

การรู้จำรูปแบบของอารมณ์ในดนตรีด้วยโครงข่ายประสาทเทียมเชิงพหุนาม

A Polynomial neural network Approach to Music Emotion Recognition

อัฐพงศ์ รัตนประพันธ์ คำรณ สุชาติ เผ่าก๊ก ศิริสุข

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีมหานคร

บทคัดย่อ

บทความนี้นำเสนอแนวคิดเกี่ยวกับการรู้จำทางอารมณ์ในดนตรี โดยนำเอาหลักการของโครงข่ายประสาทเทียมเชิงพหุนาม (Polynomial neural network) หรือรู้จักกันในอีกชื่อหนึ่งว่า Group Method of Data Handling (GMDH) มาประยุกต์ใช้เพื่อทำนายค่าของการกระตุ้นและผลกระทบต่ออารมณ์ของผู้ฟังที่มีในดนตรี โดยการเปรียบเทียบกับวิธีการวิเคราะห์แบบถดถอยแบบต่างๆ ด้วยการอาศัยชุดข้อมูลทดสอบชุดเดียวกัน โดยใช้ GMDH ทำนายค่าของการกระตุ้นและผลกระทบต่ออารมณ์ในดนตรี ผลลัพธ์ที่ได้จากการทดลองได้แสดงให้เห็นว่า GMDH มีการวางนัยโดยทั่วไปดีกว่าการวิเคราะห์แบบถดถอยชนิดอื่นๆ

คำสำคัญ : การรู้จำอารมณ์ในดนตรี, การวิเคราะห์ถดถอย, โครงข่ายประสาทเทียมเชิงพหุนาม

Abstract

In this paper, we focus on the challenging issue of music emotion recognition (MER). Specifically, we use the Polynomial Neural Network algorithm, also known as Algorithm of Group method of data handling (GMDH), to predict the arousal and valence (AV) values of each music sample directly. To compare the performance with other regression techniques, we use the same dataset with all data feature, The performance of generalization is better than other method when using GMDH to predicting AV values.

Keyword : Music emotion recognition (MER), arousal, valence, regression, polynomial neural network, GMDH

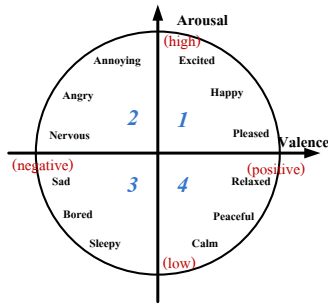
1. บทนำ

ในปัจจุบันการจำแนกประเภทของดนตรียังคงยากเกินกว่าความสามารถที่มนุษย์จะทำได้อย่างสมบูรณ์ ซึ่งต้องอาศัยผู้เชี่ยวชาญทางดนตรี รวมไปถึงข้อมูลที่มีอยู่เป็นจำนวนมาก หากมีเทคโนโลยีที่ช่วยให้สามารถเข้าถึงข้อมูลทางดนตรีได้ง่ายขึ้นจะส่งผลต่อการเพิ่มความต้องการในการเสกสรรทางดนตรีของผู้บริโภคได้สูงขึ้นอาจทำได้โดยอาศัยหลักการรู้จำรูปแบบของอารมณ์ที่มีในดนตรีหรือที่เรียกว่า Music Emotion Recognition (MER)

การรู้จำอารมณ์ในดนตรีมีองค์ประกอบในการพิจารณาสองส่วนได้แก่ ส่วนของการกระตุ้นอารมณ์ของผู้ฟัง (Arousal) และผลกระทบต่ออารมณ์ที่มีผลต่อจิตใจ

ของผู้ฟัง (Valence) ซึ่งจะเรียกค่าทั้งสองนี้ว่าค่า AV ในปัจจุบันยังไม่มีเกณฑ์มาตรฐานในการวัดและวิเคราะห์คุณลักษณะของอารมณ์ที่มีในดนตรีมาจำแนกประเภทที่ชัดเจนด้วยเหตุนี้งานวิจัยที่เกี่ยวข้องส่วนใหญ่จึงเลือกใช้แผนภูมิอารมณ์ที่เรียกว่า Thayer's arousal-valence emotion plane [1] ดังที่แสดงไว้ในรูปที่ 1

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1 Thayer’s arousal-valence emotion plane

จากแผนภูมิอารมณ์ในรูปที่ 1 แบ่งอารมณ์ออกมาได้เป็น 9 ประเภทแต่ความรู้สึกของอารมณ์ทั้ง 9 ประเภทนั้นมีระดับที่แตกต่างกันในแต่ละบุคคล ทำให้ต้องพิจารณาถึงรูปแบบการทดสอบทางอารมณ์ว่าควรใช้การแก้ปัญหาในลักษณะที่มองให้แต่ละอารมณ์ที่ได้แต่ละส่วนแยกต่างหากออกจากกัน (Discrete Analysis) หรือมองถึงความเกี่ยวเนื่องกันของอารมณ์ที่ได้ (Regression Analysis)

