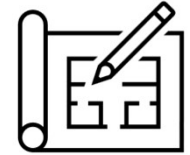
Planung des Projektes



Mail-Server



Planung des Projektes



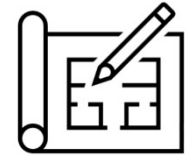
Mail-Server



Datenbank



Planung des Projektes



Mail-Server



Datenbank



KI



Planung des Projektes



Mail-Server



Semi



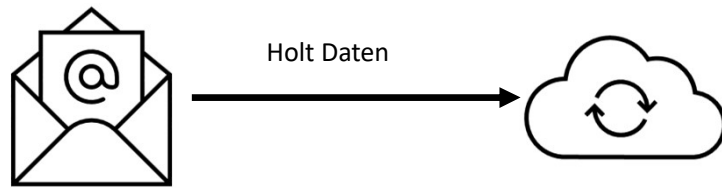
Datenbank



KI



Planung des Projektes



Mail-Server

Semi



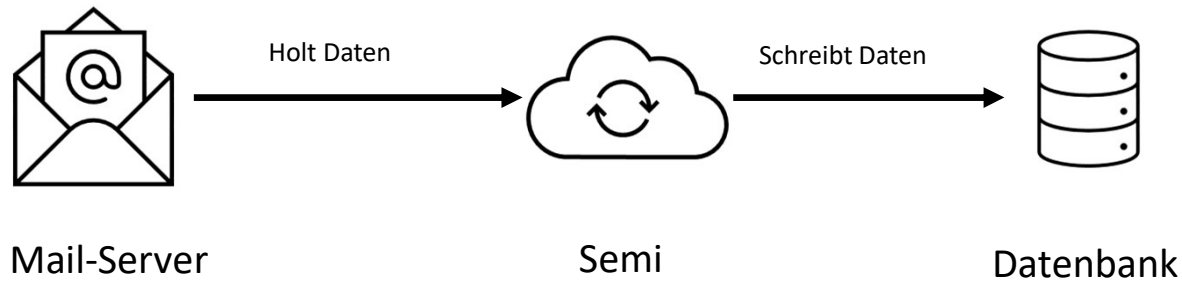
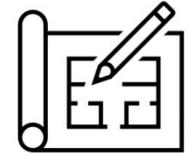
Datenbank



KI



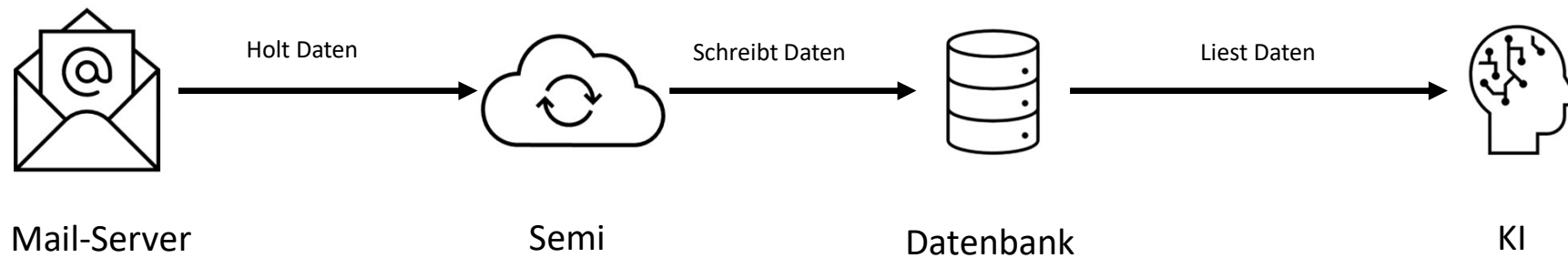
Planung des Projektes



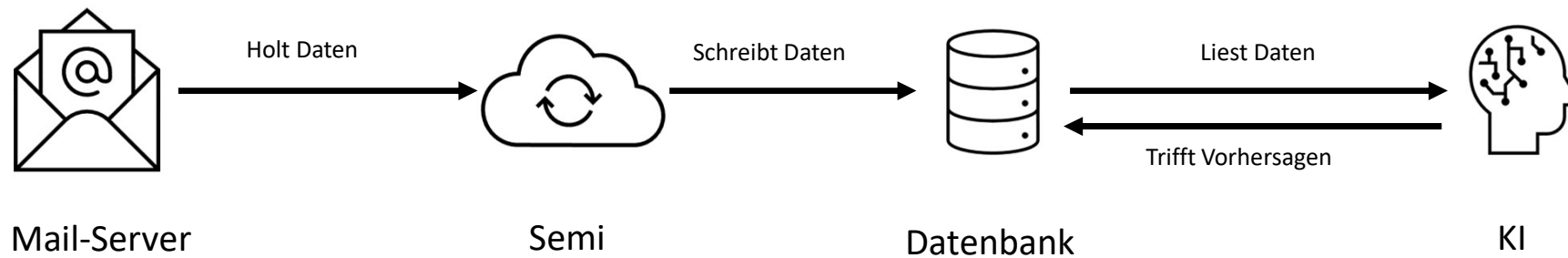
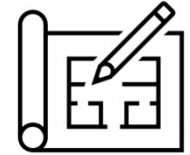
KI



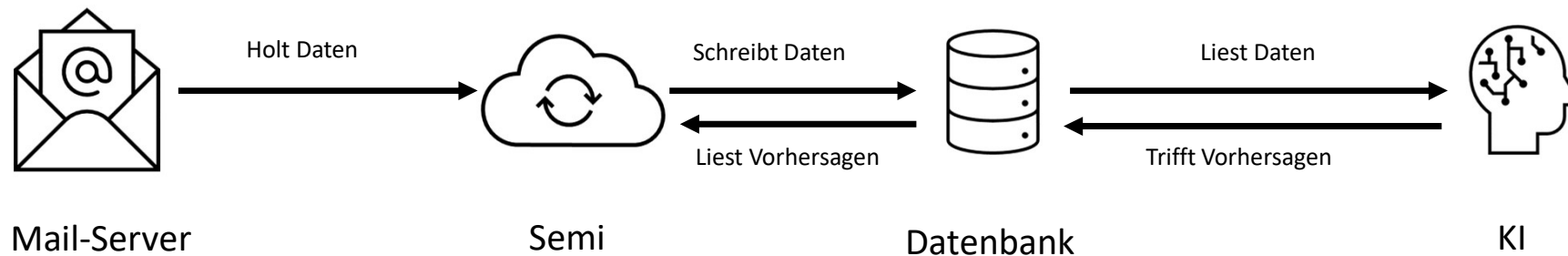
Planung des Projektes



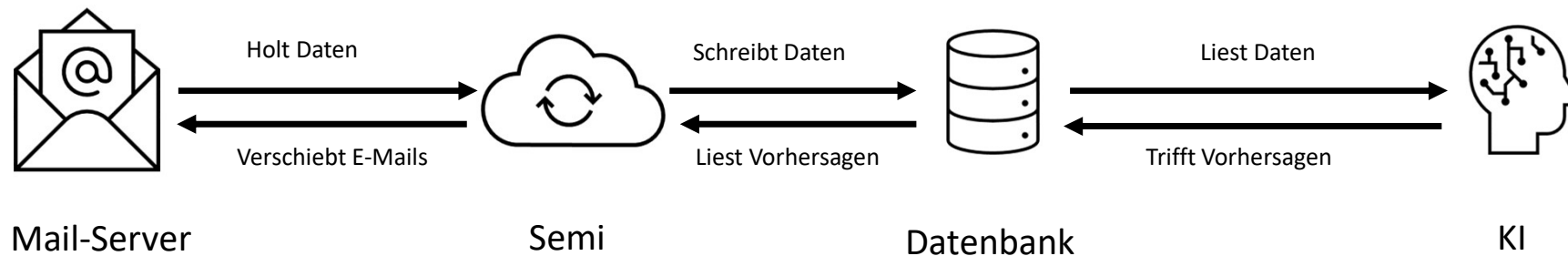
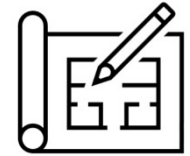
Planung des Projektes



Planung des Projektes



Planung des Projektes



Programmiersprache: Python



Programmiersprache: Python

- Weit verbreitet



Programmiersprache: Python



- Weit verbreitet



- Viele Bibliotheken



Programmiersprache: Python



- Weit verbreitet



- Viele Bibliotheken



- KI-Bibliotheken



Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Datenbank: MySQL



Datenbank: MySQL



- Visualisierung der Datenverarbeitung



Datenbank: MySQL



- Visualisierung der Datenverarbeitung
- Performance



Datenbank: MySQL



- Visualisierung der Datenverarbeitung
- Performance
- Netzwerkschnittstelle (verschlüsselt)



Datenbank: MySQL



- Visualisierung der Datenverarbeitung
- Performance
- Netzwerkschnittstelle (verschlüsselt)

→ Zentrale Schnittstelle zwischen Mail-Server & KI-Model



Google: Colaboratory (Colab)



Google: Colaboratory (Colab)



+ Vorteile:

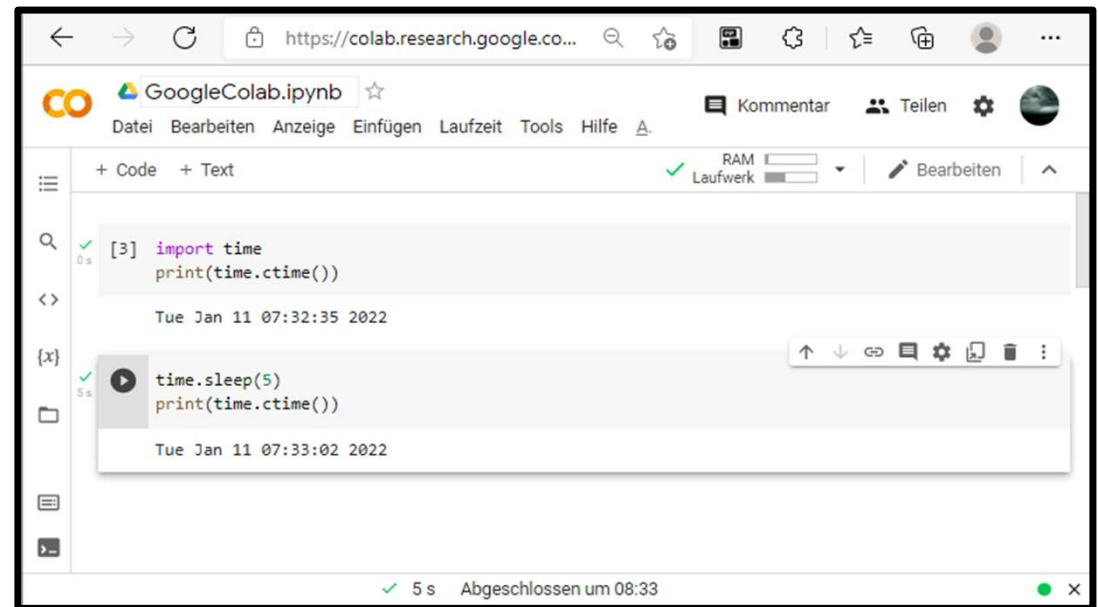


Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung



Einleitung



Analyse



Planung



Design



Durchführung



Test



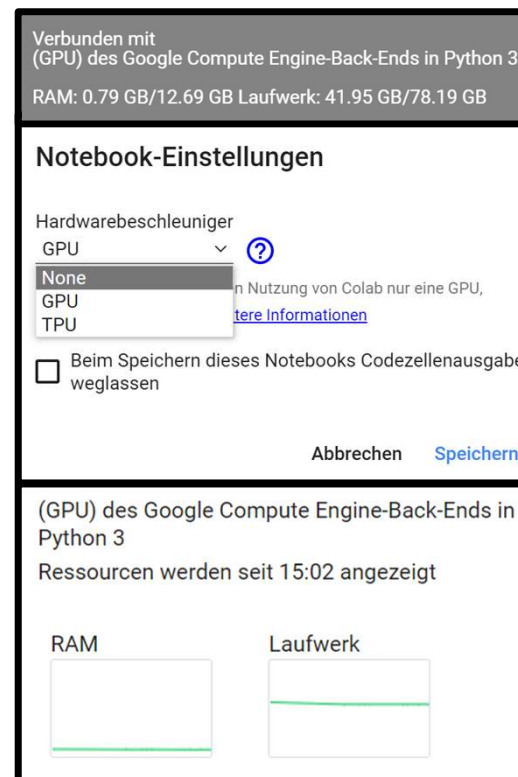
Fazit

Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung



Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung
- Konfigurationszeit



Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung
- Konfigurationszeit
- Aktueller Standard



Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung
- Konfigurationszeit
- Aktueller Standard

- Nachteile:



Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung
- Konfigurationszeit
- Aktueller Standard

- Nachteile:

- Datenschutz (DSGVO)



Google: Colaboratory (Colab)



+ Vorteile:

- Cloud Computing Lösung
- Kostenlose Rechenleistung
- Konfigurationszeit
- Aktueller Standard

- Nachteile:

- Datenschutz (DSGVO)
- Serverstandort nicht frei wählbar

```
"ip": "104.155.213.201",  
"hostname": "201.213.155.104.bc.googleusercontent.com",  
"city": "Taipei",  
"region": "Taiwan",  
"country": "TW",  
"loc": "25.0478,121.5319",  
"org": "AS15169 Google LLC",  
"timezone": "Asia/Taipei",  
"readme": "https://ipinfo.io/missingauth"
```



Textvorverarbeitung



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)

Mit Rauschen	Geräusch entfernt
#Ärger	Ärger
<a>Ärger<a>	Ärger
1.Ärger	Ärger



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)

- Normalisierung

- Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen

Mit Stoppwörter → „Wann ist der Banner fertig ?“
Ohne Stoppwörter → „Wann Banner fertig“



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)

Unbearbeitet	Lemma
Gut	Gut
Besser	Gut



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)

Unbearbeitet	Lemma
Gut	Gut
Besser	Gut

Unbearbeitet	Stammform
Reduktion	Redu
Reduzieren	Redu
reduziertes	Redu



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)
 - Tokenisierung (Tokenization)

Text	Der	Hund	Sitzt	In	Hütte
Der Hund sitzt					
Der Hund sitzt in seiner Hütte					



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)
 - Tokenisierung (Tokenization)

Text	Der	Hund	Sitzt	In	Hütte
Der Hund sitzt					
Der Hund sitzt in seiner Hütte					



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)
 - Tokenisierung (Tokenization)

Text	Der	Hund	Sitzt	In	Hütte
Der Hund sitzt					
Der Hund sitzt in seiner Hütte					



Textvorverarbeitung



- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)
 - Tokenisierung (Tokenization)

Text	Der	Hund	Sitzt	In	Hütte
Der Hund sitzt	1	1	1	0	0
Der Hund sitzt in seiner Hütte	1	1	1	1	1



