Estimating Remaining Useful Life using variational autoencoders

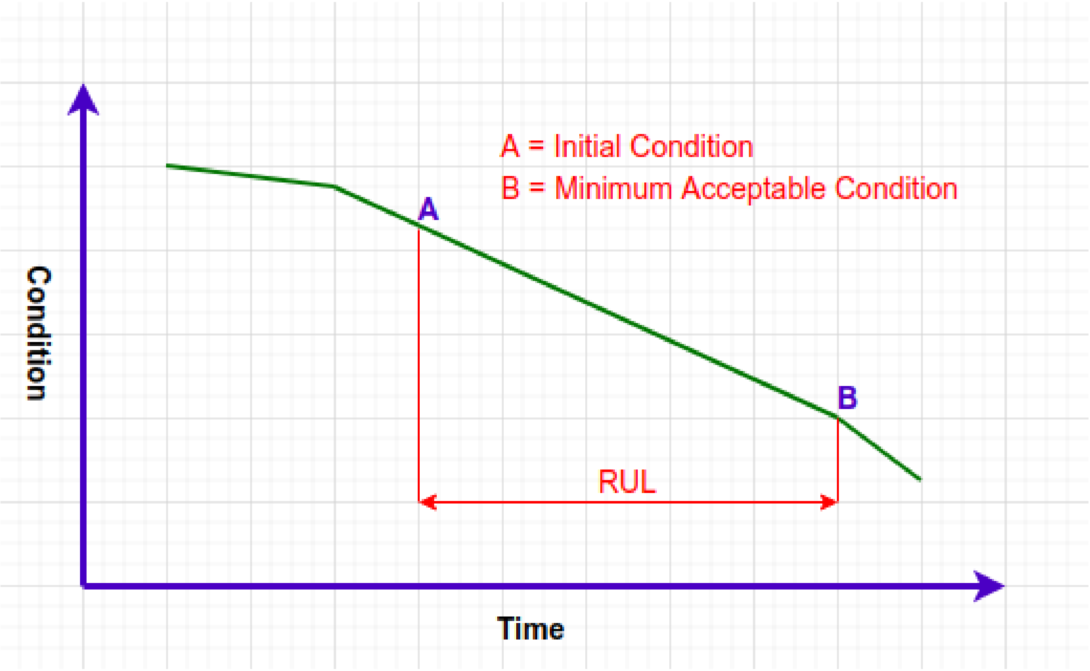
Industrial Process Monitoring

Shakiba Sharifi

Summer 2023

Introduction: RUL

- Remaining Useful Life (RUL) is a predictive maintenance metric that estimates the remaining operational lifespan of a machine or component based on its current condition and historical data.

