

Familienname, Vorname (bitte durch eine Leerspalte trennen)

[illegible]

IHK

Bereich	Berufsnummer	IHK-Nummer	Prüfungsnummer
7 3	1 2 0 3		
Sp. 1-2	Sp. 3-6	Sp. 7-9	Sp. 10-14

Abschlussprüfung Sommer 2024

1203

2 Sicherstellen der Datenqualität

Fachinformatiker
Fachinformatikerin
Daten- und Prozessanalyse

100 Punkte

1. Bevor Sie mit der Bearbeitung der Aufgaben beginnen, überprüfen Sie bitte die **Vollständigkeit** dieses Aufgabensatzes. Die Anzahl der zu bearbeitenden Aufgaben ist auf dem Deckblatt links angegeben. Wenden Sie sich bei Unstimmigkeiten sofort an die Aufsicht, weil Reklamationen am Ende der Prüfung nicht anerkannt werden können.
2. Füllen Sie zuerst die **Kopfzeile** aus. Tragen Sie Ihren Familiennamen, Ihren Vornamen und Ihre Prüfungs-Nr. in die oben stehenden Felder ein.
3. Lesen Sie bitte den **Text** der Aufgaben ganz durch, bevor Sie mit der Bearbeitung beginnen.
4. Halten Sie sich bei der Bearbeitung der Aufgaben genau an die **Vorgaben der Aufgabenstellung** zum Umfang der Lösung. Wenn z. B. vier Angaben gefordert werden und Sie sechs Angaben anführen, werden nur die ersten vier Angaben bewertet.
5. Tragen Sie die frei zu formulierenden **Antworten dieser offenen Aufgaben** in die dafür lt. Aufgabenstellung vorgesehenen Bereiche (Lösungszeilen, Formulare, Tabellen u. a.) des Arbeitsbogens ein.
6. Sofern nicht ausdrücklich ein Brief oder eine Formulierung in ganzen Sätzen gefordert werden, ist eine **stichwortartige Beantwortung** zulässig.
7. Schreiben Sie deutlich und gut lesbar. Ein nicht eindeutig zuzuordnendes oder **unleserliches Ergebnis** wird als **falsch** gewertet.
8. Zur Lösung der Rechenaufgaben darf ein nicht programmierter, netzunabhängiger **Taschenrechner** ohne Kommunikationsmöglichkeit mit Dritten verwendet werden.
9. Wenn Sie ein **gerundetes Ergebnis** eintragen und damit weiterrechnen müssen, rechnen Sie (auch im Taschenrechner) nur mit diesem gerundeten Ergebnis weiter.
10. Für **Hilfsaufzeichnungen** können Sie das in der Tasche beigelegte Konzeptpapier verwenden. Bewertet werden jedoch grundsätzlich nur Ihre Eintragungen in diesem Aufgabensatz.

Bewertung

Für die Bewertung gilt die Vorgabe der Punkte in den Lösungshinweisen.

1. Aufg.

--	--

 Punkte 2. Aufg.

--	--

 Punkte 3. Aufg.

--	--

 Punkte 4. Aufg.

--	--

 Punkte

15 16 17 18 19 20 21 22

Prüfungszeit

23

Die entsprechende Ziffer (1, 2 oder 3) finden Sie in der Abfrage nach der Prüfungszeit im Anschluss an die letzte Aufgabe.

Gesamtpunktzahl

24	25	26

Prüfungsort, Datum

Unterschrift

Gemeinsame Prüfungsaufgaben der Industrie- und Handelskammern. Dieser Aufgabensatz wurde von einem überregionalen Ausschuss, der entsprechend § 40 Berufsbildungsgesetz zusammengesetzt ist, beschlossen. Hinweis: Im Interesse einer besseren Lesbarkeit wird in der Aufgabenstellung und in den Angaben zur Aufgabenstellung nur die männliche Form (generisches Maskulinum) verwendet. Die verkürzte Sprachform beinhaltet keine Wertung und die gewählten männlichen Formulierungen gelten uneingeschränkt auch für die weiteren Geschlechter. Die Vervielfältigung, Verbreitung und öffentliche Wiedergabe der Prüfungsaufgaben und Lösungen ist nicht gestattet. Zuwiderhandlungen werden zivil- und strafrechtlich (§§ 97 ff., 106 ff. UrhG) verfolgt. – © ZPA Nord-West 2024 – Alle Rechte vorbehalten!

Die Aufgaben 1 bis 4 beziehen sich auf die folgende Ausgangssituation:

Ein Geschäftsfeld der AMAG Soft GmbH liegt im Bereich der mobilen Kassensysteme.

Die hierfür benötigten mobilen Datenerfassungsgeräte (MDE-Geräte), wie z. B. Kartenlesegeräte, Scanner, Handhelds, die notwendige Infrastruktur sowie Verbrauchsgüter werden von der AMAG Soft GmbH vertrieben.

1. Aufgabe (25 Punkte)

Der technische Zustand des Datenspeichers der AMAG Soft GmbH wird mittels zweier Messwerte (x_1 , x_2) überwacht. Die Messwerte sollen von einer künstlichen Intelligenz (KI) beurteilt werden, um den Gesundheitszustand des Speichers (good, bad) zu ermitteln.

a) Beschreiben Sie die nachfolgenden Begriffe und wie sie zueinander in Beziehung stehen.

6 Punkte

Künstliche Intelligenz	
Maschinelles Lernen	
Deep Learning	

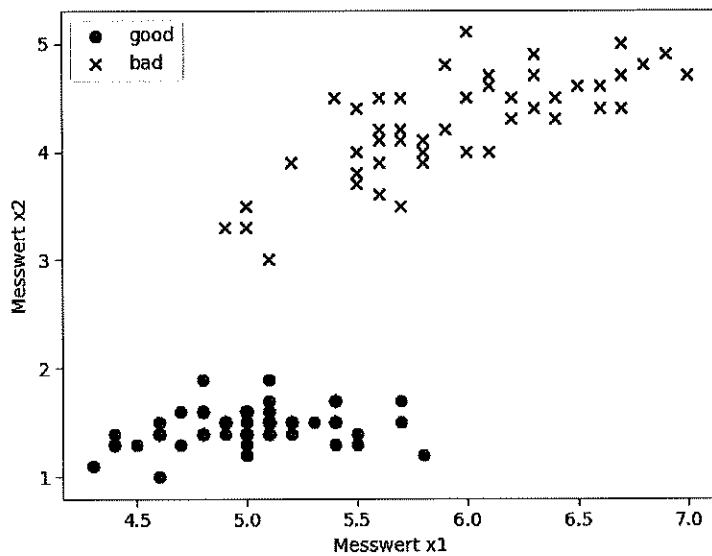
b) Beschreiben Sie die nachfolgenden Arten maschinellen Lernens.

4 Punkte

Supervised Learning	
Unsupervised Learning	

c) Folgende Trainingsdaten stehen zur Verfügung:

Korrekturrand



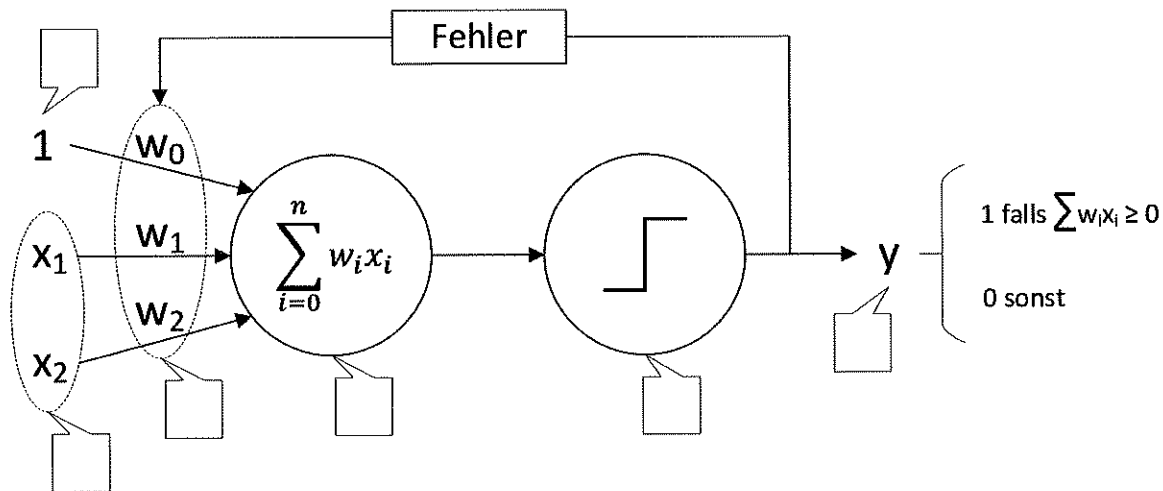
Beschreiben Sie, warum in diesem Fall eine Klassifizierung durch maschinelles Lernen möglich ist.

2 Punkte

d) Die Klassifizierung soll mit einem einfachen Perzeptron (Neuron) erfolgen.

da) Tragen Sie in die Beschriftungssprechblasen die zutreffenden Zahlen ein.

3 Punkte



1	Sprungfunktion
2	Output
3	Gewichte
4	Bias
5	Nettoeingabefunktion
6	Input

db) Erläutern Sie die Funktionsweise des Perzeptrons anhand der gegebenen Skizze.

5 Punkte

Fortsetzung 1. Aufgabe →

Fortsetzung 1. Aufgabe

- e) Sie sollen den Perzeptron-Lernalgorithmus, mit dem die Gewichte w gelernt werden, planen.

Die Trainingsdaten sind in einem zweidimensionalen Array *daten* gespeichert. Auf das x_1 -Feature wird mit dem Spaltenindex 0, auf das x_2 -Feature mit dem Spaltenindex 1 und auf das zugehörige Label mit Spaltenindex 2 zugegriffen. Beachten Sie auch das Klassendiagramm des Perzeptrons.

Perzeptron
<ul style="list-style-type: none"> – weights : Double[3] – anzEpochen : Integer
<ul style="list-style-type: none"> + Perzeptron(anzEpochen : Integer) + lernen(daten : Double[[[]]]) : void + vorhersagen(x : Double[]) : Integer

Weisen Sie die gegebenen Eintragsmöglichkeiten folgerichtig zu, indem Sie die entsprechenden Zahlen in das nachstehende Struktogramm eintragen. 5 Punkte

5 Punkte

1	$x := \text{daten}[j][0, 1]$
2	$\text{weights}[0, 1] := \text{weights}[0, 1] + (\text{update} * x)$
3	$i := i + 1$
4	solange $i < \text{anzEpochen}$
5	$\text{update} := \text{label} - \text{vorhersagen}(x)$
6	$\text{weights}[2] := \text{weights}[2] + \text{update}$
7	wenn $\text{update} \neq 0$
8	$\text{weights} := [0.0, 0.0, 0.0]$
9	$\text{label} := \text{daten}[j][2]$
10	für $j := 0$, solange $j < \text{Zeilenzahl von daten}$, $j := j + 1$

```
+ lernen(daten : Double[][]): void
```

i := 0	
ja	nein
	∅

Korrekturrand

Seriennr	Typ	Zertifikat	letztes_Update	Kunde	Kunde_Kontakt	Kunde_Tel
669-818	ScanFix 5L	IP44	01.01.2021	SoftKey GmbH	Panter, Rosa	030/22459326
289-449	ScanFix 5L	IP44	08.02.2024	SoftKey GmbH	Panter, Rosa	030/22459326
248-581	Cisao DX 4	IP67	30.10.2022	SoftKey GmbH	Panter, Rosa	030/22459326
413-983	DonseM60	IP54	26.11.2018	Elephani AG	Witzka, Heide	040/12345678
247-128	DonseM60	IP54	29.07.2023	Strewitz UG	Otter, Rob	050/71736772
787-639	Cisao DX 4	IP67	10.11.2023	Strewitz UG	Otter, Rob	050/71736772

