



FAKULTÄT FÜR
INFORMATIK

Responsible Data Science

Bias Simulations Studie

Agenda

1. Das Team: BIASpects
2. Projektidee
3. Theorie
4. Das Szenario
5. Unsere Hypothesen
6. Das Experiment
7. Die Ergebnisse
8. Reduzierung des Bias
9. Zusammenfassung
10. Ausblick

1. Das Team: BIASpects

- **B.Sc. Gesa Götte**
Statistik Master (4.Semester)
- **B.Eng. Marcel Öfele**
Digital Engineering (4.Semester)
- **B.Eng. Viviane Lisa Wolters**
Digital Engineering (4.Semester)

2. Projektidee

Agenda

1. Das Team: BIASpects

2. Projektidee

3. Theorie

4. Das Szenario

5. Unsere Hypothesen

6. Das Experiment

7. Die Ergebnisse

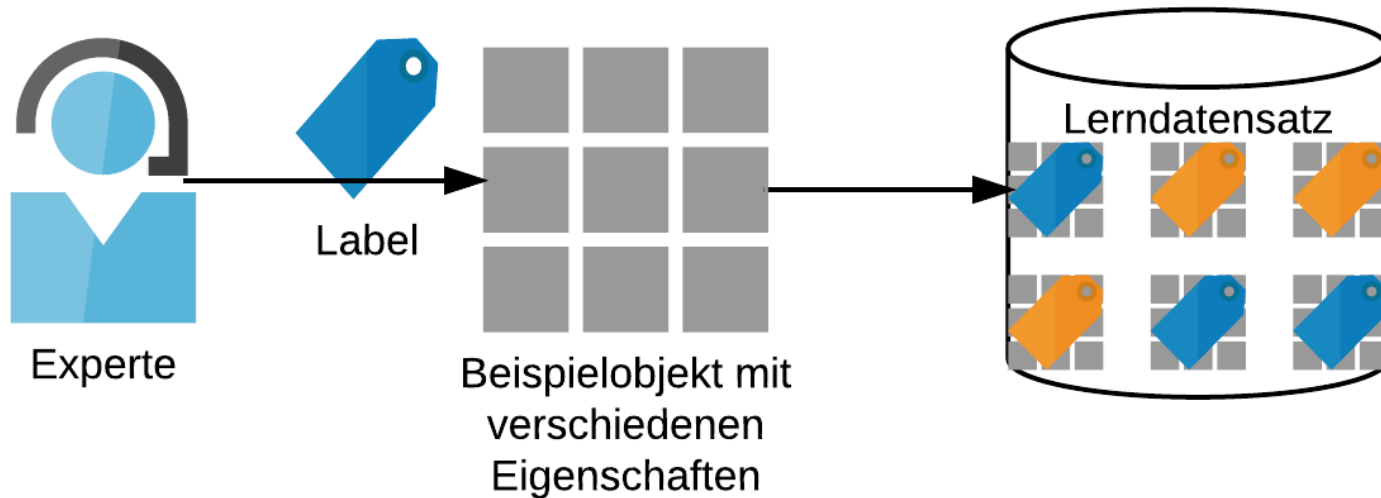
8. Reduzierung des Bias

9. Zusammenfassung

10. Ausblick

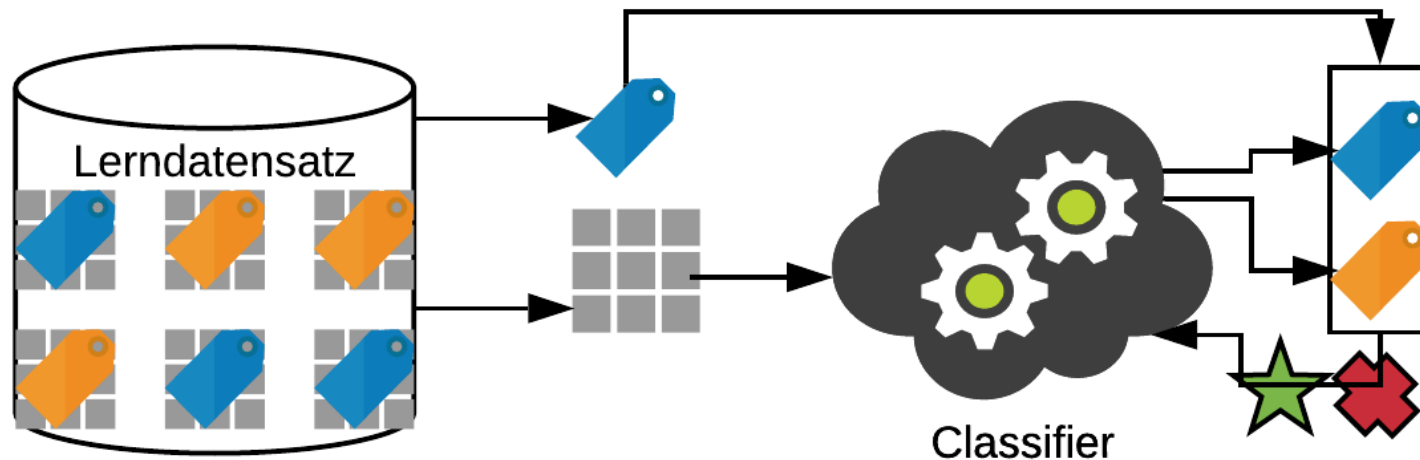
3.1 Arbeitsweise eines Machine Learning Classifier

