

# Ein Vergleich von Klassifikatoren am Beispiel von Tinnitus Patientengruppen

## Beschreibung

Dieses Projekt handelt von einem Vergleich von Machine Learning Klassifikatoren, die auf einem Tinnitus Patientendatensatz verglichen werden.

## Vorgehensweise

Zuallererst wurden die Daten vorverarbeitet. Danach wurden die Daten in Training-, Test und Validationsdaten aufgeteilt. Die Trainingsdaten wurden dann mittels RandomOverSampler aufgefüllt. Daraufhin wurde RFECV angewandt (zu finden im Ordner Vorverarbeitung\_Daten). Anschließend wurde Hyperparametertuning der Modelle auf dem Trainingsdatensatz betrieben. Als Nächstes wurden die Modelle mit optimalen Hyperparametern trainiert und mittels Metriken verglichen.

## Ordnerstruktur und Dateienstruktur:

Der Ordner src enthält jeglichen Quellcode zum Projekt aufgeteilt in den Ordnern Hyperparametertuning, Klassifikatoren, Merkmalauswahlverfahren und Vorverarbeitung Daten.

## Ordner Vorverarbeitung Daten:

-In CompletedPreprocessing.ipnb ist die Vorverarbeitung hinsichtlich Behandlung fehlender Werte, das löschen bzw. transformieren von Datenpunkten, Kombination von Merkmalen und erstes Verwerfen von Merkmalen.

-Data\_split.ipynb enthält das Aufteilen der Daten in Trainings- Test und Validationsdaten sowie das Auffüllen des Trainingsdatensatzes mittels RandomOverSampler. Außerdem ist dort der Baseline Klassifikator zu finden.

## Ordner Merkmalauswahlverfahren:

Feature\_selection\_by\_RFECV.ipynb enthält das Merkmalauswahlverfahren RFECV

## Ordner Hyperparametertuning:

Enthält das Hyperparameterertuning der Klassifikatoren. Das Hyperparametertuning wurde mittels GridSearchCV durchgeführt.

## Ordner Klassifikator:

Enthält die jeweiligen Klassifikatoren mit Training und Performancetest.

## Autor:

Filmon Mesfun

.