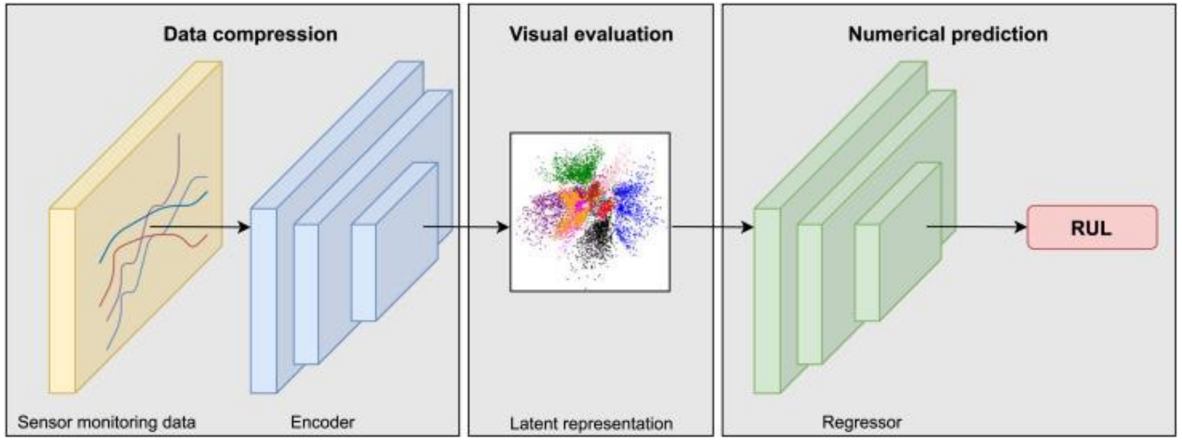


Problem Definition

- In the paper "Variational encoding approach for interpretable assessment of remaining useful life estimation" (Costa and Sánchez, 2022), authors addresses the problem of RUL estimation for aircraft engines. The authors propose a variational encoding approach to provide interpretable assessment of RUL.

