Group 3: Regression

1. Introduction

2. Coding

3. Locally Weighted Regression in Detail

4. Runtime Performance

5. Live Demo

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

Introduction

A Brief Recap

- **Initial Plan:**

  - Focus: Linear Regression

  - Task: Implement five regression models

  - Validation: Compare with established implementations

- **Received Feedback:**

  - Focus on 1 or 2 models

  - Add educational value to the project

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

2

Introduction

Our New Approach

- Chose OLS and LWR as focus models

- Retained comparative study

- Developed two web applications

 - One for education

  - One for model visualization

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

3

2. Coding

- Dev setup
- SWE principals

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

4

2.1 Coding - Dev setup

- Fixed python version via pyenv (python 3.11)

- Dependencies loaded via pip into a virtual env of the `venv` module of the standard lib

- Code formatting with `black`

- Linting with `pylint`

- Reviews of pushed code (rather informal)

- Usage of github issue tracker

2.2 SWE principals -- Separation of concerns

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

6
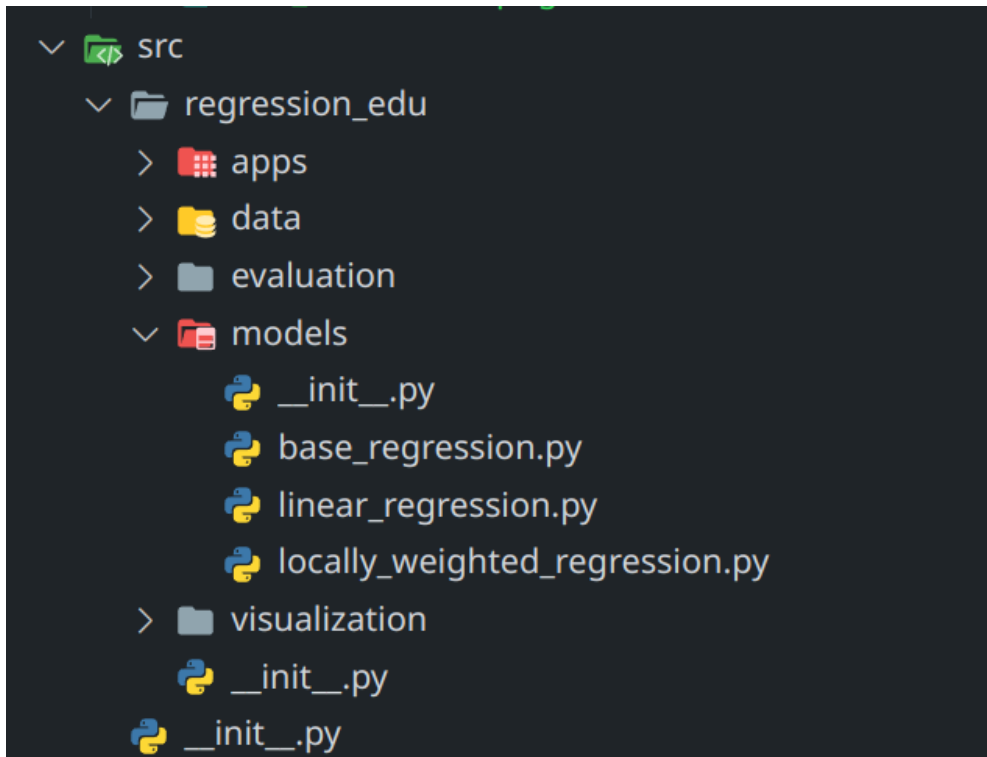
2.3 SWE principals -- Inheritance
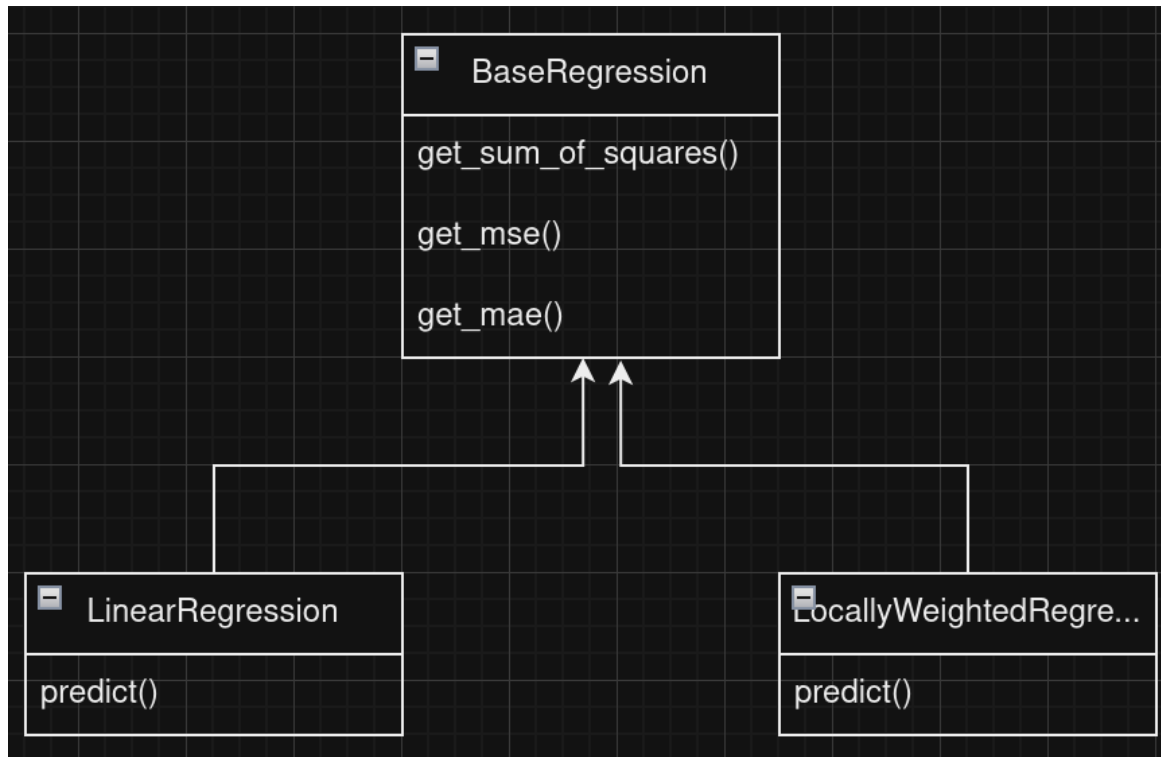
3. Locally Weighted Regression (LWR) in Detail

