หัวข้อดุษฎีนิพนธ์	ตัวจำแนกแบบทวินไฮเปอร์อีลิปซอยดอลซัพพอร์ตเวกเตอร์	
ผู้เขียน	นาย เกษมสิทธิ์ ตียพันธ์	
ปริญญา	ปรัชญาดุษฎีบัณฑิต (วิศวกรรมไฟฟ้า)	
คณะกรรมการที่ปรึกษา	รองศาสตราจารย์ ดร. นิพนธ์ ธีรอำพน	อาจารย์ที่ปรึกษาหลัก
	รองศาสตราจารย์ ดร. เสริมศักดิ์ เอื้อตรงจิตต์	อาจารย์ที่ปรึกษาร่วม
	รองศาสตราจารย์ ดร. ศันสนีย์ เอื้อพันธ์วิริยะกุล	อาจารย์ที่ปรึกษาร่วม

บทคัดย่อ

การออกแบบตัวจำแนกข้อมูลเป็นหนึ่งก้าวสำคัญในการสร้างให้คอมพิวเตอร์สามารถที่จะเรียนรู้และจำแนกรูปแบบทางสถิติที่ซ่อนอยู่ภายในข้อมูลได้ โดยปกติแล้วตัวระบบการรู้จำจะเลือกกฎของการแยกแยะจากฟังก์ชันตั้งต้นที่ถูกกำหนดไว้ก่อนล่วงหน้า โดยฟังก์ชันดังกล่าวมักจะอยู่ในรูปของฟังก์ชันเชิงเส้น ดังเช่นในกรณีของซัพพอร์ตเวกเตอร์แมชีน (SVM) ที่อยู่ในรูปของไฮเปอร์เพลน สาเหตุที่ไฮเปอร์เพลนถูกนำมาใช้เนื่องจากไฮเปอร์เพลนเป็นฟังก์ชันที่ง่ายที่สุดและการใช้ฟังก์ชันที่ซับซ้อนมากเกินไปก็ไม่จำเป็นที่จะให้ผลลัพธ์ที่ดีที่สุด อย่างไรก็ตามในงานวิจัยต่าง ๆ ได้มีการพิจารณานำไฮเปอร์เซอร์เฟสซึ่งเป็นฟังก์ชันที่ซับซ้อนกว่าไฮเปอร์เพลนเพียงหนึ่งขั้นมาใช้ในการสร้างตัวระบบการรู้จำด้วย ยกตัวอย่างเช่น ได้มีการนำทรงกลมมาเป็นพื้นฐานในการสร้างตัวจำแนกข้อมูลในชื่อซัพพอร์ตเวกเตอร์ดาต้าเดสคริปชัน (SVDD) ด้วยเหตุนี้จึงมีความน่าสนใจในการที่จะนำทรงรีมาประยุกต์ใช้ในลักษณะเดียวกันเนื่องจากทรงรีถือเป็นรูปทรงที่มีความซับซ้อนมากกว่าทรงกลมอีกขั้นหนึ่ง

วิทยานิพนธ์ฉบับนี้มุ่งที่จะออกแบบตัวระบบการรู้จำที่มีไฮเปอร์อีลิปซอยด์เป็นพื้นฐานโดยได้เสนอตัวจำแนกข้อมูลชนิดใหม่ทั้งหมด 2 วิธีสำหรับการแก้ปัญหาการจำแนกข้อมูลแบบ 1 กลุ่ม และ 2 กลุ่มตามลำดับ โดยวิธีทั้ง 2 วิธีดังกล่าวมีพื้นฐานมาจากการหาทรงรีที่มีปริมาตรน้อยที่สุดในการล้อมรอบกลุ่มข้อมูล สำหรับการแยกแยะข้อมูล 1 กลุ่มนั้น วิทยานิพนธ์ฉบับนี้ได้เสนอตัวจำแนกข้อมูลในชื่อ อีลิปซอยดอลซัพพอร์ตเวกเตอร์ดาตาเดสคริปชัน (eSVDD) ที่ใช้การสร้างทรงรีที่มีปริมาตรที่น้อยที่สุดมาล้อมรอบกลุ่มข้อมูลในขณะเดียวกันก็พยายามให้จำนวนข้อมูลที่ไม่เกี่ยวข้องให้อยู่ภายในทรงรีให้น้อยที่สุด

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

สำหรับการแยกแยะข้อมูล 2 กลุ่มนั้น วิทยานิพนธ์ฉบับนี้ได้นำเสนอตัวจำแนกในชื่อ ทวินไฮเปอร์อีลิปซอยดอลซัพพอร์ตเวกเตอร์แมชีน (TESVM) ที่ได้รับแนวคิดมาจากทวินซัพพอร์ตเวกเตอร์แมชีน (TWSVM) และ ทวินไฮเปอร์สเฟียร์ซัพพอร์ตเวกเตอร์แมชีน (THSVM) สำหรับหลักการของตัวจำแนกที่นำเสนอนี้คือการสร้างทรงรีที่มีปริมาตรน้อยที่สุดล้อมรอบข้อมูลแต่ละกลุ่ม โดยที่มีเงื่อนไขเพิ่มเติมคือทรงรีหนึ่ง ๆ จะต้องอยู่ห่างจากกลุ่มข้อมูลที่ไม่เกี่ยวข้องให้มากที่สุดเท่าที่จะเป็นไปได้

