

## Senior Project Review (Group 39)

**LMAO , Language model for anomalies detection in monitoring system O<sup>2</sup> of Alice facilities**

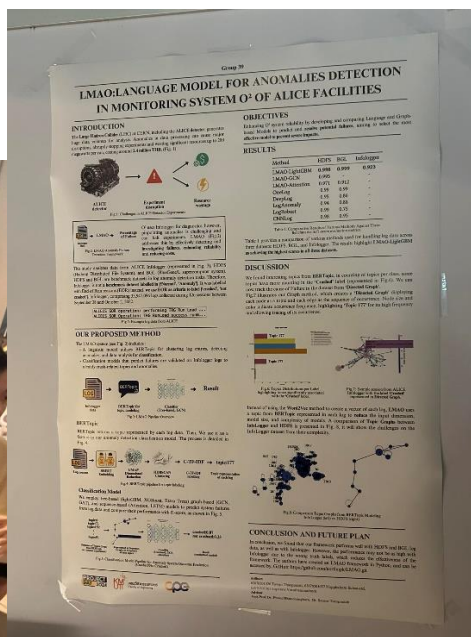
**Member:** Mr.Tawan Thaeprasit 66070501026,

**Advisor:** Asst.Prof.Dr.Phond Phunchongharn,

Mr.Nopphakorn Subsa-ard 66070501037,

Dr.Sansiri Tarnpradab

Mr.Supasate Vorathamthorn 66070501063



เนื่องจากสถาบันวิจัยนิวเคลียร์แห่งชาติ CERN นั้นมีการทดลองการเร่งอนุภาค โดยใช้เครื่องชนอนุภาคแฮดรอนขนาดใหญ่ และ ตรวจจับการทำงานของเครื่องด้วย Alice detector ที่เมื่อนำมาประมวลผลผ่าน Super Computer แล้ว จึงได้ไฟล์ Log ที่แสดง Action ทั้งหมดของตัวเครื่องอนุภาคแฮดรอนขนาดใหญ่

การพัฒนาโมเดลสำหรับวิเคราะห์ข้อมูล Action Log ที่ได้จาก Alice detector เพื่อตรวจหาความผิดปกติ หรือแนวโน้มความล้มเหลวที่อาจจะเกิดขึ้นได้ โดยมีเป้าหมายสุดท้าย คือการเลือก Machine learning model ที่ทำให้การวิเคราะห์นั้นมีประสิทธิภาพมากที่สุด ทั้งนี้ก็เพื่อการป้องกันการสูญเสีย Resource ไปอย่างไม่เกิดประโยชน์

ในการพัฒนา เนื่องจากข้อมูล Action Log ที่ได้จากการทดลองแต่ละครั้ง นั้นมีจำนวน และความซับซ้อนของข้อมูลที่ยาวนาน จึงต้องใช้เทคนิค topic modeling อย่าง BERTopic มาเพื่อทำการ ลดความซับซ้อนของข้อมูล, Clustering และ Labelling กลุ่มของ Cluster จากนั้นก็นำ Topic ที่ได้จากการทำ BERTopic มาเข้ากระบวนการ Classification ด้วย Machine learning model ซึ่งนั่นคือ tree-based , graph-based และ sequence-based จากนั้นก็ทำการเทียบผลลัพธ์ เพื่อให้ได้โมเดลที่มีประสิทธิภาพมากที่สุด

สุดท้ายแล้วทำให้ได้ผลลัพธ์ออกมาดังนี้ HDFS และ BGL ของ model ทั้งสามตัวของผู้พัฒนานั้นมีความแม่นยำกว่า model อื่นๆที่เคยมีอยู่ และโมเดลที่แม่นยำมากที่สุดคือ LMAO-LightBGM ประเภท tree-based

โดย ภูรินทร์ จินพวด 660070501043