การประเมินความปลอดภัยแบบอัตโนมัติด้วย BERT สำหรับการปฏิบัติตามข้อกำหนดในสถานที่ทำงานทางอุตสาหกรรม.

นายภูวดล นุ่นภักดี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมอัตโนมัติ
บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2567

การประเมินความปลอดภัยแบบอัตโนมัติด้วย BERT สำหรับการปฏิบัติตามข้อกำหนดในสถานที่ทำงานทางอุตสาหกรรม.

นายภูวดล นุ่นภักดี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตร
วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมอัตโนมัติ
บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2567

ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การประเมินความปลอดภัยแบบอัตโนมัติด้วย BERT สำหรับการปฏิบัติตามข้อกำหนดในสถานที่ทำงานทางอุตสาหกรรม.

โดย นายภูวดล นุ่นภักดี

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมอัตโนมัติ

คณบดีบัณฑิตวิทยาลัย / หัวหน้าภาควิชา

คณะกรรมการสอบวิทยานิพนธ์

..... อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.จิรพันธ์ุ อินเทียม)

..... อาจารย์ที่ปรึกษาร่วม

(อาจารย์ ดร.เทพภากร สิทธิวันชัย)

ชื่อ : นายภูวดล นุ่นภักดี
ชื่อวิทยานิพนธ์ : การประเมินความปลอดภัยแบบอัตโนมัติด้วย BERT สำหรับการปฏิบัติตามข้อกำหนดในสถานที่ทำงานทางอุตสาหกรรม.
สาขาวิชา : วิศวกรรมอัตโนมัติ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : ผู้ช่วยศาสตราจารย์ ดร.จิรพันธุ์ อินเทียม
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม : อาจารย์ ดร.เทพภากร สิทธิวันชัย
ปีการศึกษา : 2567

บทคัดย่อ

การศึกษานี้เน้นการเข้าใจขอบเขตของวิกฤตการจัดการความปลอดภัยในกระบวนการ (PSM) โดยเน้นความสำคัญของความปลอดภัยทั้งของบุคคลและกระบวนการในการลดความเสี่ยงจากอันตราย เช่น ไฟไหม้ การระเบิด และเหตุการณ์ทางเคมีภายในโรงงานเคมีและที่เกี่ยวข้อง การปรับปรุงการประเมินความปลอดภัยเพื่อลดความเสี่ยงแก่ผู้ปฏิบัติงานและสถานที่ทำงานรวมถึงสิ่งแวดล้อมเป็นจุดเน้นหลัก โดยปกติ JSA ถูกใช้ในการประเมินความเสี่ยงและมักใช้มนุษย์ในกระบวนการ แต่อาจมีความไม่แม่นยำและเสี่ยงในการประเมิน ซึ่งอาจส่งผลต่ออุตสาหกรรมในระยะยาว เช่น โรงกลั่นน้ำมัน หรือ โรงงานปิโตรเคมี การใช้โมเดล BERT เพื่อประมวลผลภาษาธรรมชาติในการวิเคราะห์ JSA อัตโนมัติมีประสิทธิภาพสูง ซึ่งช่วยลดการเกิดอุบัติเหตุในอุตสาหกรรมโดยเฉพาะจากการบ่งชี้อันตรายที่ไม่เพียงพอและข้อบกพร่องในการประเมินความเสี่ยง การเน้นความสำคัญของการปฏิบัติ JSA ที่เข้มงวดและการประเมินความปลอดภัยอย่างละเอียดในอุตสาหกรรมเพื่อลดความเสี่ยงแก่พนักงานและสิ่งแวดล้อมเป็นจุดสำคัญในการจัดการความปลอดภัยในอุตสาหกรรม การตรวจสอบโดยอัตโนมัติโดยผู้เชี่ยวชาญก่อนอนุญาตให้การปฏิบัติงานต่อไปเป็นขั้นตอนสำคัญในการรักษาความปลอดภัยและป้องกันอุบัติเหตุในอุตสาหกรรมในอนาคต การละเลยในการประเมินความเสี่ยงและข้อบกพร่องในการประเมินอาจนำไปสู่ผลลัพธ์ที่มีผลกระทบอย่างรุนแรงต่อพนักงานและสิ่งแวดล้อม

(มีจำนวนทั้งสิ้น 60 หน้า)

คำสำคัญ : JSA-BERT, BERT,JSA

_____ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Name : Mr. Phuwadon Nunphakdee
Thesis Title : BERT-Based Automation of Job Safety Analysis for industrial workplace compliance.
Major Field : Automation Engineering
King Mongkut's University of Technology North Bangkok
Thesis Advisor : Jiraphan Inthiam
Co-Advisor : Teppakorn Sittiwanchai
Academic Year : 2024

ABSTRACT

Our center focuses on understanding the critical role of Process Safety Management (PSM), emphasizing the safety of both personnel and control systems through software that detects hazards such as fires and unauthorized movements within chemical plants. As a leader in safety control for the supervision of workers and the environment, we emphasize fundamental control principles and the application of Job Safety Analysis (JSA) to manage human factors in safety. We address risks in key industrial operations that can be audited, such as inspections of oil and petrochemical plants. By integrating a BERT-based module, we provide natural language descriptions for automated JSA analysis, determining specific audits from key indicators. This approach highlights potential hazards identified through data analysis and offers insights into control systems, emphasizing JSA practices and safety information to keep employees and critical control systems informed. Ensuring safety in performing duties involves expert inspections prior to further investigation, which is an important step in identifying and preventing accidents. Addressing inspection omissions and deficiencies is crucial, as their consequences can have a lasting impact on employees and the environment.

(Total 60 Pages)

Keywords: JSA-BERT, BERT-JSA

_____ Advisor

กิตติกรรมประกาศ

ผมขอแสดงความขอบคุณจากใจจริงต่อทุกคนที่มีส่วนทำให้การศึกษาวิจัยครั้งนี้สำเร็จ ก่อนอื่น ผมขอแสดงความขอบคุณอย่างสุดซึ้งต่อ ดร.จิรพันธุ์ อินเทียม หัวหน้างานของผม และ ดร.เทพภากร สิทธิวันชัย อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม สำหรับคำแนะนำอันล้ำค่า การสนับสนุนอย่างไม่เปลี่ยนแปลง และกำลังใจตลอดกระบวนการวิจัยทั้งหมด ความเชี่ยวชาญ ข้อมูลเชิงลึก และผลตอบรับเชิงสร้างสรรค์ของพวกเขามีส่วนสำคัญในการกำหนดรูปแบบงานนี้ ผมขอขอบคุณคณาจารย์และเจ้าหน้าที่ของภาควิชาวิศวกรรมเครื่องมือวัดและอิเล็กทรอนิกส์ ซึ่งมีทรัพยากรและสิ่งอำนวยความสะดวกที่สร้างสภาพแวดล้อมที่ดีเยี่ยมสำหรับการวิจัยและการเรียนรู้ สุดท้าย ผมขอแสดงความขอบคุณอย่างสุดซึ้งต่อครอบครัวและเพื่อนๆ ของผมสำหรับความรัก กำลังใจ และความเข้าใจอันไม่เปลี่ยนแปลงตลอดการเดินทางครั้งนี้ ความอดทนและการสนับสนุนของพวกเขาเป็นบ่อเกิดของความเข้มแข็งและแรงจูงใจ ขอขอบคุณทุกคนที่มีส่วนร่วมในความพยายามนี้ การมีส่วนร่วมของคุณได้รับการชื่นชมอย่างสุดซึ้ง

ภูวดล นุ่นภักดี

สารบัญ

หน้า

สารบัญตาราง

สารบัญรูปภาพ

ประมวลศัพท์และคำย่อ

PSM	=	Process Safety Management
JSA	=	Job Safety Analysis
BERT	=	Bidirectional Encoder Representations from Transformers
SBERT	=	Siamese BERT
SAP	=	Systems, Applications & Products in Data Processing
LOPC	=	Loss of Primary Containment
LLM	=	Large Language Model

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของงานวิจัย

การจัดการความปลอดภัยของกระบวนการ (PSM) ถือเป็นสิ่งสำคัญในการรับรองความปลอดภัยของโรงงานเคมีและโรงงานที่เกี่ยวข้อง โดยการบรรเทาอันตราย เช่น ไฟไหม้ การระเบิด และเหตุการณ์ทางเคมี บทนำนี้เน้นย้ำถึงการบูรณาการความปลอดภัยส่วนบุคคลและกระบวนการภายในสภาพแวดล้อมทางอุตสาหกรรม โดยเน้นความสำคัญของการระบุอันตราย การประเมินความเสี่ยง และการนำมาตรการด้านความปลอดภัยไปใช้ โดยจะกล่าวถึงบทบาทของเทคโนโลยีขั้นสูง เช่น Transformer Model และโมเดล เช่น BERT ในการประมวลผลภาษาธรรมชาติ (NLP) ในการประเมินความปลอดภัยโดยอัตโนมัติและปรับปรุงกลยุทธ์การลดความเสี่ยง นอกจากนี้ ยังให้ความสำคัญของการวิเคราะห์ความปลอดภัยของงาน (JSA) และผลที่ตามมาของการละเลยการประเมินความปลอดภัยอย่างละเอียดผ่านกรณีศึกษาในโลกแห่งความเป็นจริง โดยรวมแล้ว การแนะนำนี้เป็นการปูทางสำหรับการสำรวจแง่มุมที่สำคัญของการจัดการความปลอดภัยของกระบวนการและกลยุทธ์ในการส่งเสริมความปลอดภัยในการปฏิบัติการทางอุตสาหกรรม

ในความพยายามในการทำงานวิจัยของเรา ได้ใช้ข้อดีของโมเดลประมวลผลทางภาษา ได้แก่ BERT และ SBERT เพื่อตรวจสอบมาตรฐานความปลอดภัยที่แพร่หลายในอุตสาหกรรมน้ำมันและก๊าซ อย่างถูกต้อง วัตถุประสงค์หลักของเราคือการประมาณเปอร์เซ็นต์ของความปลอดภัยที่ทำได้ โดยใช้วิธีการทางคณิตศาสตร์ที่เมื่อเปรียบเทียบกับระดับความสามารถประเมินจากทั้งผู้เชี่ยวชาญในอุตสาหกรรมและพนักงาน ยิ่งไปกว่านั้น เราพยายามที่จะก้าวข้ามเกณฑ์ความปลอดภัยที่กำหนดไว้ล่วงหน้าซึ่งกำหนดโดยผู้เชี่ยวชาญในเครือบริษัทโรงกลั่นที่มีชื่อเสียง การประเมินของเราดำเนินการผ่าน Google Colab ซึ่งเป็นแพลตฟอร์มที่ได้รับเลือกเนื่องจากความสามารถในการจำลองขั้นตอนการประเมินของ SAP เป็นที่น่าสังเกตว่าข้อจำกัดโดยธรรมชาติของการพึ่งพาการประเมินโดยมนุษย์ภายใน SAP เพียงอย่างเดียว ซึ่งอาจทำให้เกิดข้อผิดพลาดและความไม่สอดคล้องกันในกระบวนการประเมิน ด้วยการใช้ประโยชน์จากวิธีการอัตโนมัติ เราพยายามที่จะบรรเทาความท้าทายเหล่านี้ และรับประกันการประเมินมาตรฐานความปลอดภัยที่ครอบคลุมและแม่นยำยิ่งขึ้น

1.2 วัตถุประสงค์

1.2.1 ออกแบบชุดข้อมูล (Dataset) โดยอ้างอิงจากข้อมูลที่ได้จาก SAP เพื่อใช้ในการวิเคราะห์ JSA (Job Safety Analysis)

1.2.2 ทำนายผลการจัดประเภท (Classification) ของระดับความเสี่ยงในแต่ละลักษณะงานที่เกี่ยวข้องกับ JSA

1.2.3 ประเมินความเสี่ยงอย่างครอบคลุมในทุกกิจกรรมที่ต้องใช้ JSA เช่น อุปกรณ์เครื่องมือวัดในโรงงานเพื่อเพิ่มความปลอดภัยในการทำงาน

1.2.4 ปรับปรุงกระบวนการวิเคราะห์ JSA ให้มีความแม่นยำและมีประสิทธิภาพมากขึ้น

1.2.5 เพิ่มความปลอดภัยและความน่าเชื่อถือของกระบวนการทำงานในโรงงาน โดยใช้ข้อมูลจาก JSA ในการลดความเสี่ยงที่อาจเกิดขึ้น

1.3 ขอบเขตของการวิจัย

ในวิทยานิพนธ์นี้ผู้เขียนจะจำลองแบบการประเมินความเสี่ยงซึ่งสามารถปรับปรุงจากการประมาณการแบบเดิมๆ ได้ด้วยการประเมินของมนุษย์เท่านั้น วิธีนี้ไม่เพียงพอ ดังนั้นการประมวลผลทางภาษาจะครอบคลุมมากกว่าการประเมินผลแบบเดิมๆ

1.4 วิธีการวิจัย

1.4.1 ใช้ภาษาไพทอนในการเขียน

1.4.2 การคำนวณของประแต่ละประโยคจะถูกคำนวณผ่าน Google Colab

1.4.3 ใช้ BERT ในการทำ Sentence Embeddings.

1.4.4 ใช้ SBERT ในการหา Similarity ของผู้เชี่ยวชาญ และผู้ปฏิบัติงาน

1.4.5 ใช้ Google Colab ในการทำ Simulator ระหว่างผู้เชี่ยวชาญและผู้ปฏิบัติงาน

1.5 ประโยชน์ของการวิจัย

1.5.1 เรียนรู้การใช้ภาษาไพทอน ในการเขียน

1.5.2 เรียนรู้การคำนวณจากการประมวลผลทางธรรมชาติ

1.5.3 เรียนรู้การใช้ BERT โมเดลจาก Huggingface Community

1.5.4 ช่วยให้การประเมินความเสี่ยงในทางอุตสาหกรรมมีความปลอดภัยมากยิ่งขึ้น

1.5.5 ช่วยให้ผู้ปฏิบัติงานสามารถประเมินความเสี่ยงตามงานที่ได้รับมอบหมายได้อย่างถูกต้องและครอบคลุมมากขึ้น

1.5.6 สามารถประเมินความเสี่ยงได้อย่างอัตโนมัติ

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะนำเสนอทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาษาธรรมชาติ (NLP) โดยใช้ Large Language Models (LLMs) ในการแก้ไขปัญหาการประเมินความเสี่ยงในโรงงานอุตสาหกรรม ซึ่งครอบคลุมทฤษฎีหลัก ๆ เช่น Transformer, BERT, SBERT และการใช้ Cloud computing ในการประมวลผลข้อมูลTransformer เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมที่ช่วยให้การประมวลผลภาษามีประสิทธิภาพมากขึ้น โดยใช้กลไกการใส่ใจ (attention mechanism) ทำให้สามารถเรียนรู้การเชื่อมโยงระหว่างคำได้อย่างแม่นยำ BERT (Bidirectional Encoder Representations from Transformers) เป็นโมเดลที่ต่อยอดจาก Transformer ซึ่งสามารถทำความเข้าใจบริบทของคำได้จากทั้งสองทิศทาง (ซ้ายไปขวาและขวาไปซ้าย) ในขณะที่ SBERT (Sentence-BERT) เป็นการประยุกต์ใช้ BERT เพื่อคำนวณความคล้ายคลึงระหว่างประโยคได้อย่างมีประสิทธิภาพ นอกจากนี้ การใช้ Cloud computing ยังมีบทบาทสำคัญในการประมวลผลข้อมูลขนาดใหญ่ เนื่องจากช่วยลดข้อจำกัดทางด้านทรัพยากรฮาร์ดแวร์และช่วยให้การเทรนโมเดลเป็นไปอย่างรวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น

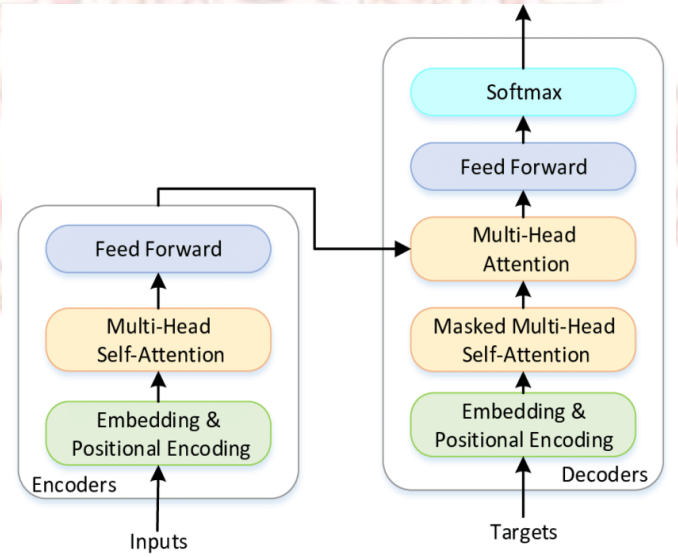
2.1 Transformer model

ทรานฟอร์เมอร์ โมเดล เป็นสถาปัตยกรรมการเรียนรู้เชิงลึกที่พัฒนาโดย Google และใช้กลไกความสนใจแบบหลายหัว ซึ่งนำเสนอในงานวิจัยปี 2017 เรื่อง "Attention Is All You Need" และโมเดลจำเป็นต้องมีข้อมูลขนาดใหญ่ จึงต้องใช้เวลาในการทำ pre-trained model หน่วยความจำระยะสั้นระยะยาว (LSTM) ชุดข้อมูล (ภาษา) เช่น คลังข้อมูล Wikipedia และ Common Crawl ข้อความจะถูกแปลงเป็นรูปแบบตัวเลขที่เรียกว่าโทเค็น และแต่ละโทเค็นจะถูกแปลงเป็นเวกเตอร์โดยการค้นหาจากตารางการฝังคำ ในแต่ละเลเยอร์ แต่ละโทเค็นจะถูกปรับบริบทภายในขอบเขตของหน้าต่างบริบทด้วยโทเค็นอื่นๆ ผ่านกลไกที่เรียกกว่า Attention ซึ่งช่วยให้ข้อมูลสำหรับโทเค็นหลักสามารถขยายได้ และโทเค็นที่สำคัญน้อยกว่าจะลดลง ทรายฟอร์เมอร์ ที่ตีพิมพ์ในปี 2017 อิงตามกลไกความสนใจแบบ Softmax ที่เสนอโดย Bahdanau et อัล ในปี 2014 สำหรับการแปลด้วยเครื่องและตัวควบคุมเวท แบบเร็ว สถาปัตยกรรมนี้ไม่เพียงแต่ใช้ในการประมวลผลภาษาธรรมชาติและคอมพิวเตอร์วิทัศน์เท่านั้น แต่ยังใช้ในการประมวลผลเสียง และการประมวลผลหลายรูปแบบด้วย นอกจากนี้ยังนำไปสู่การพัฒนาระบบที่ได้รับการฝึกอบรมล่วงหน้า เช่น ทรานฟอร์เมอร์ที่ได้รับการฝึกอบรมล่วงหน้าแบบกำเนิด GPT และ BERT การนำโมเดลทรานฟอร์เมอร์ไปใช้ในหลากหลายโดเมนการประมวลผลข้อมูลได้ทำ

ให้เกิดการเปลี่ยนแปลงอย่างมากในวงการวิจัยและการใช้งานจริง ไม่ว่าจะเป็นการประมวลผลภาษาธรรมชาติที่สามารถสร้างข้อความที่มีความสมเหตุสมผลมากขึ้น การประมวลผลภาพที่สามารถตรวจจับวัตถุได้อย่างแม่นยำมากขึ้น หรือการประมวลผลเสียงที่สามารถแยกแยะคำพูดหรือเสียงต่างๆ ได้อย่างถูกต้อง โมเดลทรานฟอร์เมอร์ยังมีความยืดหยุ่นในการทำงานกับข้อมูลหลายรูปแบบพร้อมกัน ซึ่งเป็นประโยชน์อย่างมากในงานวิจัยที่ต้องการการบูรณาการข้อมูลจากหลายแหล่งนอกจากนี้ การที่ทรานฟอร์เมอร์สามารถทำงานได้อย่างมีประสิทธิภาพในระดับใหญ่ยังเป็นการเปิดโอกาสในการพัฒนานวัตกรรมใหม่ๆ ที่อาจยังไม่เคยมีมาก่อน การวิจัยในปัจจุบันยังคงดำเนินการเพื่อปรับปรุงสถาปัตยกรรมและประสิทธิภาพของโมเดลทรานฟอร์เมอร์ให้ดียิ่งขึ้น เช่น การลดขนาดของโมเดลเพื่อให้สามารถนำไปใช้ในอุปกรณ์ที่มีทรัพยากรจำกัด หรือการปรับปรุงกลไกความสนใจให้มีประสิทธิภาพมากขึ้นเพื่อลดความซับซ้อนในการคำนวณ ทั้งนี้เพื่อให้การประยุกต์ใช้งานของโมเดลทรานฟอร์เมอร์สามารถขยายตัวไปในวงกว้างยิ่งขึ้นในอนาคต

2.1.1 Transformer Architecture

ทรานฟอร์เมอร์สามารถแบ่งได้ 2 ฝั่งคือฝั่งเข้ารหัส (Encoder) และถอดรหัส (Decoder) โดยสามารถใช้ในการประมวลผลทางภาษาได้ทั้ง 2 ฝั่ง แต่จะต้องแลกมาด้วยการคำนวณที่มีราคาแพงทำให้บริษัทใหญ่ๆในโลกเลือกใช้แค่ฝั่งเดียวแต่ก็ยังสามารถให้ผลลัพธ์ออกมาได้อย่างมีประสิทธิภาพเช่น BERT ใช้แค่การเข้ารหัสและ GPT ใช้แค่การถอดรหัส ดังรูป



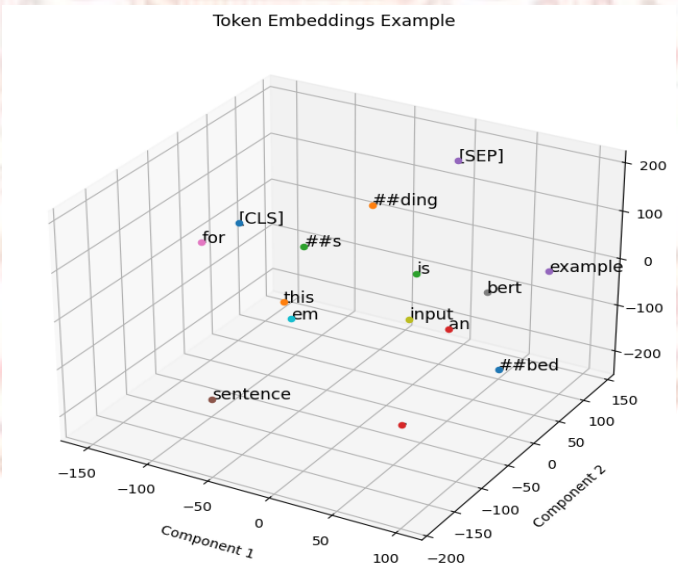
ภาพที่ 2-1 Transformer Architecture

การเลือกใช้ฝั่งใดฝั่งหนึ่งของทรานฟอร์เมอร์ขึ้นอยู่กับประเภทของงานที่ต้องการประมวลผล สำหรับงานที่ต้องการความเข้าใจในบริบทของข้อความ เช่น การค้นหาข้อมูล การตอบคำถาม หรือการวิเคราะห์ความรู้สึก BERT ซึ่งใช้ฝั่งเข้ารหัสเป็นหลักจะเหมาะสมมากกว่า เนื่องจากสามารถสร้างการแทนค่าที่มีบริบทอย่างลึกซึ้งและเข้าใจความหมายของข้อความได้อย่างดีเยี่ยม ในทางตรงกันข้าม สำหรับงานที่ต้องการการสร้างข้อความใหม่จากข้อมูลที่มีอยู่ เช่น การเขียนบทความ การแปลภาษา หรือการสร้างข้อความที่เป็นบทสนทนา GPT ซึ่งใช้ฝั่งถอดรหัสเป็นหลักจะมีประสิทธิภาพสูงกว่า เนื่องจากสามารถสร้างข้อความที่มีความสมเหตุสมผลและเป็นธรรมชาติ ทั้งนี้ การที่ทรานฟอร์เมอร์มีการคำนวณที่ซับซ้อนและต้องการทรัพยากรสูงในการฝึกอบรมและประมวลผลทำให้มีการพัฒนาและปรับปรุงโมเดลใหม่ๆ เพื่อเพิ่มประสิทธิภาพและลดความซับซ้อนในการคำนวณ เช่น การพัฒนาโมเดลที่มีขนาดเล็กลงแต่ยังคงรักษาความสามารถในการประมวลผลได้ดี หรือการใช้เทคนิคการฝึกอบรมที่มีประสิทธิภาพมากขึ้นเพื่อลดเวลาการฝึกอบรมและลดการใช้ทรัพยากร ในอนาคต การพัฒนาและการปรับปรุงโมเดลทรานฟอร์เมอร์ยังคงดำเนินไปอย่างต่อเนื่อง เพื่อให้สามารถตอบสนองต่อความต้องการที่หลากหลายและเพิ่มขีดความสามารถในการประมวลผลข้อมูลอย่างมีประสิทธิภาพมากยิ่งขึ้น ทำให้เทคโนโลยีนี้เป็นพื้นฐานสำคัญในวงการปัญญาประดิษฐ์และการประมวลผลภาษาธรรมชาติต่อไป

2.1.1.1 Input embeddings

Input embeddings เป็นการฝังลงข้อมูลเข้าตัวหนังสือหมายถึงการแทนคำหรือโทเค็นในข้อความเป็นเวกเตอร์หนาแน่นในพื้นที่เวกเตอร์ต่อเนื่อง การฝังเหล่านี้มักถูกใช้เป็นอินพุตให้กับโมเดลเรียนรู้ของเครื่องเฉพาะอย่างยิ่งในงานประมวลผลภาษาธรรมชาติ Word embeddings เช่น Word2Vec, GloVe หรือ fastText เป็นการใช้งานฝังลงที่มีความนิยมในการสร้างการแสดงเหล่านี้ แต่ละคำถูกแมพเข้าไปในเวกเตอร์มิติสูงที่คำที่คล้ายกันในเวกเตอร์พื้นที่ ระบบสามารถเข้าใจและประมวลผลภาษาได้อย่างมีประสิทธิภาพเนื่องจากการแสดงเหล่านี้จับความหมายที่เกี่ยวข้องและเชิงประสงค์ของคำ การฝังลงเป็นสิ่งสำคัญในงานเช่นการจำแนกประเภทข้อความ การวิเคราะห์ความรู้สึก การแปลภาษาเครื่อง และงานอื่น ๆ เพราะมันให้วิธีการแสดงข้อมูลข้อความในรูปแบบที่อัลกอริทึมเรียนรู้ของเครื่องสามารถทำงานได้อย่างมีประสิทธิภาพเป็นสิ่งที่สำคัญในกระบวนการประมวลภาษาธรรมชาติ (NLP) ซึ่งทำหน้าที่เป็นสะพานระหว่างข้อมูลข้อความที่เป็นตัวหลังกับอัลกอริทึมเครื่องจักร ความสำคัญและการใช้งานของ input embeddings โดยการจับสมรรถนะทางสัณฐานและไวยากรณ์ของข้อความไทยอย่าง

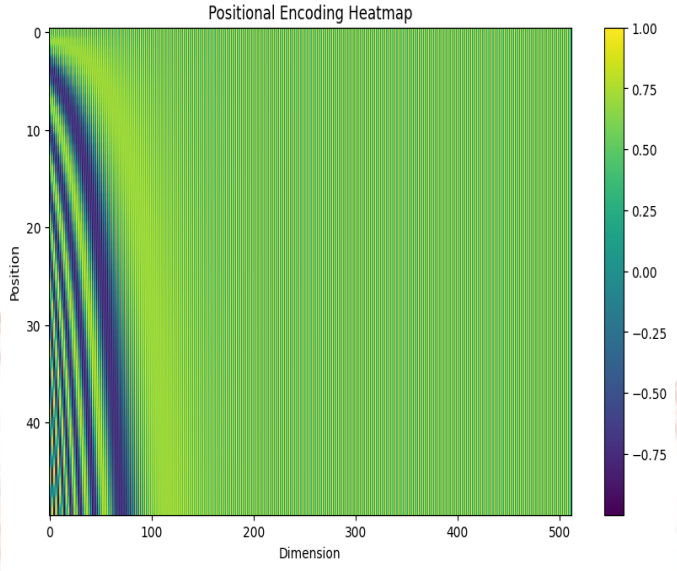
มีประสิทธิภาพ เหล่านี้ทำให้ NLP models สามารถทำงานในงานต่างๆ เช่น การจำแนกประเภทข้อความ การวิเคราะห์อารมณ์ และการแปลเครื่องข่าย ได้ด้วยความแม่นยำและความเร็วมากขึ้นข้อความที่เข้ามายังโมเดลจะถูกเข้ารหัสเรียกกว่าโทเค็นโดยโทเค็นไนเซอร์แต่ละโทเค็นจะถูกแปลงเป็นเวกเตอร์จากการที่โมเดลถูกเทรนไว้ล่วงหน้าขึ้นอยู่กับกลุ่มคำของคำที่ถูกป้อนเข้ามา นอกจากการใช้งานในงานประมวลผลภาษาธรรมชาติแล้ว input embeddings ยังมีบทบาทสำคัญในการทำงานร่วมกับโมเดลทรานฟอร์เมอร์ต่างๆ เนื่องจากการฝังลงสามารถปรับแต่งให้เหมาะสมกับบริบทของแต่ละงานได้ ตัวอย่างเช่น ในโมเดล BERT การฝังลงคำแต่ละคำจะถูกนำไปใช้ในกระบวนการเข้ารหัสเพื่อสร้างบริบทที่ละเอียดและแม่นยำ ซึ่งช่วยในการทำความเข้าใจและตีความข้อมูลข้อความได้ดีขึ้น ในขณะเดียวกัน GPT ใช้การฝังลงในกระบวนการถอดรหัสเพื่อสร้างข้อความใหม่ที่มีความสมเหตุสมผลและสอดคล้องกับบริบทที่ได้รับ การพัฒนาวิธีการฝังลงที่มีประสิทธิภาพมากขึ้นยังคงเป็นเป้าหมายสำคัญของนักวิจัยในสาขา NLP การทดลองและปรับปรุงเทคนิคต่างๆ เช่น การใช้ contextual embeddings ที่สามารถจับความหมายของคำในบริบทที่แตกต่างกัน หรือการพัฒนา embeddings ที่สามารถใช้กับภาษาต่างๆ ได้อย่างเท่าเทียม สิ่งเหล่านี้จะช่วยเพิ่มความสามารถในการประมวลผลภาษาและทำให้งานต่างๆ ในด้าน NLP มีความแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น



ภาพที่ 2-2 Input embeddings

2.1.1.2 Positional encoding

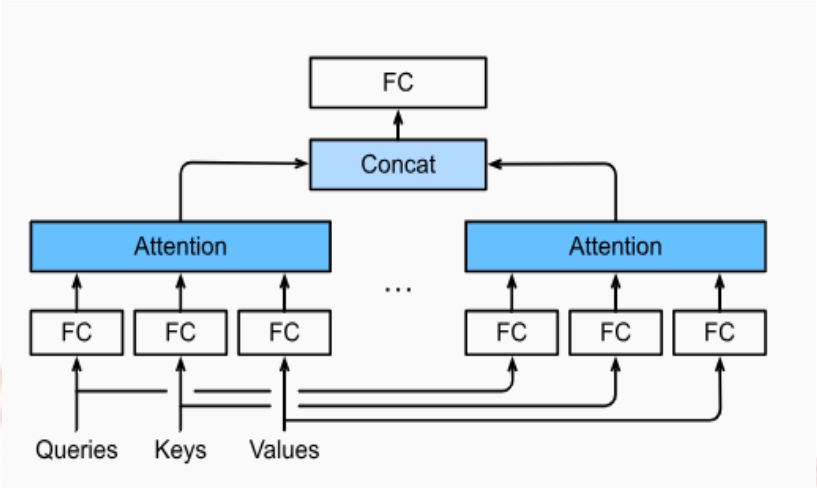
ตัวเข้ารหัสแต่ละตัวประกอบด้วยองค์ประกอบหลักสองส่วน self-attention และโครงข่ายประสาทเทียมแบบฟีดไปข้างหน้า self-attention จะรับการเข้ารหัสอินพุตจากตัวเข้ารหัสก่อนหน้า และ weight ความเกี่ยวข้องซึ่งกันและกันเพื่อสร้างการเข้ารหัสเอาต์พุต โครงข่ายประสาทเทียมแบบฟีดไปข้างหน้าจะประมวลผลการเข้ารหัสเอาต์พุตแต่ละรายการแยกกัน จากนั้นการเข้ารหัสเอาต์พุตเหล่านี้จะถูกส่งไปยังตัวเข้ารหัสถัดไปเป็นอินพุต เช่นเดียวกับตัวถอดรหัส ตัวเข้ารหัสตัวแรกจะใช้ข้อมูลตำแหน่งและการฝังลำดับอินพุตเป็นอินพุต แทนที่จะเข้ารหัส ข้อมูลตำแหน่งเป็นสิ่งจำเป็นสำหรับหม้อแปลงเพื่อใช้ลำดับของลำดับ เนื่องจากไม่มีส่วนอื่นของทรานฟอร์มอร์ใช้ประโยชน์จากสิ่งนี้ ตัวเข้ารหัสเป็นแบบสองทิศทาง สามารถให้ความสนใจกับโทเค็นก่อนและหลังโทเค็นปัจจุบันได้ มีการใช้โทเค็นแทนคำเพื่ออธิบายถึงความหลากหลาย ในตัวถอดรหัส (decoder) แต่ละตัวก็มีการทำงานที่คล้ายกันแต่มีความแตกต่างเล็กน้อยจากตัวเข้ารหัส (encoder) ตัวถอดรหัสจะประกอบด้วยสามส่วนหลัก: self-attention, encoder-decoder attention และโครงข่ายประสาทเทียมแบบฟีดไปข้างหน้า self-attention ในตัวถอดรหัสจะทำงานเหมือนกับในตัวเข้ารหัส แต่ encoder-decoder attention จะทำหน้าที่เชื่อมต่อข้อมูลที่ถูกประมวลผลจากตัวเข้ารหัสกับข้อมูลที่กำลังประมวลผลในตัวถอดรหัส เพื่อสร้างการเข้ารหัสเอาต์พุตที่มีความหมายในบริบทของข้อความทั้งหมด โครงข่ายประสาทเทียมแบบฟีดไปข้างหน้าจะทำการประมวลผลการเข้ารหัสเอาต์พุตเพื่อสร้างผลลัพธ์สุดท้าย ความสามารถในการประมวลผลแบบสองทิศทางของตัวเข้ารหัสทำให้โมเดลทรานฟอร์เมอร์สามารถเข้าใจบริบทของคำทั้งก่อนและหลังคำปัจจุบันได้ ซึ่งช่วยให้โมเดลสามารถตีความความหมายของข้อความได้อย่างแม่นยำยิ่งขึ้น ในขณะที่ตัวถอดรหัสที่มีการประมวลผลแบบลำดับต่อเนื่อง (sequential processing) จะให้ผลลัพธ์ที่สอดคล้องกับบริบทของข้อมูลที่มีการนำเข้าในลำดับนั้นๆ การใช้ข้อมูลตำแหน่ง (positional encoding) ในทรานฟอร์เมอร์ยังเป็นสิ่งสำคัญมาก เนื่องจากทรานฟอร์เมอร์ไม่มีการใช้ลำดับเชิงลึกเหมือนกับโมเดล RNN หรือ LSTM ข้อมูลตำแหน่งช่วยให้โมเดลทรานฟอร์เมอร์เข้าใจลำดับของโทเค็นในข้อความ และสามารถประมวลผลข้อความได้อย่างมีประสิทธิภาพและแม่นยำ สรุปได้ว่า การออกแบบของตัวเข้ารหัสและตัวถอดรหัสในโมเดลทรานฟอร์เมอร์ทำให้สามารถประมวลผลข้อมูลที่ซับซ้อนได้อย่างมีประสิทธิภาพ ไม่ว่าจะเป็นการประมวลผลภาษาธรรมชาติ การแปลภาษา การวิเคราะห์ข้อความหรือการสร้างข้อความ ทำให้ทรานฟอร์เมอร์เป็นเครื่องมือที่ทรงพลังในวงการปัญญาประดิษฐ์และการประมวลผลข้อมูลที่กำลังพัฒนาอย่างรวดเร็วในปัจจุบัน



ภาพที่ 2-3 Positional Encoding

2.1.1.3 Multi-head attention

ในโมเดลที่ใช้ทรานฟอร์เมอร์เพื่อการประมวลผลภาษาธรรมชาติ เช่น GPT ของ OpenAI มีโครงสร้างที่เน้นการถอดรหัสเพื่อสร้างข้อความอย่างถดถอย โดยในตัวถอดรหัสแต่ละตัวจะประกอบด้วย self-attention และโครงข่ายประสาทเทียมที่ส่งต่อ การทำงานของตัวถอดรหัสจะคล้ายกับตัวเข้ารหัส แต่มีการใช้ความสนใจเพิ่มเติมเพื่อดึงข้อมูลที่เกี่ยวข้องจากการเข้ารหัสที่สร้างโดยตัวเข้ารหัส กลไกนี้ช่วยให้โมเดลสามารถเรียนรู้และเข้าใจความสัมพันธ์ในข้อความได้อย่างลึกซึ้งมากขึ้น ตัวถอดรหัสแต่ละตัวจะใช้ข้อมูลตำแหน่งและการฝังลำดับเอาต์พุตเป็นอินพุต เพื่อช่วยในการคาดการณ์เอาต์พุตที่ถูกสร้างขึ้น โดยหม้อแปลงไฟฟ้าจะไม่ใช้เอาต์พุตปัจจุบันหรืออนาคตเพื่อป้องกันการไหลข้อมูลย้อนกลับ สิ่งนี้ทำให้สามารถสร้างข้อความที่มีความน่าจะเป็นสูงและเข้าใจได้ในมุมมองของบริบทของข้อความทั้งหมดได้ การใช้ Attention Heads ในตัวถอดรหัสเป็นส่วนสำคัญที่ช่วยให้โมเดลสามารถตรวจจับความสัมพันธ์และความสำคัญของโทเค็นในข้อความได้อย่างแม่นยำ โดยสามารถจัดการกับคำถามซับซ้อนหรือข้อความที่มีโครงสร้างซับซ้อนได้อย่างมีประสิทธิภาพ ในทางทฤษฎีการประมวลผลภาษาธรรมชาติ (NLP) การใช้ทรานฟอร์เมอร์ที่มีการถอดรหัสที่ทันสมัยและมีประสิทธิภาพเป็นสิ่งสำคัญ ซึ่งช่วยให้โมเดลสามารถทำงานได้มากขึ้นในงานต่างๆ เช่น การสร้างคำถามหรือการตอบคำถามอัจฉริยะ การแปลภาษา และการสร้างข้อความที่สมเหตุสมผล



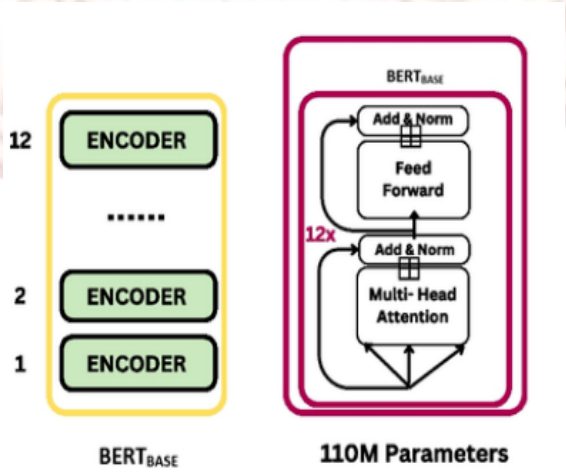
ภาพที่ 2-4 Multi-head attention

2.2 BERT model

BERT (Bidirectional Encoder Representations from Transformers) เป็นโมเดลที่พัฒนาขึ้นโดยนักวิจัยจาก Google และเปิดตัวในตุลาคม ค.ศ. 2018 เป็นฮิตในการปรับปรุงโมเดลการประมวลผลภาษาธรรมชาติ (NLP) โดยที่มีการวิจัยมากมายที่ตีพิมพ์เพื่อวิเคราะห์และปรับปรุงโมเดลนี้เป็นจำนวนมาก ตั้งแต่นำ BERT มาใช้ในภาษาอังกฤษมีสองขนาดของโมเดล คือ BERTBASE และ BERTLARGE

2.2.1 BERTBASE

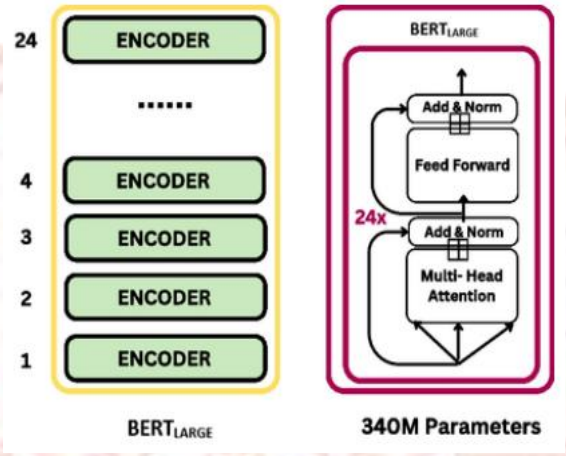
BERTBASE ประกอบด้วยตัวเข้ารหัส 12 ล้านตัวพร้อม self-attention แบบสองทิศทาง 12 หัว โดยมีทั้งหมด 110 ล้านพารามิเตอร์



ภาพที่ 2-5 BERTBASE

2.2.2 BERTBASE LARGE

BERTLARGE ประกอบด้วยตัวเข้ารหัส 24 ล้านตัวพร้อม self-attention แบบสองทิศทาง 16 หัว โดยมีทั้งหมด 340 ล้านพารามิเตอร์

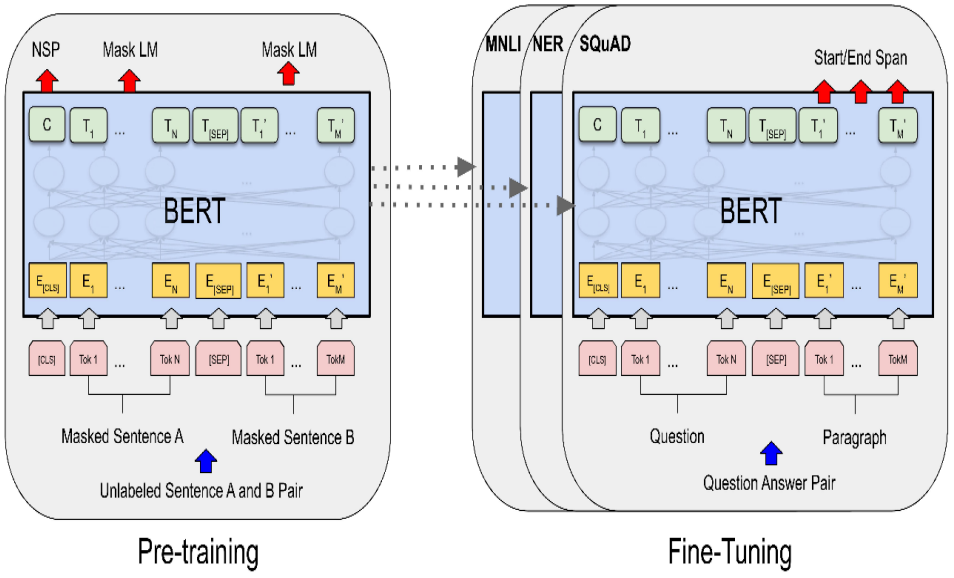


ภาพที่ 2-6 BERTLARGE

ทั้งสองโมเดลนี้ได้รับการฝึกอบรมล่วงหน้าด้วยข้อมูลจาก Toronto BookCorpus (800 ล้านคำ) และวิกิพีเดียภาษาอังกฤษ (2,500 ล้านคำ) ซึ่งทำให้ BERT มีความสามารถในการเรียนรู้และเข้าใจภาษาธรรมชาติได้อย่างลึกซึ้งและมีประสิทธิภาพสูง ทั้ง BERTBASE และ BERTLARGE เป็นโมเดลที่มีความสามารถในการปรับปรุงและประมวลผลข้อมูลทางภาษาอย่างมีประสิทธิภาพ โดย BERTBASE มีขนาดที่ใหญ่พอสมควรที่จะใช้ในงานที่ต้องการประมวลผลทางภาษาอย่างรวดเร็ว ในขณะที่ BERTLARGE มีขนาดที่ใหญ่กว่าเพื่อให้สามารถประมวลผลข้อมูลที่ซับซ้อนและมีขนาดใหญ่ขึ้นได้อย่างมีประสิทธิภาพยิ่งขึ้น การฝึก BERT ด้วยข้อมูลจาก Toronto BookCorpus และวิกิพีเดียภาษาอังกฤษช่วยเสริมความเข้าใจในภาษาธรรมชาติของโมเดลอย่างมาก ทำให้ BERT สามารถจำแนกและทำนายข้อความได้อย่างแม่นยำ การวิจัยและการพัฒนา BERT ยังคงต่อยอดไปอย่างต่อเนื่อง โดยนักวิจัยทั่วโลกใช้ BERT เพื่อพัฒนาและปรับปรุงโมเดลในหลายสาขาของ NLP เช่น การสร้างคำถามและตอบ การแปลภาษา และการวิเคราะห์เนื้อหาออนไลน์ ผลงานที่เกี่ยวข้องกับ BERT ได้ถูกนำไปใช้ประโยชน์ในการพัฒนาแอปพลิเคชันและเครื่องมือต่างๆ ที่ใช้งานในชีวิตประจำวัน เช่น การค้นหาบนเว็บ การแนะนำสินค้า และการตอบคำถามอัตโนมัติ