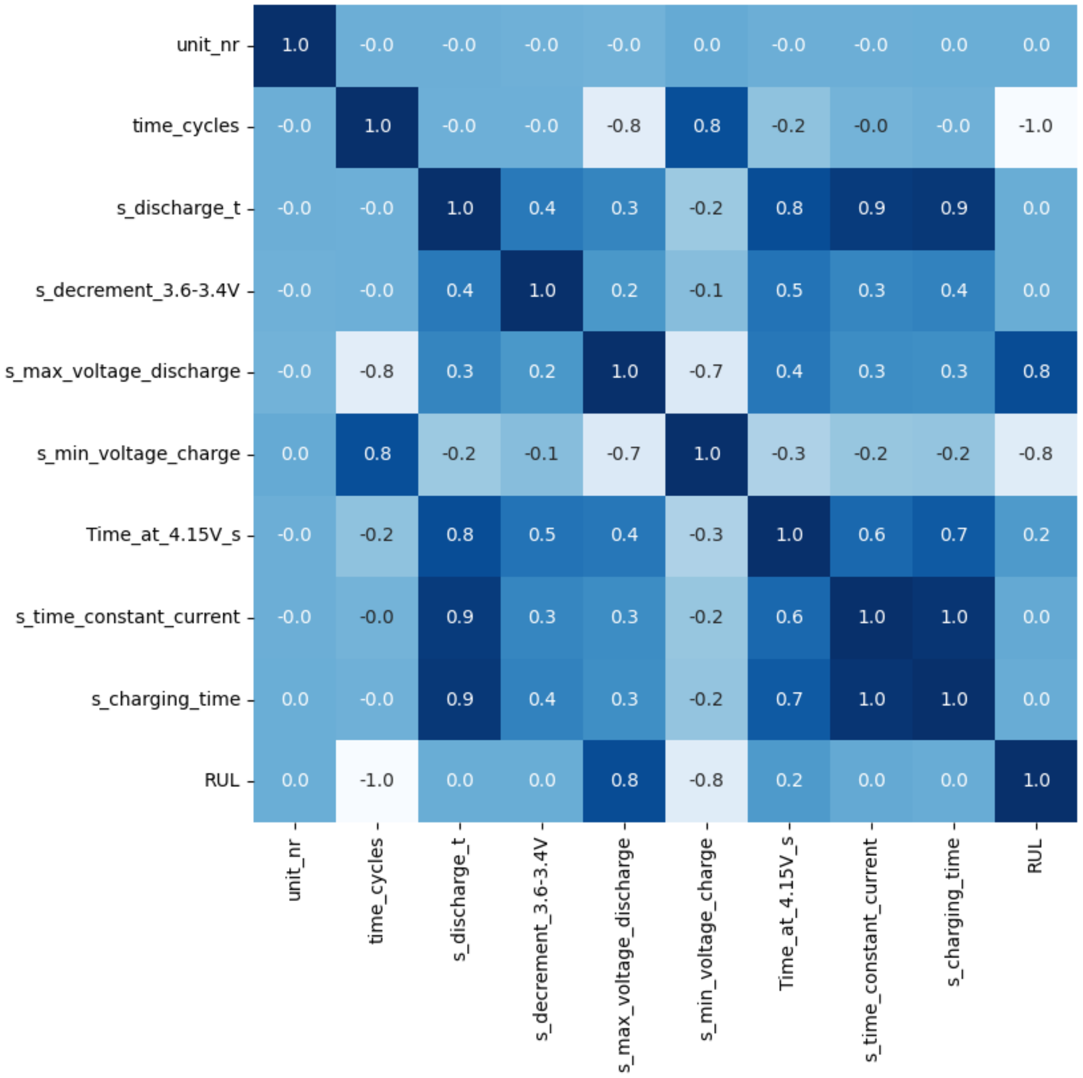
Method

- Workflow followed for the proposed approach: aircraft data is fed into the encoder, which learns a latent representation based on deterioration patterns to build a graphical map reflecting the evolution of their trajectories. The regressor directly influences obtaining such a latent space and allows reporting numerically the RUL of each engine.
- I use the same method for predicting RUL of a new dataset.



Applying the Method to New Dataset

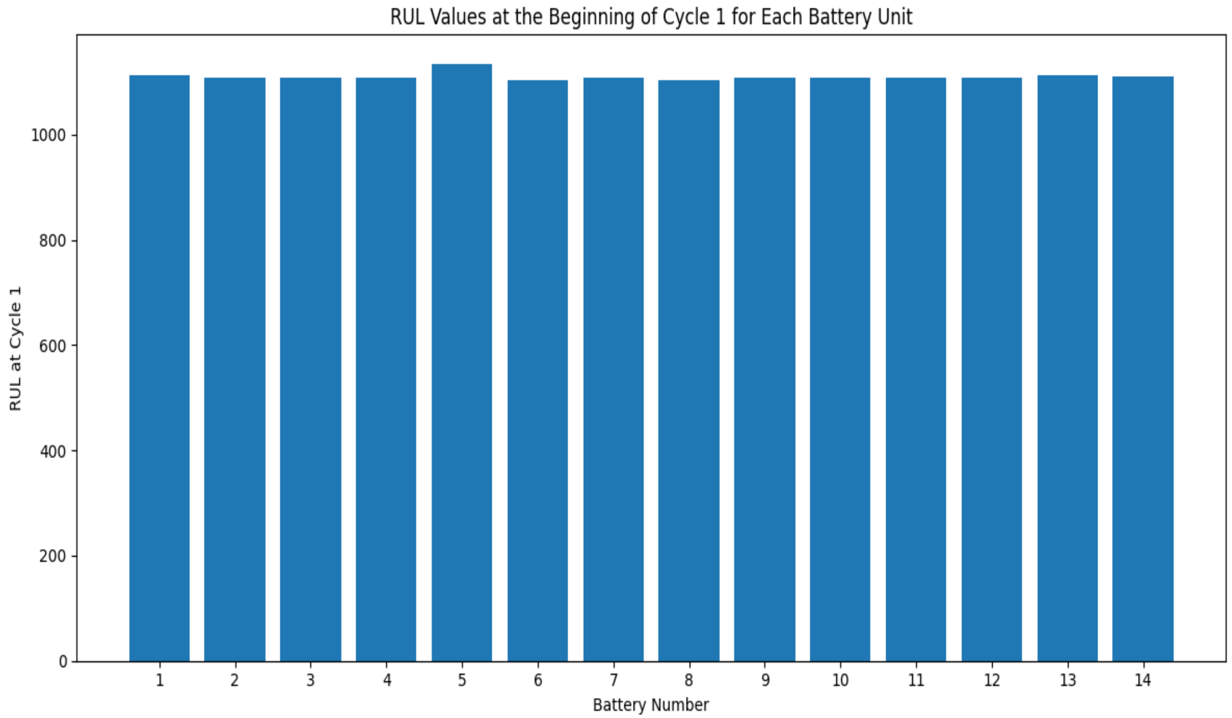
- The Hawaii Natural Energy Institute provides a battery cycling dataset which tracks voltage and current over time to model degradation and estimate RUL.
- They examined 14 batteries with a nominal capacity of 2.8 Ah, which were cycled over 1000 times.
- Similar to the original C-MAPSS aviation dataset which contains multivariate time series sensor data to model aircraft engine degradation for RUL prediction.