Erstellung Lerndatensatz



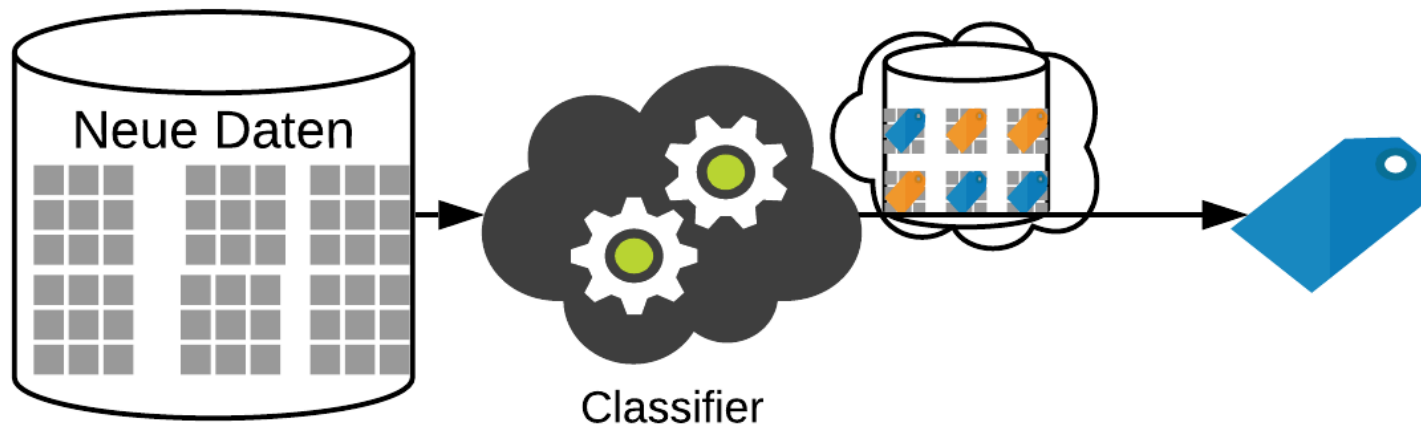
3.1 Arbeitsweise eines Machine Learning Classifier

