HPC LocalVoice ระบบสื่อสารอัจฉริยะภาษาไทยกับภาษาม้ง ครั้งแรกของโลก

บทคัดย่อ

โครงการ HPC LocalVoice มุ่งพัฒนาระบบแปลภาษาเสียงพูดแบบเรียลไทม์ ระหว่าง ภาษาไทยและภาษาม้ง เพื่อแก้ปัญหาการสื่อสารในด้านการศึกษา การแพทย์ และการติดต่อราชการ โดยใช้เทคโนโลยี HPC (High Performance Computing) สร้างโมเดล AI Speech Recognition และ Text-to-Speech ช่วยให้การสื่อสารไร้กำแพงภาษา เกิดประโยชน์ต่อสังคมและเป็นต้นแบบต่อยอดสู่ภาษาอื่น ๆ

บทนำ

ชาวม้งจำนวนมากในประเทศไทยยังคงประสบปัญหาในการสื่อสารภาษาไทย ซึ่งส่งผลโดยตรงต่อการศึกษา การรักษาพยาบาล และการเข้าถึงบริการของรัฐ ปัญหาดังกล่าวก่อให้เกิด “กำแพงภาษา” ที่จำกัดโอกาสทางสังคมและซ้ำเติมความเหลื่อมล้ำ การพัฒนาเทคโนโลยีฟัง–แปล–พูด (Listen–Translate–Speak) จึงมีบทบาทสำคัญในการเชื่อมช่องว่างทางภาษา เสริมสร้างความเท่าเทียม และสามารถต่อยอดสู่การแก้ปัญหาการสื่อสารของกลุ่มชาติพันธุ์อื่น ๆ ได้ในอนาคต

วัตถุประสงค์

1. พัฒนาระบบสื่อสารเสียงพูด ไทย - ม้ง แบบเรียลไทม์
2. ลดอุปสรรคด้านภาษาในบริบทการศึกษา การแพทย์ และงานราชการ
3. สร้างต้นแบบนวัตกรรมที่สามารถขยายผลสู่ภาษาอื่น

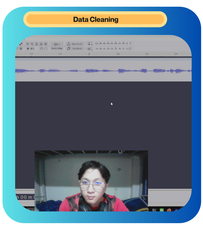
วิธีดำเนินงาน

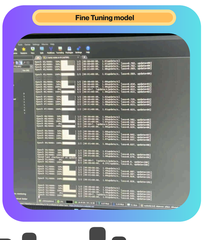
1. การเตรียมข้อมูล (Data Preparation)
   - ทำการรับบริจาคข้อมูลที่เป็นคำศัพท์และประโยค
   - ทำการตรวจเช็คประโยคให้ถูกต้อง โดย ผู้ที่เคยศึกษาภาษาม้ง
   - ทำการบันทึก เสียงพูดภาษาม้ง
   - ทำ Data Cleaning: ลบเสียงรบกวน ปรับคุณภาพไฟล์เสียง
   - สร้าง คู่ข้อมูล (Parallel Corpus) สำหรับฝึกโมเดลแปลภาษา

2. การทำโมเดล Speech to Text (STT)
   - โมเดล AI ฟังเสียงพูด → แปลงเป็นข้อความ
   - รองรับ ภาษาม้ง
   - ใช้เทคนิค Automatic Speech Recognition (ASR)
   - ใช้เทคนิคการ Fine Tune Model w2v-bert-2.0/whisper small
   - รันบน LANTA Supercomputer

3. การทำโมเดล Text to Speech (TTS)
   - แปลงข้อความเป็นเสียงพูดคุณภาพสูง
   - ใช้เทคโนโลยี F5TTS ในการ Fine tune Model TTS
   - รองรับ เสียงที่เป็นธรรมชาติและชัดเจน
   - Output: เสียงพูดที่ผู้ฟังเข้าใจง่าย
   - รันบน LANTA Supercomputer

4. การทำโมเดล Machine Translation (MT)
   - แปลงข้อความภาษาไทย → ภาษาม้ง หรือ ภาษาม้ง → ภาษาไทย
   - ใช้โมเดลฝึกจาก Parallel Corpus
   - โมเดลสามารถจับ บริบทและความหมาย ของประโยค
   - ใช้เทคนิคการ Fine tune Mode Helsinki-NLP/opus-mt-th-en
   - รันบน LANTA Supercomputer



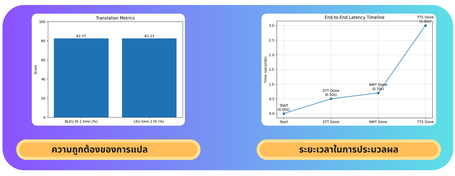


ผลการวิจัย

1. ระบบสามารถ ฟัง – แปล – พูดตอบกลับทันที
   - ระบบต้นแบบทำงานครบวงจร ตั้งแต่ Speech to Text → Machine Translation → Text to Speech
   - ผู้ใช้สามารถพูดภาษาไทยหรือม้ง แล้วระบบแปลและตอบกลับแบบ Real-time
   - ประสิทธิภาพของระบบขึ้นอยู่กับ ความถูกต้องของการแปล กับความเร็วในการประมวลผล