Data Pre-process

More data is not always useful:

- I Checked data corelation, keep correlated sensors with RUL and dropped others.



Data Pre-process

Balancing sequence of time series data for train and test:

- The sequence of the original data includes sycles around 100 to 140.
- New dataset has cycles more than 1000.
- So it is necesary to update sequence length in the model.

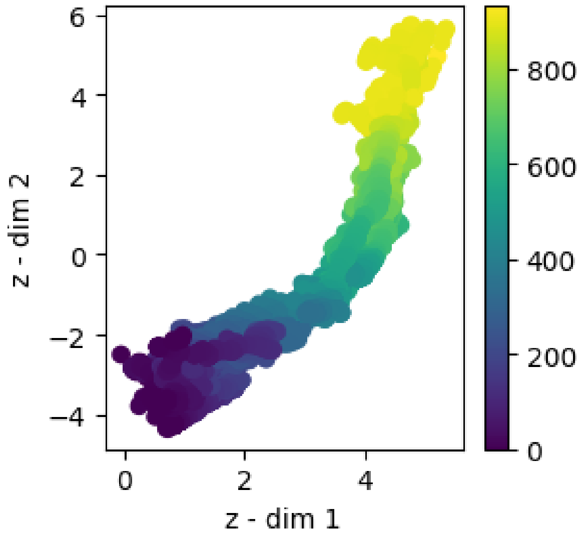
Model, Train, Validate

Apply some changes in the original model to better fit the new data:

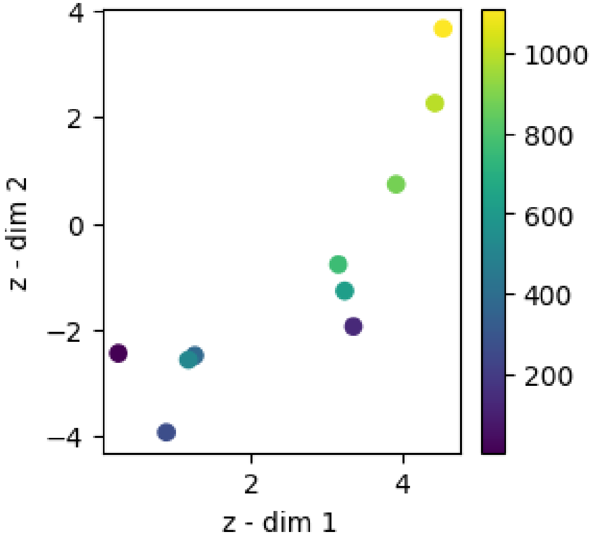
- Add Dropout and regularizations to prevent overfitting
- Decrese complexity of the model
- Try different model parameters
- Increase epochs

Visualize, and the results

- The pictures shows latent representation of data, which is learned by the encoder, based on deteriotion pattern in the sensors.
- This graphical map is then fed into a regressor which allows to report numerically the RUL of each battery.
- The color bar indicates the RUL value from 0 to the maximum possible number.



- Train data



- Test data

Discuss the results

- Comparing the RMSE and R2 metrics, the lower test accuracy compared to training implies overfitting due to insufficient training data diversity.

- Additional contributing factors may include inconsistencies in test data ground truth.

- Potential solutions include expanding training data, trying other advanced architectures, and investigating dataset mismatches. Further analysis is required to determine the core issues and develop robustness strategies to improve generalization across datasets.

- **train set:**
  - RMSE:26.65
  - R2:0.98

- **test set:**
  - RMSE:234.17
  - R2:0.56

References

- https://github.com/NahuelCostaCortez/Remaining-Useful-Life-Estimation-Variational/tree/main
- https://paperswithcode.com/paper/variational-encoding-approach-for
- https://www.kaggle.com/datasets/ignaciovinuales/battery-remaining-useful-life-rul