**SUMMARY OF ASSIGNMENT 3**

The model with GRU (32 units and 0.5 dropout) performs best in terms of MAE. Overfitting remains a challenge across all variations.

**Initial Model (GRU with 32 units and 0.5 dropout)**

* **MAE**: 2.40
* **Overfitting**: Starts after 6 epochs
* **Explanation**:

The initial model, trained from class, performs reasonably well with an MAE of 2.40. However, overfitting begins relatively early, which indicates that the model is learning the training data too well and may not generalize effectively to unseen data.

*I attempt the following techniques to improve the model.*

**Reduced Units in Recurrent Layer (GRU with 18 units)**

* **MAE:** 2.43
* **Overfitting:** Starts around 10 epochs
* **Explanation:**

By simplifying the model, we hoped to reduce overfitting. Unfortunately, the slight increase in MAE suggests that the original complexity (32 units) was beneficial, but overfitting remains an issue.

**Increased Units in Recurrent Layer (GRU with 50 units)**

* **MAE**: 2.47
* **Overfitting**: Starts around 2 epochs
* **Explanation**:

Increasing the units introduces more complexity. The early overfitting is a concern. The model may be too expressive for the available data.

**Switching from GRU to LSTM**

* **MAE:** 2.61
* **Overfitting**: Starts around 2 epochs
* **Explanation**:

LSTM (Long Short-Term Memory) cells have different dynamics than GRU (Gated Recurrent Unit) cells. The higher MAE suggests that the LSTM architecture may not be as suitable for this task. Overfitting behavior remains consistent.

**Adding Conv1D on LSTM Layer**

* **MAE**: 2.99
* **Overfitting**: Starts around the 7th epoch
* **Explanation**:

Introducing Conv1D layers aims to capture spatial features. However, the higher MAE and delayed overfitting suggest that this complexity did not improve performance significantly. The higher MAE and delayed overfitting indicate that this additional complexity did not improve performance significantly.

**Reducing Dropout (from 0.5 to 0.2)**

* **MAE**: 2.46
* **Overfitting**: Begins around 2 epochs
* **Explanation**:

It appears reducing the dropout impeded the model’s ability to generalize and instead increased overfitting.

**Increasing Dropout (from 0.5 to 0.7)**

* **MAE**: 2.42
* **Overfitting**: Begins around 10 epochs
* **Explanation**:

Higher dropout was expected to increase regularization. While MAE suggests better generalization, it does not improve on the initial model.

**Replacing RMSprop with Adam, Adjusting Learning Rate, and Increasing GRU Units (First Layer Only)**

**MAE**: 2.86

**Overfitting**: Starts at 2 epochs

**Explanation**:

Adam optimizer adapts learning rates per parameter. The increased GRU units was intended to improve expressiveness. However, the higher MAE and early overfitting indicate that this combination did not yield better results.