2. ความถูกต้องของการแปล (Translation Accuracy)
   - BLEU Score hmn 2 th : 82.33 %
   - BLEU Score th 2 hmn : 82.37 %
   - WER (Word Error Rate ใน STT) : 24.77 %
     ความพึงพอใจผู้ฟัง (MOS Test) : -

3. ระยะเวลาในการประมวลผล (Processing Time)
   - ทดสอบบนระบบ บน I 5 12400F GPU RTX 3080 10G Ram 32
   - เวลาเฉลี่ยตั้งแต่ผู้ใช้พูด ได้เสียงตอบกลับ:
     - STT (Speech to Text) : ~ 0.5 วินาที
     - MT (Translation) : 0.1 ~ 0.2 วินาที
     - TTS (F5TTS) : ~ 3 วินาที
     - รวมทั้งหมด (End-to-End Latency) : ~ 3.7 วินาที
       แสดงว่าผู้ใช้สามารถสนทนาได้ เกือบ Real-time



Web Application สื่อสารอัจฉริยะภาษาไทย-ม้ง

แหล่ง รวมโมเดล ภาษาม้ง

Huggin face : https://huggingface.co/Pakorn2112

GitHub: https://github.com/YangNobody12

WebSite : https://nlpforhmong.yangnobody.com

บทสรุป

HPC LocalVoice เป็นการบรรจบกันของนวัตกรรมและความเท่าเทียม ที่นำเทคโนโลยีการประมวลผลสมรรถนะสูงมาช่วยลดกำแพงภาษาไทยม้ง สร้างสะพานแห่งความเข้าใจและการเข้าถึงโอกาสทางสังคมอย่างเท่าเทียม พร้อมทั้งมีศักยภาพในการขยายผลสู่ภาษาอื่น ๆ ในท้องถิ่น เพื่อผลักดันการเปลี่ยนแปลงสังคมที่ยั่งยืนผ่านพลังของเทคโนโลยีขั้นสูง

เอกสารอ้างอิง

OpenAI.(2566).Whisper,สืบค้นจากhttps://github.com/openai/whisper
Hugging Face .(2567).Transformers documentation,สืบค้นจากhttps://github.com/openai/whisper
OpenAI. (2566). Whisper. GitHub. สืบค้นจาก https://github.com/openai/whisper
Hugging Face. (2567). เอกสารคำกับ Transformers . สืบค้นจาก https://huggingface.co/docs/transformers
Helsinki-NLP. (2568). Helsinki-NLP องค์กร . สืบค้นจาก https://huggingface.co/Helsinki-NLP
Helsinki-NLP. (2568). Opus-MT . GitHub. สืบค้นจาก https://github.com/Helsinki-NLP/Opus-MT
Chen, Y. และคณะ. (2567). F5-TTS . GitHub. สืบค้นจาก https://github.com/SWivid/F5-TTS
VYNCX. (2567). F5-TTS-THAI. GitHub. สืบค้นจาก https://github.com/VYNCX/F5-TTS-THAI
SWivid. (2568). F5-TTS-Plus. GitHub. สืบค้นจาก https://github.com/gjnave/F5-TTS-Plus

กิตติกรรมประกาศ

โครงการนี้ได้รับการสนับสนุนจาก โครงการระบบนิเวศสำหรับการพัฒนากำลังคนทักษะเชี่ยวชาญด้านการคำนวณสมรรถนะสูงเพื่อเสริมความพร้อมรับความท้าทายท้องถิ่นด้วยนวัตกรรมเทคโนโลยีชั้นนำ (Ecosystem for Developing High Performance Computing Talents to Boost Preparedness for Local Challenges Through Cutting-Edge Technological Innovation) ศูนย์ High Performance Computing (HPC) ทีมภาคสนาม และชุมชนชาวม้งออนไลน์ที่ให้ข้อมูลและข้อเสนอแนะอันมีค่า

ขอขอบคุณทุกฝ่ายที่ร่วมผลักดันให้เกิดนวัตกรรม HPC LocalVoice ซึ่งเป็นก้าวสำคัญในการสร้างสะพานเชื่อมความเข้าใจระหว่างภาษาและวัฒนธรรมท้องถิ่น

อาจารย์ที่ปรึกษา
อาจารย์ศิริกรณ์ กันขัติ์
มหาวิทยาลัยราชภัฏเชียงใหม่

นาย ปกรณ์ อาชาคีรี
นักศึกษา ชั้นปีที่ 2
สาขา คอมพิวเตอร์ศึกษา
คณะ ครุศาสตร์
มหาวิทยาลัยราชภัฏเชียงใหม่