

Evaluation und Interpretation

Definieren Sie für Ihr Modell die Frequenzbaseline bzw. die Mittelwertsbaseline

```
Gesamtzahl Datenpunkte Training: 1288
Anzahl Kategorie 0 58
Anzahl Kategorie 1 526
Anzahl Kategorie 2 474
Anzahl Kategorie 3 230
```

Die häufigste Klasse ist 1 (Normalgewicht) mit der Frequenz $F=0,41$.

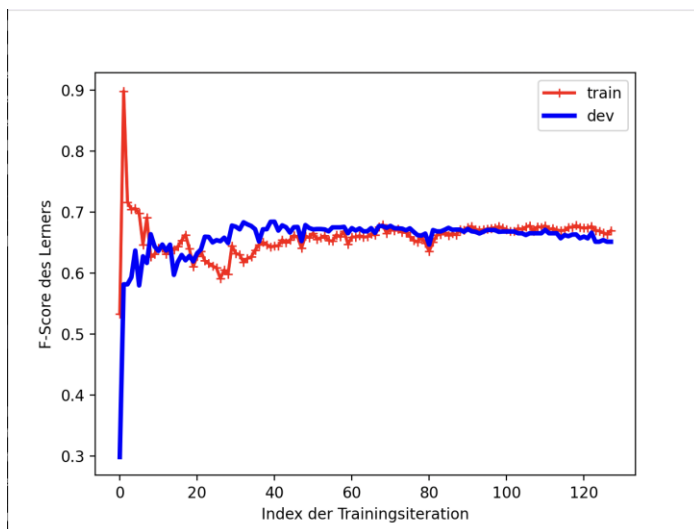
Definieren einer einfachen Vergleichsbaseline

	precision	recall	f1-score	support
0	0.10	0.93	0.18	15
1	0.66	0.20	0.31	132
2	0.67	0.41	0.51	118
3	0.52	0.65	0.58	57
accuracy			0.39	322
macro avg	0.49	0.55	0.39	322
weighted avg	0.61	0.39	0.42	322

Gemessen an dieser Baseline sind die Ergebnisse mit dem SVM, Random Forest und Logistic Regression deutlich besser: Dort war der weighted average F-Score 0.67, 0.86 und 0.68.

Prüfen Sie mittels einer Lernkurve, ob Ihr Modell zu Over- oder Underfitting neigt und evaluieren Sie entsprechend des Ergebnisses ein mächtigeres oder weniger mächtiges Modell. Wenn Ihr Modell weder Over- noch Underfitting zeigt: Herzlichen Glückwunsch, es ist nichts weiter zu tun.

Für die Lernkurven wurden für die Schrittweite 10er Schritte gewählt.

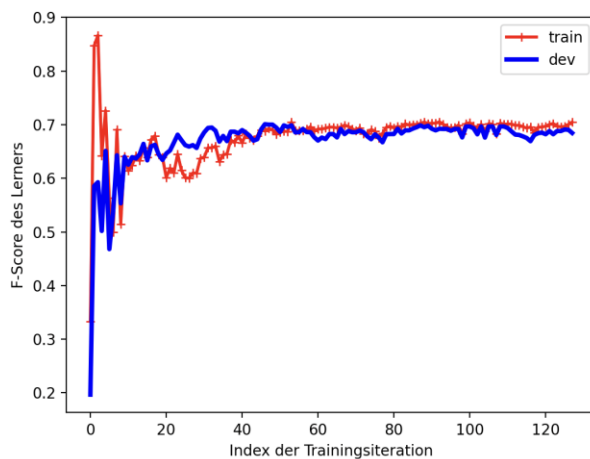


Analyse:

Die rote und die blaue Kurve liegen eng beieinander und stabilisieren sich auf einem ähnlichen Niveau

Die F1-Scores sind jedoch nicht besonders hoch (~0.7).

Diagnose: Das Modell zeigt weder Overfitting noch starkes Underfitting



Analyse:

Die rote und die blaue Kurve liegen ebenfalls eng beieinander und stabilisieren sich auf einem ähnlichen Niveau

Beide stabilisieren sich auf einem Niveau von etwa 0.7

Diagnose: Das Modell zeigt weder Overfitting noch starkes Underfitting

Interpretieren Sie Ihr Modell: Entweder mit Hilfe von LIME oder bei transparenten Algorithmen aufgrund des gelernten Modells selber.

LIME

```
[[0.007 0.956 0.035 0.002]]
Document id: 1
Predicted class: Normalgewicht
True class: Normalgewicht
Explanation for class Normalgewicht
('Age <= -0.82', 0.5805750876300088)
('-1.11 < Sex <= 0.90', 0.08075238061155644)
('-1.10 < Number_of_Main_Meals_Daily <= 0.21', 0.05581942703778003)
('-1.36 < Height <= 0.21', 0.04740590873429989)
('-1.52 < Overweight_Obese_Family <= 0.66', 0.02232837181867114)
Explanation for class Übergewicht
('Age <= -0.82', -0.4284552358255697)
('-1.11 < Sex <= 0.90', -0.06330472011384741)
('-1.36 < Height <= 0.21', 0.045218871395017894)
('-1.52 < Overweight_Obese_Family <= 0.66', 0.032071809355377345)
('-1.10 < Number_of_Main_Meals_Daily <= 0.21', -0.02071139135627209)
```

Wir betrachten nur das Modell Logistic Regression.

Die erste Zeile zeigt, dass das Modell die Klasse „Normalgewicht“ mit einer Wahrscheinlichkeit von 95,6% vorhergesagt hat, „Übergewicht“ mit 3,5% und die anderen mit 0,2% und 0,7%. Somit war die vorhergesehene Klasse „Normalgewicht“, was mit der echten Klasse übereinstimmt.

Die wichtigsten Merkmale, die zur Vorhersage von „Normalgewicht“ beitrugen, waren das Alter (stärkster Einfluss mit positivem Beitrag von 58,1%), gefolgt von dem Geschlecht, der Körpergröße und der Häufigkeit des Gemüseverzehrs. Für die Klasse „Übergewicht“ hatte das Alter einen negativen Einfluss (-42,8%) auf die Vorhersage, während die „Anzahl an Täglichem Essen“ und die Körpergröße geringe positive Beiträge leisten.

Auch bei anderen Datenpunkten hatte das Alter stets einen großen Beitrag zur Vorhersage. Dies macht Sinn, da im Datensatz jüngere Menschen einen tendenziell geringeren BMI hatten. Die Analyse zeigt, dass das sogar der entscheidendste Faktor im Modell ist.