ENGINEERING
W Booth School of Engineering
Practice and Technology

McMaster University

Lab 3

Time-Series Forecasting with ML, LP and SP

Course Number and Name:	
SEP 6DA3: Data Analytics and Big Data	
Semester, Year, and Group Number:	
2025 Fall, Group 6	
Name of Students:	**Name of Instructor and TA:**
Andi Dong Zhiyu Hu Foram Brahmbhatt Jiarui Yang Linghe Shen Shannon Chen	Pedro Tondo Muskan Sidana

1. Objective

This project develops an energy analytics pipeline using PJM's hourly electricity consumption data to compare the performance of **Random Forest** and **Transformer models** for demand forecasting. GPU acceleration is used to increase the efficiency of Transformer training. The most accurate forecasts are used in a Linear Programming optimization to minimize generation costs while satisfying demand, with an extension to **Stochastic Programming** to address demand uncertainty between multiple scenarios.

2. Outcomes

- Random Forest & Transformer MSE
- Result 1

```
Random Forest MSE: 1785677.7071172565
Using device: cuda
Epoch 10, Loss: 4091091.4685
Epoch 20, Loss: 2345442.2697
Epoch 30, Loss: 2332622.3675
Transformer MSE: 1854248.9885078138
```

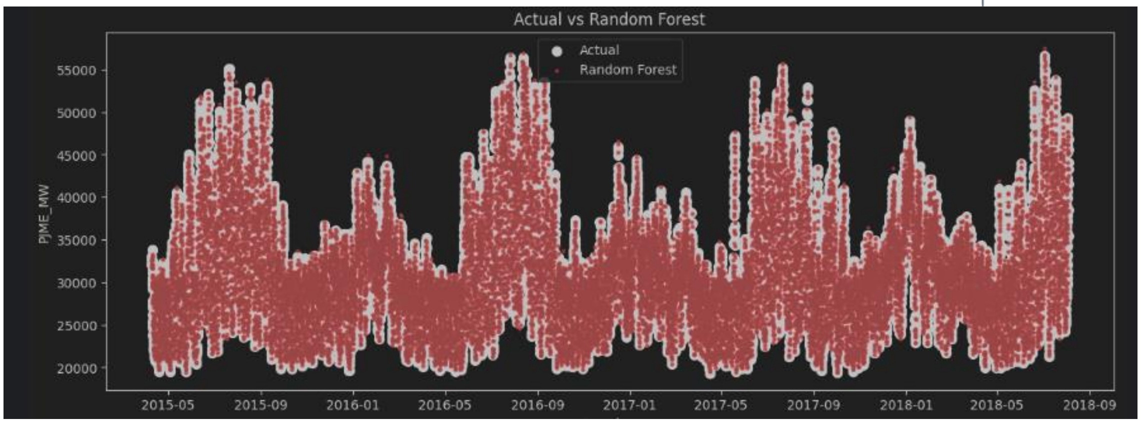
- Result 2

```
Random Forest MSE: 1777819.857003
```

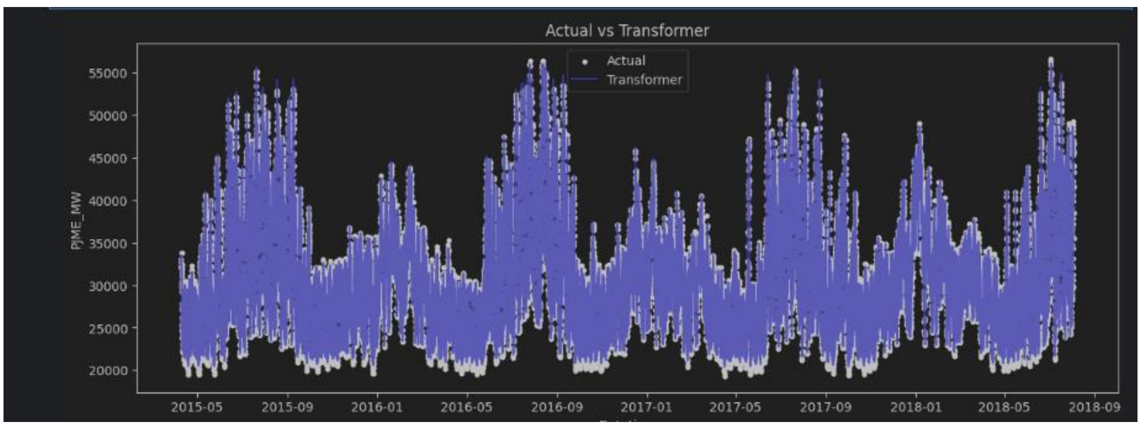
```
Epoch 10, Loss: 4447576.4659
Epoch 20, Loss: 2338986.7265
Epoch 30, Loss: 2237399.6327
Epoch 40, Loss: 2233246.9026
Epoch 50, Loss: 2217982.6863
Transformer MSE: 1911467.615969
```

- Result 3

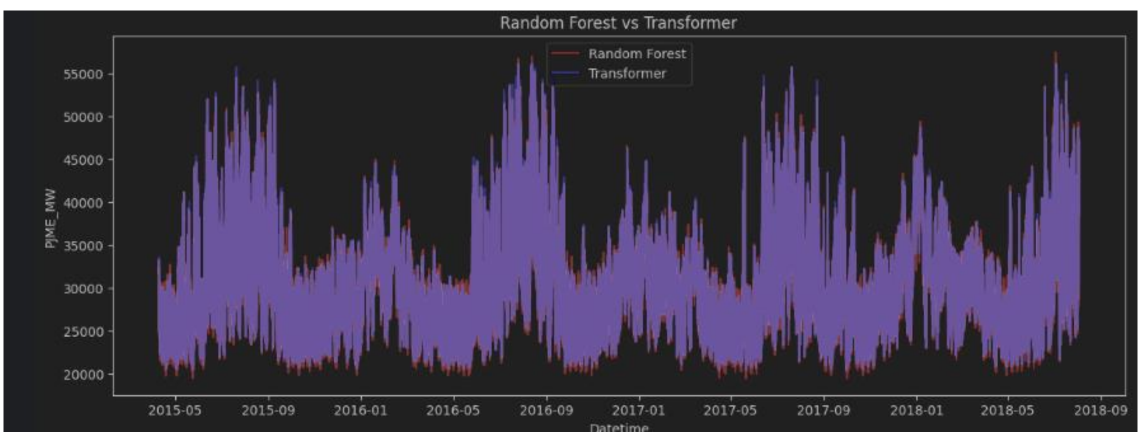
- Actual vs Random Forest



- Actual vs Transformer



- Random Forest vs Transformer



- Data analytics forecast

```
Random demand forecast (MW): [ 73.4 159.  142.   60.4 182.1 104.4  64.  140.2 174.  154.2 119.1 148.5]

Optimal Generator Schedule (LP):
Gen1: 100.00 MW
Gen2: 82.09 MW
Gen3: 0.00 MW

Expected costs across 3 SP scenarios: [10656.26 10980.24 10926.47]
```

3. Questions

1. Which model (RF or Transformer) achieved lower MSE?
   Based on the experimental results(Result 1), the **Random** Forest model achieved a lower MSE of 1785677.71 compared to the Transformer model's MSE of 1854248.99. The MSE quantifies the average squared difference between predicted and actual values; therefore, the lower MSE of the Random Forest model indicates that, for this dataset and model configuration, it shows greater predictive accuracy than the Transformer model. While both models captured the general trends in the time series data, the Random Forest model's predictions, on average, were closer to the actual power consumption values.

2. How does generator capacity influence the LP solution?
   The LP solver minimizes cost by prioritizing the cheapest generators and operating them up to their capacity limits. Generator capacities serve as hard constraints, ensuring no unit produces more than its maximum. When demand exceeds the total cheap capacity, the solver dispatches more expensive units to meet the remaining load. Conversely, if total demand is lower than the combined capacities, not all generators will be dispatched. Therefore, generator capacities directly shape the generation mix and determine the final system cost.

3. How do the three SP scenarios change expected costs?
   The three SP scenarios change expected costs by reflecting how demand uncertainty alters generation scheduling. In our results, the total costs were **10,656.26**, **10,980.24**, and **10,926.47** under different demand scenarios. Higher demand scenarios required more use of expensive generators, leading to higher costs, while lower demand scenarios relied more on cheaper units, keeping costs down. The variation among these three outcomes illustrates the financial risk of uncertain demand. The expected cost, computed as the average across scenarios, was about 10,854.32, providing a more balanced planning figure that incorporates uncertainty into decision-making.

4. Based on your results, which approach (RF or Transformer) would you recommend for operational use, and why?

Based on the results, I would recommend the Random Forest (RF) model for operational use. Although the Transformer is a more advanced deep learning architecture, in our experiment the RF model achieved a lower MSE (e.g.**1,785,677.71** vs. **1,854,248.99**), indicating better predictive accuracy for this dataset. In addition, RF is less computationally demanding, easier to train, and more interpretable, making it more practical for near-real-time forecasting in operational environments. While the Transformer may have advantages for longer-term scaling or larger datasets with richer temporal dependencies, for this task the RF provides a more reliable and efficient solution.

5. Explain why the solver chose to run Gen1 at full capacity, Gen2 at partial capacity, and Gen3 at zero. (In your answer, refer to both the cost per MW and the capacity constraints of each generator.)
   The solver's decision reflects the joint influence of **generation cost per MW** and **capacity constraints**. Generator 1 was dispatched at full capacity because it offers the lowest cost per unit of electricity. By utilizing it to its maximum limit, the model minimizes total system cost. Once Generator 1's capacity constraint was reached, additional demand had to be met by Generator 2, which has a higher marginal cost, and therefore it was operated at a partial level sufficient to cover the remaining demand. Generator 3, despite having available capacity, was not dispatched at all because its cost per MW is the highest among the three units. Allocating generation to Gen3 would have increased the overall cost, so the solver excluded it entirely. This outcome illustrates how the LP model prioritizes cheaper resources up to their capacity limits before turning to more expensive alternatives.

6. Assume your SP scenarios produced total costs of $12,200, $11,915 and $11,029.
   (a) What does this tell you about the effect of demand uncertainty on generation cost?
   The variation in total costs across the three SP scenarios demonstrates that demand uncertainty directly affects generation costs. Higher demand scenarios necessitate the dispatch of more expensive generators, thereby increasing total cost, whereas lower demand scenarios can rely primarily on cheaper units, resulting in lower costs. This variability highlights the financial risk associated with uncertain forecasts and shows that inaccurate demand predictions can significantly impact operational expenses.

   (b) If you were a grid operator, what practical action could you take to reduce cost variability across these scenarios?
   As a grid operator, one practical action to reduce cost variability would be to invest in d**emand-side management** and **energy storage systems**. Demand-side management can smooth peak loads through demand response programs, while energy storage allows surplus electricity generated during low-cost periods to be used during high-demand periods, reducing reliance on expensive units.