2.2.3 BERT Fine-tuning

การทำ Fine-tuning ของโมเดล BERT (Bidirectional Encoder Representations from Transformers) เป็นกระบวนการที่สำคัญและมีความซับซ้อนในการปรับแต่งโมเดลที่ถูกฝึกด้วยข้อมูลใหญ่มาก่อนหน้านี้ เพื่อให้สามารถทำงานได้ดีในงานที่เฉพาะเจาะจงและมีประสิทธิภาพสูงในการประมวลผลภาษาธรรมชาติ (NLP) ต่างๆ โมเดล BERT นั้นถูกฝึกด้วยข้อมูลมหภาพในลักษณะที่เป็นไปได้ทั้งหมด แต่การทำ Fine-tuning เป็นการปรับปรุงโมเดลให้เข้ากับงานที่ต้องการเฉพาะโดยใช้ข้อมูลที่เฉพาะเจาะจงเพื่อเพิ่มประสิทธิภาพของการประมวลผล ขั้นตอนแรกใน Fine-tuning คือการเลือกโมเดล BERT ที่เหมาะสมสำหรับงานที่ต้องการ โดยโมเดล BERT มีหลากหลายขนาดและความซับซ้อน เช่น BERTBASE และ BERTLARGE ที่มีจำนวนพารามิเตอร์และความสามารถในการประมวลผลที่แตกต่างกัน นักวิจัยต้องตัดสินใจเลือกโมเดลที่เหมาะสมกับขนาดของข้อมูลและความซับซ้อนของงานที่ต้องการ หลังจากนั้นคือขั้นตอนการฝึกโมเดลใหม่ด้วยข้อมูลที่เป็นไปได้ทั้งหมด (full training dataset) ซึ่งประกอบด้วยข้อมูลฝึกและข้อมูลทดสอบ โมเดลจะถูกปรับแต่งพารามิเตอร์เพื่อให้สามารถทำนายผลลัพธ์ที่ถูกต้องและเหมาะสมกับงานที่กำหนดไว้ การ Fine-tuning นี้มักใช้เทคนิคการค้นหาเชิงลึก (hyperparameter tuning) เพื่อปรับค่าพารามิเตอร์ที่เหมาะสมกับการประมวลผลที่ต้องการ เมื่อโมเดลได้รับการ Fine-tuning เพียงพอแล้ว ขั้นตอนต่อไปคือการทดสอบประสิทธิภาพของโมเดลด้วยชุดข้อมูลทดสอบที่เป็นอิสระ เพื่อวัดความแม่นยำและประสิทธิภาพในงานที่ต้องการ การทดสอบนี้ช่วยตรวจสอบว่าโมเดลที่ถูก Fine-tuning สามารถทำงานได้ดีตามที่คาดหวังหรือไม่และช่วยปรับปรุงโมเดลในกรณีที่ต้องการประสิทธิภาพเพิ่มเติม สุดท้ายคือการนำโมเดล BERT ที่ได้ผ่านกระบวนการ Fine-tuning มาใช้งานจริงในแอปพลิเคชันหรือระบบที่ต้องการการประมวลผลภาษาธรรมชาติที่มีประสิทธิภาพสูง การนำ BERT มาใช้งานจริงนั้นช่วยเพิ่มความสามารถในการเข้าใจและประมวลผลข้อความให้มีประสิทธิภาพและแม่นยำมากยิ่งขึ้น



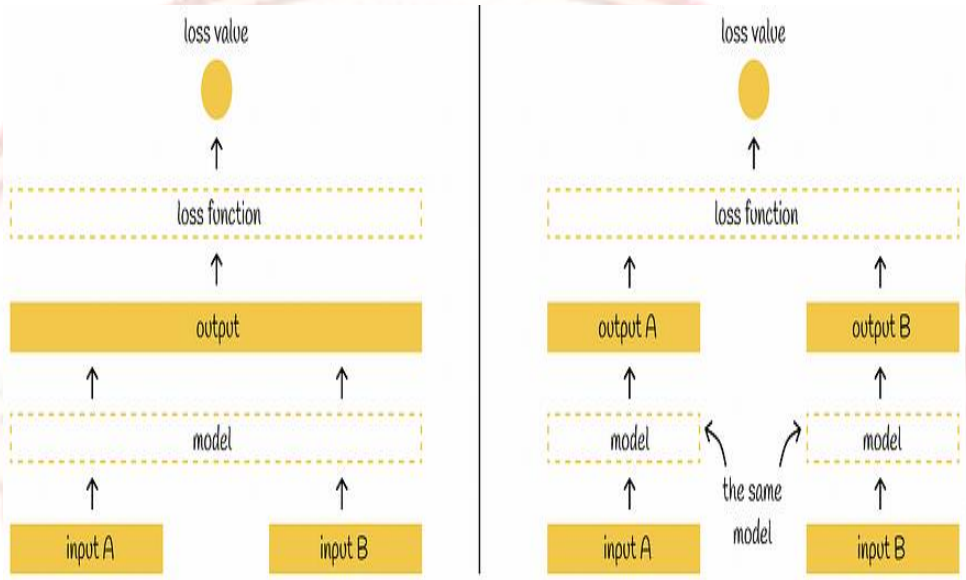
ภาพที่ 2-7 BERT โมเดล

(ที่มา: https://paperswithcode.com/method/bert)

2.3 SBERT

เครือข่ายสยาม (Siamese Network) เป็นแนวคิดที่นำมาใช้ในการสร้างแบบจำลองเชิงลำดับ (sequence model) โดยที่โมเดลจะมีโครงสร้างเดียวกันสำหรับทั้งสองประโยคที่ต้องการทำการเปรียบเทียบหรือการทำนายความคล้ายคลึงกัน ในที่นี้เราสามารถใช้เครือข่ายสยามเพื่อหาความคล้ายคลึงระหว่างประโยคที่สองประโยค โดยที่โมเดลทำงานด้วยโครงสร้างที่เหมือนกันและใช้น้ำหนัก (weights) เดียวกันในทุกโหนดของโครงข่าย เมื่อประโยคทั้งสองถูกส่งผ่านโมเดล เราจะได้ผลลัพธ์ที่บอกถึงความคล้ายคลึงของประโยคทั้งสองตามเกณฑ์ที่ได้กำหนดไว้ในโมเดล เช่น การใช้งานเครือข่ายสยามสามารถประยุกต์ใช้ในการคัดแยกประโยคที่คล้ายกันหรือไม่คล้ายกัน หรือในงานต่างๆ ที่ต้องการการจำแนกหรือการทำนายเชิงคล้ายคลึงของข้อมูล เช่น การจำแนกอารมณ์ของประโยค การจำแนกเสียงพูด หรือการตรวจจับคำและประโยคที่มีความหมายที่คล้ายคลึงกัน การออกแบบของเครือข่ายสยามมักมีโครงสร้างที่เป็นลำดับโครงข่าย (sequential architecture) ซึ่งประกอบด้วยชั้น (layers) ต่างๆ ที่เหมือนกัน แต่มีการปรับเปลี่ยนค่าของพารามิเตอร์เฉพาะตามงานที่ต้องการ ในกระบวนการฝึกอบรม เครือข่ายสยามจะถูกป้อนข้อมูลที่เป็นคู่ โดยที่ข้อมูลจะผ่านทางโครงสร้างเดียวกันของเครือข่าย ซึ่งทำให้โมเดลสามารถเรียนรู้ลักษณะเฉพาะของคู่ข้อมูลที่ต้องการเปรียบเทียบได้ ในการฝึกเครือข่ายสยาม เราจะใช้เทคนิคการปรับปรุงพารามิเตอร์ (parameter tuning) เพื่อปรับค่าของโมเดลให้สอดคล้องกับโจทย์ที่ต้องการ การปรับแต่งนี้สามารถทำได้โดยการใช้ชุดข้อมูลฝึกและชุดข้อมูล

ทดสอบที่แยกออกจากกัน เพื่อวัดประสิทธิภาพของโมเดลว่าสามารถทำนายหรือจำแนกข้อมูลที่ต้องการได้แม่นยำและเชื่อถือได้เครือข่ายสยามเป็นเครือข่ายที่มีการใช้งานอย่างกว้างขวางในหลายงานด้าน NLP และปัญญาประดิษฐ์ เช่น การจัดกลุ่มข้อความที่คล้ายคลึงกัน การจำแนกหรือคัดแยกข้อมูลที่มีลักษณะเดียวกัน หรือในงานที่ต้องการค้นหาข้อมูลที่คล้ายคลึงกัน เช่น การค้นหาเอกสารที่มีความหมายคล้ายกันจากฐานข้อมูลขนาดใหญ่



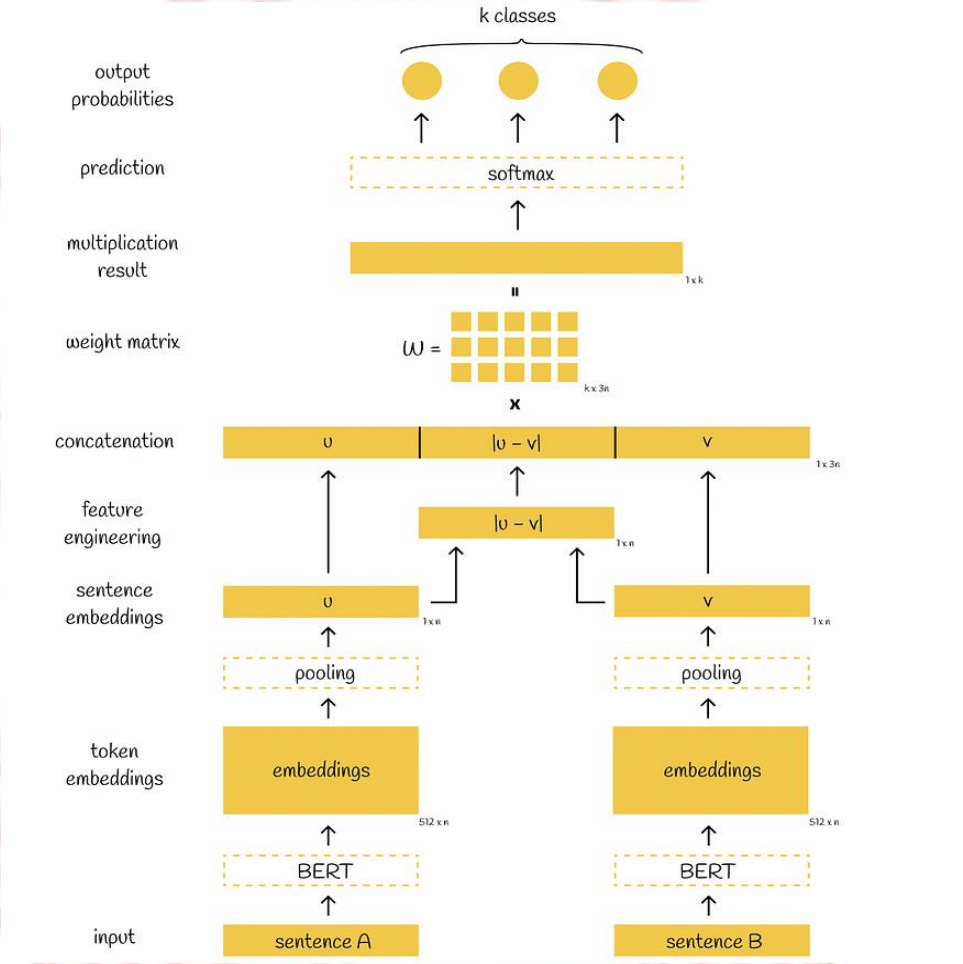
ภาพที่ 2-8 SBERT โมเดล

(ที่มา:Vyacheslav towardsdatascience , 2023)

2.3.1 SBERT Architecture

ในการใช้งาน SBERT (Sentence-BERT) หลังจากที่เราส่งประโยคทั้งสองผ่านโมเดล BERT แล้ว เราจะได้รับเวกเตอร์ที่มีขนาดเริ่มต้นที่ 768 มิติสำหรับแต่ละประโยค โดยทั่วไปเราจะนำเวกเตอร์ทั้งสองที่ได้นี้มาใช้ในเลเยอร์การรวมกลุ่ม (pooling layer) เพื่อให้ได้เวกเตอร์ซึ่งมีมิติที่ต่ำลง เช่น 768 มิติจะถูกแปลงเป็นเวกเตอร์ขนาดเดียวกัน โดยตามที่ผู้เขียน SBERT แนะนำให้ใช้เลเยอร์การรวมกลุ่มเฉลี่ยเป็นค่าเริ่มต้น เพื่อให้ได้เวกเตอร์ที่แทนลักษณะทั่วไปของประโยคทั้งสองที่ถูกเปรียบเทียบกันสำหรับเวกเตอร์ u และ v ที่ได้จากการรวมกลุ่ม เราสามารถใช้ในหลายวิธีเพื่อเพิ่มประสิทธิภาพในการประมวลผลต่อไป เช่น สามารถใช้เวกเตอร์ทั้งสองเพื่อคำนวณความคล้ายคลึงระหว่างประโยคทั้งสองด้วยวิธีการทางคณิตศาสตร์เชิงหนึ่ง เพื่อให้ได้ค่าความคล้ายคลึงที่มีความแม่นยำและเชื่อถือได้ อีกวิธีหนึ่งที่สามารถนำเวกเตอร์ u และ v ไปใช้ได้คือการนำมาใช้ในการจัดกลุ่ม (clustering) ของประโยค โดยการใช้เทคนิคการจัดกลุ่มเชิง

ลำดับ (sequential clustering) เพื่อแบ่งประโยคที่มีความคล้ายคลึงกันลงเป็นกลุ่มต่างๆ ที่มีลักษณะเดียวกัน เช่น การจัดกลุ่มประโยคที่มีความหมายคล้ายกันเพื่อใช้ในงานต่างๆ ที่ต้องการค้นหาข้อมูลที่คล้ายคลึงกัน สุดท้ายเรายังสามารถนำเวกเตอร์ u และ v ไปใช้ในการทำนายหรือจัดการกับข้อมูลประโยคที่มีลักษณะคล้ายกัน ในกระบวนการทำนายหรือการประมวลผลต่อไป เช่น การทำนายคำตอบของคำถาม การจัดเรียงคำที่มีความหมายคล้ายกัน เป็นต้น



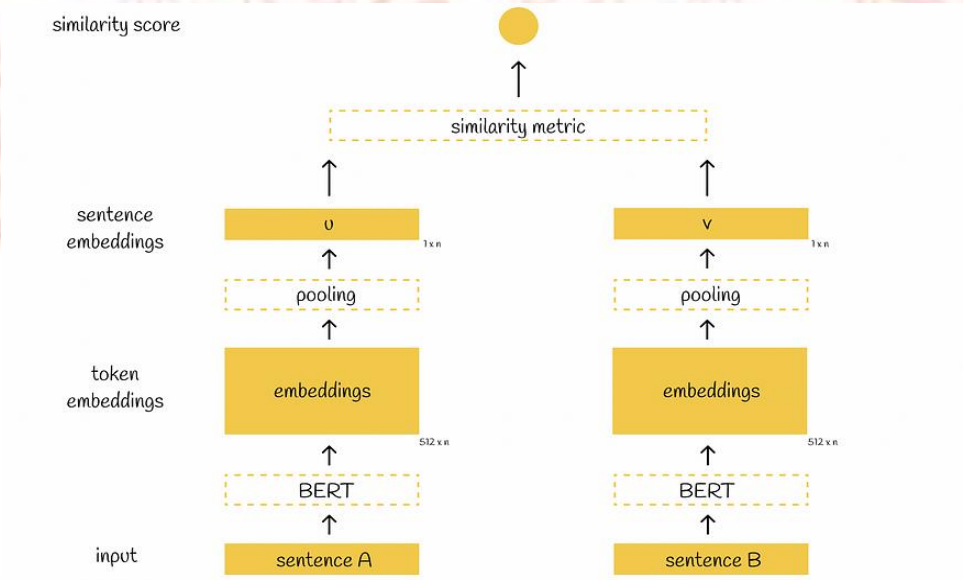
ภาพที่ 2-9 SBERT Architecture

(ที่มา:Vyacheslav towardsdatascience , 2023)

2.3.2 SBERT Similarity Metric

ในกระบวนการการทำงานของ SBERT (Sentence-BERT) หลังจากที่ได้รับเวกเตอร์ u และ v จากการรวมกลุ่มของประโยคที่ผ่านการประมวลผลของ BERT แล้ว ขั้นตอนต่อไปคือการคำนวณความคล้ายคลึงระหว่างเวกเตอร์เหล่านั้นโดยใช้หน่วยวัดความคล้ายคลึงที่เลือกไว้ ในที่นี้ผู้เขียนเลือกใช้ค่าความคล้ายคลึงโคไซน์ (cosine similarity) เนื่องจากเป็นหน่วย

วัดที่เหมาะสมและได้รับการนำมาใช้ในการวัดความคล้ายคลึงของเวกเตอร์อย่างแพร่หลายในงาน NLP (Natural Language Processing) และการทำซ้ำภายใน การคำนวณความคล้ายคลึงโดยใช้ค่าความคล้ายคลึงโคไซน์เป็นกระบวนการที่ง่ายและมีประสิทธิภาพ โดยค่าความคล้ายคลึงโคไซน์จะถูกคำนวณจากสินค้าตรงระหว่างเวกเตอร์ u และ v ซึ่งประกอบด้วยค่าความยาวของเวกเตอร์และค่าความคล้ายคลึงของมุมระหว่างเวกเตอร์สองตัวนั้น การที่ค่าความคล้ายคลึงโคไซน์มีค่ามากจะแสดงถึงประโยค u และ v มีความคล้ายคลึงกันมาก และมีค่ามุมระหว่างเวกเตอร์ที่เล็กลง เมื่อได้รับคะแนนความคล้ายคลึงแล้ว จะใช้ค่าความคล้ายคลึงที่คาดการณ์ไว้ในการเปรียบเทียบกับค่าความคล้ายคลึงที่เป็นจริง ซึ่งอาจเป็นค่าที่ได้จากข้อมูลการฝึกอบรมหรือชุดข้อมูลที่มีป้ายกำกับ โดยการเปรียบเทียบนี้ช่วยในการปรับปรุงแบบจำลองให้มีประสิทธิภาพมากยิ่งขึ้น ซึ่งสามารถทำได้โดยใช้ฟังก์ชันสูญเสีย MSE (Mean Squared Error) ในการคำนวณค่าความแตกต่างระหว่างคะแนนความคล้ายคลึงที่คาดการณ์ไว้กับค่าที่เป็นจริง ซึ่งจะทำให้เราได้แบบจำลองที่สามารถทำนายและวัดผลได้อย่างแม่นยำ ในแนวทางการใช้งานจริง ความคล้ายคลึงโคไซน์เป็นตัวชี้วัดที่สะดวกและมีประสิทธิภาพสำหรับการทำงานกับข้อมูลขนาดใหญ่ โดยเฉพาะในการทำซ้ำภายในของข้อมูลที่มีโครงสร้างซับซ้อน ซึ่งทำให้ SBERT เป็นเทคนิคที่ได้รับความนิยมในการประมวลผลข้อมูลที่เกี่ยวข้องกับภาษาธรรมชาติและใช้ในการคำนวณความคล้ายคลึงของข้อความ อย่างไรก็ตาม SBERT ยังสามารถปรับใช้หน่วยวัดความคล้ายคลึงอื่นๆ ได้ตามความเหมาะสมของการใช้งาน

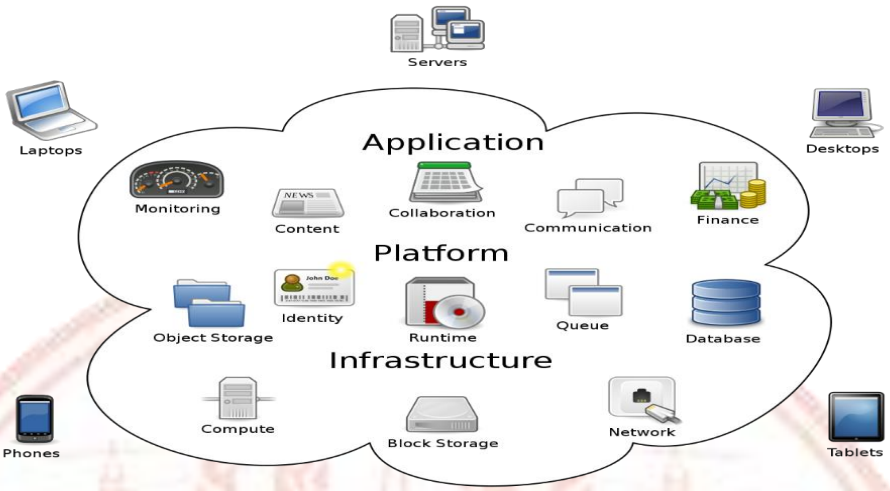


ภาพที่ 2-10 SBERT Similarity Metric

(ที่มา:Vyacheslav towardsdatascience , 2023)

2.4 Cloud computing

หนึ่งในความท้าทายหลักของการประมวลผลแบบคลาวด์เมื่อเปรียบเทียบกับการประมวลผลภายในองค์กรแบบดั้งเดิมคือความปลอดภัยของข้อมูลและความเป็นส่วนตัว ผู้ใช้คลาวด์มอบข้อมูลที่ละเอียดอ่อนของตนให้กับผู้ให้บริการบุคคลที่สาม ซึ่งอาจไม่มีมาตรการที่เพียงพอในการปกป้องข้อมูลจากการเข้าถึง การละเมิด หรือการรั่วไหลโดยไม่ได้รับอนุญาต ผู้ใช้คลาวด์ยังเผชิญกับความเสี่ยงในการปฏิบัติตามข้อกำหนดหากต้องปฏิบัติตามกฎระเบียบหรือมาตรฐานบางประการเกี่ยวกับการปกป้องข้อมูล เช่น GDPR หรือ HIPAA ความท้าทายอีกประการหนึ่งของการประมวลผลแบบคลาวด์คือการมองเห็นและการควบคุมที่ลดลง ผู้ใช้ระบบคลาวด์อาจไม่มีข้อมูลเชิงลึกโดยละเอียดเกี่ยวกับวิธีการจัดการ กำหนดค่า หรือเพิ่มประสิทธิภาพทรัพยากรระบบคลาวด์ของตนโดยผู้ให้บริการของตน พวกเขายังอาจมีความสามารถจำกัดในการปรับแต่งหรือปรับเปลี่ยนบริการคลาวด์ของตนตามความต้องการหรือความชอบเฉพาะของตน ความเข้าใจอย่างสมบูรณ์เกี่ยวกับเทคโนโลยีทั้งหมดอาจเป็นไปไม่ได้โดยเฉพาะอย่างยิ่งเมื่อคำนึงถึงขนาด ความซับซ้อน และความทึบแสงโดยเจตนาของระบบร่วมสมัย อย่างไรก็ตาม มีความจำเป็นที่จะต้องเข้าใจเทคโนโลยีที่ซับซ้อนและการเชื่อมโยงถึงกันเพื่อที่จะมีอำนาจและสิทธิ์เสรีภายในเทคโนโลยีเหล่านั้น คำอุปมาของคลาวด์ถือได้ว่าเป็นปัญหาเนื่องจากคลาวด์คอมพิวติ้งยังคงรักษารัศมีของบางสิ่งที่ไม่สำคัญและมากมาย มันเป็นสิ่งที่มีประสบการณ์โดยไม่เข้าใจอย่างแน่ชัดว่ามันคืออะไรหรือทำงานอย่างไรนอกจากนี้ การโยกย้ายระบบคลาวด์ยังเป็นประเด็นสำคัญอีกด้วย การโยกย้ายบนคลาวด์เป็นกระบวนการย้ายข้อมูล แอปพลิเคชัน หรือปริมาณงานจากสภาพแวดล้อมคลาวด์หนึ่งไปยังอีกที่หนึ่งหรือจากภายในองค์กรไปยังคลาวด์ การโยกย้ายระบบคลาวด์อาจซับซ้อน ใช้เวลานาน และมีค่าใช้จ่ายสูง โดยเฉพาะอย่างยิ่งหากมีปัญหาความไม่เข้ากันระหว่างแพลตฟอร์มคลาวด์หรือสถาปัตยกรรมที่แตกต่างกัน การโยกย้ายระบบคลาวด์อาจทำให้เกิดการหยุดทำงาน ประสิทธิภาพการทำงานลดลง หรือข้อมูลสูญหายได้ หากไม่ได้วางแผนและดำเนินการอย่างเหมาะสม



