Model Analysis | Predict New Data

Health Equality Dashboard

Model performance and fairness across subgroups (Gender × Sleep)

Subgroup Performance (Training Results)

	Subgroup	Sample_Size	Threshold	Accuracy	F1_Score	ROC_AUC
0	Female_Normal	6	0.52	0.8333	0.9091	
1	Female_Short	4	0.52	0.75	0.8571	
2	Other_Short	4	0.52	0.75	0.8571	
3	Other_Normal	7	0.52	0.7143	0.8333	
4	Male_Normal	7	0.52	0.7143	0.8333	
5	Male_Short	4	0.52	0.5	0.6667	

Summary Metrics

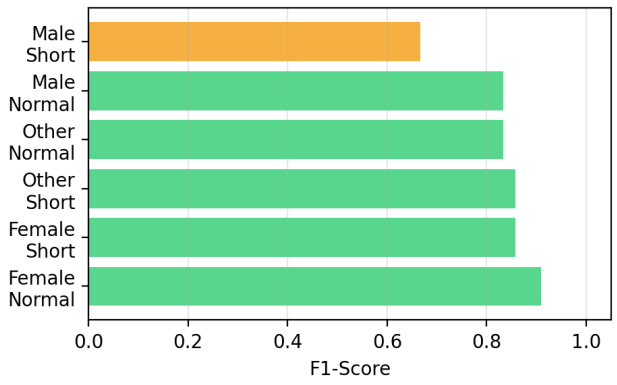
Mean F1
0.83

Mean Accuracy
0.71

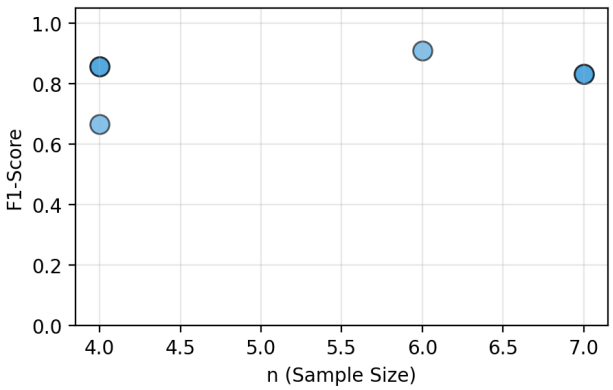
Std F1 (fairness)
0.083

Live Visualizations (from training data)

