## *Conclusions for the Gas Price Prediction Model*

It was determined that the best CNN model has parameters of [8,8,32] with a learning rate of 0.0006, 4 layers and the Adam optimizer. The SGD and RMSprop optimizers have a higher loss than the Adam optimizer and get the same result with each other with little improvement. The model was then tested with minibatch and it led to a faster convergence to the loss value. The validation loss was 0.45 $USD^2$ and MAE was 0.64 USD. The baseline model's, a 9 degree linear regression, MAE was 0.34 USD which outperformed the deep learning model and is recommended for the dataset. As recurrent neural network(RNN) is good to be used on time series data, the gasoline price can be predicted using previous price information only. The validation loss was 0.0074 $USD^2$ and MAE for test data was 0.1 USD.

| **Model** | **Architecture** | **Features** | **Error [USD]** |
|---|---|---|---|
| Polynomial Regression | Degree 9 Polynomial | Year, Weekly U.S. Exports, Season | 0.34 |
| CNN | 4-Layer Neural Network | Year, Weekly U.S. Exports, Season | 0.64 |
| RNN | LSTM-based model with 10 lookback data | Gas Price | 0.10 |

The project faces notable technical challenges, primarily stemming from the constraint of a dataset limited to only three features. This constraint makes the model struggle to predict accurate results. Additionally, there's a risk of overfitting, where the model, despite achieving a low training loss, does not perform well on unseen data during testing. This discrepancy emphasizes the need to carefully manage the trade-off between the simplicity imposed by a limited feature set and the risk of overfitting. Addressing these challenges is crucial for improving the model's reliability and predictive capabilities.

Through our project, we have learned important lessons, especially regarding the suitability of Convolutional Neural Networks (CNNs) in specific situations. We found that CNN may not be the best option when working with datasets that have limited features or lack clear patterns. Additionally, the challenge of changing patterns over time suggests that relying on the entire dataset for learning might lead to less accurate predictions. Another crucial takeaway is the substantial impact of model selection on overall performance. In our project, comparing different models revealed that choices such as Recurrent Neural Network (RNN) and Linear Regression tend to yield more promising results than CNN. This emphasizes the critical role of carefully choosing the right model to achieve favorable outcomes.

Accurate predictions of gasoline prices bring significant advantages to a number of stakeholders. For consumers, this means better budget planning and the ability to decide when it's most cost-effective to refuel their vehicles. Businesses, especially those heavily reliant on transportation, benefit from improved efficiency and better financial planning due to more accurate cost projections. Investors in energy-related sectors gain the ability to make informed decisions, leading to better risk management by understanding how prices might fluctuate. At the government and industry levels, these predictions assist in economic planning, policy adjustments, and encourage strategic planning for energy companies, promoting the adoption of sustainable practices. Altogether, these benefits contribute to smarter decision-making across various sectors of the economy.

**Remark:**

- Github
  Gasoline-Price-Prediction: University of Illinois : IE434 Deep Dive 5 (github.com)
- Presentation
  Youtube