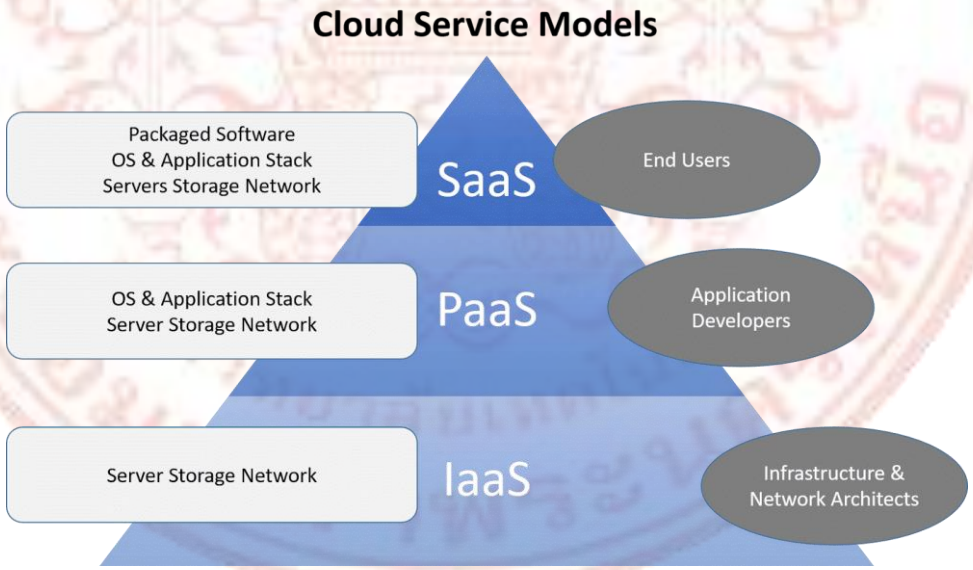
ภาพที่ 2-11 Cloud Infrastructure

2.4.1 Cloud models

สถาปัตยกรรมเชิงการบริการ (Service-Oriented Architecture, SOA) เป็นแนวคิดที่ส่งเสริมแนวคิด "ทุกสิ่งที่เป็นการบริการ" (Everything as a Service, EaaS หรือ XaaS) ซึ่งหมายถึงการให้บริการต่างๆ ผ่านโมเดลบริการที่ถูกกำหนดขึ้นโดยสถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (National Institute of Standards and Technology, NIST) โดยทั่วไปแบ่งเป็น Infrastructure as a Service (IaaS), Platform as a Service (PaaS), และ Software as a Service (SaaS) ซึ่งแต่ละโมเดลนี้เป็นการให้บริการที่มีระดับนามธรรมที่แตกต่างกันตามความต้องการของผู้ใช้และแอปพลิเคชันที่พัฒนาขึ้น เริ่มจาก Infrastructure as a Service (IaaS) ซึ่งเป็นการให้บริการที่ใช้งานฐานข้อมูลและพื้นที่จัดเก็บข้อมูลบนอินเทอร์เน็ต โดยทั่วไปจะรวมถึงการใช้งานเซิร์ฟเวอร์ เครือข่าย และพื้นที่จัดเก็บข้อมูลที่ให้บริการโดยผู้ให้บริการ IaaS ผู้ใช้สามารถเช่าพื้นที่พร้อมพื้นที่สำรองที่เหมาะสมกับความต้องการใช้งานของตนเอง โดยไม่จำเป็นต้องเป็นเจ้าของพื้นที่เหล่านั้น ซึ่งช่วยให้ธุรกิจสามารถลดต้นทุนการดำเนินงานและสามารถเข้าถึงพื้นที่นี้ได้ทุกเมื่อที่ต้องการ ต่อจากนั้นคือ Platform as a Service (PaaS) ซึ่งเป็นการให้บริการที่อนุญาตให้ผู้ใช้สามารถสร้างและดำเนินแอปพลิเคชันบนพื้นฐานของพื้นที่ข้อมูลที่เช่าเอาไว้ โดยไม่จำเป็นต้องกังวลเรื่องของการดูแลรักษาพื้นที่หรือสภาพแวดล้อมโฮสต์ ผู้ใช้สามารถใช้เครื่องมือพัฒนาซอฟต์แวร์ที่ให้มาพร้อมกับพื้นที่ที่บริการให้ได้อย่างเต็มประสิทธิภาพ ซึ่งช่วยลดเวลาในการพัฒนาและจัดการแอปพลิเคชันให้มีประสิทธิภาพมากยิ่งขึ้น สุดท้ายคือ Software as a Service (SaaS) ซึ่งเป็นการให้บริการโดยตรงแก่ผู้ใช้ที่สามารถเข้าถึงและใช้งาน

ซอฟต์แวร์ผ่านอินเทอร์เน็ต โดยไม่ต้องติดตั้งหรือดาวน์โหลดซอฟต์แวร์ลงบนอุปกรณ์ของตนเอง ส่วนใหญ่มักเป็นแอปพลิเคชันที่มีการใช้งานกว้างขวาง เช่น การจัดการเอกสารออนไลน์ หรือบริการอีเมล์ในรูปแบบคลาวด์ ที่สามารถเข้าถึงผ่านเบราว์เซอร์ได้ทุกที่ทุกเวลา โดยผู้ใช้จะได้รับประโยชน์จากค่าใช้จ่ายรายเดือนหรือรายปีตามการจัดหาของผู้ให้บริการ

SOA และแนวคิด EaaS หรือ XaaS นี้ไม่ได้ทำให้เลเยอร์แต่ละชั้นของบริการต้องพึ่งพากันอย่างมั่นใจ ซึ่งผู้ใช้สามารถเลือกใช้เฉพาะบริการที่ต้องการและเหมาะสมกับความต้องการใช้งานของตนเองได้โดยตรง ยกตัวอย่างเช่น การใช้ SaaS ซึ่งสามารถทำงานโดยตรงบนพื้นที่ IaaS โดยไม่ต้องผ่าน PaaS และมีความยืดหยุ่นในการใช้งานและปรับเปลี่ยนตามความต้องการของธุรกิจได้.ด้วยความสะดวกสบายและความยืดหยุ่นในการใช้งาน แนวคิด SOA และ EaaS หรือ XaaS ได้เป็นที่ยอมรับอย่างกว้างขวางในวงการธุรกิจและอุตสาหกรรมเทคโนโลยีสารสนเทศ ทำให้ธุรกิจสามารถเพิ่มประสิทธิภาพและลดต้นทุนในการดำเนินงานได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

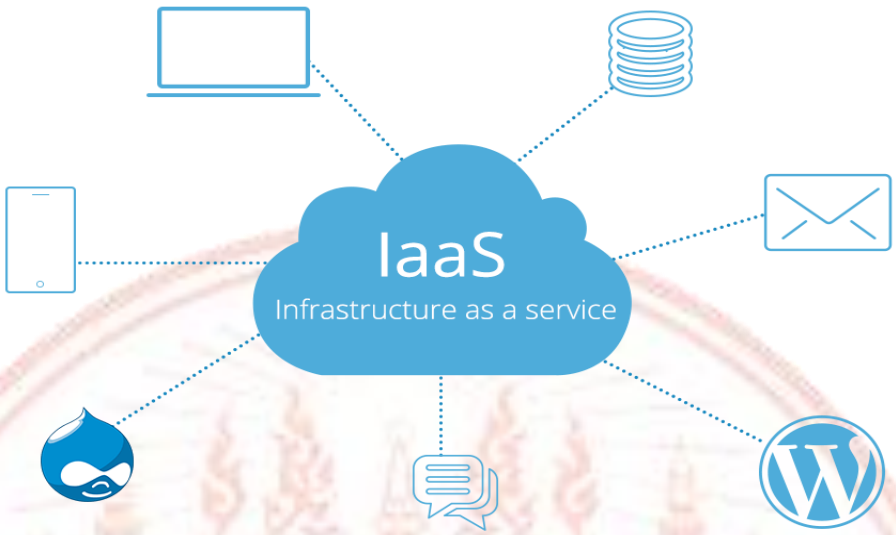


ภาพที่ 2-12 Cloud Service Models

2.4.1.1 Infrastructure as a service(IaaS)

โครงสร้างพื้นฐานเป็นบริการ (IaaS) หมายถึงบริการออนไลน์ที่ให้บริการ API ระดับสูงที่ใช้ในการสรุปรายละเอียดระดับต่ำต่างๆ ของโครงสร้างพื้นฐานเครือข่ายพื้นฐาน เช่น ทรัพยากรการประมวลผลทางกายภาพ ตำแหน่ง การแบ่งพาร์ติชันข้อมูล การปรับขนาด ความ

ปลอดภัย การสำรองข้อมูล ฯลฯ ไฮเปอร์ไวเซอร์รันเครื่องเสมือนในฐานะแขก กลุ่มไฮเปอร์ไวเซอร์ภายในระบบปฏิบัติการคลาวด์สามารถรองรับเครื่องเสมือนจำนวนมาก และความสามารถในการขยายขนาดบริการขึ้นและลงตามความต้องการที่แตกต่างกันของลูกค้า คอนเทนเนอร์ Linux ทำงานในพาร์ติชันแยกของเคอร์เนล Linux เดียวที่ทำงานบนฮาร์ดแวร์กายภาพโดยตรง cgroups และเนมสเปซ Linux เป็นเทคโนโลยีเคอร์เนล Linux พื้นฐานที่ใช้ในการแยก รักษาความปลอดภัย และจัดการคอนเทนเนอร์ การใช้คอนเทนเนอร์ให้ประสิทธิภาพที่สูงกว่าการจำลองเสมือน เนื่องจากไม่มีค่าใช้จ่ายไฮเปอร์ไวเซอร์ คลาวด์ IaaS มักนำเสนอทรัพยากรเพิ่มเติม เช่น ไลบรารีดิสก์-อิมเมจของเครื่องเสมือน พื้นที่จัดเก็บแบบบล็อกดิบ พื้นที่จัดเก็บไฟล์หรืออ็อบเจ็กต์ ไฟร์วอลล์ โหลดบาลานเซอร์ ที่อยู่ IP เครือข่ายท้องถิ่นเสมือน (VLAN) และชุดซอฟต์แวร์ คำจำกัดความของการประมวลผลแบบคลาวด์ของ NIST อธิบาย IaaS ว่า "โดยที่ผู้บริโภคสามารถติดตั้งและรันซอฟต์แวร์ที่กำหนดเองได้ ซึ่งอาจรวมถึงระบบปฏิบัติการและแอปพลิเคชันต่างๆ ผู้บริโภคไม่ได้จัดการหรือควบคุมโครงสร้างพื้นฐานคลาวด์ที่ซ่อนอยู่ แต่สามารถควบคุมระบบปฏิบัติการ พื้นที่เก็บข้อมูล และปรับใช้แอปพลิเคชัน และอาจมีการควบคุมส่วนประกอบเครือข่ายที่เลือกอย่างจำกัด (เช่น ไฟร์วอลล์โฮสต์)" ผู้ให้บริการคลาวด์ IaaS จัดหาทรัพยากรเหล่านี้ตามความต้องการจากแหล่งอุปกรณ์ขนาดใหญ่ที่ติดตั้งในศูนย์ข้อมูล สำหรับการเชื่อมต่อบริเวณกว้าง ลูกค้าสามารถใช้อินเทอร์เน็ตหรือระบบคลาวด์ของผู้ให้บริการ (เครือข่ายส่วนตัวเสมือนเฉพาะ) ในการปรับใช้แอปพลิเคชัน ผู้ใช้คลาวด์จะติดตั้งอิมเมจระบบปฏิบัติการและซอฟต์แวร์แอปพลิเคชันบนโครงสร้างพื้นฐานคลาวด์ ในรุ่นนี้ ผู้ใช้ระบบคลาวด์จะแพตช์และบำรุงรักษาระบบปฏิบัติการและแอพพลิเคชั่นซอฟต์แวร์ โดยทั่วไปผู้ให้บริการคลาวด์จะเรียกเก็บเงินบริการ IaaS บนพื้นฐานการประมวลผลแบบอรรถประโยชน์: ต้นทุนสะท้อนถึงจำนวนทรัพยากรที่จัดสรรและใช้ไป

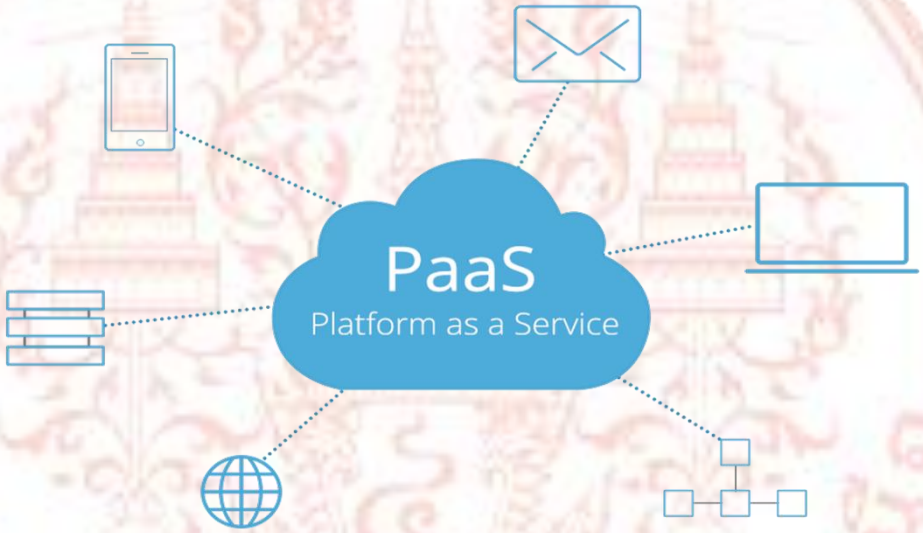


ภาพที่ 2-13 Infrastructure as a service

2.4.1.2 Platform as a service (PaaS)

คำจำกัดความของการประมวลผลแบบคลาวด์ของ NIST กำหนดแพลตฟอร์มว่าเป็นบริการดังนี้ ความสามารถที่มอบให้กับผู้บริโภคคือการปรับใช้บนโครงสร้างพื้นฐานคลาวด์แอปพลิเคชันที่ผู้บริโภคสร้างขึ้นหรือได้มาซึ่งสร้างขึ้นโดยใช้ภาษาการเขียนโปรแกรม ไลบรารี บริการ และเครื่องมือที่ผู้ให้บริการสนับสนุน ผู้บริโภคไม่ได้จัดการหรือควบคุมโครงสร้างพื้นฐานคลาวด์พื้นฐาน รวมถึงเครือข่าย เซิร์ฟเวอร์ ระบบปฏิบัติการ หรือที่เก็บข้อมูล แต่สามารถควบคุมแอปพลิเคชันที่ปรับใช้และการตั้งค่าการกำหนดค่าสำหรับสภาพแวดล้อมการโฮสต์แอปพลิเคชั่น ผู้จำหน่าย PaaS เสนอสภาพแวดล้อมการพัฒนาให้กับนักพัฒนาแอปพลิเคชัน โดยทั่วไปผู้ให้บริการจะพัฒนาชุดเครื่องมือและมาตรฐานสำหรับการพัฒนาและช่องทางในการจัดจำหน่ายและการชำระเงิน ในโมเดล PaaS ผู้ให้บริการคลาวด์จะมอบแพลตฟอร์มคอมพิวเตอร์ ซึ่งโดยทั่วไปจะรวมถึงระบบปฏิบัติการ สภาพแวดล้อมการดำเนินการด้วยภาษาโปรแกรม ฐานข้อมูล และเว็บเซิร์ฟเวอร์ นักพัฒนาแอปพลิเคชันพัฒนาและใช้งานซอฟต์แวร์ของตนบนแพลตฟอร์มคลาวด์ แทนที่จะซื้อและจัดการเลเยอร์ฮาร์ดแวร์และซอฟต์แวร์ที่เกี่ยวข้องโดยตรง ด้วย PaaS บางตัว คอมพิวเตอร์พื้นฐานและทรัพยากรการจัดเก็บข้อมูลจะปรับขนาดโดยอัตโนมัติเพื่อให้ตรงกับความต้องการของแอปพลิเคชัน เพื่อให้ผู้ใช้ระบบคลาวด์ไม่ต้องจัดสรรทรัพยากรด้วยตนเอง[41] [ต้องการใบเสนอราคาเพื่อตรวจสอบผู้ให้บริการบูรณาการและการจัดการข้อมูลบางรายยังใช้แอปพลิเคชันเฉพาะของ PaaS เป็น

แบบจำลองการส่งมอบข้อมูล ตัวอย่าง ได้แก่ iPaaS (แพลตฟอร์มการบูรณาการเป็นบริการ) และ dPaaS (แพลตฟอร์มข้อมูลเป็นบริการ) iPaaS ช่วยให้ลูกค้าสามารถพัฒนา ดำเนินการและควบคุมโฟลว์การบูรณาการได้ภายใต้โมเดลการรวม iPaaS ลูกค้าขับเคลื่อนการพัฒนาและการปรับใช้การรวมระบบโดยไม่ต้องติดตั้งหรือจัดการฮาร์ดแวร์หรือมิดเดิลแวร์ใดๆ dPaaS นำเสนอผลิตภัณฑ์การบูรณาการและการจัดการข้อมูลในรูปแบบบริการที่มีการจัดการเต็มรูปแบบภายใต้โมเดล dPaaS ผู้ให้บริการ PaaS ไม่ใช่ลูกค้า จะจัดการการพัฒนาและการดำเนินโปรแกรมโดยการสร้างแอปพลิเคชันข้อมูลสำหรับลูกค้า ผู้ใช้ dPaaS เข้าถึงข้อมูลผ่านเครื่องมือแสดงภาพข้อมูล



ภาพที่ 2-14 Platform as a Service

2.4.1.3 Software as a service (SaaS)

คำจำกัดความของการประมวลผลแบบคลาวด์ของ NIST กำหนดซอฟต์แวร์เป็นบริการเป็น:ความสามารถที่มอบให้แก่ผู้บริโภคคือการใช้แอปพลิเคชันของผู้ให้บริการที่ทำงานบนโครงสร้างพื้นฐานคลาวด์ แอปพลิเคชันสามารถเข้าถึงได้จากอุปกรณ์ไคลเอ็นต์ต่างๆ ผ่านทางอินเทอร์เฟซไคลเอ็นต์แบบบาง เช่น เว็บเบราว์เซอร์ (เช่น อีเมลบนเว็บ) หรืออินเทอร์เฟซโปรแกรม ผู้บริโภคไม่ได้จัดการหรือควบคุมโครงสร้างพื้นฐานคลาวด์พื้นฐาน รวมถึงเครือข่าย เซิร์ฟเวอร์ ระบบปฏิบัติการ พื้นที่เก็บข้อมูล หรือแม้แต่ความสามารถของแอปพลิเคชันแต่ละรายการ ยกเว้นการตั้งค่าการกำหนดค่าแอปพลิเคชันเฉพาะผู้ใช้ที่จำกัด ในซอฟต์แวร์ในรูปแบบบริการ (SaaS) ผู้ใช้จะสามารถเข้าถึงแอพพลิเคชั่นซอฟต์แวร์และฐานข้อมูลได้ ผู้ให้บริการคลาวด์จัดการโครงสร้างพื้นฐานและแพลตฟอร์มที่รันแอปพลิเคชัน