โดยส่วนใหญ่แล้วสำหรับการวิเคราะห์การรู้จำรูปแบบของอารมณ์ทางดนตรีมักเลือกใช้มุมมองที่มีลักษณะที่เกี่ยวเนื่องของอารมณ์มากกว่าที่จะระบุถึงประเภทอารมณ์ลงไปอย่างชัดเจนเนื่องจากปัญหาความกำกวมของอารมณ์มนุษย์ซึ่งรูปแบบการแก้ปัญหานี้ได้รับการยอมรับจากนักจิตวิทยาให้เป็น โมเดลสำหรับสนองความตอบรับทางอารมณ์ซึ่งเสนอไว้ [12-13]

ด้วยสาเหตุที่กล่าวมาบทความนี้จึงได้เลือกลักษณะการแก้ปัญหาโดยใช้การวิเคราะห์แบบต่อเนื่องของอารมณ์โดยใช้การสร้างตัวแบบที่อาศัยหลักการของโครงข่ายประสาทเทียมเชิงพหุนาม [3] โดยมีลักษณะการทำงานแบบก่อรูปเอง [4] และอาศัยการแบ่งกลุ่มของตัวแปรอินพุตมาทดสอบกันเป็นคู่ ซึ่งแต่ละชั้นของโครงข่ายภายในจะเชื่อมต่อกันด้วยสมการ โพลีโนเมียลทำให้เหมาะกับการแก้ปัญหาที่มีลักษณะไม่เป็นเชิงเส้นได้ดี

ในงานวิจัยที่ผ่านมาจะพบว่ามีปัญหาการเกิด Overfitting ก่อนข้างสูงซึ่งหมายถึงแบบจำลองที่สร้างขึ้นสามารถนำมาใช้ในการเรียนรู้ได้ดีแต่นำไปใช้ทำนายข้อมูลทดสอบได้ไม่ดี ซึ่งบทความนี้ได้ตั้งสมมุติฐานว่าหากทำการสร้างตัวแบบโดยใช้ GMDH น่าจะมีผลลัพธ์ที่ดีกว่าการวิจัยที่ผ่านมาเนื่องจากปัญหาที่พบในขั้นตอนการสร้างตัวแบบจะมีการคัดเลือกความสำคัญของพารามิเตอร์ที่จะนำมาใช้สร้างตัวแบบตามลักษณะการทำงานของโครงข่ายประสาทเทียมแบบก่อรูปเอง [4]

บทความนี้ได้แบ่งออกเป็น 6 ส่วน โดยส่วนแรกกล่าวถึงความสำคัญของการรู้จำอารมณ์ในดนตรีและเหตุผลของการนำเอา GMDH มาแก้ไขปัญหาคำว่าที่สอง

ได้กล่าวถึงงานวิจัยที่เกี่ยวข้องกับการรู้จำรูปแบบของอารมณ์ในดนตรี ส่วนที่สามกล่าวถึงทฤษฎีพื้นฐานของ GMDH ส่วนที่สี่กล่าวถึงการออกแบบระบบและข้อมูลที่นำมาใช้ในการทดสอบ ส่วนที่ห้าเป็นผลที่ได้จากการทดลองและส่วนสุดท้ายเป็นการสรุปผลที่ได้จากการทดลอง

2. งานวิจัยที่เกี่ยวข้องกับการรู้จำรูปแบบของอารมณ์ในดนตรี

มีงานวิจัยจำนวนหนึ่งพยายามจำแนกอารมณ์ของคนแต่พบกับปัญหาในการจัดการกับความคลุมเครือของอารมณ์จึงจะเห็นได้ว่างานวิจัย [5-6] พยายามทำการดึงเอาอารมณ์พื้นฐานต่างๆออกมาจากดนตรีเช่น ความสุข ความโกรธ ความเศร้าและความหวาดกลัว เป็นต้น สำหรับใน [7-11] ได้วิเคราะห์ถึงความกำกวมของอารมณ์โดยทำการแปลงรูปแบบของอารมณ์ให้ออกมาอยู่ในค่าที่สามารถวัดได้ซึ่งมีด้วยกัน 2 ค่า คือค่าระดับความตื่นตัวของอารมณ์ซึ่งทำให้ผู้ฟังรู้สึกตื่นเต้นหรือเฉื่อยชาและค่าของผลกระทบต้อจิตใจของผู้ฟังทั้งในเชิงบวกหรือลบ