F1 by Subgroup

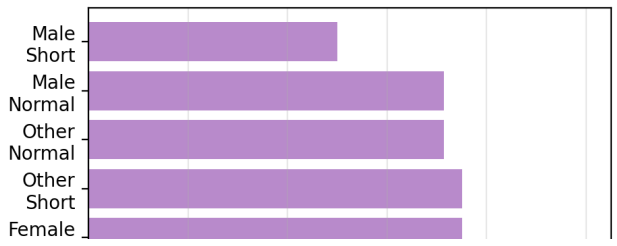


Sample Size vs F1

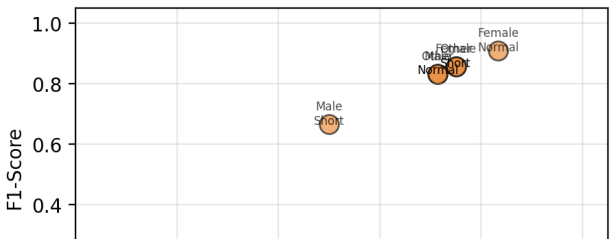


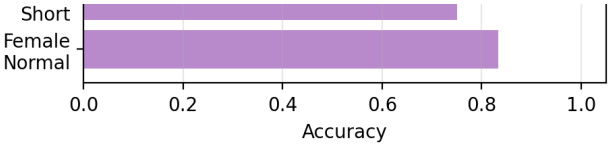
Additional Analysis

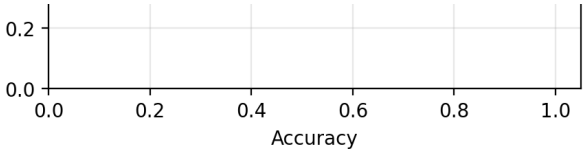
Accuracy by Subgroup



F1 vs Accuracy





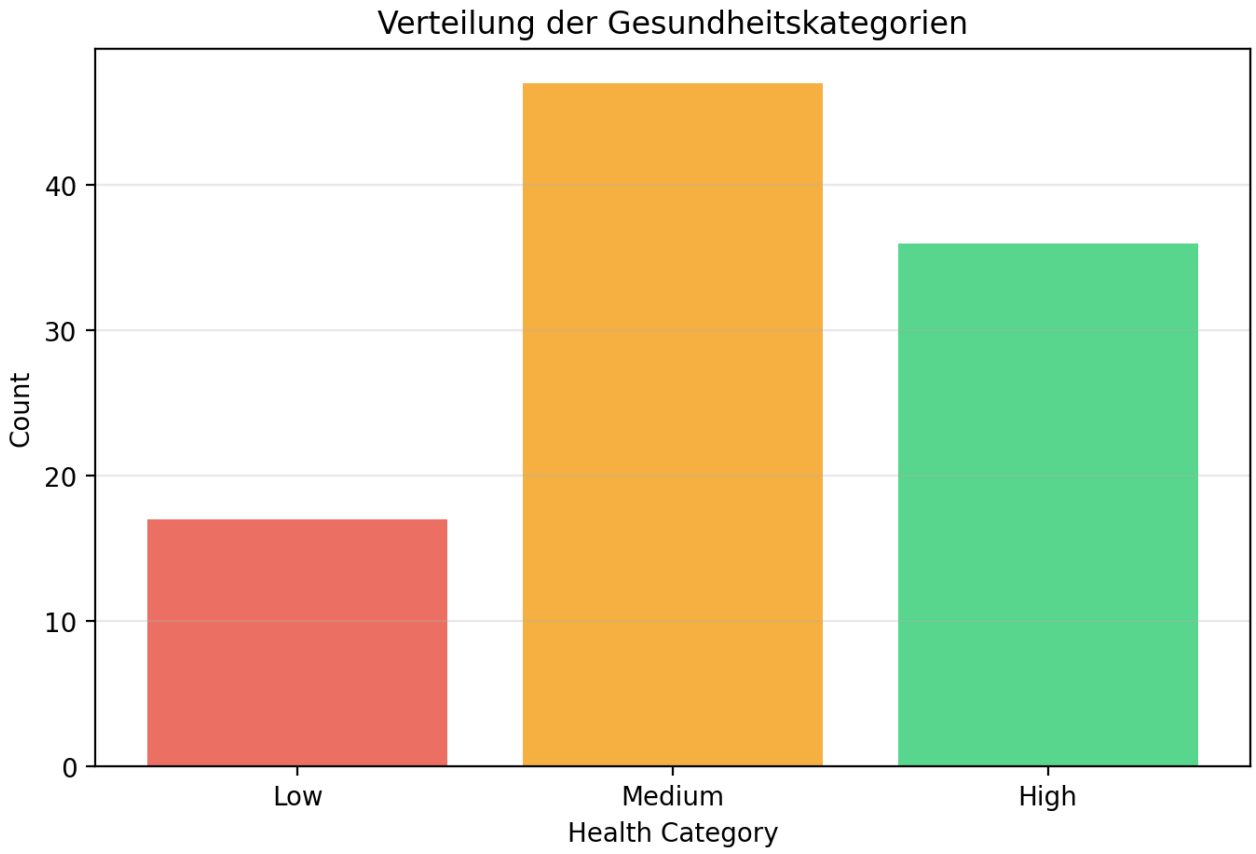


Random Forest Model Analysis

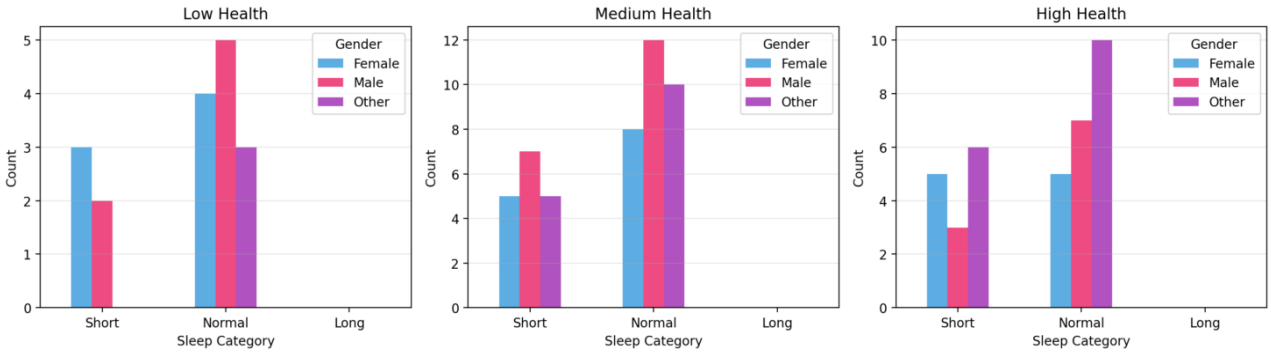
Generated from health.py - Random Forest subgroup analysis with feature engineering