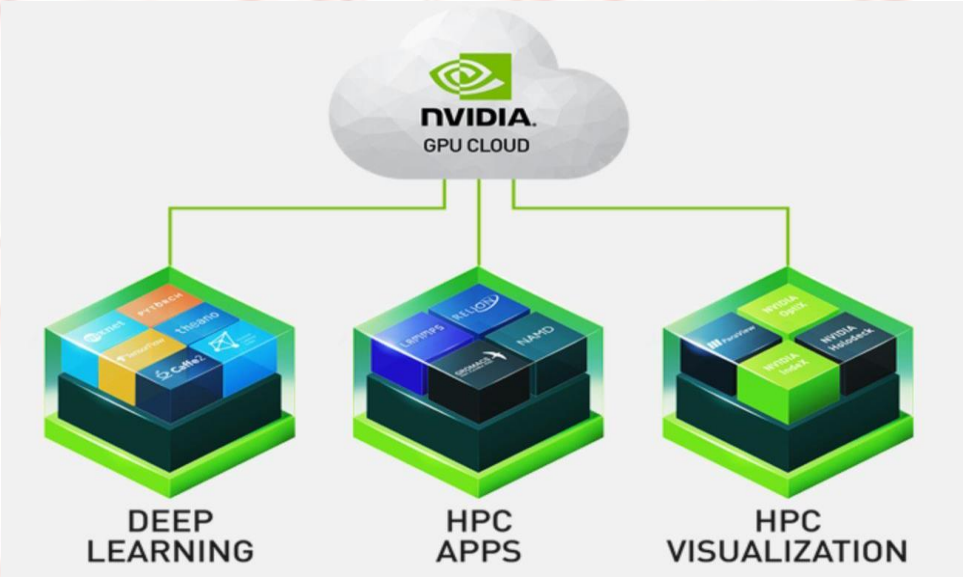
SaaS บางครั้งเรียกว่า "ซอฟต์แวร์ตามความต้องการ" และโดยปกติจะมีการกำหนดราคาแบบจ่ายต่อการใช้งานหรือค่าธรรมเนียมการสมัครสมาชิก ในโมเดล SaaS ผู้ให้บริการระบบคลาวด์จะติดตั้งและใช้งานซอฟต์แวร์แอปพลิเคชันในระบบคลาวด์ และผู้ใช้ระบบคลาวด์จะเข้าถึงซอฟต์แวร์จากไคลเอ็นต์ระบบคลาวด์ ผู้ใช้ระบบคลาวด์ไม่ได้จัดการโครงสร้างพื้นฐานระบบคลาวด์และแพลตฟอร์มที่แอปพลิเคชันทำงาน ซึ่งช่วยลดความจำเป็นในการติดตั้งและเรียกใช้แอปพลิเคชันบนคอมพิวเตอร์ของผู้ใช้ระบบคลาวด์ ซึ่งช่วยให้การบำรุงรักษาและการสนับสนุนทำได้ง่ายขึ้น แอปพลิเคชันระบบคลาวด์แตกต่างจากแอปพลิเคชันอื่นๆ ในเรื่องความสามารถในการปรับขนาด ซึ่งสามารถทำได้โดยการโคลนงานบนเครื่องเสมือนหลายเครื่อง ณ รันไทม์ เพื่อตอบสนองความต้องการงานที่เปลี่ยนแปลงไปโหลดบาลานเซอร์จะกระจายงานผ่านชุดเครื่องเสมือน กระบวนการนี้โปร่งใสสำหรับผู้ใช้ระบบคลาวด์ที่เห็นจุดเชื่อมต่อเพียงจุดเดียว เพื่อรองรับผู้ใช้คลาวด์จำนวนมาก แอปพลิเคชันบนคลาวด์สามารถมีหลายผู้เช่าได้ ซึ่งหมายความว่าเครื่องใดก็ตามอาจให้บริการได้มากกว่าหนึ่งองค์กรที่มีผู้ใช้คลาวด์ รูปแบบการกำหนดราคาสำหรับแอปพลิเคชัน SaaS โดยทั่วไปจะเป็นค่าธรรมเนียมรายเดือนหรือรายปีต่อผู้ใช้ ดังนั้นราคาจึงสามารถปรับขนาดและปรับได้หากมีการเพิ่มหรือลบผู้ใช้ ณ จุดใดก็ตาม มันอาจจะฟรีก็ได้ ผู้เสนออ้างว่า SaaS ช่วยให้ธุรกิจมีศักยภาพในการลดต้นทุนการดำเนินงานด้านไอทีโดยจ้างบุคคลภายนอกในการบำรุงรักษาฮาร์ดแวร์และซอฟต์แวร์และสนับสนุนผู้ให้บริการคลาวด์ ช่วยให้ธุรกิจสามารถจัดสรรต้นทุนการดำเนินงานด้านไอทีให้ห่างจากการใช้จ่ายด้านฮาร์ดแวร์/ซอฟต์แวร์ และจากค่าใช้จ่ายบุคลากร เพื่อบรรลุเป้าหมายอื่นๆ นอกจากนี้ ด้วยแอปพลิเคชันที่โฮสต์จากส่วนกลาง การอัปเดตจึงสามารถเผยแพร่ได้โดยที่ผู้ใช้ไม่จำเป็นต้องติดตั้งซอฟต์แวร์ใหม่ ข้อเสียเปรียบประการหนึ่งของ SaaS มาพร้อมกับการจัดเก็บข้อมูลของผู้ใช้บนเซิร์ฟเวอร์ของผู้ให้บริการคลาวด์ ด้วยเหตุนี้[ต้องการอ้างอิง] จึงมีการเข้าถึงข้อมูลโดยไม่ได้รับอนุญาตได้ [50] ตัวอย่างของแอปพลิเคชันที่นำเสนอในรูปแบบ SaaS ได้แก่ เกมและซอฟต์แวร์เพิ่มประสิทธิภาพการทำงาน เช่น Google Docs และ Office Online แอปพลิเคชัน SaaS อาจรวมเข้ากับที่เก็บข้อมูลบนคลาวด์หรือบริการโฮสต์ไฟล์ ซึ่งเป็นกรณีที่ Google Docs ถูกรวมเข้ากับ Google Drive และ Office Online ถูกรวมเข้ากับ OneDrive

รูปที่ 2.15 Software as a service

2.5 Graphics processing Unit (GPU)

หน่วยประมวลผลกราฟิกหรือ GPU (Graphics Processing Unit) เป็นเทคโนโลยีที่มีความสำคัญอย่างมากทั้งสำหรับคอมพิวเตอร์ส่วนบุคคลและสำหรับอุตสาหกรรมธุรกิจในปัจจุบัน การออกแบบ GPU ได้เน้นการประมวลผลแบบขนานที่สามารถทำงานพร้อมกันได้หลายงานพร้อมๆ กัน ซึ่งเหมาะสำหรับการใช้ในแอปพลิเคชันที่มีความต้องการในการประมวลผลที่ซับซ้อน เช่น กราฟิก และการเรนเดอร์วิดีโอที่ต้องการการประมวลผลที่รวดเร็วและมีประสิทธิภาพสูง นอกจากนี้ GPU ยังมีความสามารถที่ดีในการประมวลผลข้อมูลเชิงตัวเลขในการสร้างปัญญาประดิษฐ์ (AI) ซึ่งเป็นตัวอย่างหนึ่งที่ชัดเจนในการใช้งานที่กำลังได้รับความนิยมอย่างแพร่หลายในระดับโลก GPU ได้รับการพัฒนาขึ้นมาเพื่อใช้ในการเร่งการเรนเดอร์กราฟิก 3 มิติและการประมวลผลกราฟิกอื่นๆ โดยเฉพาะในอุตสาหกรรมเกม ซึ่งต้องการการแสดงผลที่สมจริงและเป็นภาพเพื่อประสบการณ์การเล่นที่ดีที่สุด โดย GPU สามารถปรับปรุงคุณภาพของเอฟเฟกต์ภาพและการจัดแสงเงาได้ให้มีความสมจริงมากยิ่งขึ้น นักพัฒนาซอฟต์แวร์กราฟิกได้ใช้ GPU เพื่อสร้างภาพที่มีความสมจริงและฉากที่น่าสนใจมากขึ้นด้วยเทคนิคการจัดแสงและเงาที่ขั้นสูง ทำให้ผู้ใช้สามารถได้รับประสบการณ์ที่สมจริงแบบเสมือนจริงได้อย่างใกล้เคียงกับความเป็นจริง นอกจากนี้ GPU ยังมีการนำไปใช้ในงานที่ต้องการประมวลผลที่มีประสิทธิภาพสูง (HPC) เช่น การเรียนรู้เชิงลึกและการคำนวณทางวิทยาศาสตร์ที่ซับซ้อน โดย GPU ช่วยลดเวลาในการฝึกโมเดลและการทำนายอย่างมีประสิทธิภาพมากขึ้น นักวิทยาศาสตร์และนักวิจัยใช้

GPU เพื่อประมวลผลข้อมูลขนาดใหญ่และการวิเคราะห์ข้อมูลที่ซับซ้อน เช่น การพยากรณ์สภาพอากาศหรือการศึกษาการเมือง ทำให้สามารถทำนายและวิเคราะห์ข้อมูลได้โดยมีความแม่นยำและความเร็วที่สูงขึ้น การพัฒนา GPU ไม่เพียงแต่ช่วยเพิ่มประสิทธิภาพในการประมวลผลแบบขนานและกราฟิกเท่านั้น แต่ยังเป็นเครื่องมือสำคัญที่ช่วยขับเคลื่อนนวัตกรรมในหลายด้านของเทคโนโลยี ทำให้การพัฒนาและการวิจัยต่างๆ ได้รับประโยชน์มากมาย นอกจากนี้ยังมีผลต่อการพัฒนาซอฟต์แวร์ที่สามารถทำงานได้อย่างมีประสิทธิภาพและได้ผลลัพธ์ที่มีคุณภาพสูงขึ้นด้วยการใช้ GPU ในการประมวลผลทางวิทยาศาสตร์และงานวิจัยต่างๆ ที่มีความซับซ้อนและมีข้อมูลมากมาย

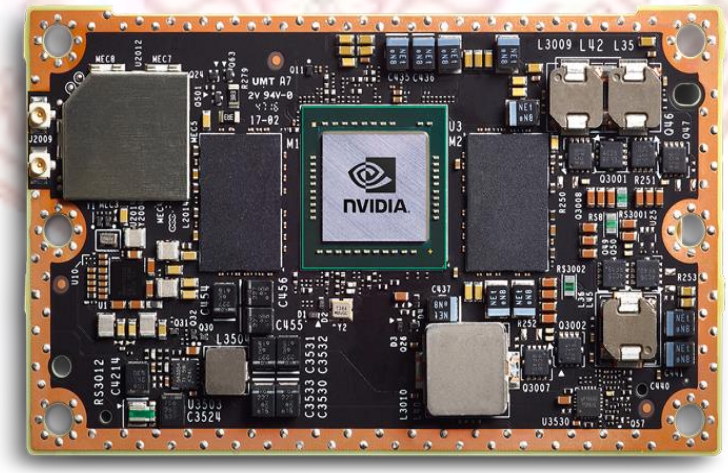


ภาพที่ 2-16 GPU Cloud

2.5.1 GPU for deep learning

การประมวลผลด้วย GPU (Graphics Processing Unit) เป็นเทคโนโลยีที่สำคัญและได้รับความนิยมอย่างมากในการใช้งานทางด้านการเรียนรู้เชิงลึก (Deep Learning) โดยเฉพาะเมื่อเทียบกับการใช้งานบน CPU (Central Processing Unit) ทั่วไป โดย GPU เป็นอุปกรณ์ที่ออกแบบมาเพื่อประมวลผลกราฟิกและภาพเคลื่อนไหวที่ซับซ้อนอย่างรวดเร็ว ซึ่งมีโครงสร้างแบบพิเศษที่ช่วยให้สามารถประมวลผลข้อมูลพร้อมกันได้มากมายพร้อมๆ กัน ทำให้เหมาะสำหรับการทำงานที่ต้องการคำนวณสูงและซับซ้อนเช่น Deep Learning ซึ่งมีการฝึกอบรมและการทดสอบโมเดลที่ต้องการประมวลผลจำนวนมากพร้อมกันในขั้นตอนการเรียนรู้ การประมวลผลด้วย GPU เหมาะสำหรับ Deep Learning เนื่องจาก GPU มีความสามารถในการทำงานแบบขนานสูง ซึ่งหมายความว่า GPU สามารถทำงานพร้อมกันได้หลายกระบวนการ

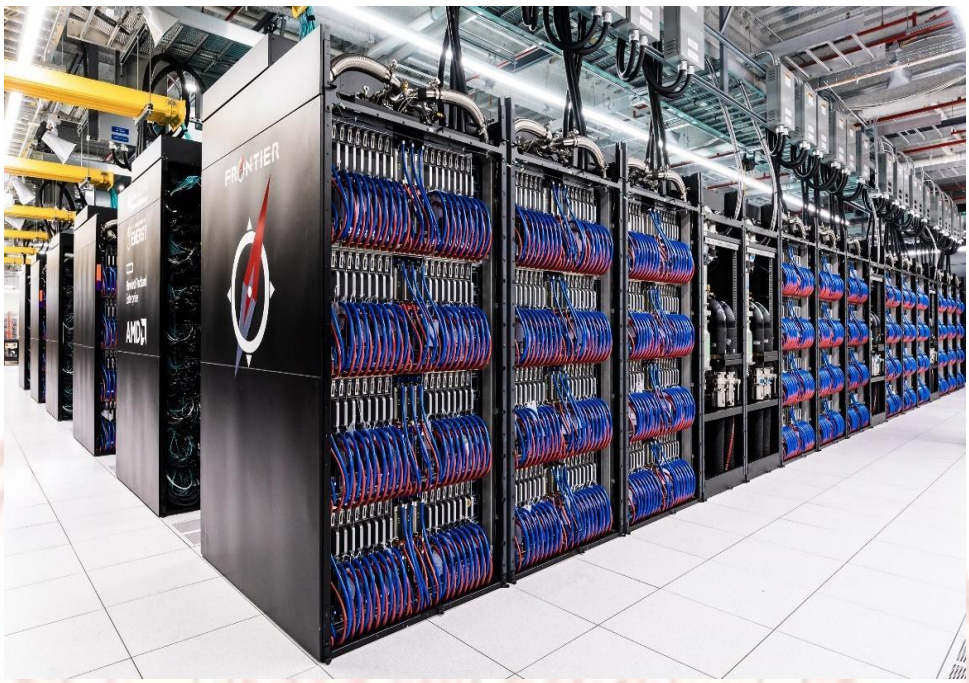
ในเวลาเดียวกัน ส่งผลให้เวลาในการฝึกโมเดลที่มีข้อมูลมากหรือมีความซับซ้อนสูงลดลงอย่างมาก เมื่อเทียบกับการใช้งานบน CPU ที่มีความสามารถในการทำงานแบบขนานน้อยกว่า การประมวลผลด้วย GPU ช่วยลดเวลาในการฝึกโมเดลจากเดือนหลายๆ เดือนลงมาเป็นเพียงสัปดาห์หรือวันเดียวได้ นอกจากนี้ GPU ยังมีความสามารถในการจัดการกับข้อมูลที่มีขนาดใหญ่ได้อย่างมีประสิทธิภาพ ซึ่งเป็นสิ่งสำคัญในการทำงานกับชุดข้อมูลที่ใหญ่ขึ้นทุกวันนี้ เช่น ข้อมูลทางการแพทย์ ข้อมูลภาพถ่ายทางดาวเคราะห์ หรือการประมวลผลสื่อสารมวลชน (Social Media) ที่มีปริมาณข้อมูลที่มากขึ้นอย่างต่อเนื่อง GPU ช่วยให้สามารถดำเนินการต่อข้อมูลเหล่านี้ได้อย่างมีประสิทธิภาพและรวดเร็ว ซึ่งเป็นประโยชน์อย่างมากในการพัฒนาและปรับปรุงโมเดล Deep Learning ให้มีประสิทธิภาพมากยิ่งขึ้นในด้านความสามารถทางเทคโนโลยี การประมวลผลด้วย GPU ยังสามารถนำไปใช้งานในหลากหลายงานที่ไม่ใช่เฉพาะในด้าน Deep Learning เท่านั้น เช่น การประมวลผลทางการแพทย์เพื่อการวินิจฉัยโรคหรือการวิเคราะห์ภาพทางดาวเคราะห์ ที่ต้องการการประมวลผลที่มีประสิทธิภาพสูง นอกจากนี้ ยังมีการประยุกต์ใช้ GPU ในงานวิจัยทางวิทยาศาสตร์ เช่น การจำลองทางคลินิกและการพยากรณ์ทางอากาศ ซึ่งการใช้ GPU ทำให้สามารถทำงานได้อย่างรวดเร็วและมีประสิทธิภาพ สรุปได้ว่า การประมวลผลด้วย GPU เป็นเทคโนโลยีที่มีความสำคัญและได้รับความนิยมอย่างมากในการใช้งานทางด้าน Deep Learning และงานที่ต้องการการคำนวณที่ซับซ้อนและมากมาย ทำให้เกิดประสิทธิภาพในการฝึกและปรับปรุงโมเดลอย่างมีประสิทธิภาพมากขึ้น นอกจากนี้ ยังสามารถประยุกต์ใช้ในหลากหลายงานที่ต้องการความสามารถในการประมวลผลที่มีประสิทธิภาพสูงได้อย่างมีประสิทธิภาพ



ภาพที่ 2-17 GPU for Deeplearning

2.5.2 HPC APPS

การประมวลผลเชิงสูงหรือ High Performance Computing (HPC) คือการใช้เทคโนโลยีคอมพิวเตอร์ที่มีการคำนวณที่รวดเร็วและมีประสิทธิภาพสูง เพื่อแก้ไขปัญหาที่มีความซับซ้อนและต้องการการคำนวณที่มากมาย ซึ่งไม่สามารถทำได้ด้วยคอมพิวเตอร์ทั่วไป โดยประการสำคัญ HPC นำไปใช้ในหลากหลายงานที่ต้องการการคำนวณแบบขนาน ๆ ซึ่งเป็นเทคโนโลยีที่สำคัญและมีการพัฒนาอย่างต่อเนื่องในปัจจุบัน หนึ่งในการประยุกต์ใช้งาน HPC ที่สำคัญคือในงานวิจัยทางวิทยาศาสตร์และวิศวกรรม เช่น การจำลองและการคำนวณทางวิทยาศาสตร์ที่ต้องการการประมวลผลที่ซับซ้อน เช่น การจำลองทางวัตถุประสงค์ทางชีวภาพเพื่อศึกษาภาพรวมของโครงสร้างโปรตีนหรือการทดสอบยา เรียกว่า Molecular Dynamics Simulation ที่ต้องการการคำนวณที่มีประสิทธิภาพสูงเพื่อศึกษาพฤติกรรมของอนุภาคต่าง ๆ ในระดับมาโมเลกุล การศึกษาและวิจัยทางวัฒนธรรมก็เป็นอีกส่วนหนึ่งที่ใช้ HPC อย่างแพร่หลาย เช่น การศึกษาประวัติศาสตร์ด้านวัฒนธรรมของมนุษย์ การประมวลผลข้อมูลทางสังคม การศึกษาด้านความเชื่อและพฤติกรรมของกลุ่มคน การคำนวณทางวิจัยด้านภูมิศาสตร์และสภาพภูมิอากาศ เพื่อทำนายและวิเคราะห์ข้อมูลที่เกี่ยวข้องกับสิ่งแวดล้อมและภัยพิบัติ การประมวลผลที่มีประสิทธิภาพสูงยังมีความสำคัญในงานวิจัยทางการแพทย์ เช่น การศึกษาและวิจัยในด้านยา เพื่อค้นหายาใหม่หรือการพัฒนายาที่มีประสิทธิภาพสูง การจำลองและการคำนวณทางชีวภาพ เพื่อศึกษาสมบัติของยาและตัวเสพติด หรือการพัฒนาเทคโนโลยีทางการแพทย์เชิงลึก เช่น การพัฒนาเทคนิคที่ใช้ AI ในการวินิจฉัยโรคหรือการทำนายผลการรักษาของโรคต่าง ๆ นอกจากนี้ ในอุตสาหกรรมเทคโนโลยีสารสนเทศและการสื่อสาร การใช้ HPC เป็นสิ่งที่สำคัญเพื่อการพัฒนาแอพลิเคชันและบริการใหม่ ๆ ที่ต้องการประมวลผลข้อมูลที่มีปริมาณมากและทันใจ เช่น การพัฒนาแพลตฟอร์มการค้าอิเล็กทรอนิกส์ที่ใช้งานง่ายและมีประสิทธิภาพ การทำนายทางธุรกิจด้วยการวิเคราะห์ข้อมูลทางการตลาด เป็นต้น สรุปได้ว่า HPC หรือ High Performance Computing เป็นเทคโนโลยีที่สำคัญและมีความสำคัญมากในหลากหลายสาขาอุตสาหกรรมและการวิจัย โดยเฉพาะในงานที่ต้องการคำนวณที่ซับซ้อนและมากมาย เช่น วิจัยทางวิทยาศาสตร์และวิศวกรรม การแพทย์ วิจัยทางการแพทย์ และอุตสาหกรรมเทคโนโลยีสารสนเทศและการสื่อสาร

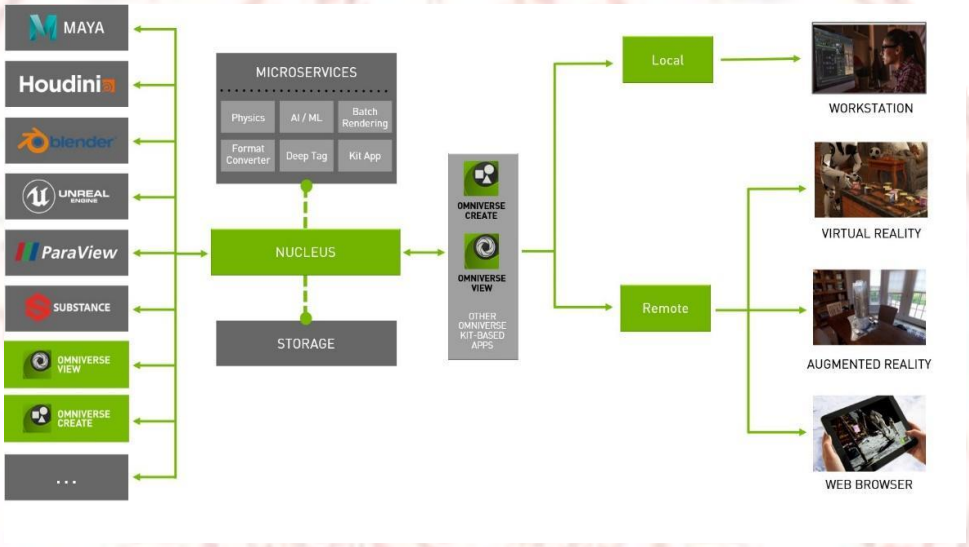


ภาพที่ 2-18 HPC

2.5.3 HPC VISUALIZATION

การแสดงผลทางความสูงของ HPC หรือ High Performance Computing เป็นกระบวนการที่มีความสำคัญอย่างมากในการสร้างภาพประกอบหรือข้อมูลที่ซับซ้อนให้เป็นรูปแบบที่เข้าใจง่ายและมีประสิทธิภาพสำหรับการวิเคราะห์และการตัดสินใจในหลากหลายสาขาอุตสาหกรรมและวิชาการต่าง ๆ ที่ใช้งาน HPC ได้แก่ วิทยาศาสตร์, วิศวกรรม, การแพทย์, ภูมิสารสนเทศและการสื่อสาร, และการพัฒนาโซลูชันในอุตสาหกรรมที่ต้องการความแม่นยำและประสิทธิภาพสูงในการประมวลผลข้อมูลที่มากมาย การใช้ HPC ในงานแสดงผลทางความสูง (Visualization) เป็นเครื่องมือที่สำคัญในการสร้างภาพที่มีความซับซ้อนจากข้อมูลที่ได้รับจากการคำนวณ HPC ซึ่งสามารถนำมาใช้ในหลากหลายรูปแบบ เช่น การแสดงผลข้อมูลทางวิทยาศาสตร์ที่ได้จากการจำลองทางชีวภาพหรือการจำลองโครงสร้างทางโมเลกุล เพื่อให้นักวิทยาศาสตร์และนักวิจัยสามารถทำความเข้าใจและวิเคราะห์ข้อมูลได้อย่างรวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น ในงานวิศวกรรมและการออกแบบผลิตภัณฑ์ เครื่องมือการแสดงผล HPC มีความสำคัญอย่างมากในการพัฒนาผลิตภัณฑ์ใหม่ ๆ และการทดสอบความคงทนของวัสดุ การใช้ HPC ช่วยในการสร้างและประเมินผลิตภัณฑ์ให้กับตลาดอย่างที่สุด ซึ่งการแสดงผลที่มีความคล้ายคลึงกับผลิตภัณฑ์จริงช่วยให้นักออกแบบและวิศวกรสามารถทำการปรับปรุงและ

