MiniProject1: Energy consumption prediction for Ro-Ro vessels from sensor data

The data on ship energy consumption were obtained from the Danish ro-ro passenger ship (MS Smyril). The MS Smyril has a length of 135 m, a ship width of 22.7 m, a design draft of 5.6 m, and four 3,360 kW main engines. The system collects data from the following sensor devices: the Doppler speed log, gyrocompass, Global Positioning System (GPS), main energy pipe flow meter, rudder angles, wind, propeller pitches, inclinometer, and level measurement device. The dataset was collected from February to April 2010, with a total of 246 voyages and 1,627,324 data records.
The goal is to predict the energy consumption of the ship. Your task is to choose the interval for calculating the energy consumption, which can be given by the rate of the fuel consumption. For example, on the daily basis can be calculated by:

EC=(fuelDensityxfuelVolumeFlowRatex3600x24)/1000 (tons/day)

A description of the recorded data is shown in the table below.

fuelDensity.csv	Fuel density	kg/l
fuelTemp.csv	Fuel temperature	C
fuelVolumeFlowRate.csv	Fuel volume flow rate	L/s
inclinometer-raw.csv	Inclinometer trim angle	degrees
latitude.csv	Latitude	
longitude.csv	Longitude	
level1median.csv	Port level measurements	m

level2median.csv	Starboard level measurements	m
longitudinalWaterSpeed.csv	Speed through water (STW)	kn
portPitch.csv	Port propeller pitch	-10 - 10 V
portRudder.csv	Port rudder angle	-10 - 10 V
speedKmh.csv	Speed over ground (SOG)	Km/h
speedKnots.csv	Speed over ground (SOG)	Kn
starboardPitch.csv	Starboard propeller pitch	-10 - 10 V
starboardRudder.csv	Starboard rudder angle	-10 - 10 V
trackDegreeMagnetic.csv	Track degree magnetic	degrees
trackDegreeTrue.csv	Track degree true	degrees
trueHeading.csv	True heading	degrees
windAngle.csv	Wind angle	degrees
windSpeed.csv	Wind speed	m/s

Task Description:

The task is to train **at least two** different machine learning algorithms to predict the energy consumption of the ship in terms of fuel consumption rate.
You will have to:
1- Perform data preprocessing

a. What features do you use for your models and why? What is the window interval you choose to predict the target variable?

b. How do you prepare the data to be used in your models?

c. Feature Engineering

2- Train the models with at least 85% accuracy.

3- Visualize the performance of your models.

4- Compare the results of different models, why is there a difference?

You must submit your Python code in a Jupyter notebook along with a written report. Write a scientific report that includes, but is not limited to, the following points:

- Introduction (what is the problem you are solving, why do you think it is relevant?)
- Data Processing (What are the data processing decisions you made and how did you implement them?)
- Modeling (What algorithms did you choose and why? How did you run them? How did you improve the performance of your models? Were you able to achieve the required accuracy in the first round of training?)
- Conclusion (What were the "scientific" bottlenecks? How did you overcome them? Which algorithm was better? Possible improvements? etc...)

Note: Please understand that the submitted report should include both text and figure explanations to cover the solution of the task. Only textual or graphical reports will not receive points. Also, the solution notebook should be completely independent of any local libraries or local installations in order to be reproducible by the peer reviewer.