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

8

3.1 LWR in Detail

1. Divide into sections

2. For each section, calculate the weighted regression with weight
$$w_i(x) = e^{-\frac{(centre_i - x)^2}{2\tau^2}}$$

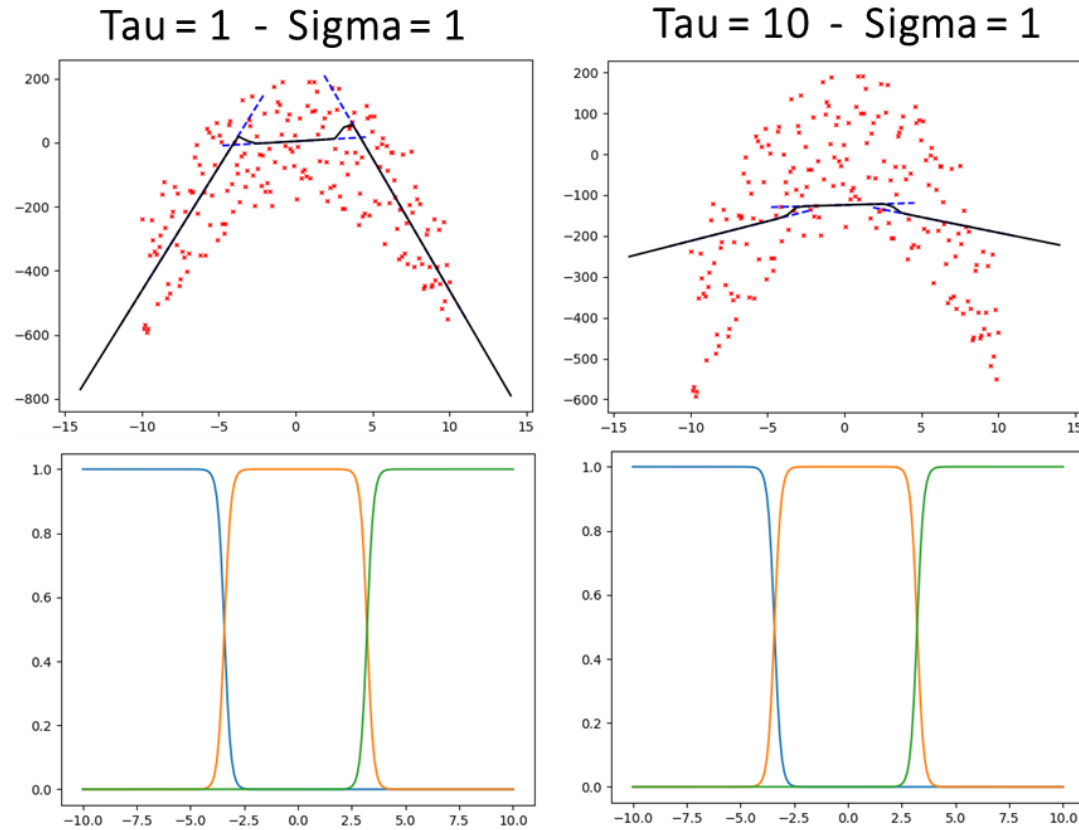3. Smoothen the