ปรับใช้งานได้อย่างมีประสิทธิภาพ ในด้านการแพทย์และวิทยาศาสตร์ชีวภาพ การแสดงผล HPC เป็นเครื่องมือที่สำคัญในการศึกษาและวิเคราะห์ข้อมูลทางการแพทย์ เช่น ภาพถ่ายทางการแพทย์, การทดสอบยาและการวินิจฉัยโรค ที่ต้องการการแสดงผลที่ถูกต้องและคล้ายคลึงกับสภาพจริง เพื่อให้แพทย์และนักวิจัยสามารถทำนายและวิเคราะห์ผลได้อย่างถูกต้อง สุดท้ายนี้ การใช้ HPC ในการแสดงผลยังมีบทบาทสำคัญในอุตสาหกรรมภาพยนตร์และสื่อสารสื่อสาร การใช้เทคโนโลยี HPC ช่วยในการสร้างภาพเคลื่อนไหวที่สมจริงและมีความสมจริงสูง การประมวลผลกราฟิกที่ซับซ้อนและการจัดแสงที่มีความซับซ้อน เพื่อให้ภาพมีความสมจริงและน่าสนใจมากยิ่งขึ้น



ภาพที่ 2-19 HPC Visualization

2.5.4 Nvidia T4

NVIDIA® T4 GPU เป็นหนึ่งในเทคโนโลยี GPU ที่ได้รับความนิยมอย่างมากในโลกของคอมพิวเตอร์ มันถูกออกแบบมาเพื่อรองรับหลายประเภทของการประมวลผลที่ต้องการประสิทธิภาพสูงและการใช้พลังงานที่มีประสิทธิภาพในระดับสูง ด้วยสถาปัตยกรรม NVIDIA Turing™ ล่าสุดและการรองรับแพลตฟอร์ม PCIe ขนาดเล็กที่ใช้พลังงานเพียง 70 วัตต์ T4 GPU มีความสามารถในการปรับปรุงประสิทธิภาพในหลายด้านอย่างมาก โดยเฉพาะในงานการเรียนรู้เชิงลึก, การฝึกอบรมและการอนุมาน, การวิเคราะห์ข้อมูล, และงานกราฟิกต่าง ๆ NVIDIA T4 มาพร้อมกับ Tensor Cores ที่ออกแบบมาเพื่อการคำนวณที่แม่นยำและเร่งความเร็วของงานที่ต้องการการคำนวณที่ซับซ้อน เช่น การประมวลผลทางวิทยาศาสตร์, การสร้างแบบจำลอง AI, หรือการคำนวณทางการแพทย์ที่ต้องการความสมจริงและความรวดเร็วในการตอบสนอง

นอกจากนี้ T4 ยังมี RT Cores ที่ช่วยในการสร้างภาพที่มีความสมจริงและการแสดงผลที่ทันสมัย ทำให้มันเป็นเครื่องมือที่สำคัญสำหรับงานกราฟิกและการวิเคราะห์ภาพที่ต้องการความคล้ายคลึงกับสภาพจริงอย่างสูง เมื่อเปรียบเทียบกับ GPU อื่น ๆ ในตระกูลเดียวกัน NVIDIA T4 มีความสามารถในการปรับปรุงประสิทธิภาพและการใช้พลังงานที่ดีกว่าอย่างมาก เนื่องจากมีการออกแบบให้เหมาะสมสำหรับการใช้งานที่ต้องการการประมวลผลกระแสหลักอย่างมาก เช่น ในงานวิทยาศาสตร์, วิจัยและพัฒนาที่ต้องการการคำนวณที่สูงและเร่งการพัฒนาผลิตภัณฑ์ และการใช้งานในระบบธุรกิจที่ต้องการความเร็วและความแม่นยำในการวิเคราะห์ข้อมูล ทั้งนี้ NVIDIA T4 ยังมีความสามารถในการรองรับแพลตฟอร์มการพัฒนาที่หลากหลายอย่างมาก เช่น การใช้งานร่วมกับซอฟต์แวร์คอนเทนเนอร์ที่มีการเร่งความเร็วจาก NGC (NVIDIA GPU Cloud) ซึ่งทำให้ผู้ใช้สามารถเข้าถึงและนำเสนอการพัฒนาแอปพลิเคชันและการประมวลผลข้อมูลที่มีประสิทธิภาพสูงได้อย่างมีประสิทธิภาพ

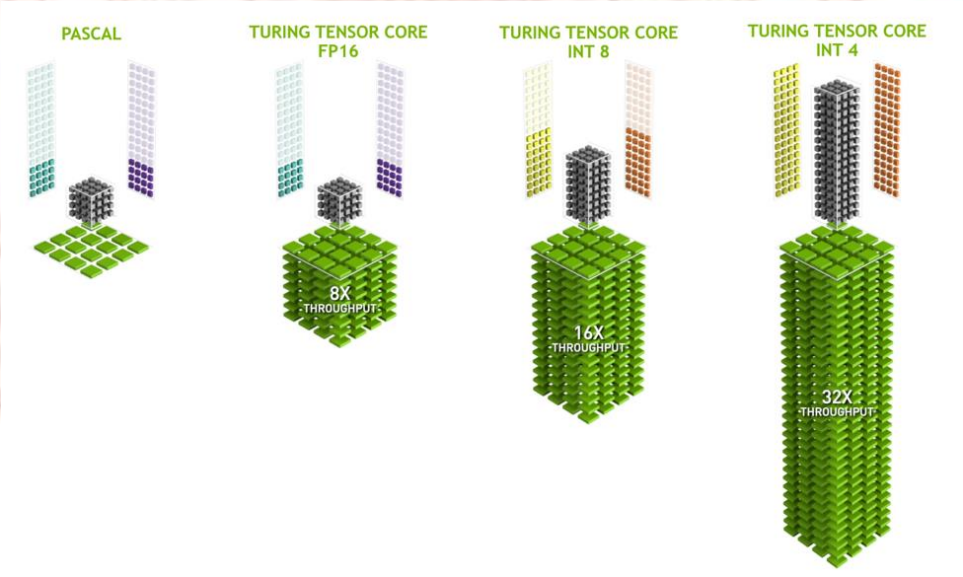
ภาพที่ 2-20 Nvidia T4

(ที่มา: advantech.com)

2.5.5 Tensor cores

Tensor Cores เป็นเทคโนโลยีที่ถูกพัฒนาขึ้นโดย NVIDIA เพื่อใช้ในหน่วยประมวลผลกราฟิกเพื่อการประมวลผลที่เกี่ยวข้องกับการเรียนรู้เชิงลึกและการคำนวณที่ใช้เทนเซอร์ (tensor) ในการทำงาน โดยเฉพาะการคำนวณทางคณิตศาสตร์ที่ซับซ้อนที่ใช้ในการฝึกอบรมโมเดลเรียนรู้เชิงลึกที่มีข้อมูลขนาดใหญ่ และการประมวลผลทางวิทยาศาสตร์เชิง

คำนวณที่ซับซ้อน การทำงานของ Tensor Cores มุ่งเน้นไปที่ความสามารถในการทำคำนวณเมทริกซ์ที่มีขนาดใหญ่อย่างมีประสิทธิภาพสูง โดย Tensor Cores สามารถทำการคูณเมทริกซ์ขนาด 4x4 ได้โดยมีความแม่นยำและประสิทธิภาพที่สูงกว่าการคำนวณแบบดั่งเดิม การใช้งานนี้มีความสำคัญอย่างมากในงานที่ต้องการความเร็วในการประมวลผลเชิงลึกเช่น การฝึกอบรมและประมวลผลข้อมูลที่ซับซ้อน ที่ต้องการการคำนวณที่แม่นยำและเร็วกว่าเพื่อให้สามารถวิเคราะห์และนำข้อมูลไปใช้ประโยชน์ได้อย่างทันท่วงที เทคโนโลยี Tensor Cores นั้นได้รับการปรับปรุงให้มีประสิทธิภาพและความสามารถที่ดีขึ้นตามการพัฒนาและการใช้งานในฟิลด์ต่างๆ โดยมีการนำเอา Tensor Cores มาประยุกต์ใช้ในหลายงาน เช่น ในด้านการประมวลผลทางวิทยาศาสตร์เชิงคำนวณที่ซับซ้อน เพื่อการวิเคราะห์ข้อมูลที่มีความซับซ้อน เพื่อให้ได้ข้อมูลที่มีความแม่นยำและเชื่อถือได้มากยิ่งขึ้น นอกจากนี้ การนำเทคโนโลยี Tensor Cores ไปใช้ในงานประมวลผลที่ต้องการความรวดเร็วสูงยังสามารถช่วยลดเวลาในการฝึกอบรมและพัฒนาโมเดลเชิงลึกได้มากขึ้น เนื่องจากมีการทำคำนวณที่เร็วและมีประสิทธิภาพสูงในการประมวลผลข้อมูลที่ซับซ้อน ดังนั้น เทคโนโลยี Tensor Cores จึงเป็นสิ่งที่สำคัญและมีความสำคัญต่อการพัฒนาและปรับปรุงงานต่างๆ ที่ต้องการความเร็วและประสิทธิภาพในการประมวลผลที่มีความซับซ้อน



ภาพที่ 2-21 Tensor cores

(ที่มา: nvidia.coma)

2.6 Risk assessment

วิธีการประเมินความเสี่ยงภายในการจัดการความปลอดภัยของกระบวนการ (PSM) โดยเน้นย้ำถึงความสำคัญของการตรวจจับอันตราย ความน่าเชื่อถือของระบบควบคุม และการวิเคราะห์ความปลอดภัยในการทำงาน (JSA) ในสภาพแวดล้อมทางอุตสาหกรรม เช่น โรงงานเคมีและโรงงานปิโตรเคมี นอกจากนี้ ยังมีการสำรวจการผสานรวมเครื่องมือที่ใช้ปัญญาประดิษฐ์ (AI) โดยเฉพาะโมเดลที่ใช้ BERT สำหรับการทำงานอัตโนมัติของ JSA เพื่อปรับปรุงแนวทางการจัดการความเสี่ยง

2.6.1 Process Safety Management (PSM) Framework

การจัดการความปลอดภัยของกระบวนการ (PSM) เป็นกรอบงานเชิงระบบที่ออกแบบมาเพื่อป้องกันอุบัติเหตุที่เกี่ยวข้องกับสารเคมีอันตรายในโรงงานอุตสาหกรรม เช่น โรงงานเคมี โรงกลั่น และอุตสาหกรรมปิโตรเคมี เป้าหมายหลักของ PSM คือการลดความเสี่ยงที่เกี่ยวข้องกับวัสดุอันตราย ปกป้องพนักงานและสิ่งแวดล้อม และรับรองความต่อเนื่องของการดำเนินงาน หน่วยงานกำกับดูแล เช่น สำนักงานความปลอดภัยและอาชีวอนามัย (OSHA) ศูนย์ความปลอดภัยกระบวนการทางเคมี (CCPS) และสถาบันปิโตรเลียมแห่งอเมริกา (API) กำหนดแนวทางสำหรับการนำ PSM มาใช้ กรอบงานนี้บูรณาการการประเมินความเสี่ยง การระบุอันตราย การควบคุมกระบวนการ และการเตรียมพร้อมรับมือเหตุฉุกเฉินเข้าไว้ในระบบการจัดการที่มีโครงสร้าง

2.6.1 Risk Assessment Methodologies

การประเมินความเสี่ยงเป็นองค์ประกอบสำคัญของการจัดการความปลอดภัยของกระบวนการ (PSM) ช่วยให้ภาคอุตสาหกรรมสามารถระบุ ประเมิน และบรรเทาอันตรายที่อาจเกิดขึ้นได้ มีการนำวิธีการต่างๆ มาใช้ในการประเมินความเสี่ยงในโรงงานเคมี โรงกลั่นน้ำมัน และการดำเนินการทางอุตสาหกรรม เพื่อให้แน่ใจว่าบุคลากร อุปกรณ์ และสิ่งแวดล้อมจะได้รับความปลอดภัย

2.6.1.1 Hazard and Operability Study (HAZOP)

HAZOP เป็นวิธีการที่เป็นระบบในการระบุความเบี่ยงเบนของกระบวนการและอันตรายที่อาจเกิดขึ้น โดยเกี่ยวข้องกับทีมสหวิชาชีพที่วิเคราะห์พารามิเตอร์ของกระบวนการ (เช่น แรงดัน อุณหภูมิ) เพื่อตรวจจับความเสี่ยงและแนะนำมาตรการควบคุม

2.6.1.2 Failure Mode and Effects Analysis (FMEA)

FMEA ประเมินโหมดความล้มเหลวที่อาจเกิดขึ้นของอุปกรณ์และกระบวนการ โดยประเมินผลกระทบต่อความปลอดภัยและความน่าเชื่อถือ โดยจัดลำดับความสำคัญของความเสี่ยงตามความรุนแรง การเกิดขึ้น และความน่าจะเป็นในการตรวจจับ เพื่อแนะนำแนวทางในการป้องกัน

2.6.1.3 Job Safety Analysis (JSA)

การวิเคราะห์ความปลอดภัยในการทำงาน (Job Safety Analysis: JSA) เป็นกระบวนการเชิงระบบที่เน้นที่การระบุและประเมินอันตรายที่อาจเกิดขึ้นที่เกี่ยวข้องกับงานเฉพาะ โดยเกี่ยวข้องกับการแบ่งงานออกเป็นขั้นตอนแต่ละขั้นตอน การวิเคราะห์ความเสี่ยงที่อาจเกิดขึ้นในแต่ละขั้นตอน และการนำมาตรการควบคุมที่เหมาะสมมาใช้เพื่อลดข้อผิดพลาดของมนุษย์และป้องกันอุบัติเหตุในสถานที่ทำงาน มาตรการควบคุมเหล่านี้อาจรวมถึงการควบคุมทางวิศวกรรม เช่น อุปกรณ์ป้องกันเครื่องจักรหรือระบบระบายอากาศ การควบคุมด้านการบริหาร เช่น ขั้นตอนด้านความปลอดภัยและโปรแกรมการฝึกอบรม และอุปกรณ์ป้องกันส่วนบุคคล (PPE) เช่น หมวกกันน็อค ถุงมือ หรือเครื่องช่วยหายใจ การดำเนินการ JSA ช่วยให้องค์กรต่างๆ สามารถเพิ่มความปลอดภัยในสถานที่ทำงาน ปรับปรุงประสิทธิภาพการทำงาน และรับรองการปฏิบัติตามกฎระเบียบด้านอาชีวอนามัยและความปลอดภัย

2.6.1.4 Quantitative Risk Assessment (QRA)

QRA ใช้การสร้างแบบจำลองทางคณิตศาสตร์และข้อมูลทางสถิติเพื่อประมาณความน่าจะเป็นและผลที่ตามมาของเหตุการณ์อันตราย โดยจัดทำกรอบการตัดสินใจตามความเสี่ยงเพื่อความปลอดภัยของกระบวนการ

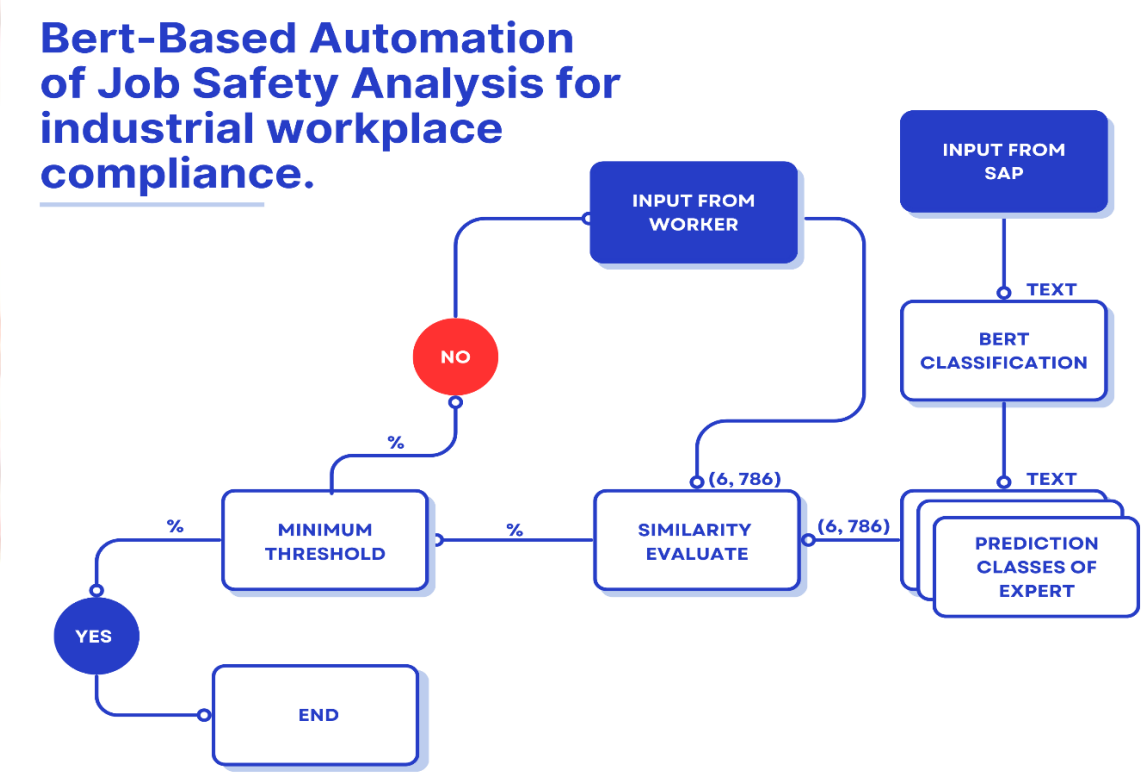
2.6.1.5 Layer of Protection Analysis (LOPA)

LOPA ระบุชั้นความปลอดภัยอิสระที่ป้องกันความล้มเหลวร้ายแรง และประเมินประสิทธิภาพของอุปสรรคด้านความปลอดภัย เช่น ระบบปิดระบบอัตโนมัติและแผนตอบสนองเหตุฉุกเฉินด้วยการบูรณาการวิธีการเหล่านี้ อุตสาหกรรมต่างๆ สามารถพัฒนากลยุทธ์การบรรเทาความเสี่ยงที่แข็งแกร่ง รับรองการปฏิบัติตามกฎระเบียบและความปลอดภัยในการปฏิบัติงาน

บทที่ 3
วิธีการดำเนินงาน

การวิจัยนี้ใช้โมเดล BERT และ SBERT ที่ฝึกด้วยการเรียนรู้เชิงลึกเพื่อประเมินความเสี่ยงในอุตสาหกรรมน้ำมันและก๊าซอย่างเป็นระบบ โดยจัดการข้อมูลอย่างมีระเบียบและวิเคราะห์จุดข้อมูลสำคัญ โมเดลสามารถแยกข้อมูลที่เกี่ยวข้องจากชุดข้อมูลซับซ้อนได้ การพัฒนาและทดสอบบน Google Colab ด้วย Python ช่วยเพิ่มประสิทธิภาพ ลดข้อผิดพลาด และเพิ่มความแม่นยำ การผสานเทคโนโลยีนี้ช่วยให้สอดคล้องกับมาตรฐานความปลอดภัยและข้อกำหนดทางกฎหมายของอุตสาหกรรม

3.1 แผนผังแสดงลำดับการทำงาน



ภาพที่ 3-1 Flowchart

3.2 รูปแบบการจำแนกประเภท

การวิจัยนี้มุ่งเน้นการรวบรวมและวิเคราะห์ข้อมูลจากอุตสาหกรรมน้ำมันและก๊าซผ่านรหัส SAP ซึ่งช่วยให้มั่นใจได้ถึงความสอดคล้องและความน่าเชื่อถือของข้อมูล ข้อมูลที่ได้รับถูกนำไปใช้ในการวิเคราะห์ความเสี่ยง การคาดการณ์ความปลอดภัย และการจำแนกประเภทของงานที่เกี่ยวข้องกับ JSA เราใช้ข้อมูล 81,978 ตัวอย่าง โดยแบ่งเป็นชุดฝึกอบรม 80% และชุดทดสอบ 20% พร้อมนำ Google Colab มาใช้จำลองกระบวนการประเมิน SAP เพื่อลดข้อผิดพลาดและเพิ่มความแม่นยำ กระบวนการพัฒนาแบบจำลองครอบคลุมการประมวลผลข้อมูลล่วงหน้า การเลือกแบบจำลอง การประเมินผล และการบันทึกกระบวนการ เป้าหมายหลักของงานวิจัยคือการพัฒนาแบบจำลองทำนายความปลอดภัย ประเมินปัจจัยเสี่ยง และใช้ AI วิเคราะห์ข้อมูลเพื่อเพิ่มความปลอดภัยในโรงงาน การวิจัยนี้ได้รับการสนับสนุนจากเกณฑ์ความปลอดภัยที่กำหนดไว้ล่วงหน้าและข้อมูลเชิงลึกจากผู้เชี่ยวชาญในอุตสาหกรรมดังภาพที่ 3-2 รูปด้านล่าง

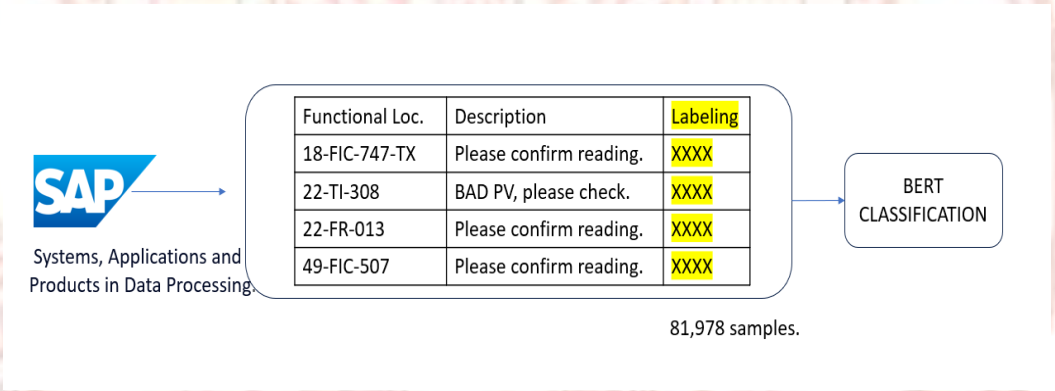
ภาพที่ 3-2 BERT CLASSIFICATION

3.2.1 BERT Classification datasets

การเตรียมข้อมูลสำหรับการวิเคราะห์และการประมวลผลด้วยโมเดล BERT ในงาน Classification นั้นเป็นขั้นตอนที่สำคัญ เนื่องจากความสำเร็จของโมเดลจะขึ้นอยู่กับคุณภาพของข้อมูลที่ใช้ในการฝึก ในกรณีของเรื่องราวที่กล่าวถึง ข้อมูลที่ใช้มาจากระบบ SAP ซึ่งได้จากการแจ้งงานของ Operator ในระบบ เป็นข้อมูลที่มีความหลากหลายและเป็นการเรียนรู้จากการดำเนินงานจริงที่เกิดขึ้นในองค์กร เมื่อได้รับข้อมูลมาแล้ว ขั้นตอนถัดไปคือการวิเคราะห์ข้อมูลเหล่านั้น เพื่อเตรียมข้อมูลให้พร้อมสำหรับการใช้ในโมเดล BERT โดยวิธีการวิเคราะห์อาจมีการทำความเข้าใจเนื้อหาของข้อความ การตัดคำ (tokenization) และการเตรียมข้อมูลใน

