

- In diesem Modul gilt es, das in allen Modulen Erlernte selektiv anzuwenden und das Verständnis darüber zu vertiefen
- Ausgehend von seinem selbst gewählten Fallbeispiel klärt der Teilnehmende die Ausgangssituation. Die dahinterliegenden geschäftsrelevanten Fragestellungen werden herausgearbeitet. Der Lösungsansatz wird entworfen, Technologieentscheidungen werden getroffen, die einzusetzenden Werkzeuge bestimmt
- Das Praxisprojekt wird durch die beteiligten Professoren begleitet und der Teilnehmende entsprechend gecoacht
- In einer Abschlusspräsentation wird das Projekt vorgestellt

Stundenplan

vorbehaltlich Änderungen!

Tag	Datum	Uhrzeit	Krz.	Fach / Bemerkung	Trainer	TE	Ort
MI	15.05.2024	08:30 - 16:30		Einführung in Data Science und Python (M1)	Prof.Dr. Brauer Jürgen	8	online/zoom
DO	16.05.2024	08:30 - 16:30		Einführung in Data Science und Python (M1)	Prof.Dr. Brauer Jürgen	8	online/zoom
FR	31.05.2024	08:30 - 16:30		Einführung in Machine Learning (M2)	Prof.Dr. Brauer Jürgen	8	online/zoom
FR	07.06.2024	08:30 - 16:30		Einführung in Machine Learning (M2)	Prof.Dr. Brauer Jürgen	8	online/zoom
MO	10.06.2024	08:30 - 16:30		Deep Learning (M3)	Prof.Dr. Brauer Jürgen	8	online/zoom
MO	17.06.2024	08:30 - 16:30		Deep Learning (M3)	Prof.Dr. Brauer Jürgen	8	online/zoom
MI	19.06.2024	08:30 - 16:30		Machine Learning und Deep Learning Vertiefung (M4)	Prof.Dr. Brauer Jürgen	8	online/zoom
DO	20.06.2024	08:30 - 16:30		Machine Learning und Deep Learning Vertiefung (M4)	Prof.Dr. Brauer Jürgen	8	online/zoom
MI	26.06.2024	08:30 - 16:30		Big Data und Datenbanken (M5)	Prof.Dr. Brauer Jürgen	8	online/zoom
DO	27.06.2024	08:30 - 16:30		Big Data und Datenbanken (M5)	Prof.Dr. Brauer Jürgen	8	online/zoom
MI	03.07.2024	08:30 - 16:30		AutoML und Machine Learning der Cloud (M6)	Prof.Dr. Brauer Jürgen	8	online/zoom
DO	04.07.2024	08:30 - 16:30		AutoML und Machine Learning der Cloud (M6)	Prof.Dr. Brauer Jürgen	8	online/zoom
FR	05.07.2024	08:30 - 12:30		Projektpräsentation	Prof.Dr. Brauer Jürgen	4	online/zoom

Projekte - Ablauf

- **Überlegen eines kleinen Projektes auf Basis der bisherigen Themen**
z.B. Datenanalyse auf eigenen oder Kaggle-Daten
z.B. Machine Learning auf eigenen oder Kaggle-Daten
- **Im Modul 5: Vorstellungsrunde der Projektideen / Zwischenstand**
- **Am Fr, 05.07.24: Abschlusspräsentation – Was ist rausgekommen?**
Jeder: ca. 5-6 Minuten Vorstellung der Ergebnisse + 4 Minuten Diskussion = 10 Minuten

Projektidee-/Zwischenstandsvorstellung: Andreas Braumann

■ Problemstellung:

- Visuelles Evaluieren und Vergleichen von verschiedenen Modellen und Meta-Parametern zur Evaluierung der Genauigkeit mit Hilfe einer Streamlit App

■ Daten:

Scikit-learn Datensatz –Nachrichten aus einer Auswahl von 20 Newsgroups (https://scikit-learn.org/stable/modules/generated/sklearn.datasets.fetch_20newsgroups.html)

■ (Weitere) Geplante Vorgehensweise:

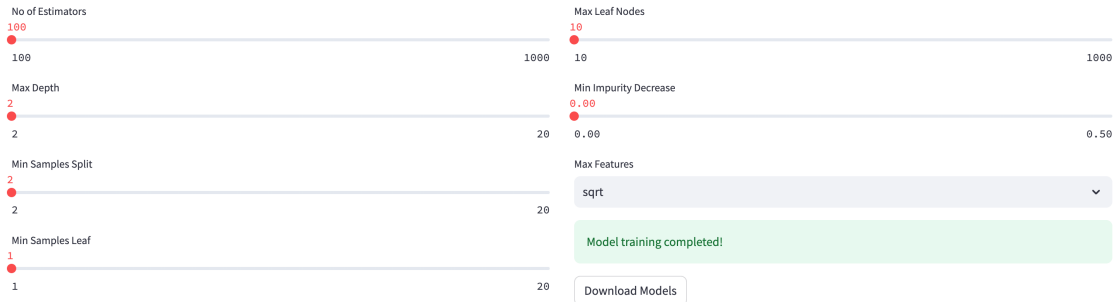
- Streamlit App Template verwenden und erweitern, weitere Parameter hinzufügen, 3 Modelle trainieren, Auswertungen anzeigen, Hosting in der Cloud
- Erweiterung des Dashboards mit mehreren Modellen, zusätzliche Auswertungen und Parametern
- Vergleich von Parametern
- Verfügbar hier: <https://streamlit-ml-dashboard.c.divisionbyheroes.com/>
- Public Github Repo: https://github.com/brewmanandi/streamlit_ml_dashboard

Projekt-Bericht: Andreas Braumann

- Was ist rausgekommen? (Ergebnisse)
 - Siehe folgende Slides
- Was habe ich dabei gelernt?
 - Erstellung und Deployment einer Streamlit App mittels Docker und Caprover auf einem virtuellen Server auf DigitalOcean
 - Grafische Darstellung von Metriken
 - Vergleich von mehreren Classifiern
- Was muss ich noch lernen?
 - Ergänzung von zusätzlichen Meta Parametern
 - Ergänzung von neuronalen Netzen
 - Zusätzliche Visualisierungen

Classifier Experiments

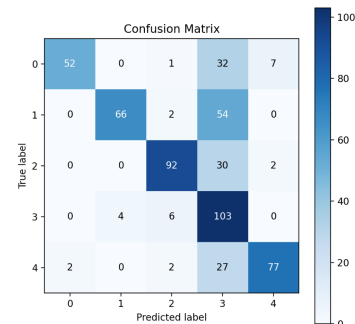
Try different meta parameters. Select widget values and submit model for training. Various ML metrics will be displayed after training. Change any value to retrain.



RandomForestClassifier DecisionTreeClassifier

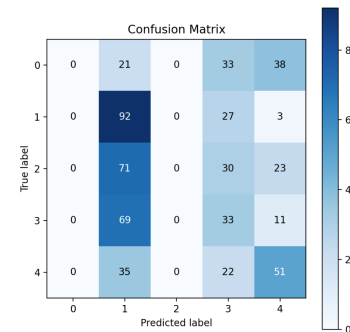
Test Accuracy: 69.77 %
Train Accuracy: 70.26 %

Confusion Matrix



Classification Report:

Confusion Matrix

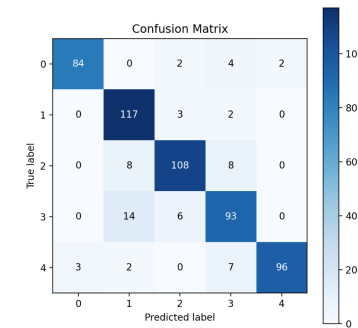


Classification Report:

GradientBoostingClassifier

Test Accuracy: 89.09 %
Train Accuracy: 93.07 %

Confusion Matrix



Classification Report:

Classification Report:

	precision	recall	f1-score
alt.atheism	0.96	0.57	0.71
comp.graphics	0.94	0.54	0.69
rec.autos	0.89	0.74	0.81
sci.electronics	0.42	0.91	0.57
talk.politics.guns	0.90	0.71	0.79
accuracy			0.70
macro avg	0.82	0.69	0.72
weighted avg	0.82	0.70	0.72

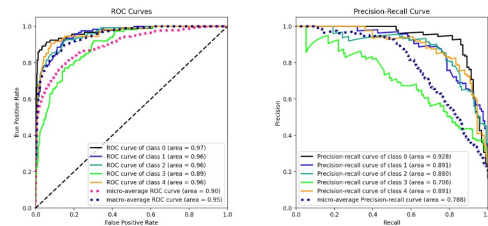
Classification Report:

	precision	recall	f1-score
alt.atheism	0.00	0.00	0.00
comp.graphics	0.32	0.75	0.45
rec.autos	0.00	0.00	0.00
sci.electronics	0.23	0.29	0.26
talk.politics.guns	0.40	0.47	0.44
accuracy			0.31
macro avg	0.19	0.30	0.23
weighted avg	0.19	0.31	0.23

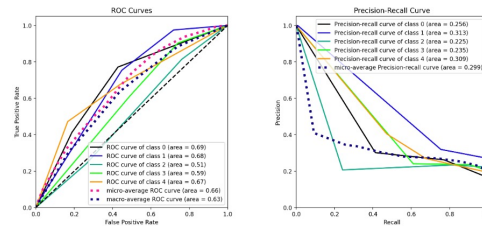
Classification Report:

	precision	recall	f1-score
alt.atheism	0.97	0.91	0.94
comp.graphics	0.83	0.96	0.89
rec.autos	0.91	0.87	0.89
sci.electronics	0.82	0.82	0.82
talk.politics.guns	0.98	0.89	0.93
accuracy			0.89
macro avg	0.90	0.89	0.89
weighted avg	0.90	0.89	0.89

ROC & Precision-Recall Curves



ROC & Precision-Recall Curves



ROC & Precision-Recall Curves