function with $gauss_{centre}(x) = e^{-\frac{(centre - x)^2}{2\sigma^2}}$ and normalising it by dividing through $\Sigma_{centre}\ gauss_{centre}(x)$
$$f(x) = \frac{1}{\Sigma_i\ gauss_i(x)}\Sigma_i gauss_i(x) \cdot f_i(x)$$
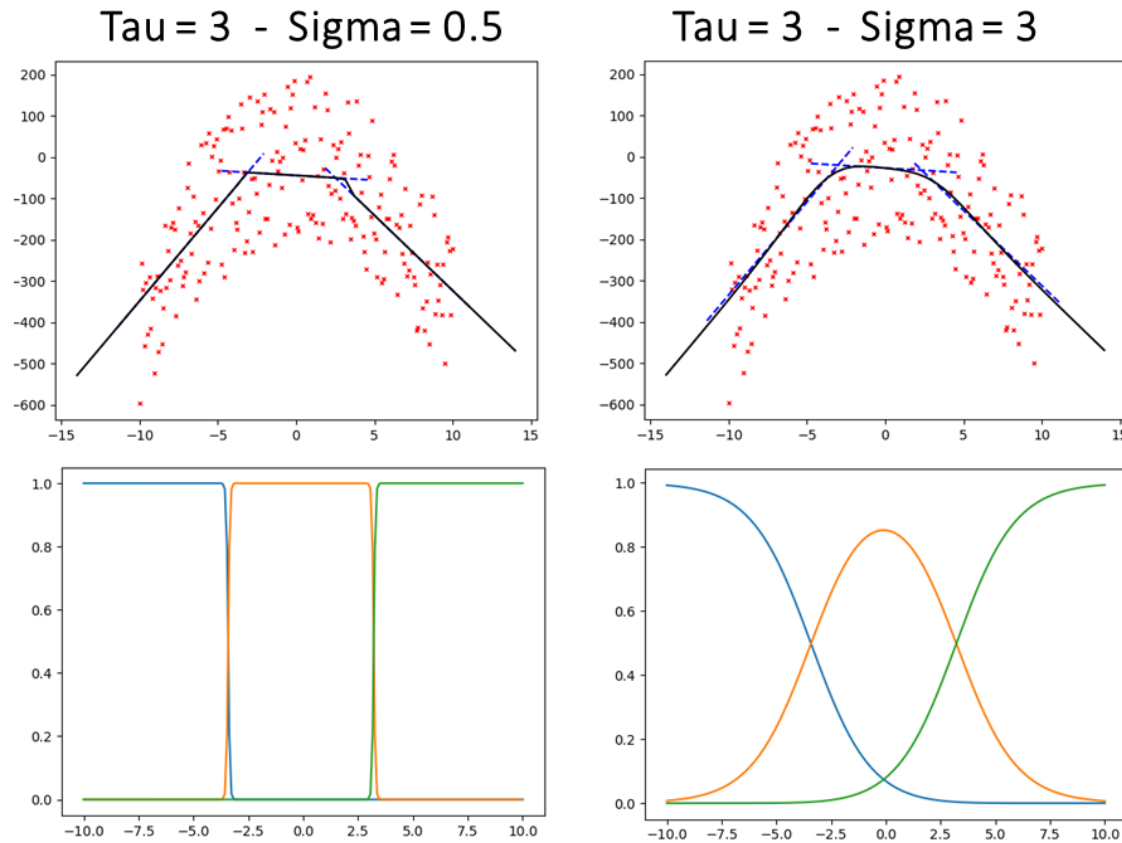
Hyperparameters:

- *amount sections*

- *tau*

- *sigma*

3. LWR in Detail - Influence of Tau

3. LWR in Detail - Influence of Sigma

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

11

4. Runtime Performance - Setup
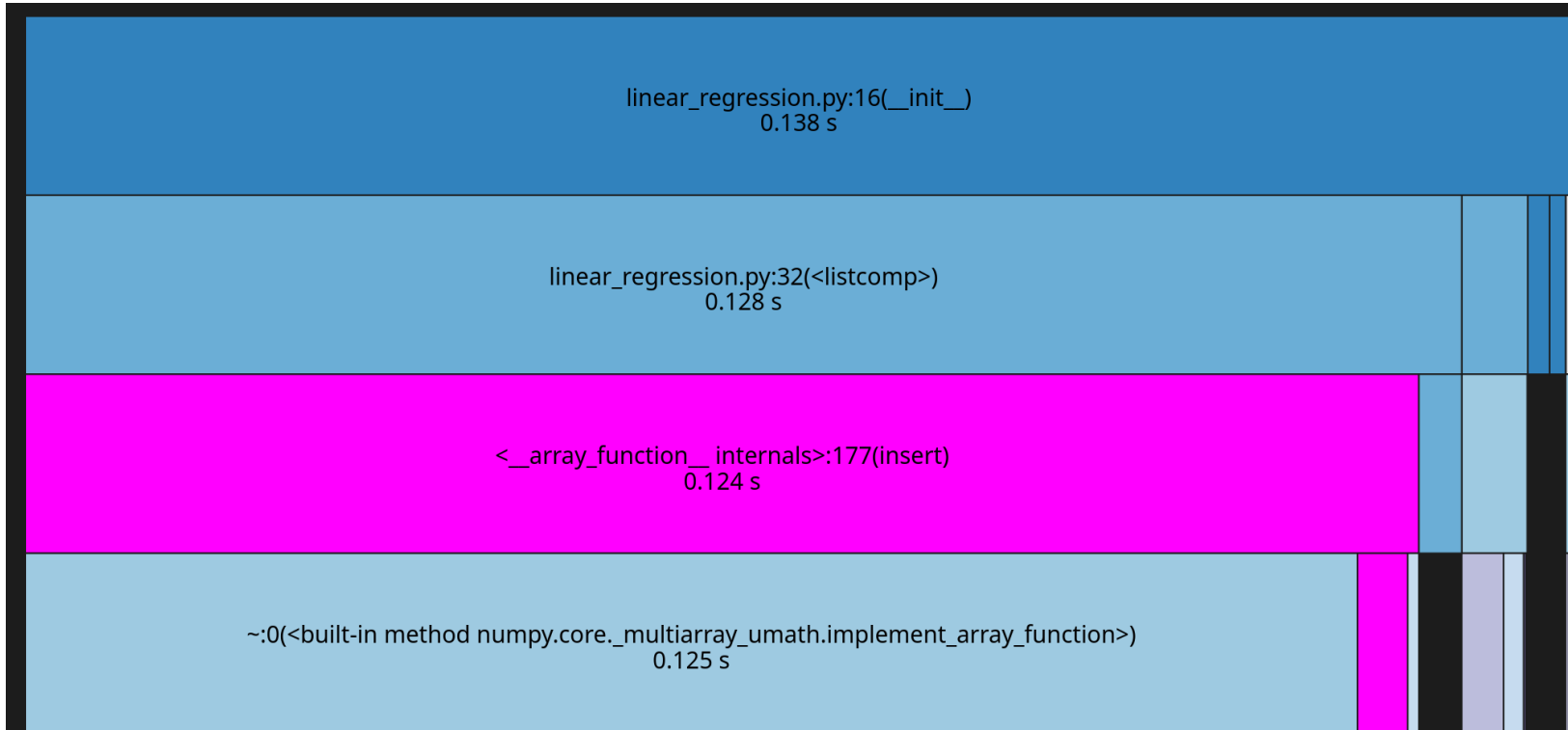
Time Complexity:

- 50 iterations

- Data:
    - 50 data points
    - 2 independent variables
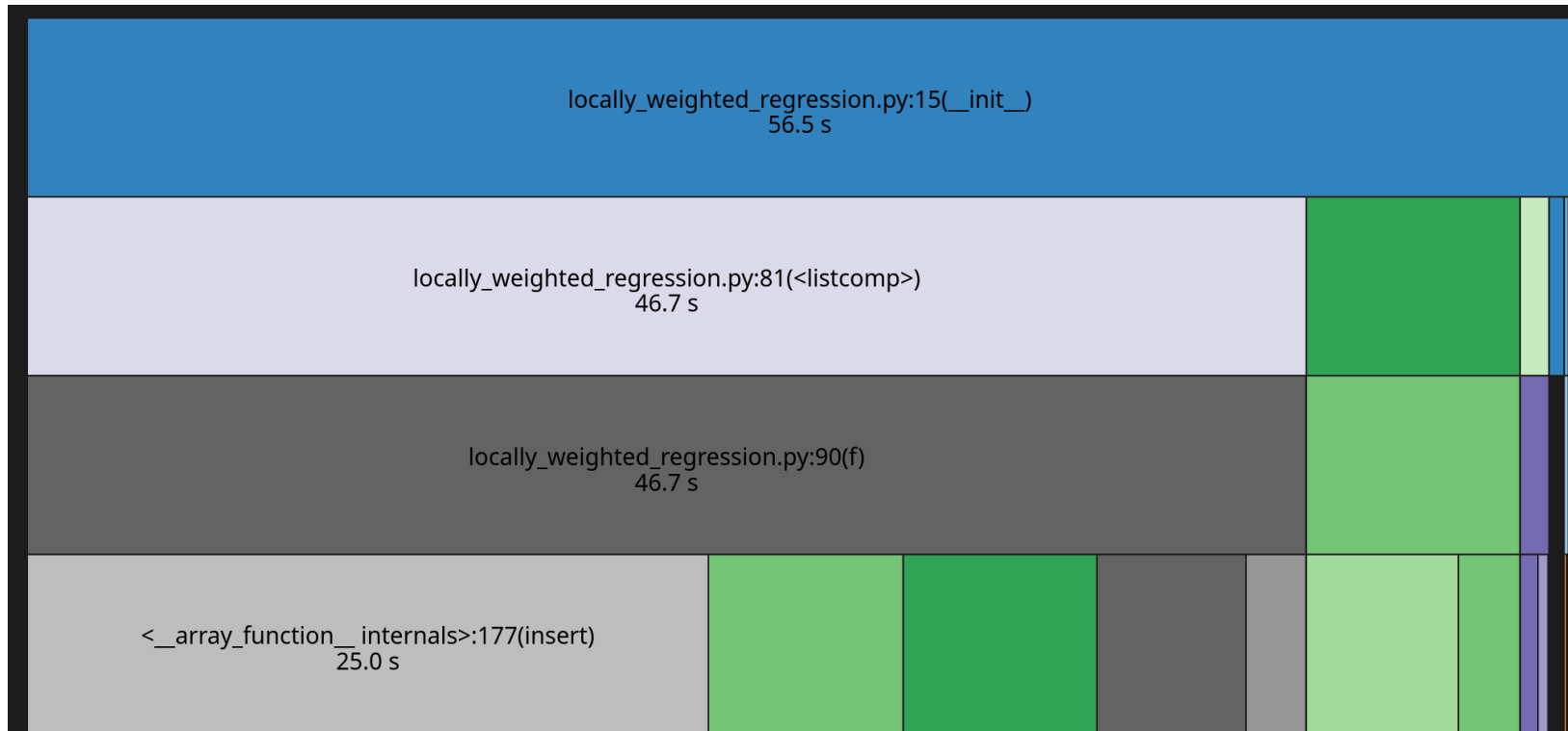
4. Runtime Performance - Results

	Runtime in seconds
LinearRegression	0.138
skearn.linear_model	0.060
LocallyWeightedRegression	56.500
localreg	2.170

4.1 LinearRegression

flame graph generated with `snakeviz`

4.2 LocallyWeightedRegression

flame graph generated with `snakeviz`

5. Live Demo

Pieer A. | Philipp R. | Jerome W. | Johannes T. | Tomislav P.

16