รูปแบบที่เหมาะสมสำหรับการใช้ในโมเดล BERT ซึ่งรวมถึงการแปลงข้อความเป็นรูปแบบของเวกเตอร์ที่เข้าใจง่ายสำหรับโมเดลทำนาย หลังจากที่ได้ทำการวิเคราะห์และเตรียมข้อมูลเสร็จสิ้นแล้ว ขั้นตอนต่อไปคือการทำการ Labeling ข้อมูล ซึ่งหมายถึงการกำหนดป้ายกำกับ (label) หรือคำตอบที่ถูกต้องสำหรับแต่ละข้อมูล เพื่อให้โมเดลสามารถเรียนรู้และทำนายผลลัพธ์ได้อย่างถูกต้อง การ Labeling นั้นเป็นกระบวนการที่ต้องใช้ความรอบคอบและความชัดเจน เนื่องจากมีผลต่อประสิทธิภาพของโมเดลที่ถูกสร้างขึ้น

หลังจากที่นำข้อมูลบน SAP มาทำการเรียบเรียงข้อมูลใหม่โดยการใช้พนักงานที่มีประสบการ์ณทำการจัดการข้อมูลและทำการติดฉลากข้อมูลซึ่งข้อมูลทั้งหมดจะถูกที่วิเคราะห์จากข้อความที่ Operator เป็นคนร้องขอจากนั้นจะทำการทำนายข้อความที่ถูกเขียนโดยผู้ร้องขอและหลังจากนั้นจะถูกส่งต่อไปยังโมเดล BERT Classification เผื่อทำการเลือกว่าแต่ละงานนั้นจะเหมาะสมกับฉลากไหนที่ผู้มีประสบการ์ณนั้นประเมินไว้ล่วงหน้าแล้วดังภาพที่ 3-3



ภาพที่ 3-3 Data of BERT Classification from SAP

3.2.2 BERT Classification Model Training

ในการใช้ BERT สำหรับงาน Classification เช่น การจำแนกประเภทข้อความ, การประเมินความเสี่ยง, หรือการประเมินความปลอดภัยในอุตสาหกรรมน้ำมันและก๊าซ เราจะใช้เทคนิคการ Fine-Tuning ซึ่งหมายถึงการนำโมเดล BERT ที่ผ่านการฝึกมาแล้วมาปรับใช้กับข้อมูลเฉพาะของงานนั้น ๆ เช่น ข้อมูลจาก SAP หรือ JSA โดยการฝึกเพิ่มเติมให้เหมาะสมกับงานจำแนกประเภทที่เราต้องการ ขั้นตอนหลักในการใช้ BERT สำหรับ Classification ได้แก่

3.2.2.1 Preprocessing

การแปลงข้อความให้เป็นรูปแบบที่ BERT สามารถประมวลผลได้ เช่น การแปลงคำเป็น Token และการตั้งค่าความยาวสูงสุดของข้อความ

3.2.2.1 Input Representation

การเตรียมข้อมูลที่ BERT สามารถใช้งานได้ โดยข้อมูลจะถูกแปลงเป็นตัวแทนเชิงตัวเลข (numerical representation) ที่ประกอบไปด้วย Token Embeddings, Segment Embeddings, และ Position Embeddings

3.2.2.2 Model Training

การฝึกโมเดล BERT บนข้อมูลที่มีการ Label (เช่น ประเภทของงาน หรือระดับความเสี่ยง) เพื่อให้โมเดลสามารถจำแนกประเภทข้อมูลได้

3.2.2.3 Output

การใช้ BERT ในการทำนายผลลัพธ์จากข้อมูลใหม่ที่ยังไม่เคยเห็นมาก่อน โดยจะให้คะแนนหรือระดับความเชื่อมั่นในแต่ละประเภท (classification label)

การวิจัยนี้เป็นการรวบรวมและวิเคราะห์ข้อมูลจากอุตสาหกรรมน้ำมันและก๊าซผ่านรหัส SAP ซึ่งช่วยให้มั่นใจได้ถึงความสอดคล้องและความน่าเชื่อถือของข้อมูล ข้อมูลที่ได้รับถูกนำไปใช้ในการวิเคราะห์ความเสี่ยง การคาดการณ์ความปลอดภัย และการจำแนกประเภทของงานที่เกี่ยวข้องกับ JSA เราใช้ข้อมูล 81,978 ตัวอย่าง และใช้พารามิเตอร์เริ่มต้น โดยแบ่งเป็นชุดฝึกอบรม 80% และชุดทดสอบ 20% พร้อมนำ Google Colab มาใช้จำลองกระบวนการประเมิน SAP เพื่อลดข้อผิดพลาดและเพิ่มความแม่นยำ กระบวนการพัฒนาแบบจำลองครอบคลุมการประมวลผลข้อมูลล่วงหน้า การเลือกแบบจำลอง การประเมินผล และการบันทึกกระบวนการเป้าหมายหลักของงานวิจัยคือการพัฒนาแบบจำลองทำนายความปลอดภัย ประเมินปัจจัยเสี่ยง และใช้ AI วิเคราะห์ข้อมูลเพื่อเพิ่มความปลอดภัยในโรงงาน การวิจัยนี้ได้รับการสนับสนุนจากเกณฑ์ความปลอดภัยที่กำหนดไว้ล่วงหน้าและข้อมูลเชิงลึกจากผู้เชี่ยวชาญในอุตสาหกรรม

3.2 การทดสอบการใช้คำพ้องความหมาย

เราทำการจำลองบทบาทของผู้เชี่ยวชาญ (Expert Role) และคนงาน (Worker Role) ในการประเมินความเสี่ยงและความปลอดภัยในโรงงานอุตสาหกรรม การทดสอบนี้มีวัตถุประสงค์เพื่อวัดความเหมือน (Similarity) ระหว่างข้อความที่ใช้ในงานประจำวันของคนงานกับคำแนะนำหรือข้อเสนอแนะจากผู้เชี่ยวชาญ โดยการใช้โมเดลประมวลผลภาษาธรรมชาติ เราสามารถวิเคราะห์และเปรียบเทียบคำพ้องความหมายเพื่อดูว่าคนงานสามารถเข้าใจและปฏิบัติตามคำแนะนำของผู้เชี่ยวชาญได้ดีเพียงใด ในกระบวนการทดสอบนี้ เราใช้ Google Colab ซึ่งเป็นแพลตฟอร์มการประมวลผลแบบคลาวด์ที่ให้การเข้าถึงทรัพยากรการคำนวณที่ทรงพลัง เช่น GPU และ TPU สิ่งนี้ทำให้เราสามารถคำนวณค่าความเหมือนและเปอร์เซ็นต์ของความปลอดภัยได้อย่างรวดเร็วและมีประสิทธิภาพ การทดสอบเหล่านี้ช่วยให้เราสามารถประเมินได้ว่าการสื่อสารระหว่างผู้เชี่ยวชาญและคนงานมีประสิทธิภาพเพียงใด และสามารถนำข้อมูลที่ได้ไปปรับปรุงกระบวนการทำงานเพื่อเพิ่มความปลอดภัยในสถานที่ทำงานได้อย่างไร

ตารางที่ 3-1 การทดสอบการใช้คำพ้องความหมาย

Roles	**Comparison sentence**	
	sentences	***Score (%)***
Expertise	The process malfunctions when PV is bad.	98.87649
Worker	The process fails when PV is faulty	

3.3 การประเมินของผู้เชี่ยวชาญและผู้ปฏิบัติงานเพื่อหาค่าต่ำที่สุด

เราได้ทำการประเมินความเสี่ยงและความปลอดภัยในโรงงานอุตสาหกรรมโดยการใช้ประสบการณ์และความชำนาญของผู้เชี่ยวชาญและผู้ปฏิบัติงานที่มีประสบการณ์ต่างกัน โดยให้ผู้เชี่ยวชาญที่มีประสบการณ์การทำงานมากกว่า 10 ปีทำการประเมินความเสี่ยงและความปลอดภัย ซึ่งถือว่าเป็นระดับสูงสุดของความชำนาญในบริบทนี้ จากนั้นนำค่าประเมินของผู้เชี่ยวชาญมาเปรียบเทียบกับค่าประเมินที่ได้จากพนักงานที่มีประสบการณ์ประมาณ 5 ปี การประเมินนี้ทำขึ้นเพื่อหาค่าความปลอดภัยที่ต่ำที่สุดซึ่งเป็นค่าที่ผู้ปฏิบัติงานสามารถประเมินได้

จากการมีประสบการณ์และความรู้ที่น้อยกว่าผู้เชี่ยวชาญ ผลการประเมินแสดงให้เห็นว่าค่าเฉลี่ยของความปลอดภัยที่พนักงานสามารถประเมินได้อยู่ที่ประมาณ 67.644905% การใช้การเปรียบเทียบนี้ทำให้เราเห็นถึงช่องว่างระหว่างการประเมินของผู้เชี่ยวชาญและผู้ปฏิบัติงาน ซึ่งสามารถนำข้อมูลที่ได้มาใช้ในการพัฒนาและฝึกอบรมเพื่อเพิ่มความแม่นยำในการประเมินความเสี่ยงและความปลอดภัยของพนักงานในอนาคตดังรูปที่ 3-4 ด้านล่าง

ภาพที่ 3-4 การใช้ผู้เชี่ยวชาญและพนักงานที่มีประสบการณ์ผ่าน Colab

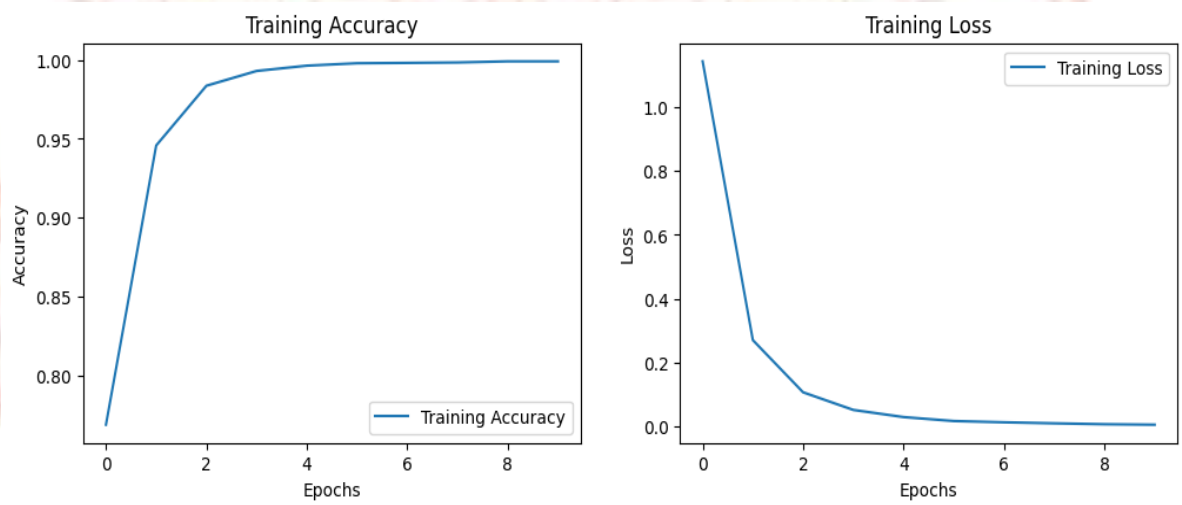
ตางรางที่ 3-2 การประเมินของผู้เชี่ยวชาญและผู้ปฏิบัติงานเพื่อหาค่าต่ำที่สุด

Roles	**Comparison sentence**		
	sentences	***estimation categories***	***Score (%)***
Expertise	Nuisance alarm on panel	What can go wrong?	67.644905
	Injured from tools		
	Process Upset		
	Pump cut-in function working.		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Not informing the panel before working		
	Not log off RCU before working		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		
	Informing the panel before working		
	Log off RCU before working		
Worker	Alarm nuisance panel man	What can go wrong?	
	Process maybe upset		
	Process leak to outside		
	Not inform panel man before work	What can go cause it go wrong?	
	Not inform panel man before work		
	Not isolate block valve before work		
	Inhibit alarm and Inform panel man before work	What can prevent it from going wrong?	
	Inform panel man before work		
	Isolate block valve and de-pressure before work		
	Wearing PPE and work carefully		

บทที่ 4
ผลการดำเนินงาน

4.1 Find-tuning BERT Classification model

ในการทดลองใช้โมเดล BERT สำหรับการจัดประเภทข้อมูล พบว่ามีข้อจำกัดด้านฮาร์ดแวร์ เนื่องจากโมเดลขนาดใหญ่ต้องใช้ RAM GPU สูง ซึ่งคอมพิวเตอร์ทั่วไปไม่สามารถรองรับได้ การใช้ Cloud Computer เป็นทางเลือกที่ดีในการเทรนโมเดลเพราะมีความยืดหยุ่นและประหยัดเวลา แต่จำเป็นต้องจัดการเวลาในการใช้งานอย่างเหมาะสม การแบ่งการเทรนออกเป็น 10 epochs บน Google Colab GPU ช่วยให้สามารถปรับพารามิเตอร์โมเดลได้อย่างละเอียดและมีประสิทธิภาพ ใช้เวลาประมาณ 9 นาทีต่อ epoch ซึ่งช่วยให้การเทรนโมเดลเป็นไปอย่างมีประสิทธิภาพและแม่นยำมากขึ้น



ภาพที่ 4-1 Training Accuracy vs Training Loss

หลังจากทำการเทรนโมเดลแล้ว จำเป็นต้องทดสอบโมเดลตามมาตรฐานเพื่อประเมินประสิทธิภาพ โดยการทดสอบประกอบด้วย 3 ตัวชี้วัดหลัก ได้แก่ 1. Precision ซึ่งเป็นการวัดความแม่นยำของโมเดลในการทำนายผลบวกที่ถูกต้อง, 2. Recall ซึ่งเป็นการวัดความสามารถของโมเดลในการทำนายผลบวกทั้งหมดที่ถูกต้อง, และ 3. F1 score ซึ่งเป็นค่าเฉลี่ยเชิงฮาร์โมนิกของ Precision และ Recall ช่วยให้เห็นภาพรวมของประสิทธิภาพโมเดลได้อย่างสมบูรณ์ การทดสอบทั้งสามอย่างนี้ช่วยให้เราเข้าใจถึงจุดแข็งและจุดอ่อนของโมเดลที่พัฒนา ทำให้สามารถ

ปรับปรุงและปรับแต่งโมเดลเพื่อให้มีความแม่นยำและประสิทธิภาพสูงสุด โดยเฉพาะในการประเมินความเสี่ยงที่มีความสำคัญต่อความปลอดภัยในอุตสาหกรรมน้ำมันและก๊าซ การใช้มาตรฐานการทดสอบนี้ช่วยให้มั่นใจได้ว่าโมเดลที่พัฒนาขึ้นมีความสามารถในการทำนายและประเมินความเสี่ยงได้อย่างมีประสิทธิภาพ

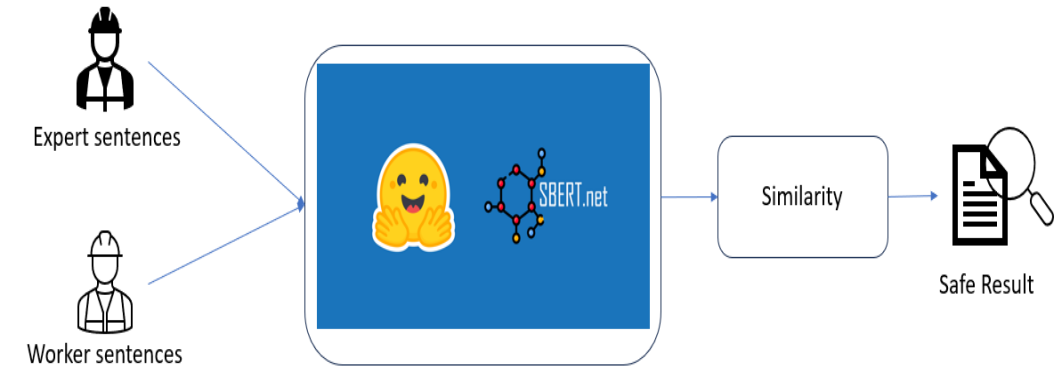
ตารางที่ 4-1 Classification Report

Iteration	Classification Report		
	Precision	***Recall***	***F1***
1	1	0.99	1
2	1	1	1
3	0.98	0.98	0.98
4	0.92	1	0.96
5	1	1	1

4.2 Similarity Scores

การใช้ Google Colab เพื่อทดสอบความสามารถของพนักงานในการทดสอบที่เข้มงวดเป็นกระบวนการที่มีความสำคัญในการพัฒนาทักษะและความเชี่ยวชาญของพนักงาน โดยที่เราใช้โมเดล BERT ที่ได้ทำการ fine-tuning มาเพื่อใช้ในการประเมินความสามารถในการตอบคำถามหรือแสดงความคิดเห็นตามกรอบงานที่กำหนดไว้ล่วงหน้า กระบวนการทดสอบที่เข้มงวดนี้ประกอบด้วยการมอบหมายให้พนักงานกรอกประโยคตามที่กำหนดในสถานการณ์ที่ต่างๆ ที่ถูกกำหนดไว้ล่วงหน้าในกรอบงานการประเมินนี้ โดยข้อมูลที่พนักงานกรอกจะถูกใช้เพื่อการประเมินคะแนน ซึ่งเป็นการทดสอบที่อ้างอิงถึงความสามารถในการเข้าใจและตอบสนองตามเนื้อหาที่กำหนด เพื่อให้การประเมินเป็นไปอย่างถูกต้องและเป็นประสบการณ์ที่มีความหมาย กรอบงานการประเมินนี้ได้รับการออกแบบโดยอ้างอิงจากแบบฟอร์ม SAP /NZEPM0012 ที่เป็นเครื่องมือที่ช่วยในการสร้าง Work Clearance จากใบสั่งงานในสถาน

ประกอบการอุตสาหกรรมต่างๆ ซึ่งได้รับการปรับแต่งเพื่อให้เหมาะสมและมีประสิทธิภาพสำหรับการใช้ในการประเมินทักษะของพนักงานในที่สุด



รูปที่ 4.2 Similarity Scheme

4.2.1 รูปแบบประโยคเหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

ตารางที่ 4-2 รูปแบบประโยคที่เหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

Roles	**Comparison sentence**		
	sentences	***estimation categories***	***Score (%)***
Expertise	The process malfunctions when PV is bad.	What can go wrong?	100
	Nuisance alarm on panel		
	Injured from tools		
	Process Upset		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Not informing the panel before working		
	Not log off RCU before working		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		
	Informing the panel before working		
	Log off RCU before working		
Worker	The process malfunctions when PV is bad.	What can go wrong?	
	Nuisance alarm on panel		
	Injured from tools		
	Process Upset		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Not informing the panel before working		
	Not log off RCU before working		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		
	Informing the panel before working		
	Log off RCU before working		

4.2.2 รูปแบบประโยคเหมือนกันครึ่งหนึ่งทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

ตารางที่ 4-3 รูปแบบประโยคที่เหมือนกันครึ่งหนึ่งทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

Roles	**Comparison sentence**		
	sentences	***estimation categories***	***Score (%)***
Expertise	The process malfunctions when PV is bad.	What can go wrong?	57.644905
	Nuisance alarm on panel		
	Injured from tools		
	Process Upset		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Not informing the panel before working		
	Not log off RCU before working		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		
	Informing the panel before working		
	Log off RCU before working		
Worker	The process malfunctions when PV is bad.	What can go wrong?	
	Nuisance alarm on panel		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		

4.2.3 รูปแบบประโยคไม่เหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

ตารางที่ 4-4 รูปแบบประโยคที่ไม่เหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงาน

Roles	**Comparison sentence**		
	sentences	***estimation categories***	***Score (%)***
Expertise	The process malfunctions when PV is bad.	What can go wrong?	59.42856
	Nuisance alarm on panel		
	Injured from tools		
	Process Upset		
	Not inhibit the alarm before starting work	What can go cause it go wrong?	
	Use tools that not suitable the type of work and not wearing PPE properly		
	Not informing the panel before working		
	Not log off RCU before working		
	Inhibit the alarm before starting work	What can prevent it from going wrong?	
	Use tools that suitable the type of work and wearing PPE properly		
	Informing the panel before working		
	Log off RCU before working		
Worker	She always uses the wrong tools for her tasks	What can go wrong?	
	He never wears the proper PPE during work		
	The workers frequently choose inappropriate equipment		
	The company mandates PPE, but some employees ignore the rules		
	They rely on incorrect equipment for their jobs	What can go cause it go wrong?	
	The safety officer inspects PPE compliance regularly		
	Some employees consistently overlook safety protocols		
	He handles hazardous materials without proper protection		
	She rarely follows guidelines for using the correct tools	What can prevent it from going wrong?	
	Employees often neglect to wear their safety helmets		
	He seldom uses the appropriate gloves for handling chemicals		
	The workers frequently fail to wear safety vests on site		

4.2.4 รูปแบบประโยคเหมือนทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงานภาษาไทย

ตารางที่ 4-5 รูปแบบประโยคเหมือนทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงานภาษาไทย

Roles	Comparison sentence		
	sentences	***estimation categories***	***Score (%)***
Expertise	กระบวนการมีปัญหาเมื่อ PV ไม่ดี	What can go wrong?	99.21787
	ระบบแจ้งเตือนรบกวนบนพาเนล		
	บาดเจ็บจากเครื่องมือ		
	กระบวนการผิดปกติ		
	ไม่ยับยั้งเสียงเตือนก่อนเริ่มงาน	What can go cause it go wrong?	
	ไม่ใช้เครื่องมือที่เหมาะสมกับงานและไม่สวมใส่ PPE อย่างเหมาะสม		
	ไม่แจ้งให้พาเนลทราบก่อนเริ่มงาน		
	ไม่ล็อกออกจากระบบ RCU ก่อนเริ่มงาน		
	ยับยั้งเสียงเตือนก่อนเริ่มงาน	What can prevent it from going wrong?	
	ใช้เครื่องมือที่เหมาะสมกับงานและสวมใส่ PPE อย่างเหมาะสม		
	แจ้งให้พาเนลทราบก่อนเริ่มงาน		
	ล็อกออกจากระบบ RCU ก่อนเริ่มงาน		
Worker	กระบวนการมีปัญหาเมื่อ PV ไม่ดี	What can go wrong?	
	ระบบแจ้งเตือนรบกวนบนพาเนล		
	บาดเจ็บจากเครื่องมือ		
	กระบวนการผิดปกติ		
	ไม่ยับยั้งเสียงเตือนก่อนเริ่มงาน	What can go cause it go wrong?	
	ไม่ใช้เครื่องมือที่เหมาะสมกับงานและไม่สวมใส่ PPE อย่างเหมาะสม		
	ไม่แจ้งให้พาเนลทราบก่อนเริ่มงาน		
	ไม่ล็อกออกจากระบบ RCU ก่อนเริ่มงาน		
	ยับยั้งเสียงเตือนก่อนเริ่มงาน	What can prevent it from going wrong?	
	ใช้เครื่องมือที่เหมาะสมกับงานและสวมใส่ PPE อย่างเหมาะสม		
	แจ้งให้พาเนลทราบก่อนเริ่มงาน		
	ล็อกออกจากระบบ RCU ก่อนเริ่มงาน		