- 13 Punkte

[illegible]

- b) Nachdem die Datenbank erstellt wurde, müssen die Tabellen angelegt und die Daten eingefügt werden. Ihnen steht dazu der Auszug aus der Pythonklasse Cursor zur Verfügung.

Cursor methods**.close()**

Close the cursor now. The cursor will be unusable from this point forward; an Error (or subclass) exception will be raised if any operation is attempted with the cursor.

.execute(operation [, parameters])

Prepare and execute a database operation (query or command). Parameters may be provided as sequence or mapping and will be bound to variables in the operation. A reference to the operation will be retained by the cursor. Return values are not defined.

An Example without parameters are

```
sql = "SELECT id FROM table"
...execute(sql)
```

An Example with parameters are

```
sql = "SELECT %s FROM %s"
val = ("id", "table")
...execute(sql, val)
```

The second example shows a prepared statement. Unlike ordinary statements, it does not yet contain any parameter values. Instead, placeholders are passed to the database system.

- ba) Die execute Methode ermöglicht, die Werte mit Parameter zu übergeben. In der Programmierung wird dabei auch von Prepared Statements gesprochen.

Erklären Sie den Begriff und einen Vorteil.

4 Punkte

Korrekturrand

9 Punkte

```
database = mysql.connector.connect(host="localhost",user="admin",
password="pw", database="geraet") # Verbindung zur DB herstellen
cursor = database.cursor() # DB öffnen
```

This image shows a full page of blank graph paper. The grid consists of small, equal-sized squares formed by thin black lines. There are no margins, text, or other markings on the page.

3. Aufgabe (28 Punkte)

Korrekturrand

Die Mitarbeiter erfassen Lagerbewegungen mithilfe von MDE-Geräten.

Die Daten werden im JSON-Format mithilfe von REST übertragen.

Die folgende Tabelle stellt Beispiele dar, die in einer Datenbank mithilfe der von den Geräten erfassten Daten gespeichert wurden.

Artikelnr	Zeitpunkt	Lagerbewegung
1	01-APR-23 12.41.08 PM	2
1	01-APR-23 11.12.19 PM	-1

a) Schreiben Sie für den ersten Beispieldatensatz das JSON-Format auf.

2 Punkte

b) Beschreiben Sie, wie Daten mit einem Restful Service bearbeitet werden.

2 Punkte

c) Die folgende Tabelle stellt verfügbare HTTP-Methoden dar, mit denen Daten bearbeitet werden können.

Geben Sie zu jeder Methode ein Beispiel für die Nutzung im Hinblick auf die Verwaltung der Lagerbewegungen an. 4 Punkte

HTTP-Methode	Operation für die Lagerbewegungen
GET	
PUT	
POST	
DELETE	

d) Die Tabelle *lagerbestaende* enthält je Artikel die Lagermengen pro Kalenderwoche.

Korrekturrand

lagerbestaende

ArtikelNr	Kalenderwoche	Menge
1	1	8
1	2	16
1	3	9
1	4	4
1	5	12
2	1	4
2	2	8
2	3	5
2	4	8
2	5	6

da) Geben Sie mithilfe eines SQL-Befehles den maximalen und den minimalen Lagerbestand pro Artikel aus.

4 Punkte

Beispielergebnis:

ArtikelNr	Max	Min
1	16	4
2	8	4

db) Schreiben Sie einen SQL-Befehl, der den Erwartungswert und die Standardabweichung für den Lagerbestand pro Artikel anzeigt. Runden Sie die Ergebniswerte jeweils auf eine ganze Zahl.