ใน [20] ได้นำเสนอการทดสอบที่เป็นระบบมากขึ้น โดยทำการมองปริมาณของอารมณ์ในรูปแบบที่ต่อเนื่องและมีการเพิ่มกระบวนการในการสร้างตัวแบบมาจากคุณลักษณะที่ถูกวิเคราะห์ออกมา นอกจากนี้ความสัมพันธ์ระหว่างการกระตุ้นและผลกระทบทางอารมณ์ยังถูกนำเข้ามาพิจารณาด้วยซึ่งจากการทดลองนี้ได้ค่าของสัมประสิทธิ์การตัดสินใจ (R^2) จากการทดสอบเป็น 78.4% สำหรับค่าของการกระตุ้นทางอารมณ์และ 21.9% สำหรับค่าของผลกระทบทางอารมณ์ของผู้ฟังอย่างไรก็ดีการทดสอบนี้เป็นการทดสอบเพื่อหาความแปรปรวนของอารมณ์ภายในเพลง เพลงเดียวเท่านั้นจึงเป็นไปได้ยากในการนำมาใช้จัดกลุ่มอารมณ์ของบทเพลง

ใน [2] ได้นำเสนอแนวทางการวิเคราะห์การแก้ปัญหาด้วยการใช้ทฤษฎีของสมการลดถอยแบบต่างๆ เพื่อแก้ปัญหาด้วยการใช้หลักการรู้จำทางอารมณ์ ซึ่งได้ผลสรุปว่าการทำการกรองเพื่อลดระดับความสัมพันธ์ของการกระตุ้นและผลกระทบทางอารมณ์โดยใช้หลักการของการวิเคราะห์หาค่าหลักประกอบหลักหรือ Principal Component Analysis (PCA) [2] กับกลุ่มข้อมูลทดสอบเพื่อกรองความสัมพันธ์ระหว่างการกระตุ้นทางอารมณ์กับผลกระทบทางอารมณ์ของผู้ฟังจากนั้นจึงนำมาทำการเรียนรู้ด้วยกรรมวิธีต่างๆดังนี้

เฟอร์เชปตรอนหลายชั้น [16] เป็นการสร้างรูปแบบการเรียนรู้โดยทำการเลียนแบบการทำงานของสมอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มนุษย์ อาศัยฟังก์ชันกระตุ้นทำงานและการปรับค่าน้ำหนักของแต่ละโหนดในโครงข่ายซึ่งอาศัยการเปรียบเทียบกับข้อมูลที่ให้กับข้อมูลเป้าหมาย

สำหรับ AdaBoost (BoostR) [18] เป็นการทำงานโดยทำการสร้างกลุ่มของตัวจำแนกย่อยๆ ($h_0, h_1, ..h_n$) ซึ่งมีการทำงานที่แตกต่างกันในแต่ละขั้นหากกฎก่อนหน้าได้สร้างรูปแบบในการเรียนรู้ได้ดีแล้วจะทำการลดน้ำหนักในการเรียนรู้กับข้อมูลนั้นลงและเพิ่มค่าน้ำหนักในการเรียนรู้กับข้อมูลที่ผิดพลาดที่ไม่ถูกต้องเพื่อใช้ในการเรียนรู้กับกฎการเรียนรู้ถัดไป

นอกจากนี้มีการนำเอาหลักการของ RReliefF [17] มาประยุกต์ใช้เพื่อทำการคัดเลือกคุณลักษณะเด่นที่สำคัญออกมาทำให้การเรียนรู้โดยอาศัยการวิเคราะห์หาองค์ประกอบหลักนี้ได้ผลลัพธ์ออกมาดีที่สุดในวิธี Support Vector Regression (SVR) [2]

สำหรับการทดสอบ Test-reTest [2] เป็นการทดสอบโดยนำกลุ่มทดสอบเดิมที่เคยทดสอบผ่านมาแล้ว 1 เดือนทำการทดสอบซ้ำอีกครั้งเพื่อตรวจสอบความน่าเชื่อถือของกลุ่มข้อมูล โดยให้อาสาสมัครที่เข้าร่วมทดสอบทำการฟังเพลงตัวอย่างจากนั้นทำการทำเครื่องหมายลงบนแผนภูมิอารมณ์ในรูปแบบที่ 1 แล้วนำมาคำนวณหาเปอร์เซ็นต์ความถูกต้องกับข้อมูลที่เก็บไว้ใน 1 เดือนก่อนหน้า

ตารางที่ 1 แสดงประสิทธิภาพของกรรมวิธี regression แบบต่างๆ [2]

กรรมวิธี	กลุ่มข้อมูล	คุณลักษณะข้อมูล	(R ²)	
			a (%)	v (%)
MLR	AV	PSY15	56.8	10.9
BoostR	AV	PSY15	55.3	11.7
SVR	AV	PSY15	57.0	22.2
SVR	PC	RRF _{18,15}	58.3	28.1
Test-reTest	N/A	N/A	80.5	58.6

3. ทฤษฎีพื้นฐานของโครงข่ายประสาทเทียมเชิงพหุนาม (GMDH)

