

COM6018 Data Science with Python

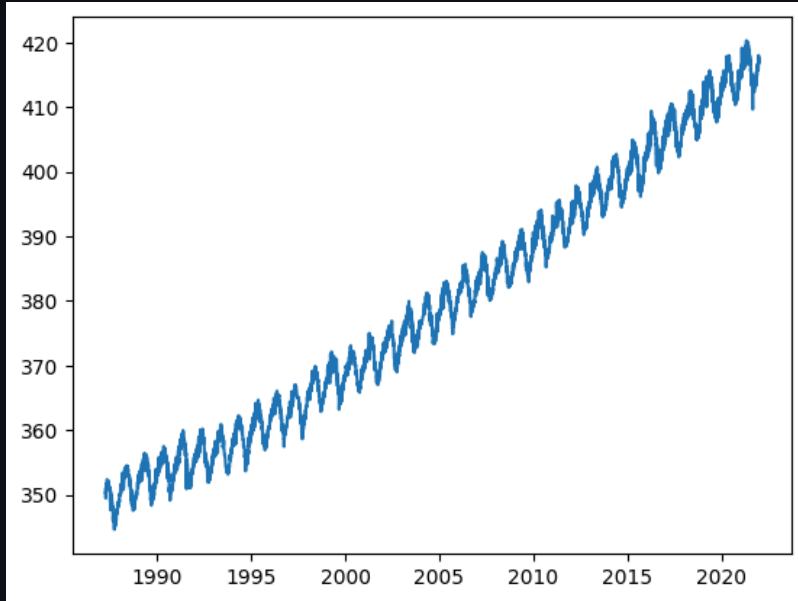
Lab 8: Curve Fitting with Scikit-Learn

Jon Barker

(c) 2023, 2024, Jon Barker, University of Sheffield

In this lab

We will use Scikit-Learn to fit the CO₂ atmospheric concentration data.



We will then use our fitting model to predict future CO₂ atmospheric concentration.

On what date will the CO₂ atmospheric concentration reach 450 ppm?

The Task

The stages of the task are as follows,

- Loading the CO₂ data.
- Fitting a polynomial curve to describe the growth trend.
- Fitting a periodic function to describe the seasonal variation.
- Tuning the model hyperparameters (order of polynomial, etc).
- Evaluating the model.
- Using the model to make a prediction.

Obtaining the Jupyter Notebook

If you have cloned and pulled the module's GitHub repository then you should see,



The lab is `080_curve_fitting_with_sckikit_learn.ipynb` and it will need the data file `data/co2.csv`.

Or you can download the notebook and data via links on Blackboard.

Getting Help

- If you are stuck just raise a hand to ask for help.
- Feel free to discuss the lab with your neighbours.
- Re-read the Scikit-Learn tutorial notes
 - In the Git repo at
`materials/tutorials/080_Curve_Fitting_with_Scikit_Learn.ipynb`
 - or online at <https://uos-com-6018.github.io/COM6018>
- Use the Scikit-Learn API documentation for reference. <https://scikit-learn.org/stable/modules/classes.html>