7 Punkte

Hinweis. Die Standardabweichung kann mit der Aggregatfunktion STDDEV() berechnet werden.

Beispieltabelle:

ArtikelNr	Erwartungswert	Standardabweichung
1	10	4
2	6	2

- dc) Rechnen Sie die Werte für den Erwartungswert und die Standardabweichung für Artikel 1 manuell nach und interpretieren Sie das Ergebnis. Der Rechenweg ist zu dokumentieren. Die Berechnungsformeln sind am Ende der Aufgabe angegeben.

6 Punkte

- dd) Berechnen Sie den Median für Artikel 1 und interpretieren Sie das Ergebnis. Die Definition für den Median finden Sie am Ende der Aufgabe. 3 Punkte

3 Punkte

[illegible]

Formeln zur 3. Aufgabe

1.1 Erwartungswert

$$\mu = E(X) = \sum_x x \cdot P(X = x)$$

Hier ist x eine bestimmte Ausprägung der Zufallsvariable X und $P(X = x)$ ist die Wahrscheinlichkeit, dass X den Wert x annimmt. Diese Formel gilt für diskrete Zufallsvariablen.

1.2 Varianz

$$\sigma^2 = \sum_i (x_i - \mu)^2 \cdot P(x_i)$$

- σ^2 ist die Varianz.
- x_i sind die möglichen Werte in der Verteilung.
- μ ist der Erwartungswert.
- $P(x_i)$ ist die Wahrscheinlichkeit des Auftretens des Wertes x_i .

1.3 Standardabweichung

$$\sigma = \sqrt{\sigma^2}$$

1.4 Median

Bei einer ungeraden Anzahl von n sortierten Datenpunkten:

$$\tilde{x} = x_{\frac{n+1}{2}}$$

Bei einer geraden Anzahl von n sortierten Datenpunkten:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

4. Aufgabe (21 Punkte)

Korrekturrand

Die Lagerverwaltung soll mithilfe einer Datenanalyse verbessert werden. Die Lagerarbeiter beschwerten sich über lange Wege. Einige Regalfächer sind überfüllt, während andere kaum genutzt werden.

Gegeben ist folgende Tabelle mit teilweise aufbereiteten Daten, die für die Analyse genutzt werden sollen. Die Regale werden nur einmal in der Woche neu befüllt.

	AID	Lagerhalle	Regal	Fach	Durchschnittliche wöchentliche Entnahmen	Maximale wöchentliche Entnahme	Regalfachkapazität
1.	1	1	4	5	253,7	356	null
2.	1	1	3	5	187,6	200	200
3.	1	1	3	5	187,6	200	20
4.	2	1	3	5	null	null	400
5.	2	1	1	2	777,2	550	500
6.	3	2	1	3	-6	300	300
7.	3	2	1	4	250	300	300
8.	4	2	2	null	40,5	55	600

a) Nennen Sie drei wichtige Schritte zur Datenaufbereitung.

3 Punkte

ba) Beschreiben Sie vier Probleme mit Beispielen, die man aus der Tabelle entnehmen kann.

8 Punkte

bb) Machen Sie anhand der Beispieldaten drei begründete Vorschläge, wie man mit den problematischen Daten umgehen kann.

6 Punkte

Fortsetzung 4. Aufgabe →

Fortsetzung 4. Aufgabe

Korrekturrand

c) Die Laufwege der Lagerarbeiter sollen zu Optimierungszwecken aufgezeichnet werden.

Beschreiben Sie zwei Aspekte, die aus datenschutzrechtlicher Sicht relevant sind.

4 Punkte

PRÜFUNGSZEIT – NICHT BESTANDTEIL DER PRÜFUNG!

Wie beurteilen Sie nach der Bearbeitung der Aufgaben die zur Verfügung stehende Prüfungszeit?

- ☐ 1 Sie hätte kürzer sein können.
- ☐ 2 Sie war angemessen.
- ☐ 3 Sie hätte länger sein müssen.

☐