Lernphase des Classifiers



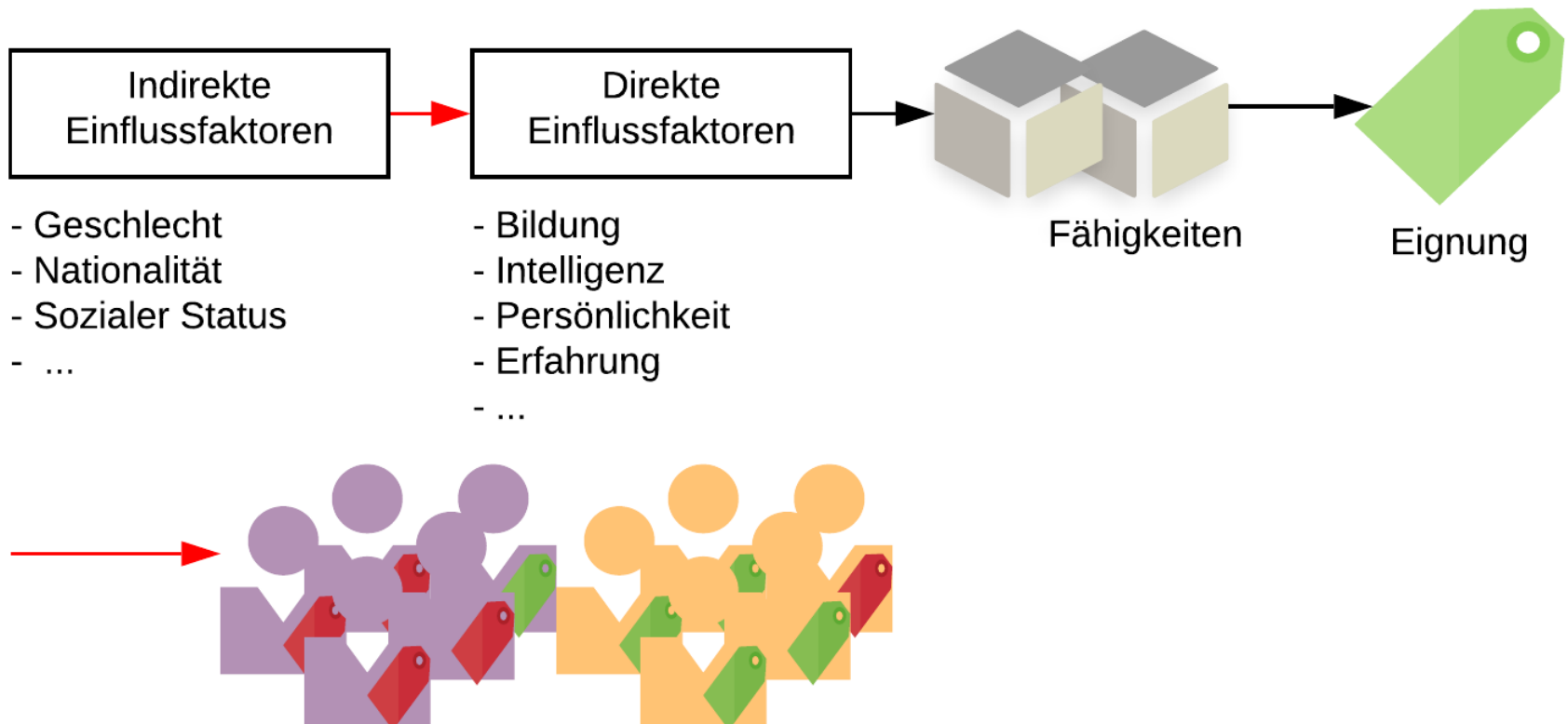
3.1 Arbeitsweise eines Machine Learning Classifier

Anwendung des Classifiers



3.2 Definition und Entstehung des Bias

Bias = Verzerrung der Wirklichkeit



3.2 Definition und Entstehung des Bias

Bias im Datensatz wird gelernt



Auswirkung auf spätere Beurteilungen



Agenda

1. Das Team: BIASpects

2. Projektidee

3. Theorie

4. Das Szenario

5. Unsere Hypothesen

6. Das Experiment

7. Die Ergebnisse

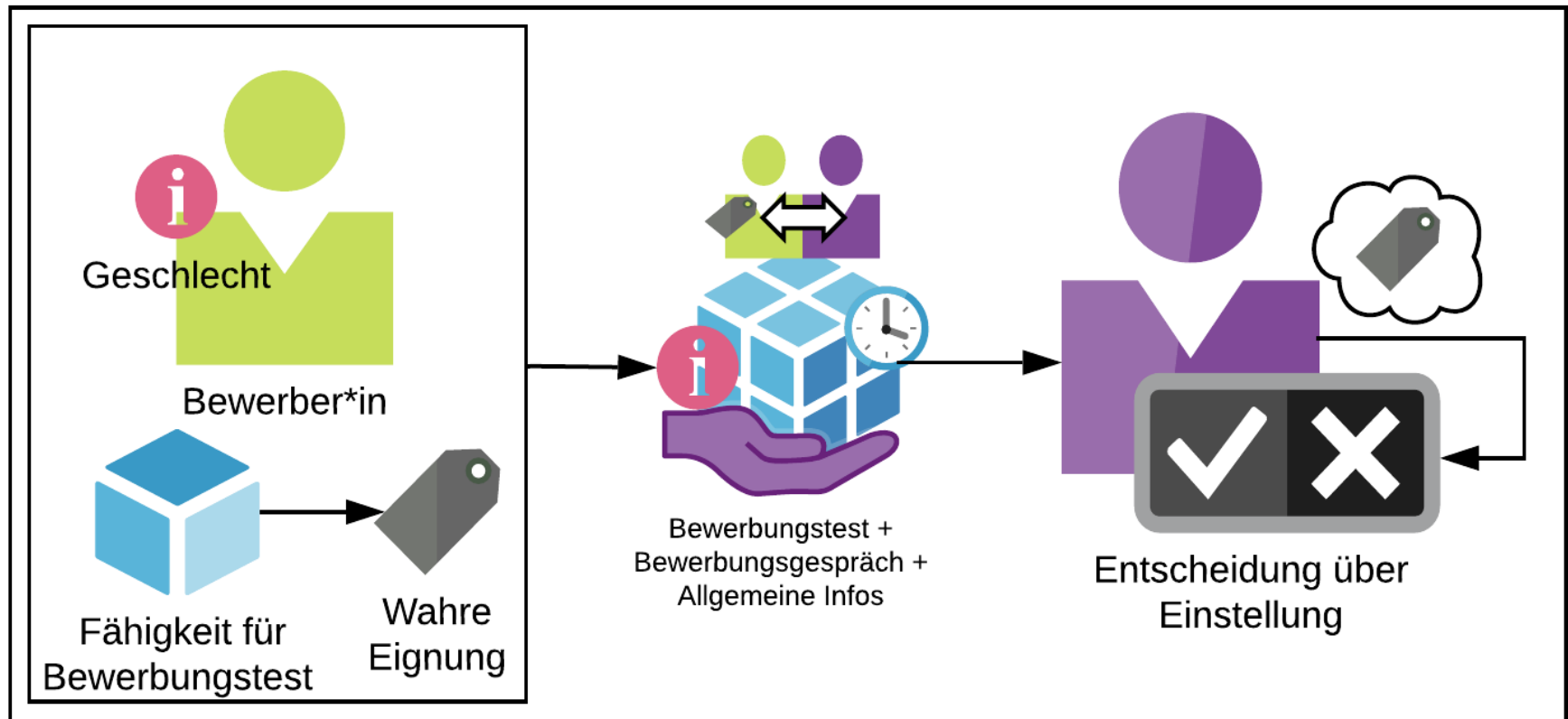
8. Reduzierung des Bias

9. Zusammenfassung

10. Ausblick

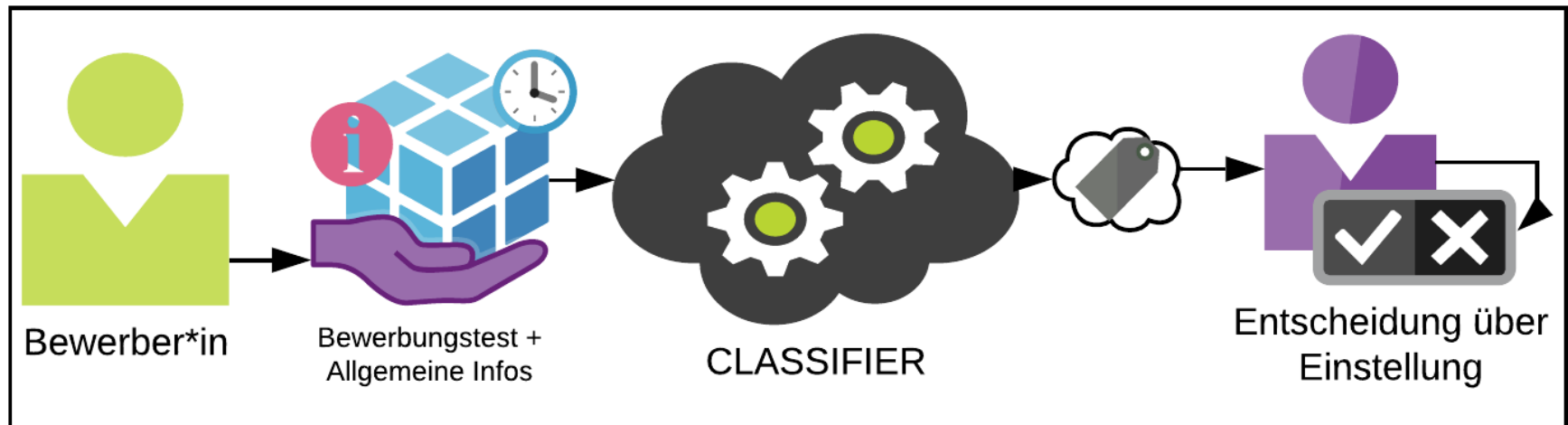
4. Das Szenario

BEWERBUNGSPROZESS BISHER



4. Das Szenario

NEUER BEWERBUNGSPROZESS



Agenda

1. Das Team: BIASpects

2. Projektidee

3. Theorie

4. Das Szenario

5. Unsere Hypothesen

6. Das Experiment

7. Die Ergebnisse

8. Reduzierung des Bias

9. Zusammenfassung

10. Ausblick

5.1 Hypothese 1

Kein Bias, wenn

- ▶ Anteil geeigneter Frauen = Anteil geeigneter Männer
(Anzahl der Frauen bzw. Männer nicht relevant)

5.2 Hypothese 2

Kein Bias, wenn

➤ Anzahl Frauen = Anzahl Männer

(Anteil geeigneter Frauen bzw. Männer nicht relevant)

5.3 Hypothese 3

Bias, wenn

- Anzahl Frauen \neq Anzahl Männer
- Anteil geeigneter Frauen \neq Anteil geeigneter Männer

5.3.1 Hypothese 3.1

- ▶ Je größer die Überrepräsentation, desto größer der Bias

5.3.2 Hypothese 3.2

- ▶ Je größer der Unterschied zwischen den Anteilen der geeigneten Männer und Frauen, desto größer der Bias.