Generate RF Visualizations

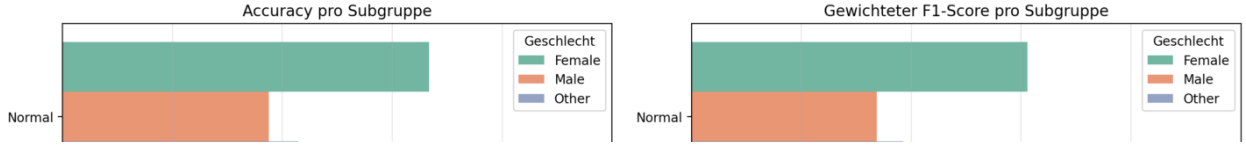
Health Category Distribution

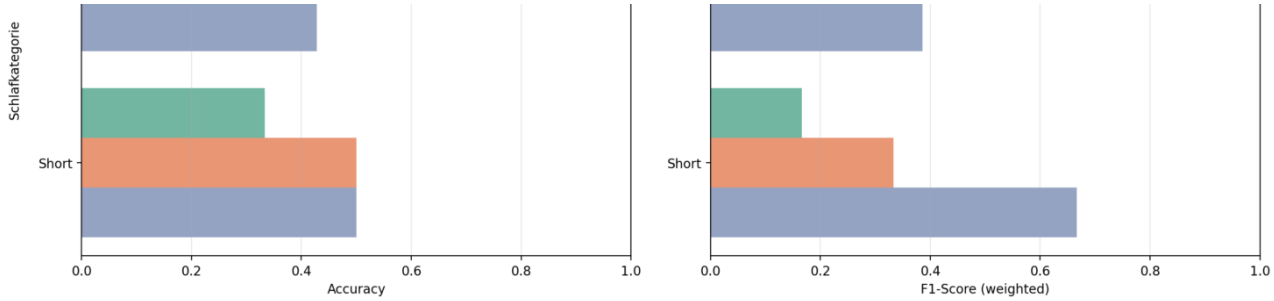


Subgroup Distribution (Gender × Sleep Category)



Random Forest: Subgroup Performance Comparison





	Gender	Sleep Category	Sample Count	Accuracy
0	Female	Normal	6	0.6667
1	Female	Short	3	0.3333
2	Male	Normal	8	0.375
3	Male	Short	4	0.5
4	Other	Normal	7	0.4286
5	Other	Short	2	0.5

✅ Random Forest analysis completed!

Probability Calibration (Validation)

Reliability curve: predicted probability vs observed frequency on validation data

Compute Calibration Curve

Training Data (CSV)

	Student ID	Age	Gender	Blood Oxygen Level (%)	Body Temperature (°C)	Heart Rate (bpm)	Physical Activity Level (METs)	Overall Health Score	Date and
5	1005	18	Male	99.5	37.1	62	2	78	2024-10-
6	1006	16	Female	97.5	36.5	85	5.2	76	2024-10-
7	1007	16	Other	98.5	36.4	80	4	78	2024-10-
8	1008	18	Other	98.2	36.8	100	4.7	96	2024-10-
9	1009	17	Female	98.8	36.7	67	6.7	82	2024-10-
10	1010	17	Female	98	36.4	84	2.6	70	2024-10-
11	1011	16	Female	98.8	36.3	67	7.1	94	2024-10-
12	1012	15	Female	95.6	37.3	80	6.8	86	2024-10-
13	1013	17	Female	100	36.3	81	2.1	93	2024-10-
14	1014	18	Other	95.7	36.2	92	7.4	84	2024-10-

📥 Download Training Data

Stacking Ensemble Model (4 Base Models)

🎯 Ensemble uses **ALL 4 MODELS**: RF + SVM + MLP + HGB → Logistic Regression → Final Prediction

Meta-Model: Logistic Regression (stacks predictions from all 4 base models)