4.2.5 รูปแบบประโยคเหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงานภาษาไทยและอังกฤษ

ตารางที่ 4-6 รูปแบบประโยคเหมือนกันทั้งผู้เชี่ยวชาญและผู้ปฏิบัติงานภาษาไทยและอังกฤษ

Roles	**Comparison sentence**		
	sentences	***estimation categories***	***Score (%)***
Expertise	กระบวนการมีปัญหาเมื่อ PV ไม่ดี (The process malfunctions when PV is bad.)	What can go wrong?	98.87649
	ระบบแจ้งเตือนรบกวนบนพาเนล (Nuisance alarm on panel)		
	บาดเจ็บจากเครื่องมือ (Injured from tools.)		
	กระบวนการผิดปกติ (Process Upset)		
	ไม่ยับยั้งเสียงเตือนก่อนเริ่มงาน (Not inhibit the alarm before starting work)	What can go cause it go wrong?	
	ไม่ใช้เครื่องมือที่เหมาะสมกับงานและไม่สวมใส่ PPE อย่างเหมาะสม (Not Use tools that not suitable the type of work and not wearing PPE properly)		
	ไม่แจ้งให้พาเนลทราบก่อนเริ่มงาน (Not informing the panel before working)		
	ไม่ล็อกออกจากระบบ RCU ก่อนเริ่มงาน (Not log off RCU before working)		
	ยับยั้งเสียงเตือนก่อนเริ่มงาน (Inhibit the alarm before starting work)	What can prevent it from going wrong?	
	ใช้เครื่องมือที่เหมาะสมกับงานและสวมใส่ PPE อย่างเหมาะสม (Use tools that suitable the type of work and wearing PPE properly)		
	แจ้งให้พาเนลทราบก่อนเริ่มงาน (Informing the panel before working)		
	ล็อกออกจากระบบ RCU ก่อนเริ่มงาน (Log off RCU before working)		
Worker	กระบวนการมีปัญหาเมื่อ PV ไม่ดี (The process malfunctions when PV is bad.)	What can go wrong?	
	ระบบแจ้งเตือนรบกวนบนพาเนล (Nuisance alarm on panel)		
	บาดเจ็บจากเครื่องมือ (Injured from tools.)		
	กระบวนการผิดปกติ (Process Upset)		
	ไม่ยับยั้งเสียงเตือนก่อนเริ่มงาน (Not inhibit the alarm before starting work)	What can go cause it go wrong?	
	ไม่ใช้เครื่องมือที่เหมาะสมกับงานและไม่สวมใส่ PPE อย่างเหมาะสม (Not Use tools that not suitable the type of work and not wearing PPE properly)		
	ไม่แจ้งให้พาเนลทราบก่อนเริ่มงาน (Not informing the panel before working)		
	ไม่ล็อกออกจากระบบ RCU ก่อนเริ่มงาน (Not log off RCU before working)		
	ยับยั้งเสียงเตือนก่อนเริ่มงาน (Inhibit the alarm before starting work)	What can prevent it from going wrong?	
	ใช้เครื่องมือที่เหมาะสมกับงานและสวมใส่ PPE อย่างเหมาะสม (Use tools that suitable the type of work and wearing PPE properly)		
	แจ้งให้พาเนลทราบก่อนเริ่มงาน (Informing the panel before working)		
	ล็อกออกจากระบบ RCU ก่อนเริ่มงาน (Log off RCU before working)		

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

การประเมินความเสี่ยงในแต่ละกิจกรรมภายในโรงงานมีความสำคัญอย่างมาก โดยเฉพาะอย่างยิ่งในภาคอุตสาหกรรมน้ำมันและก๊าซ เนื่องจากมีศักยภาพในการเกิดผลกระทบที่รุนแรง การพึ่งพาการประเมินความเสี่ยงที่ทำโดยมนุษย์เพียงอย่างเดียวในอดีตทำให้เราเห็นถึงความไม่เพียงพอและข้อจำกัด เนื่องจากยังมีการเกิดอุบัติเหตุในอุตสาหกรรมนี้อย่างต่อเนื่องและมีแนวโน้มที่จะไม่สามารถกำจัดปัญหาเหล่านี้ได้อย่างสมบูรณ์ การนำผู้เชี่ยวชาญเข้ามาประเมินความเสี่ยงในอุตสาหกรรมนี้จึงเป็นสิ่งสำคัญ เพราะการใช้ความรู้และประสบการณ์จากผู้เชี่ยวชาญสามารถช่วยให้การประเมินความเสี่ยงมีความแม่นยำมากขึ้น นอกจากนี้ การประยุกต์ใช้เทคโนโลยีการประมวลผลภาษาและการประมวลผลข้อมูลขั้นสูง เช่น โมเดลภาษาแบบ LLM (Large Language Model) และการประมวลผลแบบคลาวด์ (Cloud Computing) ยังสามารถเพิ่มประสิทธิภาพในการประเมินความเสี่ยงได้อย่างมีนัยสำคัญ การประเมินความเสี่ยงที่แม่นยำและมีประสิทธิภาพจะช่วยลดความเสี่ยงของการเกิดอุบัติเหตุและผลกระทบที่ตามมา นอกจากนี้ ยังช่วยให้อุตสาหกรรมน้ำมันและก๊าซสามารถดำเนินการได้อย่างปลอดภัยและมีประสิทธิภาพมากยิ่งขึ้น ซึ่งเป็นเป้าหมายสำคัญของการวิจัยและพัฒนาในด้านนี้

5.2 ปัญหาและอุปสรรคในการทดลอง

5.2.1 ปัญหาจาก Datasets

หนึ่งในปัญหาหลักที่พบในการทดลองครั้งนี้คือการจัดการกับข้อมูล (Datasets) ที่ได้มา เนื่องจากข้อมูลทั้งหมดที่ได้รับมาเป็นภาษาเขียน ซึ่งมีความยุ่งยากและซับซ้อนในการตีความจากข้อมูลเหล่านี้ การวิเคราะห์ข้อมูลภาษาที่ไม่เป็นทางการอาจทำให้เกิดความเข้าใจผิดพลาดหรือไม่ครบถ้วน ความจำเป็นในการตีความข้อมูลที่ซับซ้อนเหล่านี้ต้องใช้ผู้เชี่ยวชาญในด้านที่เกี่ยวข้องเท่านั้น ซึ่งเป็นข้อจำกัดที่สำคัญ เนื่องจากไม่สามารถใช้บุคคลที่ไม่มีความรู้เฉพาะทางมาทำการตีความข้อมูลได้ ทำให้กระบวนการนี้ใช้เวลานานและต้องการความละเอียดในการจัดทำ Datasets เพื่อให้ได้ข้อมูลที่มีคุณภาพสูง นอกจากนี้ การทำ Datasets ด้วยตัวผู้เชี่ยวชาญเองทั้งหมดนั้นยังเพิ่มภาระในการทำงานและมีค่าใช้จ่ายสูง การที่ต้องใช้เวลานานในการเตรียมข้อมูลนี้อาจทำให้กระบวนการทั้งหมดล่าช้าและส่งผลต่อการวิเคราะห์และการประเมินความเสี่ยงที่มีประสิทธิภาพในอุตสาหกรรม เพื่อแก้ไขปัญหาเหล่านี้ อาจจำเป็นต้องพิจารณาใช้

เทคโนโลยีและเครื่องมือที่ทันสมัยในการช่วยในการประมวลผลและตีความข้อมูล เช่น การใช้ปัญญาประดิษฐ์ (AI) และการประมวลผลภาษาธรรมชาติ (NLP) ซึ่งสามารถช่วยลดภาระและความซับซ้อนในการทำ Datasets และเพิ่มความแม่นยำในการวิเคราะห์ข้อมูลได้

5.2.2 ปัญหาจาก Hardware

การใช้โมเดลที่ผ่านการฝึกมาแล้ว (pre-trained models) ซึ่งมีขนาดใหญ่ มักจะต้องใช้ทรัพยากรฮาร์ดแวร์ที่มีประสิทธิภาพสูง โดยเฉพาะการใช้หน่วยความจำ (RAM) และการประมวลผลผ่านหน่วยประมวลผลกราฟิก (GPU) ที่มีประสิทธิภาพสูง ซึ่งคอมพิวเตอร์ทั่วไปไม่สามารถรองรับการใช้งานนี้ได้อย่างเต็มที่ การใช้งานโมเดลขนาดใหญ่นั้นจำเป็นต้องใช้ฮาร์ดแวร์ที่มีแรมสูง และอุปกรณ์คอมพิวเตอร์เฉพาะทางเช่น Super Computer ซึ่งมีค่าใช้จ่ายสูงมาก อีกทั้งยังต้องมีการดูแลและบำรุงรักษาที่ซับซ้อน เพื่อแก้ไขปัญหานี้ การใช้ Cloud Computing เป็นตัวเลือกที่ดีกว่าการลงทุนกับ Super Computer สำหรับการประมวลผลข้อมูลและการรันโมเดลขนาดใหญ่ Cloud Computing มีความยืดหยุ่นสูงและสามารถปรับขนาดทรัพยากรตามความต้องการได้ ทำให้ลดค่าใช้จ่ายและการบำรุงรักษาฮาร์ดแวร์ได้อย่างมีประสิทธิภาพ นอกจากนี้ การใช้ Cloud Computing ยังสามารถเข้าถึงเทคโนโลยีและบริการที่ทันสมัยได้อย่างรวดเร็ว โดยไม่ต้องลงทุนในโครงสร้างพื้นฐานเอง การใช้บริการจากผู้ให้บริการ Cloud Computing เช่น Amazon Web Services (AWS), Google Cloud Platform (GCP), หรือ Microsoft Azure ยังช่วยในการจัดการและประมวลผลข้อมูลขนาดใหญ่ได้อย่างมีประสิทธิภาพ การเลือกใช้ Cloud Computing ยังช่วยให้สามารถทำงานร่วมกันได้ง่ายขึ้น โดยสามารถแชร์ทรัพยากรและทำงานร่วมกันระหว่างทีมที่อยู่ในสถานที่ต่าง ๆ ได้ ซึ่งช่วยเพิ่มประสิทธิภาพในการทำงานและการประมวลผลข้อมูลได้อย่างมาก

5.3 ข้อเสนอแนะ

5.3.2 การใช้ Cloud Computing สำหรับการประมวลผล

เนื่องจากการใช้โมเดลขนาดใหญ่ (Large Language Models - LLM) ไม่จำเป็นต้องใช้ Super Computer เราสามารถใช้ Cloud Computing ในการฝึกและรันโมเดลได้ ซึ่งจะช่วยประหยัดค่าใช้จ่ายและทรัพยากรฮาร์ดแวร์ และ ควรวางแผนการใช้งานเวลาในการประมวลผลอย่างรอบคอบ เนื่องจากเวลาที่สามารถใช้ GPU บน Cloud Computing นั้นมีข้อจำกัด การวางแผนการใช้งานอย่างมีประสิทธิภาพจะช่วยลดค่าใช้จ่ายและเพิ่มประสิทธิภาพในการประมวลผล

5.3.2 การจัดเตรียม Data Sets

การที่ข้อมูลทั้งหมดเป็นภาษาเขียนทำให้การตีความซับซ้อน ควรจัดเตรียม Data Sets โดยใช้ผู้เชี่ยวชาญด้านภาษาศาสตร์และผู้เชี่ยวชาญในสาขาที่เกี่ยวข้องในการจัดการข้อมูล และควรใช้เทคนิคการทำความสะอาดข้อมูล (data cleaning) และการทำข้อมูลให้เป็นมาตรฐาน (data standardization) เพื่อให้การประมวลผลข้อมูลมีประสิทธิภาพและแม่นยำมากขึ้น

5.3.3 การจัดเตรียม Data Sets

ควรมีการฝึกอบรมทีมงานให้มีความรู้และทักษะในการใช้ Cloud Computing และการจัดการโมเดลขนาดใหญ่ และ การพัฒนาทักษะการใช้เครื่องมือและแพลตฟอร์มต่างๆ จะช่วยเพิ่มความสามารถในการประมวลผลและการจัดการข้อมูล

5.3.4 การเลือกแพลตฟอร์มและเครื่องมือที่เหมาะสม

ควรเลือกใช้แพลตฟอร์มและเครื่องมือที่เหมาะสมกับการประมวลผลและการจัดการข้อมูล เช่น Amazon Web Services (AWS), Google Cloud Platform (GCP), หรือ Microsoft Azure การเลือกเครื่องมือที่เหมาะสมจะช่วยเพิ่มประสิทธิภาพในการทำงานและลดค่าใช้จ่ายในการประมวลผลข้อมูล

5.3.5 การวางแผนและการจัดการทรัพยากร

ควรมีการวางแผนและการจัดการทรัพยากรอย่างมีประสิทธิภาพ เพื่อให้การใช้ทรัพยากรฮาร์ดแวร์และการประมวลผลข้อมูลมีประสิทธิภาพสูงสุด การจัดการทรัพยากรอย่างมีระบบจะช่วยลดข้อผิดพลาดและเพิ่มประสิทธิภาพในการทำงาน

บรรณานุกรม

[1] Alberti, C., Andor, D., Pitler, E., Devlin, J., & Collins, M. (2019). Synthetic QA Corpora Generation with Roundtrip Consistency. arXiv: Computation and Language. Retrieved 11 2, 2022, from https://arxiv.org/abs/1906.05416

[2] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv: Computation and Language. Retrieved 11 2, 2022, from https://arxiv.org/abs/1810.04805

[3] Self-Instruct: Aligning Language Models with Self-Generated Instructions

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi

[4] Mikolov, T., Chen, K., Corrado, G. S., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv: Computation and Language. Retrieved 11 2, 2022, from https://arxiv.org/abs/1301.3781

[5] Rengasamy, V., Fu, T.-Y., Lee, W.-C., & Madduri, K. (2017). Optimizing Word2Vec Performance on Multicore Systems. Intelligenza Artificiale, 3. Retrieved 11 2, 2022, from https://dl.acm.org/citation.cfm?id=3149768

[6] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. arXiv: Computation and Language. Retrieved 11 2, 2022, from https://arxiv.org/abs/1409.3215

[7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Polosukhin, I. (2017). Attention Is All You Need. arXiv: Computation and Language. Retrieved 11 2, 2022, from https://arxiv.org/abs/1706.03762

[8] Messaoudi, F., & Belmokhtar, O. (2022). Identification of Risk Features Using Text Mining and BERT-Based Models: Application to an Oil Refinery. Process

Safety and Environmental Protection. Retrieved from https://ui.adsabs.harvard.edu/abs/2022PSEP..158..382M/abstract

[9] Kim, S., & Lee, J. (2023). Development of a Knowledge Base for Construction Risk Assessments Using BERT and Graph Models. Buildings. Retrieved from https://www.mdpi.com/2075-5309/14/11/3359

[10] Ivanov, A., & Petrov, D. (2024). Assessing and Managing Risks in the Oil and Gas Industry with Machine Learning and Digital Technologies. E3S Web of Conferences. Retrieved from https://www.e3s-conferences.org/articles/e3sconf/abs/2024/04/e3sconf_icite2023_02011/e3sconf_icite2023_02011.html

[11] Macêdo, J. B., das Chagas Moura, M., Aichele, D., & Lins, I. D. (2022). Identification of Risk Features Using Text Mining and BERT-Based Models: Application to an Oil Refinery. *Process Safety and Environmental Protection, 158*, 382–392. Retrieved from https://ui.adsabs.harvard.edu/abs/2022PSEP..158..382M/abstract

[12] Canon, J., Broussard, T., Johnson, A., Singletary, W., & Colmenares-Diaz, L. (2022). A Knowledge-Based Artificial Intelligence Approach to Risk Management. *SPE Annual Technical Conference and Exhibition*. Retrieved from https://onepetro.org/SPEATCE/proceedings-abstract/22ATCE/22ATCE/509271

[13] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv preprint arXiv:1908.10084*. Retrieved from https://arxiv.org/abs/1908.10084

[14] Kulkarni, A. (2024). BERT Applications & Use Cases in Natural Language Processing. *SevenMentor*. Retrieved from https://www.sevenmentor.com/bert-applications-use-cases-in-natural-language-processing

[15] GeeksforGeeks. (2024). Transforming Language Understanding: An In-Depth Look at BERT and Its Applications. *GeeksforGeeks*. Retrieved from

https://www.geeksforgeeks.org/transforming-language-understanding-an-in-depth-look-at-bert-and-its-applications/

[16] Zhang, Y., & Chen, Q. (2023). Biomedical Abstract Sentence Classification by BERT-Based Reading Comprehension. *SN Computer Science*. Retrieved from https://dl.acm.org/doi/abs/10.1007/s42979-023-01830-0

[17] Wang, L., & Li, H. (2022). BERT-Based Scientific Paper Quality Prediction. *Artificial Neural Networks and Machine Learning – ICANN 2022*. Retrieved from https://dl.acm.org/doi/10.1007/978-3-031-15937-4_18

[18] Batra, H., Punn, N. S., Sonbhadra, S. K., & Agarwal, S. (2021). BERT-Based Sentiment Analysis: A Software Engineering Perspective. *arXiv preprint arXiv:2106.02581*. Retrieved from https://arxiv.org/abs/2106.02581

ภาคผนวก

```
!pip install -qq transformers

import pandas as pd

import numpy as np

from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score

from sklearn.preprocessing import LabelEncoder

from transformers import BertTokenizer, TFBertForSequenceClassification

import tensorflow as tf

import matplotlib.pyplot as plt



# Load your data from the Excel file

file_name = "/content/output.xlsx"

df = pd.read_excel(file_name)



# Load the pre-trained BERT tokenizer

tokenizer = BertTokenizer.from_pretrained("bert-base-uncased")

# Split your data into training and testing sets

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

X_train_encodings = tokenizer(list(X_train), padding=True, truncation=True,
            return_tensors='tf', max_length=100)

X_test_encodings = tokenizer(list(X_test), padding=True, truncation=True,
            return_tensors='tf', max_length=100)
```

```
# Encode your labels using one-hot encoding
label_encoder = LabelEncoder()
y_train_encoded = label_encoder.fit_transform(y_train)
y_test_encoded = label_encoder.transform(y_test)

num_classes = len(label_encoder.classes_)
y_train_one_hot = tf.one_hot(y_train_encoded, depth=num_classes)
y_test_one_hot = tf.one_hot(y_test_encoded, depth=num_classes)

# Load the pre-trained BERT model for sequence classification
model = TFBertForSequenceClassification.from_pretrained("bert-base-uncased",
        num_labels=num_classes)

# Define an optimizer
optimizer = tf.keras.optimizers.Adam(learning_rate=1e-5)

# Compile the model with a custom loss function
def custom_loss(y_true, y_pred):
    return tf.reduce_mean(tf.nn.softmax_cross_entropy_with_logits(labels=y_true,
        logits=y_pred))

model.compile(optimizer=optimizer, loss=custom_loss, metrics=['accuracy'])
```

```
# Define batch size and epochs

batch_size = 32

epochs = 3

# Create TensorFlow datasets

train_dataset = tf.data.Dataset.from_tensor_slices((dict(X_train_encodings),
            y_train_one_hot)).batch(batch_size)

test_dataset = tf.data.Dataset.from_tensor_slices((dict(X_test_encodings),
            y_test_one_hot)).batch(batch_size)

# Train the model

history = model.fit(train_dataset, epochs=epochs)

# Make predictions on the test set

y_pred = model.predict(test_dataset)

y_pred = label_encoder.inverse_transform(np.argmax(y_pred.logits, axis=1))


# Evaluate the model

accuracy = accuracy_score(y_test, y_pred)

print(f"Accuracy: {accuracy}")


# Plot accuracy and loss during training

plt.figure(figsize=(12, 4))

plt.subplot(1, 2, 1)

plt.plot(history.history['accuracy'], label='Training Accuracy')

plt.title('Training Accuracy')
```

```
plt.xlabel('Epochs')

plt.ylabel('Accuracy')

plt.legend()


plt.subplot(1, 2, 2)

plt.plot(history.history['loss'], label='Training Loss')

plt.title('Training Loss')

plt.xlabel('Epochs')

plt.ylabel('Loss')

plt.legend()


plt.show()

from sklearn.metrics import f1_score, recall_score


# Convert one-hot encoded y_test to categorical labels

y_test_categorical = np.argmax(y_test_one_hot, axis=1)


# Calculate F1 score

y_pred_categorical = np.argmax(model.predict(test_dataset).logits, axis=1)

f1 = f1_score(y_test_categorical, y_pred_categorical, average='weighted')


# Calculate Recall
```

```
recall = recall_score(y_test_categorical, y_pred_categorical, average='weighted')


print(f"F1 Score: {f1}")

print(f"Recall: {recall}")

from sklearn.metrics import classification_report


# Generate a classification report

report = classification_report(y_test_categorical, y_pred_categorical)


# Split the report into lines

report_lines = report.split('\n')


# Print the first 5 lines

for line in report_lines[:10]:

    print(line)


from transformers import BertTokenizer, BertModel

import torch

import torch.nn.functional as F

import numpy as np

from sklearn.metrics.pairwise import cosine_similarity
```

```
# Load pre-trained model tokenizer

tokenizer = BertTokenizer.from_pretrained('bert-base-uncased')


# Load pre-trained model

model = BertModel.from_pretrained('bert-base-uncased')


# Set evaluation mode

model.eval()

summary1 = [What1]

numemrator1 = np.exp(summary1)

denominator1 = np.sum(numemrator1)

sigma1 = numemrator1/denominator1

print(sigma1.mean())

print(sigma1)

print(np.sum(sigma1))

a = sigma1.mean()

print(a)

summary2 = [What2]

numemrator2 = np.exp(summary2)

denominator2 = np.sum(numemrator2)

sigma2 = numemrator2/denominator2

print(sigma2.mean())
```

```
print(sigma2)

print(np.sum(sigma2))

b = sigma2.mean()

print(b)

summary3 = [What3]

numemrator3 = np.exp(summary3)

denominator3 = np.sum(numemrator3)

sigma3 = numemrator3/denominator3

print(sigma3.mean())

print(sigma3)

print(np.sum(sigma3))

c = sigma3.mean()

print(c)


summary = [a,b,c]

numemrator = np.exp(summary)

denominator = np.sum(numemrator)

sigma = numemrator/denominator

np.sum(sigma)

print("Percent of a safe is:",(a+b+c)*100,"%")
```

ประวัติผู้เขียน

ชื่อ	ภูวดล นุ่นภักดี
ชื่อวิทยานิพนธ์	การประเมินความปลอดภัยแบบอัตโนมัติด้วย BERT สำหรับการปฏิบัติตามข้อกำหนดในสถานที่ทำงานทางอุตสาหกรรม
สาขาวิชา	วิศวกรรมอัตโนมัติ
ประวัติ	ปริญญาตรี วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมระบบวัดคุม