

Werkzeuge für das wissenschaftliche Arbeiten

Python for Machine Learning and Data Science
Abgabe: 15.12.2023

Inhaltsverzeichnis

1	Projektaufgabe	2
1.1	Einleitung	2
1.2	Aufbau	2
1.3	Methoden	3
2	Abgabe	3

1 Projektaufgabe

In dieser Aufgabe beschäftigen wir uns mit Objektorientierung in Python. Der Fokus liegt auf der Implementierung einer Klasse, dabei nutzen wir insbesondere auch Magic Methods.

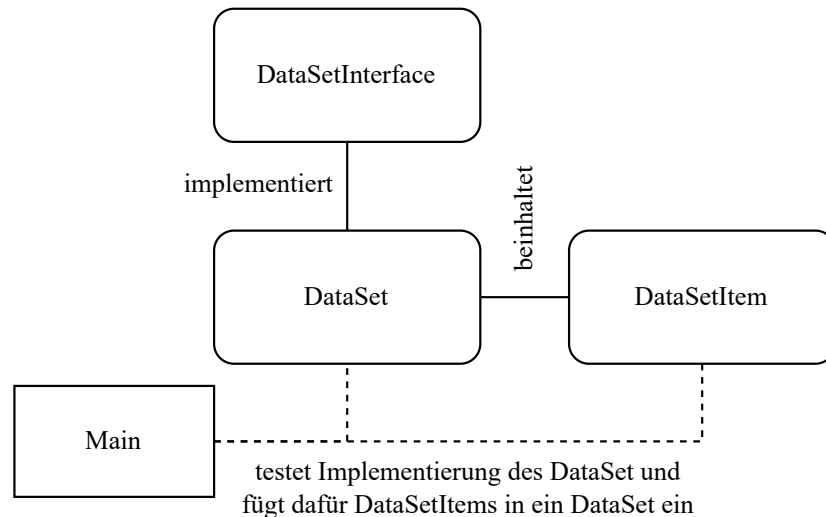


Abbildung 1: Darstellung der Klassenbeziehungen.

1.1 Einleitung

Ein Datensatz besteht aus mehreren Daten. Ein einzelnes Datum wird durch ein Objekt der Klasse `DataSetItem` repräsentiert. Jedes Datum hat:

- Einen Namen (Zeichenkette),
- Eine ID (Zahl),
- Beliebigen Inhalt.

Mehrere Daten (Objekte vom Typ `DataSetItem`) werden in einem Datensatz zusammengefasst. Die Klasse `DataSetInterface` definiert die Schnittstelle und die unterstützten Operationen eines Datensatzes.

1.2 Aufbau

Es gibt drei Dateien:

- `dataset.py`: Enthält die Klassen `DataSetInterface` und `DataSetItem`.
- `implementation.py`: Hier wird die Klasse `DataSet` implementiert.
- `main.py`: Nutzt die Klassen `DataSet` und `DataSetItem` und testet deren Funktionalität.

1.3 Methoden

Folgende Methoden der Klasse `DataSet` sind zu implementieren:

- `__setitem__(self, name, id_content)`: Hinzufügen eines Datums mit Name, ID und Inhalt.
- `__iadd__(self, item)`: Hinzufügen eines `DataSetItem`.
- `__delitem__(self, name)`: Löschen eines Datums anhand des Namens.
- `__contains__(self, name)`: Prüfen, ob ein Datum mit diesem Namen vorhanden ist.
- `__getitem__(self, name)`: Abrufen eines Datums über seinen Namen.
- `__and__(self, dataset)`: Schnittmenge zweier Datensätze bestimmen und als neuen Datensatz zurückgeben.
- `__or__(self, dataset)`: Vereinigungen zweier Datensätze bestimmen und als neuen Datensatz zurückgeben.
- `__iter__(self)`: Iteration über alle Daten des Datensatzes.
- `filtered_iterate(self, filter)`: Gefilterte Iteration mit einer Lambda-Funktion.
- `__len__(self)`: Anzahl der Daten abrufen.

2 Abgabe

Programmieren Sie die Klasse `DataSet` in der Datei `implementation.py`, um die oben beschriebenen Anforderungen zu erfüllen. Nutzen Sie die bereitgestellten Dateien aus dem Moodle.

Das VPL prüft den Code mithilfe erweiterter Testfälle in der `main.py`. Alle Testfälle sind erfolgreich, wenn `CodeIsValid` ausgegeben wird.