Meta-Model Parameters

4 Base Models in Ensemble:

RandomForestClassifier

RF Parameters


```
{
  "cv" : 3
  "ensemble" : "auto"
  "estimator__bootstr" : true
  "estimator__ccp_alp" : 0
  "estimator__class_w" : "balanced"
  "estimator__criteri" : "gini"
  "estimator__max_dep" : 9
  "estimator__max_fea" : "log2"
  "estimator__max_lea" : NULL
  "estimator__max_sam" : NULL
  "estimator__min_imp" : 0
  "estimator__min_sam" : 1
  "estimator__min_sam" : 4
  "estimator__min_wei" : 0
  "estimator__monoton" : NULL
  "estimator__n_estim" : 138
  "estimator__n_jobs" : NULL
  "estimator__oob_sco" : false
```


SVM (Calibrated)

SVM Parameters

MLPClassifier

MLP Parameters


```
{
  "cv" : 3
  "ensemble" : "auto"
  "estimator__activat" : "relu"
  "estimator__alpha" : 0.003972893133963025
  "estimator__batch_s" : "auto"
  "estimator__beta_1" : 0.9
  "estimator__beta_2" : 0.999
  "estimator__early_s" : false
  "estimator__epsilon" : 1e-8
  "estimator__hidden_" [
    0 : 64
    1 : 32
  ]
  "estimator__learnin" : "constant"
  "estimator__learnin" : 0.002741758423592905
  "estimator__max_fun" : 15000
  "estimator__max_ite" : 500
  "estimator__momentu" : 0.9
  "estimator__n_iter_" : 10
  "estimator__nestero"
```


HistGradientBoostingClassifier

HGB Parameters


```
{
  "cv" : 3
  "ensemble" : "auto"
  "estimator__categor" : "from_dtype"
  "estimator__class_w" : NULL
  "estimator__early_s" : "auto"
  "estimator__interac" : NULL
  "estimator__l2_regu" : 0
  "estimator__learnin" : 0.29999999999999993
  "estimator__loss" : "log_loss"
  "estimator__max_bin" : 255
  "estimator__max_dep" : 8
  "estimator__max_fea" : 1
  "estimator__max_ite" : 150
  "estimator__max_lea" : 31
  "estimator__min_sam" : 17
  "estimator__monoton" : NULL
  "estimator__n_iter_" : 10
  "estimator__random_" : 42
```

```
"estimator__random_
": 42
"estimator__verbose
": 0
"estimator__warm_st
": false
"estimator":
"RandomForestClassif
max_depth=9,

min_samples_split=4,

random_state=42)"
"method":
"isotonic"
"n_jobs": NULL
}
```

```
{
"cv": 3
"ensemble":
"auto"
"estimator__C":
0.7756486208064789
"estimator__break_t
": false
"estimator__cache_s
": 200
"estimator__class_w
": "balanced"
"estimator__coef0":
0
"estimator__decisio
": "ovr"
"estimator__degree
": 3
"estimator__gamma":
0.5703843027403095
"estimator__kernel
": "rbf"
"estimator__max_ite
": -1
"estimator__probabi
": true
"estimator__random_
": 42
"estimator__shrinki
": true
"estimator__tol":
0.001
"estimator__verbose
": false
"estimator":
"SVC(C=0.77564862080
class_weight='balanc
gamma=0.570384302740
probability=True
random_state=42)"
"method":
"sigmoid"
}
```

```
: true
"estimator__power_t
": 0.5
"estimator__random_
": 42
"estimator__shuffle
": true
"estimator__solver
": "adam"
"estimator__tol":
0.0001
"estimator__validat
": 0.1
"estimator__verbose
": false
"estimator__warm_st
": false
"estimator":
"MLPClassifier(alpha
hidden_layer_sizes=(

learning_rate_init=0
max_iter=500,
random
"method":
"sigmoid"
"n_jobs": NULL
}
```

```
"estimator__scoring
": "loss"
"estimator__tol":
1e-7
"estimator__validat
": 0.1
"estimator__verbose
": 0
"estimator__warm_st
": false
"estimator":
"HistGradientBoostin
max_depth=8,


"method":
"isotonic"
"n_jobs": NULL
}
```


Selected Threshold

0.520

✅ All 4 models (RF + SVM + MLP + HGB) work together in the ensemble for best prediction!

Regenerate Pipeline

🔄 Regenerate outputs (run pipeline on trained data)