การที่ SVM และ SVDD มีการใช้วิธีทางเคอร์เนลมาช่วยในการสร้างฟังก์ชันของการจำแนกข้อมูลให้มีความซับซ้อนมากขึ้น แต่วิธีการดังกล่าวไม่สามารถนำมาใช้โดยตรงกับวิธีการที่วิทยานิพนธ์ฉบับนี้ได้นำเสนอไป ด้วยเหตุผลที่ว่าวิธีการที่ได้นำเสนอถูกเขียนในรูปเอาเทอร์โปรดักส์แทนที่จะเป็นอินเนอร์โปรดักส์ ดังนั้นวิทยานิพนธ์ฉบับนี้จึงได้นำเสนอเพิ่มเติมถึงวิธีการที่จะนำวิธีการทางเคอร์เนลมาประยุกต์ใช้ โดยเสนอว่าวิธีการที่เหมาะสมทำได้โดยการใช้เทคนิคเอมพิริคอลฟีเจอร์แมป หรือเรียกอีกอย่างคือ เคอร์เนลพีซีเอแมป (kernel PCA map) ที่จะใช้การโปรเจกต์ภาพของข้อมูลไปยังปริภูมิที่ถูกกำหนดโดยเคอร์เนลฟังก์ชันก่อน แล้วจึงสร้างทรงรีในปริภูมินั้น

ทุกวิธีการจำแนกข้อมูลที่วิทยานิพนธ์ฉบับได้นำเสนอได้รับการทดสอบและเปรียบเทียบกับวิธีการอื่น ๆ ทั้งกับข้อมูลที่สร้างขึ้นและข้อมูลมาตรฐาน สำหรับการจำแนกข้อมูล 1 กลุ่ม วิธีที่ได้นำเสนอได้ถูกเปรียบเทียบกับ SVDD สำหรับการจำแนกข้อมูล 2 กลุ่ม TESVM ได้ถูกเปรียบเทียบกับ THSVM TWSVM และ SVM ผลการทดลองแสดงให้เห็นว่า การใช้ทรงรีที่ในการสร้างตัวจำแนกข้อมูลจะให้ประสิทธิภาพดีกว่าการใช้ทรงกลม และการนำข้อมูลที่ไม่เกี่ยวข้องมาใช้คำนวณร่วมด้วยจะให้ผลลัพธ์ที่ดีกว่าการไม่ได้ใช้ ยิ่งกว่านั้น TESVM ยังให้ผลการทดลองรวมที่ดีกว่า SVM และ TWSVM ในข้อมูลมาตรฐาน

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Dissertation Title	Twin Hyper-ellipsoidal Support Vector Classifier	
Author	Mr. Kasemsit Teeyapan	
Degree	Doctor of Philosophy (Electrical Engineering)	
Advisory Committee	Assoc. Prof. Dr. Nipon Theera-Umpon	Advisor
	Assoc. Prof. Dr. Sermsak Uatrongjit	Co-advisor
	Assoc. Prof. Dr. Sansanee Auephanwiriyakul	Co-advisor

ABSTRACT

Designing pattern classifiers is an important step toward creating an intelligent machine which can learn and recognize hidden statistical patterns inside the data. It is also one basic task in the field of machine learning. Generally, learning machines search for the best decision rule from its known set of functions. In the literature, such a predefined function is linear; for example, support vector machine (SVM) has hyperplanes as its underlying form of decision functions. This is because hyperplane is the simplest form of functions and using overly-complicated functions does not always guarantee the best decision boundary. Nevertheless, since hypersurfaces possess one more degree of geometrical sophistication than hyperplanes, research communities also have shown some interests in introducing hypersurfaces as a predefined set of function for learning machines. For instance, support vector data description (SVDD) searches for the best hypersphere from a given set of hyperspheres to form a descriptive boundary around the data. As a result, since hyperellipsoid is also the next complex geometrical shape to hyperspheres, it is tempting to also create learning machines which possess hyperellipsoids as a predefined set of functions.

This thesis focuses primarily on designing non-parametric learning machines whose predefined set of function is filled with hyperellipsoids. We propose two novel classifiers based on the formulation of minimum volume covering ellipsoid (MVCE), one for solving one-class classification problems and the other for binary classification problems. For one-class classification, the proposed method called “ellipsoidal support

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

vector data description (eSVDD) with negative examples" aims at constructing a descriptive boundary using the MVCE with soft margins around the target data while also trying to exclude outliers.

For binary classification, a novel MVCE-based binary classifier which tries to find two best fitting hyperellipsoids is proposed. The proposed method is named "twin hyper-ellipsoidal support vector machine (TESVM)". Its concept, as inspired by "twin support vector machine (TWSVM)" and "twin-hypersphere support vector machine (THSVM)", is to create a soft-margin MVCE around each class of data with an additional objective such that each hyperellipsoid must also be located as far as possible from the other class.

The learning machines like SVM and SVDD create more complex decision boundary through kernel tricks; however, the same kernel approach is not directly applicable for the proposed MVCE-based classifiers. In fact, applying kernel tricks to even the MVCE problem itself is not so straightforward since the formulation of MVCE comes in the form of outer products instead of inner products. As a result, we also propose that any MVCE-based classifiers can also be further equipped with kernel methods by using the so-called "empirical feature map" or "kernel principal component analysis map" (kernel PCA map). By projecting the images of the given examples into an approximate kernel-defined space, it is possible to construct a hyperellipsoids in that space, resulting in more complex descriptive boundary in the input space or the original space.

All the proposed methods are evaluated against state-of-the-art techniques in several experiments on both artificial and publicly-available standard real-world datasets. For one-class classification, we compare the proposed eSVDD with negative examples against SVDD, SVDD with negative example, as well as eSVDD. For two-class classification, the proposed TESVM is also tested with THSVM, TWSVM, and SVM. The experimental results show that the uses of hyperellipsoids in both eSVDD with negative examples and TESVM are superior to their hyper-spherical counterparts. The inclusion of negative examples into the formulations also helps improve the overall performance of the MVCE-based classifiers. Furthermore, TESVM also provides better results than SVM and TWSVM in term of average accuracy on the standard datasets.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved