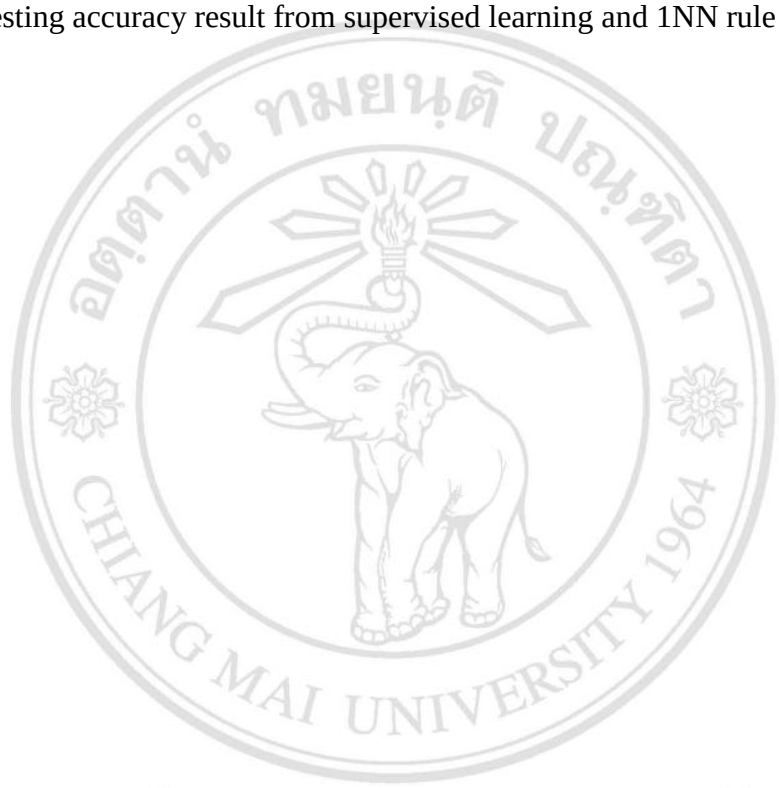
CONTENTS

LIST OF TABLES



LIST OF FIGURES

LIST OF SYMBOLS

c	Number of clusters
C	Number of classes
ℓ	Number of sample points of the data set
N	Number of attributes of the data set
$\mathbf{x}_i$	An input column vector of the output y_i
$\mathbf{T}$	The training data set matrix
$\mathbf{X}$	The input matrix
$\mathbf{y}$	The output vector
$\mathcal{T}$	The set of training data
$\mathbf{x}$	A sample point in $\mathcal{T}$
$\mathcal{B}$	The approximated class boundary set
$\mathcal{RS}$	The reduced data set
$\xi(\mathbf{x})$	The nearest enemy function of $\mathbf{x}$
$\mathcal{X}$	The temporary set

ข้อความแห่งการริเริ่ม

1) วิทยานิพนธ์นี้นำเสนอวิธีการลดข้อมูลแบบใหม่ บนพื้นฐานของการจัดกลุ่มอินพุต-เอาต์พุต ที่สามารถควบคุมความถูกต้องและอัตราการลดข้อมูล โดยกำหนดค่าเริ่มต้นให้กับการแบ่งส่วนข้อมูลและจำนวนของการแบ่งนับเพื่อใช้สอนระบบการทำนาย
2) วิทยานิพนธ์นี้ยังได้นำเสนอวิธีการลดข้อมูลแบบใหม่ บนพื้นฐานของวิธีการลดจำนวนและการแก้ไขชุดข้อมูล เพื่อใช้สอนระบบการจำแนก วิธีการที่นำเสนอนี้มีความซับซ้อนในการคำนวณน้อยกว่าวิธีการที่นำมาเปรียบเทียบ เนื่องจากมีการประมวลผลเบื้องต้นเพื่อหาข้อมูลบริเวณขอบของกลุ่มข้อมูลอื่นมาใช้ประมวลผลแทนการประมวลผลกับข้อมูลทั้งหมด โดยวิธีการนี้สามารถเลือกและคงข้อมูลที่สำคัญที่อยู่บริเวณขอบของกลุ่มข้อมูลไว้ได้

STATEMENTS OF ORIGINALITY

1. This thesis proposed a new instance reduction method for regression data based on input-output clustering with the abilities to control the accuracy and reduction ratio via setting the number of separation parts and quantization levels.

2. This thesis also proposed a new instance reduction method based on condensation and edition methods for classification. The computational complexity of the method is lower than other methods because the class boundary of the enemy class is preprocessed by calculating the class boundary instead of all enemy sample points. The proposed method can select and retain the important data which are located at the class boundary.