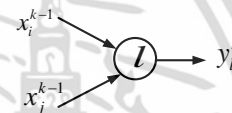
GMDH เป็นการสร้างแบบจำลองทางคณิตศาสตร์ของข้อมูลที่มีลักษณะไม่เป็นเชิงเส้น ซึ่งมีการคัดเลือกความสัมพันธ์ระหว่างข้อมูลสำหรับการวิเคราะห์ข้อมูลที่ซับซ้อนเพื่อทำการปรับปรุงอินพุตไปยังขั้นถัดไปสำหรับการทำงานของโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า [3] การทำงานพื้นฐานจะมองข้อมูลอินพุตในแต่ละชั้นว่าสามารถออกแบบโดยใช้การประมาณค่าจาก

อนุกรม Kolmogorov-Gabor polynomial ดังที่แสดงไว้ในสมการที่ (1) ได้

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k \dots \quad (1)$$

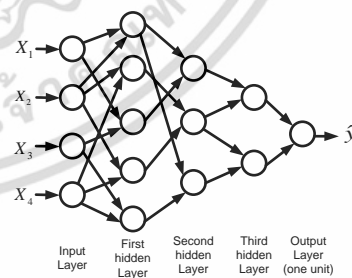
โดยปรกติในแต่ละชั้นของโครงข่ายที่มีใน GMDH จะเชื่อมต่อกันด้วยสมการ quadratic polynomial โดยชั้นซ่อนชั้นแรกที่เกิดขึ้นจะเกิดจากการจับคู่ของข้อมูลที่นำมาใช้ทดสอบซึ่งมีฟังก์ชันถ่ายโอนดังที่แสดงไว้ในสมการที่ (2) ส่วนชั้นซ่อนลำดับถัดไปจะอาศัยข้อมูลขาออกที่เกิดจากชั้นซ่อนก่อนหน้าดังที่แสดงไว้ในรูปที่ 2

$$\hat{y} = a_0 + a_1 x_{i_n} + a_2 x_{j_n} + a_3 x_{i_n} x_{j_n} + a_4 x_{i_n}^2 + a_5 x_{j_n}^2 \quad (2)$$



รูปที่ 2 แสดงองค์ประกอบของโหนด 1 ที่อยู่ในชั้นซ่อน K ของโครงข่ายประสาทเทียม

การทำงานของ GMDH จะเริ่มจากชั้นอินพุตไปยังชั้นผลลัพธ์โดยชั้นซ่อนแรกที่เกิดขึ้นจะเริ่มจากชั้น $k=1$ ไปยังชั้น $k=2$ และจากชั้น $k=2$ ไปยังชั้น $k=3$ เช่นนี้ไปเรื่อยๆจนกว่าจะได้ผลลัพธ์ \hat{y} (y_i^k ในรูปที่ 2) หรือเมื่อข้อมูลในชั้นที่ k ไม่สามารถสร้างผลลัพธ์ที่ดีกว่าชั้นที่ $k-1$ ได้ดังที่แสดงไว้ในรูปที่ 3



รูปที่ 3 แสดงโครงสร้างการทำงานของ GMDH

การหาจำนวนโหนดที่จะเกิดขึ้นในขั้นถัดไปทำได้โดยการอาศัยจำนวนโหนดในชั้นปัจจุบันดังที่แสดงไว้ในสมการ (3) เมื่อ C คือการจัดหมู่ (Combination) ซึ่งผลลัพธ์ที่ได้จะเป็นจำนวนโหนดที่เกิดขึ้นของชั้นที่ $k+1$ และ M_k คือจำนวนโหนดที่มีอยู่ในชั้น k

$$C(M_k, 2) = \binom{M_k}{2} = \frac{M_k!}{2!(M_k-2)!} \quad (3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2 เป็นภาพแสดงตัวอย่างของการสร้างโหนด l ในชั้นซ่อน k ดังนั้นในชั้นซ่อน k นี้โหนด l จะสามารถหาค่าได้จากสมการที่ (2) ดังนี้

$$y_l^k = a_0(x_i^{k-1})^2 + a_1(x_i^{k-1})(x_j^{k-1}) + a_2(x_j^{k-1})^2 + a_3(x_i^{k-1}) + a_4(x_j^{k-1}) + a_5 \quad (4)$$