Agenda

1. Das Team: BIASpects

2. Projektidee

3. Theorie

4. Das Szenario

5. Unsere Hypothesen

6. Das Experiment

7. Die Ergebnisse

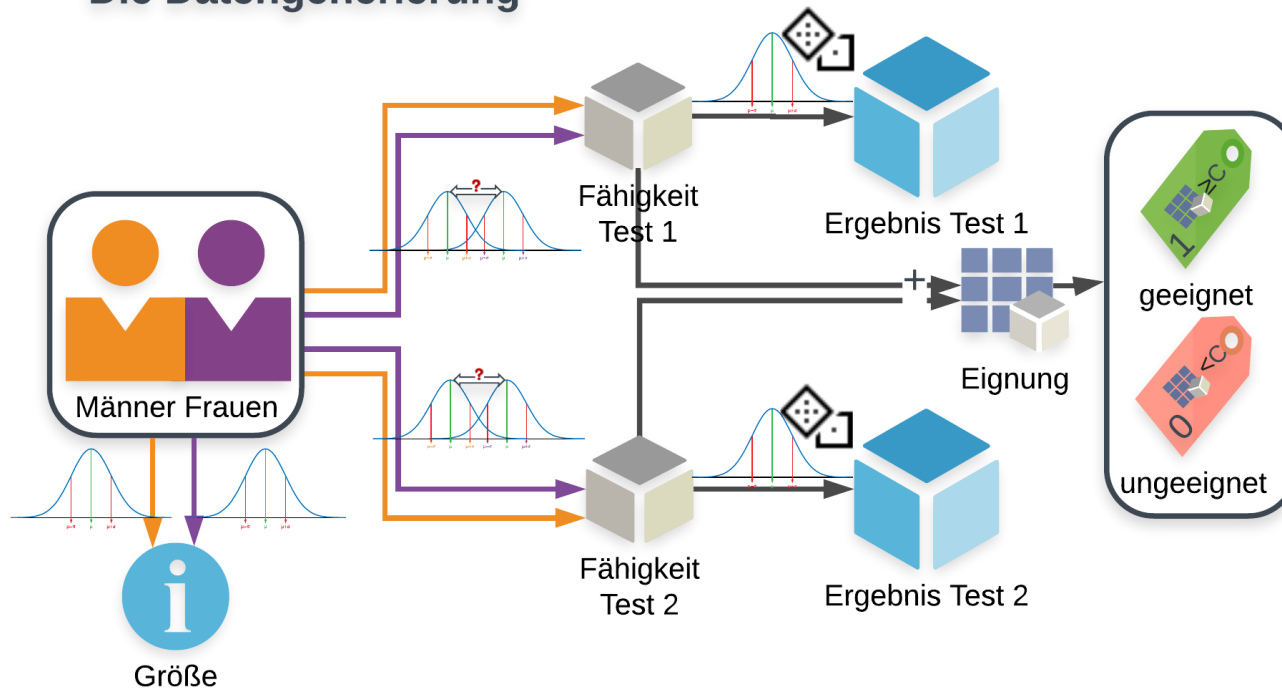
8. Reduzierung des Bias

9. Zusammenfassung

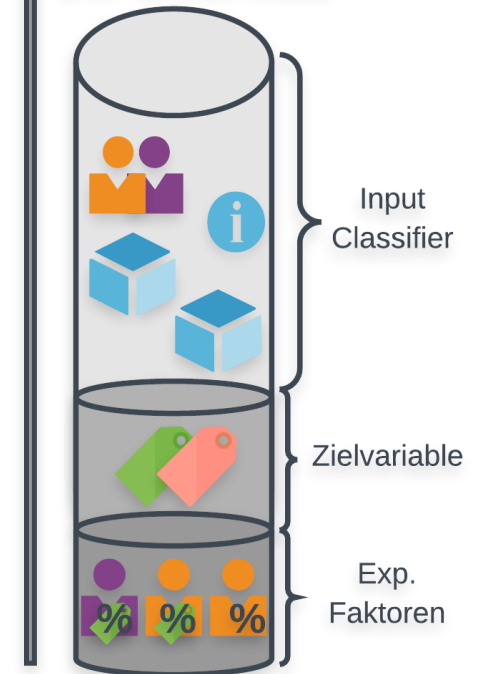
10. Ausblick

6.1 Datensatz

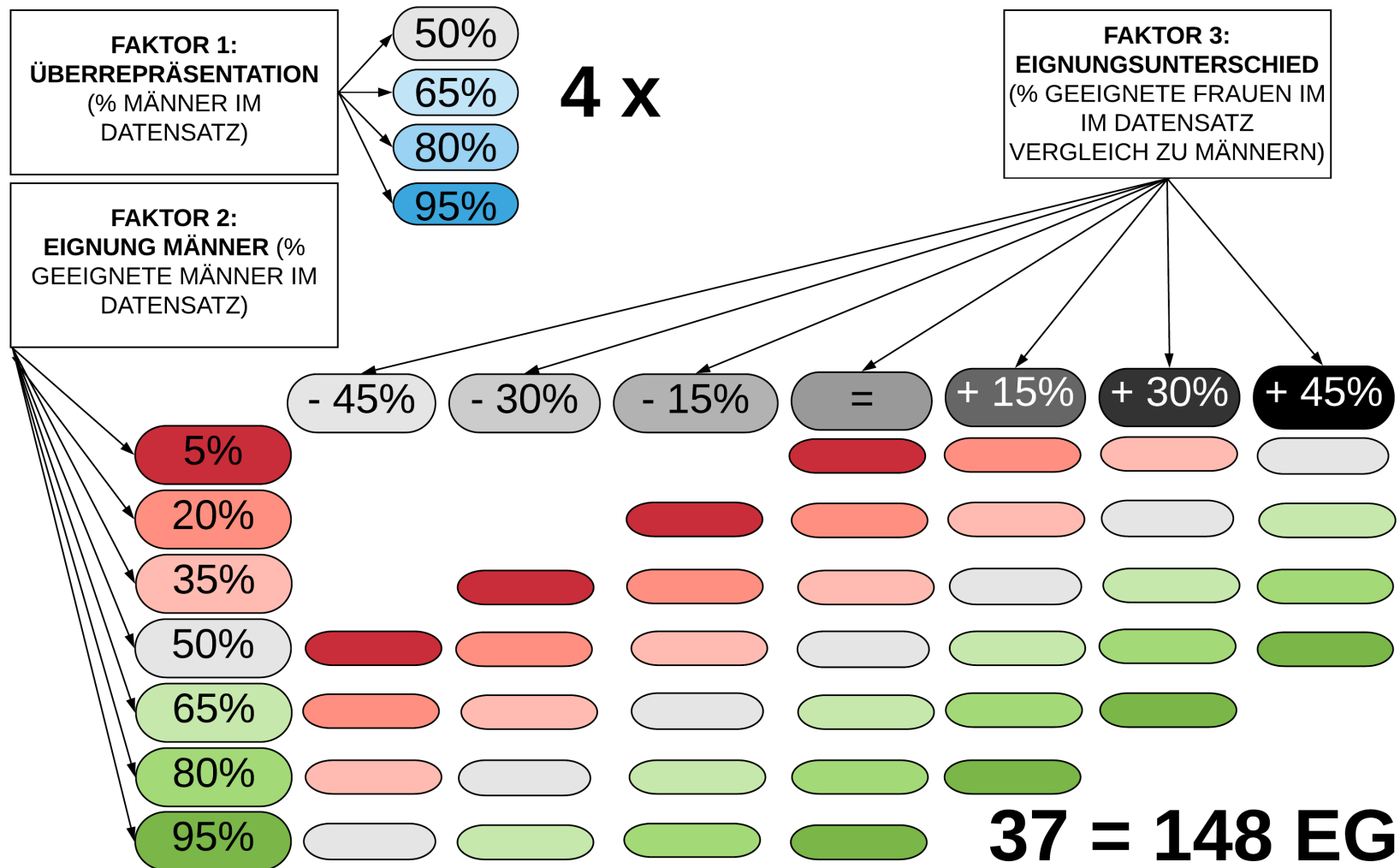
Die Datengenerierung



Der Datensatz



6.2 Aufbau/Ablauf



Agenda

1. Das Team: BIASpects

2. Projektidee

3. Theorie

4. Das Szenario

5. Unsere Hypothesen

6. Das Experiment

7. Die Ergebnisse

8. Reduzierung des Bias

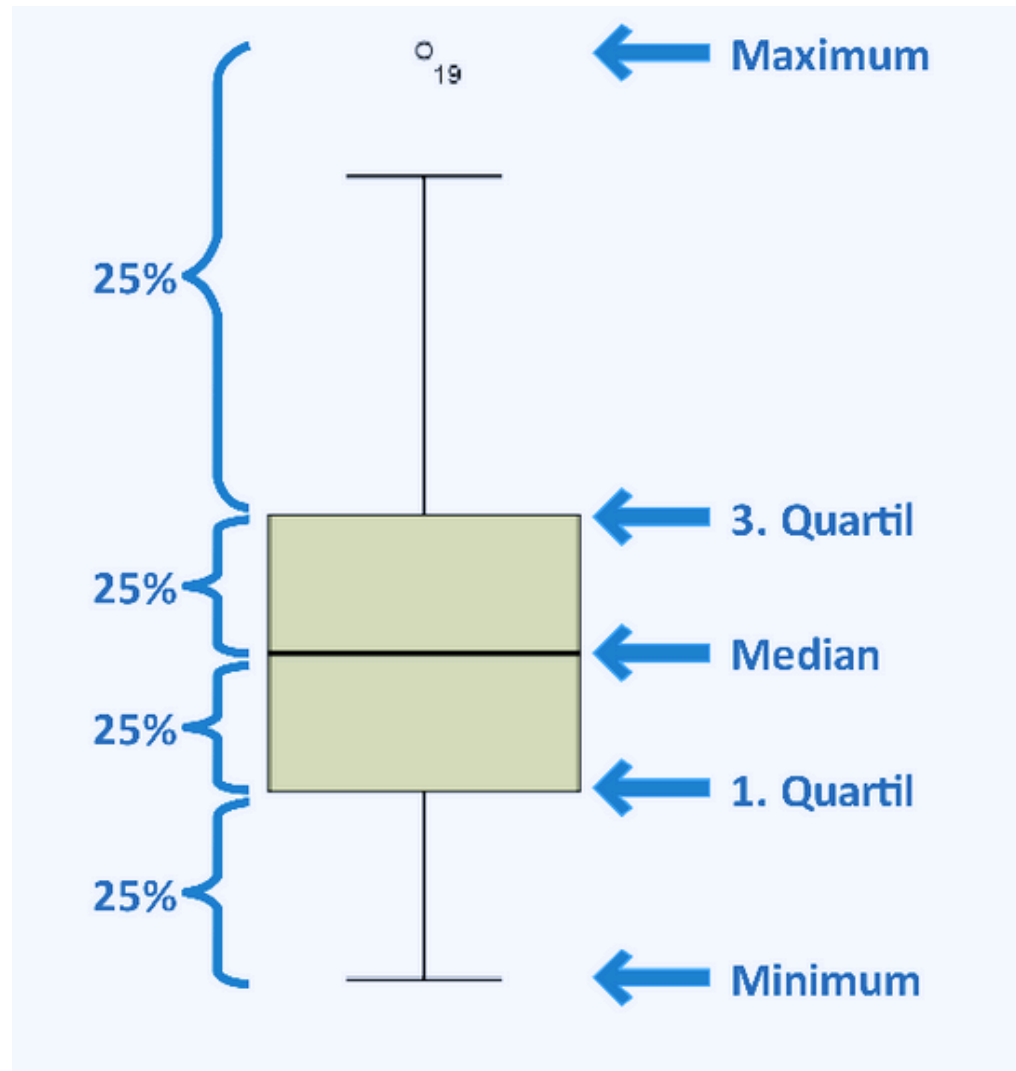
9. Zusammenfassung

10. Ausblick

7.1 Unterschied zwischen den Classifiers



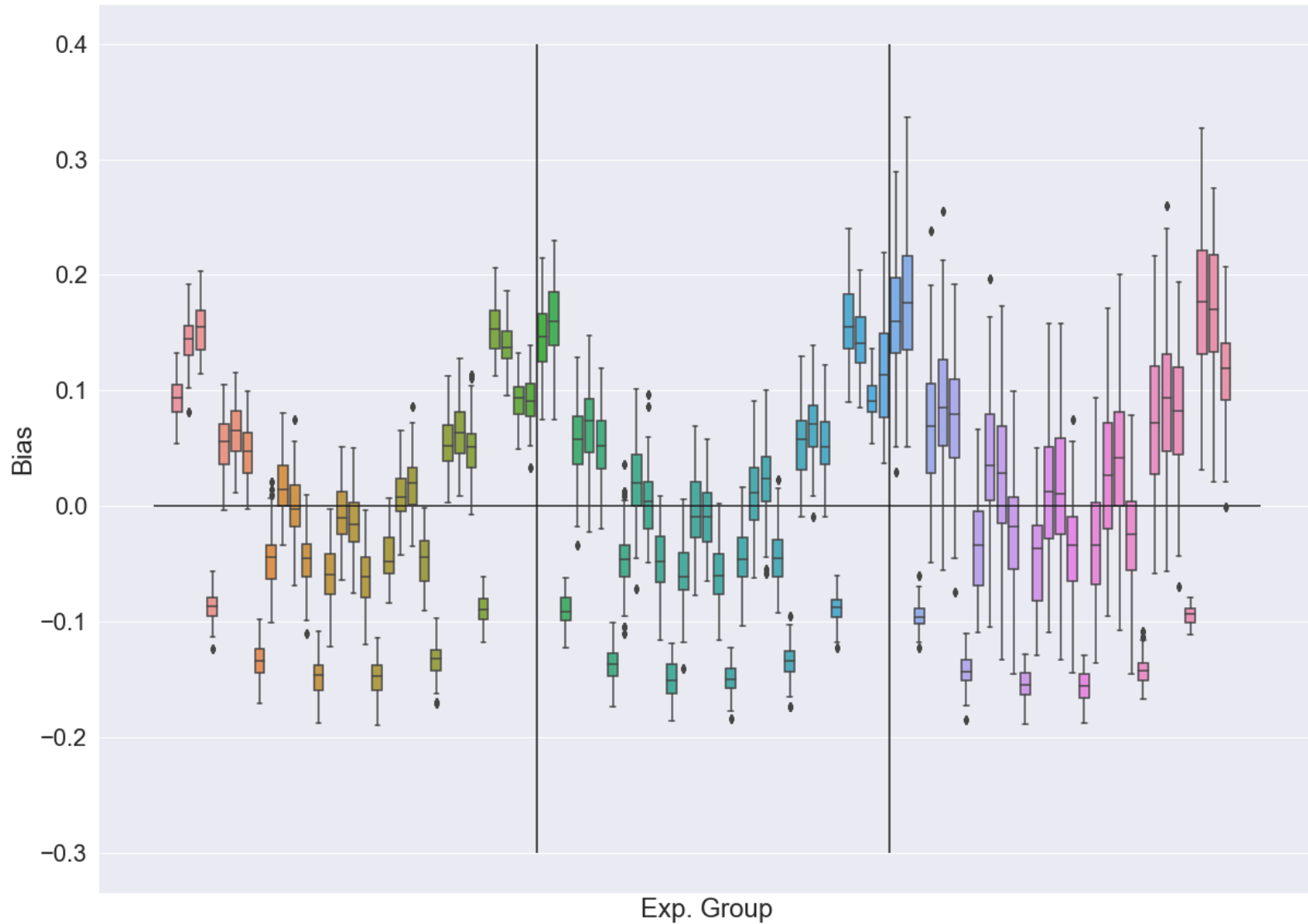
7.2 Exkurs: Boxplots



7.3 Hypothese 1

7.4 Hypothese 2

7.5 Hypothese 3



7.5.1 Hypothese 3.1

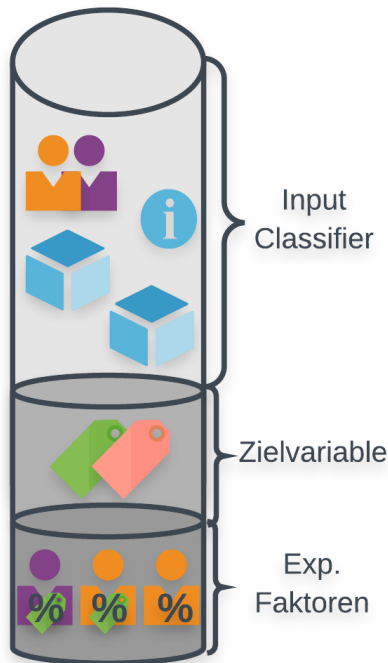
7.5.2 Hypothese 3.2

Agenda

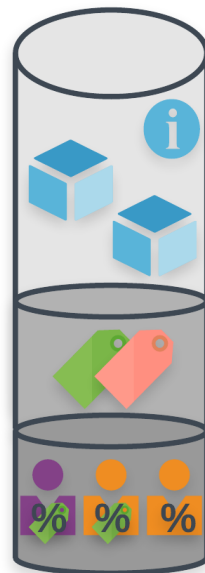
1. Das Team: BIASpects
2. Projektidee
3. Theorie
4. Das Szenario
5. Unsere Hypothesen
6. Das Experiment
7. Die Ergebnisse
- 8. Reduzierung des Bias**
9. Zusammenfassung
10. Ausblick

8.1 Methoden

Standard A0



Methode A1 ("Ohne Geschlecht"):
Entfernen des Geschlechts



Methode A2 ("Ohne geschlechtsspezifische Informationen"):
Entfernen aller Merkmale mit Geschlechtsinformation



Methode B ("50-50-50"):
Lerndatensatz mit 50% Männern, 50% geeigneten Männern und 50% geeigneten Frauen (Testdatensatz wie gehabt)



8.1 Methode 1

8.1 Methode 2

9. Zusammenfassung

10. Ausblick

Vielen Dank für Ihre Aufmerksamkeit!

www.ovgu.de