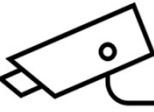
Textvorverarbeitung



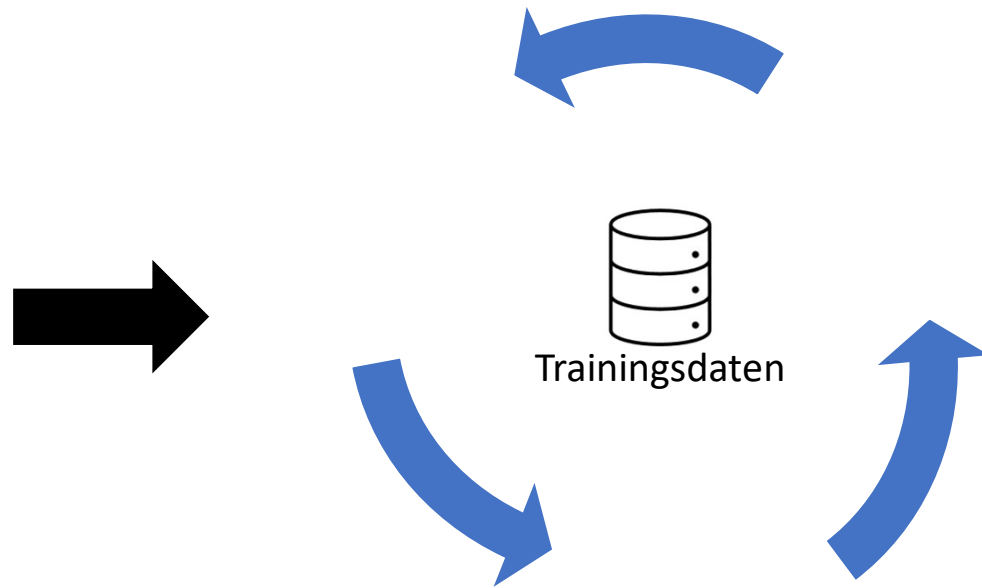
- Wichtiger Teilschritt bei KI-Projekten (NLP)
 - Normalisierung
 - Geräuschentfernung (Noise Removal)
 - Stoppwörter entfernen
 - Lemmatisierung (Lemmatisation)
 - Stammformreduktion (Stemming)
 - Tokenisierung (Tokenization)

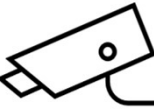
Text	Der	Hund	Sitzt	In	Hütte	Mit
Der Hund sitzt	1	1	1	0	0	0
Der Hund sitzt in seiner Hütte	1	1	1	1	1	0
Der Hund mit Hütte	1	1	0	0	1	1



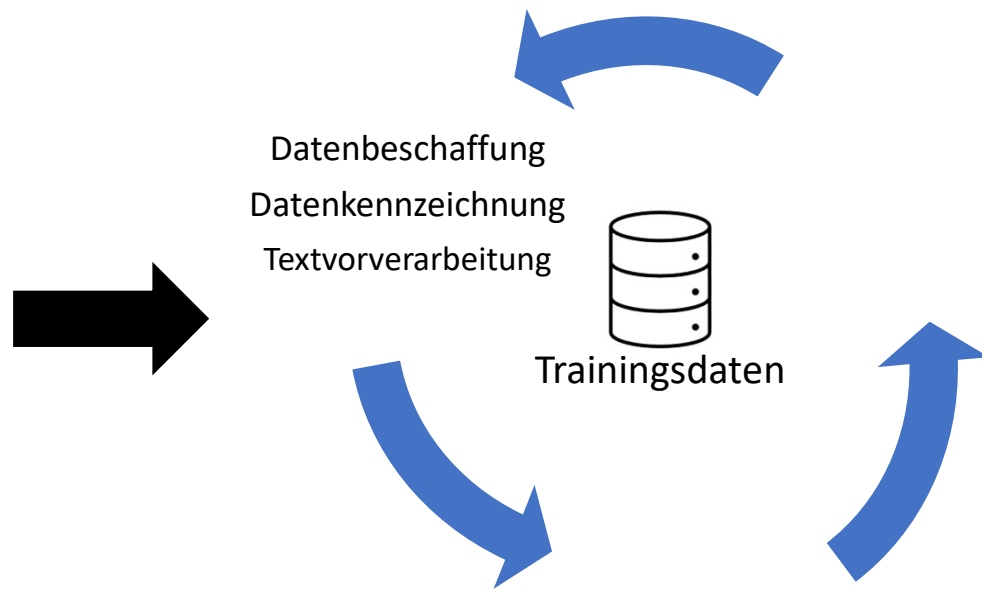


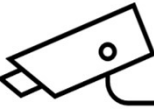
Überwachtes Lernen (Supervised Learning)



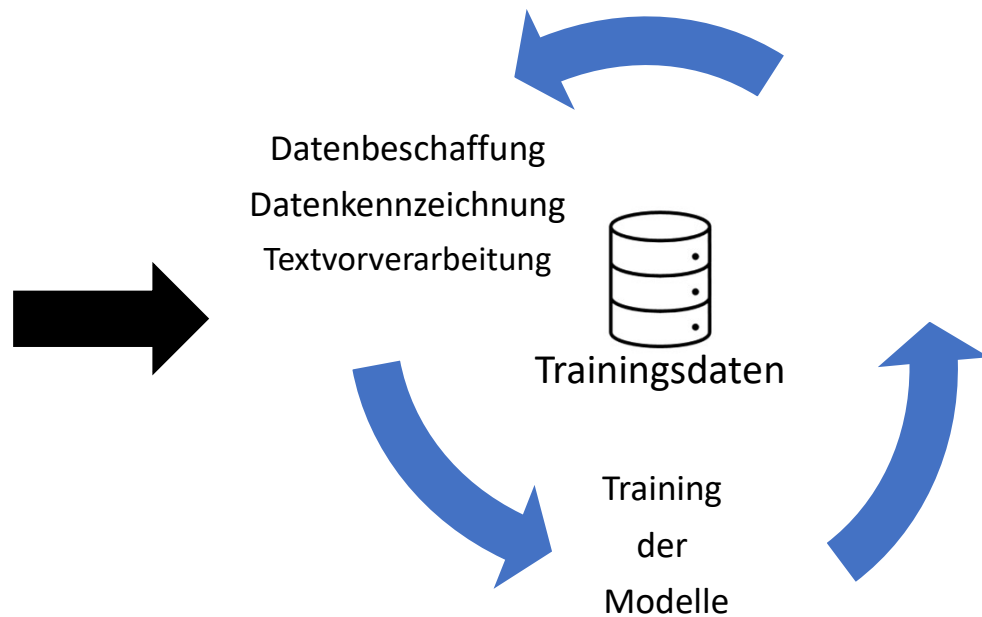


Überwachtes Lernen (Supervised Learning)



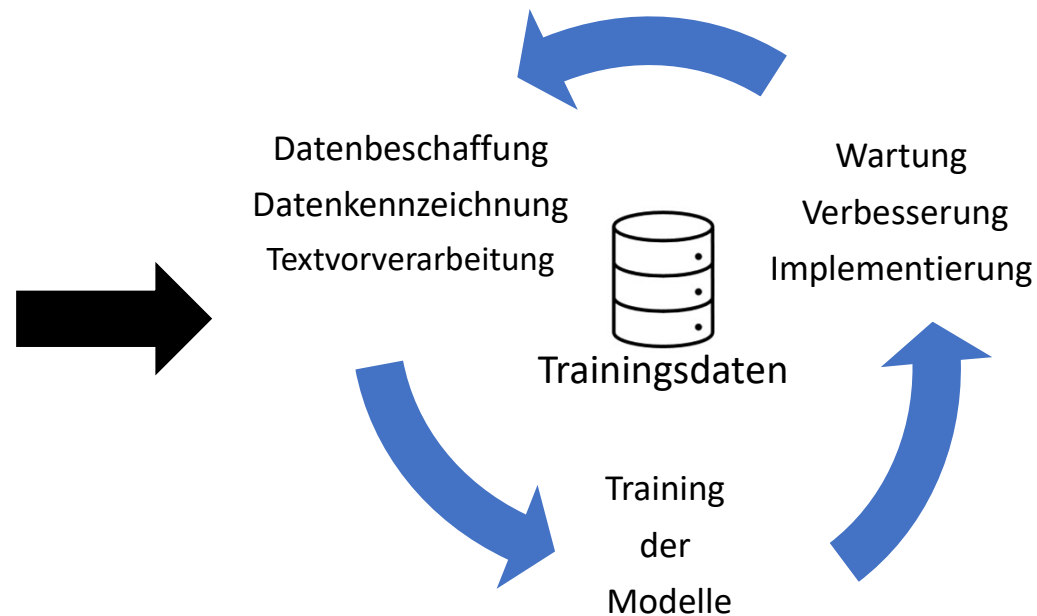


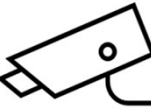
Überwachtes Lernen (Supervised Learning)



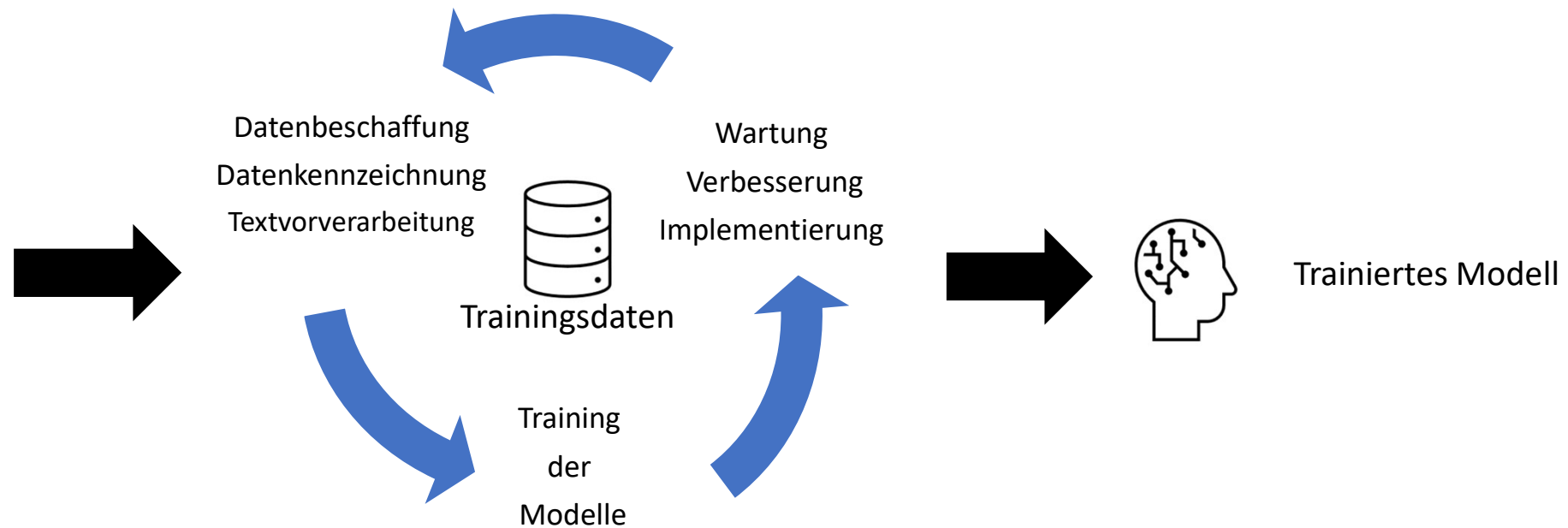


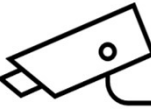
Überwachtes Lernen (Supervised Learning)



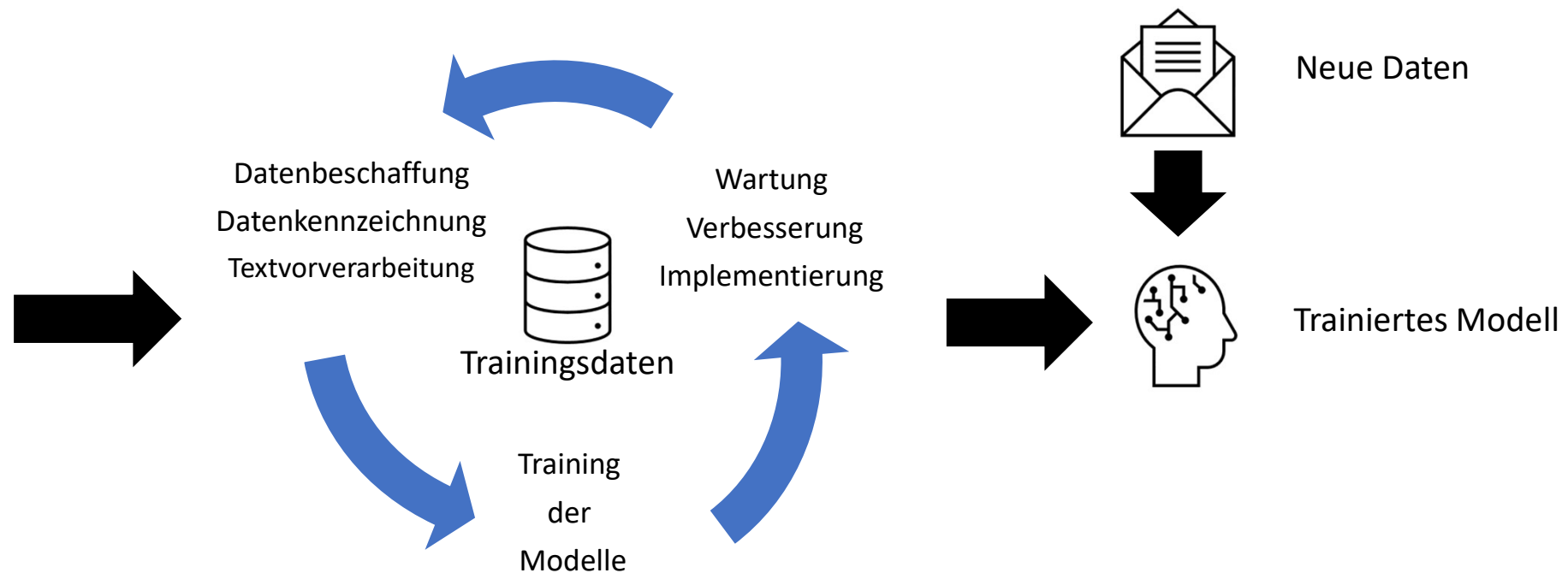


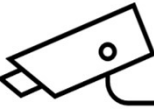
Überwachtes Lernen (Supervised Learning)



