Group 10

Report:

- Well structured report with good use of visuals.
- Detailed explanation of the Spike LASSO model with a clear demonstration of deep technical understanding. The focus on addressing the challenge of price spikes with a complex model is a very original approach to the challenge.
- Insightful comparison of training and testing errors to demonstrate the risk of overfitting.
- Good understanding of energy market changes to inform the choice of time period.
- Good to see an attempt to relate energy market understanding to modelling decisions and results, though it's not clear how the OCGT generation would be used when projecting the frequency auction prices for the previous day – generation on the day of delivery would be a lagged indicator, whereas generation from the previous day may not be strongly correlated with the generation at the time of delivery.
- It would have been interesting to expand on the conclusions and future work to give more insight into what was learnt from the project. Some sections did go into extensive detail which limited the remaining space for conclusions (perhaps due to time constraints also).
- It would also be interesting to see which markets the models performed best and worst on, and to suggest reasons why this might be.
- Ultimately, the group ran out of time to produce results for the Spike LASSO model – it would be great to see what they could do given more time as this really differentiated them from other teams.

Model:

- Impressive and unique idea to use the Spike LASSO model to deal with the key issue of predicting price spikes. Demonstrates very deep technical knowledge of machine learning.
- Excellent attention to detail in data preparation, eg timezone adjustment, day-ahead prices.
- Granularity alignment was handled well, particularly the down-sampling by creating a range of aggregated metrics.
- Excellent targeted feature selection using a range of techniques to reduce redundant features.
- A reasonable choice to focus only on 2024 data onwards given the change to EAC, though it may have been good to see some experimenting on the 2023 data to justify this choice.
- Overall a very strong time-series methodology with consistent, causality-aware feature setup.

Code:

- Well organised code, with clear variable tracking and comments.
- Data preparation was excellent - clean joins, proper lagging and a good structure.
- Excellent use of pipeline structure with training, tuning and prediction handled well.