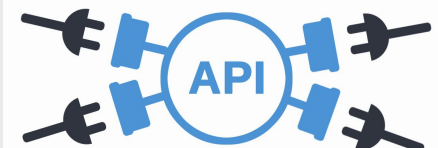


## University of Information Technology, Vietnam

- Choosing the optimal Spark configuration is challenging and directly impacts processing time and resource costs.
- Currently, this process relies on data engineers' experience or traditional heuristic-based methods, which have limitations.
- Using LLM automates the process, reduces resource waste, optimizes performance, and improves the accuracy of Spark configuration recommendations.

## Update the system



- Apply the recommended Spark configurations to improve performance and resource efficiency.
- Monitor changes in execution speed, cost reduction, and overall system stability.
- Continuously refine the model using feedback from real-world performance data.