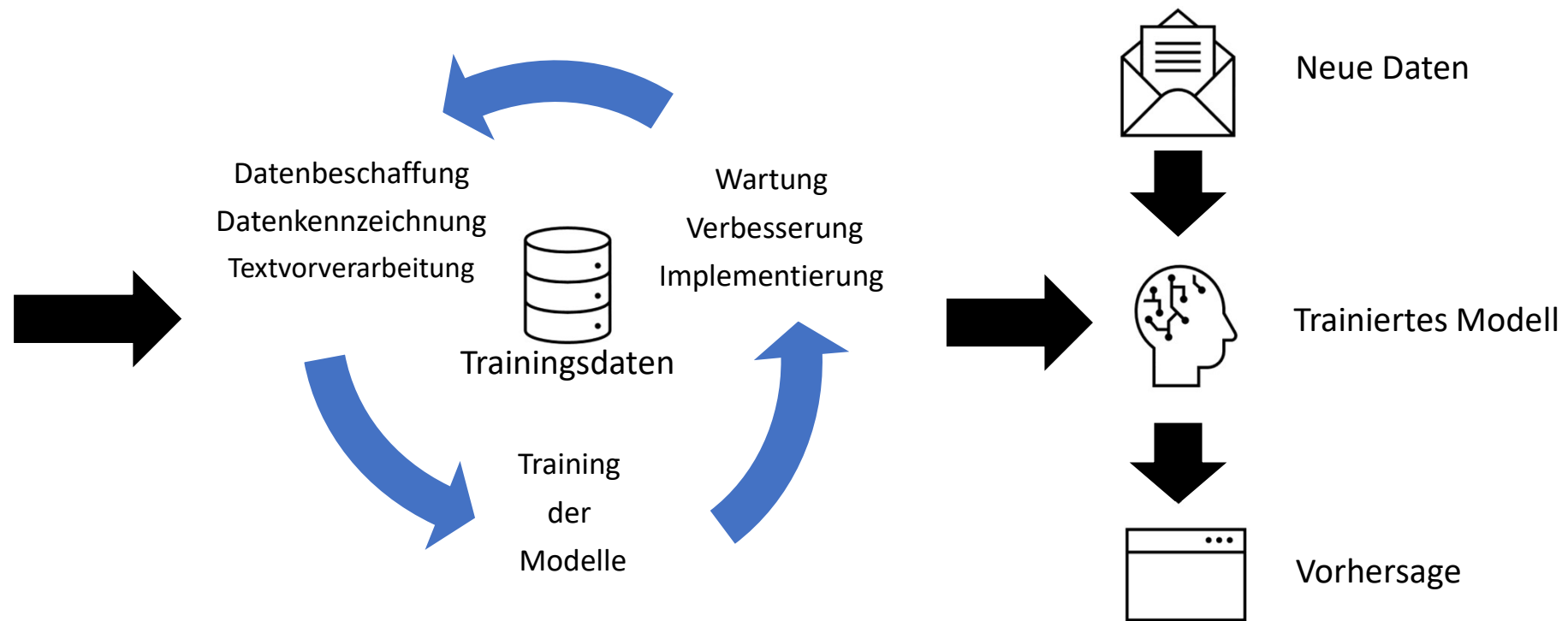


Überwachtes Lernen (Supervised Learning)





Überwachtes Lernen (Supervised Learning)



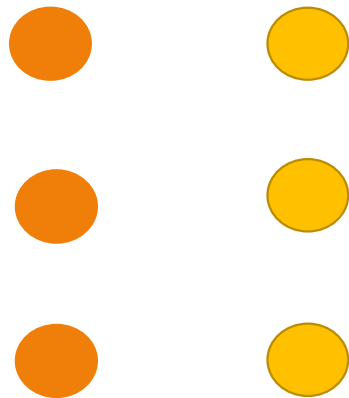
Neuronales Netzwerk



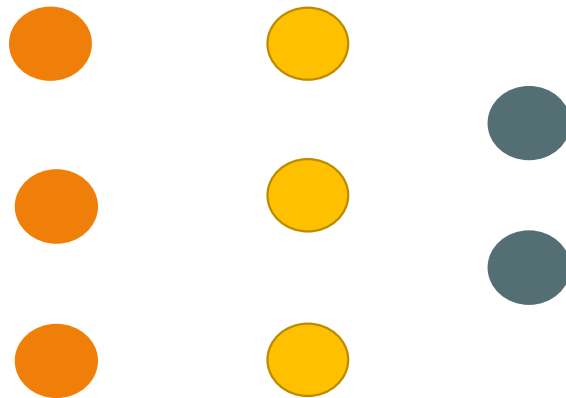
Neuronales Netzwerk



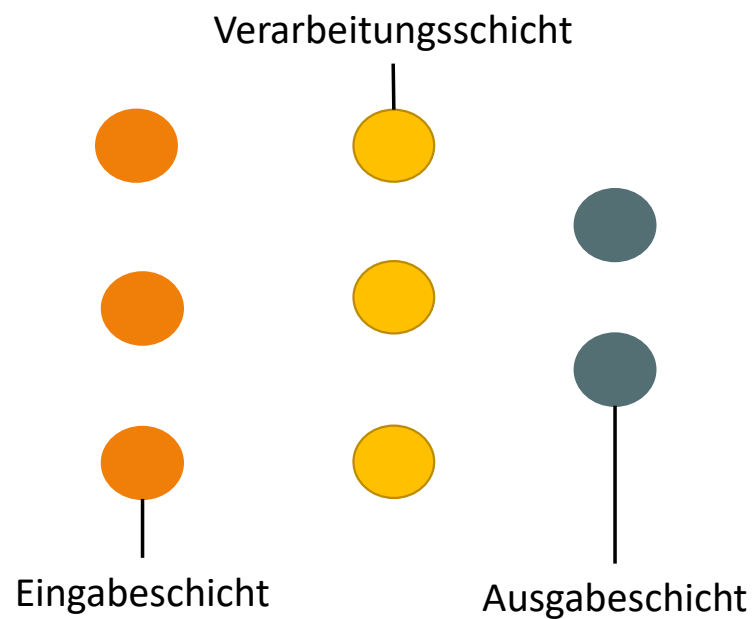
Neuronales Netzwerk



Neuronales Netzwerk



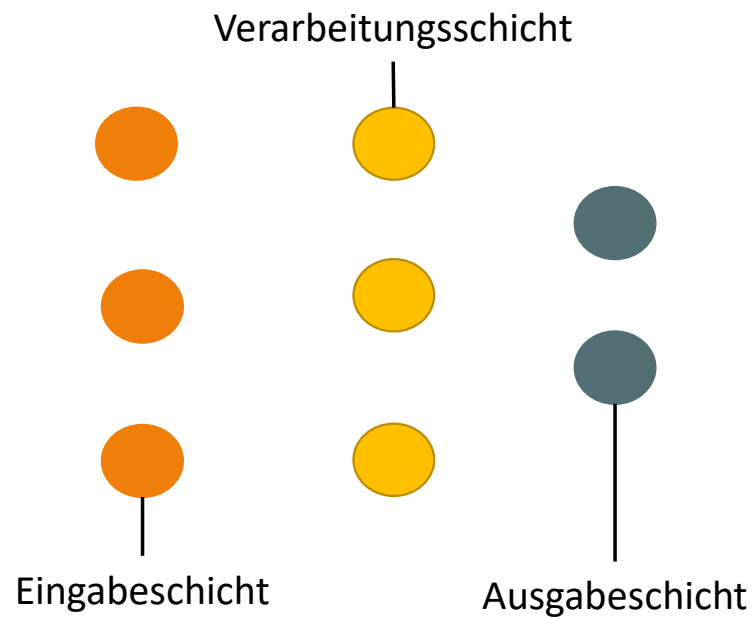
Neuronales Netzwerk



Neuronales Netzwerk



Eingabe:

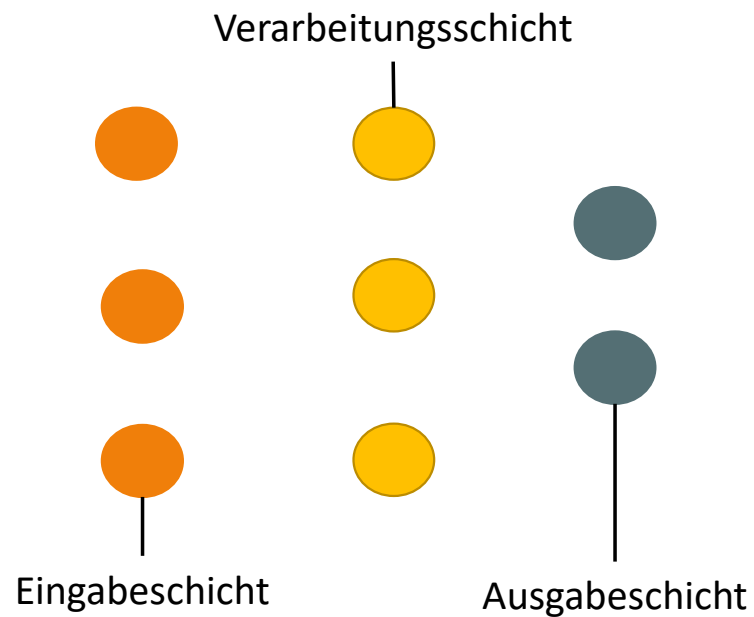


Neuronales Netzwerk



Eingabe:

Ausgabe:

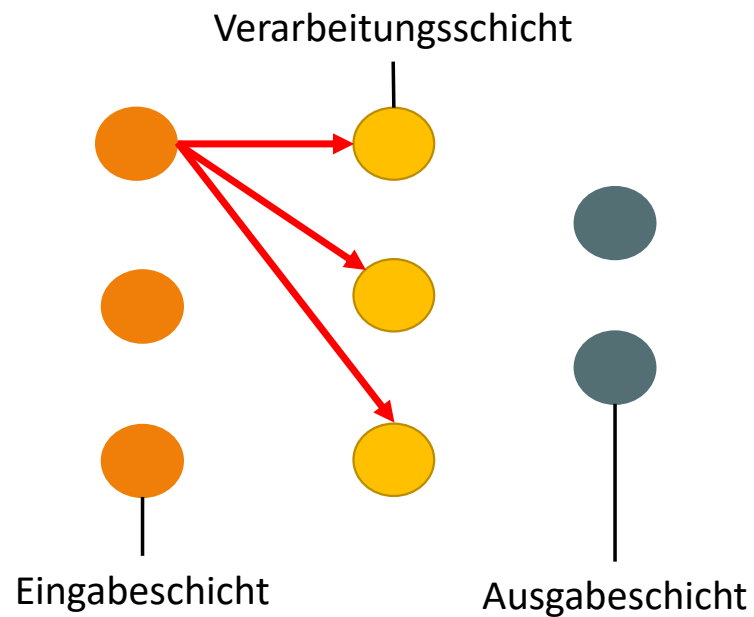


Neuronales Netzwerk



Eingabe:

Ausgabe:

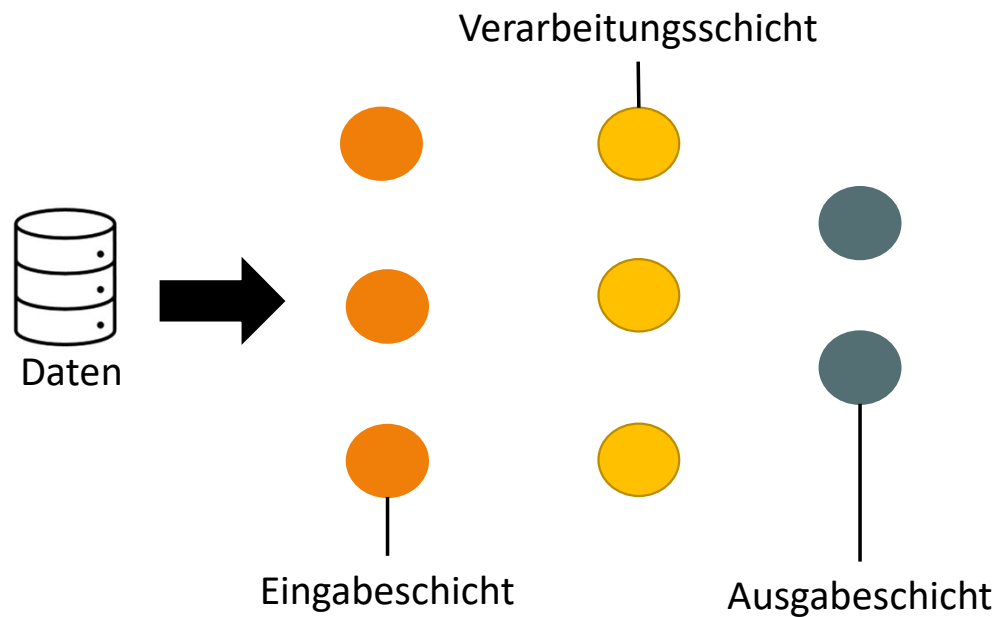


Neuronales Netzwerk



Eingabe:

Ausgabe:

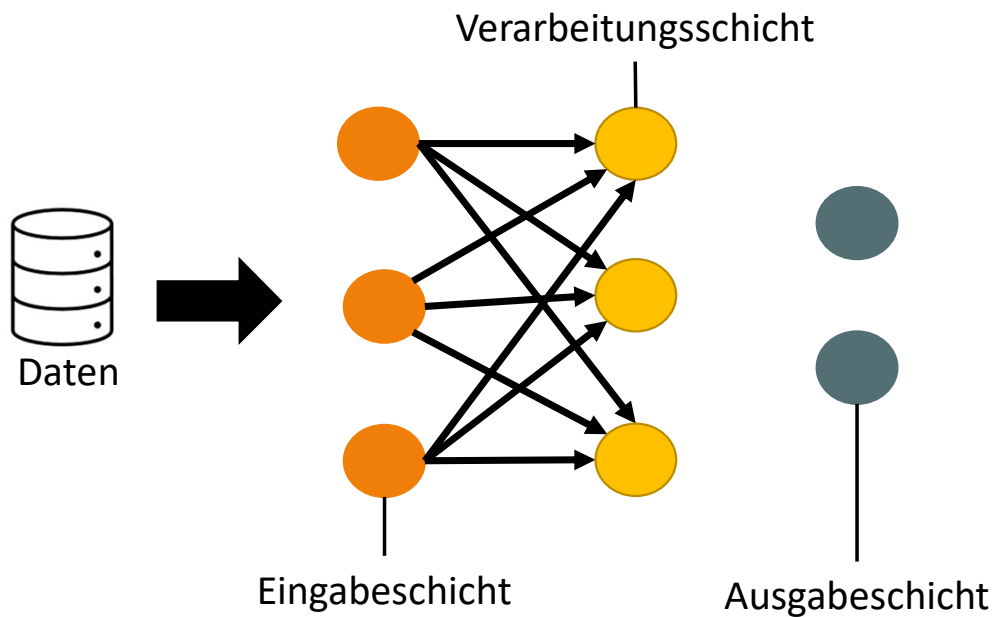


Neuronales Netzwerk



Eingabe:

Ausgabe:

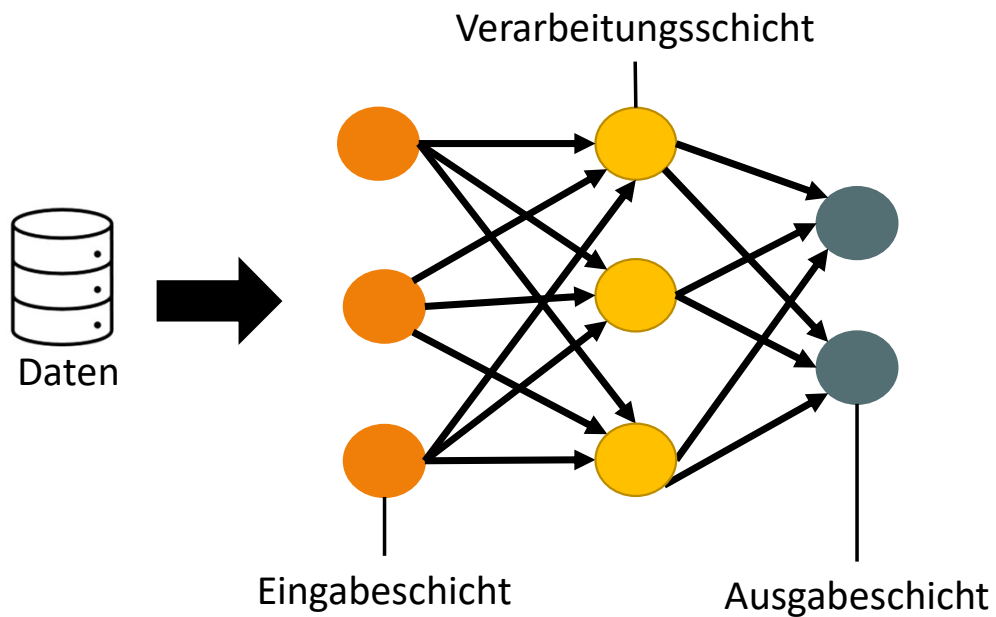


Neuronales Netzwerk



Eingabe:

Ausgabe:

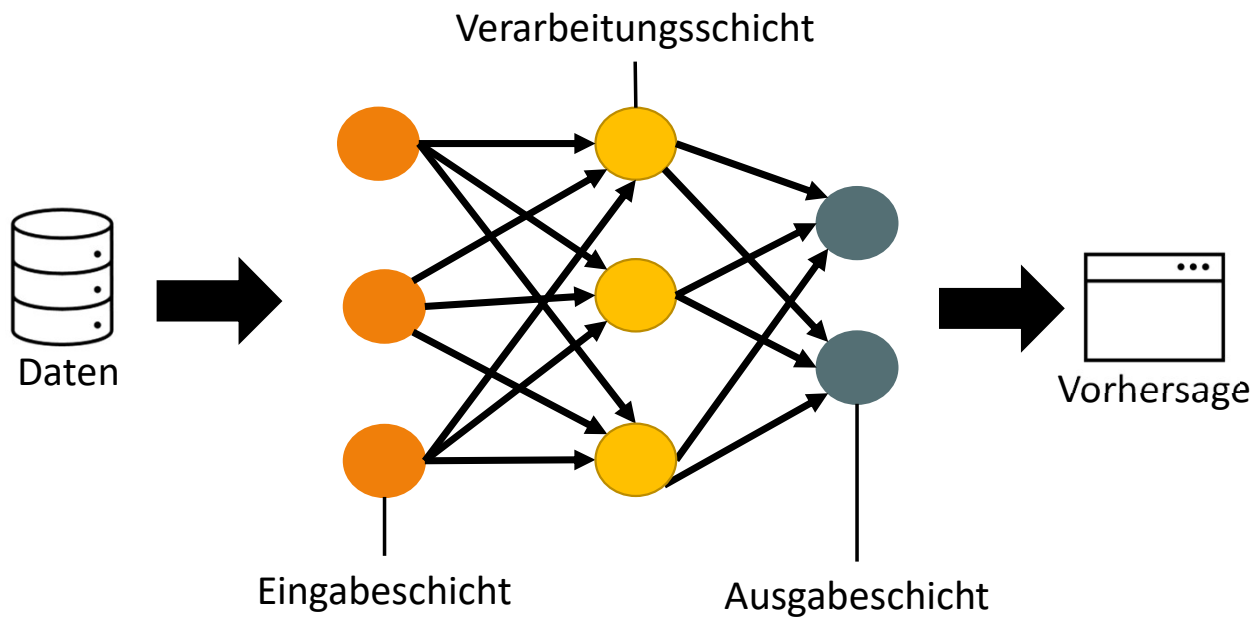


Neuronales Netzwerk



Eingabe:

Ausgabe:



Einleitung

Analyse

Planung

Design

Durchführung

Test

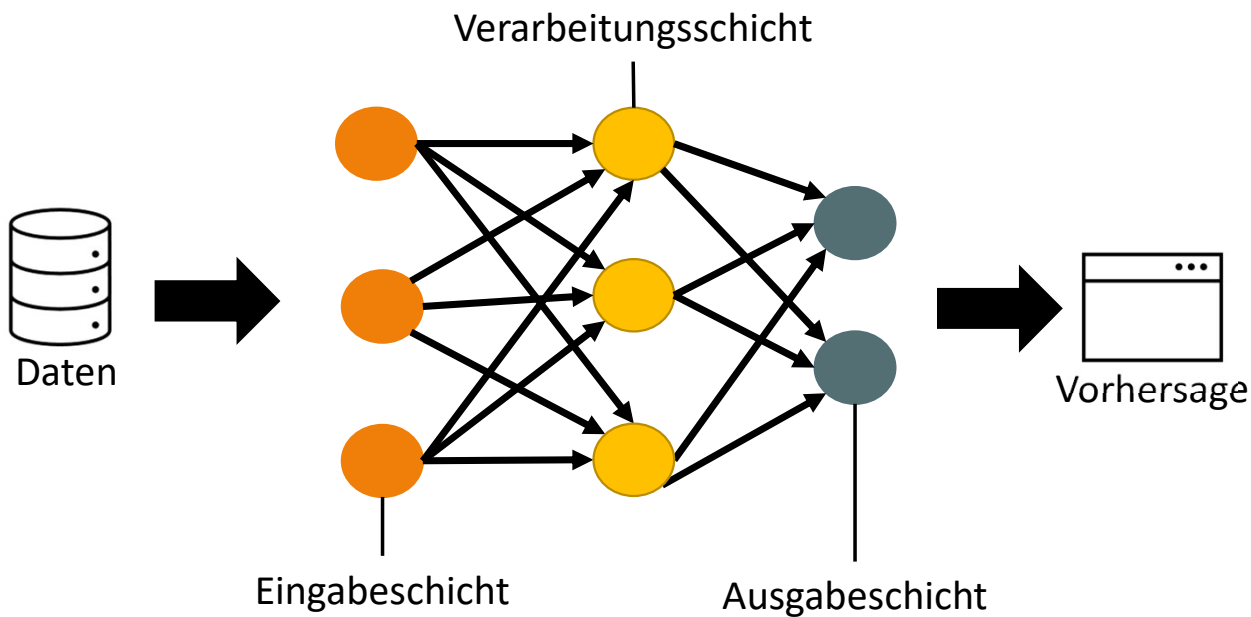
Fazit

Neuronales Netzwerk



Eingabe:

Ausgabe:



- Spam

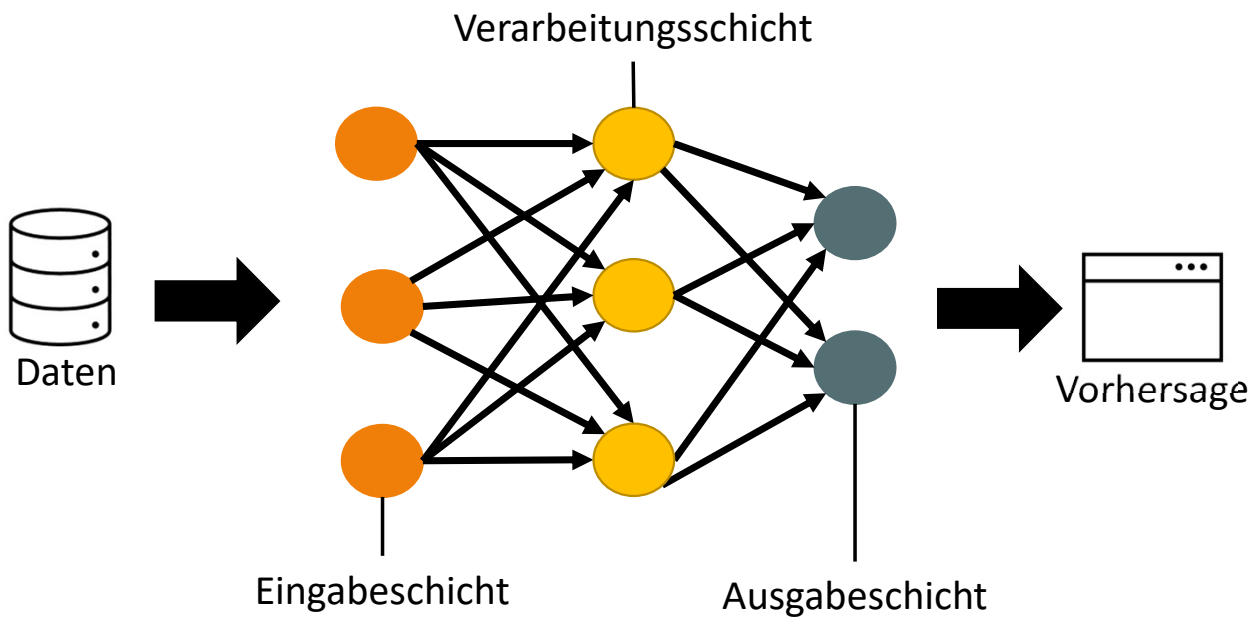


Neuronales Netzwerk



Eingabe:

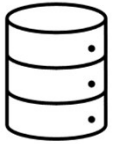
Ausgabe:



- Spam
- Themen



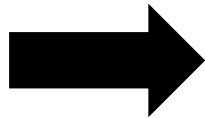
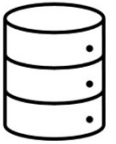
Trainingsdaten



Allgemeines Postfach



Trainingsdaten



Allgemeines Postfach

CSV

Einleitung



Analyse



Planung



Design



Durchführung

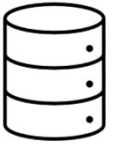


Test

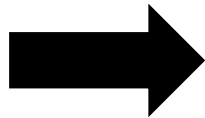


Fazit

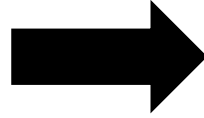
Trainingsdaten



Allgemeines Postfach



CSV



Datenbank

Einleitung



Analyse



Planung



Design



Durchführung

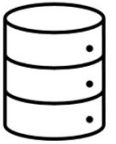


Test

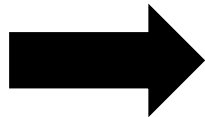


Fazit

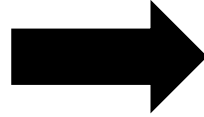
Trainingsdaten



Allgemeines Postfach



CSV



Datenbank



Google Colaboratory

Colab

Einleitung



Analyse



Planung



Design



Durchführung

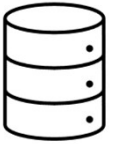


Test



Fazit

Trainingsdaten



Allgemeines Postfach



CSV



Datenbank



Google Colaboratory

Colab

→ Repräsentative Auswahl an E-Mails

Einleitung



Analyse



Planung



Design



Durchführung

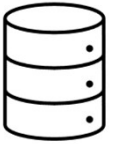


Test



Fazit

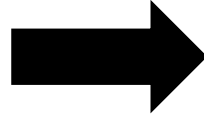
Trainingsdaten



Allgemeines Postfach



CSV



Datenbank



Google Colaboratory

Colab

→ Repräsentative Auswahl an E-Mails

→ 150 E-Mails ausgewählt

Einleitung



Analyse



Planung



Design



Durchführung

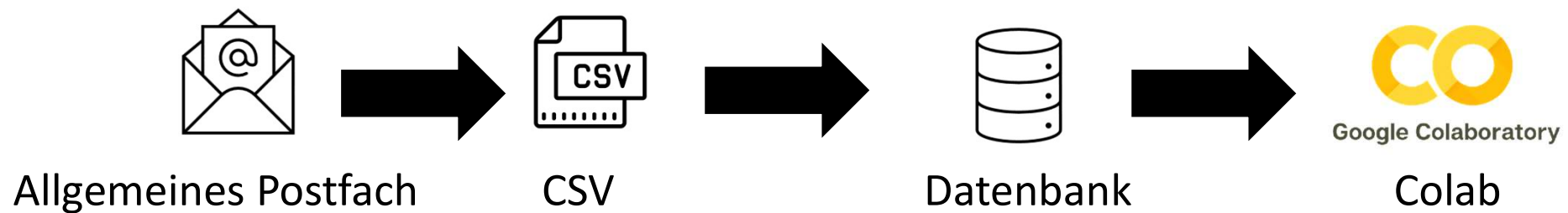
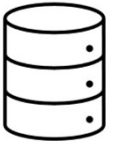


Test



Fazit

Trainingsdaten



→ Repräsentative Auswahl an E-Mails

→ 150 E-Mails ausgewählt

→ Randomisiert und vervielfältigt: 1500 Emails



Datenkennzeichnung (Data labeling)



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen
1	Braun	Apps, Kunde



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen
1	Braun	Apps, Kunde

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...
1	XY	XY	XYZ	Lorem Ipsum



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...
1	XY	XY	XYZ	Lorem Ipsum



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...
1	XY	XY	XYZ	Lorem Ipsum



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM
1	XY	XY	XYZ	Lorem Ipsum	



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM	TAGS
1	XY	XY	XYZ	Lorem Ipsum		



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM	TAGS
1	XY	XY	XYZ	Lorem Ipsum	0	



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM	TAGS
1	XY	XY	XYZ	Lorem Ipsum	0	AP, KU



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM	TAGS
1	XY	XY	XYZ	Lorem Ipsum	0	AP, KU



Datenkennzeichnung (Data labeling)



Mitarbeiter Tabelle

ID	Name	Themen	Tags
1	Braun	Apps, Kunde	AP, KU

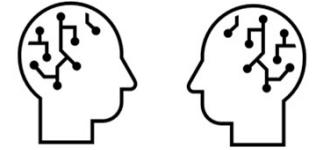
E-Mails Tabelle

ID	Absender	Empfänger	Betreff	Inhalt	...	SPAM	TAGS
1	XY	XY	XYZ	Lorem Ipsum	0	AP, KU

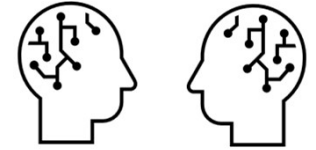
→ Skalierbarkeit



KI-Modelle: Spam & Thema



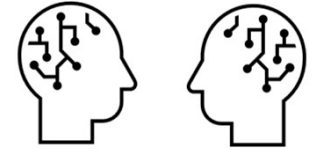
KI-Modelle: Spam & Thema



Spam:



KI-Modelle: Spam & Thema



Spam:

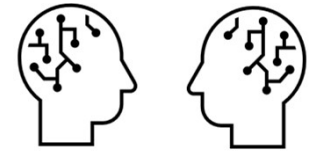
Eingabeschicht (Input-Layer)

Verarbeitungsschicht (Hidden-Layer)

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Spam:

Eingabeschicht (Input-Layer)

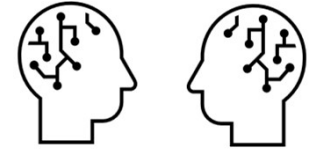
```
model.add(tf.keras.layers.Dense(256, input_shape=x_train.shape, activation='sigmoid'))
```

Verarbeitungsschicht (Hidden-Layer)

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Spam:

Eingabeschicht (Input-Layer)

```
model.add(tf.keras.layers.Dense(256, input_shape=x_train.shape, activation='sigmoid'))
```

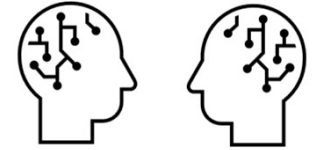
Verarbeitungsschicht (Hidden-Layer)

```
model.add(tf.keras.layers.Dense(256, activation='sigmoid'))
```

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Spam:

Eingabeschicht (Input-Layer)

```
model.add(tf.keras.layers.Dense(256, input_shape=x_train.shape, activation='sigmoid'))
```

Verarbeitungsschicht (Hidden-Layer)

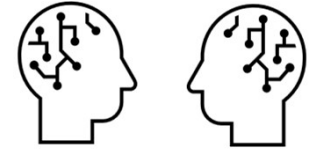
```
model.add(tf.keras.layers.Dense(256, activation='sigmoid'))
```

Ausgabeschicht (Output-Layer)

```
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
```



KI-Modelle: Spam & Thema



Spam:

Eingabeschicht (Input-Layer)

```
model.add(tf.keras.layers.Dense(256, input_shape=x_train.shape, activation='sigmoid'))
```

Verarbeitungsschicht (Hidden-Layer)

```
model.add(tf.keras.layers.Dense(256, activation='sigmoid'))
```

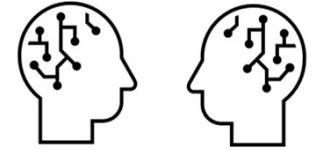
Ausgabeschicht (Output-Layer)

```
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
```

OUT: Boolean



KI-Modelle: Spam & Thema



Thema:

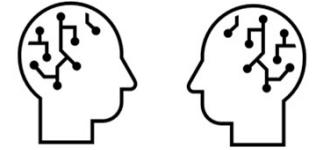
Eingabeschicht (Input-Layer)

Verarbeitungsschicht (Hidden-Layer)

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Thema:

Eingabeschicht (Input-Layer)

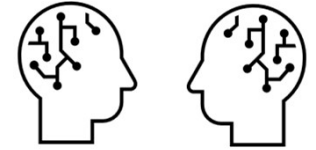
```
model.add(512, input_shape=x_train.shape, activation='sigmoid')
```

Verarbeitungsschicht (Hidden-Layer)

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Thema:

Eingabeschicht (Input-Layer)

```
model.add(512, input_shape=x_train.shape, activation='sigmoid')
```

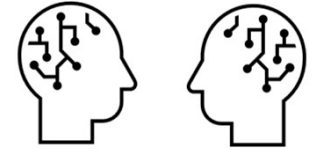
Verarbeitungsschicht (Hidden-Layer)

```
model.add(256, activation='sigmoid')
```

Ausgabeschicht (Output-Layer)



KI-Modelle: Spam & Thema



Thema:

Eingabeschicht (Input-Layer)

```
model.add(512, input_shape=x_train.shape, activation='sigmoid')
```

Verarbeitungsschicht (Hidden-Layer)

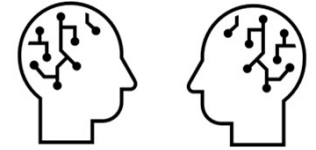
```
model.add(256, activation='sigmoid')
```

Ausgabeschicht (Output-Layer)

```
model.add(1, output_mode="multi_hot", activation="sigmoid")
```



KI-Modelle: Spam & Thema



Thema:

Eingabeschicht (Input-Layer)

```
model.add(512, input_shape=x_train.shape, activation='sigmoid')
```

Verarbeitungsschicht (Hidden-Layer)

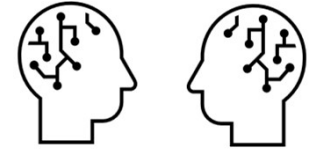
```
model.add(256, activation='sigmoid')
```

Ausgabeschicht (Output-Layer)

```
model.add(1, output_mode="multi_hot", activation="sigmoid")
```



KI-Modelle: Spam & Thema



Thema:

Eingabeschicht (Input-Layer)

```
model.add(512, input_shape=x_train.shape, activation='sigmoid')
```

Verarbeitungsschicht (Hidden-Layer)

```
model.add(256, activation='sigmoid')
```

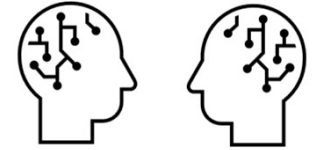
Ausgabeschicht (Output-Layer)

```
model.add(1, output_mode="multi_hot", activation="sigmoid")
```

OUT: Bag of Words



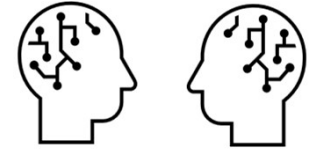
KI-Modelle: Spam & Thema



Vorhersage	<code>array([[0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]])</code>
Tags:	<code>array(['Ampelsystem', 'Apps', 'Banner', 'Bestellung', 'CitrixFehler', 'Dokumenation', 'Fake', 'Fraud', 'Gehe', 'Kalender', 'Mailserver', 'Marketing', 'Meeting', 'Mockup', 'RecuCare', 'Reporting', 'SQL', 'Salzgitter', 'Sammelbuchung', 'Semi', 'Seminare', 'Seminarkonfigurator', 'Spam', 'Sylvenstein', 'TestKunde', 'WebsiteFehler', 'WidgetVeranstaltungen', 'Zertifikat', 'elearning', 'pdfUpload'], dtype=object)</code>



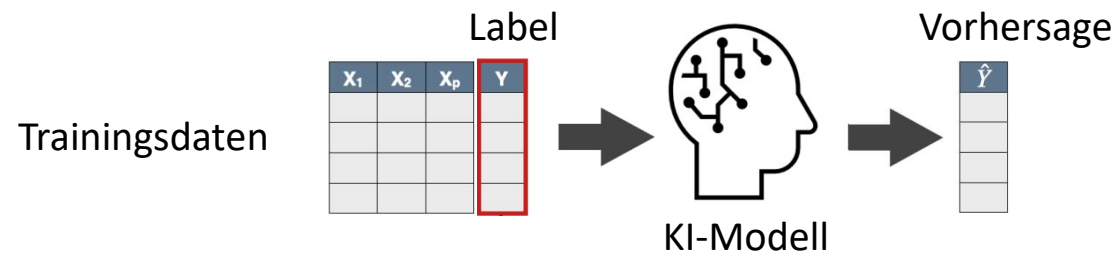
KI-Modelle: Spam & Thema



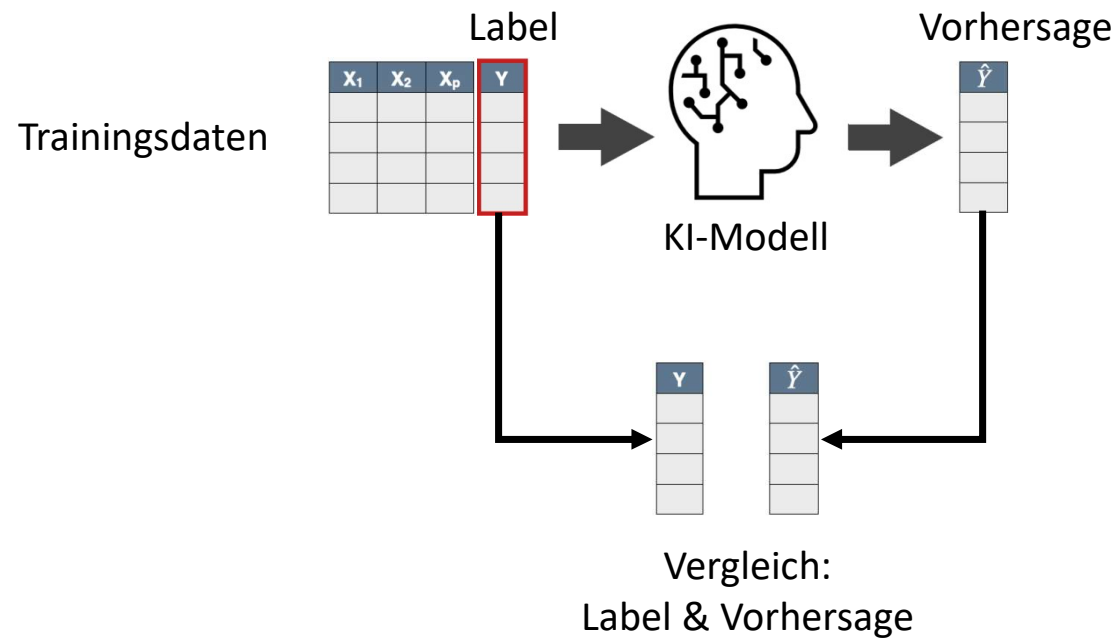
Vorhersage	<code>array([[0, 0, 0, 0, 0, 0, 0, <u>1</u>, 0, 0, 0, 0, 0, 0, <u>1</u>, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]])</code>
Tags:	<code>array(['Ampelsystem', 'Apps', 'Banner', 'Bestellung', 'CitrixFehler', 'Dokumenation', '<u>Fake</u>', 'Fraud', 'Gehe', 'Kalender', 'Mailserver', 'Marketing', 'Meeting', 'Mockup', 'RecuCare', 'Reporting', '<u>SQL</u>', 'Salzgitter', 'Sammelbuchung', 'Semi', 'Seminare', 'Seminarkonfigurator', 'Spam', 'Sylvenstein', 'TestKunde', 'WebsiteFehler', 'WidgetVeranstaltungen', 'Zertifikat', 'elearning', 'pdfUpload'], dtype=object)</code>



Iteratives Training



Iteratives Training



Einleitung



Analyse



Planung



Design



Durchführung

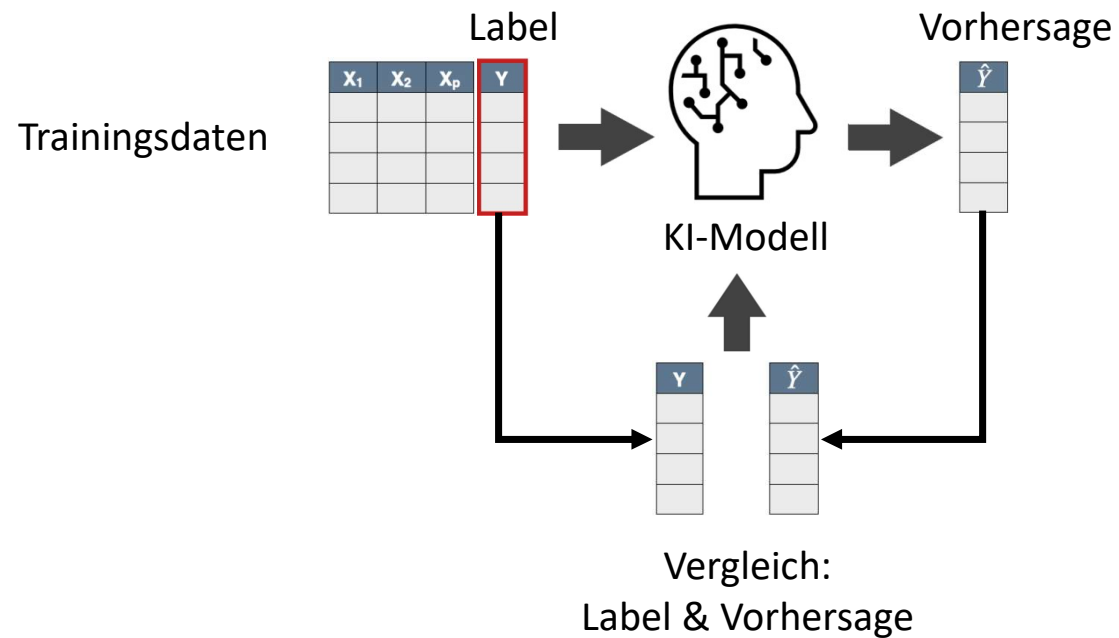


Test



Fazit

Iteratives Training



Einleitung

Analyse

Planung

Design

Durchführung

Test

Fazit

Lernkurve: Spam



Lernkurve: Spam



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312



Lernkurve: Spam



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625



Lernkurve: Spam



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625

...

...

...

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Lernkurve: Spam



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625

...

...

...

Epoch 210/500

1/1 [=====] - 0s 5ms/step - loss: 0.4533 - accuracy: 0.9688

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Lernkurve: Spam



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625

...

...

...

Epoch 210/500

1/1 [=====] - 0s 5ms/step - loss: 0.4533 - accuracy: 0.9688

Epoch 211/500

1/1 [=====] - 0s 18ms/step - loss: 0.4515 - accuracy: 0.9789

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Lernkurve: Spam



Anpassung der Parameter innerhalb des Netzes



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625

...

...

...

Epoch 210/500

1/1 [=====] - 0s 5ms/step - loss: 0.4533 - accuracy: 0.9688

Epoch 211/500

1/1 [=====] - 0s 18ms/step - loss: 0.4515 - accuracy: 0.9789

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Lernkurve: Spam



Anpassung der Parameter innerhalb des Netzes



Epoch 64/500

1/1 [=====] - 0s 4ms/step - loss: 0.6472 - accuracy: 0.5312

Epoch 65/500

1/1 [=====] - 0s 5ms/step - loss: 0.6459 - accuracy: 0.5625

...

...

...