จากสมการ (4) จะเห็นว่าแต่ละโหนดใหม่ที่เกิดขึ้นจะให้กลุ่มของค่าสัมประสิทธิ์กำหนดค่าให้อยู่ในรูปของเมตริกซ์ A เมื่อ l คือลำดับของโหนดในชั้น k ในการหาค่าสัมประสิทธิ์นั้นสามารถทำได้โดยการแทนค่าอินพุตทั้งหมดที่ใช้ในการเรียนรู้จากนั้นนำค่าเอาต์พุตมาใช้ซึ่งจุดประสงค์หลักในการเรียนรู้คือต้องการให้ y_l^k สอดคล้องกับรูปของสมการ $y = f(x)$ ดังนั้นจากสมการ (2) และ (4) จะได้ว่า

$$Y = XA \quad (5)$$

เมื่อ Y คือข้อมูลผลลัพธ์จริงที่นำมาใช้ในการเรียนรู้ โดยมีมิติเป็น $N_k \times 1$ โดยที่ N_k คือจำนวนแถวข้อมูลที่อยู่ในกลุ่มข้อมูลทดสอบที่นำมาใช้ในการเรียนรู้ ซึ่งมีลักษณะข้อมูลดังนี้

$$Y = [y_1 \ y_2 \ \dots \ y_{N_k}]^T \quad (6)$$

X เป็นเมตริกซ์ที่ได้มาจากการแทนค่าของตัวแปร x_i และ x_j ที่ทำการเลือกมาสร้างโหนดในชั้นนี้ โดยมีมิติเป็น $N_k \times 6$

$$X = \begin{bmatrix} 1 & x_{i_1} & x_{j_1} & x_{i_1} \times x_{j_1} & x_{i_1}^2 & x_{j_1}^2 \\ 1 & x_{i_2} & x_{j_2} & x_{i_2} \times x_{j_2} & x_{i_2}^2 & x_{j_2}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{i_k} & x_{j_k} & x_{i_k} \times x_{j_k} & x_{i_k}^2 & x_{j_k}^2 \end{bmatrix} \quad (7)$$

A คือค่าสัมประสิทธิ์ ของสมการ (13) โดยมีมิติเป็น 6×1

$$A = [a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ a_5] \quad (8)$$

จาก (5) หาเมตริกซ์ A เพื่อนำไปสร้างฟังก์ชันถ่วง โหนดของโหนดจะได้ว่า

$$X^T X A = X^T Y \quad (9)$$

$$A = (X^T X)^{-1} X^T Y \quad (10)$$

$$A = X^\dagger Y \quad (11)$$

เมื่อ X^\dagger เป็น Pseudo inverse ของ X หลังจากนั้นจะใช้ผลลัพธ์ที่ได้จากสมการ (11) ไปแทนค่าในสมการ (4) ทดสอบเพื่อหาค่าผิดพลาด โดยอาศัย sum square error ของค่า \hat{y}_i และค่า y_i เมื่อ \hat{y}_i เป็นค่าที่ได้จากตัวแบบและ y_i เป็นค่าจริงดังที่แสดงไว้ในสมการที่ (12)

$$\varepsilon = \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (12)$$

หลักการคัดเลือกข้อมูลที่จะผ่านไปยังขั้นต่อไปของ GMDH จะใช้หลักการของ Increasing Selection Pressure (ISP) Method [19]

4. การออกแบบระบบทดสอบ

การทดสอบในบทความนี้จะอ้างอิงชุดข้อมูลจาก [2] เนื่องจากต้องการตรวจสอบถึงความสามารถในการทำนายค่าของ GMDH ว่าแม่นยำเพียงใดซึ่งเพลงตัวอย่างจะถูกเปลี่ยนให้อยู่ในรูปแบบมาตรฐานที่ 22,050 Hz, 16 bits ในคลื่นความถี่โมโนของ PCM WAV จากนั้นทำการปรับระดับเสียงให้มีความดังเท่ากัน ทั้งนี้ดนตรีที่เลือกมาจะมีหลากหลายอารมณ์ภายในบทเพลงเดียวกัน ในการเลือกเฉพาะส่วนของบทเพลงจะใช้กรรมวิธีที่เรียกว่า Music thumbnailing algorithms [14] คัดกรองส่วนเด่นของเพลงซึ่งจะมีระยะเวลาประมาณ 24 วินาที

4.1 การเรียนรู้ของระบบ (Training)

ในการเรียนรู้ของระบบได้ทำการแบ่งข้อมูลออกมา 90% เพื่อนำมาใช้ในการเรียนรู้และ 10% สำหรับการทดสอบเช่นเดียวกับใน [2] จากนั้นจึงทำการเรียนรู้เพื่อวิเคราะห์ผลลัพธ์ที่ได้ออกมาในรูปของ R^2 ดังที่แสดงไว้ในสมการ (15) โดยทำการสุ่มจำนวนข้อมูลเพื่อนำมาใช้ในการเรียนรู้จำนวน 20 ครั้งแล้วจึงหาค่าเฉลี่ยของผลลัพธ์ที่ได้จากตัวแบบเพื่อให้สามารถนำผลลัพธ์ที่ได้นำมาใช้ในการทดสอบเปรียบเทียบกับ [2]

4.2 การออกแบบการทำงานของ GMDH

การออกแบบการทำงานของ GMDH ในรูปที่ 3 จะเริ่มจากการนำตัวแปรอินพุตมาทดสอบกันเป็นคู่ ซึ่งแต่ละชั้นจะเชื่อมต่อกันด้วยสมการโพลีโนเมียลที่สามารถลดรูปให้อยู่ในรูปของเลขยกกำลัง 2 ได้ซึ่งประกอบไปด้วยโหนดที่ถูกสร้างขึ้นจากการอาศัยเอาต์พุตของชั้นก่อนหน้า หากเป็นชั้นแรกจะใช้ข้อมูลอินพุตแทนลงในสมการโพลีโนเมียลดังกล่าว ทั้งนี้แต่ละโหนดจะให้กลุ่มของค่าสัมประสิทธิ์ a_i เมื่อ $i \in \{0, 1, 2, 3, 4, 5\}$ หลังจากนั้นจะใช้ผลลัพธ์ที่ได้จากสมการ (4) ไปทดสอบเพื่อหาค่าผิดพลาดโดยใช้สมการ (12) แล้วจึงเข้าสู่กระบวนการคัดเลือกข้อมูลสำหรับขั้นต่อไปด้วยวิธี ISP Method [19] จนกว่าจะได้ชั้นสุดท้ายที่เหลือเพียงโหนดเดียวหรือจำนวนชั้นที่เกิดขึ้นมากกว่าที่กำหนด

4.3 การวิเคราะห์ประสิทธิภาพของตัวแบบ

การประมาณค่าประสิทธิภาพของการวิเคราะห์แบบถดถอยหรือค่าสัมประสิทธิ์ในการตัดสินใจ (R^2) จากองค์ประกอบที่แตกต่างกันซึ่งประกอบไปด้วย การสุ่มชุดข้อมูลที่นำมาใช้ในการทดสอบ คุณลักษณะของข้อมูลที่

ต่างกันเพื่อนำมาใช้สำหรับการวิเคราะห์แบบถดถอยต่างๆ เพื่อตรวจสอบประสิทธิภาพโดยทำการแปลงให้อยู่ในรูปของ R^2 ซึ่งเป็นวิธีการวัดมาตรฐานของการวิเคราะห์แบบถดถอย [16]

$$R^2 = \left(1 - \frac{N\epsilon}{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2} \right) \times 100\% \quad (13)$$

เมื่อค่า \bar{y} คือค่าเฉลี่ยของค่าผลลัพธ์ของข้อมูลจริง (ground truth) และ $N\epsilon$ คือค่า normalization ของ total

5. ผลการทดลอง

ตารางที่ 2 ผลการทดลองที่ได้จากตัวแบบ GMDH ที่สร้างขึ้น

กรรมวิธี	กลุ่มข้อมูล	คุณลักษณะข้อมูล	การเรียนรู้		การทดสอบ			
			R^2		R^2		MAE	
			a(%)	v(%)	a(%)	v(%)	a	v
MLR [2]	AV	All	88	81	-8	-130	0.266	0.272
MLR [2]	AV	Psy15	65	28	57	11	0.181	0.193
SVR [2]	AV	All	68	43	56	9	0.190	0.213
SVR [2]	AV	Psy15	63	33	60	19	0.175	0.180
AdaBoost.RT [2]	AV	All	81	57	54	8	0.173	0.203
AdaBoost.RT [2]	AV	Psy15	78	52	58	11	0.183	0.189
GMDH	AV	ALL	68	18	64	18	0.222	0.295
GMDH	AV	PSY15	80	48	76	46	0.177	0.228
GMDH	PC	RRF _{18,15}	82	56	77	52	0.175	0.220

จากตารางที่ 2 ผลที่ได้ของค่าสัมประสิทธิ์ในการตัดสินใจ (R^2) ของการกระตุ้นและผลกระทบทางอารมณ์ที่ได้จากการสร้างตัวแบบของ GMDH ในการทดสอบข้อมูลกับตัวแบบที่ได้จากการเรียนรู้มีค่าที่ดีขึ้นกว่าการวิเคราะห์ถดถอยแบบอื่นทั้งนี้ค่าสัมประสิทธิ์การตัดสินใจที่ได้จากกระบวนการคัดเลือกข้อมูลและคัดกรองคุณสมบัติที่สำคัญจะมีค่าที่ดีกว่าการเลือกใช้คุณสมบัติทั้งหมดเป็น 64%, 76% และ 77% ตามลำดับสำหรับค่าการกระตุ้นอารมณ์และ 18%, 46% 52% สำหรับผลกระทบทางอารมณ์กับชุดข้อมูลที่เลือกทั้งหมด, การคัดกรองข้อมูลด้วย PSY15 และ RRF_{18,15} ตามลำดับ

ในการทดสอบที่ใช้ชุดข้อมูลทั้งหมดที่ไม่ได้ทำการคัดกรองคุณสมบัติสำคัญ พบข้อจำกัดในเรื่องหน่วยความจำทำให้สามารถวิเคราะห์ผลลัพธ์ได้เพียงขั้นเดียวของ GMDH บนเครื่องทดสอบที่ใช้ระบบปฏิบัติการ WindowsXP 32 บิต มีหน่วยความจำ 3 GB

สำหรับการสร้างตัวแบบด้วย GMDH นั้นได้ผลลัพธ์ที่ดีกว่าการทดสอบใน [2] เนื่องจากการวิเคราะห์แบบถดถอยจะมีประสิทธิภาพก็ต่อเมื่อนำไปใช้กับข้อมูลที่

squared error ทำให้ค่าของ R^2 เป็นการเปรียบเทียบกันระหว่างการทดสอบ ค่าของ R^2 มักถูกตีความถึงสัดส่วนความผันแปรของข้อมูลที่แฝงอยู่ซึ่งเกิดจากการวิเคราะห์แบบถดถอย หากค่าของ R^2 มีค่าเป็น 100.0% หมายความว่าตัวแบบสามารถนำไปใช้ประเมินข้อมูลได้ถูกต้อง หากค่าของ R^2 เป็นค่าลบหมายความว่าทิศทางของข้อมูลที่ตัวแบบประเมินได้ไม่สอดคล้องกับทิศทางของข้อมูลจริงซึ่งอาจสรุปได้ว่าตัวแบบไม่เหมาะสมกับชุดข้อมูลที่ต้องการประเมินนั่นเอง

สามารถแยกความสัมพันธ์ของชุดข้อมูลได้อย่างถูกต้องและเกี่ยวข้องกับผลลัพธ์ที่ต้องการอย่างแท้จริง แต่ชุดข้อมูลที่นำมาใช้ทดสอบนี้ยังมีข้อมูลบางส่วนที่เกี่ยวข้องกันทำให้เมื่อองค์ประกอบหนึ่งเปลี่ยนแปลงไปในทิศทางที่ไม่สอดคล้องกับทิศทางเดิมจะส่งผลกระทบต่อผลลัพธ์เกิดขึ้นด้วยความคลุมเครือขององค์ประกอบนี้จึงทำให้ผลลัพธ์ที่ได้ไม่น่าพอใจ ด้วยเหตุนี้เมื่อนำเอา GMDH มาใช้สร้างตัวแบบผลลัพธ์ที่ได้จะดีกว่าเนื่องจากในขั้นตอนของการสร้างตัวแบบได้มีการคัดกรองเอาคุณสมบัติต่างๆ ที่มีผลกระทบทับผลลัพธ์ไว้แล้วหากคุณสมบัติใดไม่เหมาะสมจะถูกคัดกรองออกไปโดยทำการจัดคู่ของลักษณะความสัมพันธ์ของข้อมูลสำหรับแต่ละขั้นดังนั้นเมื่อชั้นข้อมูลสูงขึ้นความสัมพันธ์ของข้อมูลที่เกี่ยวเนื่องกันจะมีจำนวนมากขึ้นทำให้ได้ผลลัพธ์ที่ได้จากตัวแบบใกล้เคียงกับข้อมูลจริงมากขึ้น

สำหรับค่าสัมประสิทธิ์การตัดสินใจของผลกระทบทางอารมณ์มีค่าที่น้อยกว่าค่าของสัมประสิทธิ์การตัดสินใจของการกระตุ้นทางอารมณ์อยู่พอสมควร ทั้งนี้เพราะลักษณะของผลกระทบทางอารมณ์นั้นเป็นการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากความรู้สึกเชิงนามธรรมของจิตใจซึ่งมีลักษณะเด่นที่น้อยกว่าเมื่อเทียบกับคุณลักษณะเด่นของการกระตุ้นทางอารมณ์ซึ่งเป็นข้อมูลที่เห็นได้ชัดเจนเช่น ระดับความดังของเสียง ความเร็วของจังหวะของดนตรี เป็นต้น

6. สรุป

จากปัญหาที่มีในการรู้จำอารมณ์ในดนตรีที่ใช้การวิเคราะห์แบบถดถอยในงานวิจัยที่ผ่านมาจะพบกับปัญหา Overfitting โดยเฉพาะกับการใช้ฟอร์เซปตรอนหลายชั้น ถึงแม้ภายหลังจะทำการกรองข้อมูลเพื่อคัดเลือกให้เหลือเพียงคุณลักษณะที่สำคัญแล้วก็ตามจะเห็นว่าผลลัพธ์ที่ได้จากการเรียนรู้และทดสอบมีค่าสัมประสิทธิ์ในการตัดสินใจที่แตกต่างกันมาก เมื่อนำเอาการจำแนกการรู้จำอารมณ์ในดนตรีด้วยการวิเคราะห์แบบถดถอยจากการใช้ GMDH ซึ่งมีลักษณะพิเศษคือมีการคัดกรองพารามิเตอร์ระหว่างที่ทำการสร้างตัวแบบ ทำให้การวางนัยโดยทั่วไป (Generalization) ดีกว่าการวิเคราะห์แบบถดถอยอื่นๆ ส่งผลให้ลดการเกิดปัญหา Overfitting ได้ดีกว่างานวิจัยที่ผ่านมา แต่การสร้างตัวแบบด้วย GMDH จะพบปัญหาเมื่อมีจำนวนมิติของข้อมูลเข้ามากเกินไปทำให้ไม่สามารถทำการวิเคราะห์ข้อมูลต่อไปได้เนื่องจากโหนดที่เกิดขึ้นในแต่ละชั้นทดสอบจะมากขึ้นเป็นทวีคูณ ดังนั้นการเลือกเอา GMDH ไปใช้งานควรทำการลดมิติของข้อมูลเพื่อคัดกรองเฉพาะข้อมูลที่มีความสำคัญเสียก่อน เช่นการนำเอาหลักการของ PCA ทำการกรองลักษณะเด่นของข้อมูลออกมานั้นจึงทำการเรียนรู้เพื่อทำการสร้างตัวแบบ GMDH ต่อไป

7. เอกสารอ้างอิง

- [1] R.E. Thayer, *The Biopsychology of Mood and Arousal*, Oxford University Press.
- [2] Y.H. Yang, Y.C. Lin, Y.F. Su, and H.H. Chen, "A Regression Approach to Music Emotion Recognition," *Audio, Speech, and Language Processing, IEEE Transactions*, vol.16, no.2, 2008, pp. 448-457
- [3] A.G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics*, 1971, pp. 364-378.
- [4] A.G. Ivakhnenko, "Heuristic self-organization in problems of engineering cybernetics," *Automatica*, vol.6, 1970, no.2, pp. 207-219.
- [5] Y.Feng, Y.Zhuang, and Y.Pan, "Popular music retrieval by detecting mood," *Proc. ACM SIGIR*, 2003, pp. 375-376.
- [6] T.Li and M.Ogihara, "Content-based music similarity search and emotion detection," *Proc. Int. Conf. Acoustic, Speech, and Signal Processing*, Toulouse, France, 2006, pp. 17-21.
- [7] L.Lu, D.Liu, and H.J.Zhang, "Automatic mood detection and tracking of music audio signals," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 14, no. 1, 2006, pp. 5-18.
- [8] M.Y. Wang, N.Y.Zhang, and H.C.Zhu, "User-adaptive music emotion recognition," *Proc. Int. Conf. Sig. Processing*, 2004, pp. 1352-1355.
- [9] T.L. Wu and S.K.Jeng, "Extraction of segments of significant emotional expressions in music," *Proc. Int. Workshop on Computer Music and Audio Technology*, 2006, pp. 76-80.
- [10] D.Yang and W.Lee, "Disambiguating music emotion using software agents," *Proc. Int. Conf. Music Information Retrieval*, 2004, pp. 52-58.
- [11] V.Carvalho and C.Chao, "Sentiment retrieval in popular music based on sequential learning," *Proc. ACM SIGIR*.
- [12] P.J.Lang, "The emotion probe," *American Psychologist*, vol. 50, no. 5, 1995, pp. 372-385.
- [13] J.A.Russell, A.Weiss, and G.A. Mendelsohn, "Affect grid: A single-item scale of pleasure and arousal," *Journal of Personality and Social Psychology*, vol. 57, no. 3, 1989, pp. 493-502.
- [14] N.C.Maddage, C.Xu, M.S. Kankanalli, and X. Shao, "Content-based music structure analysis with applications to music semantics understanding," *Proc. ACM Multimedia*, 2004, pp. 112-119.
- [15] D.Cabrera, "PSYSOUND: A computer program for psychoacoustical analysis," *Proc. Australian Acoustic Society Conf*, 1999, pp. 47-54.
- [16] A.Sen and M.Srivastava, *Regression Analysis: Theory, Methods, and Applications*, Springer.
- [17] M.R. Šikonja and I. Kononenko, "Theoretical and empirical analysis of Relief and RRelief," *Machine Learning Journal*, vol. 53, pp. 23-69, 2003.
- [18] D.P. Solomatine and D.L. Shrestha, "AdaBoost.RT: A boosting algorithm for regression problems," *Proc. IEEE Int. Joint Conf. Neural Networks*, 2004, pp. 1163-1168.
- [19] M. BULMER, *The Mathematical Theory of Quantitative Genetics*. Oxford University Press.
- [20] M.D. Korhonen, D. A.Clausi, and M. E. Jernigan, "Modeling emotional content of music using system identification," *IEEE Trans. Sys. Man. and Cyber.*, vol. 36, no. 3, 2006, pp. 588-599.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้