# รายงานความก้าวหน้าวิชา Computer Engineering Project Preparation

**ครั้งที่ 4**

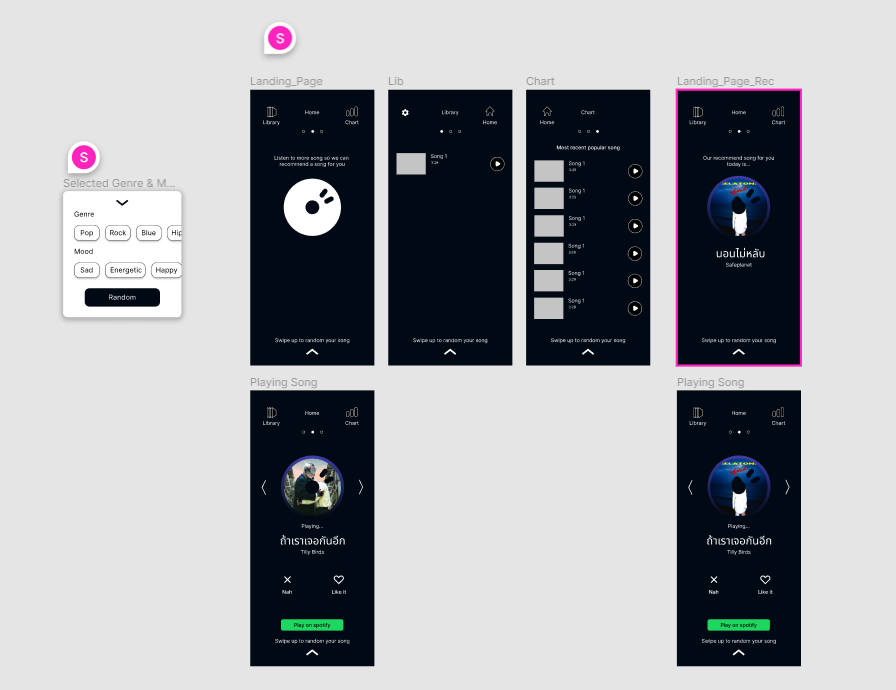
1. ชื่อโครงงาน (อังกฤษ) Music genre and mood classification with Machine learning

## การดำเนินงานมีความก้าวหน้า 80 %

## ความก้าวหน้าระหว่างวันที่ 01 เม.ย. 65 ถึงวันที่ 22 เม.ย. 65

## รายละเอียดความก้าวหน้า

* ทำการออกแบบ ส่วนของหน้าตา Application ในส่วนที่เป็นหน้าสำคัญหลักๆ ได้แก่ Landing page, Library, Chart, Random song interface ซึ่งโดยขั้นต้นได้มีการพยายามออกแบบให้ใช้งานได้ง่ายที่สุดและไม่ทำให้ผู้ใช้งานรู้สึกสับสนกับการเข้ามาใช้งานในครั้งแรก แต่อย่างไรก็ตาม อาจจะยังจำเป็นที่จะต้องทำ Instruction เพื่อแนะนำการใช้งานให้แก่ผู้ใช้ในการใช้งานครั้งแรก แต่ถึงกระนั้น หากสามารถทำให้ผู้ใช้สามารถเข้าใจการใช้งานได้ตั้งแต่แรกโดยที่ไม่จำเป็นต้องมีการแนะนำการใช้งานได้ก็จะเป็นสิ่งที่ดีที่สุดในการออกแบบ User interface ของ Application



การออกแบบหน้าตาของแอปพลิเคชันในส่วนสำคัญ

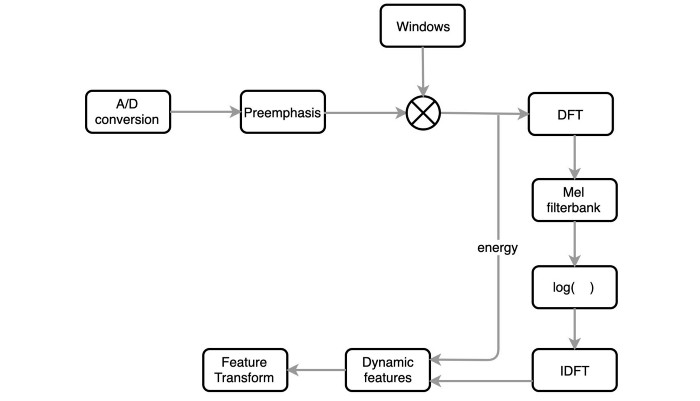
* ในการแนะนำเพลงและส่วนของการเล่นเพลงนั้น ได้คาดเอาไว้ว่า จะให้ใช้การใช้งานของสี เพื่อบ่งบอกถึงอารมณ์ในเพลงนั้นๆ ยกตัวอย่างเช่น เพลงเศร้า-โทนสีน้ำเงิน, เพลง Energetic-โทนสีเหลือง ฯลฯ เพื่อที่จะได้ไม่จำเป็นต้องใช้คำในการบ่งบอกถึงอารมณ์ของเพลง โดยเมื่อเพลงที่เล่นเปลี่ยนไป สีก็จะเปลี่ยนไปด้วย
* ทำการออกแบบ Data Flow Diagram ของแอปพลิเคชั่น

Diagram

Description automatically generated

* ศึกษาเกี่ยวกับเครื่องมือที่สำคัญในการทำ Feature Extraction ซึ่งก็คือ MFCCs หรือ Mel-frequency cepstral coefficients โดยจะเป็นเทคนิคที่ใช้ทำการดึงข้อมูลที่เหมาะสมของไฟล์เสียงออกมา โดยความเหมาะสมคือการที่ข้อมูลที่ดึงออกมาไม่ถูกรบบกวนด้วยค่า noise ของไฟล์เสียง ซึ่งเป็นเทคนิคที่ถูกทำมาใช้กับ Music Processing เมื่อต้นปี 2000 สำหรับขั้นตอนในการทำ MFCCs จะมีขั้นตอนทั้งหมด 8 ขั้นตอน

1. A/D conversion ซึ่งก็คือการแปลงสัญญาณ Analog เป็น Digital สำหรับกรณีทีรับ input เป็นเสียง Analog
2. Preemphasis เป็นการ filter เสียงในช่วงความถี่สูง โดยจะเพิ่มขนาดของพลังงานเสียงในช่วงความถี่นั้น
3. Windowing เป็นการแบ่งช่วงของไฟล์เสียงเพื่อดึง feature ต่างๆออกมาและทำการลด noise
4. DFT(Discrete Fourier Transform) เป็นการแปลงสัญญาณจากสัญญาณเทียบกับเวลาเป็นสัญญาณเทียบกับความถี่ โดยหลักการนี้บอกว่าการวิเคราะห์ไฟล์เสียงเทียบความถี่จะง่ายกว่า
5. Mel-Filter Bank เป็นการ Apply Mel-Filter กับไฟล์เสียงโดยจะอ้างถึงหลักการที่หูของมนุษย์มีประสิทธิภาพในการแยกแยะความแตกต่างที่เสียงย่านความถี่ต่ำได้ดีกว่าเสียงที่ย่านความถี่สูง ต่างกันกับคอมพิวเตอร์จึงใช้ Filter นี้เพื่อที่จะทำให้การเรียนรู้ของเครื่องในการแยกความแตกต่างใกล้เคียงกับมนุษย์
6. Applying Log โดยจะอิงถึงการที่มนุษย์มี sensitivity ในความเปลี่ยนแปลงของไฟล์เสียงที่พลังงานสูงน้อยกว่าไฟล์เสียงที่พลังงานต่ำ เช่นเดียวกับ Mel-Filter ที่จะ filter ค่าสัญญาณให้ใกล้เคียงกับที่มนุษย์ได้ยิน
7. IDFT เป็นการแปลงคลื่นเสียงโดยอิงจากค่าคลื่นเสียงต่างๆของมนุษย์ในการพูด เช่นการออกเสียงคำว่า tea, cat และ moo รูปปากและตำแหน่งลิ้นจะต่างกัน รวมไปถึงความสั่นของคลื่นเสียงจะแตกต่างกันด้วย
8. Dynamic Features จาก 7 ขั้นตอนที่ผ่านมา MFCC จะตัดสินเลือก feature ของเสียงทั้งหมด 13 features ที่ทำการ derivative อีกสองรอบทำให้ Feature ทั้งหมดที่จะได้จากการทำ MFCCs คือ 39 features



## ปัญหาที่เกิดขึ้นและแนวทางการแก้ไข

**Problem No. 2 เพลงมีการแบ่งประเภทที่ค่อนข้างหลากหลายมาก ทำให้ยากต่อการจำแนกประเภทได้แบบเฉพาะเจาะจง**

**พบปัญหาในการรายงานความก้าวหน้า ครั้งที่ 3**

**สถานะ**   กำลังดำเนินการ  แก้ไขสำเร็จ

**รายละเอียดปัญหา**

* ในเพลงๆหนึ่งนั้น แม้ว่าจะเป็นเพลงเพียงเพลงเดียว แต่ก็สามารถจำแนกออกมาได้หลากหลายอย่าง หากต้องการที่จะจำแนกทั้ง ประเภท และอารมณ์ ยกตัวอย่างเช่น เพลงแนว Pop เพลงหนึ่ง ก็จะมีท่วงทำนองที่แตกต่างกัน และด้วยทำนองที่แตกต่างกัน ก็ทำให้ได้อารมณ์ที่แตกต่างกันอีกด้วย นอกจากนี้อาจที่จะมีประเภทที่สามารถแยกย่อยได้อีกในเพลงนั้นๆ

**แนวทางแก้ไข/การแก้ไข**

* ทำการกำหนดเอาไว้ล่วงหน้า และตัดส่วนที่เราไม่ได้ทำการกำหนดเอาไว้ออก
* ทำการศึกษาเกี่ยวกับการจัดหมวดหมู่ของแอปพลิเคชั่นสตรีมมิ่งต่างๆ ที่มีอยู่

## สิ่งที่จะดำเนินการต่อไป

* ศึกษาเกี่ยวกับทฤษฏีของสีที่มีผลต่ออารมณ์หรือความรู้สึก
* เริ่มทำรายงานข้อเสนอ
* เริ่มทำเอกสารสำหรับการนำเสนอ