Epoch 210/500

1/1 [=====] - 0s 5ms/step - loss: 0.4533 - accuracy: 0.9688

Epoch 211/500

1/1 [=====] - 0s 18ms/step - loss: 0.4515 - accuracy: 0.9789



Genauigkeit der Vorhersagen



Lernkurve: Thema



Einleitung



Analyse



Planung



Design



Durchführung

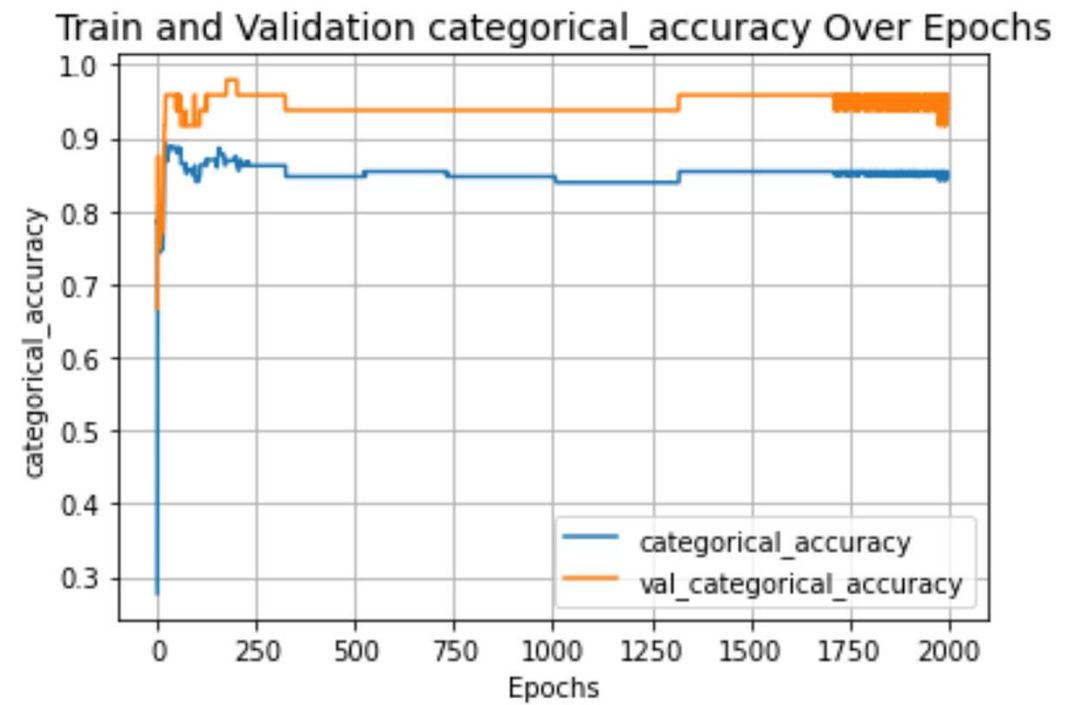
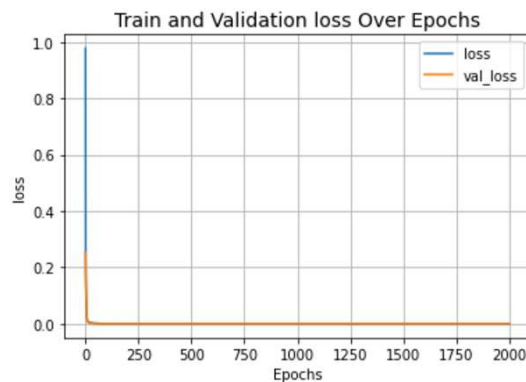


Test



Fazit

Lernkurve: Thema



Einleitung



Analyse



Planung



Design



Durchführung

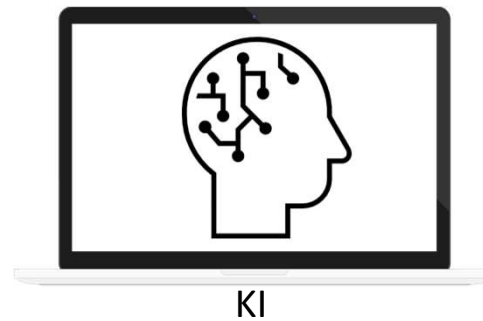


Test



Fazit

KI Testen



KI Testen



Trainingsdaten



KI

Einleitung



Analyse



Planung



Design



Durchführung

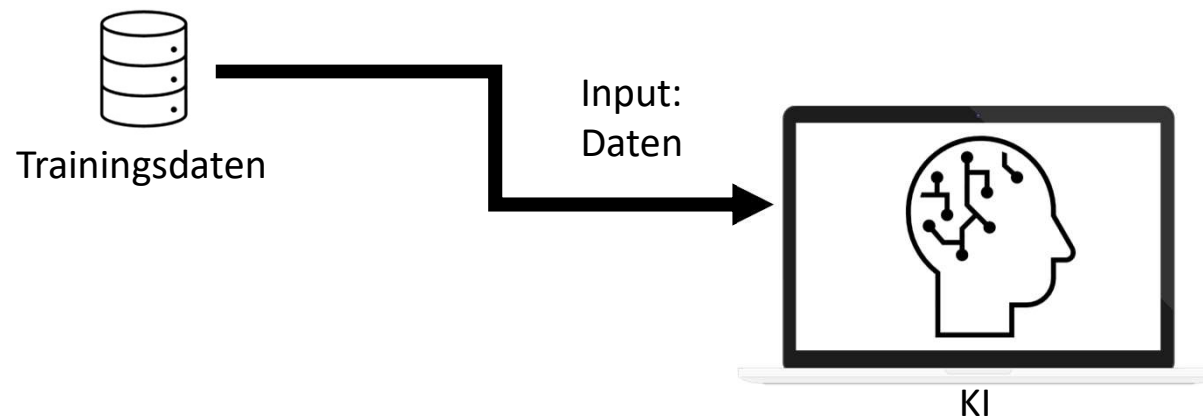


Test

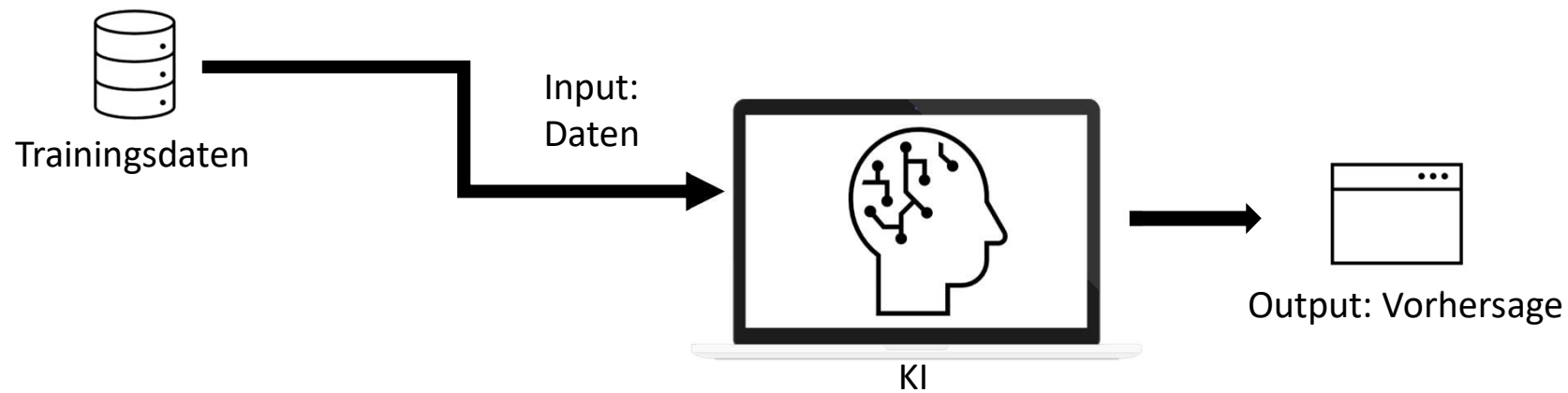


Fazit

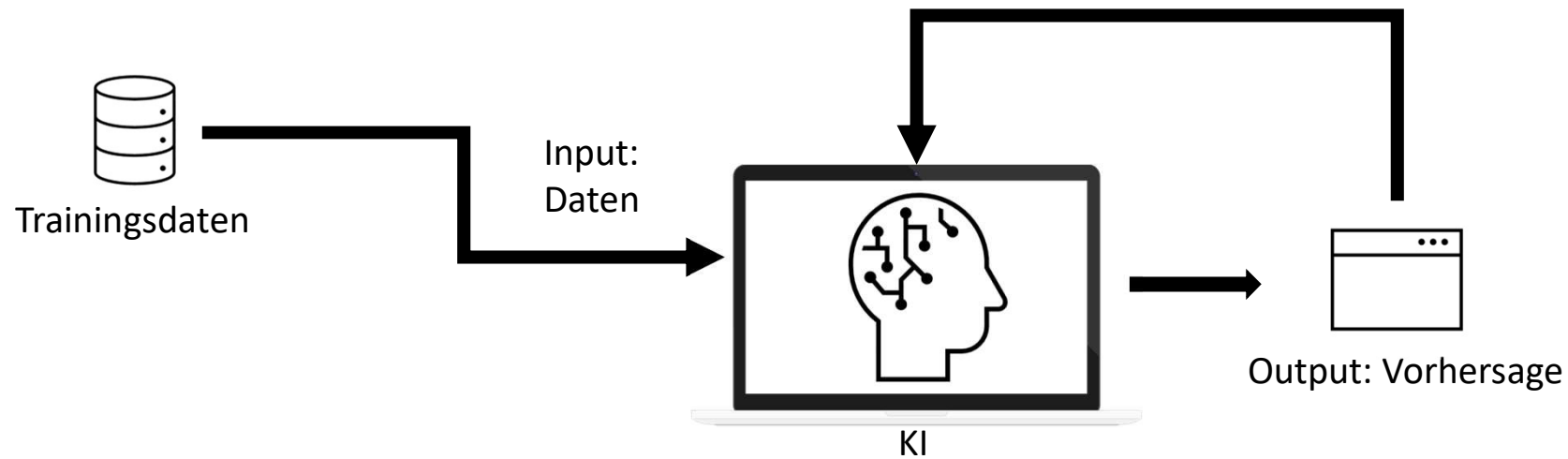
KI Testen



KI Testen



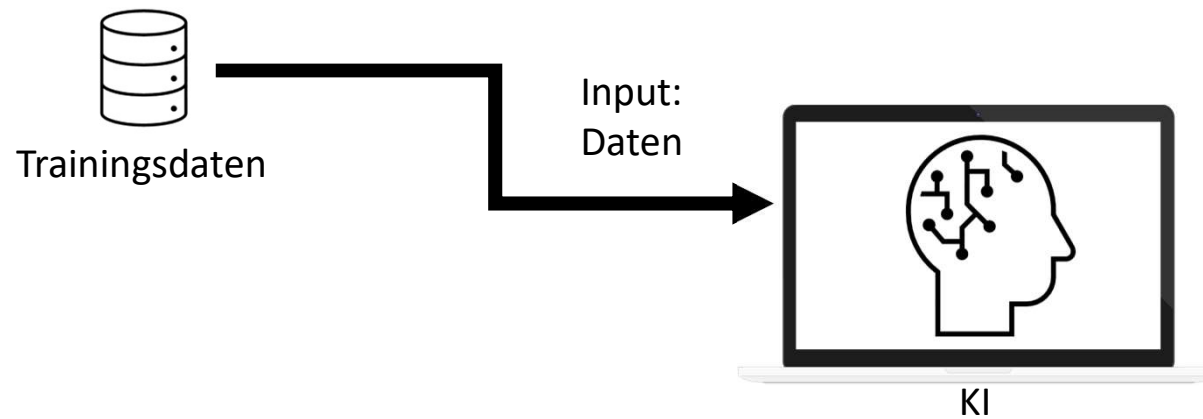
KI Testen



KI Testen



Whitebox Test



Einleitung



Analyse



Planung



Design



Durchführung



Test

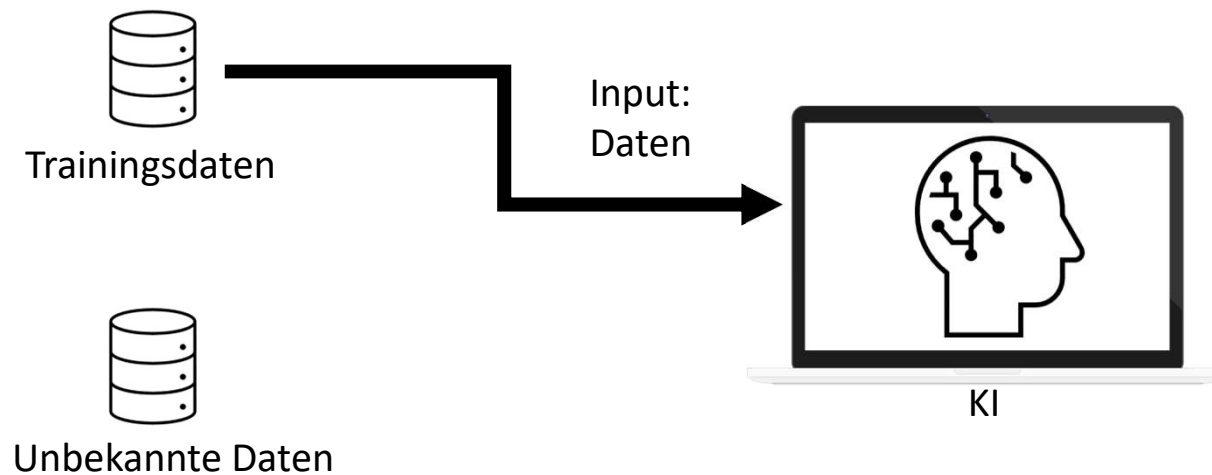


Fazit

KI Testen



Whitebox Test



Einleitung



Analyse



Planung



Design



Durchführung



Test

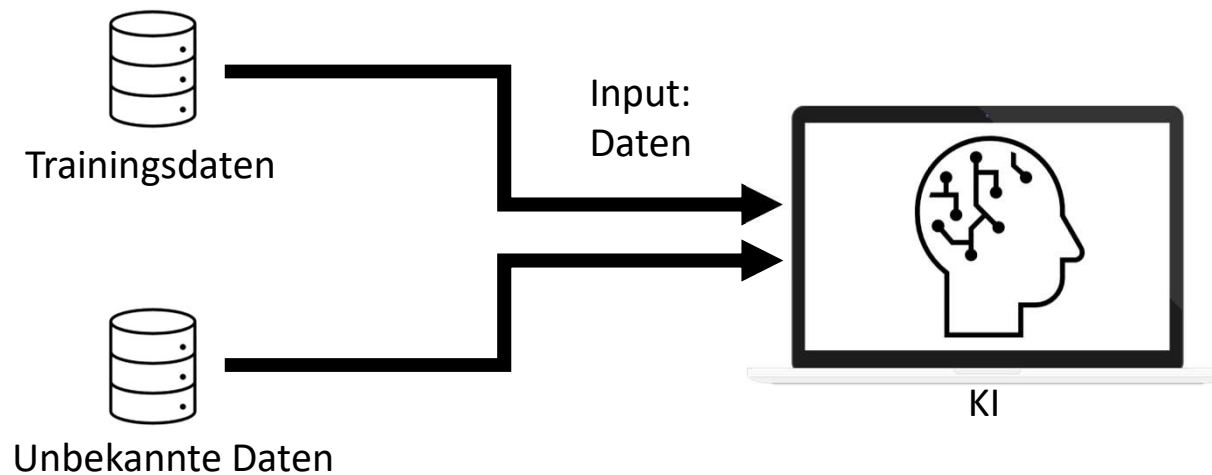


Fazit

KI Testen



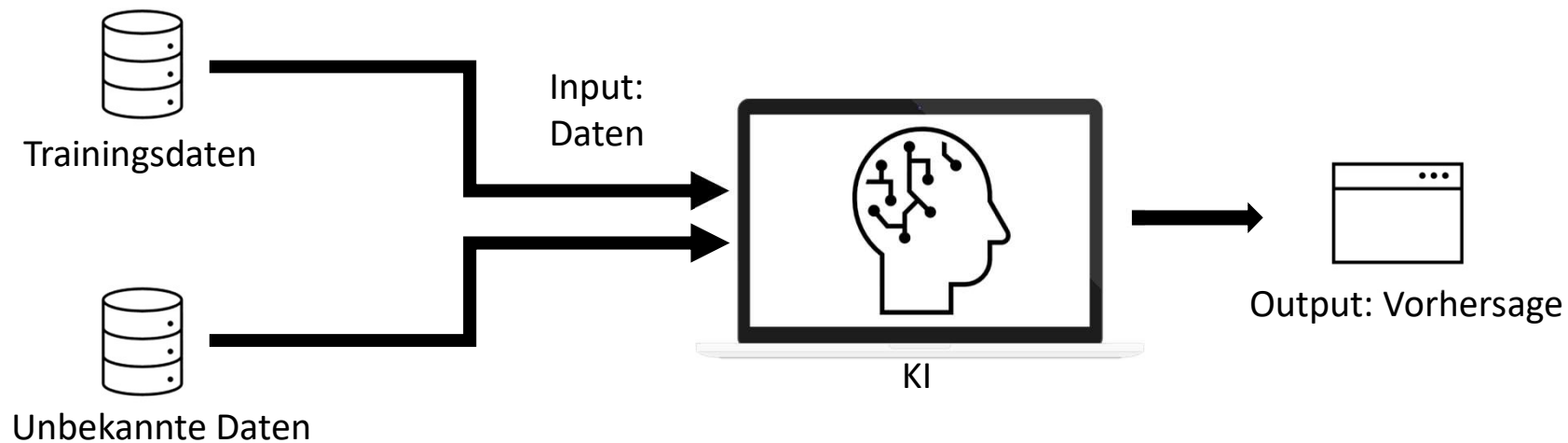
Whitebox Test



KI Testen



Whitebox Test



Einleitung



Analyse



Planung



Design



Durchführung

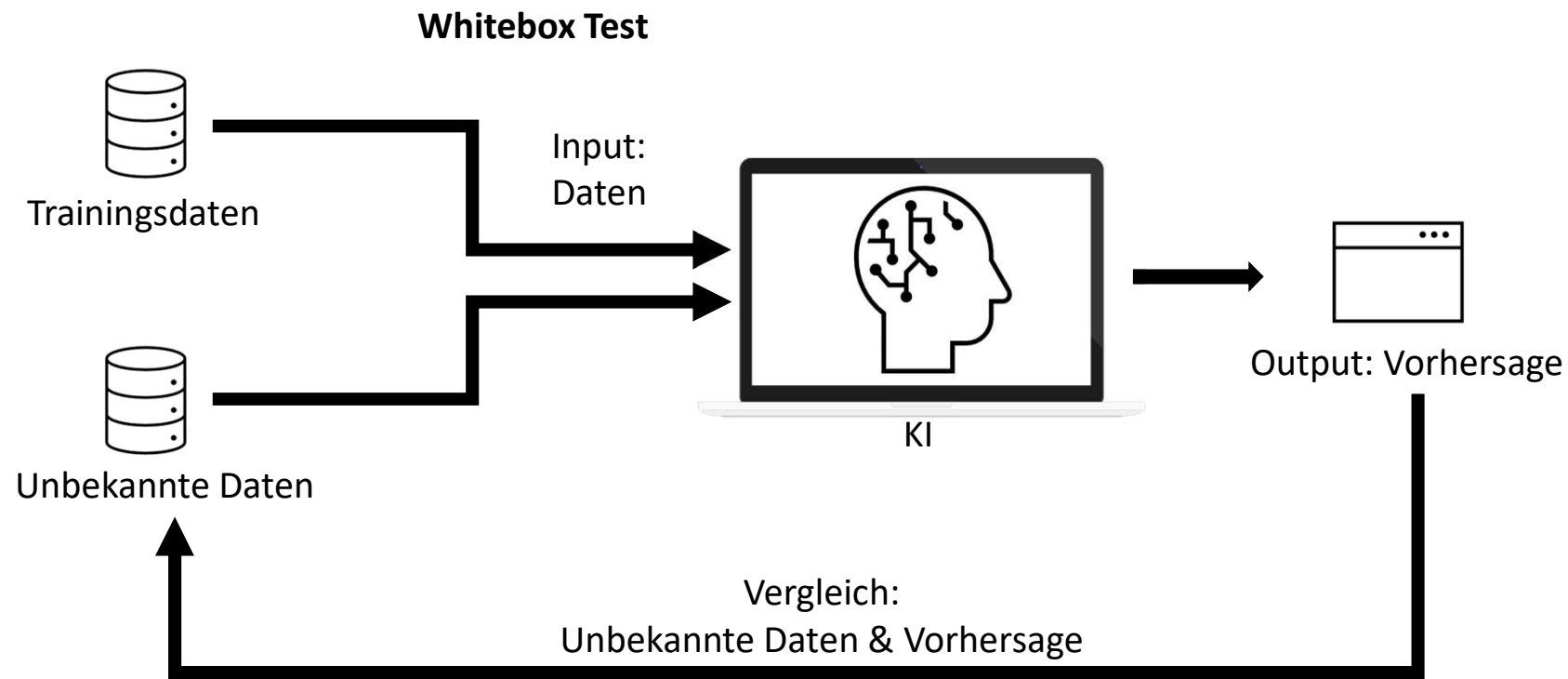


Test



Fazit

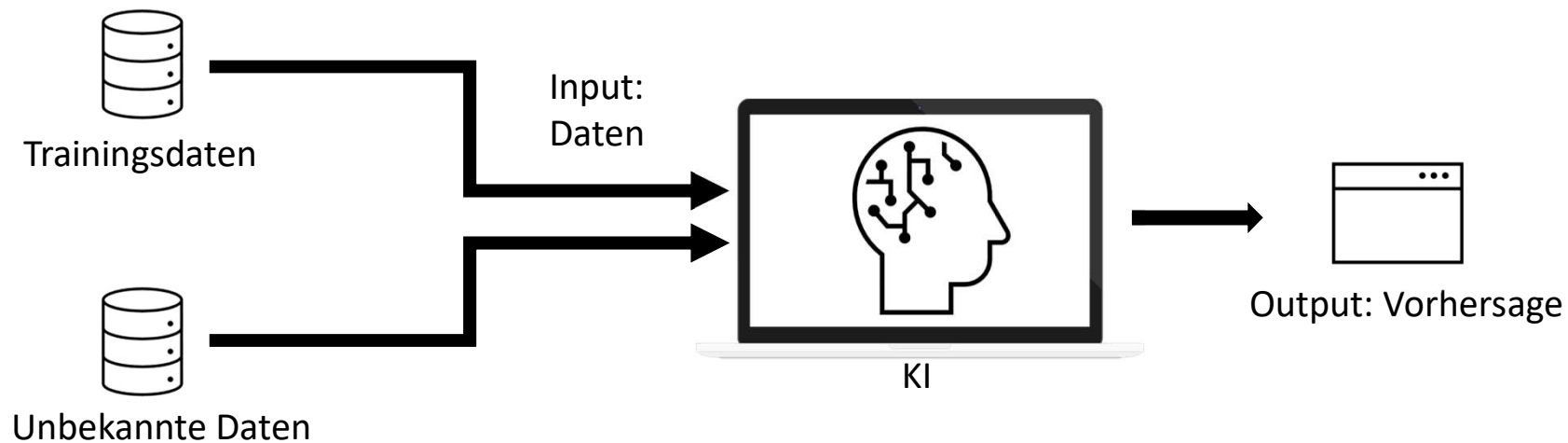
KI Testen



KI Testen



Whitebox Test



Blackbox Test



Blackbox Problematik



Model: „Sequential_Thema“

Layer (type)	Output Shape	Param #
Input (Dense)	(None, 512)	2203136
Hidden (Dense)	(None, 256)	131328
Output (Dense)	(None, 31)	7967
Total params: 2,342,431		
Trainable params: 2,342,431		
Non-trainable params: 0		

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Blackbox Problematik



Model: „Sequential_Thema“

Layer (type)	Output Shape	Param #
Input (Dense)	(None, 512)	<u>2203136</u>
Hidden (Dense)	(None, 256)	131328
Output (Dense)	(None, 31)	7967
Total params: 2,342,431		
Trainable params: 2,342,431		
Non-trainable params: 0		

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Blackbox Problematik



Model: „Sequential_Thema“

Layer (type)	Output Shape	Param #
Input (Dense)	(None, 512)	<u>2203136</u>
Hidden (Dense)	(None, 256)	<u>131328</u>
Output (Dense)	(None, 31)	7967
Total params: 2,342,431		
Trainable params: 2,342,431		
Non-trainable params: 0		

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Blackbox Problematik



Model: „Sequential_Thema“

Layer (type)	Output Shape	Param #
Input (Dense)	(None, 512)	<u>2203136</u>
Hidden (Dense)	(None, 256)	<u>131328</u>
Output (Dense)	(None, 31)	<u>7967</u>
Total params: 2,342,431		
Trainable params: 2,342,431		
Non-trainable params: 0		

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Blackbox Problematik



Model: „Sequential_Thema“

Layer (type)	Output Shape	Param #
Input (Dense)	(None, 512)	<u>2203136</u>
Hidden (Dense)	(None, 256)	<u>131328</u>
Output (Dense)	(None, 31)	<u>7967</u>
Total params: 2,342,431		
Trainable params: 2,342,431		
Non-trainable params: 0		

Einleitung



Analyse



Planung



Design



Durchführung



Test



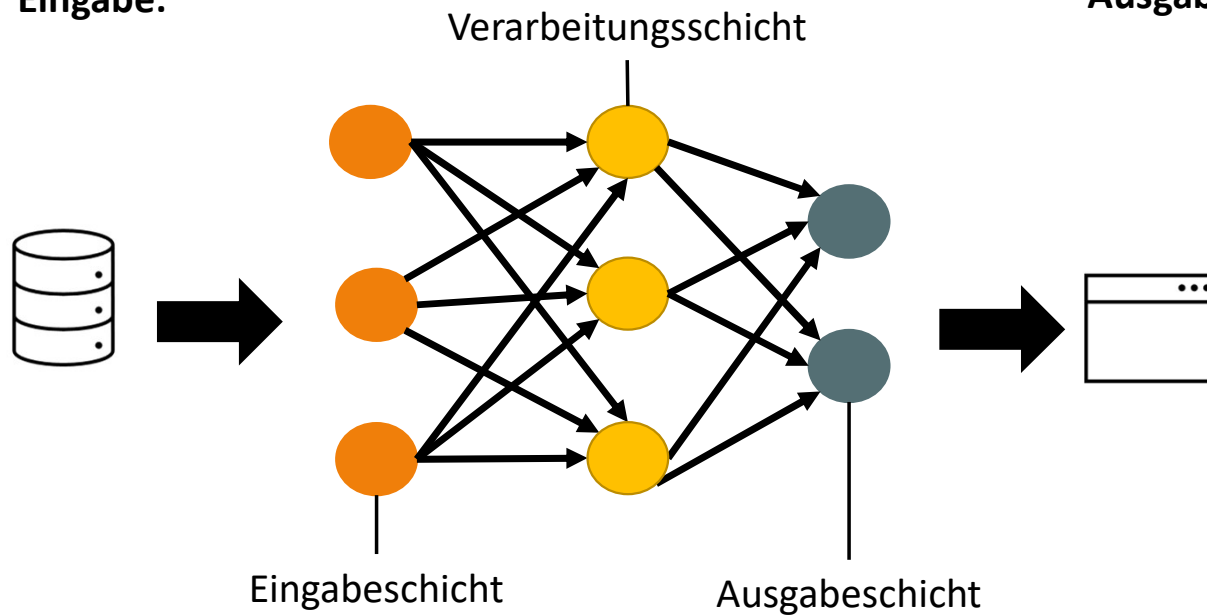
Fazit

Blackbox Problematik



Eingabe:

Ausgabe:



Einleitung

Analyse

Planung

Design

Durchführung

Test

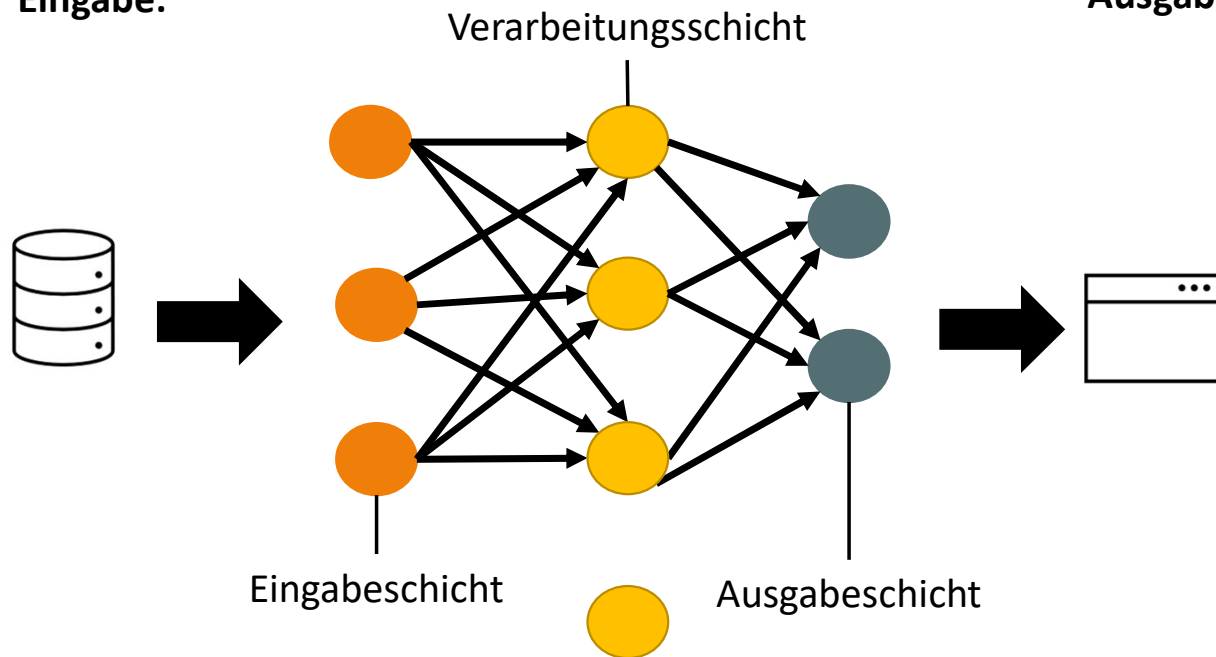
Fazit

Blackbox Problematik



Eingabe:

Ausgabe:



Einleitung

Analyse

Planung

Design

Durchführung

Test

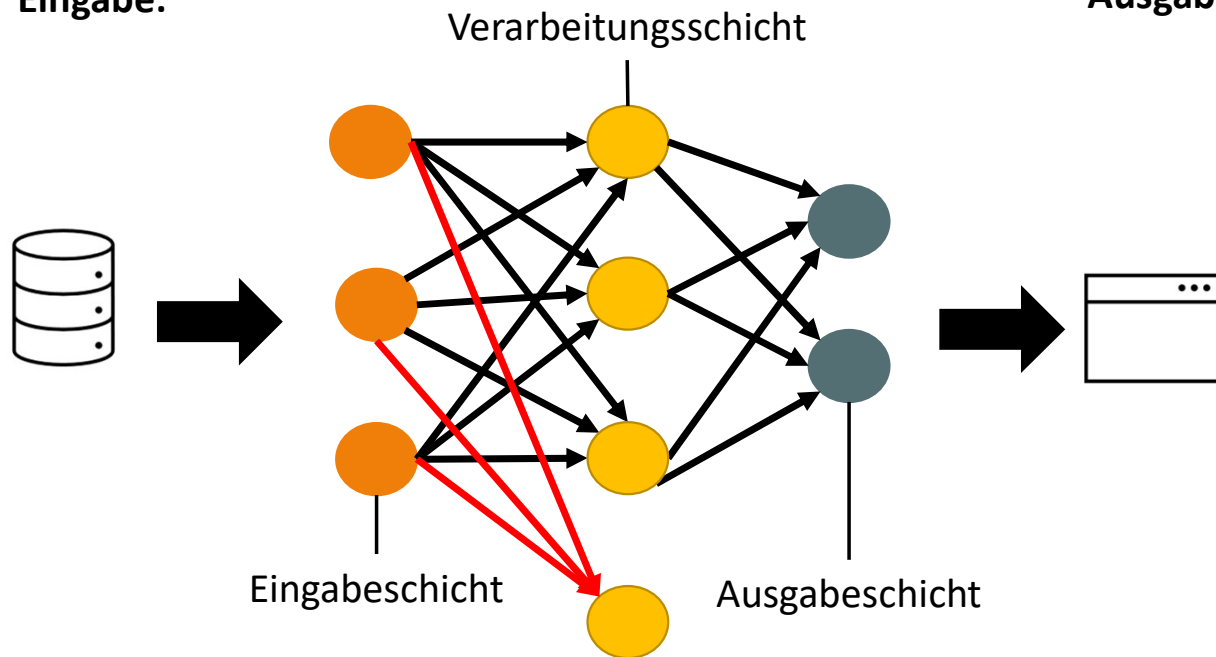
Fazit

Blackbox Problematik



Eingabe:

Ausgabe:



Einleitung



Analyse



Planung



Design



Durchführung



Test



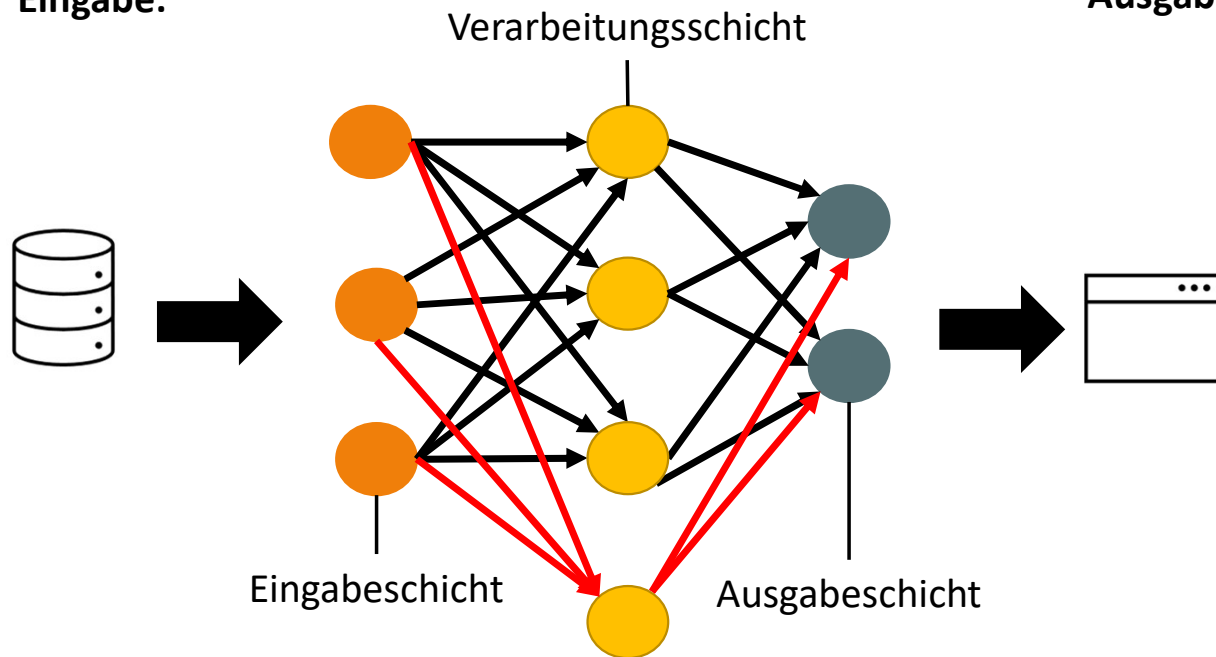
Fazit

Blackbox Problematik



Eingabe:

Ausgabe:



Einleitung



Analyse



Planung



Design



Durchführung

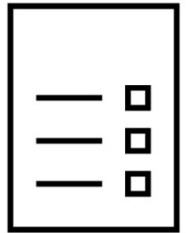


Test

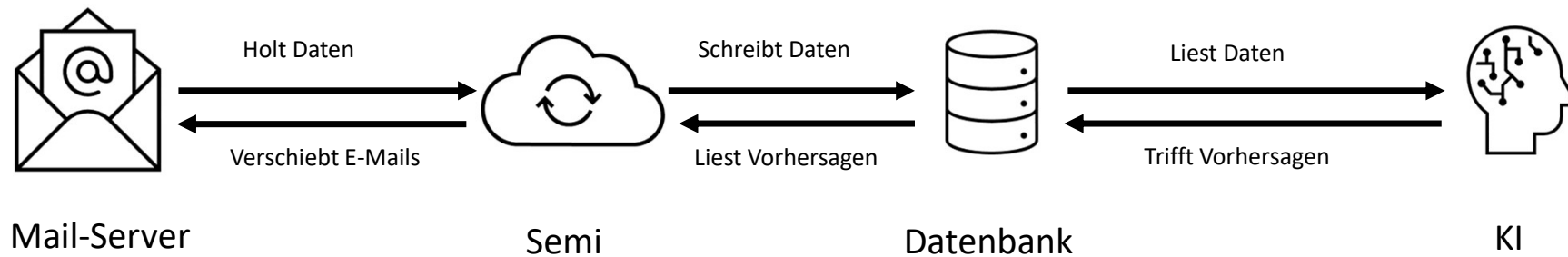
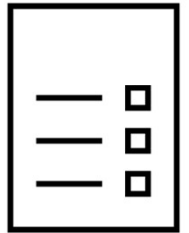


Fazit

Test- & Produktivsystem

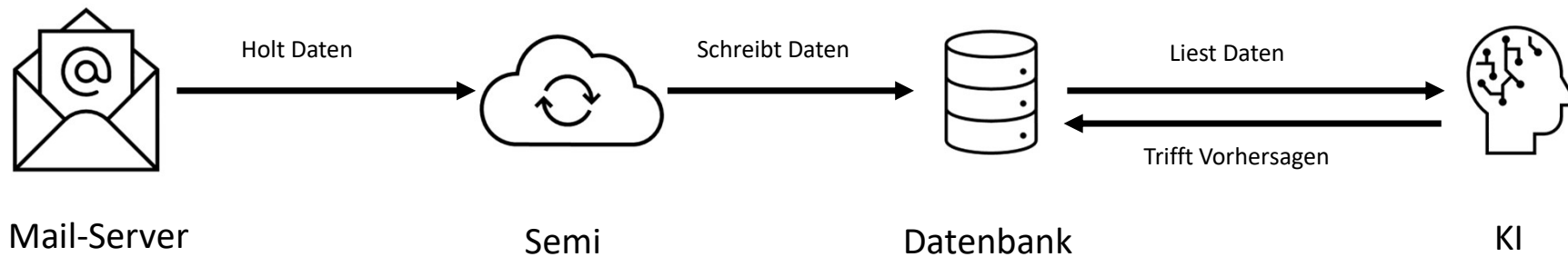
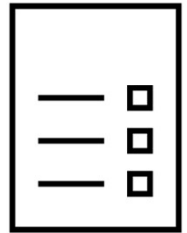


Test- & Produktivsystem



Test- & Produktivsystem

Test-System:



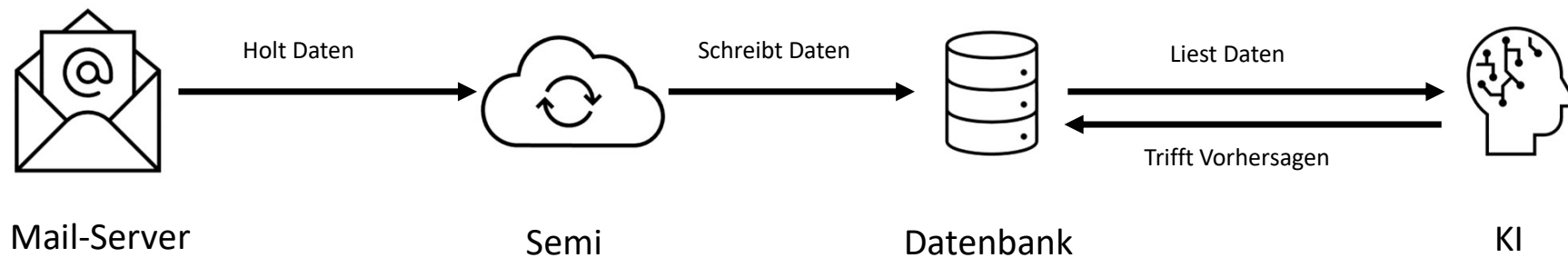
→ Semi verschiebt keine E-Mails



Test- & Produktivsystem



Test-System:



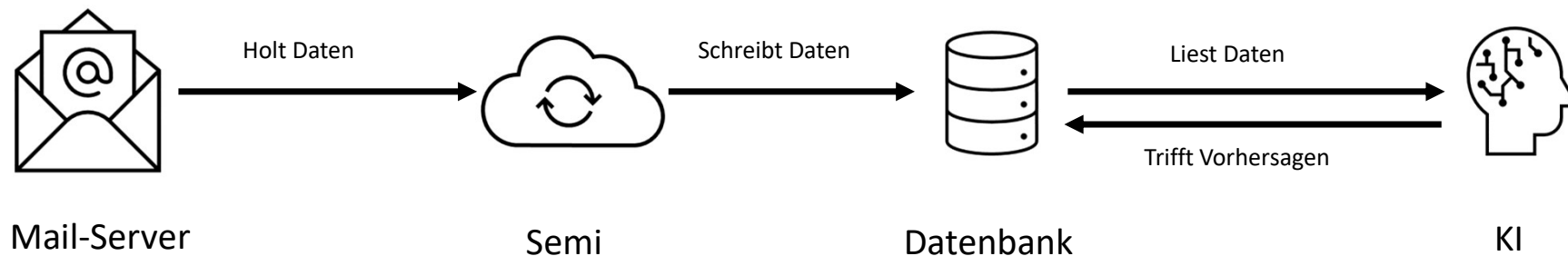
- Semi verschiebt keine E-Mails
- Training (Whitebox Test)



Test- & Produktivsystem



Test-System:



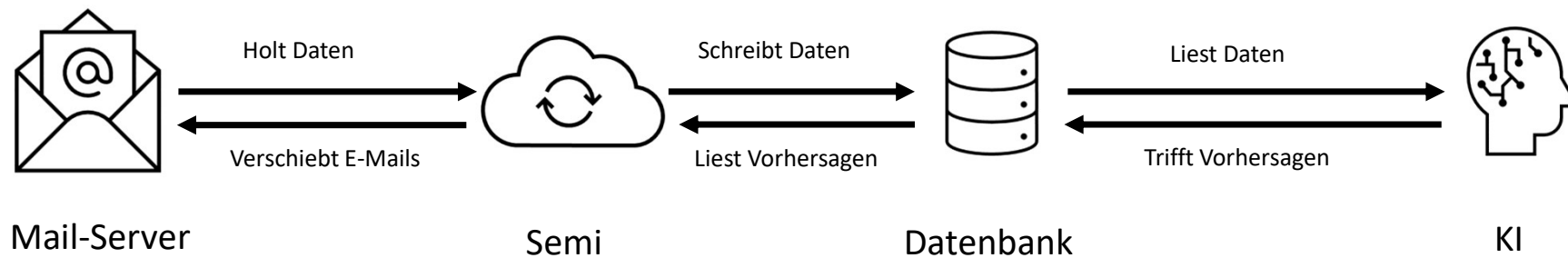
- Semi verschiebt keine E-Mails
- Training (Whitebox Test)
- Unbekannte Daten (Blackbox Test)



Test- & Produktivsystem



Produktiv-System:



→ Semi verschiebt E-Mails



Projektergebnis:



Projektergebnis:



→ Spam: ca. 99%



Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Projektergebnis:



→ Spam: ca. 99% ✓

→ Themen: ca. 85% ✓

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Projektergebnis:



→ Spam: ca. 99% ✓

→ Themen: ca. 85% ✓

- Vorhersagen entsprechen den Vorgaben



Projektergebnis:



→ Spam: ca. 99% ✓

→ Themen: ca. 85% ✓

- Vorhersagen entsprechen den Vorgaben
- Wissenserwerb hat stattgefunden

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Projektergebnis:



→ Spam: ca. 99% ✓

→ Themen: ca. 85% ✓

- Vorhersagen entsprechen den Vorgaben
- Wissenserwerb hat stattgefunden
- Blaupause für zukünftige KI-Projekte

Einleitung



Analyse



Planung



Design



Durchführung



Test



Fazit

Ausblick & Lernerfolg



Ausblick & Lernerfolg



Umsetzung der meisten KI Projekte scheitern:
→ Implementierung des KI-Modells



Ausblick & Lernerfolg



Umsetzung der meisten KI Projekte scheitern:

→ Implementierung des KI-Modells

```
1 server.run_app(app=app)

... Public URL: NgrokTunnel: "http://4c97946419c2.ngrok.io" -> "http://localhost:10000"
INFO: Started server process [129]
INFO: Waiting for application startup.
INFO: Application startup complete.
INFO: Uvicorn running on http://127.0.0.1:10000 (Press CTRL+C to quit)
```

→ KI-Modell als
Webmethode

→ FastAPI + Colab

Quelle: <https://towardsdatascience.com/colabcode-deploying-machine-learning-models-from-google-colab-54e0d37a7b09>



FIN

B. Still